International Journal of Computer Science & Emerging Technologies

ISSN 2044-6004 (Online)

Volume 1 Issue 3

October, 2010

www.sprinterglobal.org/ijcset

ijcset@gmail.com

IJCSET BOARD MEMBERS

TABLE OF CONTENTS

K.A.Suryanarayana
Department, Department of Economics, M.R.P.G.College, Vizianagaram, Andhra Pradesh, India

15. Global eXtreme Programming, a Software Engineering Framework for Distributed Agile Software Development (pp-106:112)

Ridi Ferdiana, Lukito Edi Nugroho, Paulus Insap Santoso
Department of Electrical Engineering and Information Technology, Gadjah Mada University, Yogyakarta – Indonesia

Ahmad Ashari
Department of Computer Science and Electronics, Gadjah Mada University, Yogyakarta – Indonesia

16. The Metaplectic Sampling of Quantum Engineering (pp-113:118)

Walter J. Schempp
Lehrstuhl für Mathematik I, University of Siegen, 57068 Siegen, Germany

17. KVISIMINE++ Applied to Problems in Geographical Information System (pp-119:124)

Sarita Patel, Prof. Deepa Chaurse
Computer Sceince and Engg.Technocrats Institute of Tech, Bhopal, M.P. India

18. An Exploratory and Feasibility Study of Implementing Online Based Voting System in Bangladesh (pp-125:132)

Mohammad Shabbir Hasan, Quazi Farhan and Abdullah Al Mahmood
Panacea Research Lab, Dhaka, Bangladesh

19. Characterization of QoS Based Routing Algorithms (pp-133:141)

Dr. Shuchita Upadhyaya
Department of Computer Science and Applications, Kurukshetra University, Kurukshetra, Haryana, India

Gaytri Devi
GVM Institute of Technology and Management, M.D. University, Sonipat, Haryana, India

Buffer Overrun: Techniques of Attack and Its Prevention

Mahtab Alam[1], Prashant Johri[2], Ritesh Rastogi[3]
1: Asst. Prof. & Head of Dept. Computer Science, Aryabhatt College of Engineering and Technology, Baghpat
Email: alam_mahtab@rediffmail.com
2: Asst. Prof, Dept. of Computer Sc., NIET, Gr. Noida, Email: johri.prashant@gmail.com
3: Asst. Prof, Dept. of Computer Sc., NIET, Gr. Noida, Email: rit_raj@hotmail.com

Abstract: Buffer Overflow attack has been considered as one of the important security breaches in modern software systems that has proven difficult to mitigate. This attack allows the attacker to get the administrative control of the root-privilege by using the buffer overflow techniques by overwriting on the address of a returned function, function pointer stored on the memory and overflow a buffer on the heap. In this paper, we present the different buffer overflow techniques used by the exploiters and the methodologies applied to mitigate the buffer overflow.

Keywords: Buffer Overrun, Heap Smashing, Pointer Subterfuge, Arc Injection

1. Introduction

The complexity and opportunity of software systems vulnerabilities are regularly growing with the use of computer system. Almost every software system is insecure because of the high growth rate of expertise of the malicious users. Software system is considered insecure because of its existing security holes. Buffer Overflow attacks are the most common security intrusion attack [3,5] Software security holes related to buffer overflow accounts the largest share of CERT advisories. David Wagner from University of California at Berkeley shows that buffer overflows stand for about 50% of the vulnerabilities reported by CERT [3]. In the memory allocation table, variables with similar properties are assigned into the same buffer area, and their locations are adjacent to each other. A buffer overflow condition occurs when a program attempts to read or write outside the bounds of a block of allocated memory or when a program attempts to put data in a memory area past a buffer [1]. A buffer overflow may happen accidentally during the execution of a program [2]. Buffer overflow is best known for software security vulnerability, as buffer overflow attack can be performed in legacy as well as newly developed application. Buffer overflows are applicable to most operating systems [2]. In particular the attacks are quite successful in Windows NT and Windows 2000 system [4,6,7,8,9,10]

A buffer is a sequential section of computer memory that holds more than one instances of the same data type. It is allocated to contain anything from a character string to an array of integers. An extremely common kind of buffer is simply an array of character type. Overflow occurs when data is added to the buffer outside the block of memory allocated to the buffer. Buffer overflow can be conducted either by locally or remotely. In a local attack the attacker already has access to the machine and acquires the access privileges. On the other hand in remote attack the attacker deliver commands through network port, and simultaneously gains the unauthorized access privilege.

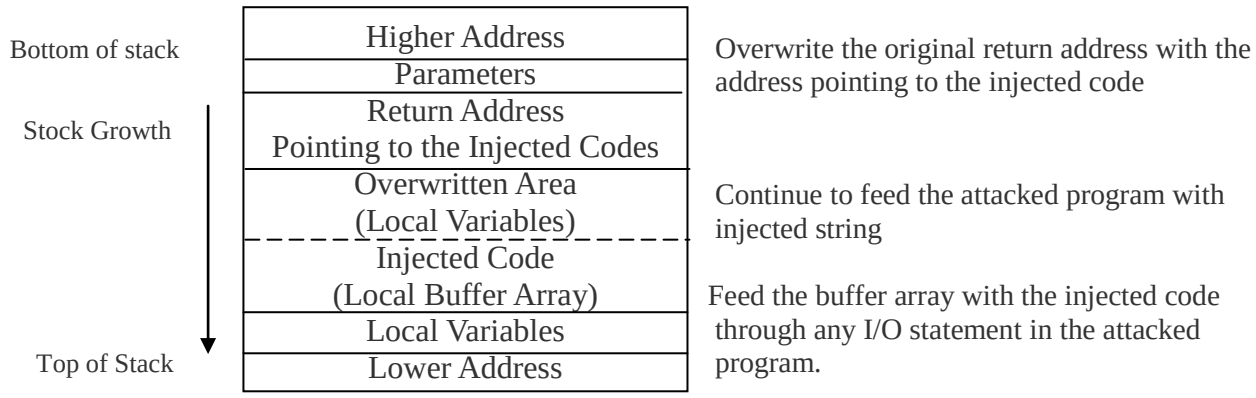


Figure -1 Fragment of a Stack

Buffer overrun is characterized as a stack overrun or heap overrun depending on what memory gets overrun. Stack memory is used in C and C++ compilers when local variables as well as parameters have been used. Heap memory in this context refers to the dynamically allocated memory uses new/delete or *alloc*(), *malloc*().

Buffer overrun mainly consist the following three steps [2]: Planting the attack code into the program, copying into the buffer which overflows it and corrupts adjacent data structures, and hijacking the program to execute code.

Commonly buffer overflow can be executed by using the stack smashing: modifying the return address saved on the stack to point to code the attacker supplies that resides in stack buffer overrun exploitation as shown in the Figure-1 [11].

2. Buffer Overrun Methods

In recent years, hackers have developed some other approaches of buffer overrun to exploit software such as Arc Injection, Pointer Subterfuge and Heap Smashing.

2.1 Arc Injection

Arc Injection sometimes also called as return–into-libc transfer control of the code that already exists in the memory space. These types of injection insert a new arc using the installation of an existing functions such *system*(), *execl*() or *printf()* as into the program's control flow graph and create a shell on the compromised machine with the permission of the compromised program. An exploiter uses the arc injection to invoke a number of functions in a small program that includes chained functions in sequence with arguments that are supplied by them.

Example: Following are the main functions used in arc injection buffer overrun vulnerability [23].

2.1.1 system():

system takes a single argument and executes that argument with /bin/sh.

2.1.2 execl():

execl() requires an argument list that is null terminated. This will end our string early, so there is a need to chain multiple calls to libc.

2.1.3 printf():

printf is very popular output function used in C language, but it can be used for exploitation of a program using following techniques:

- The %n parameter prints how many characters have been written so far to a location specified by the argument.
- By using n$ inside a parameter, one can read the value of the nth argument.
- Combining these, %3$n will write the number of characters printed so far to the address specified in the 3rd argument.

2.2 Pointer Subterfuge

Pointer Subterfuge is a general expression for exploitation by using modification of pointer address. This approach used by an attacker to divert the control flow of a program by using function pointers (a variable whose value is used as an address) as an alternative to the saved return address, or modify the program flow by subverting data pointers [1]. A pointer subterfuge software exploitation is illustrated as below [24]:

```
void SomeFunc() {
        // do something
}
typedef void (*FUNC_PTR )(void);

int MalFunc(char *ptString) {
        char buf[32];
        strcpy(buf,szString);
        FUNC_PTR fp = (FUNC_PTR)(&SomeFunc);
        // Other code
        (*fp)();
        return 0;
}
```

If the malicious user uses the ptString argument in function MalFunc, then the buffer in the stack buf is ready to overpower. If the attacker overwrites the function pointer fp, then this pointer points to another address and exploits code and the function (*fp)() is invoked. To overcome the problem caused by pointer subterfuge we have to protect the function pointer.

2.3 Heap Smashing

Heap Smashing attack overruns a heap buffer to change the control flow of a program. Such overflow could overwrite

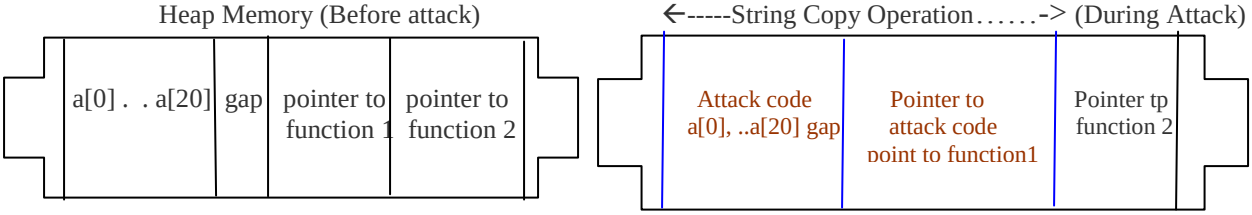


Figure-2 Heap smashing attack

function pointers stored on the heap to redirect the control flow. Heap Smashing allows an attacker to exploit the software by implementing some assumed variants in dynamically allocated memory. Although this type of attack is less common in practice but can be dangerous. Attacker typically is not aware the heap block's location ahead of time and the standard

trampoline approaches are not effective. A typical example of heap smashing is shown in the Figure-2 [25]

3. Mitigation Techniques

A user may overwrite the input buffer by providing more data for storage within the buffer than the programmer has expected. Errors in string manipulation have long been recognized as a leading source of buffer overflows in C and C++. A number of mitigation strategies have been devised. These include mitigation strategies at requirement and design levels to prevent buffer overflows from occurring and strategies that are Static analysis techniques should be employing to find the common coding problems that could expose buffer overrun. A through interface testing will further reduce the risk by providing the existence of buffer overruns and allowing the development team to fix them as they are found.

By using all the above mentioned techniques in a layered approach at secure software requirement analysis phase, it may be possible to reduce the risk of buffer overruns at some extent

3.1 Layered Approach

Buffer overruns are generally caused by introducing bugs during application implementation. These bugs can be mitigated by using following three techniques [12]:

3.1.1 Using Interpreted language:

Developed an application using a interpreted language that reduces the potential for buffer overruns, such as C# or Java. The interpreted code eventually calls into support code that is written into a compiled language such as C/C++ that could contain buffer overrun.

3.1.2 High Quality Code:

Buffer overrun to some extent can be mitigated by ensuring development in an environment that encourages a high-quality code that requires developers to participate in code review, running unit test, and educating them about buffer overruns. Buffer overrun sneak into the code either through inexperience or designed to detect buffer overflows and securely recover without allowing the failure to be exploited. Rather than completely relying on a given mitigation strategy, it is often advantageous to follow a defense-in-depth strategy of combining multiple strategies. Some approaches to prevent the buffer overrun in a program are described in this sequence.

3.1.3 Testing Public Interfaces:

Static analysis techniques should be employing to find the common coding problems that could expose buffer overrun. A through interface testing will further reduce the risk by providing the existence of buffer overruns and allowing the development team to fix them as they are found.

By using all the above mentioned techniques in a layered approach at secure software requirement analysis phase, it may be possible to reduce the risk of buffer overruns at some extent

misunderstanding on the part of the developer regarding how the code works within the large application. Unfortunately, there are a large number of dangerous functions that come with C and C++. Any place a program uses them is a warning signal, because unless they are used carefully, they become vulnerable [26].

3.2 Traditional Approaches

Traditionally, buffer overruns caused by unsafe functions in the C library, like *strcpy*() have been identified and replace them with safe function like *strncpy*(). In this approach the static intrusion prevention method in which the software bugs can be eliminated by examining the large number of program carefully is applied. Removing all security bugs from a program is considered infeasible [17] which makes the static prevention incomplete. There are some tools available that one can use to automate the search for the vulnerability [13, 14, 15], but still manual auditing of the code which makes this massive and very expensive approach [2]. While the value of this systematically auditing code has been successfully executed, the approach is not guaranteed to produce buffer-overrun-free code [16].

3.3 Compiler Approaches

Almost all the buffer overruns problem take place in the compiler-based programming languages. Range checking indices are very effective against the buffer overrun attacks. Buffer overrun attack is not possible in Java programming language because Java automatically checks that an array index within the proper bounds. In C language it is not possible because of the dichotomy

between arrays and pointers [2]. When a compiler compiles a function *strcpy*(char* a, char* b) the two arguments are pointers and it is impossible for a complier to know the length of the corresponding array, and complier cannot generate code for range checking inside the function. Compiler Approach is a kind of dynamic intrusion prevention techniques which allow changing the run-time environment or system functionality making program at some extent less vulnerable.

C compiler allocates memory space for a local variable and a function return address in the same stack frame and adjacent to each other as shown in Figure-1. To mitigate the possibility of this type of problem some types of safe compilers are invented and implemented which are as follows [18}:

3.3.1 StackGuard:

The StackGuard complier was invented and implemented by Crispin Cowan [18]. The main objective of the StackGuard is to prevent the dynamic intrusion prevention by detecting and stopping stack based buffer overflow and return address. The overhead for StackGuard can reportedly be as high as 40% [19, 20].

In buffer overrun arrack, the stack is target to fill the higher address area and then overwrite the other local variable below the area specified for local variables The key ideas to mitigate this technique is to place a dummy value known as canary, in between the return address and local variables as shown in Figure -3. If the attacker try to overrun the buffer area, the canary intact the changing the return address, either by overwriting the canary with its correct value and thus not changing the actual one, or by overwriting the return address through a pointer.

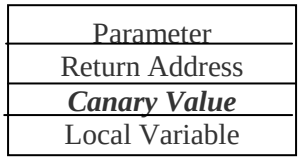


Figure-3 StackGuard Frame

Although these techniques only stop the buffer overflow attack that generally attack against the return address, but attacker still have potential to abuse the pointers variables, making it point at the return address and writing a new address to the memory position.

3.3.2 Stack Shield

Stack Shield is a tool for adding protection to programs from this kind of attacks at compile time without changing a line of code [21]. Stack Shield is also a compiler extension mechanism that protects the return address. Stack shield is more secure protection system than tool like Stack Guard. In the latest version 0.7 of stack shield there are two techniques which protect against writing of the return address and one against overwriting of function pointers.

(a) Global Ret Stack

In this mechanism the return address upon calling a function has been copied to Global Ret Stack array of 32-bit entries. Whenever a malicious user alters the address of the function, it has no effect since the original return address is remembered. In this method only prevention not attack detection is possible in this technique.

(b) Ret Range Check:

In this mechanism the value of the return address of a current function is store in the global variable. While calling the function the return address on the stack is compared with the value copied in the global variable. If there is any difference the program execution is halted. It can detect the attack too.

(c) Protection of Function Pointer:

Function pointer normally points to the text segment of the process memory. If the process ensure that no pointer is allowed to point the other parts of the memory except text segment, it is impossible for an attacker to inject a payload (Combination of data and code) into the process. Protection of function pointer can be possible by declaring a global variable in the data segment and its address is used as a boundary value. If the function points above or below the boundary the process is terminated.

4. Future Work

The results obtained from our work shows that buffer overrun can be easily conduct by smashing heap and stack memory or by overrunning the bytes available in the memory. Our motive of future work is to reduce these types of problem by using pointer encryption and resolve the techniques to mitigate the very small size even 'one- byte' of buffer overrun.

5. Conclusion

Buffer overrun is the most important software security breach. There are several techniques for stopping the common security buffer overrun. But we have presented some mitigation techniques related to requirement and design level of software development life cycle. Applying the above mentioned approaches one can mitigate the buffer overrun problem at some extent.

Reference

[1]. Jonathan Pincus, Brandon Baker, Beyond Stack Smashing: Recent Advances in Exploiting Buffer Overruns", IEEE Coputer Society, 2004, pp.20-27.

[2]. Istvan Simon, "A Comparative Analysis of Methods of Defense against Buffer Overflow Attacks", January 31, 2001.

[3]. D. wagner, J. S. Foster, E.A. Brewer, and A Aiken, "A First Step towards automated detection of Buffer Overrun Vulnerabilities. In Proceedings of the IEEE Symposium on Security and Privacy, pp.-3-17, February 2000.

[4]. DilDog, The TAO of Windows Buffer Overflow, http://www.cultdeadcow.com/cDc_files/cDc-351/, 1998

[5]. D. Evans and D. Larochelle, "Improving Security Using Extensible Lightweight Static Analysis", IEEE Software, 19(1), pp.-42-51, February 2002

[6]. Nishad Herath, (Joey_) Advanced Windows NT Security, in Black Hat Asia Conference, 2000. URL: http://www.blackhat.com/html/bh-asia-00/bh-europe-00-speakers.html#Joey

[7]. Barnaby Jack, (dark spyrit), Win32 Buffer Overflows (Location, Exploitation, and Prevention), Phrack Magazine 55,(15), URL: http://phrack.infonexus.com/search.phtml?view&article=p55- 15, 1999.

[8]. Mudge, How to Write Buffer Overflows, http://l0pht.com/advisories/buffero.html, 1997

[9]. Ryan Russel, Rain Forest Puppy, Elias Levy, Blue Boar, Dan Kaminsky, Oliver Friedrichs, Riley Eller, Greg Hoglund, Jeremy Rauch, and Georgi Guninski, "Hack Proffing Your Network Internet Tradecraft", Syngress, 2000.

[10]. Joel Scambray, Stuart McClure, George Kurtz, Hacking Exposed, Network Security Secrets & Solutions , Second Edition, Osborne/McGrawHill , 2001

[11]. Fu-Hau-Hsu, "RAD: A Compile-Time Solution to Buffer Overflow Attacks", Department of Computer Science, State University of New York at Stony Brook

[12]. Jason Taylor, "Webservices Risk Assessment and Recommendation",Security Innovation Commercial Tools Division 1318 S, Babcock Street, Melbourne, 2003/2004

[13]. HalVar Flake, "Auditing Binaries for Security Vulnerabilities", in Black Hat Europe Conference, 2000, URL: http://www.blackhat.com/html/bh-europe-00/bh-europe-00-speakers.html

[14]. Nishad Herath (Joey_) "Advanced Windows NT Security", In Black Hat Asia Conference, 2000

[15]. David Wagner, Jeffrey S Foster, Eric A. Brewer, and Alexander Aiken, "A First Step Towards Automated Detection of Buffer Overrun Vulnerabilities", in Proceedings 7th Network and Distributed System Security Symposium 2000

[16]. Chris Evans, Nasty Security Hole in lprm, posted in BugTraq, April 18, 1998

[17]. D. Larochelle and D. Evans, "Statically Detecting Likely Buffer Overflow Vulnerabilities", in Proceedings of the 2001 USENIX Security Symposium, Washington DC, USA, August 2001.

[18]. C. Cowan, C. Pu, D. Maier, J. Walpote, P. Bokke, S. Beatie, A. Grier, P. Wagle, Q. Zhang, and H. Hinton, "STachGuard: Automatic adaptive detection and prevention of buffer-overrun attacks", in Proceedings of the 7th USENIX Security Conference, pp-63-78, San Antonio, Texas, Jannuary 1998.

[19]. David Llewellyn-Jones, Madjid Merabti, Qi Shi, Bob Askwith, "Buffer Overrun Prevention Through Component Composition Analysis", Proceedings of the 29th Annual International Computer Software and Application Conference (COMPSAC'05), IEEE Transaction 2005.

[20]. C. Cowan, S. Beattie, R. Finnin Day, C. Pu, P. Wagle, and E. Walthinsen, "Protecting Systems from Stack Smashing Attacks with StackGuard", In Linux Expo., May 1999.

[21]. Stack Shield A "stack smashing" technique protection tool for Linux: URL: http://www.angelfire.com/sk/stackshield/

[22].http://www.acm.uiuc.edu/sigmil/talks/general_exploitation/arc_injection/arc_injection.html

[23]. John Wilander, Mariam Kamkar, "A Comparison of Publicly Available Tools for Dynamic Buffer Overflow Prevention", Published at 10th Network and Distributed System Security Symposium (NDSS), 2003.

[24].http://blogs.msdn.com/michael_howard/archive/2006/01/30/520200.aspx

[25]. Christof Petzer, Zhen Xiao, "Detecting Heap Smashing Attacks Through Fault Containment Wrappers.

[26]. David Wheeler, "Secure Programmer: Countering Buffer Overflows", 27 Jan 2004

Author Biographies

Mahtab Alam: Mr. Mahtab Alam is working as an Assistant Professor and Head of Dept. of Computer Science in Aryabhatt College of Engineering and Technology, Baghpat. He has having 12 years experience in the field of teaching and6 years in software industry. He has published Books in Cryptography and Information Security and currently pursuing Ph. D. in the field of Information Security. His areas of interest are Software Engineering, Information Security, Database Management System and computer graphics.

Prashant Johri : Dr. Prashant Johri is working as as Assistant Professor, Dept. of Computer Sc. In Noida Institute of Engineering and Technology, greater Noida, He has more than 10 years experience in teaching. He has completed his research work in software reliability and his areas of interest are Software Engineering, Design and Analysis of Algorithm, Simulation and Modeling.

Ritesh Rastogi: Mr. RItesh Rastogi is working as a Assistant Professor and Head of Department (MCA) at NIET Gr.Noida for last Eight years. He is having an experience of around ten years in the field of teaching. He has published number of Books and is continuously involved in the field of publication. His main areas of concern are Software Engineering, Information Security, Database and Information System.

Video Encryption using Wavelet Transform with Shift Number Coding

Mohammed Mustafa Siddeq[1], Abdulrahman Ikram Siddiq[2]
Software Engineering Deprt. – Technical College/Kirkuk
[1]mamadmmx76@yahoo.com , [2]dr_aisiddiq@yahoo.com

***Abstract*:-** The nature of playing video streams over a network in real time requires that the transmitted frames are sent in a bounded delay. Then, the most suitable encryption for real time video transmission is that with the shortest execution time for a given security level. In this paper, a new video encryption algorithm is suggested for real time applications, by using the Discrete Wavelet Transform (DWT), the Shift Number Coding (SNC) and a secret key. The SNC converts each 6 bytes into a single floating point number. The time of execution and decrypted video quality of the proposed algorithm is evaluated for distinct frames of a video file. This performance is then compared with other encryption algorithms that use the Discreet Cosine Transform (DCT), Huffman coding, and Arithmetic Coding. The tests showed that the proposed algorithm is faster than the other tested algorithms. Therefore, it seem to be more suitable for real time video encryption.

***Keywords*:-** *Discrete Wavelet Transform, Shift Number Coding*

1. Introduction

The advent of networked multimedia systems made continuous media streams, such as real time audio and video, increasingly pervasive in computing and communications environments. It is thus important to secure networked media from potential threats such as hackers, eavesdroppers, etc [1,2]. The nature of playing video streams over a network in real time requires that the transmitted frames are sent in a bounded delay. Also, video frames need to be played at a certain rate. Therefore, the processes of sending and receiving encrypted packets must be performed within a certain amount of time limited by the admissible delay [3]. The size of the multimedia data before encryption is hence a concerning factor. That is, the smaller multimedia data size, the shorter time is required to encrypt/decrypt it, and vice versa. Generally, the encryption and decryption of a video stream can be performed In two ways:

1. Secret key encryption: a single secret key can be used to encrypt and decrypt the video stream. Only the sender and the receiver have this key.
2. Public key encryption: there are two keys, one for encryption and the other for decryption. The public key, which is known for all senders, is used for encryption. While the private key, which is owned by the receiver is used for decryption.

Public key cryptography is not applicable for securing real time video because its operations require an amount of time not suitable for this application. However, the idea of public key encryption is applicable for other multimedia security aspects such as signature and authentication [4]. On the other hand, secret key encryption can provide the encrypted data within a limited time [3,5]. Therefore, secret key encryption is usually used to secure real time video transmission. In the literature, video encryption techniques may be classified as Native, Selective, Zigzag, and pure permutation algorithms.

- Native Algorithms

This technique deals with the video stream as text data and encrypts every byte in the stream individually [6]. It guarantees high security level, but it is not applicable for large amounts of data, as the case of video streams, because its operations will result in a delay not suitable for real time video encryption.

- Selective Algorithms

Tang in [7] suggested employing different levels of encryption for selected parts of the video streams. This is based on the fact that the nature of the different components of a video file is not the same. Tang has suggested four levels of selective algorithms. These levels are: encrypting all headers, encrypting all headers and I frames, encrypting all I frames and all I blocks in P and B frames, and finally, encrypting all frames individually as in Native algorithms to guarantee the highest security. The encryption time is controllable and it depends on the number of used levels.

- Zig-zag Algorithms

The idea of zig-zag algorithm is basically encrypting the video stream before compressing it. Explicitly, rather than mapping an 8x8 block to a 1x64 vector each time in the same order, a predetermined permutation can be used. However, the concept of the encryption key does not exist in the zig-zag algorithm. Once the permutation list is known the algorithm is not secure any more [8].

- Pure Permutations

The idea of pure permutation is simply to apply a permutation technique for the I frames. Both the sender and the receiver have the correct permutation order to encrypt and decrypt the video stream. Later works [9,10,11] proved that it is not secure to use pure permutations.

In this paper, a new video encryption algorithm is suggested for real time applications, by using the Discrete Wavelet Transform (DWT), the Shift Number Coding (SNC) and a secret key, the proposed algorithm deals with the successive frames individually.

2. The Proposed Algorithm

The proposed algorithm deals with the successive frames of a video stream individually. That is, each frame is encrypted and decrypted independently of the other frames in the same video sequence. This property results in a small buffering memory requirement limited in size to the size of a single frame, which is suitable for real time operation.

The encryption process consists of two stages as shown in Figure. 1. Firstly, it uses the DWT to decompose a frame into its low and high frequency components. The low frequency part is encoded with the efficient coding technique SNC. Then the result is XORed with the secret key. In parallel with this operation, only the nonzero pixels of the high frequency parts, which usually consist of a great number of zeros, are XORed with the same secret key. At the receiver, the reverse operations are performed to decrypt the received frames.

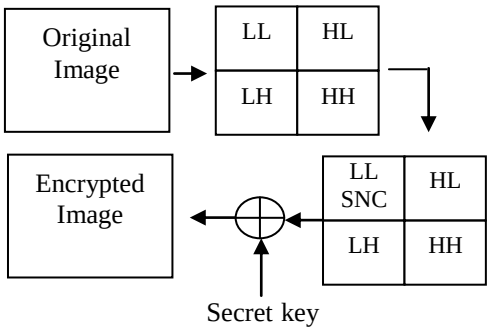
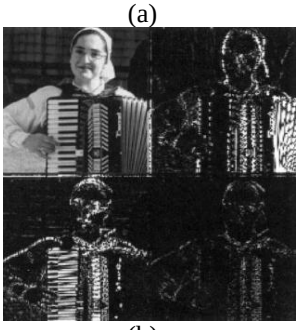


Figure – 1 Encryption Algorithm

2.1 Discrete Wavelet Transform

A single-stage wavelet transformation consists of a filtering operation that decomposes an image (frame) into four frequency bands as shown in Figure. 2. The original image is shown in Figure.2 (a), and Figure.2 (b) is the result of a single-stage wavelet transform. The top-left corner of the transformed image "LL" is the original image, low-pass filtered and sub-sampled in the horizontal and vertical dimensions. The top-right corner "HL" consists of residual vertical frequencies (i.e. the vertical component of the difference between the sub-sampled "LL" image and the original image). The bottom-left comer "LH" contains residual horizontal frequencies (the accordion keys are very visible here), whilst the bottom-right corner "HH" contains residual diagonal frequencies [11,12,13].

(a)



(b)

Figure – 2 (a) Original image (b) Single stage wavelet Transform

The reason of using the wavelet transform is to minimize the number of operation in the encryption and decryption processes. Because most transformed image regions are zeros which can be ignored when it is XORed with the secret key, as a worked example, let the 8x8 square matrix "A" represent an image, and defined as follows:

$$A=\begin{bmatrix} 136, 136, 137, 137, 139, 134, 129, 134 \\ 136, 136, 137, 137, 139, 134, 129, 134 \\ 136, 136, 137, 137, 139, 134, 129, 134 \\ 133, 140, 133, 131, 133, 128, 128, 130 \\ 128, 132, 133, 129, 138, 132, 135, 128 \\ 130, 131, 130, 126, 132, 130, 133, 131 \\ 130, 131, 130, 132, 135, 132, 128, 128 \\ 131, 133, 128, 133, 127, 122, 128, 128 \end{bmatrix}$$

The matrix A is converted to four sub-images by using the DWT. Then,

$$LL=\begin{bmatrix} 272, 274, 273\ 263 \\ 273, 269, 267, 261 \\ 261, 259, 266, 264 \\ 263\ 262, 258, 256 \end{bmatrix} \quad HL=\begin{bmatrix} 0, 0, 0, 0 \\ -1, 5, 6, 3 \\ -1, 3, 4, -1 \\ -2, 1, 9, 0 \end{bmatrix}$$

$$LH = \begin{bmatrix} 0, 0, 5, -5 \\ -3, 1, 5, -3 \\ -3, 4, 4, 4 \\ -2, -3, 4, 0 \end{bmatrix} \quad HH = \begin{bmatrix} 0, 0, 0, 0 \\ 3, -1, 0, -2 \\ -1, 0, 2, 3 \\ 0, 2, -1, 0 \end{bmatrix}$$

The matrices HL, LH, and HH are encrypted by XORing them with the secret key, while the LL matrix is first encoded by the SNC technique, and then the result is XORed with the same secret key.

2.2 Shift Number Coding

The implementation of SNC begins with reducing the number of data levels from M levels by using the following equation:

$$\text{Pixels}_{(new)} = \frac{(\text{Pixels}_{(old)} * \text{Threshold})}{\text{MAX}} \quad (1)$$

Where Threshold is the maximum value of the new pixels, and MAX is the maximum value of the old pixels.

Reduction of pixel levels leads to a reduction in the pixels size. The Threshold decides new pixels range. For example, when Threshold=60, the new pixels range becomes [0 - 60]. The same threshold value is used for all of the frames in the given video stream. Next, the new pixel values are normalized to there maximum value (Threshold) to fall within the new range [0 - 1]. Then, the interval [0 - 1] is divided into 60 parts, which means that the interval between any two pixel values is 0.016, as shown in Table 1.

Table 1 Value for each pixel

Pixels	Values	Pixels	Values
0	0.0	….	….
1	0.016	51	0.816
2	0.032	52	0.832
3	0.048	53	0.848
4	0.064	54	0.864
5	0.08	55	0.88
6	0.096	56	0.896
7	0.112	57	0.912
8	0.128	58	0.928
9	0.144	59	0.944
10	0.160	60	1.0

The SNC algorithm encodes each 6 bytes of the new pixels into a single floating number by shifting and summing pixel values. This technique does not need any statistical computations on the image file, and also does not need to store any extra information for the decoding process.

$$\text{Coded Value} = \sum_{i}^{n} \text{Shift}(i) * \text{Value}(i) \quad (2)$$

For $i = 1,2 \ldots 6$

The coded value is a single floating point number resulted from combining 6 bytes as described above. The values of Shift(i) are given in Table 2 for each index value i.

Table 2 Shift function

index	Shift(index)
1	1
2	0.01
3	0.0001
4	0.000001
5	0.00000001
6	0.0000000001

The SNC algorithm for each 6-Byte data is shown below:

Set Coded_Value to 0.0;
Set index to 1;
While (index $\leq$ 6) **Do**
  Read (symbol)
  Value = **interval** (Symbol);
  Coded_Value=Coded_Value+Shift(index)* Value;
  Index = index + 1;
End *While*

The SNC of the LL matrix resulted from the DWT decomposition of matrix A, defined before, is achieved first by changing the levels of the pixels to the new range using Threshold=60. The result is as follows

$$LL = \begin{bmatrix} 272, 274, 273, 263 \\ 273, 269, 267, 261 \\ 261, 259, 266, 264 \\ 263, 262, 258, 256 \end{bmatrix} \quad LL_{(new)} = \begin{bmatrix} 60, 60, 60, 58 \\ 60, 59, 59, 57 \\ 57, 57, 58, 58 \\ 58, 58, 57, 56 \end{bmatrix}$$

Next, when equation 2 is applied on $LL_{(new)}$ every 6 bytes (6 new pixels) are transformed to a single floating point number as shown in Table 3.

Table 3 Applied SNC

Data	Value	Shift(index)	Coded_Value
60	0.96	1	0.96
60	0.96	0.01	0.9696
60	0.96	0.0001	0.969696
58	0.928	0.000001	0.969696928
60	0.96	0.00000001	0.9696969376
59	0.944	0.0000000001	0.969696937694
Data	**Value**	**Shift(index)**	**Coded _Value**
59	0.944	1	0.944
57	0.912	0.01	0.95312
57	0.912	0.0001	0.0532112
57	0.912	0.000001	0.953212112
58	0.928	0.00000001	0.953212121279
58	0.928	0.0000000001	0.953212121372
Data	**Value**	**Shift(index)**	**Coded _Value**
58	0.928	1	928.
58	0.928	0.01	0.93728
57	0.912	0.0001	0.9373712
56	0.996	0.000001	0.937372096

Therefore the original sub-matrix $LL_{(new)}$ is converted to:

$$LL_{SNC} = \begin{bmatrix} 0.969696937694 \\ 0.953212121372 \\ 0.937372096000 \end{bmatrix}$$

The final step in the encryption process it is XOR matrices: LL_{SNC}, HL, LH, and HH with the secret key, the secret key is a 16-byte integer number. It is XORed with each 16-byte from matrices: LL_{SNC}, HL, LH, and HH. If a pixel value is zero, XOR operation ignores this pixel. For example: the matrix HL contains zeros, XOR ignored zeros, and other bytes from secret key XORed with other elements from the matrix. Also the matrix LL_{SNC} is XORed with secret key, by converting each floating point number to integer numbers.

For example: in matrix LLSNC the floating point number is [0.969696937694] convert to integer numbers as: [96, 96, 96, 93, 76,94].

The decryption algorithm is reverse for encryption, the operation start with XOR with secret key to return original matrices: LL,HL, LH, and HH. Finally the Inverse SNC applied on matrix LL to get original pixels. The Inverse SNC illustrated in the following steps, and Table 4.

While (**Coded_Value**> 0) **Do**
  For I =1 to Threshold **Do**
    IF Pixel_Value(**I**)≤ **Coded_Value**<
      Pixel_Value(**I+1**) **THEN**
        Let K= Pixel_Value(**I**);
        Pixel=Get_Pixel (**I**);
    End *IF*
  End *For*
 Coded_Value=**Coded_Value** -K;
Coded_Value=**Coded_Value***100;
End *While*

Table 4 Inverse SNC

Data Range	Value	Selected Data
[0.96 - 0.976]	0.969696937694	60
[0.96 - 0.976]	0.9696937694	60
[0.96 - 0.976]	0.96937694	60
[0.92 – 0.94]	0.937694	58
[0.96 - 0.97]	0.9694	60
[0.94-0.96]	0.94	59
Data Range	**Value**	**Selected Data**
[0.94 – 0.96]	0.953212121372	59
[0.91 – 0.92]	0.9212121372	57
[0.91 – 0.92]	0.92121372	57
[0.91 – 0.92]	0.921372	57
[0.92 – 0.94]	0.9372	58
[0.92 – 0.94]	0.92	58
Data Range	**Value**	**Selected Data**
[0.92 – 0.94]	0.937372096	58
[0.92 – 0.94]	0.9372096	58
[0.912-0.928]	0.92096	57
[0.86-0.91]	0.896	56

The LL_{SNC} matrix converts approximately to original matrix by using equation (1), the following steps illustrates conversion:

$$LL_{SNC}=\begin{bmatrix}60, 60, 60, 58\\60, 59, 59, 57\\57, 57, 58, 58\\58, 58, 57, 56\end{bmatrix}$$

→ Apply Inverse Equation (1)

$$LL_{(new)}=\begin{bmatrix}273, 273, 273, 264\\273, 268, 268, 259\\259, 259, 264, 264\\264, 264, 259, 255\end{bmatrix}$$

Finally the matrices LL, HL, LH and HH are used by Inverse DWT to compose into one matrix (Decryption Matrix), as shown below:

$$A=\begin{bmatrix}137, 137, 137, 137, 139, 134, 130, 135\\137, 137, 137,1 37, 139, 134, 130, 135\\135, 138, 137, 136, 140, 135, 130, 133\\136, 139, 132, 131, 134, 129, 127, 130\\128, 131, 133, 129, 136, 132, 134, 130\\129, 132, 130, 126, 132, 128, 135, 131\\130, 132, 131, 134, 136, 132, 128, 128\\132, 134, 130, 133, 127, 123, 128, 128\end{bmatrix}$$

Decryption matrix "A" is approximately same as original matrix, the Encryption and Decryption algorithm is lossly algorithm. Because at encryption algorithm all data are floating point numbers, and our algorithm operates on integer numbers. The conversion between integers and floating point's numbers affects on decryption algorithm, and this leads to change at image quality (See Section 3).

3. Computer Simulation

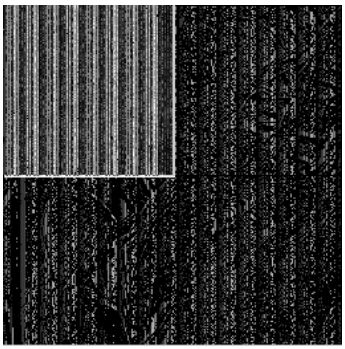
Our algorithm applied on (MATLAB Language), by using "**Pentium4 - 1.73MHz**, 1MB Cache Memory". First test for our algorithm on "Lena" image (256 x 256), at first the gray level for "Lena" image are reduced by equation(1), where MAX=256, and Threshold=55. The encrypted image shown on Figure.-3, used DWT and SNC, finally, make XORed with each sub-image, where secret key = {48, 65, 129, 51, 151, 34, 34, 167, 178, 100, 128, 9, 1, 45, 56, 48}. The decryption algorithm start with using secret key with each encrypted sub-image, then using Inverse SNC and Inverse DWT to obtain approximately original image.

Peak signal to noise ratio (PSNR) can be calculated very easily and is therefore a very popular quality measure. It is widely used as a method of comparing the "Quality" of Original and Decrypted video images [6,13]. PSNR calculated using equation (3) it is measured on a logarithmic scale and is based on the mean squared error (MSE) between an original image and decrypted image, relative to $(255)^2$ (i.e. the square of the highest possible signal value in the image).

$$PSNR = 10\log_{10}\frac{255^2}{MSE} \quad (3)$$

(a) Original Image

(b) Wavelet Transform



(c) SNC and XORed with Secret Key

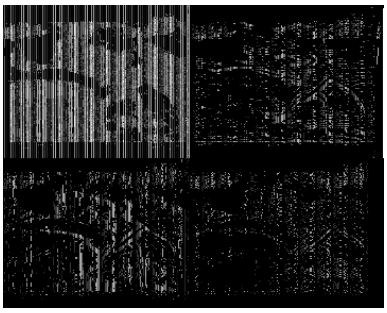
(d) Decrypted Image *PSNR=40dB*

Figure. – 3 Encryption and Decryption Algorithm applied on "Lena" image

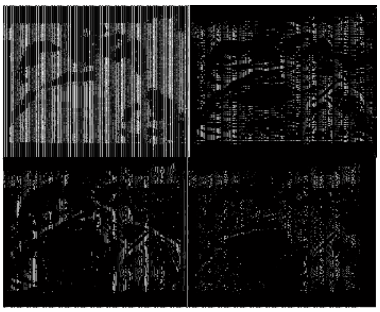
Second test for our approach on two frame of football match, frame's size (288 x 352), these frames encrypted by using same secret key, and Threshold=50, as shown on Figure - 4. Table 5 and Table 6 has shown Time execution and PSNR respectively.

(a) Original Frame 1

(b) Original Frame 2



(c) Encrypted Frame 1



(d) Encrypted Frame 2

(e) Decrypted Frame1 *PSNR=37.9dB*

(f) Decrypted Frame 2
PSNR=39.5dB

Figure. – 4 Encryption and Decryption Algorithm applied on "Frame1 and Fram2

Table 5 Time execution for our algorithm

Image Name	Encryption Time	Decryption Time	Total Time
Lena	0.178 **Sec**.	0.384**Sec.**	0.562**Sec.**
Frame 1	0.306**Sec.**	0.662**Sec.**	0.968**Sec.**
Frame 2	0.310**Sec.**	0.64**Sec.**	0.95**Sec.**

Table 6 PSNR for our algorithm

Image Name	Image Size	PSNR
Lena	64-KByte (256 x 256)	40dB
Frame 1	100-KByte (288 x 352)	37.9dB
Frame 2	100-KByte (288 x 352)	39.5dB

4. Comparison Methods

Our approach compared with (DCT), Arithmetic coding, and Huffman coding which these methods are used at most in image compression. Table 7 and Table 8 shown the comparison between our approach and DCT combined with (Arithmetic coding) and Huffman coding. Also our approach is compared with DWT combined with Arithmetic coding and Huffman coding [7,10,12,13].

Table 7 Comparison Methods with time execution

Algorithm	Image Name	Encryption Time (Sec.)	Decryption Time (Sec.)
Our Approach	Lena	0.178	0.384
	Frame1	0.306	0.662
	Frame2	0.310	0.64
DCT & Arithmetic Coding	Lena	2.71	1.138
	Frame1	3.062	0.90
	Frame2	3.0	0.92
DCT & Huffman Coding	Lena	31.7	408.53
	Frame1	52.21	508.2
	Frame2	56.21	505.2
DWT & Arithmetic Coding	Lena	1.12	0.625
	Frame1	3.45	1.0
	Frame2	3.40	1.0
DWT & Huffman Coding	Lena	29.8	212.14
	Frame1	21.9	382.9
	Frame2	20	383.0

Table 8 Comparison Methods with PSNR

Algorithm	Image Name	Total Time (Sec.)	PSNR (dB)
Our Approach	Lena	0.562	40dB
	Frame1	0.968	37.9dB
	Frame2	0.95	39.5dB
DCT & Arithmetic Coding	Lena	3.84	58.8dB
	Frame1	4.28	58.8dB
	Frame2	4.20	58.8dB
DCT & Huffman Coding	Lena	440.23	58.8dB
	Frame1	560.41	58.8dB
	Frame2	561.41	58.8dB
DWT & Arithmetic Coding	Lena	1.74	54.4dB
	Frame1	4.45	58.2dB
	Frame2	4.4	62.4dB
DWT & Huffman Coding	Lena	241.94	54.4dB
	Frame1	404.8	58.2dB
	Frame2	403	62.4dB

5. Conclusion

This research introduces a new idea for video encryption, consist from two parts: 1) DWT, used to compose an image into four sub-images. 2) SNC used to convert each 6-byte data into single floating point. Finally make XOR with secret key. Also our approach compared with DCT, Arithmetic coding and Huffman coding. The main reason for using DCT compared with DWT, because DCT and DWT used Widely in image compression [3,6,13], in this paper we use them in encryption and compare them by time execution and PSNR (See Table 7,8). The main difference between DWT and DCT, the DWT decompose one image into four sub-images, and most of sub-images regions containing on zeros [12,13]. While in DCT contains negative numbers in most regions [1,13]. In our approach the DWT take less time than DCT. Because in this paper not need make XOR with zero elements. Also compared SNC with Arithmetic and Huffman coding, the SNC not need to compute the probability for an image, while arithmetic and Huffman coding using probability computations for an image, for this reason SNC more speedily than Arithmetic coding and Huffman coding (See Table 7,8).

RERERENCES

[1] S. Li, G. Chen, A. Cheung, B. Bhargava, K.-T. Lo, "On the design of perceptual MPEG-video encryption algorithms", *IEEE Transactions on Circuits and Systems for Video Technology* 17 (2) 214–223, (2007).

[2] Chengqing Li, Guanrong Chen, "On the security of a class of image encryption schemes," in Proc. *IEEE International Symposium on Circuits and Systems 2008 (ISCAS08)*, pp. 3290-3293, 2008.

[3] Zhenyu Yang, Yi Cui, Zahid Anwar, Robert Bocchino, Nadir Kiyanclar, Klara Nahrstedt, Roy H. Campbell, William Yurcik, "Real-Time 3D Video Compression for Tele-Immersive Environments", *in Proc. of SPIE/ACM Multimedia Computing and Networking (MMCN'06),* San Jose, CA, 2006.

[4] C. Li, S. Li, G. Chen, G. Chen, L. Hu, "Cryptanalysis of a new signal security system for multimedia data transmission", *EURASIP Journal on Applied Signal Processing* 1277–1288, 2007.

[5] Klara Nahrstedt, Lintian Qiao, "Multimedia Secure Gateway", *in Proc. of Multimedia Security Workshop, IEEE International Conference on Multimedia Computing and Systems (ICMCS'98),* pp. 1-8, Austin, TX, June, 1998.

[6] K. Sayood, *Introduction to Data Compression*, Morgan Kaufmann Publishers, Inc., San Francisco, CA, 2nd dition, 2000.

[7] Witten, Ian H., Neal, Radford M., and Cleary, John G. , "Arithmetic Coding for Data Compression", *Communications of the ACM,* pp 520-540, June, 1987.

[8] Lintian Qiao, Klara Nahrstedt, Ivan Tam, "Is MPEG Encryption Using Random Lists instead of Zig Zag Order Secure?", in Proc. of IEEE International Symposium on Consumer Electronics '97, Singapore, December, 1997.

[9] G. Alvarez, S. Li, L. Hernandez, "Analysis of security problems in a medical image encryption system", *Computers in Biology and Medicine* 37 (3) 172 424–427, 2007.

[10] J. Zhou, Z. Liang, Y. Chen, A. O. C., "Security analysis of multimedia encryption schemes based on multiple Huffman table", *IEEE Signal Processing Letters 14* (3)166 (2007) 201–204.

[11] N. Saito and J.-F. Remy, ."A new local sine transform without overlaps: A combination of computational harmonic analysis and PDE", in *Wavelets: Applications in Signal and Image Processing X*, M. A. Unser, A. Aldroubi, and A. F. LaineEds, vol. Proc. SPIE 5207, pp. 495.506, 2003.

[12] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression",. *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1532.1546, 2000.

[13] Rafael C. Gonzalez, Richard E. Woods *Digital Image Processing*, Addison Wesley publishing company – 2001.

First Author: He is works on image processing lab. and programming lab. at software Engineering Department in Technical College – Kirkuk –Iraq. He is finished his (B.Sc.) degree at Alnahhrain University –Baghdad-Iraq, and he is finished his (M.Sc.) degree at university of Technology at same country. His fields; on Digital Image Processing, Cryptography, Neural Network, Genetic Algorithms and Algorithm Design.

Second Author: He is head of department of software Engineering Department in Technical College – Kirkuk – Iraq. He is finished his (B.Sc.) degree at Alnahhrain University –Baghdad- Iraq, and he is finished his (M.Sc.) and (PhD) degree at same university. His fields; on Digital Image Processing, Cryptography, FPGA and DSP.

Sign Language to Speech Converter Using Neural Networks

Mansi Gupta, Meha Garg, Prateek Dhawan

Lingaya's Institute of Management & Technology, Faridabad, India
{manasigupta18@gmail.com, mehagarg.be@gmail.com,prateek_3212@yahoo.com}

Abstract*:* The normal community has a limited fluency in sign language and because of this a communication barrier persists between the normal and the hearing-impaired people. This Barrier is diminishing as projects of the past two decades have unfolded. These not only help in interpreting the signs but also ease the communication between deaf and general communities. Through the use of artificial intelligence, researchers are striving to develop hardware and software that will impact the way deaf individuals communicate and learn. In an attempt towards the same, a converter has been proposed in this paper. This converter would act as a medium by recognizing the signed images made by the signer and then convert those into text and subsequently into speech. The signed images are classified to increase the accuracy and efficiency of the algorithm.

Key words: Sign converter, Sign language, neural networks, image processing

1. Introduction

A sign language (also signed language) is a language which, instead of acoustically conveyed sound patterns, uses visually transmitted sign patterns (manual communication, body language and lip patterns) to convey meaning—simultaneously combining hand shapes, orientation and movement of the hands, arms or body, and facial expressions to fluidly express a speaker's thoughts. Sign languages commonly develop in deaf communities, which can include interpreters, friends and families of deaf people as well as people who are deaf or hard of hearing themselves. [8]

Wherever communities of deaf people exist, sign languages develop. [9] In fact, their complex spatial grammars are markedly different from the grammars of spoken languages. Hundreds of sign languages are in use around the world and are at the cores of local deaf cultures. Some sign languages have obtained some form of legal recognition, while others have no status at all. In addition to sign languages, various signed codes of spoken languages have been developed, such as Signed English and Warlpiri Sign Language.[1] These are not to be confused with languages, oral or signed; a signed code of an oral language is simply a signed mode of the language it carries, just as a writing system is a written mode. Signed codes of oral languages can be useful for learning oral languages or for expressing and discussing literal quotations from those languages, but they are generally too awkward and unwieldy for normal discourse. For example, a teacher and deaf student of English in the United States might use Signed English to cite examples of English usage, but the discussion of those examples would be in American Sign Language.

Several culturally well developed sign languages are a medium for stage performances such as sign-language poetry. Many of the poetic mechanisms available to signing poets are not available to a speaking poet.

1.1 List of sign languages

Sign language is not universal. Like spoken languages, sign languages emerge naturally in communities and change through time. The following list is grouped into three sections:

- Deaf sign languages, which are the preferred languages of Deaf communities around the world
- Signed modes of spoken languages, also known as Manually Coded Languages;
- Auxiliary sign systems, which are not "native" languages, but signed systems of varying complexity used in addition to native languages.

1.2 British Sign Language

British Sign Language (BSL) is the sign language used in the United Kingdom (UK), and is the first or preferred language of deaf people in the UK; the number of signers has been put at 30,000 to 70,000. The language makes use of space and involves movement of the hands, body, face and head. Many thousands of people who are not deaf also

use BSL, as hearing relatives of deaf people, sign language interpreters or as a result of other contact with the British deaf community.

2. Literature review

There are various approaches that have been used for converting sign language images into text or speech.

The threshold model with Conditional Random Field (CRF) is an excellent mechanism for distinguishing between vocabulary signs and non sign patterns (which include out-of vocabulary signs and other movements that do not correspond to signs). A short-sign detector, a hand appearance-based sign verification method, and a sub-sign reasoning method are included to improve sign language spotting accuracy. [2]

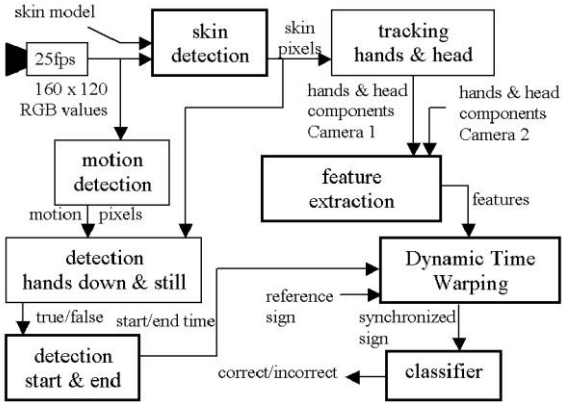


Fig. 1 Block diagram of sign detection [3]

Another method is automatic sign recognition. Its unique features are an adaptive skin model, DTW on a reference sign for synchronization, robust recognition method which is real-time and person-independent statistics, automatic feature selection for finding the best sign representation and a tolerance parameter TF that changes the behaviour of the base classifiers instead of the threshold on the total likelihood. DTW was used only for finding the best path, to synchronize the signal as in Figure 1. The method is able to generalise well over different persons, which is troublesome for many other systems. [3]

In another technique computer vision method has been used for recognizing sequences of human-hand gestures within a gloved environment. Vectors are utilized for representing the direction and displacement of the fingertips for the gesture. Modeling gestures as a set of vectors with a motion key allows the reduction of complexity in modern form and matching, which may otherwise contain multiple and lengthy datasets. [4]

The hand shape was used in recognizing people with high accuracy. It was believed that the scorecard of the hand geometry modality could be promoted to "high" in the distinctiveness and performance attributes of person recognition in that the interface is user-friendly and it is not subject to variability to the extent faces are under confounding factors of accessories, illumination effects and expression. [10] Preliminary tests indicate that hand biometric accuracy is maintained over a span of time. For any hand-based recognition scheme, it is imperative, however, that the hand image be pre processed for normalization so that hand attitude in general, and fingers in particular be aligned to standard positions[5]

Another comprehensive approach to robust visual sign language recognition system aims to signer-independent operation and utilizes a single video camera for data acquisition to ensure user friendliness. In order to cover all aspects of sign languages, sophisticated algorithms were developed that robustly extract manual and facial features, also in uncontrolled environments. The classification stage is designed for recognition of isolated signs as well as of continuous sign language. For statistical modelling of reference models, a single sign can be represented either as a whole or as a composition of smaller subunits—similar to phonemes in spoken languages. In order to overcome the problem of high interpersonal variance, dedicated adaptation methods known from speech recognition were implemented and modified to consider the specifics of sign languages. [6]

A novel algorithm to extract signemes, i.e. the common pattern representing a sign, from multiple long video sequences of American Sign Language was implemented. A signeme is a part of the sign that is robust to the variations of the adjacent signs and the associated movement epenthesis. Iterative Conditional Modes (ICM) to sample the parameters, i.e. the starting location and width of the signeme in each sentence in a sequential manner were used. In order to overcome the local convergence problem of ICM, it was run repetitively with uniformly and independently sampled initialization vectors. The results on ASL video sequences that do not involve any magnetic trackers or gloves, and also on a corresponding audio dataset were shown. [7]

Yet in another approach, an application's speech and audio output is translated into text using existing speech-to-text conversion programs. The system translates key text words

or phrases into the appropriate sign language. For this translation, pre captured gesture database and Java 3D were used to construct the simple 3D hand model, achieving a rich, interactive, animated environment focusing more on the hand's degrees of freedom (DOF) rather than texture. However this approach focuses on communicating by hand gestures that can be captured by only fingers and palms. For gestures that require other hand motions, such as wrist rotations and hand translations or facial, expressions, more data needs to be incorporated into the system [11].

A demonstrator for generating VRML animation sequences from Sign Language notation, based on MPEG-4 Body Animation has been developed. The system is able to convert almost all hand symbols as well as the associated movement, contact and movement dynamics symbols contained in any ASL sign-box. [12]

3. Problem definition

Sign language is a non-verbal language used by the hearing impaired people for everyday communication among themselves. It is not just a random collection of gestures; it is a full-blown language in its own right, complete with its own grammatical rules.

Linguistic research has to find easy but efficient strategies for the real-time adaptation of the wording in order to make a message understandable also for an audience with limited language proficiency. In order to improve communication between deaf and hearing people, more exhaustive research in automatic sign language recognition is needed. Research on human–computer interaction could also benefit from gesture and mimic analysis algorithms, originally developed for sign language recognition systems.

Euclidean distance is the "ordinary" distance between two points that one would measure with a ruler, and is given by the Pythagorean formula. On the basis of this distance between images of various signs, they are recognized for the correct output. In order to simplify the computation of this distance, the images for signs are converted into a binary format.

The pattern of the sign in binary form is carefully observed and they are then classified on the basis of Multi Layer Perceptron architecture which is a feed forward artificial neural network model that maps sets of input data onto a set of appropriate output as in Figure 2.

It uses ANN which is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. They are usually used to model complex relationships between inputs and outputs or to find patterns in data.

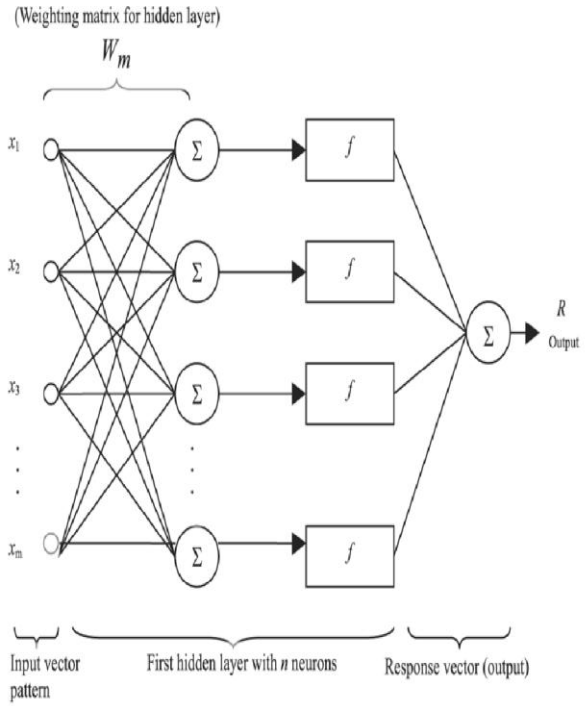


Figure 2: Multi Layer Perceptron Architecture

4. Proposed methodology

This converter recognizes the signed images made by the signer and converts them into text as well as speech. The process followed is described in Figure 3.

In this approach, five sample images per alphabet were taken in a controlled environment. These images were stored in a database. After that they were converted into LAB format as it is considered the most accurate format and can be used as an intermediary for color space conversions. Then the images were converted in binary form and 10X10 blocks were imposed on each binary image. After this the number of black pixels was found in each block. Based on this, an average number of black pixels for all the blocks from all the samples were calculated for each sign.

The signed image is then captured and it undergoes a same process of conversion from RGB to LAB to binary form. Then the number of black pixels for each block is computed and saved.

After this the Euclidean distance between the signed input image and those in the database is calculated. And on the basis of this the image is matched for a particular sign which is then displayed and converted into speech.

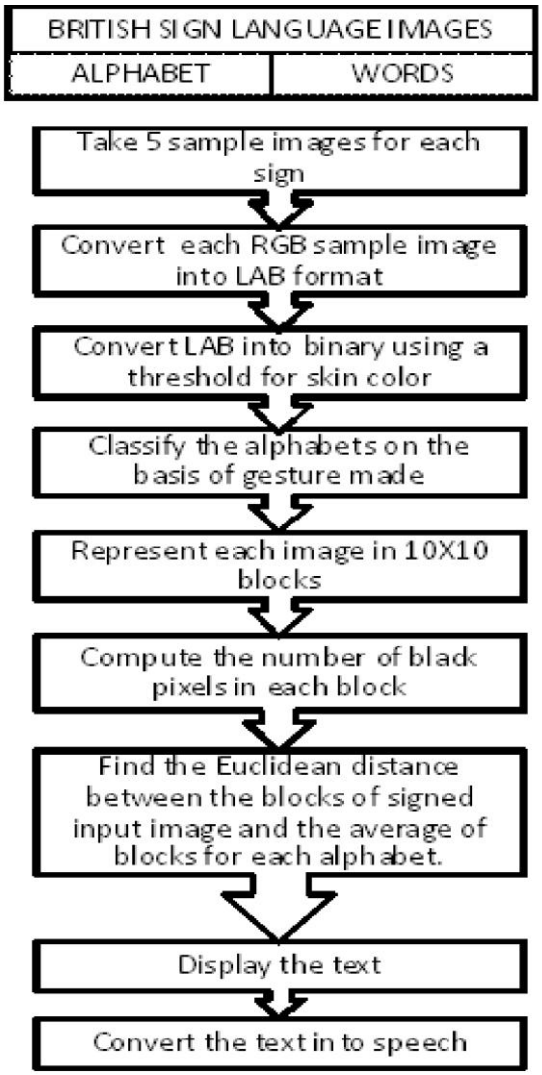


Fig. 3 Steps Followed to convert signed image into speech.

The above algorithm was implemented under the following constraints:

1. The camera is at a fixed position and at a fixed distance from the signer.
2. The signs are made in a controlled environment keeping a fixed background.
3. The RGB images are first converted into LAB and then into binary so as to reduce the distortion.
4. Only static signs have been used, i.e., there should be no movement of hands to depict a sign.
5. The size of image is kept constant.

5. Conclusions

Sign language recognition and translation is an active area of research. People with limited fluency in sign language can easily communicate with hearing impaired people with the converter that has been proposed in this paper. As this converter recognizes the signed images made by the signer and converts them into text as well as speech without any use of data gloves or other equipment. Thus, interaction gets simplified between people with or without hearing or speech impairments.

For further work, videos of hand gesture could be captured and recognized through the implementation of the same algorithm.

References

[1] Sign Language Recognition and Translation: A Multidiscipline Approach from the Field of Artificial Intelligence, Becky Sue Parton .

[2] Sign Language Spotting with a Threshold Model Based on Conditional Random Fields Hee-Deok Yang, Member, IEEE, Stan Sclaroff, Senior Member, IEEE, and Seong-Whan Lee, Senior Member, IEEE

[3] SIGN LANGUAGE DETECTION USING 3D VISUAL CUES J.F. Lichtenauer G.A. ten Holt E.A. Hendriks M.J.T. Reinders Information and Communication Theory Group Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS) Delft University of Technology, The Netherlands

[4] Visual Gesture Recognition by J. Davis and M. Shah

[5] Shape-Based Hand Recognition by Erdem Yörük, Ender Konuko˘glu, Bülent Sankur, Senior Member, IEEE, and Jérôme Darbon

[6] Recent developments in visual sign language recognition by Ulrich von Agris Æ Jo¨rg Zieren Æ Ulrich Canzler Æ Britta Bauer Æ Karl-Friedrich Kraiss

[7] Automated Extraction of Signs from Continuous Sign Language Sentences using Iterated Conditional Modes by Sunita Nayak, Sudeep Sarkar, Barbara Loeding

[8] Sign Tutor: An Interactive System for Sign Language Tutoring Oya Aran, Ismail Ari, Lale Akarun, and Bu¨lent Sankur

[9] A Multimodal Framework For The Communication Of The Disabled by Savvas Argyropoulos , Konstantinos Moustakas , Alexey A. Karpov , Oya Aran, Dimitrios Tzovaras , Thanos Tsakiris , Giovanna Varni , Byungjun Kwon

[10] A New Color Image Database for Benchmarking of Automatic Face Detection and Human Skin Segmentation Techniques
Abdallah S. Abdallah, Mohamad A bou El-Nasr, and A. Lynn Abbott

[11] Hand-Gesture Computing for the Hearing and Speech Impaired Gaurav Pradhan and Balakrishnan Prabhakaran University of Texas at Dallas Chuanjun Li Brown University

[12] TEXT-TO-SIGN LANGUAGE SYNTHESIS TOOL

Maria Papadogiorgaki, Nikos Grammalidis, Dimitrios Tzovaras and Michael G. Strintzis Informatics and Telematics Institute,
Centre for Research and Technology Hellas 1st Km Thermi Panorama str., 57001, Thermi Thessaloniki, Greece phone: +(30) 2310464160, fax: +(30) 2310464164, email: mpapad@iti.gr;ngramm@iti.gr web: http://www.iti.gr/

Author Biographies

Mansi Gupta, a final year student at Lingaya's Institute of Mgt. & Tech., Faridabad, Haryana, India. Her areas of interest include Image processing, Artificial Neural Networks, Computer organization and Operating System.

Meha Garg, a final year student at Lingaya's Institute of Mgt. & Tech., Faridabad, Haryana, India. Her areas of interest include Image processing, Artificial Neural Networks, Computer organization and Operating System.

Prateek Dhawan, a final year student at Lingaya's Institute of Mgt. & Tech., Faridabad, Haryana, India. His areas of interest include Image processing, Artificial Neural Networks, Computer organization and Operating

Real Time Tactic for Knowledge Placement & Management in Wireless Sensor Networks

Arabinda Nanda*, Saroj Kumar Rout*, Amiya Kumar Rath †

* Department of Computer Science, Krupajal Engineering College, Bhubaneswar, India
† Department of Computer Science & IT, College of Engineering , Bhubaneswar, India
aru.nanda@rediffmail.com , rout_sarojkumar@yahoo.co.in , amiyaamiya@rediffmail.com

Abstract: Knowledge Placement & Management in wireless sensor networks has been a research area in recent years. Many existing Sensor Knowledge Management Systems view sensor data as a continuous stream that is sensed, filtered, processed, and aggregated as it "flows" from sensors to users. We argue that the Knowledge Management infrastructure will need substantial redesign to fully exploit the presence of local storage and processing capability in order to reduce expensive communication. In this paper we propose an architectural framework for Knowledge Placement and Management. We propose a reliable and real time tactic for Knowledge placement and achieving data integrity using self organized sensor clusters. Instead of storing information in individual cluster heads as suggested in some protocols, in our architecture we suggest storing of information of all clusters within a cell in the corresponding base station and finally we are trying to calculate Sensor Network Knowledgebase Query Processing time and space complexity.

Keywords: Cluster Head, Data Reliability, Knowledge Management, Real Time Communication, Wireless Sensor Networks.

1. Introduction

The fact of knowing about the world is known as knowledge. Data store in a database where as knowledge store in knowledge base. Knowledge is closely related with intelligence. A person having more knowledge is called highly intelligence person. The artificial intelligence area has developed techniques for representing knowledge in forms that can be exploited by computational procedures and heuristics. Database management systems research produced techniques that support the representation and management of large amounts of relatively simple
Knowledge. Knowledge management is concerned with the representation, organization, creation, acquisition, usage, and evolution of knowledge in its many forms. To build successful technologies for knowledge management, we need to further our understanding of how individuals, groups and organizations use knowledge. Given that more and more knowledge is becoming encoded in computer readable forms, we also need to build tools which can effectively search through databases, files, web sites and the like, to extract information, capture its meaning, organize it and make it available. The main goal of our research paper is to development of a Knowledge Placement & Management System (KPMS) to store knowledge. Wireless Sensor Networks (WSNs) provide a new model for sensing and collecting information from various environments, with the possible to serve many and different applications. Current WSNs typically communicate directly with a centralized controller or satellite. On the other hand, a elegant WSN consists of a number of sensors spread across a geographical area; each sensor has wireless communication capability and sufficient intelligence for signal processing and networking of the data. Accessing and processing knowledge and data produced in a WSN using a database-like approach [2]–[3] has several advantages. Sensors can be deployed in the physical environment and applications that manipulate their knowledge and data can be created, refined, and modified afterwards without any physical involvement on the sensors themselves. The data management activity performed in the network can be remotely controlled by interactively issuing queries, expressed in a high level language, which specify what knowledge and data are of interest for a certain task, and how they should be manipulated. In this paper we propose an architectural framework for reliable and real time placement and dissemination of knowledge in WSN. We focus on middleware technology, and describe details of some existing research prototypes, then address challenges and future perspectives on the middleware. This study highlights that middleware needs to provide a common interface for various functional components of WSN: Detection and knowledge collection, signal processing, knowledge and data aggregation, and notification. By integrating sensing, signal processing, and communication functions, a WSN provides a natural platform for hierarchical information processing.

2. Related Work

[4] and [1] are two proposals providing real time data delivery services in sensor networks. They are based on an idealized model that takes into account both the remaining

lifetime before a data report expires and the remaining distance the report has to travel. Given the physical distance to the destination and the lifetime of a packet, a desired velocity can be calculated. If the packet travels along a straight line towards the destination at desired speed it will not miss the deadline. They both use greedy geographical forwarding, but differ in the exact mechanism used. In RAP, the required velocity is updated at each hop to reflect the urgency of the report. A node uses multiple FIFO queues with different priorities to schedule reports of varying degrees of urgency. Each queue accepts reports of velocities within a certain range. RAP further adjusts the waiting and back-off times at MAC layer based on the priority of the report being transmitted, so that one with higher priority has greater probability of accessing the channel. SPEED is an adaptive, location-based real-time routing protocol, which can be effectively used if the location information is available in all sensor nodes and the location updates can be delivered to the source sensors regularly. In SPEED, a node actively controls the data rate to avoid congestion by maintaining a relatively stable relay-speed to each neighbor. The node measures the delay of sending a report to each neighbor using exponential weighted moving average. Given a report with velocity v to be maintained across the sensor network, it computes the speed vi of the report, if neighbor Ni is chosen as the next hop. It chooses one neighbor from the group of neighbors with vi>v to forward the report to the next node. If no such neighbor exists, the report is forwarded to the next node with some probability. Nodes in a congested region also feedback back pressure messages upstream, so that data are detoured to bypass the region.

3. Design Approach

The basic requirements for a knowledge management system mean to build a so called information workspace [4], which can be seen as a special extension of the classical information retrieval systems [5]. The differences can be summarized as follows:

- Easy knowledge management (information work) for both content creators and information consumers (no more effort is required for storing and modifying content than for retrieving information)
- The information retrieval is not restricted to term-based queries
- The collection structure or document hierarchy is not necessarily implied by terms
- There is a possibility for automatic information exchange between the actors (e.g., by agents)

WSN Constraints

- Available bandwidth;
- Processing and memory capabilities of WSN nodes;
- Availability of energy resources;
- High error rates (from 2% to 30% per link).
- Number of nodes (network size and depth).

4. Network Topology

One important property of a WSN is its diameter, that is, the maximum number of hops between any two nodes in the network. In its simplest form, a WSN forms a single hop network, with every sensor node being able to communicate directly with every other node. An infrastructure based network with a single base station forms a star network with a diameter of two. A multi-hop network may form an arbitrary graph, but often an overlay network with a simpler structure is constructed such as a tree or a set of connected stars. The topology affects many network characteristics such as latency, robustness, and capacity. The complexity of data routing and processing also depends on the topology. Given the large number of nodes and their potential placement in difficult locations, it is essential that the network is able to self organize; manual configuration is not feasible.

4.1 Density and Network Size

The effective range of the sensors defines the coverage area of a sensor node. The density of the nodes indicates the degree of coverage of an area of interest by sensor nodes. The network size affects reliability, accuracy, and data processing algorithms. The density can range from a few sensor nodes to a hundred in a region, which can be less than 10m in diameter.

4.2 Connectivity

The communication ranges and physical locations of individual sensor nodes define the connectivity of a network. If there is always a network connection (possibly over multiple hops) between any two nodes, the network is said to be connected. Connectivity is broken if the network may be partitioned occasionally. If nodes are isolated most of the time and enter the communication range of other nodes only occasionally, we say that communication is irregular. Connectivity influences the design of communication protocols and data dissemination mechanisms

4.3 Node Addressability

This indicates whether or not the nodes are individually addressable. Ex: - the sensor nodes in a parking lot network should be individually addressable, so that the locations of all the free spaces can be determined. Thus, it may be necessary to broadcast a message to all the nodes in the network. If one wants to determine the temperature in a corner of a room, then addressability may not be so important. Any node in the given region can respond.

4.4 Data Aggregation

Data aggregation is the task of data summarization while data are traveling through the sensor network. An excessive number of sensor nodes can easily congest the network, flooding it with information. The popular solution to this problem is to aggregate or fuse data within the WSN then transmits an aggregate of the data to the controller.

There are three major ways of performing data aggregation:

1. Diffusion algorithms assume that data are transmitted from one node to the next, thus propagating through the network to the destination. Along the way data may be aggregated, mostly with simple aggregation functions and assuming homogeneous data.
2. Streaming queries are based on SQL extensions for continuous querying. Here data are considered to be transient while the query is persistent.
3. Event graphs work on streams of events and compose simple events into composite events based on event algebra

4.5 Query Capability and Propagation

There are two types of addressing in sensor network; data centric, and address centric. In a data centric paradigm, a query will be sent to specific region in the network, whereas in addressing centric, the query will be sent to an individual node. A user may want to query an individual node or a group of nodes for information collected in the region. Depending on the amount of data fusion performed, it may not be feasible to transmit a large amount of the data across the network. Instead, various local sink nodes will collect data from a given area and create summary messages. A query may be directed to the sink node nearest to the desired location.

4.6 Data Dissemination

The ultimate goal of a WSN is to detect specified events of interest in a sensor field and to deliver them to subscribers. Because of the overlap in the proximity ranges of sensors, the same phenomenon might be recorded by multiple sensor nodes. Alternatively, systematic aggregation might lose all the data on the same phenomenon. End-to-end event transfer schemes that fit the characteristics of WSNs are needed, in the same way that delivery semantics of asynchronous communication, such as publish/subscribe, is needed for wired distributed systems. The electric power consumed depends substantially on how the sensor data is handled and communicated. Because the capacity of the battery of the sensor node is very limited, it is necessary to consider the extent to which the demands of applications can be met. Adaptive communication protocols (Power aware Protocols that consider power consumption, are actively being researched.

4.7 Real-Time Support

Object tracking applications may need to correlate events from different source nodes in real-time. Real-time support (e.g. a physical event must be reported within a certain period of time) may be critical in WSNs. This aspect affects time synchronization algorithms, which may be affected by the network topology and the communication mechanism deployed.

4.8 Reliability

The reliability or fault tolerance of a sensor node is modeled [2] in using the Poisson distribution to capture the probability of not having a failure within the time interval (0, t). The fault tolerance level depends on the application of the WSN.

4.9 Security

Threats to a WSN are described in [3] and classified into the following categories:

• Passive Information Gathering: If communications between sensors or between sensors and intermediate nodes or collection points are in clear, then an intruder with an appropriately powerful receiver and well designed antenna can passively pick off the data stream.
• Subversion of a Node: If a sensor node is captured it may be tampered with, interrogated electronically and perhaps compromised. Once compromised, the sensor node may disclose its cryptographic keying material, and access to higher levels of communication and sensor functionality may be available to the attacker. Secure sensor nodes, therefore, must be designed to be tamper proof and should react to tampering in a fail-complete manner where

cryptographic keys and program memory are erased. Moreover, a secure sensor needs to be designed for its emanations not causing sensitive information to leak from it.

• False Node: An intruder might add a node to a system and feed false data or block the passage of true data. Typically, a false node is a computationally robust device that impersonates a sensor node. While such problems with malicious hosts have been studied in distributed systems, as well as ad-hoc networking, the solutions proposed there (group key agreements, quorums and per-hop authentication) are in general too computationally demanding for sensors.

• Node Malfunction: A node in a WSN may malfunction and generate inaccurate or false data. Moreover, if the node serves as an intermediary, forwarding data on behalf of other nodes, it may drop or garble packets in transit. Detecting and culling these nodes from the WSN becomes an issue.

• Node Outage: If a node serves as an intermediary or collection and aggregation point, what happens if the node stops functioning? The protocols employed by the WSN need to be robust enough to mitigate the effects of outages by providing alternate routes.
• Message Corruption: Attacks against the integrity of a message occur when an intruder inserts itself between the source and destination and modifies the contents of a message.

4.10 Knowledge accumulation and dissemination

In real meaning, management can be said to be based on the gathering and transforming knowledge into marketable products and services. Thus the management of this knowledge collection and transformation is critical to success. A good starting point for the exploration of this area is Wiig's (1995) three volume discourse "Knowledge Management Methods". Of especial interest to us was the first volume describing the way in which people and organizations create, represent, disseminate and use knowledge? Wiig argues that it is essential that the key knowledge domains are identified and incorporated within the organization's activities and structures. Wiig defines this as the concept of Critical Knowledge Function (CKF) and goes on to describe various CKFs such as Business, Constraints, Vulnerabilities, Opportunities etc. Wiig describes this panoply of CKFs as "multiple CKFs" leading to complexity escalation, causing problems in the management and organization of people and activities in order to extract maximum value from the knowledge base. The problem of identifying the key areas of knowledge conversion and dissemination has also been addressed by Hall and Andriani (1998) in their knowledge capability framework.

This growth of internal complexity is extended with external influences. Markets shift, uncertainty dominates, technologies proliferate, competition multiplies and products and services have shorter life cycles. Thus knowledge management becomes key to organizational development, learning and innovation (McElroy (2003)). At the heart of this organizational development is the relationship between tacit knowledge (informal), which is held within individuals and explicit knowledge (formal), which is articulated, codified and distributed (Davenport & Prusak (2000), Nonaka & Ta-kuechi (1995)). Knowledge creation, dissemination and use within an organization are dependent on the conversion of tacit to explicit knowledge. Nonaka and Takuechi (1995) propose four ways in which this happens:

- Socialisation (tacit to tacit)
- Externalisation (tacit to explicit)
- Internalisation (explicit to tacit)
- Combination (explicit to explicit).

Whilst technology plays a role in this conversion (e.g. use of an Intranet system), it is essentially a human based process. Von Krogh et al (2000) argue that knowledge exchange and development is best achieved formally through clear organizational goals. In contrast SainteOnge and Wallace (2003) assert that this conversion issue is best addressed informally through human processes with groups of like-minded individuals coming together in Communities of Practice (CoP) to share knowledge. It is this informal aspect that we have tried to develop within our complexity model (see later). As if ensuring that the organization "knows what it knows" is not complex enough, defining what it "needs to know but does not know" adds yet another layer of complexity. Koulopoulos (1997) proposes the concept of knowledge gaps, which can be evaluated by using the "knowledge chain" comprising:

- Internal awareness leading to
- Internal responsiveness
- External awareness leading to
- External responsiveness.

Again there is both a formal (e.g. patent searching) and informal (e.g. networking) elements to this problem. What is essential in either case is the ability for the organization/individuals to learn, develop and adapt. Within organizations there are critical knowledge functions, with people accumulating (from internal and external sources), converting and disseminating knowledge for practical

application. Some of this is achieved through formal systems, but the scale and complexity of the processes means that much of this is achieved through informal processes such as CoPs. It is the combination of for-mal and informal (based on complexity theory) organizational structures and systems that is critical. This provides the balance between freedom and control required within a commercial operation. The informal aspects are too important to be left to chance and need to be encouraged and "incorporated" into formal organizational processes.

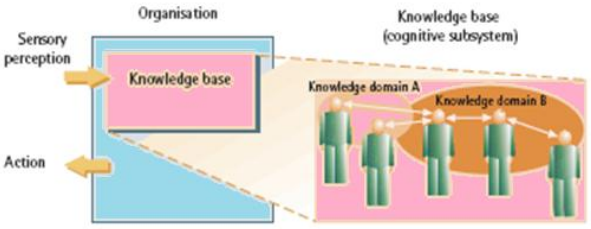


Fig-3 Organizational Knowledge

5. System Structure

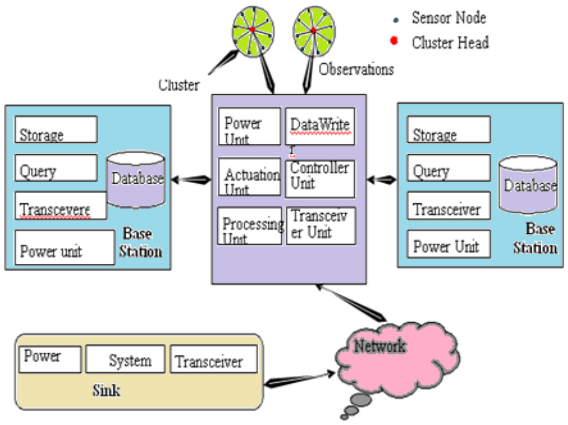


Fig-1 System Structure

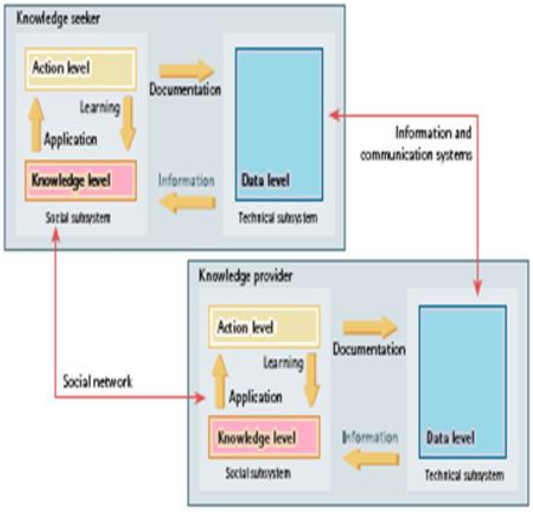


Fig-4 Knowledge transfer between knowledge seeker and knowledge provider

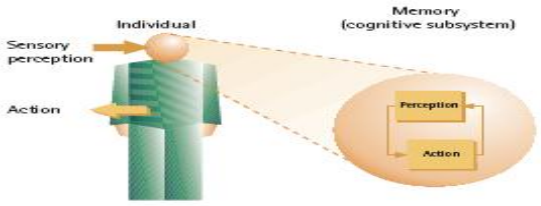


Fig-2 Individual Knowledge

The main purpose of storing the entire information of the cell in the base station is that better data integrity can be achieved as ARS need not spend time on data aggregation before taking an effective decision.

The main advantage of replication is that in case due to some severe disaster some base station collapses. The ARS can obtain the information from the adjacent cell's base station. Our main aim is to prevent the collapse of the communication system.

The functional components of the base station can be broadly divided into storage manager component, query processor component, transceiver unit and power unit. The storage manager component provides the interface between the low level data stored in the database and queries submitted to the system. Its main goal is to simplify and facilitate access to data. The storage manager component may include authorization and integrity manager, file manager, buffer manager etc. The query manager

component handles queries received from ARSs; it may further include DML compiler, metadata manager, query evaluation engine etc.

5.1 Action and Relay Station

The ARS are resource rich nodes equipped with better processing capabilities, higher transmission powers and longer battery life. The ARS nodes are placed on the bordering areas of cells and are responsible for data dissemination in a time efficient manner. An ARS unit consists of six basic components: an actuation unit, a processing unit, controller (decision unit), data writer, a power unit, and a transceiver unit. The decision unit functions as an entity that takes sensor readings as input and generates action commands as output. These action commands are then converted to analog signals and transformed into actions via the actuation unit. The ARS nodes are placed on each pair of shared edges along the border between two cells. The total number of seed ARS nodes needed for a N-cell system are computed using (5). Every ARS supports two types of interfaces: ad-hoc relay interface and cellular interface. By ad-hoc interface, ARS may communicate with other ARSs and sink nodes. It uses cellular interface to communicate with base stations of cellular network. During disaster any ARS may be collapsed. But there is little chance of collapsing all ARSs of a cell. Only one ARS is enough to convey data from sensor network of a cell to a base station. The data writer component of ARS can pass the information received from various CHs of WSN immediately to the corresponding base stations or it may store the information received in its local buffer and after some time send the combined information to further conserve power and communication bandwidth.

6. Sensor Network Knowledge Based Query processing time and space complexity

The issue of estimating database query space and time complexities. Initially, queries without joins are considered and classified into five categories in accordance with complexity (type and number of clauses) in a progressive manner. The storage space and execution time complexity measures for each category are then derived after translating the queries into their algebraic representations and then deriving possible relations that accounts for the different factors (i.e., clauses found in the statement). Joins were then considered and similar complexity expressions were derived.

Modes of evaluation.- Important modes of query evaluation: *filtering*, *full-fledged*, and *pattern matching*. The primary notion in all modes is a *match*. For a query Q and a node $r \in Q$, we denote by Qr the sub-query rooted at r. Similarly, for a document D and an element $e \in D$, we denote by D the sub-document rooted at o.

6.1 Query Complexity

On the World Wide Web, effective query processing used to be impossible due to the lack of data structures and scheme information. The time is proportional to the size of QUERY, the sum of the times required for each iteration:

$$\sum_{i=1}^{x} \mathbf{QUERY}$$

The change in QUERY with each iteration is dependent on the data structure, QUERY grows exponentially, and the total number of iterations equals log2C, where C is the number of concepts in E-R (Entity Relation) Concepts that are related to one another in E-R Relationships. For a list, QUERY remains constant, while the total number of iterations equals C. But any particular concept will be queried for atmost once, and every concept can potentially be related to the concept(s) of interest.
The time when the size of QUERY is one tuple is equal to Q, then the upper bounds on the time required for the algorithm to terminate becomes:

$$C*Q.$$

For large databases that are searched on indexed fields, Q may be approximately on the order of Log R, where R is the number of tuples in the table being searched. Q, being constant, is dropped, resulting in:

$$0(C).$$

The amount of space required to run the Query O(R), where R equals the number of tuples in table E-R Relationships, since TUPLES may equal the size of R, but cannot exceed R*(transitive closure).

7. Performance Evaluation

The performance is measured by average energy indulgence, system lifetime, successful data delivery and

number of live nodes. To improve the longevity of the system it may be desired that only minimal set of sensors sense and report the environment. It is expected that our architecture will achieve greater energy savings, enhanced availability and fault tolerance. The scalability is very easy to achieve is our architecture and it will prolong network operation lifetime.

8. Conclusion

We assert that the best implementation for reliability in Real time Knowledge Placement & Management in Wireless Sensor Networks architectures involves both the knowledge based Query processing and network management. We propose a reliable and real time tactic for knowledge placement and achieving data integrity using self organized sensor clusters. In our proposed architecture we suggest storing of information of all clusters within a cell in the corresponding base station.

References

[1] C. Lu, B. M. Blum, T. F. Abdelzaher, J. A. Stankovic, and T. He. "RAP:A real-time communication architecture for large-scale wireless sensor networks," in 8th IEEE Real-Time and Embedded Technology and Applications Symposium, 2002, pp. 55- 66.

[2] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "The design of an acquisitional query processor for sensor networks", Proc. of the 2003 ACM SIGMOD Int. Conf. on Management of Data, 2003, pp. 491- 502.

[3] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "TAG: a tiny aggregation service for ad-hoc sensor networks", Proc. of the 5th symposium on Operating Systems Design and Implementation, vol. 36, 2002, pp. 131 – 146.

[4] T. He, J.A. Stankovic, C. Lu, T. Abdelzaher. "SPEED: A stateless protocol for real-time communication in sensor networks." Proc. of 23rd Int. Conf. on Distributed Computing Systems (ICDCS'03), Providence, Rhode Island, USA, May, 2003

[5] Y. Yao and J. E. Gehrke, "The Cougar approach to in-network query processing in sensor networks", ACM Sigmod Record, vol. 31, no. 3.

Authors Profile

Prof. Arabinda Nanda: Received M. Tech (CS) from Utkal University in the year 2007.Currently working as Assistant Professor in the Department of Computer Science & Engineering at Krupajal Engineering College, Bhubaneswar, Orissa, India. Contributed more than 12 research level papers to many National and International journals and conferences. Having research interests include Sensor Network, Adhoc Network, Soft Computing, Artificial Intelligence and Data Mining.

Prof. Saroj Kumar Rout: Received M. Tech (CS) from Utkal University in the year 2007.Currently working as Assistant Professor in the Department of Computer Science and Engineering at Krupajal Engineering College, Bhubaneswar, Orissa, India. Contributed more than 08 research level papers. Having research interests include Sensor Network, Adhoc Network, Embedded System, and Network Distributed System.

Prof (Dr) Amiya Kumar Rath: Obtained B.E. degree in Computer Science & Engg. from Marathwada University, Maharastra in the year 1990, MBA degree in Systems Management from Shivaji University in the year 1993, M.Tech in Computer Science from Utkal University in year 2001 and Ph.D in Computer Science in the year 2005 from Utkal University for the work in the field of Embedded system. Served in various positions in different premier institutes namely College of Engineering, Osmanabad, Maharastra, Orissa Engineering College, Bhubaneswar, Kalinga Institute of Industrial technology (KIIT), Bhubaneswar, Krupajal Engineering College, and Bhubaneswar in the Department CS & IT Engg. Presently working with College of Engineering Bhubaneswar (CEB) as Professor of Computer Science & Engg. Cum Director (A&R) and is actively engaged in conducting Academic, Research and development programs in the field of Computer Science and IT Engg. Contributed more than 35 research level papers to many national and International journals. and conferences Besides this, published 4 books by reputed publishers. Having research interests include Embedded System, Adhoc Network, Sensor Network, Power Minimization, Biclustering, Evolutionary Computation and Data Mining.

Global Exponential Stability of Cellular Neural Network with Mixed Discrete and Distributed Delays

Hong-xing Yao and Jian Tang*
[1] Faculty of Science, Jiangsu University, Zhenjiang Jiangsu, 212013, P.R. China
Corresponding author Tel.: +86-511-8879 1998; fax: +86-511-8879 1467.
E-mail address:jstangjian@126.com

Abstract: This paper is concerned with analysis problem for the global exponential stability of recurrent neural networks (RNNs) with mixed discrete and distributed delays, unlike other papers, the nodes are associated with the topology of network. By using Lyapunov-Krasovskii functional and Young inequality, we give the sufficient condition of global exponential stability of the new cellular neural network, in addition, the example is provided to illustrate the applicability of the result.

Keywords: topology; Lyapunov-Krasovskii function; Young inequality; global exponential stability

1. Introduction

In 1992, Chua and Roska introduced a delay cellular neural network (CNN). Delay cellular neural network has been applied in many fields nowadays, such as moving target detection, identification and classification, further, along with the rapid development of biological information in cell simulation; image analysis has been widely used in recent years. We must first concern the stability of system, because the main function of CNN is about to change an input image into a corresponding output image. The various generalizations of neural networks have attracted attention of the scientific community due to their promising potentiality for tasks of classification, associative memory, parallel computation and the ability to difficult optimization [1-5]. Such applications rely on the existence of equilibrium points and the qualitative properties of neural networks. The time delay is commonly existed in various engineering systems such as chemical processes, hydraulic and rolling mill systems, etc. These effects are unavoidably existed in the implementation of neural networks, and may cause undesirable dynamic network behaviors such as oscillation and instability. Therefore, it is important to investigate the stability of delayed neural networks. The stability analysis of neural networks plays an important role in the designs and applications. A large number of the criteria on the stability of neural networks have been derived in the literature.

The global asymptotic stability results of different classes of delayed neural networks were proposed in [6-10, 22]. However, these results are only concerned with the asymptotic stability of networks without providing any conditions for exponential stability and any information about the decay rates of the delayed neural networks. Therefore, it is particularly important, when the exponential convergence rate is used to determine the speed of neural computations. Thus, it is important both theoretically and practically to determine the exponential stability and to estimate the exponential convergence rate for delayed neural networks. Considering this, the corresponding research results of many researchers have been reported in the literatures [11-15]. Neural network usually has a spatial nature due to the presence of various parallel pathways with a variety of axon sizes and lengths, so it is desirable to model them by introducing unbounded delays. Thus, there will be a distribution of conduction velocities along these pathways and a distribution of propagation delays. In these circumstances the signal propagation is not instantaneous and cannot be modeled with discrete delays and a more appropriate way is to incorporate continuously distributed delays in neural network models. In recent years there has been a growing research interest in study of neural networks with distributed delays. In fact, both discrete and distributed delays should be taken into account when modeling a realistic neural network [16-20].

Based on the above discussions, we consider a new mixed discrete and distributed delays cellular neural network described by a neutral integro-differential equation. The main purpose of this paper is to study the global exponential stability for neutral-type delayed neural networks. The structure of the neutral-type neural networks with distributed delays

under consideration is more general than the other papers existed in the literature. To the best of the author's knowledge, there were no global stability results for neutral-type neural networks. This paper is an attempt to this goal. By utilizing the Lyapunov-Krasovskii functional and Young inequality, we give the sufficient condition of global exponential stability of mixed discrete and distributed delays cellular neural network, In addition, the example is provided to illustrate the applicability of the result.

The paper is organized as follows: In Section 2, Problem formulation Model is stated and some definitions and lemmas are listed. Based on the Lyapunov stability theory and Young inequality, Main results and proofs about global exponential stability of multi-delay and distributed delay cellular neural network are listed in Section 3. In section 4, we give an example. We give the conclusion of this paper in Section 5.

2. Problem formulation Model

Consider the following multi-delay and distributed delay cellular neural network model：

$$\begin{cases} \dot{x}_i(t) = -d_i x_i(t) + \sum_{j=1}^{r} a_{ij} f_j(x_j(t)) + \sum_{j=1}^{r} b_{ij} g_j(x_j(t-\tau_{ij}(t))) \\ \qquad + \sum_{j=1}^{r} c_{ij} \int_{-\infty}^{t} K_{ij}(t-s) h_j(x_j(s)) ds + I_i, \quad i = 1,2,\ldots,r \\ \dot{x}_i(t) = k_i(\bar{x}_r(t)) + \sum_{j=r+1}^{n} c_{ij} x_j(t), \ \ i = r+1, r+2, \ldots n, \\ x_i(\theta) = \phi_i(\theta), \ \rho = \max(\tau_{ij}(t)), -\rho \le \theta \le 0 \end{cases} \tag{1}$$

where $\phi_i(\theta)$ is bounded and continuous in the sub [0, +∞)， n is the number of the neurons in the neural network, , the constants a_{ij}, b_{ij} and c_{ij} denote, respectively, the connection weights, the discretely delayed connection weights and the distributively delayed connection weighted, of the jth neuron on the I neuron. $\bar{x}_r(t) = (x_1(t), x_2(t), \ldots, x_r(t))^T$ denotes the state of the ith neural neuron at time t, $f_j(x_j(t))$, $g_j(x_j(t))$ and $h_j(x_j(t))$ are the activation functions of the jth neuron at time t , I_i is the external bias on the ith neuron, d_i denotes the rate with which the ith neuron will reset its potential to the resting state in isolation when disconnected from the network and external inputs. $\tau_{ij} \ge 0$ is a bounded time-varying Delay，Kerner coefficient K_{ij}:[0,∞)→[0,∞) is continuous in the sub [0,∞) and satisfies $\int_0^{\infty} K_{ij}(s)ds = 1, \quad i, j = 1,2,\ldots,n$. $k_i(\cdot) \in C[R^r, R]$ and $k_i(0)=0$, $C=(c_{ij})_{(n-r)\times(n-r)}$ are real matrixes, which denote the strength of neuron interconnections.

In this paper, we make the following assumptions and definitions for the neuron activation functions.

Definition.1 x_i^* $(i = 1,2,\ldots,n)$ is the equilibrium point of (1) associated with a given I_i (i=1,2,…,n) is said to be globally exponentially stable, if there are positive constants k > 0and $\mu > 0$ such that every solution x_i^* (i=1,2,… n) of (1) satisfies as follows

$$\left|x_i(t) - x_i^*\right| \le \mu e^{-kt} \sup_{-\rho \le \theta \le 0} \left|\phi_i(\theta) - x_i^*\right|, \forall t > 0.$$

Definition.2 $\forall\, \phi(\theta) \in C\left([-\rho, 0], R^n\right)$, we definite

$|\phi| = \max\left\{\left\|\phi(\theta)\right\| : \theta \in [-\rho, 0]\right\}$, then we can get as follows

$$\left\|\phi - x^*\right\|_r^r = \sup_{-\infty \le \theta \le 0} \sum_{i=1}^{n} \left|\phi_i(\theta) - x_i^*\right|^r, r > 1$$

Assumption.1 (A1) For $i = 1,2,\ldots,n$, the neuron activation functions in (1) satisfy

$\left|f_j(s_1) - f_j(s_2)\right| \le \alpha_j^+ \left|s_1 - s_2\right|, \ \forall s_1 \ne s_2$

$$\left|g_j(s_1)-g_j(s_2)\right| \le \beta_j^+ \left|s_1-s_2\right|,\ \forall s_1 \ne s_2$$
$$\left|h_j(s_1)-h_j(s_2)\right| \le \gamma_j^+ \left|s_1-s_2\right|,\ \forall s_1 \ne s_2$$

where $\alpha_j^+, \beta_j^+, \gamma_j^+$ are constants.

Assumption.2 (A2)
The neuron activation functions $f_j(x_j(t))$, $g_j(x_j(t))$, $h_j(x_j(t))$ j =1,2,…,n are bounded.

Remark.1 The constants $\alpha_j^+, \beta_j^+, \gamma_j^+$ in Assumption 1 are allowed to be positive or zero. Thus, the resulting activation functions could be non-monotonic, and more general than the usual sigmoid functions.

Lemma.1 [21] (Rogers-Holder Inequality)

If $p>1, \frac{1}{p}+\frac{1}{q}=1,$ and $a_k>0, b_k>0\ (k=1,2,...,n)$, Then

$$\sum_{k=1}^{n} a_k b_k \le (\sum_{k=1}^{n} a_k^p)^{\frac{1}{p}} (\sum_{k=1}^{n} b_k^q)^{\frac{1}{q}}$$

Lemma.2 [22] (Young Inequality) if e > 0, h > 0, P > 1, $\frac{1}{P}+\frac{1}{q}=1,$ then we can get as follows

$$eh \le \frac{1}{p}e^p + \frac{1}{q}h^q = \frac{1}{p}e^p + \frac{p-1}{p}h^{\frac{p}{p-1}}$$

3. Main results and proofs

Theorm.1 f_j, g_j, h_j are Lipschitz continuous and $\dot{\tau}_{ij}(t)<0$, if there are constants
$\vartheta, \omega_i, q_{ij}, n_{ij}, h_{ij}, j_{ij}, l_{ij}, p_{ij}, q_{ji}, n_{ji}, h_{ji}, j_{ji}, l_{ji}, p_{ji}, \mu_1 \in R\ (i=1,2,...,n), \upsilon_{ij}=\upsilon_{ji}=\chi_{ij}=\chi_{ji}=1$
$(i, j=r+1,...,n), \omega_i>0, \varpi_i>0, \vartheta \ge 1$(when ϑ =1 =1, we must let
$q_{ij}=n_{ij}=h_{ij}=j_{ij}=l_{ij}=p_{ij}=q_{ji}=n_{ji}=h_{ji}=j_{ji}=l_{ji}=p_{ji}=1\ (i, j=1,2,...,r),$
$\upsilon_{ij}=\upsilon_{ji}=\chi_{ij}=\chi_{ji}=1\ (i, j=r+1,...,n)$

$$-\omega_i \vartheta d_i + (\vartheta-1)\sum_{j=1}^{r} \omega_i \left|a_{ij}\right|^{\frac{\vartheta-q_{ij}}{\vartheta-1}} \left|\alpha_j^+\right|^{\frac{\vartheta-n_{ij}}{\vartheta-1}} + \sum_{j=1}^{r} \omega_i \left|a_{ij}\right|^{q_{ij}} \left|\alpha_j^+\right|^{n_{ij}} + (\vartheta-1)\sum_{j=1}^{r} \omega_i \left|b_{ij}\right|^{\frac{\vartheta-h_{ij}}{\vartheta-1}} \left|\beta_j^+\right|^{\frac{\vartheta-j_{ij}}{\vartheta-1}}$$
$$+(\vartheta-1)\sum_{j=1}^{r} \omega_i \left|c_{ij}\right|^{\frac{\vartheta-l_{ij}}{\vartheta-1}} \left|\gamma_j^+\right|^{\frac{\vartheta-p_{ij}}{\vartheta-1}} + \sum_{j=1}^{r} \omega_i \left|c_{ij}\right|^{l_{ij}} \left|\gamma_j^+\right|^{p_{ij}} + \sum_{j=1}^{r} \omega_i \left|b_{ij}\right|^{h_{ij}} \left|\beta_j^+\right|^{j_{ij}} + \sum_{j=1}^{r} \omega_i \left|c_{ij}\right|^{l_{ij}} \left|\gamma_j^+\right|^{p_{ij}} \left|\beta_j^+\right|^{\vartheta} < 0, \quad (i=1,...,r)$$
$$\varpi_i \sum_{j=r+1}^{n} [(\vartheta-1)\left|k_i\right|^{\frac{\vartheta-\upsilon_{ij}}{\vartheta-1}} + \left|k_i\right|^{\upsilon_{ij}}]\ \mu_1 < 0, \varpi_i (n-r) \sum_{j=r+1}^{n} [(\vartheta-1)\left|c_{ij}\right|^{\frac{\vartheta-\chi_{ij}}{\vartheta-1}} + \left|c_{ij}\right|^{\chi_{ij}}] + \varepsilon(n-r) < 0 \quad (i=r+1,...,n)$$

Then, the equilibrium point of multi-delay and distributed delay cellular neural network x^* is global exponential stability.

Proof. We shift the equilibrium point $x^* = (x_1^*, x_2^*, \ldots, x_n^*)^T$ of (1) to the equation

$$u(t) = x(t) - x^* = [u_1(t), u_2(t), \ldots, u_n(t)]^T$$

Thus we can get as follows

$$\begin{cases} \dot{u}_i(t) = -d_i u_i(t) + \sum_{j=1}^{r} a_{ij} f_j^0(u_j(t)) + \sum_{j=1}^{r} b_{ij} g_j^0(u_j(t-\tau_{ij}(t))) \\ \qquad + \sum_{j=1}^{r} c_{ij} \int_{-\infty}^{t} K_{ij}(t-s) h_j^0(u_j(s)) ds, \quad i = 1, 2, ..., r \\ \dot{u}_i(t) = k_i(\bar{u}_r(t)) + \sum_{j=r+1}^{n} c_{ij} u_j(t), \ i = r+1, r+2, \ldots n, \end{cases} \tag{2}$$

where

$$\begin{aligned} f_j^0(u_j(t)) &= f_j(u_j(t) + x_j^*) - f_j(x_j^*), \\ g_j^0(u_j(t)) &= g_j(u_j(t) + x_j^*) - g_j(x_j^*), \\ h_j^0(u_j(t)) &= h_j(u_j(t) + x_j^*) - h_j(x_j^*), \\ \bar{u}_r(t) &= (u_1(t), u_2(t), \ldots, u_r(t))^T \end{aligned} \tag{3}$$

Consider multi-delay and distributed delay cellular neural networks associated with the problem of nonlinear equations

$$\begin{cases} d_i x_i^* = \sum_{j=1}^{r} a_{ij} f_j(x_i^*) + \sum_{j=1}^{r} b_{ij} g_j(x_i^*) + \sum_{j=1}^{r} c_{ij} \int_{-\infty}^{t} K_{ij}(t-s) h_j(x_i^*) ds + I_i, \ i = 1, 2, \ldots, r, \\ d_i x_i^* = k_i(\bar{x}_r^*(t)) + \sum_{j=r+1}^{n} c_{ij} x_j^*(t), \ i = r+1, r+2, \ldots n, \end{cases}$$

We design the following Lyapunov functional as follows

$$V(u,t) = \sum_{i=1}^{r} \omega_i \left\{ |u_i(t)|^{\vartheta} e^{\varepsilon t} + \sum_{j=1}^{r} |b_{ij}|^{h_{ij}} |\beta_j^+|^{j_{ij}} \int_{t-\tau_{ij}}^{t} |u_j(s)|^{\vartheta} e^{\varepsilon(s+\tau_{ij}(t))} ds + \sum_{i=r+1}^{n} \varpi_i \sum_{j=r+1}^{n} |u_i(t)|^{\vartheta} e^{\varepsilon t} \right.$$
$$\left. + \sum_{j=1}^{r} |c_{ij}|^{l_{ij}} |\gamma_j^+|^{p_{ij}} |\beta_j^+|^{\vartheta} \int_{0}^{t} |u_j(s-\tau_{ij})|^{\vartheta} e^{\varepsilon s} ds \right\} + \sum_{j=1}^{r} |c_{ij}|^{l_{ij}} |\gamma_j^+|^{p_{ij}} \int_{t-\tau_{ij}}^{t} |g_j(u_j(\xi))|^{\vartheta} e^{\varepsilon(\xi+\tau_{ij}(t))} d\xi$$

By (2), we calculate the Dini upper right derivative of the solution $V(u, t)$,

$$\begin{aligned} & D^+V(u,t) \\ &= \sum_{i=1}^{r} \omega_i \left\{ e^{\varepsilon t} [\varepsilon |u_i(t)|^{\vartheta} + \vartheta |u_i(t)|^{\vartheta-1} sign(u_i(t)) \dot{u}_i(t)] \right. \\ &\quad + e^{\varepsilon t} \left[\sum_{j=1}^{r} |b_{ij}|^{h_{ij}} |\beta_j^+|^{j_{ij}} |u_j(t)|^{\vartheta} e^{\varepsilon \tau_{ij}(t)} - \sum_{j=1}^{r} |b_{ij}|^{h_{ij}} |\beta_j^+|^{j_{ij}} |u_j(t-\tau_{ij})|^{\vartheta} (1 - \dot{\tau}_{ij}(t)) \right] \\ &\quad + e^{\varepsilon t} \left[\sum_{j=1}^{r} |c_{ij}|^{l_{ij}} |\gamma_j^+|^{p_{ij}} |g_j(u_j(t))|^{\vartheta} e^{\varepsilon \tau_{ij}(t)} - \sum_{j=1}^{n} |c_{ij}|^{l_{ij}} |\gamma_j^+|^{p_{ij}} |g_j(u_j(t-\tau_{ij}))|^{\vartheta} (1 - \dot{\tau}_{ij}(t)) \right] \\ &\quad \left. + \sum_{j=1}^{r} |c_{ij}|^{l_{ij}} |\gamma_j^+|^{p_{ij}} |\beta_j^+|^{\vartheta} e^{\varepsilon t} |u_j(t-\tau_{ij})|^{\vartheta} + \sum_{i=1}^{r} \sum_{j=1}^{r} |u_j(-\tau_{ij})|^{\vartheta} \dot{\tau}_{ij}(t) \right\} \\ &\quad + e^{\varepsilon t} \sum_{i=r+1}^{n} \varpi_i \left\{ \sum_{j=r+1}^{n} \vartheta |u_i(t)|^{\vartheta-1} sign(u_i(t)) \dot{u}_i(t) + \varepsilon \sum_{j=r+1}^{n} |u_i(t)|^{\vartheta} \right\} \end{aligned}$$

By $\dot{\tau}_{ij}(t) \le 0$, therefore we can get as follows

$$
\begin{aligned}
&D^{+}V(u,t)\\
&\le \sum_{i=1}^{r} \omega_i \Big\{ e^{\varepsilon t}[\varepsilon |u_i(t)|^{\vartheta} + \vartheta |u_i(t)|^{\vartheta-1} sign(u_i(t))\dot{u}_i(t)] + e^{\varepsilon t}\left[\sum_{j=1}^{r} |b_{ij}|^{h_{ij}} |\beta_j^{+}|^{j_{ij}} |u_j(t)|^{\vartheta} e^{\varepsilon\tau_{ij}(t)} - \sum_{j=1}^{r} |b_{ij}|^{h_{ij}} |\beta_j^{+}|^{j_{ij}} |u_j(t-\tau_{ij})|^{\vartheta} \right]\\
&+e^{\varepsilon t}\left[\sum_{j=1}^{r} |c_{ij}|^{l_{ij}} |\gamma_j^{+}|^{p_{ij}} |g_j(u_j(t))|^{\vartheta} e^{\varepsilon\tau_{ij}(t)} - \sum_{j=1}^{r} |c_{ij}|^{l_{ij}} |\gamma_j^{+}|^{p_{ij}} |g_j(u_j(t-\tau_{ij}))|^{\vartheta} \right] + \sum_{j=1}^{r} |c_{ij}|^{l_{ij}} |\gamma_j^{+}|^{p_{ij}} |\beta_j^{+}|^{\vartheta} e^{\varepsilon t} |u_j(t-\tau_{ij})|^{\vartheta} \Big\}\\
&+e^{\varepsilon t} \sum_{i=r+1}^{n} \varpi_i \left\{ \sum_{j=r+1}^{n} \vartheta |u_i(t)|^{\vartheta-1} sign(u_i(t)) \left(k_i(\bar{u}_r(t)) + \sum_{j=r+1}^{n} c_{ij} u_j(t) \right) + \varepsilon \sum_{j=r+1}^{n} |u_i(t)|^{\vartheta} \right\}\\
&\le \sum_{i=1}^{r} \omega_i \Big\{ e^{\varepsilon t}[\varepsilon |u_i(t)|^{\vartheta} + \vartheta |u_i(t)|^{\vartheta-1} sign(u_i(t))(-d_i u_i(t) + \sum_{j=1}^{r} a_{ij} f_j^{\varepsilon}(u_j(t)) + \sum_{j=1}^{r} b_{ij} g_j^{\varepsilon}(u_j(t-\tau_{ij}(t)))\\
&\quad + \sum_{j=1}^{r} c_{ij} \int_{-\infty}^{t} K_{ij}(t-s) h_j^{\varepsilon}(u_j(s)) ds)] + e^{\varepsilon t}\left[\sum_{j=1}^{r} |b_{ij}|^{h_{ij}} |\beta_j^{+}|^{j_{ij}} |u_j(t)|^{\vartheta} e^{\varepsilon\tau_{ij}(t)} - \sum_{j=1}^{r} |b_{ij}|^{h_{ij}} |\beta_j^{+}|^{j_{ij}} |u_j(t-\tau_{ij})|^{\vartheta} \right]\\
&\quad + e^{\varepsilon t}\left[\sum_{j=1}^{r} |c_{ij}|^{l_{ij}} |\gamma_j^{+}|^{p_{ij}} |g_j(u_j(t))|^{\vartheta} e^{\varepsilon\tau_{ij}(t)} - \sum_{j=1}^{r} |c_{ij}|^{l_{ij}} |\gamma_j^{+}|^{p_{ij}} |g_j(u_j(t-\tau_{ij}))|^{\vartheta} \right] + \sum_{j=1}^{r} |c_{ij}|^{l_{ij}} |\gamma_j^{+}|^{p_{ij}} |\beta_j^{+}|^{\vartheta} e^{\varepsilon t} |u_j(t-\tau_{ij})|^{\vartheta} \Big\}\\
&\quad + e^{\varepsilon t} \sum_{i=r+1}^{n} \varpi_i \left\{ \sum_{j=r+1}^{n} \left(\vartheta |u_i(t)|^{\vartheta-1} k_i(\bar{u}_r(t)) + \vartheta \sum_{j=r+1}^{n} |c_{ij}| |u_i(t)|^{\vartheta-1} |u_j(t)| \right) + \varepsilon \sum_{j=r+1}^{n} |u_i(t)|^{\vartheta} \right\}
\end{aligned}
$$

By

$$
\begin{aligned}
&\vartheta \sum_{j=r+1}^{n} |c_{ij}| |u_i(t)|^{\vartheta-1} |u_j(t)| \le \vartheta \sum_{j=r+1}^{n} [|c_{ij}|^{\frac{\vartheta-\chi_{ij}}{\vartheta-1}} |u_i(t)|^{\vartheta}]^{\frac{\vartheta-1}{\vartheta}} [|c_{ij}|^{\chi_{ij}} |u_j(t)|^{\vartheta}]^{\frac{1}{\vartheta}}\\
&\le \vartheta \sum_{j=r+1}^{n} \{ \frac{\vartheta-1}{\vartheta} [|c_{ij}|^{\frac{\vartheta-\chi_{ij}}{\vartheta-1}} |u_i(t)|^{\vartheta}]^{\frac{\vartheta-1}{\vartheta}\cdot\frac{\vartheta}{\vartheta-1}} + \frac{1}{\vartheta} [|c_{ij}|^{\chi_{ij}} |u_j(t)|^{\vartheta}]^{\frac{1}{\vartheta}\cdot\vartheta} \}
\end{aligned}
\tag{4}
$$

and

$$
\begin{aligned}
&\vartheta \sum_{j=r+1}^{n} |k_i| |u_i(t)|^{\vartheta-1} \le \vartheta \sum_{j=r+1}^{n} [|k_i|^{\frac{\vartheta-\upsilon_{ij}}{\vartheta-1}} |u_i(t)|^{\vartheta}]^{\frac{\vartheta-1}{\vartheta}} [|k_i|^{\upsilon_{ij}} |u_j(t)|^{\vartheta}]^{\frac{1}{\vartheta}}\\
&\le \vartheta \sum_{j=r+1}^{n} \{ \frac{\vartheta-1}{\vartheta} [|k_i|^{\frac{\vartheta-\upsilon_{ij}}{\vartheta-1}} |u_i(t)|^{\vartheta}]^{\frac{\vartheta-1}{\vartheta}\cdot\frac{\vartheta}{\vartheta-1}} + \frac{1}{\vartheta} [|k_i|^{\upsilon_{ij}} |u_j(t)|^{\vartheta}]^{\frac{1}{\vartheta}\cdot\vartheta} \}
\end{aligned}
\tag{5}
$$

Then, by (4) and (5), we can obtain

$$
\begin{aligned}
&D(u,t)\\
&\le e^{\varepsilon t}\sum_{i=1}^{r}\omega_i\left\{(\varepsilon-\vartheta d_i)|u_i(t)|^{\vartheta}+\vartheta\sum_{j=1}^{r}|a_{ij}|\,|\alpha_j^{+}|\,|u_i(t)|^{\vartheta-1}|u_j(t)|+\vartheta\sum_{j=1}^{r}|b_{ij}|\,|\beta_j^{+}|\,|u_i(t)|^{\vartheta-1}|u_j(t-\tau_{ij}(t))|\right.\\
&\quad+\vartheta\sum_{j=1}^{r}|c_{ij}|\,|\gamma_j^{+}|\,|u_i(t)|^{\vartheta-1}|u_j(t)|\int_0^{+\infty}K_{ij}(s)ds\Big]+\left[\sum_{j=1}^{r}|b_{ij}|^{h_{ij}}|\beta_j^{+}|^{j_{ij}}|u_j(t)|^{\vartheta}e^{\varepsilon\tau_{ij}(t)}-\sum_{j=1}^{r}|b_{ij}|^{h_{ij}}|\beta_j^{+}|^{j_{ij}}|u_j(t-\tau_{ij})|^{\vartheta}\right]\Bigg\}\\
&\quad+e^{\varepsilon t}\sum_{i=r+1}^{n}\varpi_i\left\{\vartheta\sum_{j=r+1}^{n}\{\frac{\vartheta-1}{\vartheta}[|k_i|^{\frac{\vartheta-\upsilon_{ij}}{\vartheta-1}}|u_i(t)|^{\vartheta}]+\frac{1}{\vartheta}[|k_i|^{\upsilon_{ij}}|u_j(t)|^{\vartheta}]\}(\bar{u}_r(t))\right.\\
&\quad+\vartheta(n-r)\sum_{j=r+1}^{n}\{\frac{\vartheta-1}{\vartheta}[|c_{ij}|^{\frac{\vartheta-\chi_{ij}}{\vartheta-1}}|u_i(t)|^{\vartheta}]+\frac{1}{\vartheta}[|c_{ij}|^{\chi_{ij}}|u_j(t)|^{\vartheta}]\}+\varepsilon\sum_{j=r+1}^{n}|u_i(t)|^{\vartheta}\Bigg\}\\
&\le e^{\varepsilon t}\sum_{i=1}^{r}\omega_i\left\{(\varepsilon-\vartheta d_i)|u_i(t)|^{\vartheta}+\vartheta\sum_{j=1}^{r}|a_{ij}|\,|\alpha_j^{+}|\,|u_i(t)|^{\vartheta-1}|u_j(t)|+\vartheta\sum_{j=1}^{r}|b_{ij}|\,|\beta_j^{+}|\,|u_i(t)|^{\vartheta-1}|u_j(t-\tau_{ij}(t))|\right.\\
&\quad+\vartheta\sum_{j=1}^{r}|c_{ij}|\,|\gamma_j^{+}|\,|u_i(t)|^{\vartheta-1}|u_j(t)|\Big]+\left[\sum_{j=1}^{r}|b_{ij}|^{h_{ij}}|\beta_j^{+}|^{j_{ij}}|u_j(t)|^{\vartheta}e^{\varepsilon\tau_{ij}(t)}-\sum_{j=1}^{r}|b_{ij}|^{h_{ij}}|\beta_j^{+}|^{j_{ij}}|u_j(t-\tau_{ij})|^{\vartheta}\right]\Bigg\}\\
&\quad+e^{\varepsilon t}\sum_{i=r+1}^{n}\varpi_i\left\{\sum_{j=r+1}^{n}\{(\vartheta-1)[|k_i|^{\frac{\vartheta-\upsilon_{ij}}{\vartheta-1}}|u_i(t)|^{\vartheta}]+[|k_i|^{\upsilon_{ij}}|u_j(t)|^{\vartheta}]\}(\bar{u}_r(t))\right.\\
&\quad+(n-r)\sum_{j=r+1}^{n}\{(\vartheta-1)[|c_{ij}|^{\frac{\vartheta-\chi_{ij}}{\vartheta-1}}|u_i(t)|^{\vartheta}]+[|c_{ij}|^{\chi_{ij}}|u_j(t)|^{\vartheta}]\}+\varepsilon\sum_{j=r+1}^{n}|u_i(t)|^{\vartheta}\Bigg\}
\end{aligned}
\tag{6}
$$

We are divided into two kinds of discussion

1. when $\vartheta>1$, by Young inequality

$$
\begin{aligned}
&\vartheta\sum_{j=1}^{r}|a_{ij}|\,|\alpha_j^{+}|\,|u_i(t)|^{\vartheta-1}|u_j(t)|\\
&=\vartheta\sum_{j=1}^{r}[|a_{ij}|^{\frac{\vartheta-q_{ij}}{\vartheta-1}}|\alpha_j^{+}|^{\frac{\vartheta-n_{ij}}{\vartheta-1}}|u_i(t)|^{\vartheta}]^{\frac{\vartheta-1}{\vartheta}}[|a_{ij}|^{q_{ij}}|\alpha_j^{+}|^{n_{ij}}|u_j(t)|^{\vartheta}]^{\frac{1}{\vartheta}}\\
&\le\vartheta\sum_{j=1}^{r}[|a_{ij}|^{\frac{\vartheta-q_{ij}}{\vartheta-1}}|\alpha_j^{+}|^{\frac{\vartheta-n_{ij}}{\vartheta-1}}|u_i(t)|^{\vartheta}]^{\frac{\vartheta-1}{\vartheta}}[|a_{ij}|^{q_{ij}}|\alpha_j^{+}|^{n_{ij}}|u_j(t)|^{\vartheta}]^{\frac{1}{\vartheta}}\\
&\le(\vartheta-1)\sum_{j=1}^{r}[|a_{ij}|^{\frac{\vartheta-q_{ij}}{\vartheta-1}}|\alpha_j^{+}|^{\frac{\vartheta-n_{ij}}{\vartheta-1}}|u_i(t)|^{\vartheta}]+\sum_{j=1}^{r}[|a_{ij}|^{q_{ij}}|\alpha_j^{+}|^{n_{ij}}|u_j(t)|^{\vartheta}]
\end{aligned}
\tag{7}
$$

$$
\begin{aligned}
&\vartheta\sum_{j=1}^{r}|c_{ij}|\,|\gamma_j^{+}|\,|u_i(t)|^{\vartheta-1}|u_j(t)|\\
&=\vartheta\sum_{j=1}^{r}[|c_{ij}|^{\frac{\vartheta-l_{ij}}{\vartheta-1}}|\gamma_j^{+}|^{\frac{\vartheta-p_{ij}}{\vartheta-1}}|u_i(t)|^{\vartheta}]^{\frac{\vartheta-1}{\vartheta}}[|c_{ij}|^{l_{ij}}|\gamma_j^{+}|^{p_{ij}}|u_j(t)|^{\vartheta}]^{\frac{1}{\vartheta}}\\
&\le\vartheta\sum_{j=1}^{r}[|c_{ij}|^{\frac{\vartheta-l_{ij}}{\vartheta-1}}|\gamma_j^{+}|^{\frac{\vartheta-p_{ij}}{\vartheta-1}}|u_i(t)|^{\vartheta}]^{\frac{\vartheta-1}{\vartheta}}[|c_{ij}|^{l_{ij}}|\gamma_j^{+}|^{p_{ij}}|u_j(t)|^{\vartheta}]^{\frac{1}{\vartheta}}\\
&\le(\vartheta-1)\sum_{j=1}^{r}[|b_{ij}|^{\frac{\vartheta-h_{ij}}{\vartheta-1}}|\beta_j^{+}|^{\frac{\vartheta-j_{ij}}{\vartheta-1}}|u_i(t)|^{\vartheta}]+\sum_{j=1}^{r}[|b_{ij}|^{h_{ij}}|\beta_j^{+}|^{j_{ij}}|u_j(t-\tau_{ij}(t))|^{\vartheta}]
\end{aligned}
\tag{8}
$$

$$
\begin{aligned}
&\vartheta\sum_{j=1}^{r}\left|c_{ij}\right|\left|\gamma_j^{+}\right|\left|u_i(t)\right|^{\vartheta-1}\left|u_j(t)\right| \\
&=\vartheta\sum_{j=1}^{r}[\left|c_{ij}\right|^{\frac{\vartheta-l_{ij}}{\vartheta-1}}\left|\gamma_j^{+}\right|^{\frac{\vartheta-p_{ij}}{\vartheta-1}}\left|u_i(t)\right|^{\vartheta}]^{\frac{\vartheta-1}{\vartheta}}[\left|c_{ij}\right|^{l_{ij}}\left|\gamma_j^{+}\right|^{p_{ij}}\left|u_j(t)\right|^{\vartheta}]^{\frac{1}{\vartheta}} \\
&\le\vartheta\sum_{j=1}^{r}[\left|c_{ij}\right|^{\frac{\vartheta-l_{ij}}{\vartheta-1}}\left|\gamma_j^{+}\right|^{\frac{\vartheta-p_{ij}}{\vartheta-1}}\left|u_i(t)\right|^{\vartheta}]^{\frac{\vartheta-1}{\vartheta}}[\left|c_{ij}\right|^{l_{ij}}\left|\gamma_j^{+}\right|^{p_{ij}}\left|u_j(t)\right|^{\vartheta}]^{\frac{1}{\vartheta}} \\
&\le(\vartheta-1)\sum_{j=1}^{r}[\left|c_{ij}\right|^{\frac{\vartheta-l_{ij}}{\vartheta-1}}\left|\gamma_j^{+}\right|^{\frac{\vartheta-p_{ij}}{\vartheta-1}}\left|u_i(t)\right|^{\vartheta}]+\sum_{j=1}^{r}[\left|c_{ij}\right|^{l_{ij}}\left|\gamma_j^{+}\right|^{p_{ij}}\left|u_j(t)\right|^{\vartheta}]
\end{aligned}
\tag{9}
$$

By (6) (7) (8) and (9), we can get as follows

$$
\begin{aligned}
&D^{+}V(u,t) \\
&\le e^{\varepsilon t}\sum_{i=1}^{r}\omega_i\Big\{(\varepsilon-\vartheta d_i)\left|u_i(t)\right|^{\vartheta}+(\vartheta-1)\sum_{j=1}^{r}\left|a_{ij}\right|^{\frac{\vartheta-q_{ij}}{\vartheta-1}}\left|\alpha_j^{+}\right|^{\frac{\vartheta-n_{ij}}{\vartheta-1}}\left|u_i(t)\right|^{\vartheta}+\sum_{j=1}^{r}\left|a_{ij}\right|^{q_{ij}}\left|\alpha_j^{+}\right|^{n_{ij}}\left|u_j(t)\right|^{\vartheta} \\
&\quad+(\vartheta-1)\sum_{j=1}^{r}\left|b_{ij}\right|^{\frac{\vartheta-h_{ij}}{\vartheta-1}}\left|\beta_j^{+}\right|^{\frac{\vartheta-j_{ij}}{\vartheta-1}}\left|u_i(t)\right|^{\vartheta}+\sum_{j=1}^{r}\left|b_{ij}\right|^{h_{ij}}\left|\beta_j^{+}\right|^{j_{ij}}\left|u_j(t-\tau_{ij}(t))\right|^{\vartheta} \\
&\quad+e^{\varepsilon t}\sum_{i=r+1}^{n}\varpi_i\Bigg\{\sum_{j=r+1}^{n}\{(\vartheta-1)[\left|k_i\right|^{\frac{\vartheta-\upsilon_{ij}}{\vartheta-1}}\left|u_i(t)\right|^{\vartheta}]+[\left|k_i\right|^{\upsilon_{ij}}\left|u_j(t)\right|^{\vartheta}]\}(\overline{u}_r(t)) \\
&\quad+(n-r)\sum_{j=r+1}^{n}\{(\vartheta-1)[\left|c_{ij}\right|^{\frac{\vartheta-\chi_{ij}}{\vartheta-1}}\left|u_i(t)\right|^{\vartheta}]+[\left|c_{ij}\right|^{\chi_{ij}}\left|u_j(t)\right|^{\vartheta}]\}+\varepsilon\sum_{j=r+1}^{n}\left|u_i(t)\right|^{\vartheta}\Bigg\}
\end{aligned}
$$

$$
\begin{aligned}
&\le e^{\varepsilon t}\sum_{i=1}^{r}\omega_i\Big\{(\varepsilon-\vartheta d_i)+(\vartheta-1)\sum_{j=1}^{r}\left|a_{ij}\right|^{\frac{\vartheta-q_{ij}}{\vartheta-1}}\left|\alpha_j^{+}\right|^{\frac{\vartheta-n_{ij}}{\vartheta-1}}+\sum_{j=1}^{r}\left|a_{ij}\right|^{q_{ij}}\left|\alpha_j^{+}\right|^{n_{ij}}+(\vartheta-1)\sum_{j=1}^{r}\left|b_{ij}\right|^{\frac{\vartheta-h_{ij}}{\vartheta-1}}\left|\beta_j^{+}\right|^{\frac{\vartheta-j_{ij}}{\vartheta-1}} \\
&\quad+(\vartheta-1)\sum_{j=1}^{r}\left|c_{ij}\right|^{\frac{\vartheta-l_{ij}}{\vartheta-1}}\left|\gamma_j^{+}\right|^{\frac{\vartheta-p_{ij}}{\vartheta-1}}+\sum_{j=1}^{r}\left|c_{ij}\right|^{l_{ij}}\left|\gamma_j^{+}\right|^{p_{ij}}+\sum_{j=1}^{r}\left|b_{ij}\right|^{h_{ij}}\left|\beta_j^{+}\right|^{j_{ij}}e^{\varepsilon\rho}+\sum_{j=1}^{r}\left|c_{ij}\right|^{l_{ij}}\left|\gamma_j^{+}\right|^{p_{ij}}\left|\beta_j^{+}\right|^{\vartheta}e^{\varepsilon\rho} \\
&\quad+\sum_{j=1}^{r}\left|c_{ij}\right|^{l_{ij}}\left|\gamma_j^{+}\right|^{p_{ij}}\left|\beta_j^{+}\right|^{\vartheta}\Big\}\left|u_i(t)\right|^{\vartheta}+e^{\varepsilon t}\sum_{i=r+1}^{n}\varpi_i\Bigg\{\sum_{j=r+1}^{n}[(\vartheta-1)\left|k_i\right|^{\frac{\vartheta-\upsilon_{ij}}{\vartheta-1}}+\left|k_i\right|^{\upsilon_{ij}}]\mu_1e^{-\delta_1t}(\underbrace{1,1,\ldots,1}_{r})^{T} \\
&\quad+(n-r)\sum_{j=r+1}^{n}[(\vartheta-1)\left|c_{ij}\right|^{\frac{\vartheta-\chi_{ij}}{\vartheta-1}}+\left|c_{ij}\right|^{\chi_{ij}}]+\varepsilon(n-r)\Bigg\}\left|u_i(t)\right|^{\vartheta}<0
\end{aligned}
$$

where δ_1, μ_1 are certain constants.

2, when $\vartheta=1$, we must let

$$
\begin{aligned}
&q_{ij}=n_{ij}=h_{ij}=j_{ij}=l_{ij}=p_{ij}=q_{ji}=n_{ji}=h_{ji}=j_{ji}=l_{ji}=p_{ji}=1\,(i,j=1,2,\ldots,r), \\
&\upsilon_{ij}=\upsilon_{ji}=\chi_{ij}=\chi_{ji}=1\,(i,j=r+1,\ldots,n)
\end{aligned}
$$

By (6)

$$
\begin{aligned}
&D^{+}V(u,t)\\
&\le e^{\varepsilon t}\sum_{i=1}^{r}\omega_i\left[(\varepsilon-\vartheta d_i)+\sum_{j=1}^{r}\left|a_{ij}\right|\left|\alpha_j^{+}\right|+\sum_{j=1}^{r}\left|c_{ij}\right|\left|\gamma_j^{+}\right|+\sum_{j=1}^{r}\left|c_{ij}\right|\left|\gamma_j^{+}\right|\left|\beta_j^{+}\right|+\sum_{j=1}^{r}\left|b_{ij}\right|\left|\beta_j^{+}\right|e^{\varepsilon\tau_{ij}(t)}\right]\left|u_i(t)\right|\\
&+e^{\varepsilon t}\sum_{i=r+1}^{n}\varpi_i\left\{\sum_{j=r+1}^{n}\left|k_i\right|^{\upsilon_{ij}}\mu_1 e^{-\delta_1 t}\underbrace{(1,1,...,1)}_{r}^{T}+(n-r)\sum_{j=r+1}^{n}\left|c_{ij}\right|^{\chi_{ij}}+\varepsilon(n-r)\right\}\left|u_i(t)\right|<0
\end{aligned}
$$

where δ_1, μ_1 are certain constants, thus, we can learn that when ϑ =1, the conclusions are valid.

4. An example

Consider the following cellular neural network model

$$
\begin{cases}
\dfrac{dx_1(t)}{dt}=-5x_1(t)+\sum_{j=1}^{2}a_{1j}f_j(x_j(t))+\sum_{j=1}^{2}b_{1j}g_j(x_j(t-\tau_{ij}(t)))+0.1\int_{-\infty}^{t}e^{-(t-s)}\tanh(x_1(s))ds\\
\qquad -0.3\int_{-\infty}^{t}2e^{-(t-s)}\tanh(x_2(s))ds+2\\
\dfrac{dx_2(t)}{dt}=-4x_2(t)+\sum_{j=1}^{2}a_{2j}f_j(x_j(t))+\sum_{j=1}^{2}b_{2j}g_j(x_j(t-\tau_{ij}(t)))-0.1\int_{-\infty}^{t}2e^{-(t-s)}\tanh(x_1(s))ds\\
\qquad +0.5\int_{-\infty}^{t}e^{-(t-s)}\tanh(x_2(s))ds+3\\
\dfrac{dx_3(t)}{dt}=k_3(\overline{x}_2(t))+c_{33}x_3(t)
\end{cases}
$$

where $K=\left\{k_{ij}\middle|c_{ij}\right\}_{2\times 2}$, $\int_0^{\infty}e^{\lambda t}K_{ij}(t)dt=k_{ij}<\infty$, $c_{33}=-1, k_3(\overline{x}_2(t))=x_1(t)$, λ is exponential convergence rate estimate.

$$A=\begin{bmatrix}1.001 & 0.801\\ 0 & -1.201\end{bmatrix},\ B=\begin{bmatrix}-2.301 & 1.720\\ 1.102 & 0\end{bmatrix}$$

Activation function as follows

$f_j(x_j(t))=g_j(x_j(t))=\frac{1}{2}(|x+1|-|x-1|)$, $(j=1,2)$, so $\left|\alpha_j^{+}\right|=\left|\beta_j^{+}\right|=\left|\gamma_j^{+}\right|=1$,

If we let

$\omega_i=q_{ij}=n_{ij}=h_{ij}=j_{ij}=l_{ij}=p_{ij}=q_{ji}=n_{ji}=h_{ji}=j_{ji}=l_{ji}=p_{ji}=-\varpi_i$

$=\mu_1=\upsilon_{ij}=\upsilon_{ji}=\chi_{ij}=\chi_{ji}=1, \varepsilon=0, \vartheta=2$

We can easily obtain

$$-\omega_i\vartheta d_i+\sum_{j=1}^{n}\omega_i\left|a_{ij}\right|\left|\alpha_j^{+}\right|+\sum_{j=1}^{n}\omega_i\left|a_{ij}\right|\left|\alpha_j^{+}\right|+\sum_{j=1}^{n}\omega_i\left|b_{ij}\right|\left|\beta_j^{+}\right|+\sum_{j=1}^{n}\omega_i\left|c_{ij}\right|\left|\gamma_j^{+}\right|+\sum_{j=1}^{n}\omega_i\left|c_{ij}\right|\left|\gamma_j^{+}\right|+\sum_{j=1}^{n}\omega_i\left|b_{ij}\right|\left|\beta_j^{+}\right|<0,$$

$$\varpi_i\sum_{j=r+1}^{n}[(\vartheta-1)\left|k_i\right|^{\frac{\vartheta-\upsilon_{ij}}{\vartheta-1}}+\left|k_i\right|^{\upsilon_{ij}}]\mu_1<0,\ \varpi_i(n-r)\sum_{j=r+1}^{n}[(\vartheta-1)\left|c_{ij}\right|^{\frac{\vartheta-\chi_{ij}}{\vartheta-1}}+\left|c_{ij}\right|^{\chi_{ij}}]+\varepsilon(n-r)<0$$

Because the exponential convergence rate and specific nuclear function are not known in advance, so we can only prove the conclusion's correctness in mathematics.

5. Conclusion

A new sufficient condition is derived to guarantee the global exponential stability of the equilibrium point for multi-delay and distributed delay cellular neural network. To the best of our knowledge, compared with traditional methods, our approach is effective.

6. Acknowledgment

Project supported by the National Natural Science Foundations of China (Grant Nos 70871056), the Society Science Foundation from Ministry of Education of China (Grant 08JA790057) and the Advanced Talents' Foundation and Student's Foundation of Jiangsu University (Nos 07JDG054 and 07A075).

7. References

[1] Balasubramaniam P, Samath J A, Kumaresan N and Kumar V A, A Solution of matrix Riccati differential equation for the linear quadratic singular system using neural networks, Appl. Math. Comput.182(2006), pp.1832-1839.

[2] Balasubramaniam P, Samath Abdul J and Kumaresan N, Optimal control for nonlinear singular systems with quadratic performance using neural networks, Appl. Math. Comput.187(2007), pp.1535-1543.

[3] Bouzerdoum A and Pattison T R, Neural networks for quadratic optimization with bound constraints, IEEE. Trans. Neural Networks.4(1993), pp.293-303.

[4] Forti M and Tesi A, New conditions for global stability of neural networks with application to linear and quadratic programming problems, IEEE. Trans. Circuits Syst. I.42(1995), pp.354-366.

[5] Kennedy M P and Chua L O, Neural networks for non-linear programming, IEEE. Trans. Circuits Syst.35(1988), pp.554-562.

[6] Arik S, Global asymptotic stability of a large class of neural networks with constant time delays, Phys. Lett. A.311(2003), pp.504-511.

[7] Cao J and Zhou D, Stability analysis of delayed cellular neural networks, Neural Networks.11(1998), pp.1601-1605.

[8] Joy M, On global convergence of a class of functional differential equations with applications in neural network theory, J. Math. Anal. Appl.232(1999), pp.61-81.

[9] Van den Driessche P and Zou X, Global attractivity in delayed Hopfield neural network models, SIAM. J. Appl. Math.6(1998), pp.1878-1890.

[10] Zhang J, Global stability analysis in delayed cellular neural networks, Comput. Math. Appl.45 (2003), pp.1707-1720.

[11] Arik S, An analysis of exponential stability of delayed neural networks with time varying delays, Neural Networks.17 (2004), pp.1027-1031.

[12] Cao J and Li Q, On the exponential stability and periodic solutions of delayed cellular neural networks, J. Math. Anal. Appl.252(2000), pp.50-64.

[13] Chu T G, An exponential convergence estimate for analog neural networks with delays, Phys. Lett. A.283(2001), pp.113-118.

[14] Liu Z and Liao L, Existence and global exponential stability of periodic solution of cellular neural networks with time varying delays, J. Math. Anal. Appl.290(2004), pp.247-262.

[15] Zhang Q, Wei X and Xu J, Global exponential convergence analysis of delayed neural networks with time varying delays, Phys. Lett. A.318(2003), pp.537-544.

[16] Ruan S and Filfil R S, Dynamics of two-neuron system with discrete and distributed delays, Physica D.191(2004), pp.323-342.

[17] Zhao H, Global asymptotic stability of Hopfield neural network involving distributed delays, Neural Networks.17(2004), pp.47-53.

[18] Zhao H, Existence and global attractivity of almost periodic solution for cellular neural network with distributed delays, Appl. Math. Comput.154(2004), pp.683-695.

[19] Yao H and Jia Y, Global exponential stability of mixed discrete and distributed delays cellular neural network (under view), Jiangsu University. 2010.

[20] Zhang H G, Recursive time-delay neural networks and Dynamic Characteristics of a comprehensive analysis, www.sciencep.com, 2008.

[21] Maligranda L, Why holder's inequality should be called Rogers' inequality, Mathematical Inequalities and Applications.1(1998), pp.69-83.

[22] Chen A P, Cao J D and Huang L H, Global robust stability of interval cellular neural networks with time varying delays, Chaos Solitons and Fractals.23(2005), pp.787-799.

[23] Wang G, Research for Stability of Delayed Neural Dynamical Systems [PhD thesis] china Shenyang, Northeast University. 2006.

Adaptive Synchronization-Based Approach for Parameters Identification in Delayed Chaotic Network

Cui Jun-feng (崔俊峰) *, Yao Hong-xing(姚洪兴) and Shao Hai-jian(邵海见)
Nonlinear Scientific Research Center, Jiangsu University, Zhenjiang Jiangsu, 212013, P.R. China
*Corresponding author *E-mail address*: cjfxxm@yahoo.com.cn

Abstract: In this paper, an adaptive procedure to the problem of synchronization and parameters identification for chaotic networks with time-varying delay is introduced by combining the adaptive control and linear feedback. Especially, we consider that the equations $\dot{x}_i(t)$ (for i =r+1, r+2,…, n) can be expressed by the former $\dot{x}_i(t)$ (for i = 1, 2,…, r), which is not the same as the previous equation. This approach is also able to track the changes in the operating parameters of the chaotic networks rapidly and the speed of synchronization and parameters estimation can be adjusted. In addition, this method is quite robust against the effect of slight noise and the estimated value of a parameter fluctuates around the correct value.

Keywords: Parameters identification; synchronization- based Time-varying delay; chaotic network

1. Introduction

Chaos appears in time evolutions of some kinds of nonlinear equations [1]. Chaotic systems intrinsically defy synchronization due to the evolution of a chaotic system sensitive dependence on its initial condition. However, over the past few years, the surprising phenomenon of synchronization between coupled chaotic systems has generated much interest since the pioneer work of Pecora and Carrol [2,3]. The interest in understanding the synchronization characteristics of chaotic systems stems from its potential applications in a variety of areas, such as in secure communication, chemical and biological systems, information science, optics and so on [1,4–5]. Recently, a wide variety of approaches have been proposed for the synchronization of chaotic systems, which include PC method[2], OGY method [3], scalar driving method, coupling control, manifold-based method [6], fuzzy control, impulsive control method [7], active control, adaptive control [8–13], and time-delay feedback approach [14], etc. However, the aforementioned approaches and many other synchronization methods are valid for chaotic systems only when the systems' para--meters are known. But in many practical situations, the values of some systems' parameters cannot be exactly known a priori, and the synchronization will be destroyed and broken with the effects of these uncertainties. Therefore, this paper is devoted to the synchronization-based estimation of connection parameters by combining adaptive scheme and dynamical linear feedback control, provided the model is known. Recently, there is increasing interest in the study of dynamical properties of chaotic networks due to their potential applications in different fields such as combinatorial optimization, pattern recognition, signal, and image processing [15–17]. Most of existing works focused on the stability analysis and periodic oscillation of this kind of chaotic networks. In particular, the introduction of delays into chaotic networks may make their dynamical behaviors much more complicated even with strange attractor [18]. Recently, the chaos synchronization phenomenon for chaotic networks has drawn the attention of some researchers [19,20]. In above Refs. the synchronization schemes are proposed based on exactly knowing the concrete values of the connection weight matrices, and there is an assumption that the parameters in all coupled chaotic networks are identical.

Some synchronization-based strategies have been devised to estimate all unknown parameters of the master chaotic system. In Refs. [10, 21–25], some schemes such as auto-synchronization, random optimization, error minimization and geometric control method have been developed to recover unknown parameter values of a given model. By combining adaptive scheme and

dynamical linear feedback, an analytical scheme was proposed for estimating all unknown parameters from time series in Ref. [26-27], based on the invariance principle of differential equations. Moreover, Huang has given a more detailed proof and some interesting remarks on the adaptive-feedback control algorithm [28]. Some other authors have also used synchronization to estimate unknown parameters [29-32, 34]. However, the parameters estimation for delayed chaotic systems has not been explicitly considered and studied.

Motivated by the above discussions, in this paper, we consider the problem of parameters identification for delayed chaotic networks. we extend the method proposed in Refs. [29,30] to the chaotic networks with time-varying delay and show that this method is also effective in this case. In this paper, based on the invariance principle of functional differential equations , a dynamical feedback process is adopted, which is quite different from conventional feedback scheme. The feedback term vanishes once the synchronization and parameters identification have been achieved successfully.

This paper is organized as follows. In Section 2, the chaotic network model, and some necessary definition, hypotheses are given. In Section 3, by combining the dynamical feedback control and adaptive control, the synchronization-based scheme for parameters identification in chaotic networks with time-varying delay is described. Some numerical examples are given to verify the effectiveness of described parameter estimation scheme in Section 4. We conclude the paper in Section 5.

2. Chaotic Networks Model and Preliminaries

In this section, we consider the following chaotic network:

$$\begin{cases} \dot{x}_i(t) = f_i(x_i(t)) + \sum_{j=1}^{r} a_{ij} g_j(x_j(t-\tau)) - \gamma \sum_{j=1}^{r} b_{ij} x_j(t) & i = 1,2,\ldots r, \\ \dot{x}_i(t) = k_i(\bar{x}_r(t)) + \sum_{j=r+1}^{n} c_{ij} x_j(t) & i = r+1, r+2,\ldots n, \end{cases} \tag{1}$$

Or in a compact form

$$\begin{cases} \dot{x}_r(t) = f(x_r(t)) + Ag(x_r(t-\tau)) - \gamma B x_r(t) \\ \dot{x}_{n-r}(t) = K(x_r(t)) + C x_{n-r}(t) \end{cases} \tag{2}$$

where $x_i(t)(i = 1,2,\ldots n)$ denotes the state variable of the chaotic system, $\bar{x}_r(t) = (x_1(t), x_2(t), \ldots x_r(t))^T$, r is an integer and $1 \le r \le n, \tau(t) > 0.$ Functions $f_i(\cdot)$ and $g_i(\cdot): R \to R$ are continuous, and $f_i(0) = g_i(0) = 0.$ $\gamma > 0$ is a constant, $k_i(\cdot) \in C[R^r, R]$ and $k_i(0) = 0.$ $A = (a_{ij})_{r\times r}, B = (b_{ij})_{r\times r}$ and $C = (c_{ij})_{(n-r)\times(n-r)}$ are real matrixes and c_{ij} is negative constant, which denote the strength of neuron interconnections.

Throughout the paper, we have the following three assumptions:

(S_1) There exist nonnegative L_i and $\tilde{L}_i (i = 1,2,\ldots r)$ such that

$$|f_i(x) - f_i(y)| \le L_i |x-y|, \quad |g_i(x) - g_i(y)| \le \tilde{L}_i |x-y|,$$

For $\forall x, y \in R.$ and let $L = \max_{1\le i\le r} L_i, \tilde{L} = \max_{1\le i\le r} \tilde{L}_i.$

(S_2) $\tau(t)$ is a differential function with $0 \le \dot{\tau}(t) < 1$. Clearly, this assumption is certainly ensured if the transmission delay $\tau(t)$ is a constant.

(S_3) For $\forall x = (x_1, x_2, \ldots x_r)^T \in R^r, \|x\|$ denotes the norm of x defined by $\|x\| = (x^T x)^{1/2}$. For $\forall Q \in R^{r \times r}, \|Q\|$ indicates the norm of Q induced by $\|Q\| = (\lambda_{\max}(Q^T Q))^{1/2}$.

Lemma 1. [33] For any vector $x, y \in R^n$, and positive definite $Q \in R^{n \times n}$, the following matrix inequality holds:

$$2x^T y \le x^T Q x + y^T Q^{-1} y.$$

3. Description of the Parameters Identification Scheme

In order to observe the synchronization behavior of system (1), we introduce another chaotic network which is the response system of the drive system (1) . The behavior of the response system depends on the behavior of the drive system, but the drive system is not influenced by the response system.

$$\begin{cases} \dot{y}_i(t) = f_i(y_i(t)) + \sum_{j=1}^{r} \hat{a}_{ij} g_j(y_j(t-\tau)) - \gamma \sum_{j=1}^{r} \hat{b}_{ij} y_j(t) + u_i(t) & i = 1, 2, \ldots r, \\ \dot{y}_i(t) = \hat{k}_i(\bar{y}_r(t)) + \sum_{j=r+1}^{n} \hat{c}_{ij} y_j(t) & i = r+1, r+2, \ldots n, \end{cases} \tag{3}$$

or in a compact form

$$\begin{cases} \dot{y}_r(t) = f(y_r(t)) + \hat{A} g(y_r(t-\tau)) - \gamma \hat{B} y_r(t) - u_r(t) \\ \dot{y}_{n-r}(t) = K(y_r(t)) + \hat{C} y_{n-r}(t) \end{cases} \tag{4}$$

where $y_i(t)(i = 1, 2, \ldots n)$ denotes the state variable of the response system, $u_i(t) = \varepsilon_i \cdot e_i(t)$, $\varepsilon_i = -\alpha_i \cdot e_i^2(t)$ indicates the external control input that will be appropriately designed for a control objective, and $\bar{y}_r(t) = (y_1(t), y_2(t), \ldots y_r(t))^T$.

Let $e_i(t) = x_i(t) - y_i(t),$ the error dynamical system of (1) and (3) is

$$\begin{cases} \dot{e}_i(t) = F_i(e_i(t)) + \sum_{j=1}^{r} a_{ij} G_j(e_j(t-\tau)) + \sum_{j=1}^{r} \tilde{a}_{ij} g_j(y_j(t-\tau)) - \gamma \sum_{j=1}^{r} b_{ij} e_j(t) - \gamma \sum_{j=1}^{r} \tilde{b}_{ij} y_j(t) - u_i(t) & i = 1, 2, \ldots r, \\ \dot{e}_i(t) = K_i(\bar{e}_r(t)) + \sum_{j=r+1}^{n} c_{ij} e_j(t) + \sum_{j=r+1}^{n} \tilde{c}_{ij} y_j(t) & i = r+1, r+2, \ldots n, \end{cases} \tag{5}$$

where $F_i(e_i(t)) = f_i(x_i(t)) - f_i(y_i(t)), G_j(e_i(t-\tau)) = g_j(x_j(t-\tau)) - g_j(y_j(t-\tau)),$

$K_i(\bar{e}_r(t)) = k_i(\bar{x}_r(t)) - k_i(\bar{y}_r(t)).$

Model (5) can be rewritten as the following matrix form

$$\begin{cases} \dot{e}_r(t) = F(e_r(t)) + AG(e_r(t-\tau)) + \tilde{A}g(y_r(t-\tau)) - \gamma Be_r(t) - \gamma \tilde{B}y_r(t) - \varepsilon * e_r(t) \\ \dot{e}_{n-r}(t) = K(e_r(t)) + Ce_{n-r}(t) + \tilde{C}y_{n-r}(t) \end{cases} \tag{6}$$

$\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots \varepsilon_r)^T \in R^r$ are the updated feedback gain; and the mark $*$ is defined as $\varepsilon * e_r(t) = (\varepsilon_1 \cdot e_1(t), \varepsilon_2 \cdot e_2(t), \ldots \varepsilon_r \cdot e_r(t))^T$.

where $e_r(t) = (e_1(t), e_2(t), \ldots e_r(t))^T$, $e_{n-r}(t) = (e_{r+1}(t), e_{r+2}(t), \ldots e_n(t))^T$,

$F(e_r(t)) = (F_1(e_1(t)), F_2(e_2(t)), \ldots F_r(e_r(t)))^T$,

$G(e_r(t-\tau)) = (G_1(e_1(t-\tau)), G_2(e_2(t-\tau)), \ldots G_r(e_r(t-\tau)))^T$, $K(e_r(t)) = (K_{r+1}(e_r(t)), K_{r+2}(e_r(t)), \ldots K_n(e_r(t)))^T$.

Theorem 1. *Under the assumptions* (S_1) , (S_2) *and* (S_3) , *the dynamical feedback strength* $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots \varepsilon_r)^T$ *and the estimated parameters* $\tilde{A}, \tilde{B},$ *and* $\tilde{C}$ *are adapted according to the following updated law, respectively,* $\varepsilon_i = -\alpha_i e_i^2(t), \quad i = 1, 2, \ldots, r,$ *and*

$$\begin{cases} \dot{\tilde{a}}_{ij} = -\eta_{ij} e_i(t) g_j(y_j(t-\tau)) \quad i, j = 1, 2, \ldots r \\ \dot{\tilde{b}}_{ij} = \gamma \delta_{ij} e_i(t) y_i(t) \qquad i, j = 1, 2, \ldots r \\ \dot{\tilde{c}}_{ij} = -\gamma_{ij} e_i(t) y_i(t) \quad i, j = r+1, r+2, \ldots n \end{cases} \tag{7}$$

where $\eta_{ij} > 0, \delta_{ij} > 0, \gamma > 0 (i = 1, 2 \ldots r)$ *and* $\gamma_{ij} > 0 (i = r+1, r+2, \ldots n)$ *are arbitrary constants, respectively, then the controlled response chaotic network* (4), *and satisfies the following condition.*

$$\lim_{t \to \infty} e_i(t) = \lim_{t \to \infty} \left| c_{ij} - \hat{c}_{ij} \right| = 0 \qquad i,\ j = r+1,\ r+2, \ldots .$$

$$\lim_{t \to \infty} e_i(t) = \lim_{t \to \infty} \left| a_{ij} - \hat{a}_{ij} \right| = \lim_{t \to \infty} \left| b_{ij} - \hat{b}_{ij} \right| = 0 \quad i,\ j = 1,\ 2, \ldots .$$

Proof. Let $\tilde{A} = A - \hat{A}, \tilde{C} = C - \hat{C}$ and $\tilde{B} = B - \hat{B}$ be the estimation errors of the parameters A, C and $B,$ and subtracting Eq. (2.2) from (3.2). We can yield the error dynamical system as follows:

$$\begin{cases} \dot{e}_r(t) = F(e_r(t)) + AG(e_r(t-\tau)) + \tilde{A}g(y_r(t-\tau)) - \gamma Be_r(t) - \gamma \tilde{B}y_r(t) - \varepsilon * e_r(t) \\ \dot{e}_{n-r}(t) = K(e_r(t)) + Ce_{n-r}(t) + \tilde{C}y_{n-r}(t) \end{cases} \tag{8}$$

$\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots \varepsilon_r)^T \in R^r$ are the updated feedback gain; and the mark $*$ is defined as $\varepsilon * e_r(t) = (\varepsilon_1 \cdot e_1(t), \varepsilon_2 \cdot e_2(t), \ldots \varepsilon_r \cdot e_r(t))^T$

We design the following Lyapunov function:

$$V(e(t)) = \frac{1}{2}\sum_{i=1}^{r} e_i^T(t)e_i(t) + \frac{1}{2}\sum_{i=1}^{r}\frac{1}{\alpha_i}(\varepsilon_i - l)^2 + \frac{1}{2}\sum_{i=1}^{r}(\sum_{j=1}^{r}\frac{1}{\eta_{ij}}\tilde{a}_{ij}^2 + \sum_{j=1}^{r}\frac{1}{\delta_{ij}}\tilde{b}_{ij}^2) + \frac{1}{2}\sum_{i=r+1}^{n}\sum_{j=r+1}^{n}\frac{1}{\gamma_{ij}}\tilde{c}_{ij}^2 + \frac{1}{2(1-\sigma)}\int_{t-\tau}^{t} g^T(e(s))g(e(s))ds \quad (9)$$

where l is a constant to be determined.

Calculating the derivative of (9) along the trajectories of (8) , we have

$$\dot{V}(e(t)) = \sum_{i=1}^{r} e_i^T(t)\dot{e}_i(t) - \sum_{i=1}^{r}(\varepsilon_i - l)\dot{\varepsilon}_i + \sum_{i=1}^{r}\sum_{j=1}^{r}\tilde{a}_{ij}\dot{\tilde{a}}_{ij} + \sum_{i=1}^{r}\sum_{j=1}^{r}\tilde{b}_{ij}\dot{\tilde{b}}_{ij} + \sum_{i=r+1}^{n}\sum_{j=r+1}^{n}\tilde{c}_{ij}\dot{\tilde{c}}_{ij} + \frac{1}{2(1-\sigma)}g^T(e(t))g(e(t)) - \frac{1-\dot{\tau}}{2(1-\sigma)}g^T(e(t-\tau))g(e(t-\tau))$$

$$= e_r^T(t)[F(e_r(t)) + AG(e_r(t-\tau)) + \tilde{A}g(y_r(t-\tau)) - \gamma Be_r(t) - \gamma\tilde{B}y_r(t) - \varepsilon * e_r(t)] + e_{n-r}^T(t)[K(e_r(t)) + Ce_{n-r}(t) + \tilde{C}y_{n-r}(t)] + \sum_{i=1}^{r}(\varepsilon_i - l)e_i^2(t) - \sum_{i=1}^{r}\sum_{j=1}^{r}\tilde{a}_{ij}e_i(t)g_j(y_j(t-\tau)) + \sum_{i=1}^{r}\sum_{j=1}^{r}\gamma\tilde{b}_{ij}e_i(t)y_j(t) - \sum_{i=r+1}^{n}\sum_{j=r+1}^{n}\tilde{c}_{ij}e_i(t)y_j(t) + \frac{1}{2(1-\sigma)}g^T(e(t))g(e(t)) - \frac{1-\dot{\tau}}{2(1-\sigma)}g^T(e(t-\tau))g(e(t-\tau))$$

$$= e_r^T(t)[F(e_r(t)) + AG(e_r(t-\tau)) - \gamma Be_r(t)] + e_{n-r}^T(t)[K(e_r(t)) + Ce_{n-r}(t)] + \sum_{i=1}^{r} -le_i^2(t) + \frac{1}{2(1-\sigma)}g^T(e(t))g(e(t)) - \frac{1-\dot{\tau}}{2(1-\sigma)}g^T(e(t-\tau))g(e(t-\tau))$$

According to the properties of (S1) and (S2), then we can get:

$$G^T(e_r(t-\tau))G(e_r(t-\tau)) = \sum_{i=1}^{r} G_i^2(e_i(t-\tau)) \le \sum_{i=1}^{r}\tilde{L}_i^2 e_i^2(t-\tau) = \tilde{L}^2 e_r^T(t-\tau)e_r(t-\tau)$$

$$\left\|F(e_r(t))\right\|^2 \le L^2 e_r^T(t)e_r(t), \quad g^T(e(t-\tau))g(e(t-\tau)) \ge 0, \frac{1-\dot{\tau}}{2(1-\sigma)} \ge \frac{1}{2}$$

$$g^T(e(t))g(e(t)) = \sum_{i=1}^{r} g_i^2(e_i(t)) \le \sum_{i=1}^{r}\tilde{L}_i^2 e_i^2(t) = \tilde{L}^2 e_r^T(t)e_r(t)$$

and by lemma 1, we can obtain

$$e_r^T(t)AG(e_r(t-\tau)) \le \frac{1}{2}e_r^T(t)AQ_1\tilde{L}e_r(t) + \frac{1}{2}e_r^T(t-\tau)AQ_1^{-1}\tilde{L}e_r(t-\tau)$$

$$e_{n-r}^T(t)K(e_r(t)) \le \frac{1}{2}e_{n-r}^T(t)KQ_2e_{n-r}(t) + \frac{1}{2}e_r^T(t)KQ_2^{-1}e_r(t)$$

As a tool of deriving a less conservative stability criterion, we add the following one zero equation to be chosen as:

$$l_1 e_{n-r}^T(t) \times [e_r(t-\tau) - e_r(t-\tau)] = 0$$

and by lemma 1, we also have $2e_{n-r}^T(t)e_r(t-\tau) \le e_{n-r}^T(t)Q_3e_{n-r}(t) + e_r^T(t-\tau)Q_3^{-1}e_r(t-\tau)$

Thus, we will get

$$\begin{aligned}\dot{V}(t) &\le e_r^T(t)[L-\gamma B-l+\frac{1}{2}(\tilde{L}^2(1-\sigma)^{-1}+KQ_2^{-1}+AQ_1\tilde{L})]e_r(t)\\ &+e_{n-r}^T(t)(C+\frac{KQ_2}{2})e_{n-r}(t)+\frac{1}{2}e_r^T(t-\tau)(-\tilde{L}^2+AQ_1^{-1})e_r(t-\tau)\\ &+l_1e_{n-r}^T(t)Q_3e_{n-r}(t)+l_1e_r^T(t-\tau)Q_3^{-1}e_r(t-\tau)-l_1e_{n-r}^T(t)e_r(t-\tau)\\ &\le \eta^T(t)Q\eta(t)<0\end{aligned}$$

where $\eta(t)=(e_r^T(t),e_{n-r}^T(t),e_r^T(t-\tau))^T$,

$$Q=\begin{pmatrix}\Im_1 & 0 & 0\\ 0 & \Im_2 & -l_1I_r\\ 0 & 0 & \Im_3\end{pmatrix}<0$$

$$\Im_1=[L-\gamma B-l+\frac{1}{2}(\tilde{L}^2(1-\sigma)^{-1}+KQ_2^{-1}+AQ_1\tilde{L})]I_r$$

$$\Im_2=(l_1Q_3+C+\frac{KQ_2}{2})I_r,\ \Im_3=\frac{1}{2}(-\tilde{L}^2+AQ_1^{-1}+l_1Q_3^{-1})I_r$$

Q_i, (i=1,2,3) is positive matrix.

The constant l and l_1 can be properly chosen to make $\dot{V}(t)<0$

Therefore, based on the lyapunov stability theory, the errors vector $e(t)\to 0$, as $t\to 0$. Then the theorem 1 has been proofed.

It is obvious that $M=\{\dot{V}(e(t))|=0\}=\{e(t)=0\}$. Therefore the set $E=\{e(t)=0,\hat{C}=C,$ $\hat{A}=A,\hat{B}=B,\varepsilon=\varepsilon_0\}$ is the largest invariant set contained in M for system (8). In fact, if one of the following equalities cannot hold: $\hat{C}=C,\hat{A}=A,\hat{B}=B,$ then $e(t)=0$ can not be a fixed point of Eq. (8) at all, i.e., one cannot conclude that the solution $e(t)$ is equal to 0 for $t>0$ when the initial values $e(t)=0$. So according to the invariant principle of functional differential equations , starting with arbitrary initial values of Eq. (8), the trajectory converges asymptotically to the set E , i.e., $e(t)=0$, $\hat{C}=C,\hat{A}=A$ and $\hat{B}=B,$ as $t\to\infty$ .This indicates that the unknown parameters C,A and B can be successfully estimated using updated laws (5) and (6), and synchronization is achieved at the same time. This ends the proof.

In Theorem 1, we can note that the variable feedback strength ε_i is automatically adapted to a suitable strength depending on the initial values for the synchronization of chaotic networks, which is significantly different from the usual linear feedback. Whereas, by using ordinary linear feedback scheme with constant feedback strength, the unknown parameters can also be estimated with certain updated laws. In the following, a corollary will be given to show the fact.

For simplicity, we assume that all notations are the same as those mentioned before.

Corollary 1. *Under the assumptions* $(S_1),(S_2)$ and (S_3), *the fixed feedback strength* $\varepsilon=(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_r)^T$ *is large enough and the estimated parameters* $\hat{C},\hat{A},$ *and* $\hat{B}$ *are adapted according to the updated law* (7)*, then the controlled response chaotic network* (3) *is synchronized with the drive network* (1)*, and satisfies the following condition:*

$$\lim_{t\to\infty} e_i(t)=\lim_{t\to\infty}\left|a_{ij}-\hat{a}_{ij}\right|=\lim_{t\to\infty}\left|b_{ij}-\hat{b}_{ij}\right|=0 \quad i,\ j=1,\ 2,\ldots,\ r$$

$$\lim_{t\to\infty} e_i(t)=\lim_{t\to\infty}\left|c_{ij}-\hat{c}_{ij}\right|=0 \qquad i,\ j=\ r+1,\ r+\ 2,\ldots,\ n.$$

Proof. By constructing another Lyapunov function:

$$V(e(t))=\frac{1}{2}\sum_{i=1}^{r}e_i^T(t)e_i(t)+\frac{1}{2}\sum_{i=1}^{r}\Big(\sum_{j=1}^{r}\frac{1}{\eta_{ij}}\tilde{a}_{ij}^2+\sum_{j=1}^{r}\frac{1}{\delta_{ij}}\tilde{b}_{ij}^2\Big)+\frac{1}{2}\sum_{i=r+1}^{n}\sum_{j=r+1}^{n}\frac{1}{\gamma_{ij}}\tilde{c}_{ij}^2 \qquad (11)$$
$$+\frac{1}{2(1-\sigma)}\int_{t-\tau}^{t}g^T(e(s))g(e(s))ds$$

This proof is similar to the proof of Theorem 1, we can easily derive the result. Its proof is straightforward and hence omitted.

Remark 1 The scheme described in this paper can be used to identify the parameters of chaotic systems but not stable systems, because for two stable systems with identical equilibrium point, synchronization can be easily obtained even with completely different parameters and structures. However, as stated in Ref. [30] , the analysis of the results should be based on the LaSalle invariant principle, since the Lyapunov direct method can only guarantee the stability in the sense of Lyapunov but cannot guarantee asymptotic stability.

Remark 2 In practice, the linear feedback controller with large enough strength in Corollary 1 is not realizable. Although appropriate feedback strength can be ascertained by certain calculations for concrete dynamical systems, its value is different for nonidentical systems. Whereas, the feedback strength of different systems can be automatically enhanced to the required value for the synchronization of drive and response systems in Theorem 1.

Remark 3 Some sufficiently large adaptive gains

$\alpha_i, \eta_{ij}, \delta_{ij} (i, j = 1, 2, \ldots, r)$ and $\gamma_{ij} (i, j = r+1, \ldots, n)$ would lead to fast synchronization and quick parameters identification, while for sufficiently small adaptive gains, the time to achieve synchronization and parameters estimation may be quite long.

4. An example

In this section, we give a numerical example to demonstrate the effectiveness of our results.

Considering the following system:

$$\begin{cases} \dot{x}_i(t) = f_i(x_i(t)) + \sum_{j=1}^{2} a_{ij} g_j(x_j(t-\tau)) - \gamma \sum_{j=1}^{2} b_{ij} x_j(t) \quad i = 1, 2, \\ \dot{x}_3(t) = k_3(\bar{x}_2(t)) + C x_3(t), \end{cases} \tag{12}$$

where $f_i(x_i(t)) = \tanh(x_i(t)), g_i(x_i(t-\tau)) = \tanh(x_i(t-\tau)), i = 1, 2.$

$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} -\frac{1}{4} & \frac{1}{20} \\ \frac{1}{20} & -\frac{1}{10} \end{pmatrix}, B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $C = -1, k_3(\bar{x}_2(t)) = x_1(t), \gamma = 1$ and $\tau = 1.$ The system satisfies assumption S_1 with $L = \tilde{L} = L_i = \tilde{L}_i = 1, i = 1, 2.$ The initial condition $(x_1(t), x_2(t), x_3(t))^T = (0.2, 0.1, 0.3)^T$, for $t \in [-1, 0]$

The response system of network (3.1) is

$$\begin{cases} \dot{y}_i(t) = f_i(y_i(t)) + \sum_{j=1}^{2} \hat{a}_{ij} g_j(y_j(t-\tau)) - \gamma \sum_{j=1}^{2} \hat{b}_{ij} y_j(t) + u_i(t) \quad i = 1, 2, \\ \dot{y}_3(t) = \hat{k}_i(\bar{y}_2(t)) + C y_3(t). \end{cases} \tag{13}$$

where $f_i(x_i(t)) = \tanh(x_i(t)), g_i(x_i(t-\tau)) = \tanh(x_i(t-\tau)), i = 1, 2.$

$\hat{A} = \begin{pmatrix} \hat{a}_{11} & \hat{a}_{12} \\ \hat{a}_{21} & \hat{a}_{22} \end{pmatrix} = \begin{pmatrix} -\frac{1}{4} & \frac{3}{20} \\ \frac{3}{20} & -\frac{3}{10} \end{pmatrix}, \hat{B} = \begin{pmatrix} \hat{b}_{11} & \hat{b}_{12} \\ \hat{b}_{21} & \hat{b}_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $C = -1, k_3(\bar{x}_2(t)) = x_1(t), \gamma = 1$ and $\tau = 1.$ The system satisfies assumption S_1 with $L = \tilde{L} = L_i = \tilde{L}_i = 1, i = 1, 2.$ The initial condition $(y_1(t), y_2(t), y_3(t))^T = (0.4, 0.2, 0.6)^T$, for $t \in [-1, 0]$;

$\varepsilon_i(0) = \hat{a}_{ii}(0) = \hat{b}_{ii}(0) = 0, i = 1, 2,$ and parameters update gain $\eta_{ij} = 10,$. $\delta_{ij} = 10, (i, j = 1, 2),$ $\gamma_{33} = 10$. Let $e_i(t) = x_i(t) - y_i(t),$ the error dynamical system of (12) and (13) is

$$\begin{cases} \dot{e}_i(t) = F_i(e_i(t)) + \sum_{j=1}^{2} a_{ij} G_j(e_j(t-\tau)) + \sum_{j=1}^{2} \tilde{a}_{ij} g_j(y_j(t-\tau)) - \gamma \sum_{j=1}^{2} b_{ij} e_j(t) - \gamma \sum_{j=1}^{2} \tilde{b}_{ij} y_j(t) - u_i(t) \quad i = 1,2, \\ \dot{e}_3(t) = K_3(\bar{e}_2(t)) + Ce_3(t) + \tilde{C} y_3(t) \end{cases} \tag{14}$$

where $F_i(e_i(t)) = f_i(x_i(t)) - f_i(y_i(t)), G_j(e_i(t-\tau)) = g_j(x_j(t-\tau)) - g_j(y_j(t-\tau)),\ i = 1,2,$

$K_3(\bar{e}_r(t)) = k_3(\bar{x}_2(t)) - k_3(\bar{y}_2(t)),$ and $\bar{e}_2(t) = (e_1(t), e_2(t))^T$.

5. Conclusion

In summary, we have shown that a combination of synchronization based on dynamical feedback with an adaptive evolution for parameters unknown to the responser, enables the estimation of the unknown parameters for uncertain delayed chaotic network. In comparison with previous methods, time-delay is taken into account in this simple, analytical and systematic synchronization-based parameters identification scheme. In addition, it is quite robust against the effect of noise, and it is also able to rapidly track changes in the operating parameters of the experimental chaotic network. We also believe that this approach can be easily generalized to the case of other continuous and discrete time-delayed chaotic networks, and also the case of other chaotic dynamical systems with delay.

Acknowledgment

Project supported by the National Natural Science Foundations of China (Grant Nos 70871056), the Society Science Foundation from Ministry of Education of China (Grant 08JA790057) and the Advanced Talents' Foundation and Student's Foundation of Jiangsu University (Nos 07JDG054 and 07A075).

References

[1]Chen G R and Dong X 1998 *from Chaos to Order: Methodologies, Perspectives, and Applications* (Singapore: World Science Publishing Company)
[2]Pecora L M and Carroll T L 1990 *Phys. Rev. Lett.* **64** 821
[3]Ott E, Grebogi C and Yorke J A 1990 *Phys. Rev. Lett.* **64** 1196
[4]Ojalvo J G.and Roy R 2001 *Phys. Rev. Lett.* **22** 5204
[5]Zhao H Q and Zhang J S 2008 *Chinese Physics B* **8** 3996
[6]Yu X H, Chen G.R, Xia Y, Song Y X and Cao Z W 2001 *Circuits Syst*. I **8** 930
[7]P Li, Cao J D, Wang Z D 2007 *Physica A* **373** 261
[8]Xiong W J, Xie W and Cao J D 2006 *Physica A* **370** 832
[9]Cai L L and Nong L 2008 *Chinese Physics B* **12** 7500
[10]Huang D B 2005 *Phys. Rev. E* 71
[11]Li X C, Xiao Y Z and Xu W 2008 *Chinese Physics B* **8** 4721
[12]Huang D B 2004 *Phys. Rev. Lett.* 93
[13]Lei Y M and X W 2008 *Chinese Physics B* **6** 3342
[14]Cao J D, Li H X and Ho D W C 2005 *Chaos Solitons Fractals* **23** 1285
[15]Xu S Y, Lam J, Ho D W C and Zou Y 2005 *IEEE Trans.Circuits Syst.* **6** 349
[16] Chen B C, Tian Q and Zhang Y 2007 *Chinese Physics B* **1** 56
[17]Liao X X, Wang J, and Zeng Z G 2005 *IEEE Trans. Circuits Syst*.**7** 403
[18]Lu H T 2002 *Phys. Lett.* **298** 109
[19]Chen G R, Zhou J and Liu Z R 2004 *Chaos* **7** 2229
[20]Cheng C J, Liao T L and Hwang C C 2005 *Chaos Solitons Fractals* **24** 197
[21]Cao J D and Lu J Q 2006 *Chaos* 16
[22]Yu D C and Wu A 2005 *Phys. Rev. Lett.* 94
[23] Hai P, Wang X D and Yang Y X 2007 *Chinese Physics B* **1** 51
[24]Freitas U S, Macau E E N and Grebogi C 2005 *Phys. Rev. E* 71
[25] Wang X Y and Wu X J 2006 *Chinese Physics B* **2** 605
[26]Konnur R 2003 *Phys. Rev. E* 67
[27]Huang D B 2004 *Phys. Rev. E* 69
[28]Huang D B 2006 *Phys. Rev. E* 73

[29]Li S, Li R H and Xu W 2006 *Chinese Physics* **2** 598

[30]Lu J A, Tao C H, Lu J H and Liu M 2002 *Phys. Lett.* **5** 632

[31]Lu J Q and Cao J D 2007 *Phys. A* **382** 672

[32]Xiong W J, Xie W and Cao J D 2006 *Phys. A* **370** 832

[33] Lu J and Cao J 2007 *Physica A* **382** 672

[34] Cai G L and Shao H J 2009 *Chinese Physics B* **19** 060507

Implementation of Parallel Optimized ABC Algorithm with SMA Technique for Garlic Expert Advisory System

Prof. M Surendra Prasad Babu[1] N.Thirupathi Rao[2]
[1] Professor, Dept of CS & SE, Andhra University, Visakhapatnam, India
[2] M.Tech (AI & R), Andhra University, Visakhapatnam, India

Abstract- The present paper deals with the development of web based expert systems using machine learning techniques to advice the farmers in villages through online. An expert system is a computer program, with a set of rules encapsulating knowledge about a particular problem domain.

Parallel Optimized ABC Algorithm with SMA Technique is a new machine learning algorithm developed by taking ABC Algorithm as base and modified this ABC Algorithm with Shared Memory Architecture Technique (SMA Technique). As an application of this Algorithm, a new 'Garlic Expert Advisory System' was developed for advising the farmers in villages.

This system is mainly aimed at identifying the diseases for garlic crop and suggesting the farmers in the villages about disease management to obtain standardized yields. This system is developed by using JSP as front end and MYSQL as backend.

Keywords**:** Expert Systems, Machine Learning Techniques, ABC Algorithm, Shared Memory Architecture Technique, Garlic Crop, Optimization, JSP & MYSQL.

1. Introduction:

Expert systems (ES) are most popular traditional applications in the field of artificial intelligence. An Expert system can be defined as a tool for information generation from knowledge in a specific problem domain. Expert System implementations automatically perform tasks for which specially trained or talented people required. Expert systems might have extensive learning components but once the system is developed, it is proven, it can be placed in the same real world problem solving situation as the human SME (Subject Matter Expert). Typically it is an aid to human workers. The sequence of steps followed to reach any conclusion is dynamically synthesized with each new case and it is not explicitly confined to the cases, programmed when the system is built. Problem solving is accomplished by applying specific knowledge rather than specific technique. This is a key idea in expert systems technology. A wide variety of methods can be used to study the performance of an expert system.

Machine Learning is a mechanism that concerned with writing a computer program that automatically improves the knowledge with experience. It is a very young scientific discipline whose birth can be placed in the mid-seventies. The First Machine Learning Workshop was taken place in 1980 at Carnie-Mellon University (USA). The goal of machine learning is to program the computers such that to use example data or past experience to solve a given problem. There were many successful applications of machine learning exists today, including systems that analyze past sales data to predict customer behavior, recognize faces or spoken speech, optimize robot behavior so that a task can be completed using minimum resources, and extract knowledge from bioinformatics data.

1.1. ABC Algorithm:

The Artificial Bee Colony (ABC) Algorithm [1, 3, and 4], Proposed by Karaboga in 2005, is a meta-heuristic algorithm for numerical optimization. Meta-heuristics are high-level strategies for exploring search spaces. Many meta-heuristic algorithms, inspired from nature, are efficient in solving numerical optimization problems. ABC algorithm is motivated by the intelligent foraging behavior of honey bees. The ABC algorithm was first proposed for unconstrained optimization problems [5, 6]. Subsequently, the algorithm has been developed by Karaboga and Basturk and extended to constrained optimization problems. Improvements to the performance of the algorithm and a hybrid version of the

algorithm can also be seen in the literature. The ABC algorithm is a swarm-based algorithm good at solving unimodal and multimodal numerical optimization problems. It is very simple and flexible when compared to other Swarm Based algorithms such as Particle Swarm Optimization (PSO). It does not require external parameters like mutation and crossover rates, which are hard to determine in prior. The algorithm combines local search methods with global search methods and tries to attain a balance between exploration and exploitation. Researchers have come up with several real-world applications for the ABC algorithm.

1.2. About Garlic Crop:

Garlic is one of the most commonly used vegetables in India. Garlic is also known as Lassan and its botanical name is Allium Sativa Linn. It belongs to the Lilliaceae family. Garlic has germanium in it. Germanium is an anti-cancer agent, and garlic has more of it than any other herb. Another benefit of garlic is it helps regulate the body's blood pressure. Garlic is packed with vitamins and nutrients which include Protein, Potassium, Vitamins A, B1, B2 and C, Calcium, Zinc and many others. The main diseases of the garlic crop and its control measurers are given here under.

Disease1: White Rot **Cure:** Control by rotating out of Allium.

Disease2: Fusarium (basal or bottom rot) **Cure:** Proper crop rotation with non-susceptible crops for four years.

Disease3: Pink Rot **Cure:** Using at least a three- to four-year rotation without Allium.

Disease4: Botrytis **Cure:** Rapid drying during harvest.

Disease5: Penicillium Molds **Cure:** Prevent the disease by planting clean stock.

Disease6: Garlic Rust **Cure:** Use healthy seed in well-drained soil and rotate with Non-Allium crops.

Disease7: Purple Blotch **Cure:** Spraying of Mancozeb.

Disease8: Powdery Mildew **Cure:** Spraying of Sulphur Fungicides

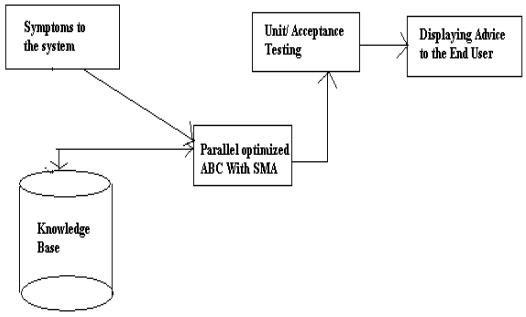
Disease9: Mosaic Disease **Cure:** Spraying of Monocrotophos

Disease10: Leaf Blight **Cure:** Spraying of Ziram or Copper Oxychloride.

Disease11: Stemphylium Blight **Cure:** Spraying of ziram or captan.

2. Proposed System:

The Architecture of the proposed Expert advisory system for garlic crop by using ABC Algorithm with SMA Technique is as follows:



2.1. Proposed Algorithm:

In general, parallel architectures may use either a shared memory or a message passing mechanism to communicate in between the multiple processing elements. Parallel metaheuristic algorithms have been developed for both these kinds of architectures. A parallel implementation of the algorithm is designed for optimized shared memory architecture. The entire colony of bees is divided equally among the available processors. Each processor has a set of solutions in a local memory. A copy of each solution is also maintained in a global shared memory. During each cycle the set of bees at a processor improves the solutions in the local memory. The output optimization can be taken as the sum of the total number of symptoms matching divided by the total number of symptoms present in the system. At the end of the cycle, the solutions are copied into the corresponding slots in the shared memory by overwriting the previous copies. The solutions are thus made available to all the processors.

The proposed Parallel ABC Optimized Algorithm with SMA Technique is as follows:

Step.1. Divide the solutions (symptoms) equally among p processors by copying SNp solutions to the local memory of each processor.

Step.2. Steps 3 to 9 are carried out in parallel at each processor Pr.

Step.3. For each solution in the local memory Mr ( ) of the range processor Pr (Processes), determine a neighbor.

Step.4. Calculate the optimization for the solutions in Mr.

Step.5. Place the onlookers on the food sources in Mr and improve the corresponding solutions.

Step.6. Determine the abandoned solution (if any) in Mr and replace it with a new randomly produced solution.

Step.7. Record the best local solution obtained till now at Pr.

Step.8. Copy the solutions in Mr to the corresponding slots in S.

Step.9. Repeat steps 4 to 9 until MCN (Maximum Cycle Number) cycles are completed.

Step.10. Determine the global best solution among the best local solutions recorded at each processor.

2.2. Database Generation:

Production rules discussed 1.2 are stored in the knowledge base in the following format.

Rule 1: **Symptoms:** S1=0, S2= 0, S3= 1, S4= 0, S5=1, S6=1, S7= 0, S8=1, S9= 0, S10= 0, S11= 0, S12= 0
Result: Disease may be D2

Rule 2:
Symptoms: S1=1, S2=0, S3=0, S4=0, S5=0, S6=0, S7=0, S8=0, S9=1, S10=0, S11=0, S12=1
Result: Disease may be D4

Rule 3:
Symptoms: S1=1, S2=1, S3= 0, S4=1, S5=1, S6=1, S7=0, S8=1, S9=0, S10=1, S11=1, S12= 0
Result: Disease may be D5.

Rule 4:
Symptoms: S1=0, S2=0, S3= 0, S4=1, S5=1, S6=0, S7=0, S8=1, S9=1, S10=1, S11=1, S12= 0
Result: Disease may be D1.

Rule 5:
Symptoms: S1=0, S2=0, S3=1, S4=1, S5=0, S6=1, S7=0, S8=1, S9=1, S10=1, S11=0, S12= 1
Result: Disease may be D2.

Rule 6:
Symptoms: S1=1, S2=1, S3=0, S4=0, S5=1, S6=0, S7=1, S8=1, S9=1, S10=1, S11=1, S12= 0
Result: Disease may be D3.

Rule 7:
Symptoms: S1=0, S2= 1, S3=0, S4=1, S5=1, S6=1, S7=0, S8=0, S9=0, S10=1, S11=0, S12=1
Result: Disease may be D1.

Rule 8:
Symptoms: S1=1, S2=1, S3=0, S4=1, S5=1, S6=1, S7=0, S8=1, S9=0, S10=1, S11=1, S12= 0
Result: Disease may be D4.

Rule 9:
Symptoms: S1=1, S2=1, S3=0, S4=1, S5=1, S6=1, S7=0, S8=1, S9=0, S10=1, S11=1, S12= 0
Result: Disease may be D2.

Rule 10:
Symptoms: S1=0, S2=1, S3=0, S4=0, S5=1, S6=0, S7=0, S8=1, S9=0, S10=0, S11=1, S12= 1
Result: Disease may be D3.

Rule 11:
Symptoms: S1=0, S2=0, S3=1, S4=1, S5=1, S6=0, S7=1, S8=1, S9=0, S10=1, S11=1, S12= 1
Result: Disease may be D4.

Rule 12:
Symptoms: S1=0, S2=1, S3=0, S4=1, S5=1, S6=0, S7=1, S8=1, S9=1, S10=1, S11=0, S12=1
Result: Disease may be D2.

Rule 13:
Symptoms: S1=1, S2=1, S3=0, S4=0, S5=1, S6=0, S7=1, S8=0, S9=1, S10=1, S11=0, S12= 1
Result: Disease may be D1.

3. Results and Discussions:

Screen shot 1: In the following screen shot the user selects the symptoms which were observed by him in the garlic crop and submits the symptoms to the expert system for processing.

Fig.1: Selection of Symptoms

This screenshot contains:

1. Was foliage brown in color? Yes or No
2. Was leaves grayish violet in color? Yes or No
3. Was the seeds violet in color? Yes or No
4. Was lesions turn violet to purple? Yes or No
5. Was the leaves fold over? Yes or No

Screen shot 2: In this screen shot the user selects the symptoms which were observed by him in the garlic crop and submits the symptoms to the expert system for processing

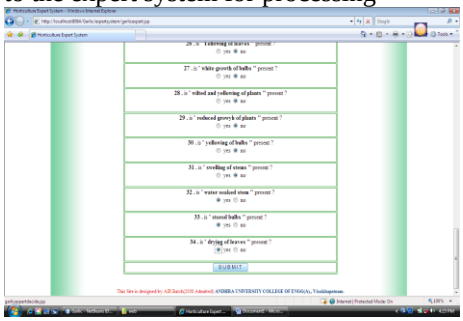


Fig.2: Selection of Symptoms

This screenshot contains:

1. Was your leaves are in Yellow? Yes or No
2. Is there any white growth in bulbs? Yes or No
3. Is there wilting of garlic plants? Yes or No
4. Reduced growth of plants? Yes or No
5. Yellowing of bulbs? Yes or No
6. Swelling of stems? Yes or No
7. Water soaked stem? Yes or No
8. Stored bulbs present? Yes or No
9. Drying of leaves present? Yes or No

Screen shot 3: In this screen shot the user can see the resultant disease affected to the garlic crop with proper cure to that particular disease as

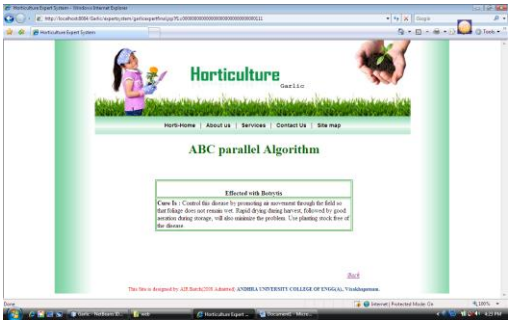


Fig: 3. Displaying Advice to the end user

affected with Botrytis and cure is rapid drying during harvest and good aeration during storage.

4. Conclusions:

In the proposed system, A Parallel Implementation of Optimized Artificial Bee Colony (ABC) Algorithm was developed which gives better results than the general ABC Algorithm. The algorithm used in the present system can be treated as quite effective; in most of the cases it finds a solution which represents a good approximation to the optimal one. Its main emphasis is to have a well designed interface for giving garlic plant related advices and suggestions to farmers by providing facilities like online interaction between expert system and the user without the need of subject expert all times. By the thorough interaction with the users and beneficiaries the functionality of the system and the algorithm can be extended further to many more areas. The results are tested by placing the system in the web portal www.bharathgramarogya.net/newindex.html.

References:

1. B.Basturk, Dervis Karaboga, An Artificial Bee Colony (ABC) Algorithm for Numeric function Optimization, IEEE Swarm Intelligence Symposium 2006, May 12-14, 2006, Indianapolis, Indiana, USA.
2. D. Karaboga, B. Basturk, A Powerful and Efficient Algorithm for Numerical Function Optimization: Artificial Bee Colony (ABC) Algorithm, Journal of Global Optimization,

Volume: 39, Issue: 3, pp: 459-471, Springer Netherlands, 2007. doi: 10.1007/s10898-007-9149.
3. D. Karaboga, B. Basturk, On The Performance Of Artificial Bee Colony (ABC) Algorithm, Applied Soft Computing, Volume 8, Issue 1, January 2008, Pages 687-697. doi:10.1016/j.asoc.2007.05.007
4. D. Karaboga, B. Basturk, Artificial Bee Colony (ABC) Optimization Algorithm for Solving Constrained Optimization Problems, LNCS: Advances in Soft Computing: Foundations of Fuzzy Logic and Soft Computing, Vol: 4529/2007, pp: 789-798, Springer- Verlag, 2007, IFSA 2007. doi: 10.1007/978-3-540-72950-1_77
5. D. Karaboga, B. Basturk Akay, Artificial Bee Colony Algorithm on Training Artificial Neural Networks, Signal Processing and Communications Applications, 2007. SIU 2007, IEEE 15th. 11-13 June 2007, Page(s):1 - 4, doi: 10.1109/SIU.2007.4298679
6. J. Vesterstrøm, J. Riget, Particle swarms extensions for improved local, multimodal and dynamic search in numerical optimization, M.Sc. Thesis, May 2002.
7. Dervis Karaboga, Bahriye Akay, A comparative study of Artificial Bee Colony algorithm. Applied Mathematics and Computation 214(1): 108-132 (2009)
8. Dervis Karaboga, Bahriye Basturk, on the performance of artificial bee colony (ABC) algorithm. Appl. Soft Comput. 8(1): 687-697 (2008) 2007.
9. D. Karaboga, An Idea Based On Honey Bee Swarm For Numerical Optimization, Technical Report-TR06, Erciyes University, Engineering Faculty, Computer Engineering Department, 2005.
10. J. Vesterstrøm, J. Riget, Particle swarms extensions for improved local, multimodal and dynamic search in numerical optimization, M.Sc. Thesis, May 2002.
11. Tereshko, V. and Loengarov, A., Collective decision-making in honey bee foraging dynamics. Comput. Inf. Syst. J. 1352-9404. v9.
12. "A Web Based Sweet Orange Crop Expert System using Rule Based System and Artificial Bee Colony Optimization Algorithm", International Journal of Engineering Science and Technology, vol.2 (6), 2010, 2408-2417.
13. A Web Based Tomato Crop Expert Information System Based On Artificial Intelligence And Machine learning algorithms, Prof.M.S.Prasad Babu, N.V. Ramana Murty, S.V.N.l.Narayana International Journal of Computer Science and Information Technologies**,** Vol. 1 (1), 2010, 6-15 (Issn: 0975-9646).
14. "A Machine doctor for Diagnosis of Gynic & Obstetric Problems using Neural Networks", Prof.M.S.Prasad Babu, Smt.Ch.Sita Kameswari, Sri.N.V.Ramana Murty, and Sri.G.Satyanarayana, Accepted for Publication in International Journal of Computer Sciences, Systems Engineering and Information Technology, July-Dec 2010.
15. Valery Tereshko, Reaction-Diffusion Model of a Honeybee Colony's Foraging Behaviour, Proceedings of the 6th International Conference on Parallel Problem Solving from Nature, p.807-816, September 18-20, 2000.
16. D. Jeya Mala, V. Mohan, ABC Tester - Artificial Bee Colony Based Software Test Suite Optimization Approach, International Journal of Software Engineering , 15-43, 2009.
Referred the Web Pages
17. www.indiakisan.net
18. www.bharathgramarogya.net
19. www.bharathgramarogya.net/newindex.html
20. www.indg.in/agriculture

Name Entity Recognition in Machine Translation

R.C. Balabantaray
International Institute of Information Technology, Bhubaneswar, India
Email: rakeshbray@gmail.com

Abstract- We propose a new name entity recognition method and its translation based on association rules and assumptions. We evaluate the performance of our method with various kinds of texts. The work presented in this paper is part of a larger effort to develop Machine Translation (MT) system which can take care of name entities.

Keywords: MT, NER, ENAMEX, NUMEX, TIMEX.

1. Introduction

Name Entity Recognition (NER) is an information extraction task that is concerned with the recognition and classification of name entity from free text [6]. Name entities classes are, for instance, locations, person named, organization named, dates, times and money amounts. To terms and expressions in the text correspond the entities they represent. For example, in the sentence: *"Indian Prime Minister Mannmohan Singh and President Dr. Abdul Kalam pose for photographers at the Palace",* a name entity recognition process looking for named persons and locations would identify the two persons *Manmohan Singh* and *Abdul Kalam* and the location *Palace*. This recognition can be based on a variety of features of the terms, the sentence, the text and its syntax and could leverage external sources of information such as thesauri and dictionaries, for instance. In the example, a system may have applied a simple rule guessing that the capitalize words directly following the terms '*President*' or '*Minister*' are names of persons. But the most important question is how to convert these name entities to appropriate words in the target language so, that the Machine translation will be fruitful.

Named Entity Recognition (NER), the identification of entity names in free text, is a well-studied problem. In most previous work, NER has been applied to news articles (e.g., (Bikel et al., 1999; McCallum and Li, 2003)), scientific articles (e.g., (Craven and Kumlien, 1999; Bunescu and Mooney, 2004)), or web pages (e.g., (Freitag, 1998)).

India is the second largest in population in the world with more than one billion populations. There are 19 constitutional languages with 10 scripts and over 1650 dialects. Orissa is a state of India situated in the eastern region, with a population of 36.7 million according to 2001 census. Orissa is the first state in India to have been formed on linguistic basis. Oriya is the official as well as spoken language of Orissa, and is one of the constitutional languages of India. We are working to develop a system (OMTrans) which will translate the source language English to target language Oriya.

In this paper, we propose a method for name entity class recognition based on such rules and some assumptions. The rest of the paper is organized as follow. We present and discuss some background and related work on name entity class recognition in the next section. The method is presented in section 3.Finally we conclude and identify the next steps of our research.

2. Related Work

Name Entity Recognition (NER) has been a well-studied problem. This problem is applicable to both formal as well as informal texts. In case of informal text just like emails it is rather easy to identify the name entities as in this case one can find some labels attached with the name entities. But in the formal texts it will be difficult to find it out. There are several classification methods which are successful to be applied on this task. Chieu and Ng[7] and Bender et al.[10] used Maximum Entropy approach as the classifier. Conditional Random Filed (CRF) was explored by McCallum and Li [1] to NER. Mayfield et al.[8] applied Support Vector Machine (SVM) to classify each name entity. Florian et al. [12] even combined Maximum Entropy and hidden Markov Model (HMM) under different conditions. Some other researches are focused more on extracting some efficient and effective features for NER. Chieu and Ng[7] successful used local features, which are near

the word, and global features, which are in the whole document together. Klein et al.[5] and Whitelaw et al.[2] reports that character-based features are useful for recognizing some special structure for the name entity.

3. Methodology

Our algorithm described below is a unique one which is identifying the name entities expressions (ENAMEX), numerical entity expressions (NUMEX), and temporal entity expressions (TIMEX). The most important thing with this technique is that the dictionary should be exhaustive and the lists corresponding to the name entity varieties (like ENAMEX_TYPES, NUMEX_TYPES and TIMEX_TYPES) should be proper. In the name entity ENAMEX there are almost eleven types such as person, organizations, location, facilities, locomotives, artifacts, entertainment, cuisine's, organisms, plants and diseases. Similarly in NUMEX there are four varieties like distance, money quantity, count and in TIMEX there are again four varieties like time, date, day, period. For the above described TYPES we are maintaining different lists consisting of words relevant to the corresponding TYPES arranged according to the ASCII collating sequence so, that the process of searching will be faster. These lists have been prepared with the help of various sources like Corpora and from the texts being used in our day to day life. So, far basically person names are concerned new names are getting created day by day and that might not appeared in the list. In this scenario our system will simply identify it as name entity and it will not be able to categorize it.

Algorithm

Step1: The root words (after removing the tag features from the corresponding words) starting with Uppercase letters which is not present in our Bilingual dictionary or the root words not present in the bilingual dictionary is treated as Name Entity (it may be ENAMEX, NUMEX or TIMEX) with some exception.

Step 2: For categorizing whether it is ENAMEX, NUMEX or TIMEX certain rules are being followed.

Step 3: After that for finding the types we have to refer to the list defined under the corresponding broad categories like ENAMEX, NUMEX and TIMEX.

For example in the sentence like "*Mannmohan Singh is the prime minister of our country.*" The word *Monnmohan* and *Singh* both are starting with uppercase letters and not present in the dictionary so both will be identified as name entity and as ENAMEX category as per our rules. Now binary search will be performed in the list defined under ENAMEX category to get the subcategory and both the word will be found as person name (Individual).

<ENAMEX TYPE="PERSON"> Mannmohan Singh</ENAMEX>

In the sentence like "*I am a citizen of India."* the word *India* is starting with uppercase letter but it is present in the dictionary so, it will not be treated as name entity but here as per our exception rules and the information retrieved from dictionary this is name of a place so, it will be treated as ENAMEX and type will be location.

<ENAMEX TYPE="LOCATION"> India</ENAMEX>

Similarly in the sentence like "*it is 5:30 AM in the morning.*" The token *5:30* will be treated as name entity but since it contains numeric figure it can be NUMEX or TIMEX but as per our rules it is followed by AM so, it will be TIMEX. Then after consulting the lists the type will be finalized as *Time*.

<TIMEX TYPE="TIME"> 5:30 AM</TIMEX>

The table given below shows the names of the lists we have maintained for identifying the types of the name entities under the three broad categories ENAMEX, NUMEX and TIMEX.

Name of broad categories	Name of Lists
ENAMEX	ENAMEX_PERSON
	ENAMEX_ORGANIZATION
	ENAMEX_LOCATION
	ENAMEX_FACILITIES
	ENAMEX_LOCOMOTIVES
	ENAMEX_ARTIFACTS ENAMEX_ENTERTAINMENT
	ENAMEX_CUISINE
	ENAMEX_ORGANISMS
	ENAMEX_PLANTS
	ENAMEX_DISEASE
NUMEX	NUMEX_DISTANCE
	NUMEX_MONEY
	NUMEX_QUANTITY
	NUMEX_COUNT
TIMEX	TIMEX_TIME
	TIMEX_DATE

	TIMEX_DAY
	TIMEX_PERIOD

4. Conclusion

We have presented a new name entity class recognition method based on association rules and assumptions. The dictionary which we are following for this task is exhaustive but the contents of the lists are growing day by day. Once the list will be exhaustive it is guaranteed that our method will show a higher degree of precession than the other methods. The technique used over here to identify the name entities can also be applicable to the translation system of English to other Indian languages.

We are in the process of transliterating these name entities to the target language Oriya so, that this will be a standard for the translation system of English to other Indian Languages as all Indian languages are phonetically linear.

References

[1]. A. McCallum and W. Li. 2003. Early results for named entity recognition with conditional random fields, feature induction and web-enhanced lexicons. In *CoNLL- 2003*.

[2]. Casey Whitelaw and Jon Patrick *2003.* Named Entity Recognition Using a Character-based Probabilistic Approach. In: *Proceedings of CoNLL-2003*, Edmonton, Canada, 2003, pp. 196-199.

[3]. D. M. Bikel, R. L. Schwartz, and R. M. Weischedel. 1999. An algorithm that learns what's in a name. *Machine Learning*, 34:211–231.

[4]. D. Freitag. 1998. Information extraction from html: application of a general machine learning approach. In
AAAI-98.

[5]. Dan Klein, Joseph Smarr, Huy Nguyen and Christopher D. Manning, *2003.* Named Entity Recognition with Character-Level Models. In: *Proceedings of CoNLL-2003*, Edmonton, Canada, pp. 180-183.

[6]. Grishman, R., 1997. "Information Extraction: Techniques and Challenges", Lecture Notes in Computer Science, Vol. 1299,Springer-Verlag.

[7]. Hai Leong Chieu and Hwee Tou Ng, 2003. Named Entity Recognition with a Maximum Entropy Approach. In: *Proceedings of CoNLL-2003*, Edmonton, Canada, pp. 160-163.

[8]. James Mayfield, Paul McNamee and Christine Piatko, 2003. Named Entity Recognition using Hundreds of Thousands of Features. In: *Proceedings of CoNLL-2003*, Edmonton, Canada, pp. 184-187.

[9]. M. Craven and J. Kumlien. 1999. Constructing biological knowledge bases by extracting information from text sources. In *ISMB-99*.

[10]. Oliver Bender, Franz Josef Och and Hermann Ney, 2003. Maximum Entropy Models for Named Entity Recognition In: *Proceedings of CoNLL-2003*, Edmonton, Canada, pp. 148-151.

[11]. R. Bunescu and R. J. Mooney. 2004. Relational markov networks for collective information extraction. In *ICML-2004 Workshop on Statistical Relational Learning*.

[12]. Radu Florian, Abe Ittycheriah, Hongyan Jing and Tong Zhang, 2003. Named Entity Recognition through Classifier Combi nation. In: *Proceedings of CoNLL-2003*, Edmonton, Canada, pp. 168-171.

Dr. Rakesh Chandra Balabantaray is currently working as Assistant Professor in the Department of Computer Science & Enginerring at IIIT, BHUBANESWAR, Orissa, India. He did his Masters in Computer Science in the year 2001 and Ph.D. in Computer Science in the year 2008 from Utkal University, Orissa, India. He was born in the year 1978. He has more than twenty publications in various journals and conferences. His major area of research is Artificial Intelligence, Natural Language Processing & Information Retrieval.

Normalization of Neutrosophic Relational Database

Smita Rajpal[1], Pariza kamboj[1]
[1]ITM University, Gurgaon , India
smita_rajpal@yahoo.co.in , pariza@gmail.com

***Abstract*:** In this paper authors have presented a method of normalizing a relational schema with Neutrosophic attributes into 1NF.This Method is called as Neutrosophic-First Normal Form(1NF(N)) a revision of First normal Form in Relational database. Authors are taking the Neutrosophic Relational database [3, 1] which is the extension of Fuzzy and Vague database to define the Neutrosophic- First Normal form.

***Keywords*:** Neutrosophic Normal Form, Query language, Vague Set, Fuzzy Set, Neutrosophic 1NF

1. INTRODUCTION

The normalization process takes a relation schema through a series of tests to check up whether it satisfies a certain normal form. Consider an instance of a relation schema. In real life situation, the data available are not always precise or crisp, rather it can be in any form like it can be in natural language, any imprecise data or you can say Neutrosophic data .Consequently if, at least one data is Neutrosophic, the relation schema can not be called to be in 1NF. The quest to manage imprecision's is equal to major driving force in the database community is the Ultimate cause for many research areas: data mining, semi structured data, and schema matching, nearest neighbor. Processing probabilistic data is fundamentally more complex than other data models. Some previous approaches sidestepped complexity.

For example, consider an attribute SALARY (in $) of a relation schema Employee. If a tuple value for this attribute SALARY is precise viz. 5000, then it is a single atomic (Indivisible) value. But if a tuple value is Neutrosophic viz. "Approximately 5000", then it can not be called an atomic value.

In this paper authors study this problem and suggest a method to normalize such relational Schemas into 1NF. Such a normal form we shall call by Neutrosophic-1NF or 1NF (N).

2. PRELIMINARIES

Out of several higher order fuzzy sets, Vague sets and Neutrosophic Sets the concept of Neutrosophic sets has been found to have enormous potential to deal with vague or imprecise data in case of engineering or technological or economical or mathematical analysis to list a few only. In this section, Author presents some preliminaries on the theory of Neutrosophic sets (NS) which will be required for the progress of this paper. The failure of the RDBMS is due to the presence of imprecise constraints in the query predicate which can not be tackled due to the limitation of the grammar in standard query languages which work on crisp environment only. But these types of queries are very common in business world and in fact more frequent than grammatical-queries, because the users are not always expected to have knowledge of DBMS and the query languages. Consequently, there is a genuine necessity for the different large size organizations, especially for the industries, companies having world wide business, to develop such a system which should be able to answer the users queries posed in natural language, irrespective of the QLs and their grammar, without giving much botheration to the users. Most of these type of queries are not crisp in nature, and involve predicates with fuzzy (or rather vague) data, fuzzy/vague hedges (with concentration or dilation). Thus, these types of queries are not strictly confined within the domains always. The corresponding predicates are not hard as in crisp predicates. Some predicates are soft because of vague/fuzzy nature and thus to answer a query a hard match is not always found from the databases by search, although the query is nice and very real, and should not be ignored or replaced according to the business policy of the industry. To deal with uncertainties in searching match for such queries, fuzzy logic and rather vague logic [7] and Neutrosophic logic by Smarandache [3] will be the appropriate tool.

Fuzzy set theory has been proposed to handle such vagueness by generalizing the notion of membership in a set. Essentially, in a Fuzzy Set (FS) each element is associated with a point-value selected from the unit interval [0,1], which is termed the grade of membership in the set. A Vague Set (VS), as well as an Intuitionistic Fuzzy Set (IFS), is a further generalization of an FS. Now take an example, when we ask the opinion of an expert about certain statement, he or she may say that the possibility that the statement is true is between 0.5 and 0.7, and the statement is false is between 0.2 and 0.4, and the degree that he or she is not sure is between 0.1 and 0.3. Here is another example, suppose there are 10 voters during a voting process. In time t1, three vote "yes", two vote "no" and five are undecided, using neutrosophic notation, it can be expressed as x(0.3,0.5,0.2); in time t2, three vote "yes", two vote "no", two give up and three are undecided, it then can be expressed as x(0.3,0.3,0.2). That is beyond the scope of the intuitionistic fuzzy set. So, the notion of neutrosophic set is more general and overcomes the aforementioned issues. In neutrosophic set, indeterminacy is quantified explicitly and truth membership, indeterminacy-membership and falsity-membership are independent. This assumption is very important in many applications such as information fusion in which we try to combine the data from different sensors. Neutrosophy was introduced by Smarandache [7].

Neutrosophic set is a powerful general formal framework which generalizes the concept of the classic set, fuzzy set [2], Vague set [1] etc.

A neutrosophic set A defined on universe U. x = x(T,I,F) ε A with T,I and F being the real standard or non-standard subsets of]0-,1+[, T is the degree of truth-membership of A, I is the degree of indeterminacy membership of A and F is the degree of falsity-membership of A.

Definition 2.1

A Neutrosophic set A of a set U with $t_A(u)$, $f_A(u)$ and $I_A(u)$, $\forall u \in U$ is called the α- Neutrosophic set of U, where $\alpha \in [0,1]$.

Definition 2.2

A Neutrosophic number (NN) is a Neutrosophic set of the set R of real numbers.

A tuple in a neutrosophic relation is assigned a measure. Will be referred to as the *truth* factor and will be referred to as the *false* factor. The interpretation of this measure is that we believe with confidence and doubt with confidence that the tuple is in the relation. The truth and false confidence factors for a tuple need not add to exactly 1. This allows for incompleteness and inconsistency to be represented. If the truth and false factors add up to less than 1, we have incomplete information regarding the tuple's status in the relation and if the truth and false factors add up to more than 1, we have inconsistent information regarding the tuple's status in the relation.

In contrast to vague relations where the grade of membership of a tuple is fixed, neutrosophic relations bound the grade of membership of a tuple to a subinterval $[\alpha, 1-\beta]$ for the case, $\alpha + \beta \leq 1$.. The operators on fuzzy relations can also be generalized for neutrosophic relations. However, any such generalization of operators should maintain the belief system intuition behind neutrosophic relations.

Definition 2.3

A neutrosophic relation on scheme R on $\sum$ is any subset of $\tau(\sum) \times [0,1] \times [0,1]$, Where $\tau(\sum)$ denotes the set of all tuples on any scheme $\sum$.

For any $t \in \tau(\sum)$, we shall denote an element of R as $\langle t, R(t)^+, R(t)^- \rangle$, where, $R(t)^+$ is the truth factor assigned to t by R and $R(t)^-$ is the false factor assigned to t by R. Let $V(\sum)$ be the set of all neutrosophic relation on $\sum$.

Definition 2.4

A neutrosophic relation on scheme R on $\sum$ is consistent if $R(t)^+ + R(t)^- \leq 1$, for all $t \in \tau(\sum)$. Let $C(\sum)$ be the set of all consistent neutrosophic relations on $\sum$. R is said to be complete if $R(t)^+ + R(t)^- \geq 1$, for all $t \in \tau(\sum)$. If R is both consistent and complete, *i.e.* $R(t)^+ + R(t)^- = 1$, for all $t \in \tau(\sum)$. Then it is a *total* neutrosophic relation, and let $T(\sum)$ be the set of total neutrosophic relation on $\sum$.

2.1 Operator Generalizations

It is easily seen that neutrosophic relations are a generalization of vague relations, in that for each vague relation there is a neutrosophic relation with the same information content, but not *vice versa*. It is thus natural to think of generalizing the operations on vague relations such as union, join, and projection etc. to neutrosophic relations. However, any such generalization should be intuitive with respect to the belief system model of neutrosophic relations. We now construct a framework for operators on both kinds of relations and introduce two different notions of the generalization relationship among their operators.

An n-ary operator on fuzzy relations with signature $\langle \sum_{1,\ldots,} \sum_{n+1} \rangle$ is a function $\Theta : F(\sum_1) \times \ldots \times F(\sum_n) \rightarrow \sum_{n+1}$, where $\sum_{1,\ldots,} \sum_{n+1}$ are any schemes. Similarly An n-ary operator on neutrosophic relations with signature $\langle \sum_{1,\ldots,} \sum_{n+1} \rangle$ is a function $\Psi : V(\sum_1) \times \ldots \times V(\sum_n) \rightarrow V(\sum_{n+1})$.

Definition 2.5

An operator Ψ on neutrosophic relations with signature $\langle \sum_{1,\ldots,} \sum_{n+1} \rangle$ is *totality preserving* if for any total neutrosophic relations $R_{1,\ldots}R_n$ on schemes $\sum_{1,\ldots,} \sum_{n+1}$, respectively. $\Psi(R_{1,\ldots}R_n)$ is also total.

Definition 2.6

A totality preserving operator Ψ on neutrosophic relations with signature $\langle \sum_{1,\ldots,} \sum_{n+1} \rangle$ is a *weak generalization* of an operator Θ on fuzzy relations with the same signature, if for any total neutrosophic relations $R_{1,\ldots}R_n$ on schemes $\sum_{1,\ldots,} \sum_n$, respectively ,we have

$\lambda_{\sum_{n+1}}(\Psi(R_{1,\ldots}R_n)) = \Theta(\lambda_{\sum_1}(R_1), \ldots, \lambda_{\sum_n}(R_n))$.

The above definition essentially requires Ψ to coincide with Θ on total neutrosophic relations (which are in One-one correspondence with the vague relations). In general, there may be many operators on neutrosophic relations that are weak generalizations of a given operator Θ on fuzzy relations. The behavior of the weak generalizations of Θ on even just the consistent neutrosophic relations may in general vary. We require a stronger notion of operator generalization under which, at least when restricted to consistent neutrosophic relations, the behavior of all the generalized operators is the same. Before we can develop such a notion, we need that of 'representation' of a neutrosophic relation.

We associate with a consistent neutrosophic relation R the set of all (vague relations corresponding to) total neutrosophic relations obtainable from R by filling the gaps between the truth and false factors for each tuple. Let the map be $reps_{\sum} : C(\sum) \rightarrow 2^{F(\sum)}$. is given by ,

$reps_{\sum}(R) = \{Q \in F(\sum) \mid \bigwedge_{t_i \in \tau(\sum)} (R(t_i)^+ \leq Q(t_i) \leq 1 - R(t_i)^-)\}$.

The set $reps_{\sum}(R)$ contains all fuzzy relations that are 'completions' of the consistent neutrosophic relation R .Observe that $reps_{\sum}$ is defined only for consistent neutrosophic relations and produces sets of fuzzy relations. Then we have following observation.

Proposition 2.1 For any consistent neutrosophic relation R on scheme $\sum$, $reps_{\sum}(R)$ is the singleton $\{\lambda_{\sum}(R)\}$, iff R is total.

Proposition 2.2 If Ψ is a strong generalization of Θ, then Ψ is also a weak generalization of Θ.

2.2 Generalized Algebra on Neutrosophic Relations

In this section, we present one strong generalization each for the vague relation operators such as union, join, and projection. To reflect generalization, a hat is placed over a vague relation operator to obtain the corresponding neutrosophic relation operator. For example, ∞ denotes the natural join among fuzzy relations, and $\overset{\wedge}{\infty}$ denotes natural join on neutrosophic relations. These generalized operators maintain the truth system intuition behind neutrosophic relations.

2.3 Set-Theoretic Operators

We first generalize the two fundamental set-theoretic operators, union and complement.

Definition 2.7 Let R and S be neutrosophic relations on scheme $\sum$. Then, The union of R and S, denoted $R \overset{\wedge}{\cup} S$, is a neutrosophic relation on scheme $\sum$,given by $R \overset{\wedge}{\cup} S=$ $\left\langle \max\{R(t)^{+}, S(t)^{+}\}, \min\{R(t)^{-}, S(t)^{-}\} \right\rangle$ for any $t \in \tau(\sum)$.

(a) The complement of R ,denoted by $\overset{\wedge}{-}R$,is a neutrosophic relation on scheme $\sum$,given by $(\overset{\wedge}{-}R)(t)=\langle R(t)^{+}, R(t)^{-} \rangle$,for any $t \in \tau(\sum)$.

An intuitive appreciation of the union operator can be obtained as follows: Given a tuple *t*, since we believed that it is present in the relation *R* with confidence $R(t)^{+}$ and that it is present in the relation *S* with confidence $S(t)^{+}$, we can now believe that the tuple *t* is present in the "either *-R* or *- S* " relation with confidence which is equal to the larger of $R(t)^{+}$ and $S(t)^{+}$. Using the same logic, we can now believe in the absence of the tuple *t* from the "either - *R* - or - *S* " relation with confidence which is equal to the smaller (because *t* must be absent from both *R* and *S* for it to be absent from the union) of $R(t)^{-}$ and $S(t)^{-}$

Proposition 2.3 : The operator $\overset{\wedge}{\cup}$ and $\overset{\wedge}{-}$ on neutrosophic relation are strong generalization of the operators $\cup$ and unary – on vague relations.

Definition 2.8 Let R and S be neutrosophic relations on scheme $\sum$. Then, The intersection of R and S denoted as $\overset{\wedge}{\cap}$, is a neutrosophic relation on scheme $\sum$,given by

$R \overset{\wedge}{\cap} S(t)=\langle \min\{R(t)^{+}, S(t)^{+}\}, \max\{R(t)^{-}, S(t)^{-}\} \rangle$, for any $t \in \tau(\sum)$ .The difference of R and S denoted as $R \overset{\wedge}{-} S$, neutrosophic relation on scheme $\sum$,given by

$(R \overset{\wedge}{-} S)(t) = \langle \min\{R(t)^{+}, S(t)^{-}\}, \max\{R(t)^{-}, S(t)^{+}\} \rangle$, any $t \in \tau(\sum)$.

The following proposition relates the intersection a difference operators in terms of the more fundamental s theoretic operators union and complement.

Proposition 2.4 : For any neutrosophic relation on the sa scheme

$R \overset{\wedge}{\cap} S = -(-R \overset{\wedge}{\cup} -S)$ and

$R \overset{\wedge}{-} S = -(-R \overset{\wedge}{\cup} S)$.

The important issue of closeness can not be addressed with the crisp mathematics. That is why author have used the Neutrosophic tools[1].

3. NEUTROSOPHIC-1NF OR 1NF (N)

In this section author will explain the method of normalizing a relational schema (with Neutrosophic attributes) into 1NF in Table 1. For the sake of simplicity, author consider a relation schema R with only one Neutrosophic attribute, all other three attributes being crisp. By "Neutrosophic attribute" author mean that at least one attribute value in a relation instance is Neutrosophic.

A_1	A_2	A_3	A_4

Table 1:Relational Schema R

This relational schema R has four attributes of which A_4 is the only Neutrosophic attribute (say).Consider a relation instance r of R given by:

A_1	A_2	A_3
A_{11}	**A21**	A_{31}
A_{12}	**A_{22}**	A_{32}
A_{13}	**A_{23}**	A_{33}
A_{14}	A_{24}	A_{34}

Table 2

Suppose that A_2 is the primary key here, all the data are precise except $\tilde{a}$, which is an Neutrosophic number. Thus all the data except $\tilde{a}$ is atomic. This is not in 1NF because of the non atomic data $\tilde{a}$ **.**

An Neutrosophic number is an Neutrosophic set of the set R of real numbers. The universe of discourse R is an infinite set. But, in our method of normalization we shall consider a finite universe of discourse, say X, whose cardinality is N.

Let us suppose that X: {x1, x_2,.. x_n} and the Neutrosophic number $\tilde{a}$,
Proposed by a database expert is an NS (Neutrosophic Set) given by:

$\tilde{a}$ **= { (x_i , μ_i ,V_i) : $X_i \in X$, I= 1 , 2 , 3 , . . .N}**

Then theTable2 can be replaced by the following table:

A_1	A_2	A_3	A_4
A_{11}	**A_{21}**	**A_{31}**	A_{41}
A_{12}	**A_{22}**	**A_{32}**	{(X_i , μ_i, V_i),(X_1, μ_1,V_1),……..(X_n, μ_n,V_N)}
A_{13}	**A_{23}**	**A_{33}**	A_{43}
A_{14}	A_{24}	A_{34}	A_{44}

Table 3: The Relation Instance r

Now remove all the Neutrosophic attributes (here A_4 only), from Table 3. Replace Table 3 by the following two tables:

A_1	A_2	A_3	A_4
A_{11}	$\mathbf{A_{21}}$	$\mathbf{A_{31}}$	A_{41}
A_{12}	$\mathbf{A_{22}}$	$\mathbf{A_{32}}$	A_{42}
A_{13}	$\mathbf{A_{23}}$	$\mathbf{A_{33}}$	A_{43}
A_{14}	A_{24}	A_{34}	A_{44}

Table 4: The Relation r_1

In table 5 we have all the attributes of the primary-key of r (here only one attribute A_2),
the Neutrosophic attribute A_4 and two new attributes which are MEMBERSHIP_VALUE(A_4) or MV(A_4)and NONMEMBERSHIP_VALUE(A_4) or NMV(A_4). Corresponding to all precise values of A_4, the MV(A_4) value is put 1 and the NMV(A_4) value is 0.

A_2	A_4	MV(A_4)	N MV(A_4)
A_{21}	$\mathbf{A_{41}}$	**1**	0
A_{22}	$\mathbf{X_1}$	μ_1	V_1
A_{22}	$\mathbf{X_2}$	μ_2	V_2
A_{22}	$\mathbf{X_3}$	μ_3	V_3
…	**…**	**…**	…
A_{22}	$\mathbf{X_n}$	μ_n	V_4
A_{23}	$\mathbf{A_{43}}$	**1**	0
A_{24}	A_{44}	1	0

Table 5: The Relation r_2

Now we see that the relation schema is in 1NF. Such a method of normalization is called Neutrosophic normalization and the normal form is called Neutrosophic 1NF or 1NF (N).

4. RESULTS

We study the method here by an example with hypothetical data. Consider a relation schema FRUIT as shown below whose primary key is FCODE and the attribute YEARLY-PRODUCTION is an Neutrosophic attribute.

FNAME	FCODE	YEARLY-PRODUCTION (in million of tones)

Table 6: The Relation Schema FRUIT

Consider a relation instance of this relation schema given by the following Table 7

FNAME	FCODE	YEARLY-PRODUCTION (in million of tones)
APPLE	**F001**	4563
MANGO	**F002**	6789
GUAVA	**F003**	Approximately 56
BANANA	F004	8987

Table 7

In this instance FNAME and FCODE are crisp attribute whereas YEARLY-PRODUCTION is attribute values for FNAME are atomic; all the attribute values for the attribute FCODE are atomic. But all the attribute values for the attribute Yearly Production are not atomic. The data" approximately 56" is an Neutrosophic number $\tilde{56}$.Suppose that for this relation, a database expert proposes the Neutrosophic number $\tilde{56}$ as an NS given by $\tilde{56}$ ={**(55,.8,.1),(56,.9,.03),(56.5,.7,.10)}.**

Therefore Table 7 could be replaced by the following Table 8

FNAME	FCODE	YEARLY-PRODUCTION (in million of tones)
APPLE	**F001**	4563
MANGO	**F002**	6789
GUAVA	**F003**	**{(**55,.8,.1),(56,.9,.03),(56.5,.7,.10)}
BANANA	F004	8987

Table 8

Now remove the Neutrosophic attribute YEARLY-PRODUCTION (YP) for this instance and divide it into two relations given as:

FNAME	FCODE
APPLE	F001
MANGO	F002
GUAVA	F003
BANANA	F004

Table 9: FRUIT-1 Relation

FCODE	YP	MV(YP)	NMV(YP)
F001	**4563**	**1**	0
F002	**6789**	**1**	0
F003	**55**	**.8**	.1
F003	**56**	**.9**	.03
F003	**56.5**	**.7**	.10
F004	8987	1	0

Table 10: FRUIT-2 Relation

Clearly, it is now in 1NF, called by 1NF (N). For FRUIT-1, the Primary Key is FCODE, but for the newly created FRUIT-2 the Primary Key is {FCODE, YP}. Let us present below the sequence of steps for Neutrosophic normalization of relation schema into **1NF(N).**

5. ALGORITHM

(l) Remove all the Neutrosophic-attributes from the relation.
(2) For each Neutrosophic-attribute create one separate table with the following attributes:
(i) All attributes in the primary key
(ii) MV (z)
(iii) NMV (Z)
(3) For every precise value of the Neutrosophic attribute put MV=1 and NMV=0.

Thus, if there is m number of attributes in the relation schema then, after normalization there will be in total (m+l) number of relations .In special case, when the hesitation or in deterministic parts are nil for every element of the universe of discourse the Neutrosophic number reduces to fuzzy number. In such cases, the attribute NMV (Z) will not be required in any reduced tables of 1NF.In future work I will consider this part.

6. CONCLUSION

In this paper we have presented a method of normalization of a relational schema with Neutrosophic attribute in 1NF (N). We have implemented the method by an example given in section 4 which proves that how the imprecise data can be handle in relational schema using First Normal Form of Neutrosophic databases. We claim that the algorithm suggested in section 5 is totally a new concept which can easily handle the neutrosophic attributes of First normal Form.

REFERENCES

[1] Smita Rajpal, M.N.Doja and Ranjit Biswas, A Method of Neutrosophic Logic to Answer Queries in Relational Database,JCS, Vol.4(4): 309-314 ,USA, 2008.

[2] Atanassov, K., Intuitionistic Fuzzy Sets : Theory and Applications, Physica-Verlag (2000), New- York.

[3] F. Smarandache (2002a), A unifying Field in Logics: Neutrosophic Logic, in Multiple- Valued Logic / An International Journal, Vol. 8, No. 3, 385- 438, 2002.

[4] Atanassov, K., Intuitionistic Fuzzy Sets : Theory and Applications,Physica-Verlag (2000), New- York.

[5] Biswas,R ., Intuitionistic fuzzy relations Bull. Sous. Ens.Flous.Appl. (BUSEFA L) 70 ,1997,22-29.

[6] Codd,E .F.,A relational model of data for large shared databanks ACM commun.l 3(6)(1970)377-387.

[7] Gau, W.L. and Buehrer, D, J., Vague sets, IEEE Transactions on Systems, Man and Cybernetics, Vol.23 (1993) 610-614

Author Biographies

Ms. Smita Rajpal is working as an Assistant Professor Computer Science and Engineering at ITM University, Gurgaon (India). She is specialized in Soft Computing and RDBMS. She is a Life Time member of ISTE, SCIP, IACSIT, WASET. She is working as an Editor / Reviewer/Technical member of various International Journals and Conferences. She is an active member of IEEE. Her biography is a part of Marquis who's who in 2010.Her total Research publications are 18 and book chapters -5. She has published two books.

Ms. Pariza Kamboj is working as an Assistant Professor Computer Science & Engg. at ITM university, Gurgaon (India). She is specialised in Mobile Ad hoc Networks and Soft Computing. Currently she is pursuing her Ph.D. in Computer science. She has published a total of 14 research papers in various International/ National Journals and Conferences. She is a Life time member of Indian Society of Technical Education (ISTE) and International Association of Computer Science and Information Technology (IACSIT). She is also serving as an Reviewer /Editor/Technical Member of various International Committee and Journals.

Coping with Emergent Skills: an Appraisal of the Provision and Integration of ICT Infrastructures in Nigerian Universities.

Dr A. O. Ogunleye., Dr. B. F. Adeoye & Dr. C. O. Oke.
Dept of Science and Technology Education,
University of Lagos, Akoka-Yaba. Lagos-Nigeria.

Abstract- Since the industrial revolution, technology has expanded to influence and impact upon more and more of the world. At work or play, the human being is increasingly faced with technological interfaces. As a result of these changes, many universities have embraced the huge challenge of technology in education based on a transition from process to information-based activities, from an individual approach to a collaborative one, from a knowledge oriented to skill oriented learning. Information and Communications Technologies (ICT) has increasingly dominated universities attention, more so now that the global concern for Higher Education is access for the vast number of people seeking university education. To this end, many universities in developed counties are adapting to the e-learning and virtual university as a panacea to this problem. The purpose of this study is to first and foremost review the state of higher education (University) in Nigeria vis-à-vis factors militating its relevance and efficiency. Secondly, it provides an empirical investigation by examining four key areas of the provision and integration of ICT in Nigerian universities, namely, its impact on the lecturers, the learners, the institutions and our higher education overall policies. Thirdly, technology has provided individuals around the world the convenience of constructing a virtual global learning community. A specific growing technology that compliments the expanding global need is E-learning. When coupling the growth of e-learning with an expanding global marketplace, it becomes imperative to gain an understanding of how Nigerian students are coping with these emergent tools.

Keywords: Emergent skills, ICT infrastructures, Nigerian universities.

1. Introduction

Technology has provided individuals around the world the convenience of constructing a virtual global learning community. Businesses, educational, and training organizations no longer must rely solely on face-to-face, teacher-led or lecture style instruction.
E-Learning and Information Communications Technologies (ICT) cannot be successful without an appropriate and adequate infrastructure. It is worth mentioning here that The ICT revolution started in Nigeria after return to democratic rule in 1999 (Ajayi, 2003). This notwithstanding several government agencies and other and other stakeholders in the private sector have initiated ICT driven projects and programs to impact all levels of the educational sector. When coupling the growth of ICT usage with an expanding global marketplace, it becomes imperative to gain a better understanding of the provision and integration of ICT infrastructures in Nigerian Universities.

Many universities and organizations are building infrastructure systems that support e-learning programs, performance, content, and resource management, but most of them focused on content and delivery. Although content and delivery are important, they alone don't equal ICT success. In fact, focusing only on content and delivery can create an unsuccessful ICT experience. Infrastructure is the permanent foundation on which ICT can function effectively. Supporting the entire infrastructure is a common vision and language for conveying effective ICT implementation across the enterprise.

Infrastructure, according to Blinco, K.,, Mason, J., McLean., N., and Wilson, S., (2004) is highly contextual in its meaning. In e-learning contexts, "e-learning infrastructure", "technical infrastructure", and "ICT infrastructure" all convey a range of meanings. For the technically inclined, "infrastructure" often describes a bottom "layer" of an architectural description or diagram, indicating network hardware components, communications processes, services and protocols. For the purpose of this study 'infrastructure' describes everything that supports both the flow and processing of information, including but not limited to, Internet connectivity, hardware, software, facilities, communication protocols.

In recent times, there has been a few e-Learning educational provisions in Nigeria. There are currently 52 university centers and colleges of education who have their domain names registered while 15 are active email users. Also, there are several initiatives; these initiatives are either being undertaken by government, civil society or the private sectors. Omwenga (2003) listed them as:

Government e-education initiatives and NGO's e-Learning initiatives:

Government e-education initiatives include the following:

1. The Nigerian Universities Network (NUNet) Project
2. The Polytechnics Network (PolyNet) Project
3. The School Net Project
4. The Nigerian Education, Academic and Research Network (NEARNet)
5. The Teachers Network (TeachNet) Project
6. National Open University
7. National Virtual (Digital) Library (Ministry of Education/ NUC)
8. National Virtual Library (Ministry of Science and Technology/NITDA)
9. National Information, communication and education programme of the presidency.
10. The NGO's e-Learning Initiatives

These include examples of tele-centers

1. Community Teaching and Learning Centers
2. Lagos Digital Village by Junior Achievement Nigeria
3. Owerri Digital Village by Youth for Technology Foundation
4. Computer Literacy for Older Persons Programme by Mercy Mission

Even though details of the various initiatives can be found in Omwenga (2003), it is important to elaborate on the objectives of two of the initiatives which are detailed below.

a) National Universities Network (NUNet) Initiatives
National Universities Network (NUNet) was established in 1994 with the following objectives:

1. To end isolation of Nigerian academic staff and students from each other and from the global academic community;
2. To ameliorate staff shortages arising from the brain-drain, by providing the ICT facility required by our academic staff-in-Diaspora to make their contributions regardless of where they live or work;
3. To encourage the sharing of resources, foster academic and research collaborations among Nigerian Universities, and with their counterparts throughout the world;
4. To provide universities with access to electronic databases, journals and books many of which are increasingly available only in digital formats;
5. To serve as vehicle to expand access to education at minimal cost of capital building expenditure; and
6. To place Nigerian universities at the forefront of the information revolution, that they might serve their proper roles as foci for national development.

(b) Virtual Library Initiatives
Virtual library initiatives in Nigeria include:

- The National Virtual (Digital) Library Project of the Ministry of Education,
- The National Virtual Library Project of the Ministry of Science and Technology is supervised by the National IT Development Agency.
- An ongoing effort by UNESCO to develop a virtual Library for all Nigerian Higher Education Institutions in Nigeria.

The objectives of national virtual library initiatives are as follows:

1. Improving the quality of teaching and research institutions through the provision of current books, journal and other library resources;
2. Enhancing access of academic libraries to global library and information resources; enhancing scholarship and lifelong learning through the establishment of permanent access to shared digital archival collections;
3. Providing academic libraries on ways of applying appropriate technologies for production of digital library resources;
4. Advancing the use and usability of globally distributed networks library resources.

To further boost distance and open education in tertiary institutions of learning the National Universities Commission also established the Virtual Institute for Higher Education (VIHEP) in year 2003 as part of the preparation for the eventual take-off of the National Higher Education Pedagogy Centre (NHEPC) in July 2004. The objectives of the institute as elaborated by Okebukola (2003) are as follows:

- To provide academic staff in tertiary institutions in Nigeria especially universities with Internet-based training on modern method of teaching and learning in higher education
- To enhance the knowledge and skills of academic staff on such issues as (a) teaching of large classes; (b) effective utilization of (meager) resource; (c) modern methods of assessment and evaluation of students' performance; (d) basic guidance and counseling techniques; (e) basic skills of curriculum development: and (f) techniques for writing winning grant proposals.
- To share experiences among academic staff in Nigerian universities on best practices in university teaching and how to deal with academic and social vices such as examination malpractice, plagiarism and cultism.
- To try out draft training modules for the National Higher Education Pedagogic Centre.

2. Problem Statement

Nigeria lacks the infrastructures for the full-scale deployment of e-learning facilities. This is besides the high cost of using the technology and maintaining it. Telephone density in the rural areas is still low while Internet access is concentrated only in urban centers and a few higher learning institutions. Furthermore, most of the telephones in the urban areas are within offices rather than households. This situation calls for institutions of higher learning and the private sector to prioritize investments in this sector in order to accelerate growth in offering e-Learning education programs as well as enhancing delivery of educational content in general. Poor ICT infrastructure in many of the campuses of our Universities in Nigeria can slow down the implementation of some of the technology-supported instructional methods. We hope that with a new e-learning platform and increased awareness, as well as the desire for more flexible delivery alternatives, our Universities will find reason to move even faster and extend the infrastructures to their campuses. Because of the problems stated above, it is imperative to explore the provision and integration of ICT Infrastructures in Nigerian Universities.

3. Purpose

Effective Internet access is dependent on the availability of technological, physical and support infrastructures. Despite their availability, the full benefits cannot be obtained if there is too little use of the facilities and if the level of customer satisfaction is low because of poor management of the facilities. The purpose of this study is to assess the provision and integration of e-Learning infrastructures in Nigerian Universities and the impact of infrastructures on both educator and the students. Specifically, we will explore the technological and physical infrastructure for ICT and Internet access in the campus and the support infrastructure, including financial assistance and the proportion of ICT-oriented courses of studies and of academic staff and officials with an ICT background working in the concerned sectors of the university; investigate some aspects of using the Internet.

4. Research Questions

To fulfil the purpose, this study seeks to answer the following research questions:

1. What is the status of e-learning and ICT infrastructures in Nigerian universities?
2. What are the current views of learners towards e-learning and ICT infrastructure?
3. What are the current views of lecturers towards e-learning and ICT infrastructure?
4. How do these views differ across various universities?

5. Research Methodology

A quantitative approach was used in this study. The study reviewed the relevant literature published in books, journals and websites. The research conducted a survey of three universities in Nigeria (University of Lagos, Covenant University and Babcock University respectively) to evaluate the provision and integration of ICT infrastructures in their respective schools.

5.1. Data Analysis

Given the parameters of this study a descriptive analysis was deemed as an appropriate statistic for data analysis. With the exception of a small number of open-ended questions, the responses were pre-coded for ease of data entry. The data, including comments, were first entered on to an Excel spreadsheet. Following checking and correction, the data were then uploaded onto SPSS for statistical analysis (frequency analysis and factor analysis). Questionnaires were distributed to lecturers and students at university of Lagos, Covenant University, and Babcock University in the Nigeria.

5.2. Findings

The distribution of the questionnaires was done directly during school period. The researchers went to the schools, asked for the permission of the schools' authorities, and distributed to the students and the lecturers that were around. A total of 215 valid questionnaires were obtained. Forty one percent of the participants were male and 59% were female; 12% of the participants are between 19 to 20 years old, 78% are between 21 to 25 years old, 3% are between 26 to 30, while 7% are over 30 years old.

When it was asked, how many of the participants have a personal computer or a laptop, 54% indicated they do not have either a personal computer or a laptop while 46% indicated they have personal computer or a laptop.

i. Demographic Information (Profile of the Respondent)

By means of a combined reading of the data obtained from the survey on population and the socio-demographical variables it is possible to identify the most characterizing elements amongst students and lecturers. Gender appears to be the first characterizing element:

44% of men stated that they had used Internet for school and academic work and 56% of women said the same thing. Using other information of a socio-demographic nature such as age, it is possible to say that, as regards the frequency of Internet usage, participant between 26 to 30 years old use the Internet the most for school and academic assignments.

ii. Users' Access to Computer and Internet

Teachers, students, and administrative staffs were identified as computer and Internet users in the campus. The study finds that (see Table 1) 69 out of 89 male respondents indicate that they use the Internet very frequently while only four indicate that they have never used it at all. Also, 81 out of 126 female respondents indicate that they frequently use the Internet, but five indicate that they have never use it. Based on this data, we concluded that most students had limited access to computers and Internet in some of the departments, institutes, and faculty labs. However, academic staffs had computer access in either in their offices or in the departmental offices.

Table 1: Frequency of Internet Usage for school and academic assignments

Variables	*Not at all*	*Seldom*	*Very frequent*	*Neutral*	*Total*
Male	4	11	69	5	89
Female	5	40	81	0	126

Based on this analysis and other data collected in this study, we were able to conclude that e-learning is fairly new in Nigerian universities; ICT has been in use for some time, but the infrastructures to support both are not where it supposes to be in many universities in Nigeria.

iii. Access to a range of technologies

Respondents were asked to indicate whether they had access to a list of ICT technologies (Desktop PC, Laptop computer, CD-ROM drive, Internet connection,
Fax, Digital TV, ISDN line, DVD Player and Mobile phone). They were given a number of options to indicate where these technologies could be accessed (no access, home only, university only, work only, other place only, or a combination of these options). While a number of respondents indicated that they had no access to PCs or the Internet, in fact, all students are provided with some access on campus through computer laboratories and Internet cafes at minimal cost. It is apparent from examination of the non-responses that some respondents were not familiar with the technology. Learners are anxious about e-learning and ICT, but the infrastructures are not available in many of their campuses.

Experience of ICT in education

The data on usage of email and accessing educational material were further analyzed by programme, to establish if there are any variations in patterns of usage. The first research question was answered by looking at various variables as shown in Table 2 below. This question was investigated in many ways; specifically, 15 Deans of Education from 15 universities in Nigeria were surveyed during a conference that was held at the University of Lagos. They were asked the following question: What is the status of e-Learning and ICT infrastructures in Nigerian universities?

Based on the survey of the Deans from the fifteen universities, 87% of the university has its own dedicated web site. Most of the university web sites contain the general information about the university, its faculties, institutes, departments, library, publications, admission, etc. The study reveals that only 67% of the universities has Internet connectivity for each academic staff, 64% indicated they could access computer network remotely, while 73% claimed they had computer for individual academic staff. The Deans claimed that in almost all cases, the existing infrastructures could support e-learning and ICT. How do these views differ across various universities?

Table 2: Percentage of Responses to e-Learning and ICT infrastructures in Nigerian Universities

	Computer for individual academic staff	Internet connectivity for each academic staff	Computer network accessible remotely	A dedicated website
Don't know	7%	0%	13%	0%
No	20%	33%	20%	13%
Yes	73%	67%	64%	87%
	100%	100%	100%	100%

Also, when the Deans were asked, "of their views and which of the statements in the table below they think best describes their department's approach to ICT and e-learning, their responses are discussed below:

Fifty three percent indicate they have not considered e-learning, 53% considered e-learning activities but do not think it is appropriate, 4% have considered e-learning activities but not taken it forward yet, 31% are putting in place plans for implementing e-learning, 62% are piloting e-learning activities, and 29% do not have appropriate infrastructures for e-learning and ICT.

Student ICT abilities, experiences and attitudes

The questions concerning ability to use different applications were answered on a four point scale, ranging from 1 ('beginner) through 2 competent) to 3 advanced and 4 no knowledge). These were a subset of the questions used in this project in three universities with a total of 215 respondents and which proved highly reliable.

In general, the students in our sample reported good skills with different applications. The highest abilities were found with Microsoft Office products, Internet usage, downloading information, sending emails, and playing games on PC. A slighter lower ability for web design and programming were reported.

Experience of e - learning methods

The questions with respect to experience of e-learning methods were answered on a four point scale; 1 ('never heard of this'); 2 ('used this once'); 3 ('never used this'); 4 ('used it several times'). Students at the different universities reported most experience of academic support and advice from a teacher using e-mail, followed by on-line discussion forums and a website with interactive features. In contrast, students had less experience with virtual learning environments and even less experience with video-conferencing. Given the known high use of email for support and low use of videoconferencing in undergraduate teaching at many of the universities these patterns are to be expected and confirm the reliability of the students' responses.

Internet Application Software

Among the Internet application software available, Internet Explorer was remarkably used by all the sectors (100%) of the university. Other software was less used including Netscape Navigator and SN Explorer.

Internet Access Benefits

General overviews of the responses indicate some common Internet access benefits, and options were given to answer more than one item. It is found that all respondents (100%) identified Internet as a platform for online communication most commonly via e-mail, while 91% considered Internet as a tool for academic and professional excellence. About 80% respondents mentioned that Internet facilitated them to access a wide range information sources including e-books, e-journals, digital theses and dissertations on diverse subject fields. Another benefit indicated by students was the access to basic and general information of different universities in home and abroad which would help them to know other universities, network, do home works, conduct research etc.

Major Constraints

Here are some of the specific issues the participants consider as challenges for ICT and e-learning usage.

Table 3: Frequency of Internet Usage for school and academic assignments (n=215)

	f	*%*
Not at all	8	3.7
Very Little	40	18.6
Extremely Important	167	77.7
Total	215	100

Table 4: Provision of sufficient funding for ICT infrastructures (n=215)

	f	*%*
Not at all	20	9.3
Very Little	14	6.5
Extremely Important	181	84.2
Total	215	100

Table 5: Obtaining different software that are specific enough for different course (n=215)

	f	*%*
Not at all	20	9.3
Very Little	14	6.5
Extremely Important	181	84.2
Total	215	100

The perceived barriers in tables 3, 4, and 5, which prevent or restrain ICT and e-learning usage indicated by the greatest number of people, are that of frequency of Internet Usage for school and academic assignments, cost (funding), and the ability to obtaining different software that are specific enough for different courses.

Conclusion

It should be remarked that the quantity and quality of e-learning initiatives in Nigeria is grossly inadequate and poorly maintained. The present situation should not be

allowed to continue more so as we are a country aspiring to be technological buoyant. The potential benefits, which the new technologies can bring to our educational system, if implemented and exploited effectively do not seem to be in question. However, fundamental issues need to be considered relating to the possibilities and practicalities of integrating them in a meaningful way into the system. There will be some major obstacles to their use in our higher educational institutions and the wider community. Some of these obstacles include poor infrastructure, professional competence, teachers' attitudes, cost, lack of technical expertise and skilled Manpower, conflict with the curriculum, lack of information, and content development.

Poor Infrastructure - Nigeria today lacks the infrastructures for the full-scale deployment of e-learning. This is besides the high cost of using the technology and maintaining it. Telephone density in the rural areas is still low while Internet access is concentrated only in urban centers and a few higher learning institutions. Furthermore, most of the telephones in the urban areas are within offices rather than households. This situation calls for institutions of higher learning and the private sector to prioritize investments in this sector in order to accelerate growth in offering e-Learning education programs as well as enhancing delivery of educational content in general. Poor ICT infrastructure in many of the campuses of our Universities in Nigeria can slow down the implementation of some of the technology-supported instructional methods. We hope that with a new e-learning platform and increased awareness, as well as the desire for more flexible delivery alternatives, our Universities will find reason to move even faster and extend the infrastructures to their campuses.

Professional Competence - Teacher education is considered to be the single most important factor in ensuring the successful use of ICTs in education. Its importance has tended to be overlooked or underestimated in the development of initiatives for introducing these technologies into our educational institutions with the result that many projects fail outright or are never developed to their full potential. Teacher education is not only vital for equipping educators with the necessary skills for using ICTs effectively in the classroom, but also for helping teachers to overcome their often strong resistance to these technologies and to develop positive attitudes towards them. Hence, there is the need to introduce two levels or types of training for our teachers. These include: 1) an introduction to the technologies and preparation to operate and manage the hardware; and 2) training in the pedagogical use of the technologies. The latter might pose a particular major challenge since it remains a relatively new area of teacher education.

Teacher attitudes - Teachers are likely to resist the introduction of ICTs into the classroom for a variety of reasons. Their unfamiliarity with the technologies, the additional time and effort necessary for their effective use, and perhaps the feeling that ICTs pose a threat to their professional role and image, are some of the reasons for this resistance. Hence, there should be a constantly involvement of teachers in the logistics of implementation, courseware preparation, experimentation and piloting of the e-learning system.

Costs - The price of hardware and software, although constantly decreasing, remains considerable for many of our higher educational institutions' budgets in Nigeria. They are also under-equipped with personal computers, CD-ROM drives, appropriate and adequate software packages. In addition to ordinary maintenance costs, the rapid evolution of information and communication technologies implies constant upgrading of equipment and facilities if our higher educational institutions are to keep abreast of these developments in the classroom. The various learning technologies have major differences in their cost structure. On the one hand, the major cost of using online communication technologies to provide flexible tutorial support is frequently the cost of tutors' wages as the number of hours which tutors can spend online is potentially very substantial. On the other hand, the major cost element in the use of online resource technologies is frequently the initial cost of producing the resources (especially as one goes up the levels of these resources). The cost of on-line telephone charges for use of the Internet remains high and it is a major deterrent to the use of Internet in many institutions. Installation of faster connections such as ISDN digital telephone lines or fiber optic cables, which would in theory reduce time spent on the network, is also expensive. It is however important to remark here that the Federal Government of Nigeria has already started to install fiber optic cables in all the hooks and corners of the country especially in major cities.

Lack of technical expertise and skilled Manpower - Nigeria, like most other developing countries, lack technical expertise in this domain at all levels. Their limited resources mean that the country usually has neither the local capacity to develop the necessary human resources in this field nor the means to attract highly skilled and expensive experts from abroad. Therefore, skilled manpower for the administration of the e-learning environment's technical issues is crucial to the success of the methodology. Content must be updated as and when it is necessary, hosted and made available for access. Where CD-ROMs must be cut for remote access, such activities must be done speedily and availed with minimum delay.

We need to train middle-level skilled manpower for the successful implementation of the e-learning programmes.

Conflict with the curriculum - Problems may arise during implementation that conflict with the present curriculum being used in our higher education institutions, especially with the development of educational software, which may be imposed, on teachers without their being involved in its selection, development or evaluation. For example, some educational CD-ROMs are often not curriculum-based and teachers will need to take permission or spend a lot of extra time devising appropriate ways to incorporate them into the regular curriculum. There is therefore the urgent need to update our curriculum in line with modern technological trends.

Lack of information - A significant obstacle to the use of ICTs in education in Nigeria hinges on the lack of information that is available both to educational decision makers and practitioners. The type of information lacking is broadly of two major categories i.e. information about the role and value of these technologies in education and more specific information relating to available hardware and software and how to use it in the curriculum. It is important therefore to sufficiently educate the populace.

Content development - Teachers will need to be trained especially in content development, which is the driving engine of any information system. It is often realized, after beautiful websites have been developed, that there is not enough content to justify the investment and hence failure to update them often. Also important is the development of such content into appropriate modes for use in the new instructional technologies.

E-learning materials can be delivered in different modes. However, choosing a mode will depend on such factors as proximity to the Internet facilities, access to the central server where the material is held, and access to computing facilities. Sometimes, a mix mode of e-learning with printed materials, classroom face-to-face teaching, chatting tools and even videoconferencing may be preferable. The following four main modes of rendering content can be used:

(i) Content on Media
This is content organized into logical units and made available in portable media such as CD-ROMs. The content may have limited interactivity inbuilt into it. Graphics, animations, audio and video forms of content can be included. This option is quite suitable for students in remote places.

(ii) Content on Intranet
Any content on CD can be made available within an intranet as a shared resource. The presentation of such content is largely similar to the one above but limited asynchronous interaction can be enabled by email.

(iii) Content on Web
This kind of content is similar to content on Intranet but is accessed using web browsers and therefore can be available on the Internet. An advantage of this is that content can be regularly updated and made more current.

(iv) Content on Electronic Learning Environment
An Electronic Learning Environment (ELE) is an elaborate Instructional platform that enables the development and delivery of content to learners. It also supports content

RECOMMENDATIONS

There is no doubt that the Internet has brought a great change in the nature of functions and activities in different sectors of Universities in Nigeria. Academicians are increasingly dependent on the Internet day by day. But as reflected in the present research, the infrastructures in Nigeria Universities are not at a satisfactory level. On the basis of the situation of Internet connectivity and accessibility in the campus, certain recommendations are made for future considerations.

- In most of the universities, the LAN should be extended to cover all academic buildings, teachers' and students' dormitories, medical centre, physical department, and other offices to ensure more connectivity and more accessibility.
- Initiatives should be taken to enhance the present bandwidth of VSAT data circuit so that the faster data transmission and quick access to Internet can be provided.
- Many authorities were not satisfied with the present Internet services and of the university computer centres. The university should establish a full-fledged cyber café equipped with a good number of latest computers, modern lab facilities with sufficient space to accommodate more users, and major Internet access services to meet the customer demand.
- All faculties, departments and institutes should be provided with sufficient computers and accessories to develop independent computer laboratories with Internet connection as well as to ensure more access facilities.
- The university authority should take a long-term plan to create Internet access opportunity for general students usage.
- Various faculties and departments should equally be provided with sufficient budget, and even special allocations only for ICTs on a priority basis.
- University authorities should take immediate action to automate their libraries and to introduce Internet library system. Moreover, the library, being a focal point, should establish an online networking and resource-sharing programme with the departmental seminar libraries,

institute libraries, faculty libraries and the libraries for students.

- Besides the common Internet access services, the authority should introduce some special type of facilities and value added services like- Internet telephony, fax-to-fax and voice over IP, etc.
- Since power failure has become a common problem in Nigeria, the university should take initiatives to ensure uninterrupted power supply within their capability so that Internet can be used without any hazard.
- The implementation of the suggested measures would ensure maximum utilization and benefits of the campus network. But it is beyond the university's means to implement all the measures at the same time. Therefore, a step-by-step priority-basis plan should be taken to make the recommendations more realistic and successful.

References

1. Ajayi, G.O. (2003) e-Government in Nigeria's e-Strategy; The Fifth Annual African Computing & Telecommunications Summit, Abuja, Nigeria.

2. Blinco, K., Mason, J., McLean., N., and Wilson, S., (2004). Trends and issues in e-learning infrastructure development. Available online: http://www.jisc.ac.uk/uploaded_documents/Altilab04-infrastructureV2.pdf.

3. Omwenga E. I, (2003) Ph.D. thesis. School of Computing and Informatics, University of Nairobi.

4. Peh, W. L. & Foo, S. (2001). Students' perceptions of online learning: A case study of Singapore Temasek Polytechnic's Virtual School of Business Project. Volume 11 Issue 2; September 30. Bi-annual LIBRES 11N2.

Genetic Algorithm of Resource Partition and Task Scheduling

Mieczyslaw Drabowski

Cracow University of Technology
Warszawska 24 Krakow 31-155 Poland
drabowski@pk.edu.pl

***Abstract*:** The paper presents an innovative approach to solving the problems of computer system synthesis based on genetic methods assisted with simulated annealing strategy. We describe algorithm realizations aimed to optimize resource partition and task scheduling, as well as the adaptation of those algorithms for coherent synthesis realization. We then present selected analytical experiments proving the correctness of the coherent synthesis concept and indicate its practical motivations.

***Keywords*:** task, resource, allocation, genetic, coherent, synthesis.

1. Introduction

The goal of high-level synthesis of computer systems (i.e. systems of type the complex of resources and operations) is to find an optimum solution satisfying the requirements and constraints enforced by the given specification of the system. The following criteria of optimality are usually considered: costs of system implementation, its operating speed, power consumption and dependability. A specification describing a computer system may be provided as a set of interactive tasks (processes, functions).
The partition of the functions between hardware and software is the basic problem of synthesis. Such partition is significant, because every computer system must be realized as result of hardware implementation for its certain tasks.
In the synthesis methods so far, the software and hardware parts were developed separately and then connected in process the co-called co-synthesis, which increased the costs and decreased the quality and reliability of the final product.
The resources distribution is to specify, what hardware and software are in system and to allocate theirs to specific tasks, before designing execution details.
The problems of tasks scheduling are one of the most significant issues occurring at the procedure synthesis of operating systems responsible for controlling the distribution of tasks and resources in computer systems.
The objective of this research is to present the concept of coherent approach to the problem of system synthesis, i.e. a combined solution to task scheduling and resource partition problems. The model and approach are new and original proposals allowing synergic design of hardware and software for performing operations of the computer system. This is approach, which we called a par-synthesis (coherent co-synthesis).
This research shows the results selected of computational experiments for different instances of system par-synthesis problems proving the correctness of the coherent synthesis concept and shows the methods solving this problems.
Due to the fact that synthesis problems and their optimizations are NP-complete we suggest meta-heuristic approach, i.e. genetic with simulated annealing.
Coherent co-synthesis of computer systems, as well as synergic design methodology their structures and scheduling procedures may have practical application in developing the tools for automatic aided for rapid prototyping of such systems.

2. Coherent synthesis of computer system

2.1 The classical process of computer system synthesis

The classical process co-synthesis [2], [14], [22] – hardware and software – for computer system consists of the following stages (Fig. 1.1):

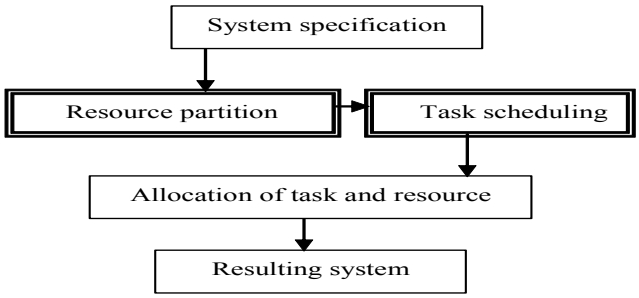


Fig. 1.1. The process co-synthesis

1. Specification of the designed system in terms functional and behavioural – requirements and constraints analysis. The system description in an high-level language, abstracting from the physical implementation.
2. Resource partition – architecture development.
3. Task scheduling – system control development.
4. Allocation the system functions to the architecture elements – generating the system modular architecture, control adaptation and the whole system integration.

The system being constructed consists of hardware elements and software components performed by selected hardware modules. The system is specified by a set of requirements to be met. In general, each requirement may be satisfied by hardware elements or software components executed by universal processors and memories. Obviously, at this stage of design, one must take into account appropriate system constraints and criteria of optimal system operation. Accordingly, the key issue in the synthesis is efficient

partitioning of system resources due to their hardware and software implementation, providing fulfilment of all requirements and the minimum implementation cost.
Such partitioning methodology [17] may accept, as a starting point, assignment of the hardware implementation to all system functions and further optimization of project costs, search for possibilities of replacing certain tasks realized by hardware with their software equivalents. Other methods [20] of the resources partitioning start with an exclusive software implementation and further search for implementation of certain tasks by hardware. In both approaches the objective is optimization of the implementation cost of the same tasks, i.e. in particular minimization of the execution time by specialized hardware [3]. Obviously the requirements and constraints, especially those regarding time and power consumption, have decisive influence upon selection of necessary hardware components.
The measure for an efficient implementation of a computer system is the degree of its modules utilization, minimized idle-time of its elements and maximized parallel operation of its elements [21].
A non-optimum system contains redundant modules or modules that are excessively efficient in comparison to the needs defined by the tasks what, consequently, increases the system cost. In high-level synthesis, the optimization of the designed system costs, speed and power consumption is usually an iterative process, requiring both changes in the architecture and task scheduling [23]. That is, why an optimum system may be created as a compromise between the system control algorithm and its hardware organization.

2.2 The general model for the problem of system synthesis

System synthesis is a multi-criteria optimization problem. The starting point for constructing our approach to the issues of hardware and software synthesis is the deterministic theory of task scheduling [4], [7], [25]. The theory may serve as a methodological basis for multiprocessor systems synthesis.
Accordingly, decomposition of the general task scheduling model is suggested, adequate to the problems of computer system synthesis. From the control point of view such a model should take into account the tasks, which may be either preemptable or nonpreemptable. These characteristics are defined according to the scheduling theory. Tasks are preemptable when each task can be interrupted and restarted later without incurring additional costs. In such a case the schedules are called to be preemptive. Otherwise, tasks are nonpreemptable and schedules nonpreemptive.
Preemptability of tasks in our approach cannot be a feature of the searched schedule – as in the task scheduling model so far. The schedule contains all assigned tasks with individual attributes: preemptive, nonpreemptive. From the point of view of the system synthesis, the implementation of certain tasks from the given set must be nonpreemptible, for the other may be preemptible (what, in turn, influences significantly selection of an appropriate scheduling algorithm) [5]. Moreover, we wish to specify the model of task scheduling in a way suitable for finding optimum control methods (in terms of certain criteria) as well as optimum assignment of tasks to universal and specialised hardware components. Accordingly, we shall discuss the system of type the complex of resources and operations:

$$\sum = \{ \mathbf{R, T, C} \} \quad (1)$$

where:
R – is the set of resources (hardware and software),
T – is the set of the system's tasks (operations),
C – is the set of optimization criteria for the system's behaviour and structure.

Resources. We assume that processor set $P = \{P_1, P_2,\ldots, P_m\}$ consists of *m* elements and additional resources $A = \{ A_1, A_2,\ldots, A_p\}$ consist of *p* elements.
Tasks. We consider a set of *n* tasks to be processed with a set of resources. The set of tasks consists of *n* elements $T = \{T_1, T_2,\ldots, T_n\}$. A feasible schedule is optimal, if its length is minimum and it is implemented using minimum resource cost.
Each task is defined by a set of parameters: resource requirements, execution time, ready time and deadline, attribute - preemptable or nonpreemptable. The tasks set may contain defined precedence constraints represented by a digraph with nodes representing tasks, and directed edges representing precedence constraints. If there is at least one precedence constraint in a task set, we shall refer it to as a set of dependent tasks, otherwise they are a set of independent tasks.
Optimality criteria. As for the optimality criteria for the system being designed, we shall assume its minimum cost, maximum operating speed and minimum power consumption.
The proposed model may be used for defining various synthesis problems for optimum computer systems.
The model of a system in our approach, [9], [11] typical for the theory of task scheduling, consists of a set of requirements (operations, tasks) and existing relationships between them (related to their order, required resources, time, readiness and completion deadlines, preemptability/nonpreemptability, priority etc.). The synthesis procedure contains the following phases: identification of hardware and software resources for task implementation, defining the processing time, defining the conflict-free task schedule and defining the level of resource co-sharing and the degree of concurrency in task performance.
The synthesis has to perform the task partitioning into hardware and software resources. After performing the partition, the system shall be implemented partially by specialized hardware in the form of integrated circuits (readily available on the resources pools or designed in accordance to the suggested characteristics) [18]. Software modules of the system are generated with the use of software engineering tools. Appropriate processors shall be taken from the resource pool. Synthesis of a system may also provide a system control, create an interface and provide synchronization and communication between the tasks implemented by software and hardware [11].
The system synthesis, i.e. defining system functions, identifying resources, defining control should be implemented in synergy and be subject to multi-criteria optimization and verification during implementation.

2.3 The coherent process of system synthesis

Modeling the joint search for the optimum task schedule and resource partition of the designed system into hardware and software parts is fully justified. Simultaneous consideration of these problems may be useful in implementing optimum

solutions, e.g. the cheapest hardware structures. Synergic approach enables also performing of all assigned tasks with the minimum schedule length. With such approach, the optimum task distribution is possible on the universal and specialized hardware and defining resources with maximum efficiency.

We propose the following schematic diagram of a coherent process of systems synthesis [10], (Fig. 1.2). The suggested coherent synthesis consists of the following steps:

1. specification of requirements for the system to be designed and its interactions with the environment,
2. specification of tasks, including evaluation of task executive parameters using available resources (e.g. execution times),
3. assuming the initial values of resource set and task scheduling – initial resource set and task schedule should be admissible, i.e. should satisfy all requirements in a non-optimum way,
4. task scheduling and resource partitioning,
5. evaluating the operating speed and system cost, multi-criteria optimization,
6. the evaluation should be followed by a modification of the resource set, a new system partitioning into hardware and software parts (step 4).

Iterative calculations are executed till satisfactory design results are obtained – i.e. optimal (or sub-optimal) system structure and schedule. The designed system should be fast and cheap.

3. The genetic method for coherent synthesis of computer system

This chapter presents a coherent approach to solving the problems of computer system synthesis based on genetic method assisted with simulated annealing strategy. We describe algorithm realizations aimed to optimize resource partition and task scheduling, as well as the adaptation of those algorithms for coherent co-synthesis realization. We then present selected analytical experiments proving the correctness of the coherent synthesis concept and indicate its practical motivations. Due to the fact that synthesis problems and their optimizations are NP-complete [6], [15] we suggest meta-heuristic approach, genetic with Boltzmann tournament selection strategy [1], [12], [24].

In order to eliminate solution convergence in genetic algorithms, we use data structures which ensure locality preservation of features occurring in chromosomes and represented by a value vector. Locality is interpreted as the inverse of the distance between vectors in an n-dimension hypersphere. Then, crossing and mutation operators are data exchange operations not between one-dimensional vectors but between fragments of hyperspheres. Thanks to such an approach, small changes in a chromosome correspond to small changes in the solution defined by the chromosome. The presented solution features two hyperspheres: task hypersphere and resource hypersphere.

The solutions sharing the same allocations form the so-called clusters. The introduction of solution clusters separates solutions with different allocations from one another. Such solutions evolve separately, which protects the crossing operation from generating defective solutions. There are no situations in which a task is being allocated to a non-allocated resource. Solution clusters define the structures of the system under construction (in the form of resources for task allocation). Solutions are the mapping of tasks allocated to resources and task scheduling. During evolution, two types of genetic operations (crossing and mutation) take place on two different levels (clusters and solutions).

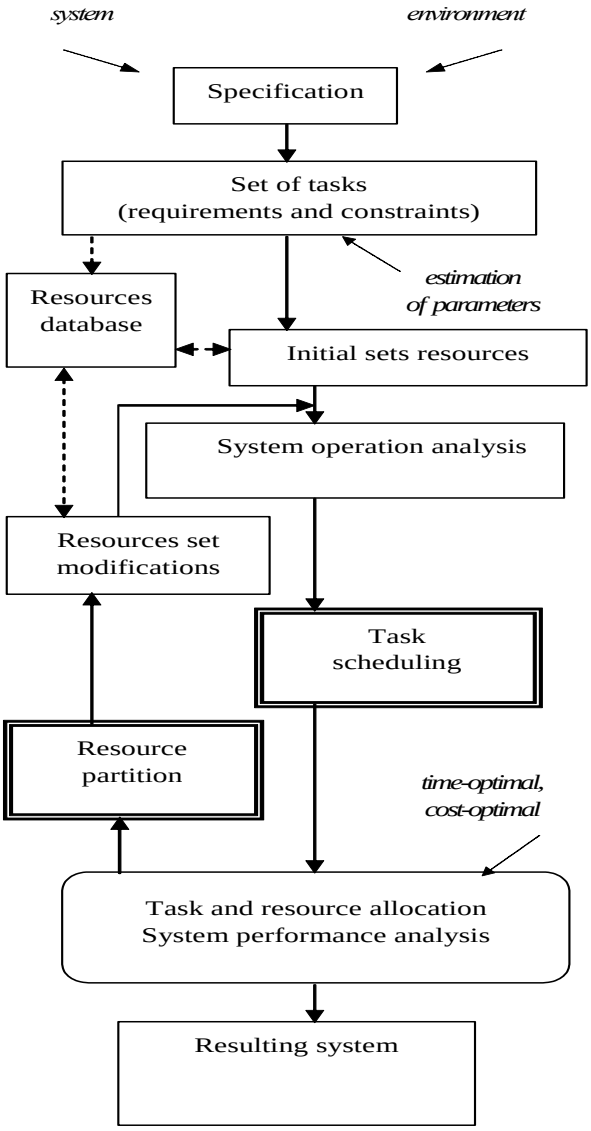


Fig. 1.2 . The coherent process of computer system synthesis

A population is created whose parameters are: the number of clusters, the number of solutions in the clusters, the task graph and resource library. For the synthesis purposes, the following criteria and values are defined: optimization criteria and algorithm iteration annealing criterion if solution improvement has not taken place, maximum number of generations of evolving solutions within clusters, as well as the limitations - number of resources, their overall cost, total time for the realization of all tasks, power consumption of the designed system and, optionally, the size of the list of the best and non-dominated individuals.

3.1 Data structures

3.1.1 The structure "Population"

This structure contains information about individuals' population:

- The table of clusters of solutions.
- Hyperspheres the resources and graph of tasks.
- Number of generation of clusters.
- Number of generation of solutions in cluster.
- Maximum quantity of generations of evolving solutions inside the clusters.
- Criterion of stop of evolution - the maximum quantity of generations without improvement of solution.
- Number of generations without obtainment of improvement of solution.
- Keeping table the the best, dominated encountered during evolution solution.
- Map of costs of optimization.
- Criterion of optimization - the maximum quantity of processors, maximum cost, maximum time, power maximum consumption.
- Dimension the list of the best solutions.
- Probability of crossing of individuals. The of mutation even probability is 1 - the probability of crossing.

3.1.2 The structure "Clusters of solutions"

The structure contains the information about cluster of solutions possessing the same the allocation of resources:

- Area describing the allocation of resources.
- Total price of allocated processors.
- Ranking of clusters (the sum of rankings of solutions inside the cluster).

Use of clusters of solutions about the same allocation has on aim the separating from me the solutions about different alokacjach. Solutions such evolve separately. In it secures oneself then the operation of crossing before production the defective solutions. It does not come to situation such that task be becomes attached to supply which he does not be allocate.

3.1.3 The structure " Solutions in demand"

The structure contains the information about the outcome structure and functionality of architecture of system:

- The area describing the attributing to resources the tasks.
- The table of optimized costs.
- Ranking of solution (the quantity of solutions in population which did not dominate this solution)

3.1.4 The structure "Allocation of resources"

The structure contains about allocated resources in frames of cluster of solutions:

- The table of solutions about the same allocation of supplies.

3.1.5 The structure "Attributing to resources tasks"

The behaviour of tasks describes in system (attributing, schedule):

- The table of list describing order in a row on individual processors tasks. Every list responds one allocated processor.
- The table including the times of beginning and end of executing the tasks.
- The table the describing allotment of tasks to allocated resource.

3.1.6 The structure "Graph of tasks"

The structure contains the information about of graph of tasks describing the functional requirements of system:

- Number of tasks in vice - count
- The table of sorted nods of graph.
- The area describing construction of graph (matrix of incidences)
- The area constains sorted in order of tasks by the BFS algorithm.
- The area of nods of graph.

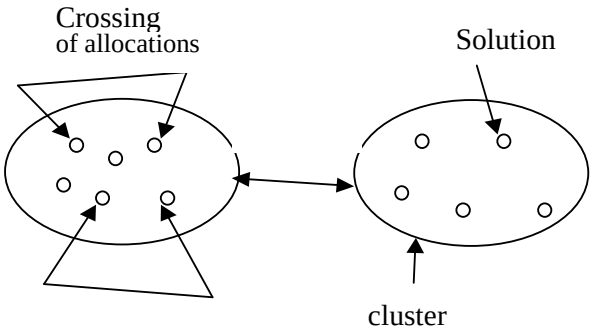


Fig. 3.2. The operations of crossing on different structures of data. Prevention of formation of defective solutions.

3.1.7 The structure "Nod of graph of tasks"

The structre describing the nod in graph:

- Number of nod.
- Level in graph.
- Predecessors' list.
- Successors' list.

3.1.8 The structure "Resources"

It contains information describing available resources:

- Number of processors
- Number of features describing the given processor.
- The area of structures describing the processor.

3.1.9 The structure "Processor"

It contains information describing processor:

- Type (universal, dedicated).
- Cost of operating memory.
- Cost of processor.
- The area of times of executed through this processor the tasks.
- The area of power average consumptions tasks.

3.1.10 The structure "Global temperature"

It contains information describing the global temperature of algorithm:

- Current temperature.
- Ratio of cooling.
- Step of temperature.

During working of algorithm, the temperature will diminish with function peaceably,

$$e^{-a \cdot x} \quad (2)$$

where *a* - the ratio of cooling. The workings about step the algorithm of reducing the temperature the argument *x* be reduced in time.

3.1.11 The Structure "Task scheduling"

It contains the functions for scheduling of the tasks:

- Initial scheduling, ASAP algorithm.
- Mutation of schedule.
- Crossing of schedule.
- Function for the counting schedule length.

3.1.12 Struktura "Hiperspher the features of system"

It contains the information the relating similarities of features of processors and tasks.

- Co-ordinates of centre hiperspher.
- Length of diameter
- Factors hiperplane cutting hiperspher.
- Distance of all vectors inside the hiperspher from centre.

The structure of system is represented by file of linear tables of data. During it crossing it comes to exchange of data among tables. It unfortunately, many problems were not it been possible was to describe with the help of the one dimension of series of data. Linear order usually forces upon on optimized data [8], [13].

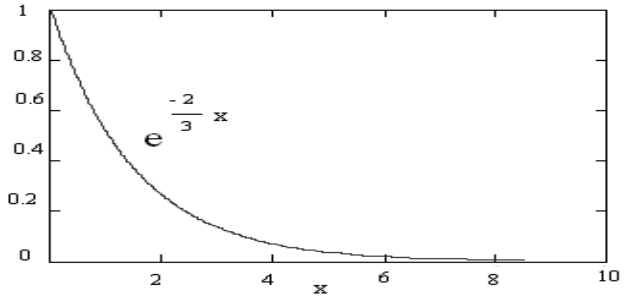


Fig. 3.3. Chart of reducing the temperature of algorithm

The evolutionary algorithm, to he could act skilfully, need in of data representing the solution structure the behaviours of lokalności. The exchange of data among individuals (the crossing) she should separate the information the describing more similar features of architecture more more seldom the than information the describing entirely different features [19]. Small changes in genotype should answer in solution which genotype represents small changes.

Putting on linear order on multidimensional information, wears out lokalność becomes. This problem the structure of data representing hipersferę in aim of solution was applied.

The multidimensional information becomes recorded in figure of vector. We interpret as reverse of local distance n - dimension vectors inside n - dimension hiperspher.

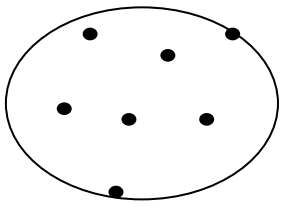


Fig. 3.4. Two dimension hipersfera (circle). Resources be described by two features here, e.g. time and cost.

The furthest distant from me vectors mark diameters and centre hyper sphere.

Algorithm keeps two hyper sphere:

- Task hyper sphere - two the dimensional, representing task graph structure. Each of the nods is defined by two coordinates: an indicator obtained through topological sorting (the tasks are "closest" if one of them is adirect successor of the other), and an indicator calculated from the BFS algorithm parallel tasks are equally distant from the beginning of the graph).
- Resources hyper sphere is three-dimensiona representing the depedencies of resource features. Each of the resources may be defined by the following coordinates: cost, speed and power consumption.

3.2 Partition of resources

It is the data the graph of tasks, pool of resources as well as criterions of optimality. The algorithm of partition of resources has determine resources, which have execute all tasks with all criterions.

3.2.1 Initial of algorithm

The aim initial of algorithm is of the construction of architecture of system the simplest and first. The architecture of system must base of accessible resources and realise required functions and set criterions. Algorithm executes following steps:

3.2.1.1 Construction of graph of tasks

On basis of input data the structure the describing graph of tasks is built. The graph of tasks represents the functionality of system. After creation of graph, nods be sorted. The topological order defines the position of tasks in graph. Equivalent levels become for nodes in graph. This features of tasks are used in scheduling algorithm. If tasks will be scheduling according to levels in graph then it will assure throwed of order constarints. The levels of tasks on following graph were marked.

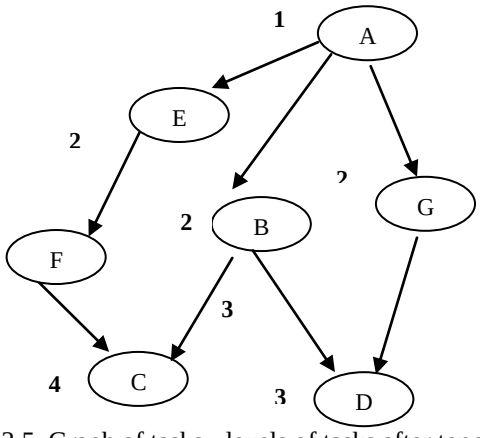


Fig. 3.5. Graph of tasks - levels of tasks after topological sorting

The searching of graph is for partition of resources the next step with the help of the algorithm the BFS [12]. The nodes of graph are assigned equivalent indices. Indices these keep the information relating the is parallel of tasks in graph. The following drawing represents the levels of tasks formed in result searching the graph with assistance of BFS.

3.2.1.2 Creation resources

On basis of data input, the representing the accessible resources object be built (the processors). Object this throws open the relating the resources information (the cost, the cost of operating memory, times of executing on processors the

tasks, averages the power consumptions, the relative speed of processor, power relative consumption).

3.2.1.3 Creation population

They are the parameters of population:

- Number of clusters in population.
- Number of solutions in clusters,
- Graph of tasks - the functionality.
- Accessible resources.
- Criterion of alloy - defines the quantity of loop of algorithm when the improvement of solution did not happen.

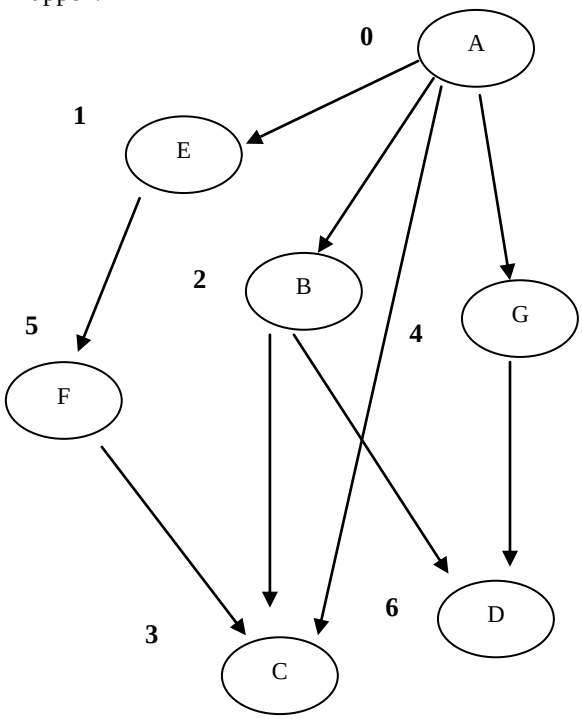


Fig. 3.6. The indices of nodes after searching the grafu by the BFS

- Map of costs - defines, which of criterions of optimization will be the taken into account during finding the optimum solution.
- Regard for universal processors the costs of operating memory during optimizing cost.
- Maxiumum number of generations of evolving solutions inside the clusters.
- Criterions of optimization - the maximum number of processors, maximum cost, maximum time, power maximum consumption.
- Size the best individuals' letters. Nodominated and the best individuals to this list be recorded. They longer list this algorithm can remember suddenly more individuals. List behaves how queue FIFO.

The created objects of clusters and solutions in clusters, and also objects hypersphere: processors and graph of tasks. The created also the object of global temperature of algorithm. The global algorithm "temperature" is initialized at this stage as well.

3.2.1.4 Initialization hypersphere

Two hypersphere be created: processors and graph of tasks. The hypersphere of processors has since 1 to 3 dimensions. The dimension depends from number of optimized features (cost, time, power consumption). Hypersphere for the graph of tasks is two dimension.

3.2.1.5 Definitions of hyper sphere

- Filling multi-dimensional vectors with data defining a given object (resources, tasks).
- Calculating the diameters of the hyper spheres, i.e. the distance between the two most remote points and determining the hyper sphere center on the basis of the extreme coordinate.

3.2.1.6 Population initialization

- Clusters and solutions are initialized randomly.
- For every task, a resource capable of completing the task is selected.
- If the resource is allocated, the algorithm proceeds to the next task.
- A resource capable of completing the task is selected and they are allocated.

3.2.1.7 Initializing the allocation of tasks to resources

- A vector of resources for allocation is taken for each task.
- Resource type and number are randomly assigned to the tasks.
- Task scheduling by the ASAP (As Soon As Possible) algorithm is initialized - eliminates the violations of sequence limitations.

3.2.1.8 Solution evaluation

- The following are calculated: resource cost, task completion time and power consumption; the cost is the sum of allocated resources' costs, the time of completed tasks is the time of completing the tasks on all allocated resources, power consumption is the sum of power inputs taken by selected resources.
- If for an individual representing a solution any of the optimized criteria exceeds the maximum value acceptable, the individual is punished and the survival chances of a punished individual diminish considerably.
- As the result of the above operations, we obtain a vector containing the values of optimized criteria (time, cost, power consumption).
- A solution ranking is determined (the rating of a given solution is the number of solutions in a population which do not dominate the solution).
- A solution is dominated if each of its costs is lesser from or equal to the costs of the dominant solution (for optimization in the Pareto sense) [5].

3.2.1.9 Cluster evaluation

The solution cluster ranking is created. The rating of a cluster is the sum of the ratings of all solutions within the cluster.

3.2.2 Resource selection

The input data for resource selection are the task graph, the library of available resources and the optimization criteria, and its goal is to partition tasks into the software and the hardware part and to select resources for the realization of all tasks consistent with the established optimization criteria. The diagram of the algorithm of resource selection is showed on Fig. 3.7.

3.2.2.1 *Cluster reproduction*

Clusters are reproduced with the use of genetic operators: crossing and mutation. At the reproduction stage, the cluster population is doubled and its initial size is restored at the elimination stage. This method was introduced arbitrarily and ensures that within a population some new individuals appear and fight for survival with their parents. The mutation operator creates one and the crossing operator two new clusters. The likelihood of using either of the genetic operators is defined by the algorithm parameters.

3.2.2.2 Genetic operators

The cluster mutation operator consists in mutating allocation vectors in the following way: a cluster with identical likelihood is picked at random and copied. The number of the resource which will be mutated in a new cluster is picked randomly Then, a number in the 0-1 range is picked - if the number is smaller than the global temperature, the resource is added, otherwise it is subtracted. Adding resources is limited by the maximum resource number parameter. At the beginning of the algorithm operation, resources will be added to the structure. As the algorithm approaches the end of the run defined by the cooling process, resources will be subtracted. This is aimed at creating a cost-effective structure. The cluster crossing operator consists in randomly picking two clusters and copying them. Crossing is achieved through cutting the resource hyper sphere with a hyper plane. The information contained on "one side" of the hyper plane is exchanged between clusters – Fig. 3.8.

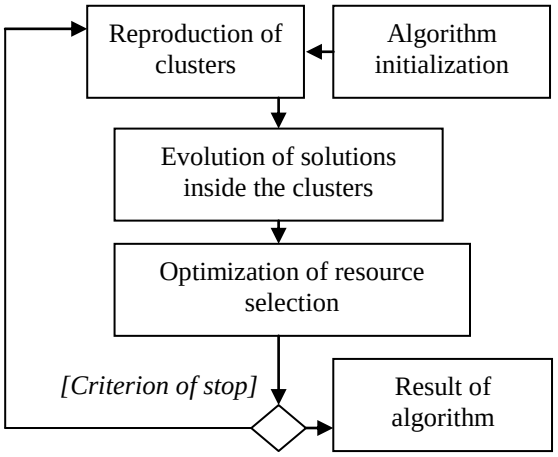


Fig. 3.7. Algorithm of resource selection

3.2.2.3 The algorithm for cutting the hyper sphere with a hyper plane

- Determining the cutting hyper plane by picking n points inside an n-dimensional hyper sphere.
- Creating a random permutation, e.g. for n = 3, the permutation can be (2, 1, 3).
- Constructing the point displacement vector in respect to the hyper sphere center; square coordinates are picked consistent with dimension permutations, e.g. for three dimensions with the permutation (2, 1, 3): y2 = rand() % r2, x2 = rand() % (r2 – y2), z2 = rand() % (r2 – (y2 + x2)), where: r – hyper sphere radius, and (x, y, z) are the coordinates of the constructed point in a three-dimensional space.
- The roots of square coordinates are calculated.
- A coordinate radical sign is picked.
- The hyper sphere center coordinates are added to the new point resulting in obtaining a new point inside the n-dimensional hyper sphere.
- The equation of the hyper plane cutting the hyper sphere is calculated and the obtained system of equations is solved.

3.2.2.4 Saving the best solutions

After solution reproduction, a new procedure is called to save the globally non-dominated solutions generated during evolution. This procedure executes:

- Searches for non-dominated solutions in the present generation.
- Creates the ranking of the best solutions saved so far and in the present generation.
- Saves the non-dominated solutions from both the "old" and the "new" solutions.
- Deletes the solutions saved in the past if they were dominated by new solutions; if there are more than one solution whose all optimized criteria values are identical, only one of those solutions is saved (the "newest" one).
- If the new solutions dominated none of the ones saved in the past, the population was not improved.
- The number of non-dominated solutions that the algorithm can save is defined by an algorithm parameter.

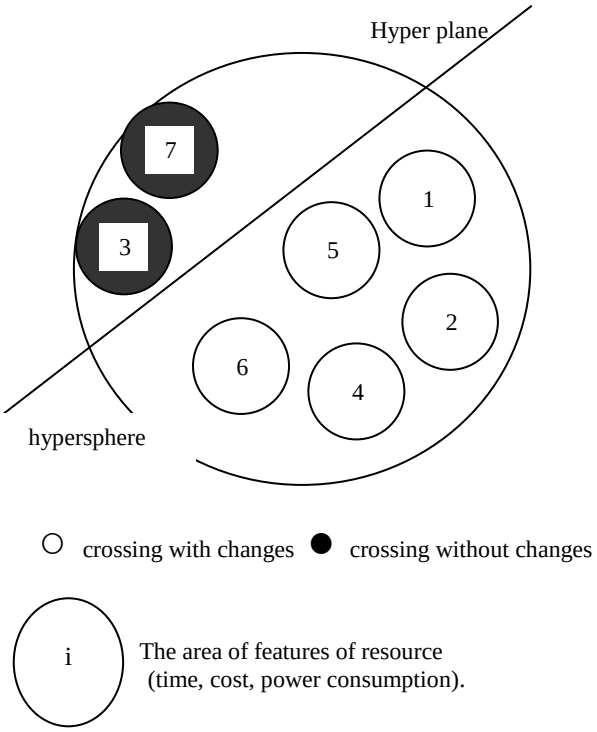


Fig. 3.8. The crossing operator with the hyper plane

3.2.2.5 Cluster evaluation

At this stage of the algorithm, half the individuals are removed from the population. The initial number of individuals is restored. The elimination of individuals is carried out using Boltzmann tournament selection strategy.

3.2.2.6 Boltzman tournament

The calculations of following equation the winner of tournament be appeared on basis of result:

$$\left[1 + e^{\frac{(r1-r2)}{T}}\right] \quad (3)$$

where:

- r1 - ranking of first solution
- r2 - ranking of second solution
- T - global temperature

They are values of this function the number from compartment from < 0,1 >. We draw in aim delimitations the winner of tournament number from compartment (0,1). If she is larger from enumerated number with example then individual about ranking is winner *r1*. Second individual in opposite incident winner is (about ranking *r2*) [1].

It the analysis of results of tournament was it been possible was to conduct on basis of graph of function (Fig. 3.9.):

$$\left(1 + e^{x}\right)^{-1} \quad (4)$$

where:

$$x = \frac{(r1-r2)}{T} \quad (5)$$

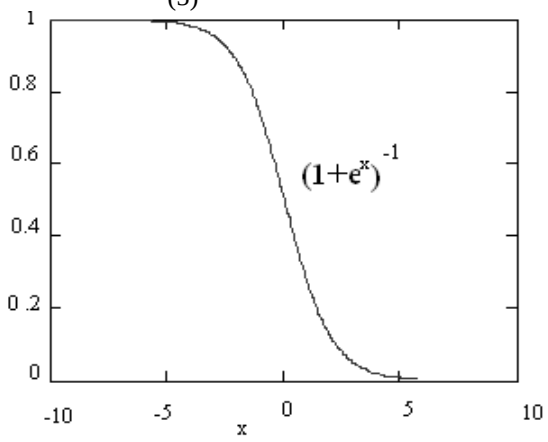


Fig. 3.9. The chart of probability of victory Boltzman tournament in dependence from global temperatur

If *r1* < *r2* this *x* is negative and for high temperature larger probability exists ,that individual about rank *r1 will* win tournament than for lower temperatures. For low temperatures winner the most often will be individual about rank *r2*.

If *r1* > *r2* this *x* is positive and for high temperature larger probability exists ,that individual *r2 will* win tournament than for lower temperatures. For low temperatures winner the most often will be individual about rank *r1*.

3.2.2.7 Report of algorithm

If the quantity of generations individuals' improvement during which did not happen, crosses the broadcast in criterion of alloy quantity, algorithm finishes his working. The dominated osobniki in scale of whole evolution become considered in report.

3.3 Scheduling of tasks

Task scheduling is aimed at minimizing the schedule length (the total tasks completion time).

3.3.1 Algorithm initialization

The scheduling algorithm initialization resembles the initialization of resource selection algorithm. The difference is that there is solely one cluster in which solutions evolve. The cluster allocation remains unchanged during the algorithm's run because all the resources are known for the task scheduling algorithm.

3.3.2 Algorithm of task scheduling

The diagram of the algorithm of task scheduling is showed on Fig. 3.10.

3.3.3 Solution reproduction

Solutions are reproduced using the genetic operators: crossing and mutation. Solutions are reproduced until their number doubles (the number of new solutions has been chosen arbitrarily).

The mutation operator produces one and the crossing operator two new solutions. The likelihood of using either of the genetic operators is defined by the algorithm parameters.

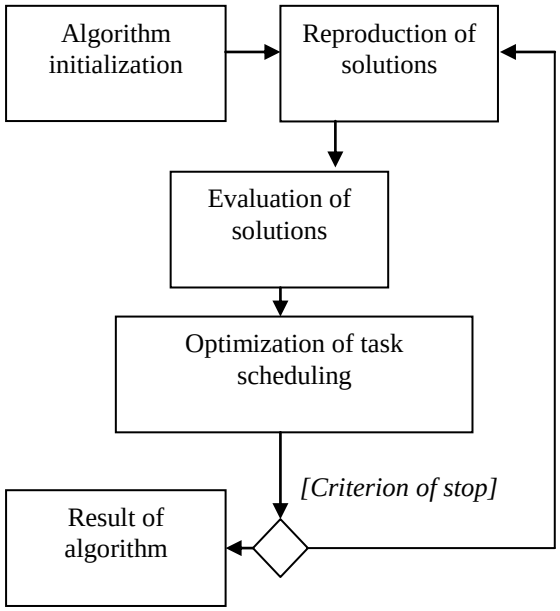


Fig. 3.10. Algorithm of task scheduling

3.3.4 Genetic operators

- **The mutation operator of task allocation to resources** acts in the following manner: a solution is randomly selected and copied. Then, the number of tasks in the system is multiplied by the global temperature. When the global temperature is high, the number of tasks changed in the allocation to resources will be greater than that in later stages of the evolution. Tasks are picked at random and allocated to resources.
- **The schedule mutation operator** acts in the following manner: if due to the mutation operation of task allocation to resources, the resource the task had been running on was changed, then the task is removed from the schedule for the "old" resource and boundaries are set on the new resource schedule between which the task may be

allocated. A location within the boundaries is picked and the task is allocated.

- **The crossing operator of task allocation to resources** resembles cluster crossing, however, the task graph hyper sphere is used for that purpose.
- **Schedule crossing operator** acts in the following way – after the allocations have been crossed, a map is created defining which parent a given feature of an offspring comes from. The offspring stores the allocation vector (obtained after crossing task allocations to resources) and the empty vector of lists with schedules of tasks on available resources. The algorithm analyzes the tasks by checking their position on the graph. For all tasks in one position, the resources on which the tasks will be performed (defined by the vector of allocation to resources) are put on the list. If in a position there is only one task ran on a given resource, the task is entered into the resource schedule, otherwise the tasks are sorted according to the starting time they had in the parent and are placed in the schedule in ascending order.

3.3.5 Solution evaluation, saving the best solutions and solution elimination

They are the same algorithms which were employed in the resource distribution algorithm. Analogical solutions are eliminated using Boltzmann tournament selection strategy [1].

3.3.6 Algorithm report

If within the number of generations determined by the annealing criterion a better individual did not appear, the evolution is stopped and the evolution report is created. The result of the algorithm operation is a set of non-dominated individuals (in the scale of the whole calculation process).

3.4 Coherent resource partition and task scheduling

The diagram of the algorithm of the coherent resources selection and tasks scheduling according to genetic approach, is showed on Fig. 3.11. The initialization of the coherent synthesis algorithm resembles the initialization of resource selection algorithm. The input parameters are the number of clusters in the population and the number of solutions in clusters. Solution clusters represent the structures sharing the same resource allocation, but with different task allocation to resources and different schedules.

The outer loop of the algorithm (realizes resource selection) is ran until the number of generations without population improvement is exceeded. This value is defined by the annealing criterion parameter. There are few outer loops at the beginning of the algorithm operation.

The number of iteration of internal loop algorithm be definite (Fig. 3.12.):

$$f(x) := -k \cdot \left(e^{-a \cdot x}\right)^3 + k \qquad (6)$$

where the k - the parameter of algorithm,

a - the annealing parameter.

Argument *x* with < 0, n >, he in every generation be enlarged about step of temperature.

N - value near which temperature is levels 0.00001.

Their number grows until it reaches the value of k with the falling of the temperature. Fewer task allocations and scheduling processes are performed at the beginning. When the temperature falls sufficiently low, each inner loop has k iterations. The number of iterations may be regulated with the temperature step parameter. The greater the step, the faster the number of inner iterations reaches the *k* value.

3.5 Computational experiments

3.5.1 The comparison coherent and non coherent synthesis with genetic algorithm. The results for tasks of dependent and nonpreemptive without cost of operating memory.

We present the analytical results obtained by testing the presented algorithms. In the tests represented by the tables and flowcharts below, we compared the results from the incoherent and coherent synthesis.

3.5.1.1 Minimize of cost

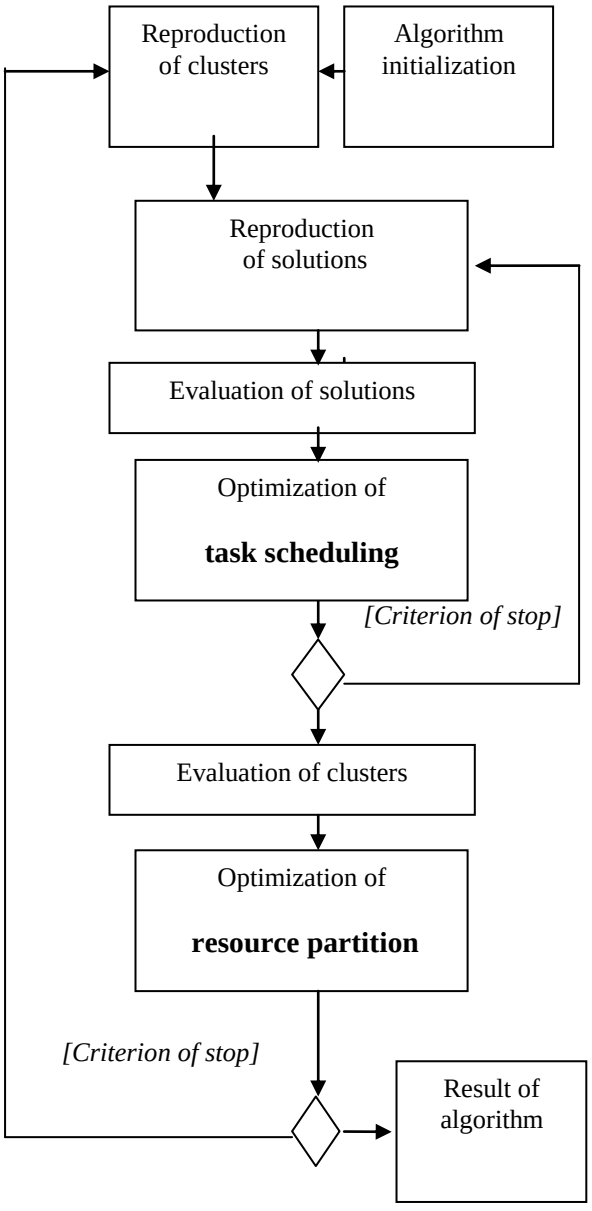


Fig. 3.11. The coherent synthesis of computer system – genetic approach

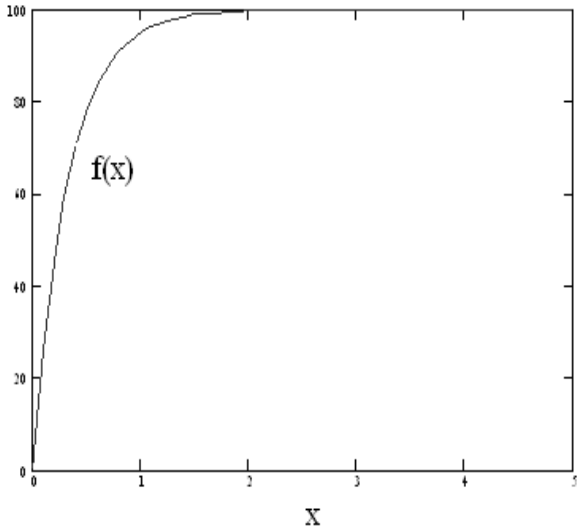


Fig. 3.12. Graph f(x) for k = 100

Table 3.1.

Tasks dependent. Minimum of cost.

Number of task	Non-coherent		Coherent		
	Cost	**Time**	**Cost**	**Time**	Power consumption
5	1.0	5.67	1.0	5.67	**11.68**
10	1,25	7.75	1.25	7.75	**28.94**
15	1.5	8.4	1.5	8.3	**54.42**
20	1.5	11.4	1.5	8.4	**42.64**
25	1.5	14.2	1.5	14	**80.18**
30	1.5	17.6	1.5	17.5	**103.12**
35	2.5	15.75	2.5	12.5	**101.04**
40	2.5	18.25	2.5	12.1	**129.17**
45	2.5	19,5	2.5	19.4	**126.95**
50	2.75	19.4	2.75	15.9	**124.67**
55	**2.75**	**18**	**2.75**	**14.7**	**147.32**

3.5.1.2 Minimize of time

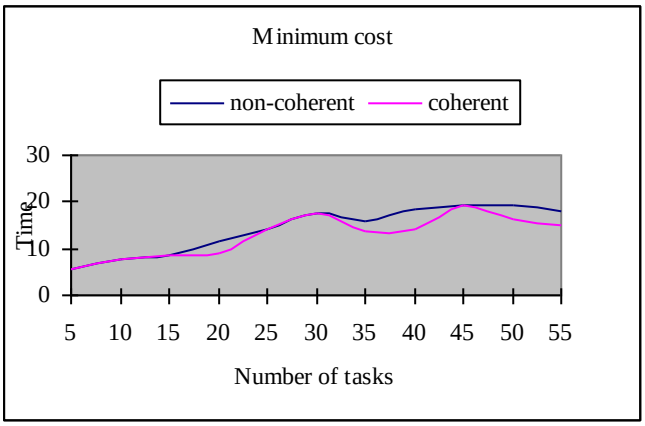
Table 3.2.

Tasks dependent. Minimum of processing time.

Number of task	Non-coherent		Coherent		
	Cost	**Time**	**Cost**	**Time**	Power consumption
5	1.00	5.57	1.8	5	**26,57**
10	1.80	7.40	1.8	7.45	**31,98**
15	3.5	7.70	3.1	6,9	**73,59**
20	3.6	8.65	3.7	7.1	**97,63**
25	4.2	7.95	3.9	7,7	**105,13**
30	4.1	9.20	2.9	7.8	**121,14**
35	5.6	8.45	5.5	7.77	**158,2**
40	7.1	8.65	6.35	7.7	**168,82**
45	8.6	8.6	5.1	5.5	**230.11**
50	8.4	8.65	7.3	7.45	**190.64**
55	**9.51**	**9.53**	**7.7**	**7.95**	**222.78**

3.5.1.3. Charts

Chart 3.1.



During cost optimization, both algorithms yielded similar cost values for all tested task sets. However, the coherent algorithm improved time optimization for graphs exceeding 30 tasks. For 50 tasks it achieved a 15% improvement of the task completion time Chart 3.1.

Chart 3.2.

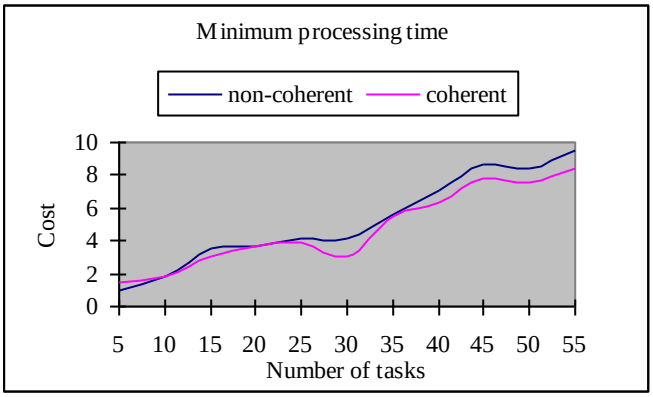
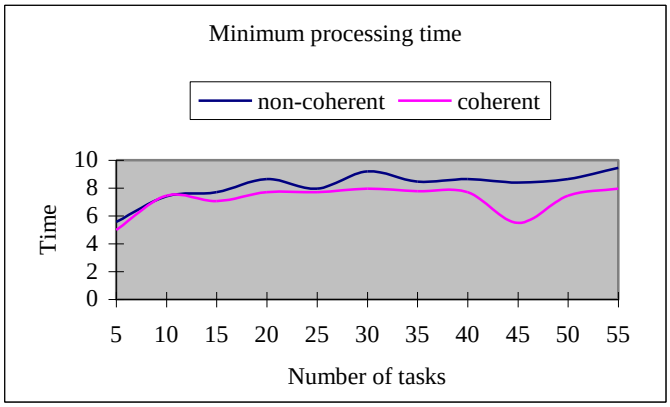


Chart 3.3.



When the flowchart reflecting the dependence of time from the number of system tasks is considered (Chart 3.2), time

minimization is comparable for both algorithms. Nevertheless, once the chart showing the interdependence of cost and the number of tasks is analyzed (Charts 3.3), it is clear that the solutions yielded by the coherent algorithm are far less expensive than those from the incoherent algorithm. The coherent algorithm achieves similar task completion times in solutions much cheaper from those found by the incoherent algorithm.

3.5.2. The comparison coherent and non coherent synthesis with genetic algorithm. The results for tasks of dependent and nonpreemptive with cost of operating memory.

3.5.2.1 Minimize of cost

Table 3.3.

Tasks dependent. Minimum of cost.

Number of task	Non-coherent		Coherent		
	Cost	**Time**	**Cost**	**Time**	Power consumption
10	1.9	7.27	1.5	7.6	**35.47**
20	2.25	11.99	2	12.25	**40.17**
30	2.25	15.33	2.25	15.4	**86.39**
40	2.5	18.67	2,5	18.66	**111.1**
50	2.75	20.6	2.73	20.25	**166.87**
60	3,25	29.25	2.7	17.8	**242.29**
70	2.5	32	2.5	32	**201.29**
80	2.75	28.8	2.39	28.8	**380.28**
90	2.25	48.25	2.25	42.1	**247.93**
100	2.6	45.6	2.2	45.6	**311.85**
110	**2.6**	**56.25**	**2.2**	**50.4**	**320.98**

3.5.2.2 Minimize of time

Table 3.4.

Tasks dependent. Minimum of time.

Number of task	Non-coherent		Coherent		
	Cost	**Time**	**Cost**	**Time**	Power consumption
10	1.2	13.95	1.5	9.8	**30.77**
20	2.5	19.29	2.2	15.,95	**59.64**
30	3.67	15.45	3.3	12.25	**107.22**
40	4.4	15.85	4.4	13.45	**159.94**
50	5.5	15.8	5.1	15.05	**197.36**
60	5.6	21.45	5.7	13.45	**252.37**
70	7.3	20.15	7.7	16.40	**340.48**
80	7.6	17.45	7.2	16.3	**242.51**
90	8.5	24.45	7.7	20.15	**302.94**
100	10	19.75	7.9	18.85	**358.76**
110	**10.33**	**24.8**	**9.1**	**20.8**	**421.98**

3.5.2.3 Charts

Chart 3.4.

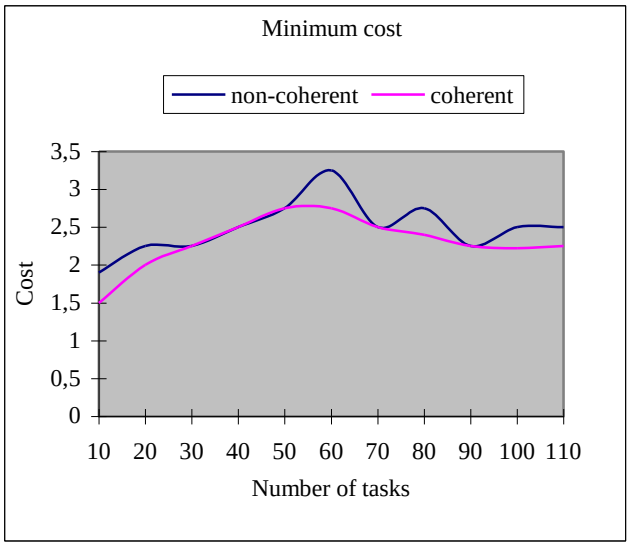


Chart 3.5.

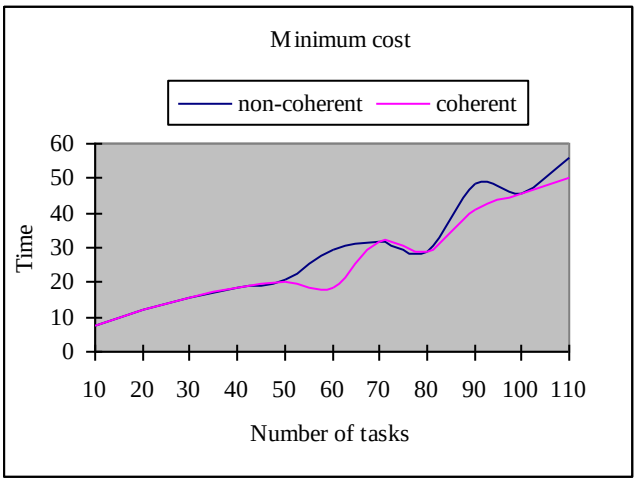


Chart 3.6.

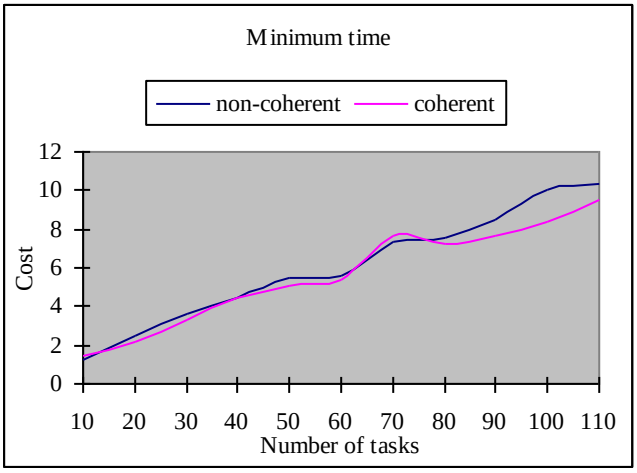
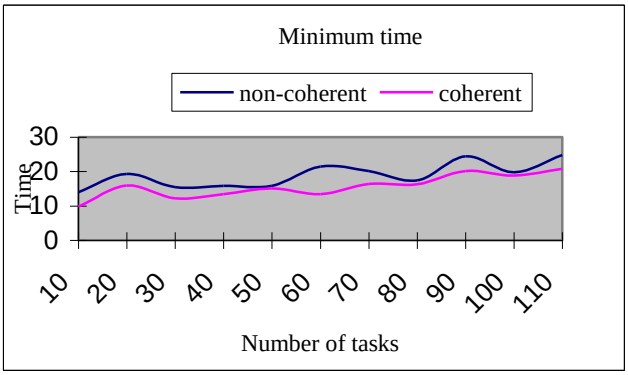


Chart 3.7.



The coherent algorithm improved time optimization. The solutions yielded by the coherent algorithm are far less expensive, as well. During minimization of time the got architectures are complex and then the scheduling algorithm has the larger possibilities for optimization. This is consequence of this that coherent algorithm gets better results than non-coherent algorithm.

3.5.3. Mult-icriterions optimization. The optimization of time of executing, power consumption and cost. Results for dependent tasks.

Tests were conducted for nonpreemptive and dependent tasks. Parameters of constraints: the maximum number of processors - 5, maximum cost - 3, maximum time 25. Optymalizowane simultaneously. It the area of optimum solutions in result was received was in sense Pareto [166]. The following tables presented solutions (in Pareto area) for the cost, time and power consumption and solution "compromissing". Searching space of solutions be led to time when global temperature reached value 0.01.

3.5.3.1 Minimize of cost

Tab. 3.5.

Multi-criterions optimization. Minimum of cost.

Number of tasks	Coherent synthesis		
	Cost	**Time**	Power consumption
5	0.5	17	**6.47**
10	0.75	15.5	**15.6**
15	1.5	8.4	**54.42**
20	1	19	**42.64**
25	2	15.75	**48.51**
30	2.25	18.4	**70.51**
35	1.5	20.8	**114.05**
40	2.75	17.75	**104.68**
45	2.25	24.67	**102.02**
50	2.25	24.25	**108.48**
55	**2.5**	**25**	**164.58**

3.5.3.2 Minimize of time

Tab. 3.6.

Multi-criterions optimization. Minimum of time.

Number of tasks	Coherent synthesis		
	Cost	**Time**	Power consumption
5	1.75	4.25	**9.56**
10	3	3.6	**35.47**
15	2.75	4.2	**77.69**
20	1.75	12.33	**37.21**
25	2	12.25	**52.24**
30	2.25	14.9	**92.18**
35	2.75	10.4	**173.83**
40	2.75	12.6	**203.57**
45	2.75	14.8	**230.11**
50	2.75	16.3	**242.29**
55	**2.75**	**18**	**268.59**

3.5.3.3 Minimize of power consumption

Tab. 3.7.

Multi-criterions optimization. Minimum of power consumption.

Number of tasks	Coherent synthesis		
	Cost	**Time**	Power consumption
5	1.75	14.75	**6.28**
10	2.5	23	**12.57**
15	2.95	18.5	**20.9**
20	1.75	21	**28.78**
25	2.5	23	**40.46**
30	2.75	24.6	**54.73**
35	3	13.33	**78.3**
40	2.75	15.85	**112.03**
45	2.25	24.67	**95.44**
50	2.25	24.5	**105.99**
55	**2.5**	**25**	**164.58**

3.5.3.4 Compromissing solution

Tab. 3.8.

Multi-criterions optimization. Compromissing solution.

Number of tasks	Coherent synthesis		
	Cost	**Time**	Power consumption
5	1.75	6.75	**9.26**
10	1,5	6.2	**35.47**
15	2.95	15.5	**24.01**
20	1.75	12.83	**35.45**

25	2	14.5	**51.25**
30	2.75	16.9	**63.58**
35	2	18	**78.3**
40	2.75	17.75	**104.68**
45	2.25	21.75	**99.5**
50	2.25	23.88	**113.26**
55	**2.5**	**25**	**164.58**

3.5.3.5 Charts

Chart. 3.8.

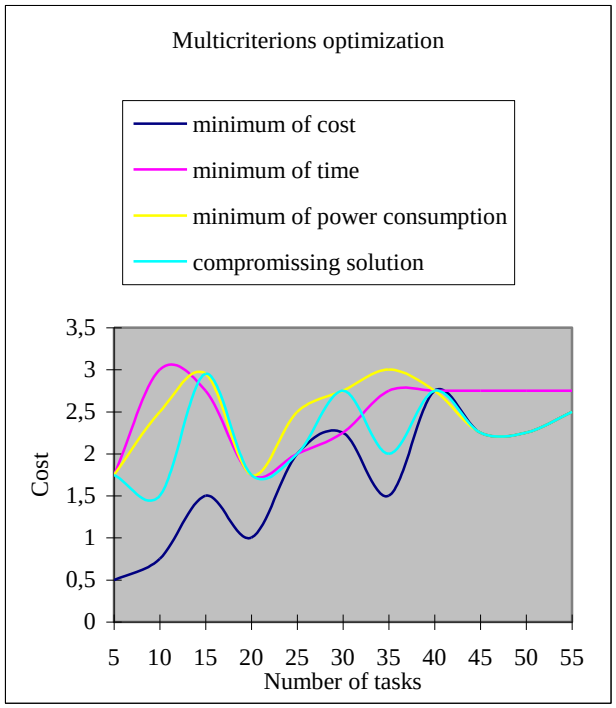


Chart. 3.9.

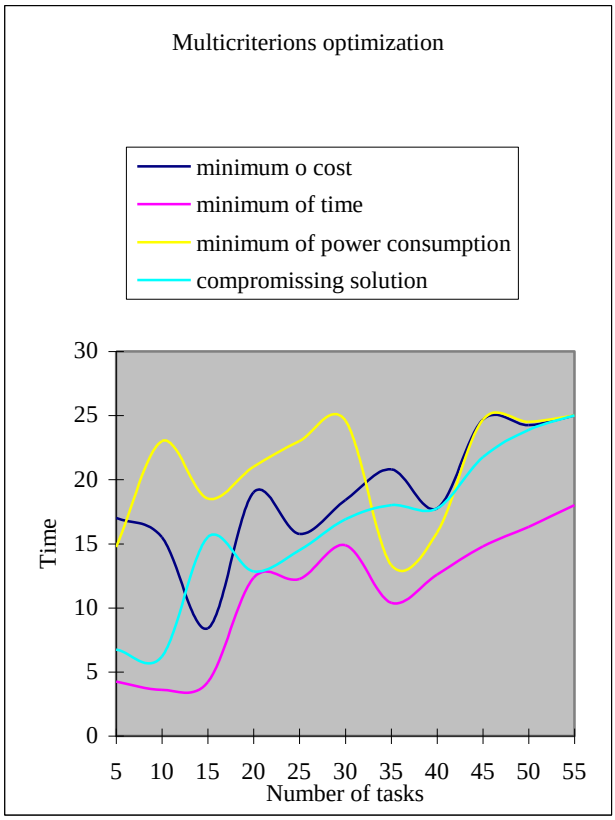
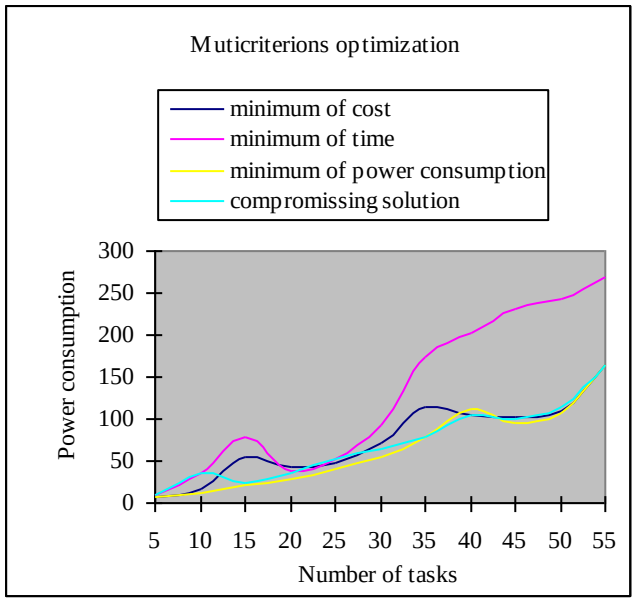


Chart. 3.10.



3.5.4. Multi-criterions optimization. The optimization of time of executing, power consumption and cost. Results for dependent tasks with cost of operating memory.

Tests were conducted for nonpreemptive and dependent tasks. Parameters of constraints: the maximum number of processors - 5, maximum cost - 3, maximum time 25, optimized simultaneously. It the area of optimum solutions in result was received was in sense Pareto. The following tables presented solutions (in Pareto area) for the cost, time and power consumption and solution "compromissing". Searching space of solutions be led to time when global temperature reached value 0.01.

3.5.4.1 Minimize of cost

Tab. 3.9.

Multi-criterions optimization. Minimum of cost.

Number of tasks	Coherent synthesis		
	Cost	**Time**	Power consumption
10	5.5	1.9	**56.07**
20	1.5	18,5	**33.15**
30	5.9	23	**82.41**
40	1.5	50	**81.73**
50	1.25	45.5	**137.2**
60	2.75	22	**299.61**
70	2.5	44.67	**158.2**
80	2.25	47	**179.99**
90	4.25	33.8	**311.6**
100	5.25	46	**291.7**
110	**4.25**	**47**	**431.76**

3.5.4.2 Minimize of time

Tab. 3.10.

Multi-criterions optimization. Minimum of time.

Number of tasks	Coherent synthesis		
	Cost	**Time**	Power consumption
10	6.5	1.9	**43.61**
20	2.75	7.3	**105.25**
30	5.9	23	**82.41**
40	7	20.67	**111.88**
50	4.25	13.4	**205.63**
60	4.25	16.2	**257.48**
70	2.5	32	**175.24**
80	3.25	31	**189.78**
90	4.25	22.2	**393.7**
100	5.75	20.06	**390.77**
110	**5.5**	**20.1**	**558.53**

3.5.4.3 Minimize of power consumption

Tab. 3.11.

Multi-criterions optimization. Minimum of power consumption.

Number of tasks	Coherent synthesis		
	Cost	**Time**	Power consumption
10	6.5	6.67	**23.93**
20	1.5	18,5	**33.15**
30	7.7	45	**54.44**
40	1.5	50	**81.73**
50	1.25	45.5	**137.2**
60	4.25	21.8	**225.56**
70	2.5	32	**175.24**
80	2.25	47	**179.99**
90	4.25	33.8	**311.6**
100	5.25	46	**291.7**
110	**4.3**	**49**	**429.31**

3.5.4.4 Compromissing solution

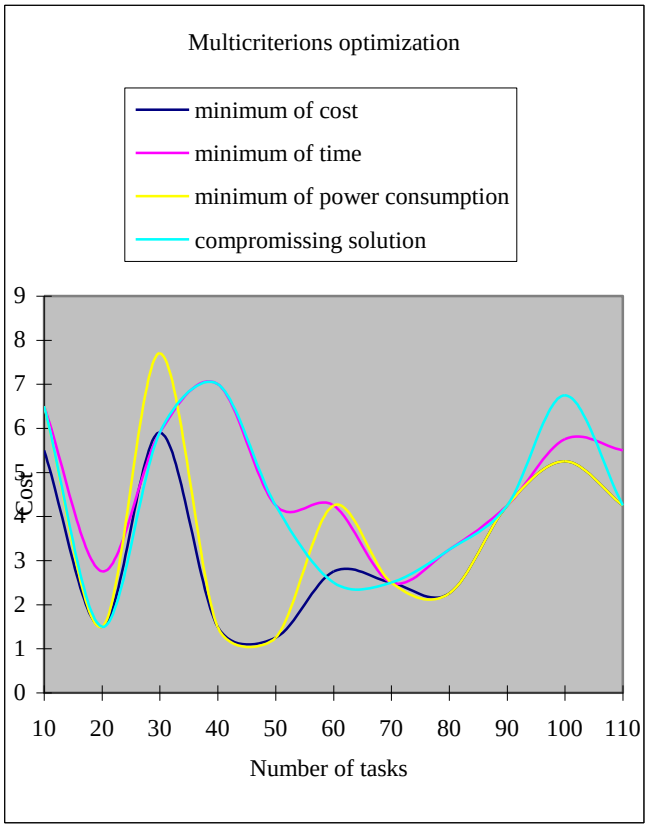
Tab. 3.12.

Multicriterions optimization. Compromisssing solution.

Number of tasks	Coherent synthesis		
	Cost	**Time**	Power consumption
10	6.5	2	**37.99**
20	1.5	18,5	**33.15**
30	5.9	23	**82.41**
40	7	23	**121.56**
50	4.25	16.2	**186.05**
60	2.5	32	**175.24**
70	2.5	38	**167.59**
80	3.25	37	**183.67**
90	4.25	28.6	**328.73**
100	6.75	30.33	**336.36**
110	**4.25**	**41.8**	**435.77**

3.5.4.5 Charts

Chart 3.11.



3.6 Conclusions

This graphs presented of multicriterions optimizationform coherent synthesis of computer system. The designer in result of working of algorithm receives in sense the gathering of optimum solutions Pareto. It stays with the designer's task the selection the most answering his requirements of solution. In dependence from this what are for system requirements it was it been possible to lean on one of got results. To to get to know for given authority of problem the specific of space of solutions well, important the use is long the list of remembering the best solutions (in tests the parameter of algorithm "it quantity *the best" it was* established was value 50). Important the settlement of slow refreshing the algorithm is equally (the parameters "the *step of temperature"* 0.1 and "the *coefficient of cooling*" - in dependent on from quantity of tasks in system; generally smaller than 0,05). We prevent thanks this sale large convergence in population [79], [98]. The algorithm searches near smaller temperature, the larger area in space of solutions. It it was noticed was also that the larger probability of mutation helps the finding the better architecture of system, and the larger probability of crossing improves the optimization of temporary criterion.

Chart. 3.12.

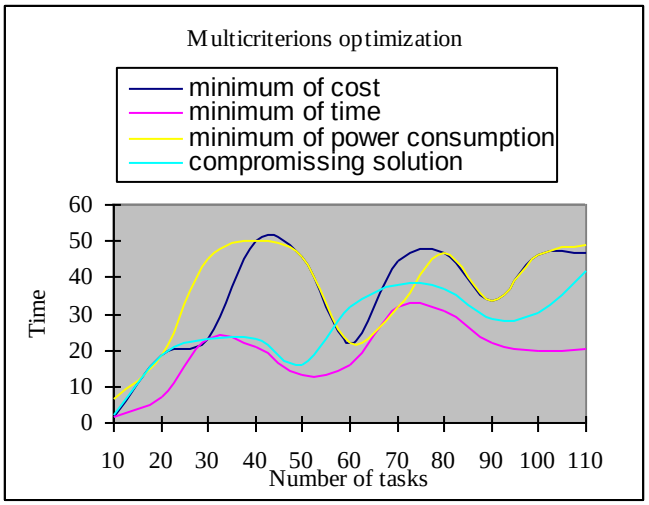
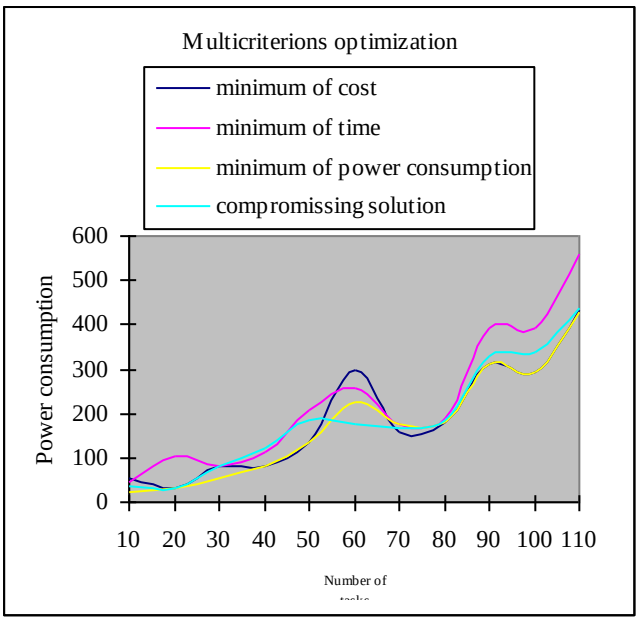


Chart 3.13.



3.7 Example coherent synthesis's applied

3.7.1 Specification of system

We assume, that algorithm of coherent synthesis has to design computer system described on this graph of tasks – Fig. 3.13. We have pools of available resources:

Tab. 3.13.

Available resources

ID	Type	Cost	Cost of memory	Power consumption
P1	universal	1	0.15	0.001
P2	universal	1.5	0.2	0.002
A1	dedicated	0.7	0	0.001
A2	dedicated	0.9	0	0.002

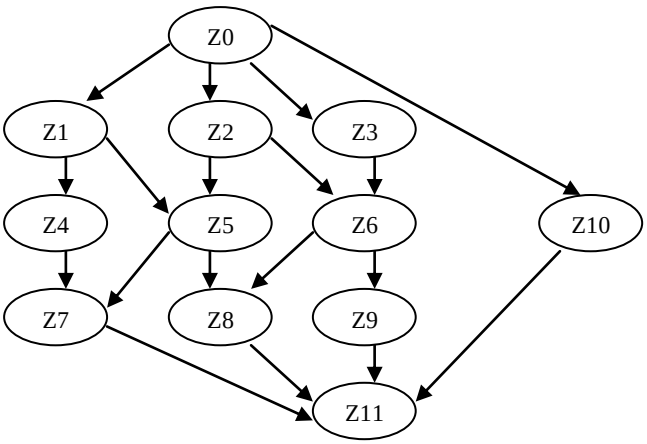


Fig. 3.13. Graph of tasks

In other requirements and constraints of system are following:

Tab. 3.14.

Times of tasks's processing.

If 0 - processor does not be adapted to realization of this task.

Tasks	Processors			
	P1	P2	A1	A2
Z0	2	1	3	2,5
Z1	4	2	1,5	0
Z2	2	1	3.5	3
Z3	4	2	0	0
Z4	5	2,5	0	0
Z5	6	3	2	1
Z6	3	1,5	0	0
Z7	2	1	0	0
Z8	1	0,5	0	0
Z9	4	2	0	1
Z10	2	1	0,5	0,4
Z11	2	1	0,5	0,4

Tab. 3.15.
Power consumption of tasks's processing.
If 0 - processor does not be adapted to realization of this task.

Tasks	Processors			
	P1	P2	A1	A2
Z0	1	1,5	1,5	1,5
Z1	2	3	1	0
Z2	1	1,5	2	2
Z3	2	3	0	0
Z4	2,5	3,5	0	0
Z5	3	4,5	2	1,5
Z6	1,5	3	0	0
Z7	1	2	0	0
Z8	0,5	1	0	0
Z9	2	3	0	1,5
Z10	1	2	0,7	0,6
Z11	1	2	0,7	0,6

Processor P2 is quicker from processor P1 but processor P2 takes more power and is dearer than P1. Dedicated processor A1 can realize tasks Z0, Z1, Z2, Z5, Z10 and Z11 and he is adequate for tasks: Z5, Z10 and Z11. Task Z0 and Z2 can be executed on this processor but time of their realization is longer than on universal processors. This processor is cheaper from dedicated processor A2, but the power consumption has greater. Dedicated processor A2 is adequate for execution of tasks Z5, Z9, Z10 and Z11. This processor is insufficient for tasks Z0 and Z1 than universal processors, but more suitable than processor specialized A1. The cost of processor A2 is greater than processor A1.

3.7.2 Results of optimization

Simplifying of analysis we assume that four is the maximum number of processors.

3.7.2.1 *Minimize of cost (without of cost of memory)*

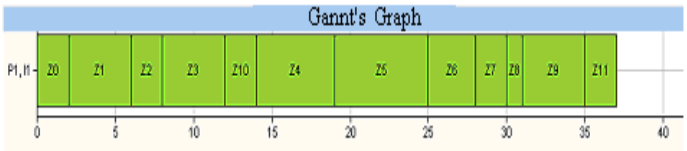


Fig. 3.14. Minimize of cost without of cost of memory

Tab. 3.16.
Minimize of cost without of cost of operating memory

Cost	Time	Power consumption
1	37	18,5

We for obvious reasons in this example have consisting system from one and the cheapest universal processor. Dedicated processors, which are cheaper possibilities of realization of all tasks have not in this system.

3.7.2.2 Minimize of cost (with of cost of memory)

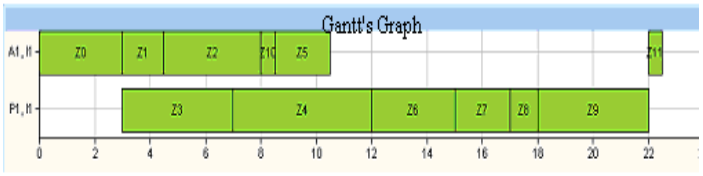


Fig. 3.15. Minimize of cost (with of cost of memory)

Tab. 3.17.
Minimize of cost (with of cost of operating memory)

Cost	Time	Power consumption
2,6	23,5	17,4

We for this example will receive system folded from one universal processor (cheaper) as well as one dedicated processor (cheaper). Here two processors more suitable is apply. Cost of realization of all tasks on the cheapest universal processor is equal 2,8 (the cost of processor and 12 the tasks the by 0,15 cost of unit of memory).

3.7.2.3 Minimize of processing time

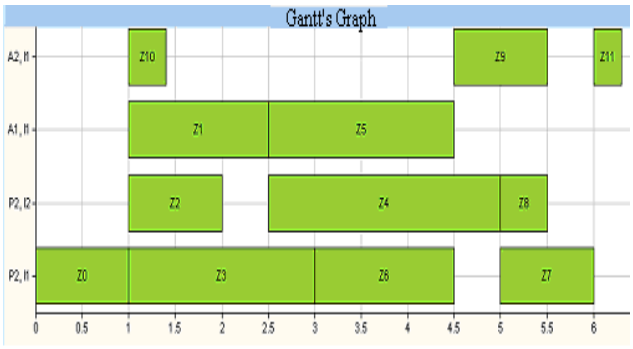


Fig. 3.16. Minimize of processing time

Tab. 3.18.
Minimize of processing time

Cost	Time	Power consumption
6	6,3	21,2

For example in which in have the minimization of time is generated system, which are two quicker universal processors and two dedicated processors A1 and A2. Algorithm counted schedule in which processors A1 and A2 executed tasks for which they are dedicated.

3.7.2.4 Minimize of power consumption

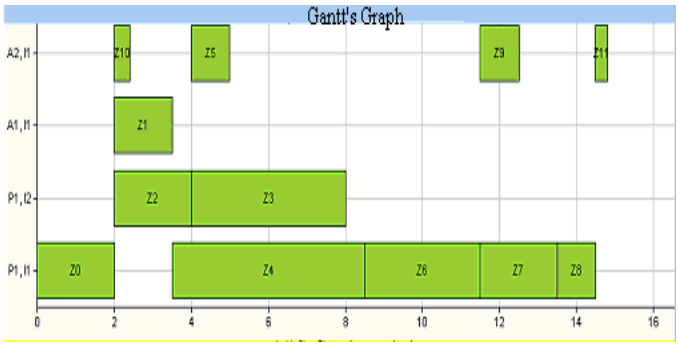


Fig. 3.17. Minimize of power consumption

Tab. 3.19.

Minimize of power consumption

Cost	Time	Power consumption
4,65	14,8	14,7

For this example algorithm generated system folded from two cheaper dedicated processors for which the power consumption is the lowest for almost all tasks. Task Z1 was executed on processor A1 because executing this task processor A1 is dedicated. Task Z10, Z5, Z9 and Z11 are executed on dedicated processor A2 on which is smallest the power consumption.

3.7.2.5 Minimize of time and cost (for given maximum of cost)

Maximum cost was given on value 4.

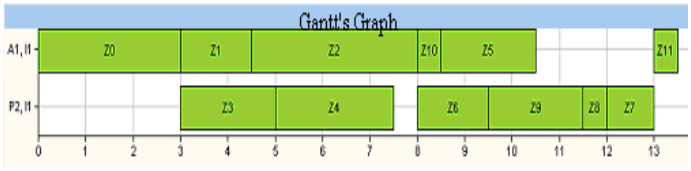
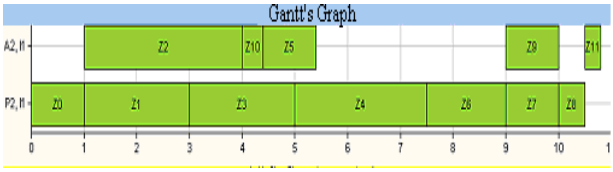


Fig. 3.18. Minimize of time and cost. Cheaper system.

Tab. 3.20.

Minimize of time and cost. Cheaper system.

Cost	Time	Power consumption
3,4	13,5	23.4



3.19. Minimize of time and cost. Quicker system.

Tab. 3.21.

Minimize of time and cost. Quicker system.

Cost	Time	Power consumption
3,8	10,8	23,2

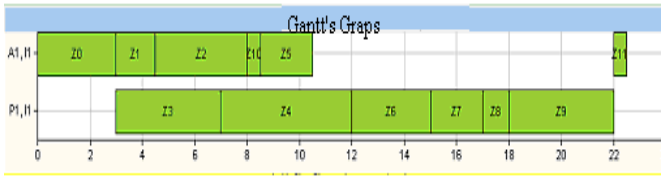


Fig. 3.20. Minimize of cost and time. The cheapest system. Improvement of time.

Tab. 3.22.

Minimize of cost and time. The cheapest system. Improvement of time.

Cost	Time	Power consumption
2,6	22,5	17,4

Algorithms for multicriterions optimization receive the area of optimum solutions in sense Pareto. First and second solution have the costs greater than settled the maximum as well as have larger time from minimum time of system. However they find among these conflicting requirements compromise. First of solutions is cheaper since second about 0,4. The time of realization of tasks is longer about 2,7. The power consumption is lower for dearer system. The solution for last case is about the smallest of cost and other criterions are simultaneously estimated. Solution this has the same cost how the cheapest system (solution for minimize of cost) and schedule length about 1 reduces simultaneously.

3.7.2.6 Minimize of cost and power consumption (for given maximum of time)

The maximum of time was given on value 20.

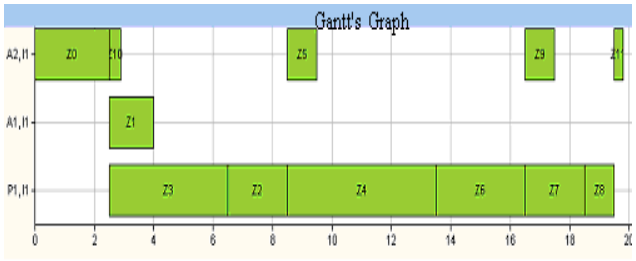


Fig. 3.21. Minimize of cost and power consumption. Maximum time equal 20. Cheaper system.

Tab. 3.23.

Minimize of cost and power consumption. Maximum time equal 20.
Cheaper system.

Cost	Time	Power consumption
3,5	19,8	15,2

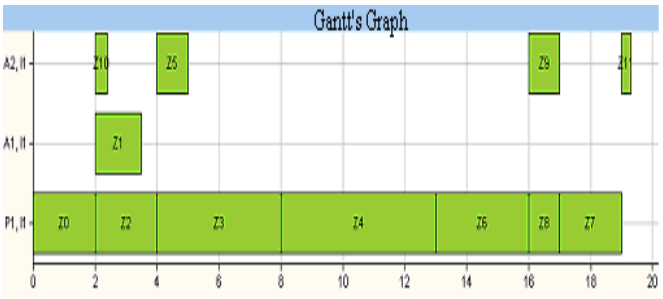


Fig. 3.22. Minimize of cost and power consumption. Maximum time equal 20. Dearer system.

Tab. 3.24.

Minimize of cost and power consumption. Maximum time equal 20.
Dearer system.

Cost	Time	Power consumption
3.65	19.3	14.7

Criterion of time was the requirement for resultant of system: he had to finish the executing the tasks before 20 units of time. Other criterions have be optimized: i.e. time and cost. Algorithm generated two optimum solutions in sense Pareto. System for first solution executed tasks to 19,8 units of time, has the larger power consumption and his cost is smaller. Systems these differ of the execution of task Z0, only. In first system this task is executed through dedicated processor, which has larger power consumption and he need not additional operating memory. Execution of this task through universal processor P1 has smaller of the power consumption and he need to execution of task the additional memory. The second solution generated system for which time executing all tasks is shorter.

3.7.2.7 Multicriterions optimalization. Coherent minimize of cost, time and power consumption (for given maximum of time).

Maximum of time was given on value 15.

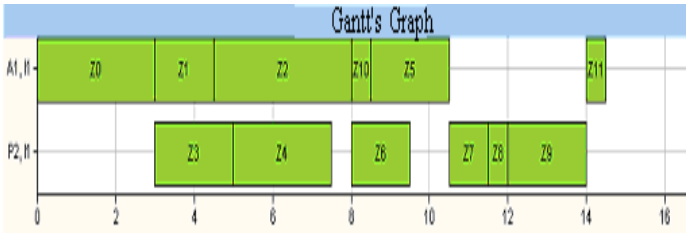


Fig. 3.23. Minimize of cost, time and power consumption. Cheaper system.

Tab. 3.25.
Minimize of cost, time and power consumption. Cheaper system.

Cost	Time	Power consumption
3,4	14,5	23,4

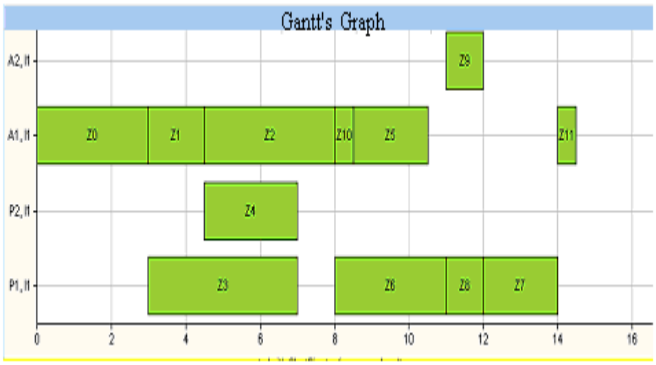


Fig. 3.24. Minimize of cost, time and power consumption. Cheaper system. System with optimized power consumption.

Tab. 3.25
Minimize of cost, time and power consumption. Cheaper system.
System with optimized power consumption.

Cost	Time	Power consumption
4,9	14,5	17,9

The algorithm generated area of optimum solutions in sense Pareto for three criterions. Generated solutions get the different compromise among conflicting criterions. Two systems to analysis were chosen. Finish execution time for all of tasks is 14,5 unit. First system consists with 2 processors and his cost is small. However executing of tasks on processor P2 causes the power consumption considerable [73]. The dedicated processor executes the tasks with smaller power consumption than processor A2. Task Z2 is except. Realization of this task on processor P2 would cause the crossing the limit of time. Second system consists with 4 processors. This system is dearer (150%) however the power consumption reduces significantly and the limit of time of realization tasks fulfils simultaneously.

4. Résumé

The paper describes genetic algorithms and their implementation flowcharts. Moreover, it presents selected results of analytical experiments for resource selection and task scheduling. The paper explores the coherent synthesis algorithm of computer systems, in which resource selection and task scheduling optimization processes are realized concurrently and coherently. The coherent approach in the synthesis generates common and interdependent solutions regarding the system structure (type and configuration of the selected resources), as well as the scheduling of tasks ran on those resources. In the presented approach, the cost of resources (system cost), the time of completing all tasks (system speed) and the power consumption of the system are optimized. The coherent algorithm yields much (up to 40%) better solutions, which is proved by analytical experiments.

References

[1] Aarts E.H.L. Korst J., "Simulated Annealing and Boltzmann Machines", J. Wiley, Chichester, 1989.
[2] D'Ambrioso J., Hu X., "Configuration Level Hardware/Software Partitioning for Real-Time Systems", Proc. of the Int. Workshop on Hardware/Software Codesign, Vol. 14, 34-41, 1994.
[3] Axelson J., "Architecture Synthesis and Partitionig of Real-Time Systems: A Camparison of Three Heuristic Search Strategies", Proc. of the Int. Workshop on Hardware/Software Codesign, 161-165, 1997.
[4] Błażewicz J., Ecker K., Pesch E., Schmidt G., Węglarz J., "Handbook on Scheduling, From Theory to Applications", Springer-Verlag Berlin Heidelberg, 2007.
[5] Błażewicz J., Ecker K., Plateau B., Trystram D., "Handbook on Parallel and Distributed Processing", Spinger-Verlag, Berlin – Heidelberg, 2000.
[6] Brucker P., Knust S., "Complex Scheduling", Springer, 2006.
[7] Coffman E. G., Jr., "Computer and Job-shop scheduling theory", John Wiley&Sons, Inc. New York, 1976.
[8] Dick R. P., Jha N. K., MOGAC: "A Multiobjective Genetic Algorithm for the Cosynthesis of Hardware-Software Embedded Systems", Proc. of the Int. Conference on Computer Aided Design, 522-529, 1997.
[9] Drabowski M., M., Rola M, Roślicki A., "Algorytmy neuronowe w sterowaniu rozdziałem zadań i zasobów w kompleksie operacji", 2rd International Congress of Intelligent Building Systems, INBus2002, Kraków, 45-52, 2002.
[10] Drabowski M., Czajkowski K., "Task scheduling in coherent, co-synthesis of computer system", Advanced Computer Systems – Computer Information Systems and Industrial Management Application (ACS-CISIM 2005), in. Image Analysis, Computer Graphics, Security Systems and Artificial Intelligence Applications, vol. 1, 53-60, 2005.
[11] Drabowski M., "Par-synthesis of multiprocessors parallel systems", International Journal of Computer Science and Network Security, Vol. 8, No. 10, 90-96, 2008.
[12] Eiben A.E. Aarts E.h.l. van Hee K.H., "Global convergence of genetic algorithms: A Markov chain analysis", LNCS 496, 4-9, 1991.
[13] Gajski D.D., Dutt N.D., Wu A.C.H., Lin S.Y., "High-level Synthesis. Introduction to Chip and System Design", Kluwer Academic Pub., Boston, MA, 1994.
[14] Gajski D., "Principles of Digital Design", Prentice Hall, Upper Saddle River, NJ, 1997.
[15] Garey M. R., Johnson D. S., "Computers and intractability: A guide to the theory of NP-completeness", San Francisco, Freeman, 1979.
[16] Gupta R.K., De Micheli G., "Hardware-Software Co-synthesis for Digital Systems", IEEE Design&Test of Computers, Vol. 10, No. 3, 29-41, 1993.
[17] Harel D., *Statecharts*: "A Visual Formalism for Complex Systems", Science of Computer Programming, Vol. 8, No. 3, 231-274, 1987.
[18] Henkel J., Ernst R., "High-Level Estimation Techniques for Usage in Hardware/Software Co-Design", Proc. of the Asia and South Pacific Automation Conference, 353-360, 1998.
[19] de Micheli G., "Computer-Aided hardware-Software Codesign", IEEE Micro, Vol. 14, No. 4, pp. 10-24, 1994.
[20] Schulz S., Rozenbilt J.W., Mrva M., Buchenrieder K., "Model-Based Codesign", IEEE Computer, Vol. 31, No. 8, 60-67, 1998.
[21] Sgroi M., Lavagno L., Sangiovanni-Vincentelli A., "Formal Models for Embedded System Design", IEEE Design&Test of Computers, Vol. 17, No. 2, 14-27, 2000.
[22] Steinhausen U., "System Synthesis Using Hardware/Software Codesign", Proc. of the Int. Workshop on Hardware-Software Co-Design, 1993.
[23] Teich J., Blickle T., Thiele L., "An Evolutionary Approach to System-Level Synthesis", Proc. of the Int. Workshop on Hardware/Software Codesign, 167-172, 1997.
[24] Węglarz J., "Project Scheduling – Recent Models, Algorithms and Applications", Kluwer Academic Publ., 1999.
[25] Węglarz J., "Sterowanie w systemach typu kompleks operacji", PWN, PAN, Oddział w Poznaniu, 1981.

Author Biographies

Mieczyslaw Drabowski, Assistant Professor of Department of Computer Engineering, Faculty of Electrical and Computer Engineering, Cracow University of Technology, received the M. Sc. degree in automatic control and communication from AGH University of Science and Technology, graduated mathematic from Jagiellonian University in Krakow and received the Ph. D. degree (with honors) in computing science from Poznan University of Technology, in 1977, 1979 and 1986, respectively.
Currently he is member of several editorial boards, among others International Association for Development of the Information Society (IADIS), International Association of Science and Technology for Development (IASTED) on Artificial Intelligence and Soft Computing, International Journal of Computer Science and Emerging Technologies.
His research interests include schedule, assignment and allocation for tasks and resources, dependable and fault tolerant systems, artificial intelligence, operating systems and software engineering, author and co-author of 3 monographs and over 80 papers in major professional journals and conference proceedings.
Dr. Drabowski is a member of the council of the Polish Information Processing Society.

Clonal Selection Algorithm for DG Sources Allocation for Minimum Loss in Distribution System

M.Padma Lalitha[1],V.C. Veera Reddy[2], N.Sivarami Reddy[3],K.Sivagangadhar Reddy[4]

[1] IEEE Member, HOD, Department of Electrical & Electronics Engineering,
Annamacharya Institute of Techonology & Sciences, Rajampet, A.P. Tel# +9848998643,
e-mail:padmalalitha_mareddy@yahoo.co.in
[2] Electrical & Electronics Engineering Dept , S.V.U.C.E,S.V.University,Tirupathi, A.P,
e-mail:veerareddy_vc@yahoo.com
[3] HOD, Department of Mechanical Engineering,
Annamacharya Institute of Technology & Sciences, Rajampet, A.P. Tel# +9848998645,
e-mail:siva.narapureddy@yahoo.com
[4] Assistant Professor,Department of Electrical & Electronics Engineering,
Annamacharya Institute of Technology & Sciences, Rajampet, A.P.

Abstract: Distributed Generation (DG) is a promising solution to many power system problems such as voltage regulation, power loss, etc. This paper presents a new methodology using Fuzzy and Artificial Immune System(AIS) for the placement of Distributed Generators(DG) in the radial distribution systems to reduce the real power losses and to improve the voltage profile. A two-stage methodology is used for the optimal DG placement. In the first stage, an analytical method is used to find the optimal DG locations and in the second stage, Clonal Selection algorithm of AIS is used to find the size of the DGs corresponding to maximum loss reduction. This algorithm is a new population based optimization method inspired by cloning principle of human body immune system. The advantage of this algorithm is that the population size is dynamic and it is determine by the fitness values of the population. The proposed method is tested on standard IEEE 33 bus test system and the results are presented and compared with different approaches available in the literature. The proposed method has outperformed the other methods in terms of the quality of solution and computational efficiency.

Keywords: *DG placement, Meta heuristic methods, Artificial Immune Systems, Clonal Selection algorithm, loss reduction, radial distribution system.*

1. Introduction

Distributed or dispersed generation (DG) or embedded generation (EG) is small-scale power generation that is usually connected to or embedded in the distribution system. The term DG also implies the use of any modular technology that is sited throughout a utility's service area (interconnected to the distribution or sub-transmission system) to lower the cost of service [1]. The benefits of DG are numerous [2, 3] and the reasons for implementing DGs are an energy efficiency or rational use of energy, deregulation or competition policy, diversification of energy sources, availability of modular generating plant, ease of finding sites for smaller generators, shorter construction times and lower capital costs of smaller plants and proximity of the generation plant to heavy loads, which reduces transmission costs. Also it is accepted by many countries that the reduction in gaseous emissions (mainly CO_2) offered by DGs is major legal driver for DG implementation [4].

The distribution planning problem is to identify a combination of expansion projects that satisfy load growth constraints without violating any system constraints such as equipment overloading [5]. Distribution network planning is to identify the least cost network investment that satisfies load growth requirements without violating any system and operational constraints. Due to their high efficiency, small size, low investment cost, modularity and ability to exploit renewable energy sources, are increasingly becoming an attractive alternative to network reinforcement and expansion. Numerous studies used different approaches to evaluate the benefits from DGs to a network in the form of loss reduction, loading level reduction [6-8].

Naresh Acharya *et al* suggested a heuristic method in [9] to select appropriate location and to calculate DG size for minimum real power losses. Though the method is effective in selecting location, it requires more computational efforts. The optimal value of DG for minimum system losses is calculated at each bus. Placing the calculated DG size for the buses one by one, corresponding system losses are calculated and compared to decide the appropriate location. More over the heuristic search requires exhaustive search for all possible locations which may not be applicable to more than one DG. This method is used to calculate DG size based on approximate loss formula may lead to an inappropriate solution.

In the literature, genetic algorithm and PSO have been applied to DG placement [10-13].In all these works either sizing or location of DGs are determined by these methods. This paper presents a new methodology using Clonal selection algorithm[14-17] for the placement of DG in the radial distribution systems. The Clonal algorithm is a new population based meta heuristic approach inspired by Clonal principle of immune system of human body. The advantage of this algorithm is that it does not require external parameters such as selection, cross over rate and mutation rate as in case of genetic algorithm and differential evolution and it is hard to determine these parameters in prior. The other advantage is that the global search ability in the algorithm is implemented by introducing hyper mutation which differs from mutation in GA in two ways. One is the mutation rate is very high that every solution is mutated here and the second one is the mutation is not a

single bit mutation. The advantage of Clonal algorithm is its dynamic population size.

In this paper, the optimal locations of distributed generators are identified based on single DG placement method[18] and Clonal optimization technique which takes the number and location of DGs as input has been developed to determine the optimal size(s) of DG to minimize real power losses in distribution systems. The advantages of relieving Clonal method from determination of locations of DGs are improved convergence characteristics and less computation time. Voltage and thermal constraints are considered. The effectiveness of the proposed algorithm was validated using 33-Bus Distribution System [19].To test the effectiveness of proposed method; results are compared with different approaches available in the literature. The proposed method has outperformed the other methods in terms of the quality of solution and computational efficiency.

2. Theoretical Background

The total I^2R loss (P_L) in a distribution system having *n* number of branches is given by:

$$P_{Lt} = \sum_{i=1}^{n} I_i^2 Ri \qquad (1)$$

Here I_i is the magnitude of the branch current and R_i is the resistance of the i^{th} branch respectively. The branch current can be obtained from the load flow solution. The branch current has two components, active component (I_a) and reactive component (I_r). The loss associated with the active and reactive components of branch currents can be written as:

$$P_{La} = \sum_{i=1}^{n} I_{ai}^2 Ri \qquad (2)$$

$$P_{Lr} = \sum_{i=1}^{n} I_{ri}^2 Ri \qquad (3)$$

Note that for a given configuration of a single-source radial network, the loss P_{La} associated with the active component of branch currents cannot be minimized because all active power must be supplied by the source at the root bus. However by placing DGs, the active component of branch currents are compensated and losses due to active component of branch current is reduced. This paper presents a method that minimizes the loss due to the active component of the branch current by optimally placing the DGs and thereby reduces the total loss in the distribution system. A two stage methodology is applied here. In the first stage optimum location of the DGs are determined by using single DG placement method and in the second stage Clonal selection algorithm is used to determine sizes of the DGs for maximum real loss reduction.

3. Identification Of Optimal DG Locations By Single DG Placement Algorithm

This algorithm determines the optimal size and location of DG units that should be placed in the system to minimize loss. First optimum sizes of DG units for all nodes are determined for base case and best one is chosen based on the maximum loss saving. This process is repeated if multiple DG locations are required by modifying the base system by inserting a DG unit into the system one-by-one.

3.1. Methodology

Assume a single-source radial distribution system with *n* branches and DG is to be placed at bus ***m*** and α be set of branches connected between source and bus ***m***. The DG produces active current I_{DG}, and for radial network it changes only active component of branch current of set α. The currents of other branches are unaffected. Thus new active current I_{ai}^{new} of i^{th} branch is given by

$$I_{ai}^{new} = I_{ai} + D_i I_{DG} \qquad (4)$$

where $D_i = 1$; if branch $i \in \alpha$
$= 0$; otherwise

The loss P_{La}^{com} associated with active component of branch currents in new system is given by

$$P_{La}^{com} = \sum_{i=1}^{n} (I_{ai} + D_i I_{DG})^2 R_i \qquad (5)$$

The saving S is difference between equation 3 and 5 and is given by

$$S = P_{La} - P_{La}^{com}$$
$$= -\sum_{i=1}^{n} (2D_i I_{ai} I_{DG} + D_{DG} I_c^2) R_i \qquad (6)$$

The DG current I_{DG} that provides maximum saving is obtained from

$$\frac{\partial S}{\partial I_{DG}} = -2\sum_{i=1}^{n} (D_i I_{ai} + D_i I_{DG}) R_i = 0 \qquad (7)$$

The DG current for maximum saving is

$$I_{DG} = -\frac{\sum_{i=1}^{n} D_i I_{ai} R_i}{\sum_{i=1}^{n} D_i R_i} = -\frac{\sum_{i\in\alpha} I_{ai} R_i}{\sum_{i\in\alpha} R_i} \qquad (8)$$

The corresponding DG size is

$$P_{DG} = V_m I_{DG} \qquad (9)$$

V_m is voltage magnitude of bus-*m*. The optimum size of DG at each bus is determined using eqn (9). Then saving for each DG is determined using eqn (6).The DG with highest saving is candidate location for single DG placement. When the candidate bus is identified and DG is placed, the process is repeated to identify subsequent buses for DG placement.

3.2. Algorithm for Single DG Placement

Step 1: Conduct load flow analysis for the original system .

Step 2: Calculate I_{DG} and DG size using equations 8 & 9 for buses *i*=2…*n*.

Step 3: Determine saving using equation 6, for buses *i*=2…*n*.

Step 4: Identify the maximum saving and the corresponding DG size.

Step5: The corresponding bus is candidate bus where DG can be placed. Modify the active load at this bus and conduct the load flow again.

Step 6: Check whether the saving obtain is more than 1kW. If yes, go to step 2. Otherwise, go to next step.

Step 7: print all the candidate locations to place DG sources and the sizes.

Since the DGs are added to the system one by one, the sizes obtained by single DG placement algorithm are local optima not global optimum solution. The global optimal solution is obtained if multiple DGs are simultaneously placed in the system by using ABC algorithm as explained in next section.

4. Identification Of Optimal DG Sizes By Clonal Algorithm

4.1. Introduction to Artificial Immune System

The 'artificial immune system' is an approach which used the natural immune system as a metaphor for solving computational problems, *not* modeling the immune system [14]. The main application domains of AIS are anomaly detection, pattern recognition, computer security, fault tolerance, dynamic environments, robotics, data mining, optimization, and scheduling. The 'immune system' (IS) can be considered to be a remarkably efficient and powerful information processing system which operates in a highly parallel and distributed manner. It contains a number of features which potentially can be adapted in computer systems; recognition, feature extraction, diversity, learning, memory, distributed detection, self-regulation, threshold mechanism, co-stimulation, dynamic protection, and probabilistic detection. It is unnecessary to replicate *all* of these aspects of the IS in a computer model, rather they should be used as general guidelines in designing a system.

There are a number of different algorithms that can be applied to many domains, from data analysis to autonomous navigation [14]. These immune algorithms were inspired by works on theoretical immunology and several processes that occur within the IS. The AISs lead to the development of different techniques, each one mapping a different mechanism of the system. For examples, the *Artificial Immune Networks* as proposed by Farmer et al. [15], the *Clonal Selection Algorithm* proposed by de Castro and Von Zuben [16], and the *Negative Selection Algorithm* introduced by Forrest et al. [17]. Immune network models are suitable to deal with dynamic environments and optimization problems, algorithms based upon the clonal selection principle are adequate to solve optimization and scheduling problems, and the negative selection strategies are successfully applied to anomaly detection.

4.2. Application of Clonal Selection Algorithm to determine DG unit sizes

The clonal selection algorithm (CSA) is inspired by the immunological processes of *clonal selection* and *affinity maturation*. When an antigen is detected, those antibodies that best recognize this antigen will proliferate by cloning. This process is called *clonal selection principle* [6]. The clonal selection principle is used to explain how the IS 'fights' against an antigen. When a bacterium invades our organism, it starts multiplying and damaging our cells. One form the IS found to cope with this replicating antigen was by replicating the immune cells successful in recognizing and fighting against this disease-causing element. Those cells reproduce themselves asexually in a way proportional to their degree of recognition: the better the antigenic recognition, the higher the number of clones (offspring) generated. During the process of cell division (reproduction), individual cells suffer a mutation that allows them to become more adapted to the antigen recognized: the higher the affinity of the parent cell, the lower the mutation they suffer. The algorithm is given below.

Step 1. ***Initialization*****:** initialize a population of antibodies (feasible sizes of DG unit at predetermined locations). Each antibody represents a solution in the search space.

Step 2. ***Selection:*** All the antibodies are selected in optimization version

Step 3. ***Affinity Evaluation*****:** determine the affinity of selected antibodies (affinity = 1/Power Loss)

Step 4. ***Cloning or proliferation*****:** The selected antibodies will be cloned (reproduced) independently and proportionally to their affinities, generating a repertoire of clones: the higher the affinity, the higher the number of clones generated for each of the selected antibodies;

Step 5. ***Hyper-mutation* :** The repertoire of clones is submitted to an affinity maturation process inversely proportional to the affinity, generating a population of matured clones: the higher the affinity, the smaller the mutation rate;

Step 6. ***Affinity evaluation:*** determine the affinity of matured clones

Step 7. ***Reselection :*** From this set of mature clones re-select the higher affinity clones

Step 8. Finally, replace the *d* lowest affinity antibodies from the population of the antibodies by new individuals generated randomly to maintain the diversity in population.

The main operators in CSA are *cloning*, *mutation and* r*eselection*. All solutions are selected for cloning and a number of clones are generated for each solution. Almost all clones will be mutated to produce new feasible solutions for the next generation since '1- selection probability' would give a high mutation rate for each clone. But only new solutions with high affinity will be selected to replace the low affinity solutions in the current population. The process will be repeated until the stopping criteria are met.

In our implementation, it was assumed that the *n* highest affinity antibodies were sorted in ascending order after Step 3, so that the amount of clones generated for all these *n* selected antibodies was given by (5):

$$N_c = \sum_{i=1}^{n} round\left(\frac{\beta . N}{i}\right), \quad (5)$$

where *N*c is the total amount of clones generated, β is a multiplying factor, *N* is the total amount of antibodies and *round*( ) is the operator that rounds its argument towards the closest integer. Each term of this sum corresponds to the clone size of each selected antibody, e.g., for $N = 100$ and $\beta = 1$, the highest affinity antibody ($i = 1$) will produce 100 clones, while the second highest affinity antibody produces 50 clones, and so on.

5. RESULTS AND DISCUSSION

First load flow is conducted for IEEE 33 bus test system[7]. The power loss due to active component of current is 136.9836 kW and power loss due to reactive component of the current is 66.9252 kW. A program is written in "MATLAB" to implement single DG placement algorithm . For the first iteration the maximum saving is occurring at bus 6. The candidate location for DG is bus 6 with a loss saving of 92.1751 kW. The optimum size of DG at bus 6 is 2.4886 MW. By assuming 2.4886 MW DG is connected at bus 6 of base system and is considered as base case. Now the candidate location is bus 15with 0.4406 MW size and the loss saving is 11.4385 KW. This process is repeated till the loss saving is insignificant. The results are shown in Table I.

The candidate locations for DG placement are taken from single DG placement algorithm i.e. 6,15,25,32. With these locations, sizes of DGs are determined by using Clonal Algorithm described in section 4. The sizes of DGs are dependent on the number of DG locations.

TABLE I.
SINGLE DG PLACEMENT RESULTS

iteration No.	Bus No.	DG Size (MW)	Saving (KW)
1	6	2.4886	92.1751
2	15	0.4406	11.4385
3	25	0.6473	7.6936
4	32	0.4345	8.1415

Generally it is not possible to install many DGs in a given radial system. Here 4 cases are considered. In case I only one DG installation is assumed. In case II two DGs, in case III three DGS and in the last case four DGs are assumed to be installed. DG sizes in the four optimal locations, total real power losses before and after DG installation for four cases are given in Table II.

TABLE II.
RESULTS OF IEEE 33 BUS SYSTEM

Case	bus locations	DG sizes (MW)	Total Size (MW)	losses before DG installation (kW)	loss after DG installation (kW)	Saving (kW)	saving/ DG size
I	6	2.5775	2.5775		105.0231	98.8857	39.9
	6	1.9707					
II	15	0.5757	2.5464		89.9619	113.9469	44.75
	6	1.7569					
	15	0.5757					
III	25	0.7826	3.1152		79.2526	124.6562	40.015
	6	1.0765					
	15	0.5757					
	25	0.7824					
IV	32	0.6538	3.0884	203.9088	66.5892	137.3196	44.86

The last column in Table II represents the saving in kW for 1 MW DG installation. The case with greater ratio is desirable. As the number of DGs installed is increasing the saving is also increasing. In case4 maximum saving is achieved but the number of DGs is four. Though the ratio of saving to DG size is maximum of all cases which represent optimum solution but the number of DGs involved is four so it is not economical by considering the cost of installation of 4 DGs. But in view of reliability, quality and future expansion of the system it is the best solution.

Table III shows the minimum voltage and % improvement in minimum voltage compared to base case for all the four cases. In all the cases voltage profile is improved and the improvement is significant. The voltage profile for all cases is shown in Figure 1.

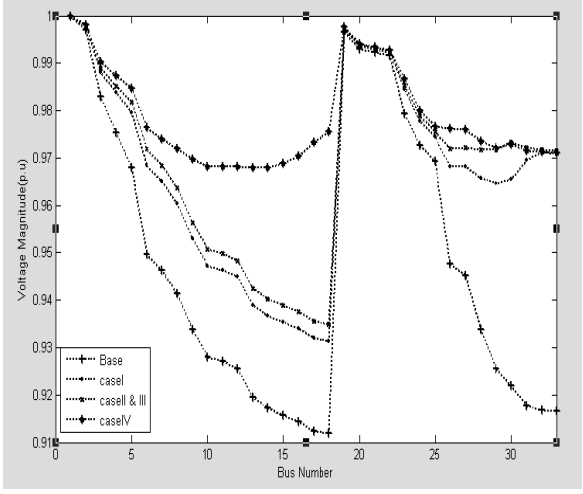


Figure 1. Voltage profile with and without DG placement for all Cases

TABLE III.
VOLTAGE IMPROVEMENT WITH DG PLACEMENT

case No.	Bus No.	Min Voltage	% change
Base case	18	0.9118	
case1	18	0.9314	2.149
case2	18	0.9349	2.533
case3	18	0.9349	2.533
case4	14	0.9679	6.153

Table IV shows % improvements in power loss due to active component of branch current, reactive component of branch current and total active power loss of the system in the four cases considered. The loss due to active component of branch current is reduced by more than 68% in least and nearly 96% at best. Though the aim is reducing the P_{La} loss, the P_{Lr} loss is also reducing due to improvement in voltage profile. From Table 5 it is observed that the total real power loss is reduced by 48.5% in case 1 and 67% in case 4.

TABLE IV.
LOSS REDUCTION BY DG PLACEMENT

case No.	P_{La} **(kW)**	**% Saving**	P_{Lr} **(kW)**	**% Saving**	P_{Lt} **(kW)**	**% Saving**
Base case	136.9836	----	66.9252	----	203.9088	----
case1	43.2029	68.46	61.9297	7.46	105.1326	48.44
case2	28.5651	79.15	61.4845	8.13	90.0496	55.84
case3	18.1166	86.77	61.2138	8.53	79.3303	61.1
case4	5.6123	95.9	61.0493	8.78	66.6616	67.31

The convergence characteristics of the solution obtained by Clonal selection algorithm for all the four cases are shown in figure 2.

Table V shows the minimum, average and maximum values of total real power loss from 100 trials of Clonal selection algorithm. The average number of iterations and average CPU time are also shown.

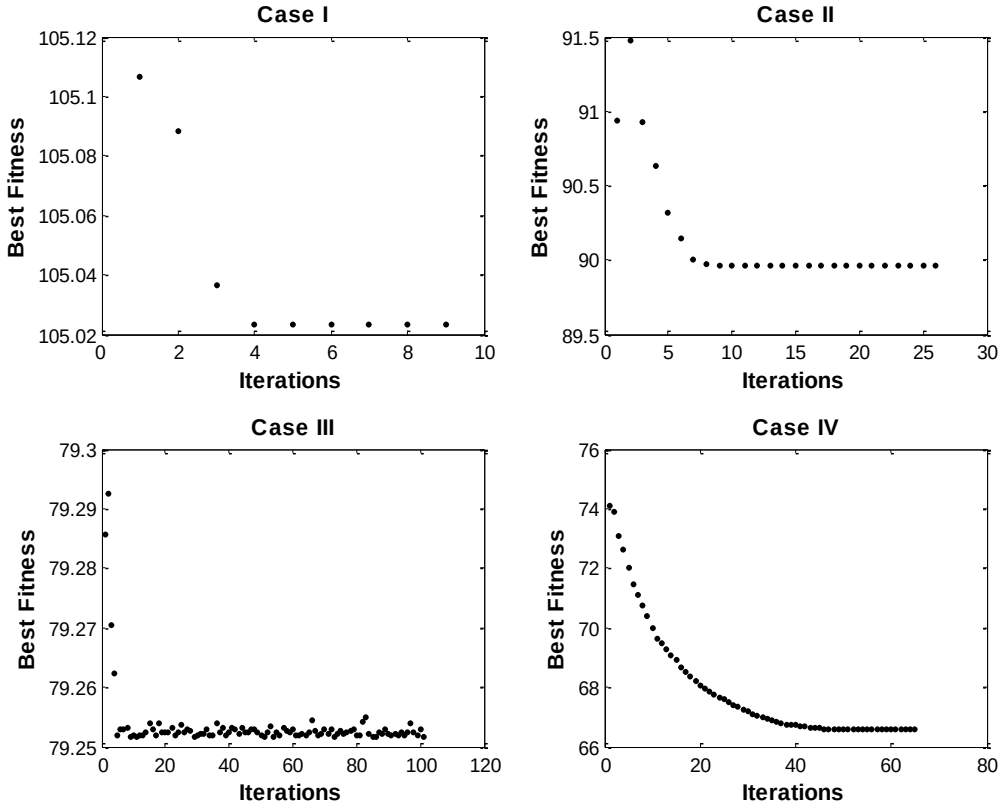


FIGURE II. Convergence characteristic of Clonal selection algorithm for 33 bus test system.

TABLE V.
PERFORMANCE OF CLONAL SELECTION ALGORITHM FOR IEEE 33 BUS SYSTEM

Total real power loss (kW)	Case I	Case II	Case III	Case IV
Min	104.813	89.752	79.012	66.029
Average	105.023	89.9619	79.2515	66.5892
Max	105.424	90.12	79.856	66. 759
No. of Antibodies	50	50	50	50
Avg. No. of iterations	8.257	24.384	62.896	67.903
Average Time (Sec.)	1.563	9	21.25	28.937

5.1. Comparison Performance

To demonstrate the validity of the proposed method the results of proposed method are compared with an existing PSO method [20]. The comparison is shown in Table VI.

TABLE IV.
COMPARISON OF RESULTS OF IEEE 33-BUS SYSTEM BY PROPOSED METHOD AND OTHER EXISTING METHOD

Case	**Bus locations**	**sizes(Mw)**		**Total Size(Mw)**		**saving(Kw)**	
		Clonalg	**PSO**	**Clonalg**	**PSO**	**Clonalg**	**PSO**
1	6	2.5775	2.5775	2.5775	2.5775	98.8857	98.8857
2	6	1.9707	1.9707	2.5464	2.5464	113.9469	113.9469
	15	0.5757	0.5757				
3	6	1.7569	1.7569	3.1152	3.1152	124.6562	124.6562
	15	0.5757	0.5757				
	25	0.7826	0.7826				
4	6	1.0765	1.0765	3.0884	3.0884	137.3196	137.3196
	15	0.5757	0.5757				
	25	0.7824	0.7824				
	32	0.6538	0.6538				

From the above table it is clear that the proposed method is producing the results that matches with those of existing method. To demonstrate the supremacy of the proposed method the convergence characteristics are compared with that of PSO algorithm as shown in Table VII. Both the number of iterations and computation time are less for Clonal selection algorithm. The only disadvantage of the proposed method is, it is producing slightly different results for each run.

6. CONCLUSIONS

In this paper, a two-stage methodology of finding the optimal locations and sizes of DGs for maximum loss reduction of radial distribution systems is presented. An analytical method is proposed to find the optimal DG locations and a Clonal Selection algorithm is proposed to find the optimal DG sizes. Voltage and line loading constraints are included in the algorithm.

The validity of the proposed method is proved from the comparison of the results of the proposed method with other existing methods. The results proved that the proposed algorithm is simple in nature than GA and PSO so it takes less computation time. This algorithm has the advantage of dynamically changing population size for each iteration depending on fitness value. By installing DGs at all the potential locations, the total power loss of the system has been reduced drastically and the voltage profile of the system is also improved. Inclusion of the real time constrains such as time varying loads and different types of DG units and discrete DG unit sizes into the proposed algorithm is the future scope of this work.

REFERENCES

[1]. G. Celli and F. Pilo, "Optimal distributed generation allocation in MV distribution networks", Power Industry Computer Applications, 2001. Pica 2001. Innovative Computing For Power - Electric Energy Meets The Market. 22nd IEEE Power Engineering Society International Conference, May 2001, pp. 81-86.

[2]. P.A. Daly, J. Morrison, "Understanding the potential benefits of distributed generation on power delivery systems," *Rural Electri Power Conference*, 29 April – 1 May 2001, pp. A211 – A213.

[3]. P. Chiradeja, R. Ramakumar, "An approach to quantify the technical benefits of distributed generation" *IEEE Trans Energy Conversion*, vol. 19, no. 4, pp. 764-773, 2004.

[4]. "Kyoto Protocol to the United Nations Framework Convention on climate change" , http://unfccc.int/resource/docs/convkp/kpeng.html

[5]. R.E. Brown, J. Pan, X. Feng, and K. Koutlev, "Siting distributed generation to defer T&D expansion," *Proc. IEE. Gen, Trans and Dist*, vol. 12, pp. 1151- 1159, 1997.

[6]. E. Diaz-Dorado, J. Cidras, E. Miguez, "Application of evolutionary algorithms for the planning of urban distribution networks of medium voltage", *IEEE Trans. Power Systems* , vol. 17, no. 3, pp. 879-884, Aug 2002.

[7]. M. Mardaneh, G. B. Gharehpetian, "Siting and sizing of DG units using GA and OPF based technique," *TENCON. IEEE Region 10 Conference*, vol. 3, pp. 331-334, 21-24, Nov.2004.

[8]. Silvestri A.Berizzi, S. Buonanno, "Distributed generation planning using genetic algorithms" *Electric Power Engineering, Power Tech Budapest* 99, *Inter. Conference*, pp.257, 1999.

[9]. Naresh Acharya, Pukar Mahat, N. Mithulanathan, "An analytical approach for DG allocation in primary distribution

network", *Electric Power and Energy Systems*, vol. 28, pp. 669-678, 2006

[10]. G.Celli,E.Ghaini,S.Mocci and F.Pilo, "A multi objective evolutionary algorithm for the sizing and sitting of distributed generation",IEEE Transactions on power systems,vol.20,no.2,pp.750-757,May 2005.

[11]. G.Carpinelli,G.Celli, S.Mocci and F.Pilo,"Optimization of embedded sizing and sitting by using a double trade-off method", IEE proceeding on generation, transmission and distribution, vol.152,no.4, pp.503-513, 2005.

[12]. C.L.T.Borges and D.M.Falcao, "Optimal distributed generation allocation for reliability,losses and voltage improvement", International journal of power and energy systems,vol.28.no.6,pp.413-420,July 2006.

[13]. Wichit Krueasuk and Weerakorn Ongsakul, "Optimal Placement of Distributed Generation Using Particle Swarm Optimization",M.Tech Thesis, AIT, Thailand.

[14]. De Castro, Leandro Nunes: Immune, Swarm, and Evolutionary Algorithms, Part I: Basic Models. Proceeding of the ICONIP, Workshop on Artificial Immune Systems, Vol. 3. November 18-22, Singapore (2002) 1464-1468.

[15]. Farmer, J.D., Packard, N.H. and Perelson, A.S.: The immune system, adaptation, and machine learning. Physica, 22D (1986) 187-204.

[16]. De Castro, L.N. and Von Zuben, F.J.: The Clonal Selection Algorithm with Engineering Applications. Proceedings of The Genetic and Evolutionary Computation Conference 2000 (GECCO'OO) - Workshop Proceedings. July 8-12, Las Vegas, USA (2000) 36-37.

[17]. Forrest, S., Perelson, A.S., Allen, L. and Cherukuri, R.: Self-nonself discrimination in a computer. Proceedings of the 1994 IEEE Symposium on Research in Security and Privacy Los Alamitos, CA: IEEE Computer Society Press (1994) 202-212.

[18]. M.Padma Lalitha,V.C.Veera Reddy, N.Usha "DG Placement Using Fuzzy For Maximum Loss Reduction In Radial Distribution System" International journal of computer applications in engineering, technology and sciences, Vol. 2,issue1, Oct 09-March 2010,pp. 50-55.

[19]. M. E. Baran and F. F. Wu, " Network reconfiguration in distribution systems for loss reduction and load balancing", IEEE Transactions on Power Delivery, Vol. 4, No 2 Apr 1989, pp. 1401-1407.

[20]. Mrs. M. Padma Lalitha , Dr.V.C. Veera Reddy, V. Usha Reddy "Optimal DG Placement For Minimum Real Power Loss In Radial Distribution Systems Using PSO" Journal Of Theoretical And Applied Information Technology ,Vol.13,No.2,March,2010,pp.107-116.

On Some Transformation Formulas for The $\overline{H}$-Function

Yashwant Singh
Department of Mathematics,
S.M.L.(P.G.)College, Jhunjhunu (Rajasthan), India
E-mail: ysingh23@yahoo.co.in

Abstract: In the present paper we establish four transformations of double infinite series involving the $\overline{H}$ -function. These formulas are then used to obtain double summation formulas for the said function. Our results are quite general in character and a number of summation formulas can be deduced as particular cases. Several interesting special cases of our main finding have been mentioned briefly.

Key words: $\overline{H}$ -function, Gauss's summation theorem, Double infinite series
(2000 Mathematics subject classification: 33C99)

1. Introduction

The $\overline{H}$ -function occurring in the paper will be defined and represented as follows:

$$\overline{H}_{P,Q}^{M,N}\left[z\right]=\overline{H}_{P,Q}^{M,N}\left[z\left|\begin{matrix}(a_j;\alpha_j;A_j)_{1,N},(a_j;\alpha_j)_{N+1,P}\\(b_j,\beta_j)_{1,M},(b_j,\beta_j;B_j)_{M+1,Q}\end{matrix}\right.\right]=\frac{1}{2\pi i}\int_{-i\infty}^{i\infty}\bar{\phi}(\xi)z^{\xi}d\xi \tag{1.1}$$

where $$\bar{\phi}(\xi)=\frac{\prod_{j=1}^{M}\Gamma(b_j-\beta_j\xi)\prod_{j=1}^{N}\left\{\Gamma(1-a_j+\alpha_j\xi)\right\}^{A_j}}{\prod_{j=M+1}^{Q}\left\{\Gamma(1-b_j+\beta_j\xi)\right\}^{B_j}\prod_{j=N+1}^{P}\Gamma(a_j-\alpha_j\xi)} \tag{1.2}$$

Which contains fractional powers of the gamma functions. Here, and throughout the paper $a_j\,(j=1,\ldots,p)$ and $b_j\,(j=1,\ldots,Q)$ are complex parameters, $\alpha_j\geq 0(j=1,\ldots,P),\beta_j\geq 0(j=1,\ldots,Q)$ (not all zero simultaneously) and exponents $A_j\,(j=1,\ldots,N)$ and $B_j\,(j=N+1,\ldots,Q)$ can take on non integer values.

The following sufficient condition for the absolute convergence of the defining integral for the $\overline{H}$ -function given by equation (1.1) have been given by Buschman and Srivastava [1].

$$\Omega\equiv\sum_{j=1}^{M}\left|\beta_j\right|+\sum_{j=1}^{N}\left|A_j\alpha_j\right|-\sum_{j=M+1}^{Q}\left|\beta_jB_j\right|-\sum_{j=N+1}^{P}\left|\alpha_j\right|>0 \tag{1.3}$$

and $$\left|\arg(z)\right|<\frac{1}{2}\pi\Omega \tag{1.4}$$

The behavior of the $\overline{H}$ -function for small values of $\left|z\right|$ follows easily from a result recently given by Rathie ([7],p.306,eq.(6.9)).
We have

$$\overline{H}_{P,Q}^{M,N}\left[z\right]=0\left(\left|z\right|^{\gamma}\right),\gamma=\min_{1\leq j\leq N}\left[\mathrm{Re}\left(b_j/\beta_j\right)\right],\left|z\right|\to 0 \tag{1.5}$$

If we take $A_j=1\left(j=1,\ldots,N\right)$, $B_j=1\left(j=M+1,\ldots,Q\right)$ in (1.1), the function $\overline{H}_{P,Q}^{M,N}$ reduces to the Fox's H-function [3].
We shall use the following notation:

$A^* = \left(a_j, \alpha_j; A_j\right)_{1,N}, \left(a_j, \alpha_j\right)_{N+1,P}$ and $B^* = \left(b_j, \beta_j\right)_{1,M}, \left(b_j, \beta_j; B_j\right)_{M+1,Q}$

2. Transformation Formulas:

In this section we establish the following four transformation Formulas for the $\overline{H}$ -function:

First formula

$$\sum_{m,n=0}^{\infty} x^m y^n \overline{H}_{P+2,Q+1}^{M,N+2}\left[z \left| \begin{matrix} (1-a-m,\rho;1),(1-b-n,\sigma;1),A^* \\ B^*,(1-a-b-m-n,\sigma+\rho;1) \end{matrix} \right. \right] = (x+y-xy)^{-1}$$

$$\sum_{s=0}^{\infty} x^{s+1} \overline{H}_{P+2,Q+1}^{M,N+2}\left[z \left| \begin{matrix} (1-a-s,\rho;1),(1-b,\sigma;1),A^* \\ B^*,(1-a-b-s,\sigma+\rho;1) \end{matrix} \right. \right] + \sum_{t=0}^{\infty} y^{t+1} \overline{H}_{P+2,Q+1}^{M,N+2}\left[z \left| \begin{matrix} (1-a,\rho;1),(1-b-t,\sigma;1),A^* \\ B^*,(1-a-b-t,\sigma+\rho;1) \end{matrix} \right. \right] \qquad (2.1)$$

The formula (2.1) is valid, if the following (sufficient) conditions are satisfied.

$$(i)\ \rho,\sigma > 0 \qquad (ii)\ \Omega - \rho - \sigma > 0\ , \left|\arg z\right| < \frac{1}{2}\left(\Omega - \rho - \sigma\right)\pi$$

$$(iii) \max\left\{\left|x\right|, \left|y\right|\right\} < 1 \quad or\ x = y = 1 \quad with \quad Re(a) > 1, \mathrm{Re}(b) > 1$$

Second formula

$$\sum_{m,n=0}^{\infty} \frac{x^m y^n}{m!n!} \overline{H}_{P+2,Q+1}^{M,N+2}\left[z \left| \begin{matrix} (1-a-m-n,u;1),(1-b-m,v;1),A^* \\ B^*,(1-c-m,\omega;1) \end{matrix} \right. \right]$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!}(1-y)^{-a} \left(\frac{x}{1-y}\right)^k \overline{H}_{P+2,Q+1}^{M,N+2}\left[z\left(1-y\right)^{-u} \left| \begin{matrix} (1-a-k,u;1),(1-b-k,v;)1,A^* \\ B^*,(1-c-k,\omega;1) \end{matrix} \right. \right] \qquad (2.2)$$

Provided that

$$(i) u,v,\omega > 0 \qquad (ii)\ \Omega - \omega > 0, \left|\arg z\right| < \frac{1}{2}\left(\Omega - \omega\right)\pi$$

$$(iii)\ \left|x\right| + \left|y\right| < 1\ and\ either \left|\frac{x}{1-y}\right| < 1\ or \left|\frac{x}{1-y}\right| = 1\ with\ \mathrm{Re}\left(c-a-b\right) > 0$$

Third formula

$$\sum_{m,n=0}^{\infty} \frac{x^m y^n}{m!n!} \overline{H}_{P+3,Q+2}^{M,N+3}\left[z \left| \begin{matrix} (1-a-m-n,u;1),(1-b-m,v;1),(1-b'-n,\omega;1)A^* \\ B^*,(1-a-m,u;1),(1-a-n,u;1) \end{matrix} \right. \right]$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!}(1-x)^{-b}(1-y)^{b'} \left(\frac{xy}{(1-x)(1-y)}\right)^k \overline{H}_{P+2,Q+1}^{M,N+2}\left[z\left(1-x\right)^{-v}(1-y)^{-u} \left| \begin{matrix} (1-b-k,v;1),(1-b'-k,\omega;1),A^* \\ B^*,(1-a-k,u;1) \end{matrix} \right. \right]$$

(2.3)

Provided that

$$(i) u,v,\omega > 0 \qquad (ii)\ \Omega - 2u > 0, \left|\arg z\right| < \frac{1}{2}\left(\Omega - 2u\right)\pi$$

$$(iii)\ \left|x\right| + \left|y\right| < 1\ and\ either \left|\frac{xy}{(1-x)(1-y)}\right| < 1\ or \left|\frac{xy}{(1-x)(1-y)}\right| = 1\ with\ \mathrm{Re}\left(a-b-b'\right) > 0$$

Fourth formula

$$\sum_{m,n=0}^{\infty} \frac{x^m y^n}{m!n!} \overline{H}_{P+3,Q+1}^{M,N+3}\left[z \left| \begin{matrix} (1-a-m-n,u;1),(1-b-m,v;1),(1-b'-n,\omega;1)A^* \\ B^*,(1-b-b'-m-n,\omega+v;1) \end{matrix} \right. \right]$$

$$=\sum_{k=0}^{\infty}(1-y)^{-a}\frac{1}{K!}\left(\frac{x-y}{1-y}\right)^{k}\overline{H}_{P+3,Q+1}^{M,N+3}\left[z(1-y)^{-u}\left|\begin{matrix}(1-a-k,u;1),(1-b-k,v;1),(1-b',\omega;1),A^{*}\\B^{*},(1-b-b'-k,\omega+v;1)\end{matrix}\right.\right] \tag{2.4}$$

Provided that

$$(i)\,u,v,\omega>0 \qquad (ii)\ \Omega-v-\omega>0,\left|\arg z\right|<\frac{1}{2}(\Omega-v-\omega)\pi$$

$$(iii)\ \max\left\{|x|,|y|\right\}<1\text{, either }\left|\frac{x-y}{1-y}\right|<1\ or\left|\frac{x-y}{1-y}\right|=1\ \ with\ \mathrm{Re}(b'-a)>0$$

In all the aforementioned formulas Ω is given by (1.3).

Derivation of the first formula: Using Mellin-Barnes type of contour integral (1.1) for the $\overline{H}$ -function occurring on the L.H.S. of (2.1) and changing the order of integration and summation, we find that L.H.S. of (2.1).

$$=\frac{1}{2\pi i}\int_{-i\infty}^{i\infty}\bar{\phi}(\xi)z^{\xi}\frac{\Gamma(a+\rho\xi)\Gamma(b+\sigma\xi)}{\Gamma(a+b+(\sigma+\rho)\xi)}F_{2}\left[a+\rho\xi,b+\sigma\xi,1,1,;a+b+(\sigma+\rho)\xi;x,y\right]d\xi \tag{2.5}$$

Now appealing to a known result due to Srivastava ([8], p.1259, eq. (2.2))

$$F_{2}\left[a,b,1,1;a+b;x,y\right]=(x+y-xy)^{-1}\left\{x,\ {}_{2}F_{1}\left[a,1;a+b;x\right]+y\ {}_{2}F_{1}\left[b,1;a+b;x\right]\right\} \tag{2.6}$$

in (2.6) , we get L.H.S. of (2.1)

$$=\frac{1}{2\pi i}\int_{-i\infty}^{i\infty}\bar{\phi}(\xi)z^{\xi}\frac{\Gamma(a+\rho\xi)\Gamma(b+\sigma\xi)}{\Gamma(a+b+(\sigma+\rho)\xi)}(x+y-xy)^{-1}$$

$$\left\{x\,{}_{2}F_{1}\left[a+\rho\xi,1;a+b+(\rho+\sigma)\xi;x\right]+y\,{}_{2}F_{1}\left[b+\sigma\xi,1;a+b+(\rho+\sigma)\xi;y\right]\right\}d\xi \tag{2.7}$$

Now expressing the ${}_{2}F_{1}$ functions in terms of their series and changing the order of integration and summation, and interpreting the result so obtained with the help of (1.1), we arrived at the formula (2.1). Derivation of the formulas (2.2) to (2.4) : The summation formulas (2.2), (2.3) and (2.4) can be developed on the lines similar to the formula (2.1) except that, in place of(2.6) , here we use the following known results ([2],p.238,eq.(2),eq.(3) and eq.(1) respectively):

$$F_{2}\left[\alpha,\beta,\beta';\gamma,\beta;x,y\right]=(1-y)^{-\alpha}\ {}_{2}F_{1}\left[\alpha,\beta;\gamma;\frac{x}{(1-y)}\right] \tag{2.8}$$

$$F_{2}\left[\alpha,\beta,\beta';\alpha,\alpha;x,y\right]=(1-x)^{-\beta}(1-y)^{-\beta}\ {}_{2}F_{1}\left[\beta,\beta';\alpha;\frac{xy}{(1-x)(1-y)}\right] \tag{2.9}$$

$$F_{2}\left[\alpha,\beta,\beta';\beta+\beta';x,y\right]=(1-y)^{-\alpha}\,{}_{2}F_{1}\left[\alpha,\beta;\beta+\beta';\frac{x-y}{(1-y)}\right] \tag{2.10}$$

3. Summation Formulas:

If we take $x = y = 1$ in (2.1) and use the well known Gauss's summation theorem, we arrived at the result

$$\sum_{m,n=0}^{\infty}x^{m}y^{n}\overline{H}_{P+2,Q+1}^{M,N+2}\left[z\left|\begin{matrix}(1-a-m,\rho;1),(1-b-n,\sigma;1),A^{*}\\B^{*},(1-a-b-m-n,\sigma+\rho;1)\end{matrix}\right.\right]=$$

$$\overline{H}_{P+2,Q+1}^{M,N+2}\left[z\left|\begin{matrix}(1-a-m,\rho;1),(2-b,\sigma;1),A^{*}\\B^{*},(2-a-b,\sigma+\rho;1)\end{matrix}\right.\right]+\overline{H}_{P+2,Q+1}^{M,N+2}\left[z\left|\begin{matrix}(2-a,\rho;1),(1-b,\sigma;1),A^{*}\\B^{*},(2-a-b,\sigma+\rho;1)\end{matrix}\right.\right] \tag{3.1}$$

Valid under the conditions of (2.1) .

Again if we put $x = y = \frac{1}{2}$ in (2.2), $y = 1-x$ in (2.3) and $x = 1$ in (2.4) and make use of well known Gauss's summation theorem therein, we shall arrive at the following results

$$\sum_{m,n=0}^{\infty} \frac{\left(\frac{1}{2}\right)^{m+n}}{m!n!} \overline{H}^{M,N+2}_{P+2,Q+1}\left[z \left| \begin{matrix} (1-a-m-n,u;1),(1-b-m,v;1),A^* \\ B^*,(1-c-m,\omega;1) \end{matrix} \right. \right]$$

$$= 2^a \overline{H}^{M,N+3}_{P+3,Q+2}\left[2^u z \left| \begin{matrix} (1-a,u;1),(1-b,v;1),(1-c+a+b,\omega-u-v;1)A^* \\ B^*,(1-c+a,\omega-u;1),(1-c+b,\omega-v;1) \end{matrix} \right. \right] \qquad (3.2)$$

Where $\omega - u - v > 0, \omega \neq u \ \ or \ \ \omega \neq v$

And valid under the conditions of (2.1)

$$\sum_{m,n=0}^{\infty} \frac{x^m (1-x)^n}{m!n!} \overline{H}^{M,N+3}_{P+3,Q+2}\left[z \left| \begin{matrix} (1-a-m-n,u;1),(1-b-m,v;1),(1-b'-n,\omega;1)A^* \\ B^*,(1-a-m,u;1)(1-a-n,u;1) \end{matrix} \right. \right]$$

$$= x^{-b'}(1-x)^{-b} \overline{H}^{M,N+3}_{P+3,Q+2}\left[zx^{-\omega}\left(1-x\right)^{-\nu} \left| \begin{matrix} (1-b,v;1),(1-b',\omega;1),(1-a+b+b',u-v;1),A^* \\ B^*,(1-a+b,u-v;1),(1-a-b',u-\omega;1) \end{matrix} \right. \right] \qquad (3.3)$$

Where $u - v - \omega > 0, \ u \neq \omega, \ u \neq v$

And valid under the conditions of (2.3).

$$\sum_{m,n=0}^{\infty} \frac{y^n}{m!n!} \overline{H}^{M,N+3}_{P+3,Q+1}\left[z \left| \begin{matrix} (1-a-m-n,u;1),(1-b-m,v;1),(1-b'-n,\omega;1),A^* \\ B^*,(1-b-b'-m-n,\omega+v;1) \end{matrix} \right. \right]$$

$$= (1-y)^{-a} \overline{H}^{M,N+3}_{P+3,Q+1}\left[z(1-y)^{-u} \left| \begin{matrix} (1-a,u;1),(1-b,v;1),(1-b'+a,\omega-u;1),A^* \\ B^*,(1-b-b'+a,\omega-u+v;1) \end{matrix} \right. \right] \qquad (3.4)$$

Where $v \neq \omega, v \neq u$ and valid under the conditions (2.4).

4. Special Cases:

(i) In (2.1), taking M=1, N=0=P, Q=2, b_1=0, b_2= - λ , β_1 =1, $\beta_2 = \nu$, the $\overline{H}$ function reduces to generalized Wright-Bessel function $\overline{J}^{\nu,\mu}_{\lambda}$ ([4],p.271,eq.(8)) and we get

$$\sum_{m,n=0}^{\infty} x^m y^n \frac{\Gamma(a+m+\rho\xi)\Gamma(b+n+\sigma\xi)}{\Gamma(a+b+m+n+(\sigma+\rho)\xi)} \overline{J}^{\nu,\mu}_{\lambda}[z] = (x+y-xy)^{-1}$$

$$\left\{ \sum_{s=0}^{\infty} x^{s+1} \overline{H}^{1,2}_{2,3}\left[z \left| \begin{matrix} (1-a-s,\rho;1),(1-b,\sigma;1) \\ (0,1),(-\lambda,\nu;\mu),(1-a-b-s,\sigma+\rho;1) \end{matrix} \right. \right] + \sum_{t=0}^{\infty} y^{t+1} \overline{H}^{1,2}_{2,2}\left[z \left| \begin{matrix} (1-a,\rho;1),(1-b-t,\sigma;1) \\ (0,1),(-\lambda,\nu;\mu),(1-a-b-t,\sigma+\rho;1) \end{matrix} \right. \right] \right\} \qquad (4.1)$$

where $(1-\nu) > 0, (1+\nu) \geq 0, |\arg z| < \frac{1}{2}(1-\nu-\rho-\sigma)\pi$ and the conditions (i) and (iii) given with (2.1) also satisfied.

(ii) In (2.1) replacing M,N,P,Q, by 1,P,P,Q+1 respectively, the $\overline{H}$ function reduces to the Wright's generalized hyper geometric function ${}_P\overline{\Psi}_Q([4], p.271, eq.(7))$ and we get

$$\sum_{m,n=0}^{\infty} x^m y^n \frac{\Gamma(a+m+\rho\xi)\Gamma(b+n+\sigma\xi)}{\Gamma(a+b+m+n+(\sigma+\rho)\xi)} {}_P\overline{\Psi}_Q\left[z \left| \begin{matrix} (a_j,\alpha_j;A_j)_{1,P} \\ (b_j,\beta_j;B_j)_{1,Q} \end{matrix} \right. \right]$$

$$= (x+y-xy)^{-1}\left\{\begin{array}{l}\displaystyle\sum_{s=0}^{\infty} x^{s+1}\overline{H}^{1,p+2}_{p+2,q+2}\left[z\left|\begin{matrix}(1-a-s,\rho;1),(1-b,\sigma;1)(a_j,\alpha_j;A_j)_{1,P}\\(0,1),(b_j,\beta_j;B_j)_{1,Q},(1-a-b-s,\sigma+\rho;1)\end{matrix}\right.\right]+\\ \displaystyle\sum_{t=0}^{\infty} y^{t+1}\overline{H}^{1,p+2}_{p+2,q+2}\left[z\left|\begin{matrix}(1-a,\rho;1),(1-b-t,\sigma;1),(a_j,\alpha_j;A_j)_{1,P}\\(0,1),(b_j,\beta_j;B_j)_{1,Q},(1-a-b-t,\sigma+\rho;1)\end{matrix}\right.\right]\end{array}\right\} \quad (4.2)$$

where

$\sum_{j=1}^{P}\alpha_j + 1 - \sum_{j=1}^{Q}\beta_j \equiv T > 0;\quad |\arg z| < \frac{1}{2}(T-\rho-\sigma)\pi,\ 1+\sum_{j=1}^{Q}\beta_j - \sum_{j=1}^{P}\alpha_j \geq 0$ and the conditions (i) and (iii) given with (2.1) also satisfied.

(iii) The function $g_1 = (-1)^p g(\gamma,\eta,\tau,p,z)$ $([4], p.271, eq.(10))$ where

$$g_1 = (-1)^p g(\gamma,\eta,\tau,p,z) = \frac{K_{d-1}\Gamma(p+1)\Gamma\left(\frac{1}{2}+\frac{\tau}{2}\right)}{2^{2+p}\pi^{1/2}\Gamma\left(\gamma-\frac{\tau}{2}\right)\Gamma(\gamma)}\overline{H}^{1,3}_{3,3}\left[-z\left|\begin{matrix}(1-\gamma,1;1),(1-\gamma+\tau/2,1;1),(1-\eta,1;1+p)\\(0,1),(-\frac{\tau}{2},1;1),(-\eta,1;1+p)\end{matrix}\right.\right]$$

Where $K_d = \frac{2^{1-d}\pi^{-d/2}}{\Gamma(d/2)}$ $\left([6], p.4121, eq.(5)\right)$ (4.3)

From this we get

$$\sum_{n,n=0}^{\infty} x^m y^n \frac{K_{d-1}\Gamma(p+1)\Gamma\left(\frac{1}{2}+\frac{\tau}{2}\right)\Gamma(a+m+\rho\xi)\Gamma(b+n+\sigma\xi)}{2^{2+p}\pi^{1/2}\Gamma\left(\gamma-\frac{\tau}{2}\right)\Gamma(\gamma)\Gamma(a+b+m+n+(\rho+\sigma)\xi)}$$

$$\overline{H}^{1,3}_{3,3}\left[z\left|\begin{matrix}(1-\gamma,1;1),(1-\gamma+\tau/2,1;1),(1-\eta,1;1+p),(1-a-m,\rho;1),(1-b-n,\sigma;1)\\(0,1),(-\frac{\tau}{2},1;1),(-\eta,1;1+p),(1-a-b-m-n,\sigma+\rho;1)\end{matrix}\right.\right]$$

$$= \left(x+y-xy\right)^{-1}\frac{K_{d-1}\Gamma(p+1)\Gamma\left(\frac{1}{2}+\frac{\tau}{2}\right)}{2^{2+p}\pi^{1/2}\Gamma\left(\gamma-\frac{\tau}{2}\right)\Gamma(\gamma)}$$

$$\left\{\sum_{s=0}^{\infty} x^{s+1}\overline{H}^{1,5}_{5,4}\left[z\left|\begin{matrix}(1-\gamma,1;1),(1-\gamma+\tau/2,1;1),(1-\eta,1;1+p),(1-a-s,\rho;1),(1-b,\sigma;1)\\(0,1),(-\frac{\tau}{2},1;1),(-\eta,1;1+p),(1-a-b-s,\sigma+\rho;1)\end{matrix}\right.\right]+\right.$$

$$\left.\sum_{t=0}^{\infty} y^{t+1}\overline{H}^{1,5}_{5,4}\left[z\left|\begin{matrix}(1-\gamma,1;1),(1-\gamma+\tau/2,1;1),(1-\eta,1;1+p),(1-a,\rho;1),(1-b-t,\sigma;1)\\(0,1),(-\frac{\tau}{2},1;1),(-\eta,1;1+p),(1-a-b-t,\sigma+\rho;1)\end{matrix}\right.\right]\right\} \quad (4.4)$$

valid under the conditions of (2.1).

Further if we take $\rho \to 0$ in (2.1), we get the following new transformation formula:

$$\sum_{m,n=0}^{\infty} x^m y^n (a)_m \overline{H}^{M,N+1}_{P+1,Q+1}\left[z\left|\begin{matrix}(1-b-n,\sigma;1),A^*\\B^*,(1-a-b-m-n,\sigma+\rho;1)\end{matrix}\right.\right]$$

$$= (x+y-xy)^{-1}\left\{\sum_{s=0}^{\infty} x^{s+1}(a)_s \overline{H}^{M,N+1}_{P+1,Q+1}\left[z\left|\begin{matrix}(1-b,\sigma;1),A^*\\ B^*,(1-a-b-s,\sigma;1)\end{matrix}\right.\right] + \sum_{t=0}^{\infty} y^{t+1}\overline{H}^{M,N+1}_{P+1,Q+1}\left[z\left|\begin{matrix}(1-b-t,\sigma;1),A^*\\ B^*,(1-a-b-t,\sigma;1)\end{matrix}\right.\right]\right\} \quad (4.5)$$

Valid under the conditions of (2.1) .

A similar type of result can be obtained by taking $\sigma \to 0$ in (2.1).

References:

1. Buschman, R.G. and Srivastava , H.M., The $\overline{H}-$function associated with a certain class of Feynman integrals, J.Phys.A:Math. Gen. 23,(1990),4707-4710.
2. Erdelyi,A.et.al.,Higher Traanscendental Functions ,Vol.1,Mc.Graw-Hill , New York (1953).
3. Fox.C., The G.and H-function as symmetrical Fourier kernels, Trans. Amer. Math.Soc. 98,(1961),395-429.
4. Gupta,K.C. , Jain, Rashmi and Sharma,Arti, A study of unified finite integral with applications,J.Rajashthan Acad. Phys. Sci. Vol.2,No.4,(2003),269-282.
5. Gupta,Rajni, Transformation formulas for the multivariable H- function , Indian Journal of Mathematics, Vol. 30,N0.2,(1988),105-118.
6. Inayat-Hussian,A.A., New properties of hypergeometric series derivable from Feynman integrals II. A generalization of the H-function,j. Phy. A: Math. 20, (1987),4119-4128.
7. Rathie,A.K.,A new generalization of generalized hypergeometric functions, Le Mathematiche Fasc.II, 52,(1997),297-310.
8. Srivastava,H.M. , Some formulas of Hermite and Carlitiz, Rev. Roumaine Math.Pures Appl.,17,(1972),1257-1263.
9. Srivastava,H.M.,Gupta, K.C. and Goyal,S.P., The H-Functions of One and Two Varibale with Applications,South Asian Publishers, New Delhi and Madras(1982).

A Web Based Rose Crop Expert Information System Based on Artificial Intelligence And Machine Learning Algorithms

Prof.M.S.Prasad Babu[1] S.V.N.L.Narayana[2] ch.viswanadha sarma[3] K.A.Suryanarayana[4]

1. Professor, Department of CS & SE, Andhra University, Visakhapatnam, Andhra Pradesh, India.
msprasadbabu@yahoo.co.in

2. Head of the Department, Department of Physics, M.R.P.G.College, Vizianagaram, Andhra Pradesh, India
laxminarayanasvn@yahoo.com

3.Sr.Asst.professor Department of mathematics, Vignan engineering college, Dwavada, visakhapatnam.
Andhra Pradesh, India
Vissu.viit@gmail.com

4. Head of the Department, Department of Economics, M.R.P.G.College, Vizianagaram, Andhra Pradesh, India
suryakas@yahoo.co.in

Abstract- The Rose Expert System is basically gives the information regarding the diseases, viruses and their symptoms and chemical controls which are used to prevent the diseases and viruses of cultivation of Rose.
This Rose Expert Advisory System is simulation of more than one expert. This system is developed in Java Server Pages. There are two modules "Information system" and "Advisory system" are embedded in present system. For advising the farmers in the disease identification and preventive measures of Rose plantation, this "Rose Expert Advisory System" uses backward chaining mechanism for building Rule Based Expert system which results in giving exact disease. If rule base system fails in giving the exact disease then Rose Expert Advisory System provides an alternative system using Machine Learning Techniques – Optimization Algorithm, which produces subset of diseases with probabilistic values. Then the system uses Particle Swarm Optimization Algorithm for each disease by taking some sign values from the user and gives a nearest disease to the user with the preventive measures.

Keywords:
Expert Advisory System – Information System – Rule Based-Optimization Algorithms- Particle Swarm Optimization Algorithm – Web Based – JSP – SQL

1. Introduction

Expert system can be defined as a tool for information generation from knowledge. Information is either found in various forms or generated from data and/or knowledge. Text, images, video, audio are forms of media on which information can be found, and the role of information technology is to invent, and devise tools to store and retrieve it. Information is generated from knowledge. The expert system gives advises.

Rose flowers are symbol of fragrance, beauty and to convey the message of love. In India roses are used to make bouquet, oil, rose water and gulkand. And also used for making incense sticks with Dry petals of roses. The native places of Roses are Asia, North America, China, Japan, Europe and Himalayas regions in India. There are about 150 species

1.1 Rose Varieties

1.2 Climate:

Rose can grow in various climates, and also good for flower production at bright sunshine It should be noted the quality rose blooms are obtained during December to April.

1.3 Temperature:

The growth of the Rose plant and flower are depending upon the temperature. For the rose the minimum and maximum temperatures are 7.9^{o} C 22.6^{o} C with minimum 6 sunshine hours during winter season. During summer and rain seasons flower quality and lifetime of rose and rose plant are to be affect.

1.4 Soil:

In fertile soil, the growth of Rose plants is well. If the proper drainage facilities in all types soil then the plant growth is well. In soil pH level upto 6.0 to 7.5, Roses grow well. If pH is less than 6 then the soil depth of minimum 45 cm is good for rose growing.

2. Proposed System:

The proposed system is Rose expert advisory system. It is divided into two aspects

1. Information System
2. Advisory System

In Information system, the user can get all the static information about different species, Diseases, Symptoms, chemical controls, Preventions, Pests, Virus of Rose flower and plants.

In Advisory System, the user is having an interaction with the expert system online; the expert system asks the questions. User has to answer which is given by the expert system. Depends on the response by the user the expert system decides the disease and displays its control measure of disease.

It is aimed at a collaborative venture with eminent Agriculture Scientist and Experts in the area of Rose Plantation with an excellent team of computer Engineers, programmers and designers.

This web application is expected to have the following features:

1) This web application provides time- to- time updates of Rose information to the users at their door steps regarding diseases, virus and its control measure which leads to good yields.

2) This site contains four major sections named Information Systems of Rose, Rose Advisory System, other services related to web application and an additional feature is links to other agriculture systems

3) The web directory service, articles and the discussion forum service provided in the website will help the floriculture fraternity in a greater way to interact each other to produce better findings in the area of floriculture field.

2.1. Functional Requirements for Rose Expert System:

2.1.1. Inputs:

The system needs the information about the symptoms from the user to produce the output.

2.1.2. Outputs:

The outputs of the system will be:

1) Information Diseases & Viruses

2) Small Description about the disease & Viruses

S.No	Name of variety	Flower color
1.	Black Velvet, Crimson Glory, Happiness	Red and dark red
2.	Eiffel Tower, First Love, First Prize	Pink
3.	Blue, Africa Star, Paradise	Lavender
4.	Hawaii, Super, Star, Duke of Windsor	Orange
5.	Virgo, White Christmas	White
6.	Careless Love	Novel Color
7.	Summer Sunshine, Golden Giant, Kiss of Fire, Double Delight	Yellow
8.	Perfecta (pink and white), Suspense (red and yellow)	Bicolor

3) Chemical controls & Nutrients

4) Preventions

2.1.3. Store:

The information collected through experts is stored as a database (Knowledge Base) that serves as a repository for quick processing and future retrieval. The system stores the following information in terms of html files.

1) About Rose system
2) About Rose Varieties
3) Climate and Soil
4) Plantation
5) Common Symptoms
6) Common Diseases
7) Chemical Control
8) Preventions

The System Stores the information related to Expert System in knowledge base in the following ways.

2.1.4. Rules A set of rules that constitute the program stored in a rule memory of production memory and on an inference engine using JSP files required to execute the rules.

2.1.5. Dataset: MySQL database can be used to store data in a database. The post-analysis of the monitoring data can be conducted by you for increases the opportunity.

3. Machine Learning

Architecture of Rose Crop Advisory Expert System

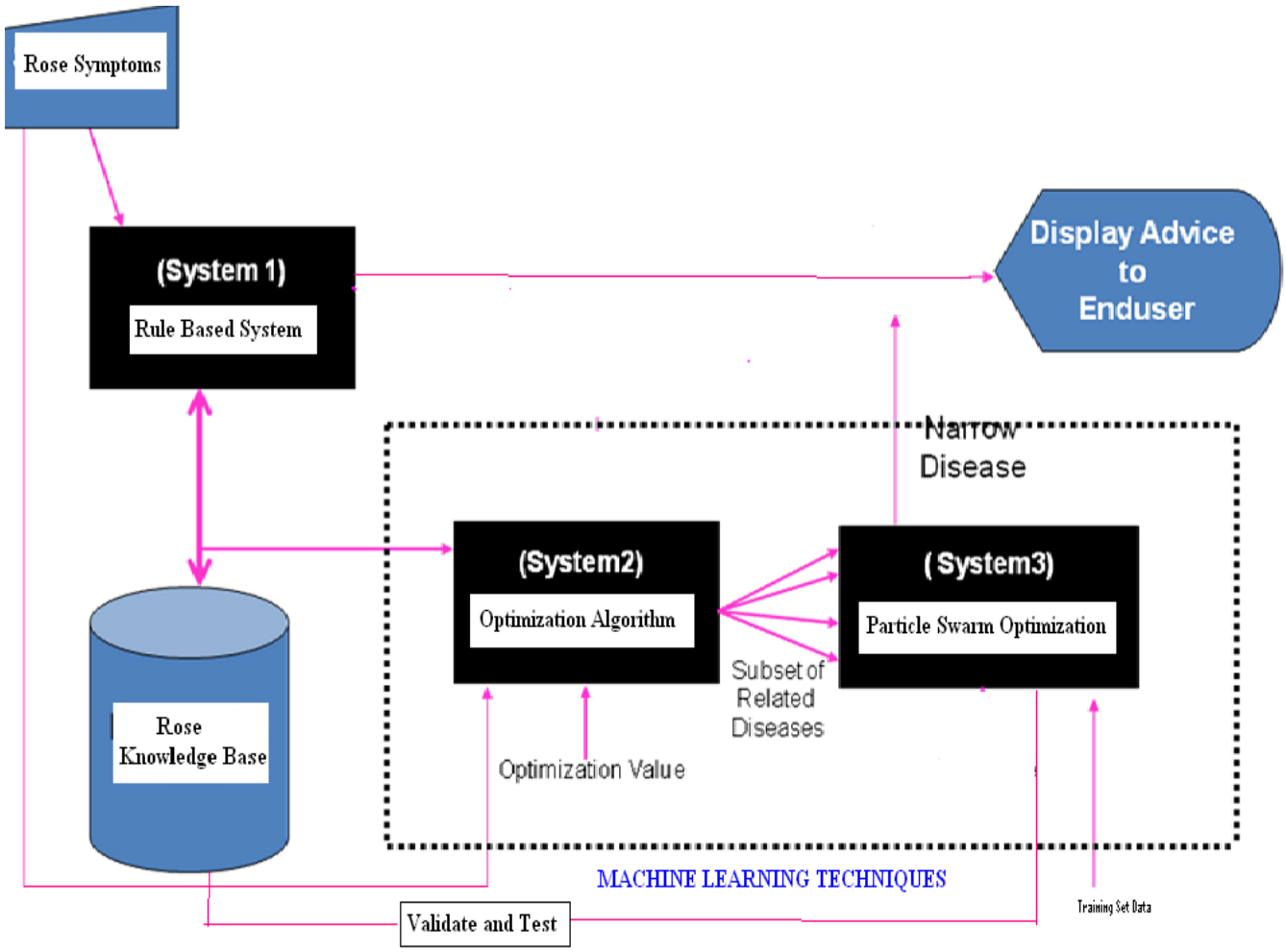


Fig 1. Rose Crop Advisory System

Fig.2 Architecture of subsystem –I (RULE BASED SYSTEM)

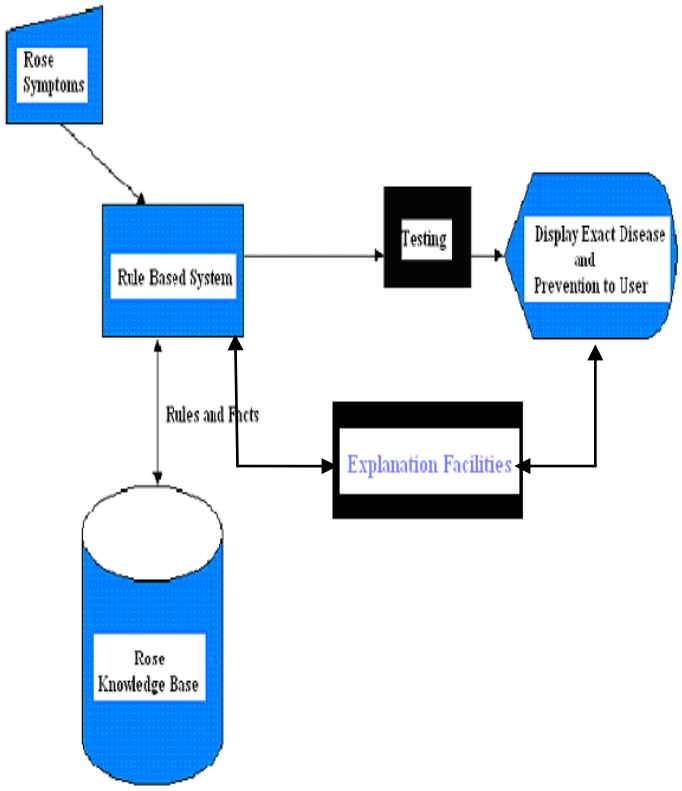


Fig 3. Architecture of subsystem –II (Optimization Algorithm):

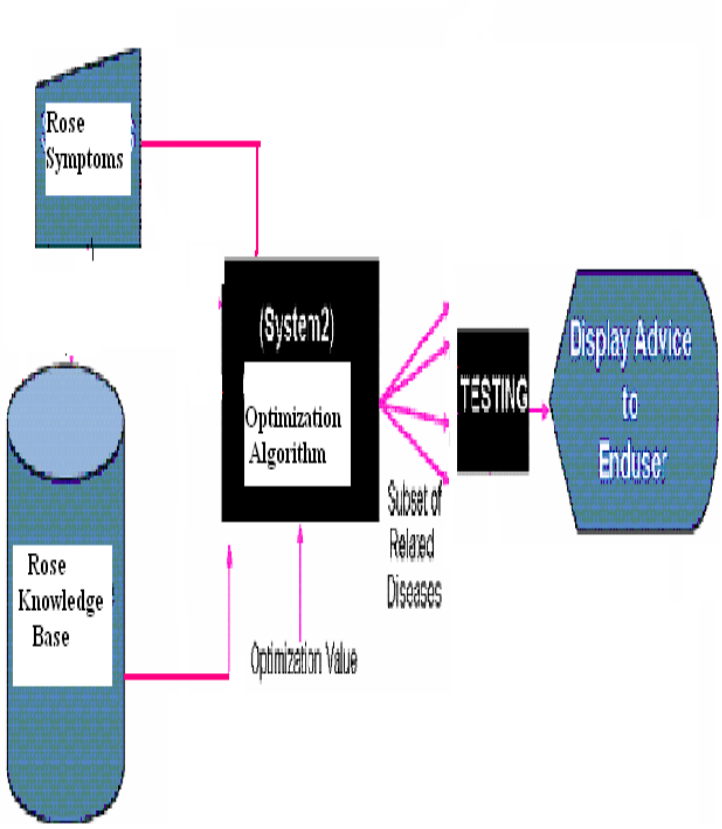
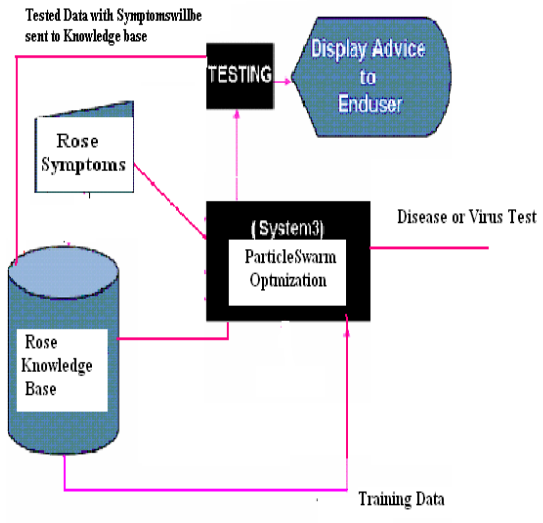


Fig.4 Architecture of subsystem –III



(Particle Swarm Optimization Algorithm)

3.1. Rule Based System
(System-1)

In the Rule Based System the System takes the Symptoms as Input and produce the Exact Disease with all the facts and Rules that matches with in the Knowledge base. This Rule Based System Consists of Knowledge Base, Inference Engine, User Interface, Expert and the User.

Algorithm:

- Repeat
- Collect the *rules* whose conditions match facts in Working Memory.
- If more than one rule matches
- Use *conflict resolution strategy* to eliminate all but one
- Do actions indicated by the rules
- (Add facts to Working Memory or delete facts from Working Memory)
- Until problem is solved or no condition match

The output of the this system produce the exact disease basing on the symptoms produced by the user which leads to a disadvantage that if any of the symptom does not match with the knowledge it will not produce any output for the further proceedings.

If the system1 (Rule Based System) unable to produce the exact disease then the system2 starts performing its work.

1.2. Optimization Algorithm: -
(SYSTEM-II)

Algorithm:

Optimization (Computation Vector, Memory Matrix, Resultant vector)

Computation Vector is the input vector that has to be mapped with the Memory Matrix and produce the result in the Resultant vector.

Computation Vector is an Input Boolean string of length n. and Memory Matrix is generated from the Knowledge base with all Boolean value of order m X n.

Step 1: Read the Computation vector as Boolean String of n length.

Step 2: Create a Resultant vector that initialize to null and counter to zero

Step 3: Construct the Memory Matrix of Boolean Value from the Knowledgebase.

Step 4: for i in 0 to m in the Memory Matrix of m X n
Make counter as zero. (Perform Step 4a and step 4b m times)

Step 4a: for j in 0 to n element in the row
Compare the jth element of the Computation vector with the jth element of the row
if both the element are equal
counter++ (increment the counter).

Step 4b: Assign the ith element of the Resultant vector with the counter.

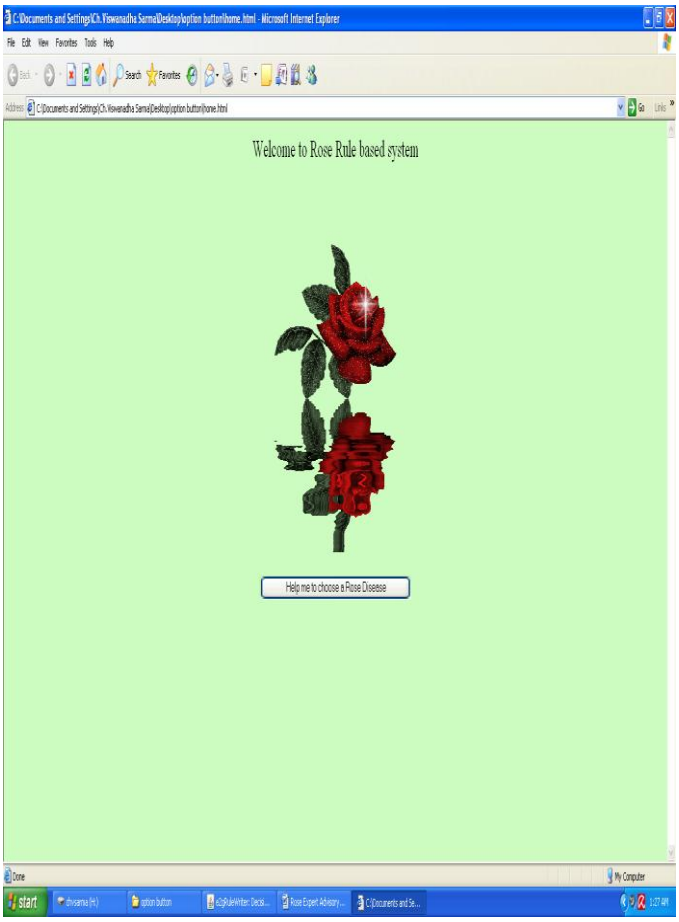
Step 5: (The resultant vector will be within the value of count which is has to be converted into
probability value of percentage value).

Step 6: For each element in the Resultant Vector replace the value by value/m*100 for Percentage
value of replace the value by value/m*1 by Probability value.

Step 7: End the process

1.3. Particle Swarm Optimization Algorithm:

For each particle
Initialize particle with feasible random number
END
Do
For each particle
Calculate the fitness value
If the fitness value is better than the best fitness value (pbest) in history
Set current value as the new pbest
End
Choose the particle with the best fitness value of all the particles as the gbest
For each particle
Calculate particle velocity according to velocity update equation
Update particle position according to position update equation
End
While maximum iterations or minimum error criteria is not attained.



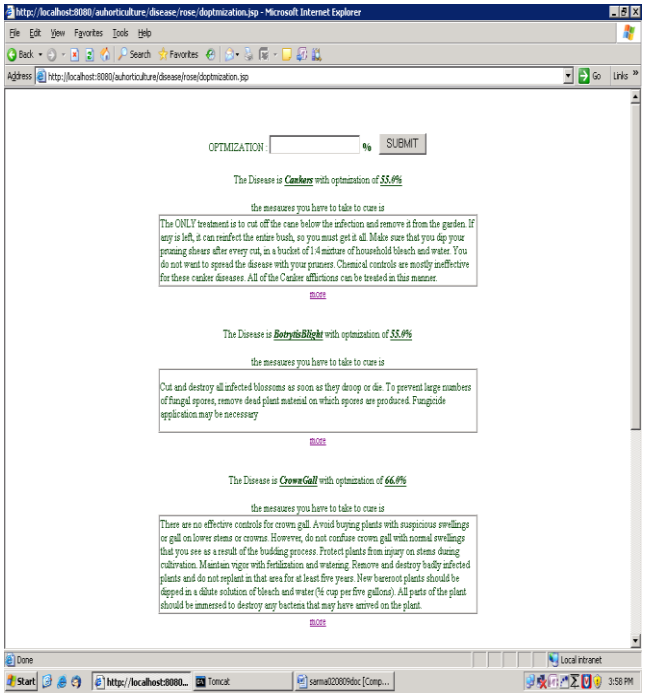
Disease

1. BotrytisBlight
2. Cankers
3. CrownGall
4. BlackSpot
5. PowderMildew
6. BotrytisBlight

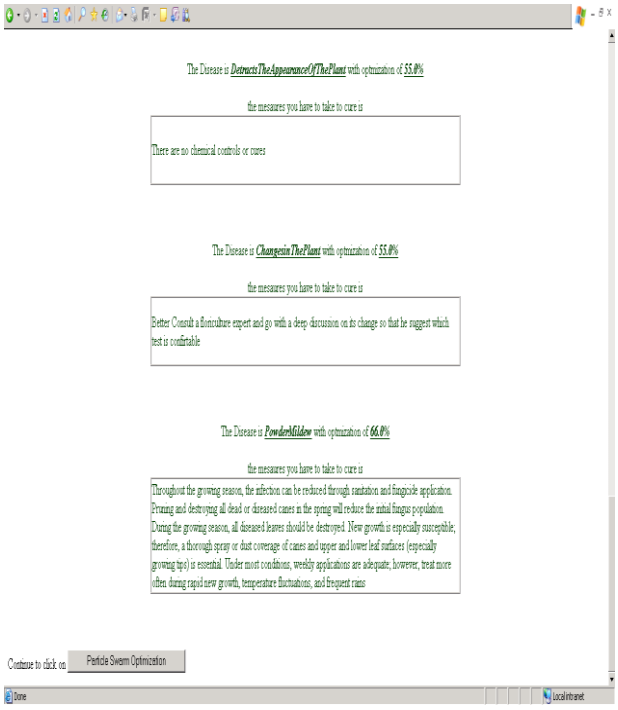
2. Results & Reports

4.1 Rule Based System Screen Shots:

4.2. Optimization Algorithm Screen shots:



4.3. Particle Swarm Optimization Screen Shots:



5. Future Work

In Rose Expert System we have impletemented three algorithms to identify the Diseases and viruses of the rose plant and also give the advices to the farmer about the prevention of that particular disease. Those algorithms are

1. Rule based Algorithm
2. Optimization Algorithm
3. Particle Swarm Optimiation

In future we want to enhance the system in such a way the it has to perform multi language (Translators) , animations , video's for the userinterface design of the system and markerting statistics in the system.

Research issues in agricultural expert systems are categorized under these topics: integration of software components with agricultural expert systems, knowledge sharing and reuse, intelligent retrieval of agricultural data, and automatic knowledge acquisition. The future trends in research and development of agricultural expert systems are expected to be using agent based approaches to solve the integration problem of different software components, developing domain specific tasks that will contribute to knowledge sharing and reuse and automatic knowledge acquisition.

6. Conclusion

The project "Rose Expert Advisory System" is a web-enabled application developed using java server pages (jsp) and MySql database is used as backend. So as to ensure the quality of the software, all software engineering concepts, including test cases are implemented.

Its main emphasis is to have a well designed interface for giving advices and suggestions in the area of horticulture (Rose) field by providing facilities like dynamic interaction between expert system and the user without the need of expert at all times.

By the thorough interaction with the users and beneficiaries the functionality of the System can be extended further to many more areas in and around the world.

7. Acknowledgements

Thank you to every one

References

1. TIMOTHY C. LETHBRIDGE ROBERT LAGANIERE*:*
Object-Oriented Software Engineering,,TMH Publications.

2. **ANN NAVARRO TODD STAUFFER**
HTML by Example, PHI Publications

3. **TIMOTHY BRIGGS, SAMEER TYAGI**
Professional JSP - Wrox Press, Shroff Publishers

4. **C.J. DATE**
An Introduction to Database System

5. Rose crop information collected from professors, **Acharya. N.G.Ranga, Agricultural University.**

6. **Lavandowska**, **L."**JSP Architectures: An explanation and comparison of the methodologiesCommonly known as 'Model I' and 'Model II',"

Global eXtreme Programming, a Software Engineering Framework for Distributed Agile Software Development

Ridi Ferdiana[1], Lukito Edi Nugroho[1], Paulus Insap Santoso[1], Ahmad Ashari[2]

[1]Department of Electrical Engineering and Information Technology, Gadjah Mada University
[2]Department of Computer Science and Electronics, Gadjah Mada University
Yogyakarta - Indonesia
ridi@te.gadjahmada.edu, lukito@mti.ugm.ac.id, insap@te.gadjahmada.edu, ashari@ugm.ac.id

***Abstract*:** Distributed software development or multi-site development may increase the development life cycle. This happens predominantly due to cross-sites communication and coordination difficulties, which have made software development an even more challenging task. The need of a simple and formal framework has been discussed through the concept of Global Software Development (GSD). However, the redundancy of development effort makes GSD need to be streamlined. This paper will make an effort to create a conceptual framework based on the well-known agile method called eXtreme Programming with the existing GSD process. The framework is called GXP, and it provides a formal framework that dedicated for distributed software development.

***Keywords*:** Distributed Software Development, Global Software Development, eXtreme Programming, Software Engineering Framework.

1. Introduction

Distributed software development comes with several types and level [7]. It can be global (different places, different organization), inter-organization (different places, same organization) or open source project development (different places, no organization). Holmstrom et al. [8] refines the types of distribution models, which are geographical, organizational, temporal, and stakeholder's boundaries. Those types are escalated through two main entities, which are people and its artifacts.

People challenges in distributed software development are caused by the dispersion of people among several locations. This is related to communication and cooperation between people. For example, the physical distance between people limited in formal communication. This might lead to a lack of information in a project. Extended communication effort, on the other hand, can lead to information overhead or too much discussion rather than develop the software. Thus, the software project should aim at a trade-off between lack information and information overhead.

Communication problems arise because several factors, which are social and cultural differences between distant sites [5], time zone separation [9], perceives distance within members of the given stakeholder group (Evaristo and Scudder, 2000), and different motivation background [7]. Those factors addressed by providing communication tools like Computer-supportive and Collaborative Work (CSCW) [6]. CSCW has been discussed long time to enhance communication in the distributed team. The issues within CSCW implementation are about the learning curve, the amount of such work is increasing, and the unusual way to communicate.

Artifact challenges are caused the need of task distribution, the level of synchronization, decision making, skills and knowledge of each member. Although it is not related directly with the people, this challenge should be answered through the organizational and software engineering process.

The rest of paper is organized as follows. First, we discuss the existing solution in the multi-site development. Secondly, we describe our research approach to synthesize the formal framework called GXP. The research then reports the result by a discussion of the implication of those results, limitation of the work and future research directions.

2. Current Research Solution

Distributed Software Development, Collaborative Software Development (CSD), and Global Software Development (GSD) process is related processes, which make an effort to manage artifacts, people, and product through software engineering disciplines for multi-site software development.

Distributed Software Development (DSD), Collaborative Software Development (CSD), and Global Software Development (GSD) are termed that interchangeably used to describe a software engineering process solution to overcome software engineering limitation in the distributed development model. Although those terms are used interchangeably, those terms have a different point of view to solve the problem.

DSD is the generic term which is used to describe management, development, and maintenance of software that being geographically distributed across the globe [13]. DSD research focuses in non-technical issues that related with distributed software development like coordination, awareness, and dependency management. DSD provides a problem-solution model that captured from field reports and adapted to the other's problem which has same context. For example, manufacturing organization is creating their production monitoring software through distributed software development model. The organization creates patterns and practices from their experience. Those patterns and practices afterward are adopted by different organization in the

different country to create the similar software.

CSD is another term that describes a set of tools that strengthen collaboration in software development [13]. CSD researches to provide several alternative tools that help collaboration in distributed software development. CSD is created based on CSCW concept, which is solving the distant and coordination problem through tools, the difference is that CSD provides a specific tool for distributed software development.

GSD is a contemporary form of software development undertaken in globally distributed locations and facilitated by advanced information and communication technology (ICT), with the predominant aim of rationalizing the development process [12]. GSD offers theoretical process to handle distributed software development. As a software engineering process, GSD offers planning strategy, organization structure, and progress control and monitoring. Rather than others approaches or terms, GSD provides more sufficient process and workflow in the software engineering framework.

Many of the GSD implementations are happened in an organization that has a software project in enterprise level. Company like Lucent, Microsoft, Philips, and Siemens is a small sample that done GSD for their software products. Nowadays, GSD is also happening in personal, small scale (1-6 people), and community software development. For example, a person can get a software development project from a freelance website, small group can get a client from different region or countries through an internet project bidding and community can build software like open source software through Sourceforge, or Codeplex system. Those opportunities give a clear view that GSD needs to be simplified.

This research has been motivating factors to deliver a simple approach in distributed software development. Simplification and effectiveness are the legacy problem for every software development. Therefore, many research focus in simplification and effectiveness. Agile process is one of the software engineering processes, which are dedicated to simplify the process and give a center of attention in delivering working software. Agile community through theirs manifesto promises a simple and standard way to build a software. However, agile process is fitting in collocated software development since the process extremely needs direct interaction without a distant.

Based on those hypotheses the research sees an opportunity to integrate the existing GSD process with the agile method. In a specific view, the research will choose one of the agile methods called eXtreme Programming (XP). XP is a lightweight methodology for small-to-medium-sized teams developing software in the face of vague or rapidly changing requirements [3]. During its execution XP gets optimistic feedback to implement in personal software development [1], and also enterprise development [4]. Those researches provide factual information that XP has been sufficient in scalability aspect. Therefore, XP is chosen as an agile process in this research and will be integrated with the GSD in this research through a formal framework design process.

3. Framework Design

3.1 Software Engineering Framework

The high level of the framework describes the entities, building block and its relation. A framework is reusable design that requires components to functions. To create a framework, a researcher should provide the components required by the framework. In order to do this effectively, the framework-component interfaces must be specified so the researcher knows what expectations the framework makes about the component, and so the components can be verified against these assumptions. The framework itself can be designed to several points of view such as technical function, software engineering, and domain-specifics process.

A component is a software unit (for example, example module, set of function, or a class) or data unit that has a defined interface for which the component provides an instantiation. As framework entities, component should be easy to understand through its interface. To do so, component in a framework can be anything includes the data or non-software component.

To understand and use a framework, the framework must be specified, the engineer must understand what framework does, what components must be provided to instantiate the framework, and how to use the framework. In order to discover this information, the research defines the framework specification that includes three tasks.

1. Specify the syntax of the framework
2. Specify the semantics of the framework
3. Specify the framework component-interfaces.

The syntax of the framework specifies how the engineer uses the framework, that is, how the framework becomes a system of a part of the engineering system. For example, if a framework provides a test driven development function, the engineer knows how the function is executed. The semantics of the framework specifies what functionality it provides. The framework-component interfaces define the syntax of the components.

The result of the framework creation from the followed processes is called as Global eXtreme Programming framework (GXP). The research result is said as an unadjusted framework. Unadjusted term means that the framework needs several details, implementation, and assessments.

3.2 GXP Framework syntax

GXP framework syntax states how the distributed software development is executed. The framework has three main operations that are processed, method and tools operation. The process is the first operation that is executed to define an asset value and principles of the software development in GXP model. After the team understood the process, the team can learn the comprehensive daily execution of GSD through method operations. This operation makes the team understood the daily how-to and technical practices of the development execution. The last syntax of the framework is a tool. The tool gives productive understanding and real experience execution through the supporting tool. Figure 1 describes the three framework syntax in a sequential block

diagram.

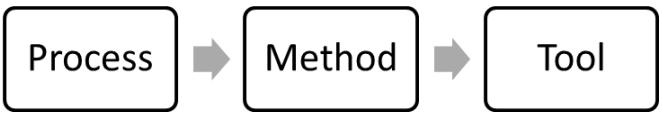


Figure 1. Framework syntax block diagram

GXP Process describes base framework activities that exist in global software development. Software development life cycles, software estimation, organization structure, and quality assurances are components that described in GSD process. GXP process works as umbrella activities for relating building block like GXP method or GXP tools.

GXP method describes the practical "how-to" in order to build software in the distributed model. GXP method includes software development practices, value, project artifacts, and project management practices. GXP method works as a guidance or step by step execution to solve distributed software development.

The last but not least is GXP tools. It is designed to give productivity tools for distributed software development. It consists of infrastructure requirements supports, communications tools, and collaboration software. The tools work as cross cutting building blocks that support GXP process and GXP method.

3.3 GXP Framework semantics

GXP semantics explicitly provides the component that contains within the operation. This step explicitly identifies components and its functionality. The aggregation of the result provides what the framework does and decision information for a software engineers to do properness and suitability of the framework based on their need.

Process semantics contains several components that related with the groundwork activities before the project is executed. Therefore, the process component displays several abstract components that related with the process. The research identifies the process semantics are preparation activities that need to be done before the project executed, the lists come up as follows.

1. Quality controls focus, since the software engineering purposes are to deliver good quality software. The quality controls focus dedicates several principles to cover good characteristic software in GXP framework. This component should be a first component to comprehend by the team.
2. Software development life cycle, this component acknowledges the SDLC of GXP. The SDLC describes the phases of the project that need to covers in GXP.
3. Software estimation, this component will allow the engineer to estimate the complexity the software by looking at the technical and experience factor. The early estimation will give the engineer better understanding about feasibility of the project based on the budget, time, and resources.
4. Team organization, this component is the last component that needs to learn in order create the jell team in GXP. Several roles are introduced and job descriptions are described.

Method semantics describes the technical "how-to" in terms of the project execution. This semantics only executed if the project is agreed in resources, budget, and time. Several components are identified as follows.

1. GXP values and principles, this component describes the GXP values and principles. This component covers what the team needs to understand the framework mindset.
2. GXP practices, this component describes the practical action which is believed to be more effective at delivering a particular outcome. GXP Practices also be defined as the most efficient (least amount of effort) and effective (best results) way of accomplishing a task.
3. GXP artifacts, this component describes any kind of tangible product that produced during the development of software. GXP artifacts consist of several document, template, and knowledge base to execute GXP software development.
4. GXP project management, this component describes the discipline of planning, organizing, and resource management to bring about the successful completion of specific project goals and objectives based on GXP point of view.

Tool semantics described the support of the software and infrastructure to improve the development productivity. The components are described as follows.

1. The communication Tools. This component discusses several communication tools and sample that appropriate to support GXP framework.
2. Infrastructure support. This component recommends some of the infrastructure specification to develop GXP environment.
3. Collaborative workspace. This component provides proof of concept recommendation by delivering suite tools for GXP framework projects.

Through the three framework syntax, the research proposes eleven components that dedicated to support the framework purposes. Those components will be rearranged in its execution through an interface identification step.

3.4 GXP Framework Interface

GXP framework component interfaces discuss the input and the output of the component that already defined in previous step. Table 1 describes the interface for GXP component.

Table 1. GXP Framework component interface

Component	Input interface	Out Interface
Quality control focus	Market intent and project vision	User stories baseline
SDLC	Project scope and time	Schedule plan
Software Estimation	User stories, technical factor, and experience factor	Software complexity, user story point, and effort rate
Team Organization	Effort rate, software complexity	Team profile
Values and principles	Team profile	Team mindset
Practices	Team mindset	Action plan
Artifacts	Schedule plan, action plan	Project artifact
Project Management	Schedule plan, action plan	Tracking execution
Communication Tool	Crosscutting interface	Crosscutting interface
Infrastructure Support	Crosscutting interface	Crosscutting interface

Collaborative workspace	Crosscutting interface	Crosscutting interface

Input interface is an input for the component, after the process of the component the output interface delivers the output channel for the process result. The process result can be processed further through another component. Several components also work as crosscutting interface, crosscutting interfaces describes a component that communicates intensively as a support component.

4. GXP framework prototype

4.1 Instantiation of the GXP Framework

Instantiation of the framework describes a step to pull the component into high level architecture. This step consists of two tasks, which are specifying the system properties and providing concrete components. The system properties define the system that implements when the framework is instantiated with all the components it needs. One of the complexities in framework use ensuring that the framework is applied to areas for which it is suitable. Stating system properties also allows for verification tasks to be undertaken ensuring that the framework with the instantiated component satisfies these properties. The second task is to provide concrete components that are instantiations of the components specified as part of framework specification. It can be concreted as file, document, or even the codes.

GXP framework properties cover several key points such as follows.

1. The input of the framework is an initiative to execute the software project distributed.
2. The output of the framework is an effective approach to manage and track the distributed project.
3. The process of the framework follows the XP phases which are exploration, planning, iteration, production, and maintenance.
4. The feedback mechanism of the framework which are executed through user acceptances test and production release feedback.
5. The environment of the framework is an environment where the ICT infrastructure like broadband the internet exists.

The concrete of the framework can be described through artifacts and the tools that support the framework. The artifact will exist in every phase of the project as main deliverables of the framework execution and tracking. The tool works as artifacts placeholder to manage, collaborate, and track the artifact.

4.2 GXP in the big picture

In a big picture, GXP provides overall building block of the framework architecture. The big picture covers the component, the dependency between components, and the semantics of the framework. As mentioned before, there are three building blocks, which are processing building block, method building block, and tool building block.

Process building block discusses about the starting point where the GXP should be started by the team to follow. It's provided several abstract values and principles that can be done by the team before the project starts and when the project is executed. The technical how to of the process is described in method building blocks, method building block discusses the implementation of the process through several values, principles, practices, and artifacts. Both process and method are supported by the existence of the tools as crosscutting layers that can communicate in the term process or method.

The proposed framework semantics and components provide probabilities to use GXP framework in several models of implementation. For the people who want to implement the GXP, it is recommended to start by seeing the project condition. GXP proposes three project conditions, which are remote model execution, virtual team model, and distributed team model.

Remote execution model is based on the situation, when the distributed context is only happened between team and the client. This execution model usually happens in a small scale project and less urgency. The team is on the same place while the client is separated by distance. The characteristic of this execution model is.

1. There is no management difference since the team is still one site. The working process can be like XP team with the small tweak from GXP framework in terms of tools and method.
2. Few number of the team members with the small complexity inside the project. The kind of projects that worked in this execution model most likely small in dependency and urgency for the client.
3. Project length between one to three months with small iteration for two weeks or less.

Virtual execution model is based on the situation where the distributed context is happen between the client and inside the team. The difference between the remote model executions is the location the team that also separated. The portion of the team is onsite with the client while the rest is separated. The others characteristics of virtual team model described as follows.

1. Management treats the team as a single virtual team. The team will have one single management but separated by distances.
2. Small to the medium numbers of the team with the medium complexity inside the project. The project typically is related with the core business for the client
3. The project length might be in three to twelve month with small medium iteration length between one into two months.

Global execution model is based on the situation where the distributed context is happening globally. The main characteristic of this execution model is huge numbers of team member's. The member might be different time zone and culture. The main characteristic of this execution model as follows.

1. Management divides clearly between central management and site management as separated instances.
2. Enterprise scale application with medium to large team numbers. The project typically is related with a core product that will be sold in the global market with different language, need, and culture.

3. The project length might be multiyear with medium to long iteration model between bimonthly into quarterly iteration.

Several execution models will make the team behave differently in terms of working model and initial configuration. Table 2 provides information about the recommended configuration for the team which wants to choose GXP as a distributed software development framework.

Table 2. GXP classification and initial recommendation

Recommendation	Remote execution model	Virtual execution model	Global execution model
Milestone length	1-3 months	3-12 months	12 months or more
Iteration length	1-2 weeks	4-8 weeks	8-16 weeks
Team model[12]	Hub-to-Spoke	Hub-to-Hub	Mesh (fully distributed)
Project management	Onsite	Offshore and centralized	Offshore and distributed

4.3 Using GXP Framework

Framework is a baseline for the team that wants to extend customized distributed development implementation without losing the essential of the framework component that need to be defined. Using GXP framework can be done through seven simple steps. These steps can be followed by the development team which wants to implement the framework for their distributed development need. The seven checklists are described as follows.

1. Verifying GXP Framework and project appropriates.
2. Choosing GXP classification.
3. Composing GXP team.
4. Learning GXP values, principles, practices, and artifacts.
5. Estimating the project.
6. Preparing the infrastructure and tools.
7. Executing and monitoring GXP through its SDLC.

GXP is not a silver bullet, there are several constraints where the framework becomes ineffective based on several conditions and constraints. The team should investigate the suitable project based on the discussion of the team. The first step is choosing the appropriateness between the project and the framework. Based on the GSD and XP concepts, GXP frameworks work appropriate when the several conditional.

1. The project at least partially distributed. Adopting GXP in the onsite project will make it work but not as efficient as just like adopting XP in the onsite project.
2. The project is not real time project or crucial project. Building software for the earth quake monitoring or other extensive resources project will not be appropriate with GXP.
3. The business culture of the client support working distributed. Some clients have a culture to appreciate the development team always onsite when this situation happens; the GXP framework is not effectively useful.
4. The business culture support continuous communication and improvement. Single meeting specification or the business of the offshore client make the GXP project won't work best when they do less communication.
5. The business culture support less document work but running software instead. GXP is dedicated to work in balance between document and software.

If the project is suitable based on those checklists, the team can start to choose the GXP as a framework and select the appropriate model of GXP such as remote, virtual, or global execution. Choosing the GXP model will give the team the initial configuration and recommendation of the project.

The initial configuration will help the team to choose the proper team scale and initial planning. GXP encourages the team follows XP roles such as the coach, developer, tracker, project manager, tester, and domain expert. These roles will be composed in a central team or a site team. Central team is a team that manages one or more site teams. In the small team, central team can be a team that has the management role for the overall team.

Just like in XP, team organization and structure creation is created several phases in the software development life cycle (SDLC). As the iteration phase begins, both site team and central team should be established and ready to build test and codes. Central team works as a monitoring team for the site team during the iteration, production, and maintenance phase. Ideally, both central team and site team has the same team composition. For example, central team can only have a management members like a project manager and domain expert, and site team has the rest technical team. There are no rules of thumb the team composition, since the situation is mostly driven by the resources.

After the team is composed, the team can start the learning curves by understanding the values, principles, practices, and the related GXP artifacts. Values, principles, and practices are lent from the XP methodology. However, several practices like pair programming and standup meeting is replaced with the equal interaction like online interaction through instant messaging or video conferences. In this step, it will be the good idea for the team to execute short term coaching workshops that learn principles, practices, and values.

The team can do estimating Estimation defines the approach to do better planning. It is executed in exploration phase. The result of the estimation technique is a quantitative result that works as a baseline for the software complexity measurement, user stories points, and effort rate. GXP estimation technique will answer how big or small is the project in a quantitative model.

The estimation result can make the team have clear visibility how complex the project, and how long it will take. After the estimation activities, team can implement the infrastructures that needed by the team. It means every communications and collaboration software should be installed, configured, and ready to use by the team. This step is including, training and testing the infrastructures between of sites.

After the infrastructures ready, the team can start the full of software development lifecycles. There are four phases in GXP SDLC. Figure 2 shows the GXP SDLC. It defines the overall lifecycles based on the well-known eXtreme Programming development life cycle. XP method adopts exploration, planning, iteration, production, and maintenance as a phase in a development cycle.

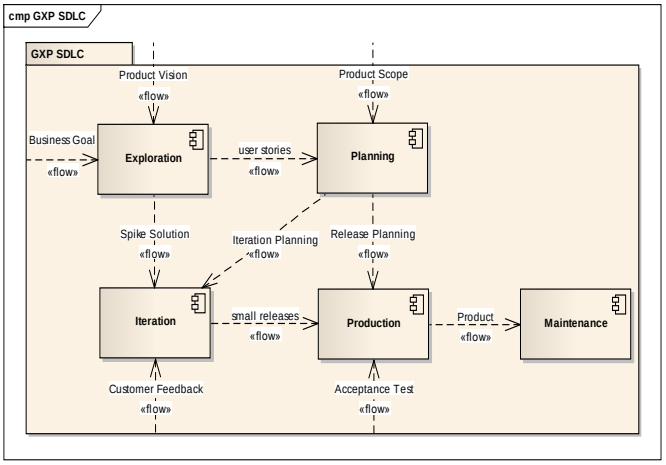


Figure 2. GXP Software Development lifecycles

Every phase has several inputs and the output results of the phase. The execution of each phase is not really sequential and one at a time. Except the exploration phase, the rest phases on the GXP SDLC are iterative and incremental. Table 3 describes the input and the output of every phase and also shows the workflow of the software development life cycle.

Table 3. GXP SDLC input and output process

Phase	Input	Output
Exploration	Product Vision	User Stories
	Business Goal	Spike Solution
Planning	Product Scope	Iteration Planning
	User Stories	Release Planning
Iteration	Spike Solution	Small Releases
	Iteration Planning	
	Customer Feedback	
Production	Small Releases	Product
	Release planning	
	Acceptances Test	
Maintenance	Product	Fixed Product

Based on the phases, the team should also prepare the artifacts for each phase. The artifacts are a useful documentation that provides a tracking history for the product evaluation that developed by the team. Unlike XP that mostly depends on the codes and comments, GXP should prepare the artifacts as a tool to exchange the knowledge between sites. Furthermore, the team should create just enough document and others resources so that the artifacts are not redundancy in terms of size and numbers.

5. Conclusion and Future Work

In the multi-site software development, software process like CSD, DSD, and GSD have begun to introduce.

The paper limits the discussion as a framework that can be fulfilled later in the specific aspect like values, principles, practices, and implementation. As a further work, the framework component should be detailed with the how an aspect that gives a framework user a detailed action that needs to be fulfilled. On the other's side, the framework also needs to be evaluated through several assessments and case studies implementation.

References

[1] Agarwal, R. and Umphress, D. 2008. Extreme programming for a single person team. In Proceedings of the 46th Annual Southeast Regional Conference on XX (Auburn, Alabama, March 28 - 29, 2008). ACM-SE 46. ACM, New York, NY, 82-87.

[2] Bass, M. 2006. Monitoring GSD projects via shared mental models: a suggested approach. In Proceedings of the 2006 international Workshop on Global Software Development for the Practitioner (Shanghai, China, May 23 - 23, 2006). GSD '06. ACM, New York, NY, 34-37.

[3] Beck, K. 1999. Extreme Programming Explained. Addison-Wesley.

[4] Cao, L., Mohan, K., Xu, P., and Ramesh, B. 2004. How Extreme Does Extreme Programming Have to Be? Adapting XP Practices to Large-Scale Projects. In Proceedings of the Proceedings of the 37th Annual Hawaii international Conference on System Sciences (Hicss'04) - Track 3 - Volume 3 (January 05 - 08, 2004). HICSS. IEEE Computer Society, Washington, DC, 30083.3.

[5] Evaristo, J. R. and Scudder, R. 2000. Geographically Distributed Project Teams: A Dimensional Analysis. In Proceedings of the 33rd Hawaii international Conference on System Sciences-Volume 7 - Volume 7 (January 04 - 07, 2000). HICSS. IEEE Computer Society, Washington, DC, 7052.

[6] Grudin, J. 1994. Computer-Supported Cooperative Work: History and Focus. Computer 27, 5 (May. 1994), 19-26.

[7] Gumm, D. C. 2006. Distribution Dimensions in Software Development Projects: A Taxonomy. IEEE Software. 23, 5 (Sep. 2006), 45-51.

[8] Holmstrom, H., Conchuir, E. O., Agerfalk, P. J., and Fitzgerald, B. 2006. Global Software Development Challenges: A Case Study on Temporal, Geographical and Socio-Cultural Distance. In Proceedings of the IEEE international Conference on Global Software Engineering (October 16 - 19, 2006). ICGSE. IEEE Computer Society, Washington, DC, 3-11.

[9] Lanubile, F., Damian, D., and Oppenheimer, H. L. 2003. Global software development: technical, organizational, and social challenges. SIGSOFT Softw. Eng. Notes 28, 6 (Nov. 2003), 2-2.

[10] Mockus, A. and Herbsleb, J. 2001. Challenges of Global Software Development. In Proceedings of the 7th international Symposium on Software Metrics (April 04 - 06, 2001). METRICS. IEEE Computer Society, Washington, DC, 182.

[11] Pilatti, L., Audy, J. L., and Prikladnicki, R. 2006. Software configuration management over a global software development environment: lessons learned from a case study. In Proceedings of the 2006 international Workshop on Global Software Development For the Practitioner (Shanghai, China, May 23 - 23, 2006). GSD '06. ACM, New York, NY, 45-50.

[12] Sangwan, R., Bass, M., Mullick, N., Paulish, D. J., and Kazmeier, J. 2007. Global Software Development Handbook (Auerbach Series on Applied Software Engineering Series). Auerbach Publications.

[13] Sengupta, B., Chandra, S., and Sinha, V. 2006. A research agenda for distributed software development. In Proceedings of the 28th international Conference on Software Engineering (Shanghai, China, May 20 - 28, 2006). ICSE '06. ACM, New York, NY, 731-740.

Author Biographies

Ridi Ferdiana. Mr. Ridi Ferdiana was born in 1983. He is a doctoral student at Gadjah Mada University, Yogyakarta since 2008. He earned his master degree from the same university in 2006. In his professional area, he holds several professional certifications such as MCP, MCTS, MCPD, MCITP, MVP and MCT. In his daily research activities he really enjoys to learn about software engineering, business platform collaboration, and programming optimization.

Lukito Edi Nugroho. Born in 1966, Dr. Lukito Edi Nugroho is an Associate Professor in the Department of Electrical Engineering and Information Technology, Gadjah Mada University. He obtained his M.Sc and PhD degrees from James Cook University in 1995 and Monash University in 2002, respectively. His areas of interest include software engineering, distributed and mobile computing, and application of ICT in education.

Paulus Insap Santosa. Insap was born in Klaten, 8 January 1961. He obtained his undergraduate degree from Universitas Gadjah Mada in 1984, master degree from University of Colorado at Boulder in 1991, and doctorate degree from National University of Singapore in 2006. His research interest including Human Computer Interaction and Technology in Education.

Ahmad Ashari Place and date of birth: Surabaya, May 2nd 1963. Get Bachelor's degree 1988 in Electronics and Instrumentation, Physics department Gadjah Mada University, Yogyakarta. Master degree 1992 in Computer Science, University of Indonesia, Jakarta Doctor Degrees 2001 in Informatics, Vienna University of Technology. Major Field of study is distributed system, Internet, Web Services, and Semantic Web.

The Metaplectic Sampling of Quantum Engineering

Walter J. Schempp

Lehrstuhl für Mathematik I, University of Siegen, 57068 Siegen, Germany

Abstract. Due to photonic visualization, quantum physics is not restricted to the microworld. Starting off with synthetic aperture radar, the paper provides a unified approach to clinical magnetic resonance tomography and the bacterial protein dynamics of structural microbiology. Its mathematical base is harmonic analysis on the three–dimensional Heisenberg Lie group with associated nilpotent Heisenberg algebra Lie(N).

Keywords: Heisenberg Lie group, Heisenberg nilpotent Lie algebra, metaplectic representation, data sampling, synthetic aperture radar imaging, magnetic resonance tomography, bacterial protein dynamics, flagellar motor, rotational switching, tracking of chemotaxis

1. THE HEISENBERG LIE GROUP

Consider a particle in a one–dimensional space. Its quantum mechanical description is obtained by introducing variables suggested by the classical theory in its Hamiltonian form. These variables which are *not* numbers and do not commute in general, form an algebraic structure generated by two basic self–adjoint elements P and Q satisfying the Heisenberg commutation relation

$$\boxed{[P,Q]=-i}$$

where the bracket denotes the commutator. The elements P and Q are supposed to be represented by self–adjoint operators in Hilbert space. These operators are unbounded and this might involve domain problems which Werner Karl Heisenberg (1901 to 1976) was unable to appreciate. Heisenberg did not recognize the difference between bounded and unbounded linear operators in infinitely dimensional Hilbert spaces, which is at the origin of distribution theory created by Laurent Schwartz (1915 to 2002). Similarly, Erwin Schrödinger (1887 to 1961) strongly believed that even thought–experiments with just one electron or atom or small molecule invariably entail ridiculous consequences. From this opinion, the advances of present–day quantum physics and quantum engineering can be easily recognized.

The three–dimensional Heisenberg Lie group N with one–dimensional center implements the Heisenberg commutation relation. As a central extension of the symplectic plane $\mathbf{R}\oplus\mathbf{R}$ it is isomorphic to the direct sum $\mathbf{C}\oplus\mathbf{R}$ under the multiplication law

$$\boxed{(w_1,z_1)(w_2,z_2)=\left(w_1+w_2,z_1+z_2+\frac{1}{2}\Im(\bar{w}_1w_2)\right)}$$

for elements $(w,z)\in\mathbf{C}\oplus\mathbf{R}$. The fact that the one–dimensional center $\{(0,z)|z\in\mathbf{R}\}$ of N is non–trivial implies dramatic deviations from commutative Fourier analysis of signal analysis and communication theory. Roughly speaking, the Fourier transform and the Fourier cotransform on the real line $\mathbf{R}$ have to be replaced by equivalence classes of irreducible unitary linear representations of N and their contragredient siblings. These equivalence classes can be visualized by inhomogeneous planes transversal to the central axis $\{(0,z)|z\in\mathbf{R}\}$, and symmetrically organized with respect to the origin so that they are tangent to the two–dimensional Bloch sphere $\mathbf{S}_2$ of nuclear magnetic resonance spectroscopy. The contact points of the Bloch sphere $\mathbf{S}_2=\mathrm{SO}(3,\mathbf{R})/\mathrm{SO}(2,\mathbf{R})$ allow to coherently control the nuclear spin choreography: Its north pole represents the spin–up state $\uparrow$, its south pole the spin–down state $\downarrow$.

The symmetries associated with the so–called coadjoint orbit picture of N in the vector space dual of its Lie algebra, the real Heisenberg algebra Lie(N), is two–fold. In the longitudinal direction it is represented by the Galois group $\mathrm{Gal}_{\mathbf{R}}\mathbf{C}$ which acts in binary terms by reflecting the central axis. In the transverse direction it is represented by the metaplectic group $\mathrm{Mp}(1,\mathbf{R})$ which acts as a two–fold cover of the unimodular Lie group $\mathrm{SL}(2,\mathbf{R})$ by letting the one–dimensional center pointwise fixed.

The main mathematical tool for a unified approach to quantum holographically data organizations is the three–dimensional real Heisenberg Lie group N which allows to bridge the gap between abelian and nonabelian harmonic analysis by showing how various important filtering results in abelian harmonic analysis may be enriched through an interpretation in terms of the Heisenberg group N and its nilpotent Lie algebra, the Heisenberg algebra Lie(N). Obviously, it needs some habituation to get familiar with these mathematical concepts. However, the broad range of ensuing results makes it worthwhile to think in the broad category of holographically encoded data organizations.

Unfortunately, harmonic analysis on the Heisenberg

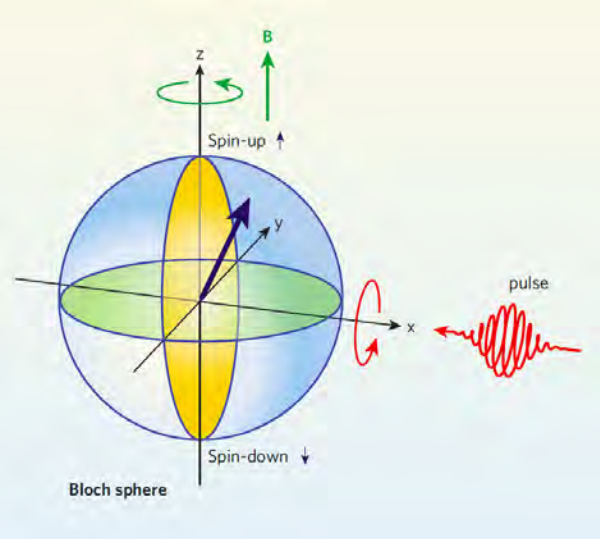


FIGURE 1. Nuclear spin excitation procedure of magnetic resonance tomography. The Bloch sphere $\mathbf{S}_2$ of nuclear magnetic resonance spectroscopy allows to control the nuclear spin choreography in a strong magnetic field B of homogeneous density.

Lie group N is not a very popular research subject today. Why has a mathematical object with applications of such a wide spectrum gone relatively unnoticed until recently? One can only speculate. Hermann Weyl (1885 to 1955), one of the pioneers in introducing the Heisenberg group into quantum physics, overlooked the natural occurrence of the Heisenberg group, exploitation of which yields results which one feels he would have liked very much. An obstacle to the appreciation of the common underlying structure may have been the very diversity of the topics, for detection of its presence in one place need not suggest its presence elsewhere. Indeed, investigators in one field may very well never have been aware that the Heisenberg Lie group had been found in some area not seemingly related to theirs ([8], [9]).

2. SYNTHETIC APERTURE RADAR

In 1903, a mere 15 years following the seminal studies of Heinrich Hertz (1857 to 1894) on the generation, reception, and scattering of electromagnetic waves, the engineer Christan Hülsmeyer (1881 to 1957) demonstrated a ship collision avoidance radar system which he later patented. The range of the first remote sensing system was about 3 km. Although there was a demand for such remote sensing modalities, time was not mature enough to recognize the importance of Hülsmeyer's discovery. It needed several generations of scientists after Hülsmeyer's breakthrough to evaluate the information contents of the sophisticated echo or response signals generated by remote sensing modalities other than radar systems. Due to theoretical insights into the application of the Hanbury Brown–Twiss phenomenon of photon bunching to astrophysics, the link between radar technology and quantum engineering has been established. Important technological advances in varied applications of photonics have made the breadth of this link by far more apparent. They demonstrated that quantum physics and quantum engineering are not restricted to the visualization of the microworld.

Astronomy is considered to be the most ancient of all the sciences. The role of planetary exploration and imaging is quite unusual in the history of science: A research subject that was created almost instantaneously when an esoteric speciality was elevated to prominence during the space race. One of the culminations in the development of remote sensing systems represented the spaceborne Magellan's synthetic aperture radar which has penetrated the Venusian clouds to furnish 100–meter resolution, false–colour images. Once considered the earth's twin because of its similar size and mass, the planet Venus which is completely shrouded in dense clouds is now known as strikingly different. Indeed, a carbon dioxide atmosphere caused a runaway greenhouse effect that today produces scorching temperatures that induced nearly all of the planet's water to escape. Venus' thick clouds are made of sulphuric acid, not water droplets like the earth's clouds. The planet's surface is surprisingly young, and volcanoes are widespread. Lava flooding is thought to have obliterated most surface features only 800 million years ago. The earth's nearest planetary sibling presents an impressive example of how a planet's surface conditions are sensitive to its atmospheric content. It demonstrates that advances in high technology enhanced considerably the range of our knowledge.

Radar systems have been operated since the early 1930s, but with rapidly increasing sophistication of technology. Nearly 60 years have passed since Carl A. Wiley first observed that a side–looking radar system can considerably improve its azimuth resolution by utilizing the Doppler spread of the echo signal. In the ensuing years, a flurry of activity followed, leading toward steady advancement in performance of both the sensor and the echo signal processor technology. Although much of the early work on remote sensing was aimed toward detection and tracking of moving targets, the potential for utilizing this technology as a tomographic imaging modality for scientific applications was widely recognized, and an apex in the development of high resolution imaging modalities has been achieved by synthetic aperture radar systems of airborne or spaceborne platforms. Later on, orbital spin geometry, fast nuclear spin–echo strategies, and photon–echo formation opened new horizons to innovative non–invasive tomographic imaging modalities

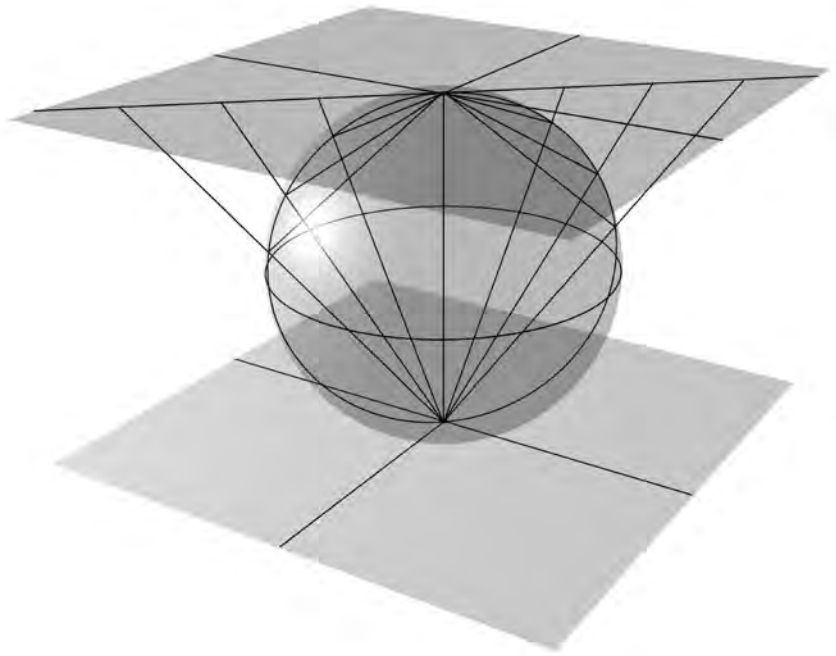


FIGURE 2. The axisymmetric and time asymmetric coadjoint orbit picture of the Heisenberg Lie group N in the three–dimensional real projective space visualizes pairs of non–equivalent symplectic tomographic slices. They are at the origin of the quantum eraser principle of quantum entanglement. The circular boundary of the closed disc in the equatorial plane $\nu = 0$ indicates the points of angular momentum transfer. The symmetry in the longitudinal direction is represented by the Galois group $\mathrm{Gal}_{\mathbf{R}}\mathbf{C}$ whereas in the transversal direction it is represented by the metaplectic group $\mathrm{Mp}(1, \mathbf{R})$.

FIGURE 3. Three–dimensional projective geometry: Rotations induced by the symplectic structure of the planar detector. The exposure time is 10 h.

FIGURE 4. Spaceborne Magellan's synthetic aperture radar observations reveal the planet's Venus landscape by mapping radar mosaics onto a computer–simulated two–dimensional sphere $\mathbf{S}_2$ in the three–dimensional real projective space. The global radar view of the shrouded surface of Venus is centered at 180^{o} east longitude. Remarkably, Galileo concluded "with absolute necessity that Venus revolves about the sun just as do all the other planets", thus consenting to Nicholas Copernicus and Johannes Keppler in discrediting geocentric cosmologies.

which offered a stepwise improved contrast resolution of the echo signals ([1]).

A mathematical cross–correlation analysis of the holographic organizations of synthetic aperture radar data

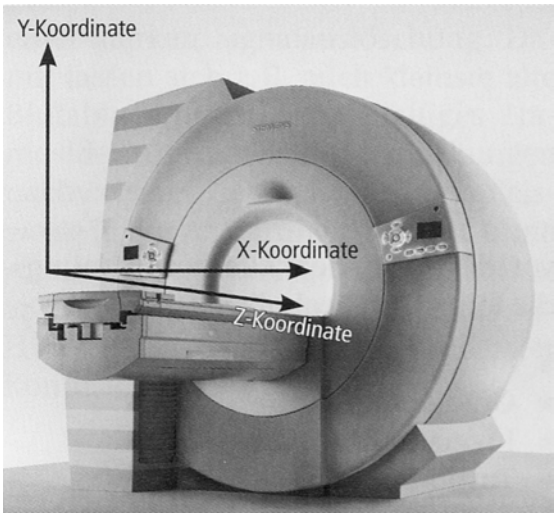


FIGURE 5. Coordinatization procedure in a magnetic resonance tomography scanner by the basis $\{X,Y,Z\}$ of the nilpotent Heisenberg algebra Lie(N), where the vector Z spans the center of Lie(N). The exponential diffeomorphism gives rise to the coordinatization of the three–dimensional real Heisenberg Lie group N by unipotent matrices. The Fourier equivalent polarization planes spanned by $\{X,Z\}$ and $\{Y,Z\}$ in Lie(N) induce the coadjoint orbit picture in the three–dimensioanl real projective space. Modern high end clinical magnetic resonance tomography scanner for routine radiological examinations implement a magnetic field density of 3 Tesla.

leads to harmonic analysis on the Heisenberg group N. The most common synthetic aperture radar mode, the spotlight mode, is operating analogously to the lighthouse model of pulsar astrophysics ([3], [4]). The word pulsar is an acronym based on the phrase "pulsating source of radio emission". The prototypical neutron–star binary pulsar PSR B1913 + 16 at a distance of 21000 light years from the earth cannot be observed by using the visible light even by means of the most sensitive telescopes. The neutron star rotates on its axis 17 times per second. Thus the pulse period is $T = 59$ ms on the time scale. The orbital period of $7\frac{3}{4}$ hours is remarkably short. The pulsar and its companion both follow highly elliptical orbits around their common center of mass. The major axis of the star's orbit is only 6.4 light seconds, and the minor axis is 5 light seconds. Each star moves on its orbit according to the Kepplerian laws of orbital motions. Notice that the eccentric Kepplerian motions are cross–sections of the corotating circular bicylinder passing through the torque ring. It is the image of the projectivized coadjoint orbit picture of the Heisenberg Lie group N under the conformal Möbius inversion.
According to the lighthouse effect, the pulsar's beams of radio waves sweep the earth producing highly regular sequences of radio pulses. The extraordinary stability of the pulse sequences makes pulsars very accurate clocks, rivaling the best atomic clocks on earth, and makes a transversal matched filter bank reconstruction possible. Pulses from the neutron star traverse the interstellar medium before being received at the radio telescope where they are de–dispersed and added to form a mean pulse profile. During the observation, the data regularly receive a time stamp, usually based on a caesium time standard or hydrogen maser at the observatory plus a signal from the global positioning system (GPS) of satellites. The time–of–arrival measurement can be accurately determined by photon cross–correlation of the observed profile with a high signal–to–noise template profile obtained from the linear superposition of many observations at the particular observing frequency label of the filter bank which reconstructs holographically the pulsar's parameters under consideration. The lattice underlying reconstruction procedure leads to the concept of compact Heisenberg nilmanifold which is associated to the Heisenberg Lie group N. It forms a circle bundle over the two–dimensional torus.

3. CLINICAL MAGNETIC RESONANCE TOMOGRAPHY

In clinical magnetic resonance tomography, the symplectic structure of planar cross–sections, called tomographic slices, is used to install the PROPELLER (Periodically Rotated Overlapping ParallEL Lines with Enhanced Reconstruction) strategy of nonuniformly data sampling to mitigate quantum holographically the signal instability present in all multiple–shot fast spin–echo pulse trains. And further, analogously to quantum holographic rotation sensing with a dual atom–interferometer Sagnac gyroscope, the PROPELLER sampling modality of magnetic resonance tomography provides greater immunity of the linear gradient contrast image formation against geometric distortion ([6], [7]).

The clinical utility of diffusion–weighted magnetic resonance tensor imaging is well established ([5], [10]). Few advances in magnetic resonance tomography have had the impact that diffusion–weighted and tensor imaging have had in the evaluation of the human brain. A significant benefit of the PROPELLER sampling modality in high–resolution diffusion–weighted magnetic resonance tensor imaging is its immunity to image warping from eddy currents. The warping phenomenon is a challenge in echo–planar imaging based protocols, since pixels in different regions of the head are warped in a manner that depends on the direction of the diffusion gradient, making image fusion problematic. In the robust PROPELLER sampling modality, tractographic images of fiber pathways corresponding to different diffusion gradient directions are very well registered.

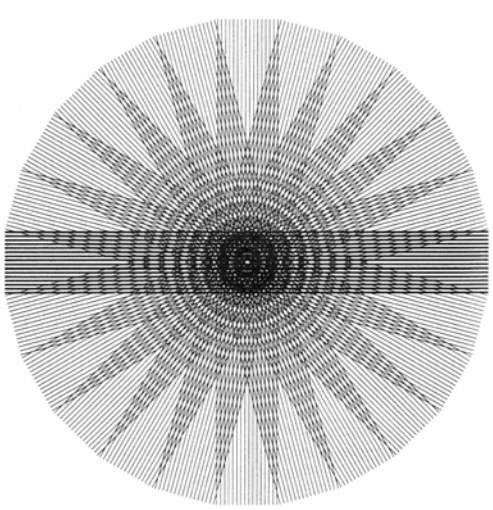


FIGURE 6. Reconstruction procedure of quantum holography for planar signal detectors: The coaxisymmetric data collection trajectory for the metaplectic PROPELLER sampling modality. The bold lines indicate the measured area, called a blade, by one echo train in a fast spin–echo experiment. In subsequent steps, the frequency and phase encode linear gradients are rotated about the tomographic slice selection coaxis. The central core data are resampled for every blade. The data are then combined to the robust formation of a high resolution magnetic resonance tomography contrast image.

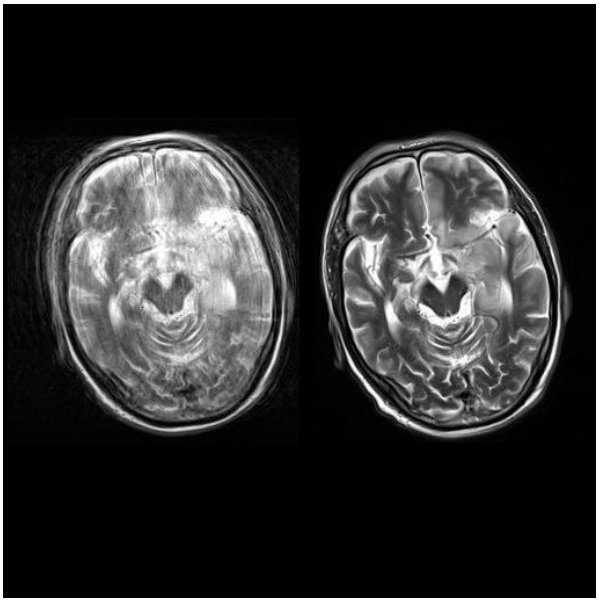


FIGURE 7. Magnetic resonance tomographic image: Transversal slice or axial cross–section of the brain with motion correction. The metaplectic PROPELLER sampling modality reduces dramatically the sensitivity to movement under high resolution magnetic resonance tomographic scanning.

4. BACTERIAL PROTEIN DYNAMICS OF STRUCTURAL MICROBIOLOGY

Bacteria and their genome sequences are evolutionarily and physiologically interesting microbiological objects. Many motile species of bacteria are propelled by flagella, which are rigid helical filaments turned by rotary motors in the cell membrane. The bacterial flagellar motor is a remarkable nanomachine: built from approximately 25 different proteins, it drives the rotation of the helical filaments at speeds of up to 2000 Hz, efficiently propelling bacteria through viscous media. The rotation rates of the flagellar motor had been reported to be, for instance, 170 Hz for *Salmonella typhimurium*, 270 Hz for *Escherichia coli*, and 1700 Hz for *Vibrio alginolyticus*. Indeed, the flagella of *V. alginolyticus* are one of the fastest molecular rotors in biological systems with a pitch of its flagellar helix of 1.58 μm ([2]).

The structure of the bacterial flagellar motor excites considerable interest because of the ordered expression of its genes, its regulated self–assembly, the sophisticated interactions of its proteins, and its startling dynamics. A protein called FliG forms a ring in the spinning rotor of the flagellar motor that is involved in the generation of torque through an interaction with the anion–channel–forming stator proteins MotA and MotB. The latter form a ring of studs within and above the inner membrane that couple the passage of protons across the membrane. A promising application of the sub–Riemannian geometry of the Heisenberg Lie group N is in the bacterial protein dynamics of molecular systems biology. By winding up the Heisenberg helix on the circular bicylinder passing through the torque ring, the central torus singularity and the Galois group $\mathrm{Gal}_{\mathbf{R}}\mathbf{C}$ model in binary terms the molecular structure of the spin switching action on the torque ring in the full–length FliG protein. The Heisenberg helix is a solution of the symplectic Hamilton–Jacobi equations, hence a geodesic trajectory for the sub–Riemannian metric of N.

The diameter of the spinning rotor is about 30 nm. Its flagellas are embedded in the cell surface and constructed by means of about eleven protofilaments of a single protein, *flagellin*. The flagellar motor of peritrichously flagellated bacteria uses the potential energy from an electrochemical gradient of anions across the cytoplasmic membrane to impress angular momentum onto the FliG torque ring. A rapid spin switch from anticlockwise to clockwise rotation determines whether a bacterium runs smoothly forward or tumbles to change its trajectory in response to chemotactic signals. A brief tumble is caused by such a quick reversal of the flagellar motor to clockwise rotation, which generates spin twisting pulses that transforms the left–handed helical path of the filament into a right–handed one. The spin twist can be mod-

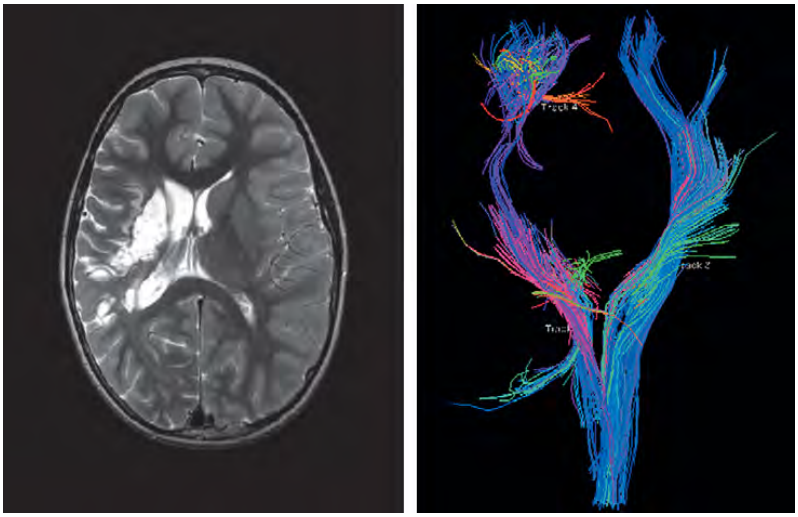


FIGURE 8. Pediatric neuropathology: The morphological magnetic resonance tomographic image demonstrates cystic encephalomalacia and volume loss due to right middle cerebral artey embolic occlusion secondary to traumatic internal carotid artery dissection. The diffusion–weighted magnetic resonance tensor image demonstrates disruption of the major white matter tracts.

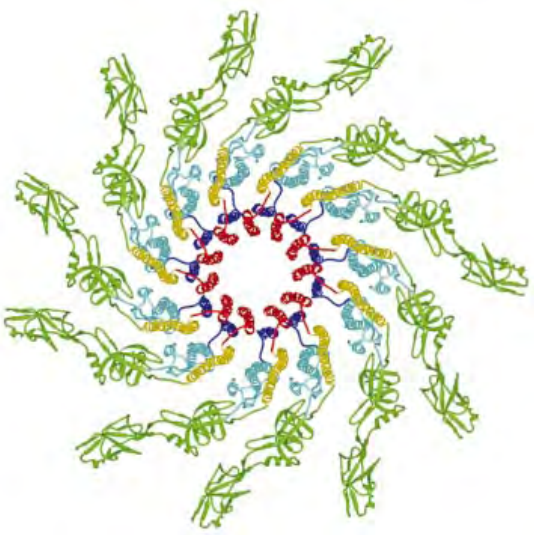


FIGURE 9. Transversal visualization of the flagellin crystal organization: Distal view on the end of the protofilament. Eleven subunits of protofilament and the central core are displayed by the tomographic cryomicroscopic image. The studs are clearly outlined. The spatial resolution is about 2 Å.

elled in binary terms by the action of the Galois group $\mathrm{Gal}_{\mathbf{R}}\mathbf{C}$. Chemotaxis is achieved by modulation of the tumbling frequency. When moving up an attracted gradient, the bacteria encounter an attractant concentration that increases with time. In response, they tumble less frequently and thus continue to move up the gradient. In this way, the analogy to the gradient control of clinical magnetic resonance tomography becomes obvious.

ACKNOWLEDGMENTS

This work has been financially supported by the research project "Geometric analysis of Lie groups amd applications" (GALA) of the European Union.

REFERENCES

1. E.N. Leith, "Synthetic Aperture Radar," in *Optical Data Processing*, edited by D. Casasent, Springer–Verlag, Berlin, Heidelberg, New York, 1978, pp. 89–117.
2. S.A. Lloyd, F.G. Whitby, D.F. Blair, C.P. Hill, *Nature* **400**, 472–476 (1999).
3. D.R. Lorimer, M. Kramer, Handbook of Pulsar Astronomy, Cambridge University Press, Cambridge, New York 2005.
4. A.G. Lyne, F. Graham–Smith, Pulsar Astronomy, Third Edition, Cambridge University Press, Cambridge, New York 2006.
5. T. Moritani, S. Ekholm, P.–L. Westesson, Diffusion–Weighted MR Imaging of the Brain, Second Edition, Springer–Verlag, Berlin, Heidelberg 2009.
6. J.G. Pipe, V.G. Farthing, K.P. Forbes, *Magn. Reson. Med.* **47**, 42–52 (2002).
7. W.J. Schempp, Magnetic Resonance Imaging: Mathematical Foundations and Applications, Wiley–Liss, New York, Chichester, Weinheim 1998.
8. W.J. Schempp, *Math. Meth. Appl. Sci.* **22**, 867–922 (1999).
9. W.J. Schempp, "The Fourier holographic encoding strategy of symplectic spinor visualization," in *New Directions in Holography and Speckles*, edited by H.J. Caulfield, C.S. Vikram, American Scientific Publishers, Stevenson Ranch, California 2008, pp. 479–522.
10. J.D. Schmahmann, D.N. Pandya, Fiber Pathways of the Brain, Oxford University Press, Oxford, New York 2009.

KVISIMINE++ Applied to Problems in Geographical Information System

SARITA PATEL, Prof. DEEPA CHAURSE
COMPUTER SCEINCE AND ENGG.TECHNOCRATS INSTITUTE OF TECH.
BHOPAL, M.P. INDIA.
(spatelvijay@gmail.com, deepa_chaurse@yahoo.co.in)

***Abstract*—** In this paper, describes an application K-MEAN++ clustering algorithm and image information mining for exploration of image information and large volumes data. Images are highly complex multidimensional signals, with rich and complicated information content in geographical information system. For this reason they are difficult to analyze through a unique automated approach. However a KVISIMINE++ scheme is helpful for the understanding of image and data content and provides the best segmentation of images and easy to store and retrieve information. In Geographical Information System, is any system that captures, stores, analyzes, manages, and presents data that are linked to location and like Image satellite sensors acquire huge volumes of imagery to be processed and stored in big archives. Technically, a GIS is a system that includes mapping software and its application to remote sensing, land surveying, aerial photography, mathematics, geography, and tools that can be implemented with GIS software Building a GIS is a fruitful area if one likes the challenge of having difficult technical problems to solve. Some problems have been solved in other technologies such as CAD or database management. However, GIS throws up new demands, therefore requiring new solutions. In this paper we have examine difficult problems, and to be solved and gives some indication of the state of the art of current solutions.

Keywords- K-MEAN++ clustering algorithm, Visimine, GIS.

I. INTRODUCTION

The purpose of this paper is to present K-MEAN++ clustering algorithm for choosing the initial values for the k-means clustering algorithm. K-MEAN++ **is** a way of avoiding the poor clustering found by the k-means algorithm. The k-means problem is to find cluster centers that minimize the sum of squared distances from each data set being clustered to its cluster center that is closest to it. Although finding an exact solution to the k-means problem for arbitrary input, using this approach we find an approximate solution is used to finds reasonable solutions quickly.

However, the k-mean algorithm has two major theoretic limitations:

- First, it has been shown that the worst case running time of the algorithm is super-polynomial in the input size.
- Second, the approximation found can be arbitrarily bad with respect to the objective function compared to the optimal clustering.

In K-MEAN++ clustering algorithm addresses the second of these obstacles by specifying a procedure to initialize the cluster centers before proceeding with the standard k-means optimization iterations. With the K-MEAN++ initialization, the algorithm is guaranteed to find a solution that is O (log k) competitive to the optimal k-means solution.

Our goal is to use this algorithm to segment or classification image in an automated fashion and supply the points and the number of clusters you expect to get, and the algorithm returns the same points, organized into clusters. And the VISIMINE system provides the overall structure of image database and methodology required for the analysis of GIS images. In other word we can say to facilitate the analysis of large amounts of image data, and extract features of images. Large images are partitioned into a number of smaller, more manageable image tiles. Visimine uses an SQL query language that enables the image and data mining task, and these features to be used in the mining process, and any additional constraints. Visimine data can be accessed from within MatLab. In addition, Visimine has the MatLab command tool, which provides for easy transfer of images and for data processing. . The rich statistical functionality of MatLab, together with the Visimine user interface and the scalability of its image mining engine, allows for easy and powerful customization of the image analysis process. The subject of Geographical Information Systems has moved a long way from the time when it was thought to be concerned only with digital mapping. Whereas digital mapping is limited to solving problems in cartography, GIS is much more concerned with the modeling, analysis and management of geographically related resources. However, there is a widespread lack of awareness as to the true

potential of GIS systems in the future. When the necessary education has been completed, will the systems be there to handle the challenge? It has to be said that the perfect GIS system has not yet been developed. Today's database technology is barely up to the task of allowing the handling of geographic data by large numbers of users with adequate performance. Serious questions have been raised as to whether the most popular form of database, the relational model, will be able to handle the geometric data with adequate response. Certainly, if this data is accessed via the approved route of SQL calls, the achievable speed is orders of magnitude less than that which can be achieved by a model structure built for the task. It is a common problem with systems that contain parts that are front ended by different languages that it is not possible to integrate them properly. Modern query languages such as SQL are not sufficient in either performance or sophistication for much of the major development required in a GIS system - but then one would argue that they were not intended for this. A problem which has to be addressed is spatial queries within the language, since trying to achieve this with the standard set of predicates provided is extremely difficult and clumsy. An example of a spatial query is to select objects "inside" a given polygon. If the route adopted is to provide two databases in parallel, a commercial one driven by SQL and a geometry database to hold the graphics, and then there is a problem constructing queries that address both databases. The rest of the paper is organized as follows: Section 2 gives the details of related work, proposed work introduced in section 3.And Preliminaries in Section 4, and we discuss our experiments and the results in section 5. Conclusions are presented in Section 6.

II. RELATED WORK

A great deal of research has been focused the use *o*f GIS in the spatial analysis of an archaeological cave site[1], according to HOLLEY MOYES have viewed geographic information systems (GIS) as a tool for the investigation of large regions, its flexibility allows it to be used in non-traditional settings such as caves. This study demonstrates the utility of GIS as a tool for data display, visualization, exploration, and generation. Clustering of artifacts was accomplished by combining GIS technology with a K-means clustering analysis, and basic GIS functions were used to evaluate distances of artifact clusters to morphological features of the cave. The use of GIS in Archaeological Settlement Research Facts, Problems and Challenges [2], Frankfurt Germany, September 26th 2008 using Free and Open Source Software (FOSS) licenses generally allow free deployment anywhere and for any purpose. No redundant licensing costs, more flexible investment options, full control over development. Stable and long-lived data formats, free and *open standards* instead of "industry" no pressure to deprecate older software or data formats. The current pool of available FOSS is gigantic and growing rapidly. Open source licenses generally allow free deployment anywhere and for any purpose. Clustering with GIS [3], Ece AKSOY, Turkey, presented there is no universally applicable clustering technique in discovering the variety of structures display in data sets. Also, a single algorithm or approach is not adequate to solve every clustering problem. In GIS environment used Self Organizing Maps (SOM) algorithm which is the best and most common spatial clustering algorithm in recent years. Geospatial Information and Geographic Information Systems (GIS): Current Issues and Future Challenges in June 8, 2009[4], according to Peter Folger**,** Geospatial information is data referenced to a place a set of geographic coordinates which can often be gathered, manipulated, and displayed in real time. A Geographic Information System (GIS) is a computer system capable of capturing, storing, analyzing, and displaying geographically referenced information. Global Positioning System (GPS) data and their integration with digital maps have led to the popular hand-held or dashboard navigation devices used daily by millions.Challenges to coordinating how geospatial data are acquired and used collecting duplicative data sets. Implementation of the Extended Fuzzy C-Means Algorithm in Geographic Information Systems [5], Ferdinand Di Martino' Salvatore Sessa, in 2009, focused on density cluster methods have elevated computational complexity and are used in spatial analysis for the determination of impact areas. We propose the extended fuzzy c-means (EFCM) algorithm like alternative method because it has three advantages: robustness to noise and outliers, linear computational complexity and automatic determination of the optimal number of clusters. We can use the EFCM algorithm in spatial analysis for the determination of circular buffer areas. These areas can be considered on the geographic map as a good approximation of classical hotspots. Applications to other frameworks like crime analysis, industrial pollution, etc. shall be tried in future works. Issues of GIS data management [6] 2007, this paper deals with current issues of spatial data modeling and management used by spatial management applications. Paper describes ways of solving this problem. Now we can summarize the problem of the GIS and CAD integration. Because of the different characteristics of the GIS/CAD worlds, firstly there's need to decide for some suitable 3D data model, which could maintain complex and structured data types. This model also must be able to maintain the large-scale 3D models produced by CAD as well as low-scale objects used by GIS.Comparative Analysis of k-mean Based Algorithms [7] 2010, in this paper they make analysis of k-mean based algorithms namely global k-means, efficient k-means, k-means++ and x-means over leukemia and colon datasets.

III. PROPOSED WORK

In this paper we proposed KVISIMINE++, is the combination of K-mean++ clustering and Visimine.

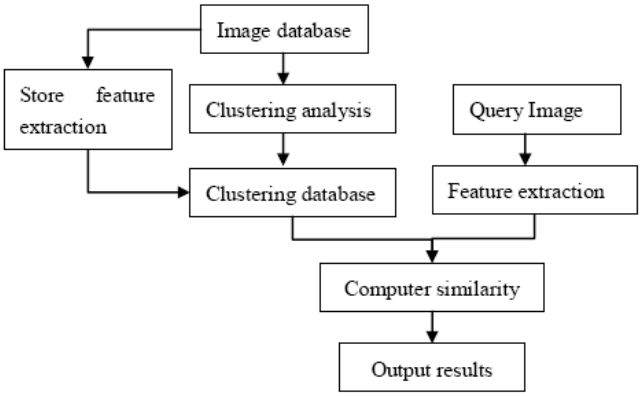


Figure 1. Combination of K-mean++ and Visimine

A. K-MEAN++ clustering algorithm

K-MEAN++ is another variation of k-means, a new approach to select initial cluster centers by random starting centers with specific probabilities is used. With the intuition of spreading the k initial cluster centers away from each other, the first cluster center is chosen uniformly at random from the image data points that are being clustered, after which each subsequent cluster center is chosen from the remaining image data points with probability proportional to its distance squared to the point's closest cluster center.

The algorithm is as follows:

1. Choose one center uniformly at random from among the data points.

2. For each data point x, compute $D(x)$, the distance between x and the nearest center that has already been chosen.

3. Add one new data point at random as a new center, using a weighted probability distribution where a point x is chosen with probability proportional to $D(x)^2$.

4. Repeat Steps 2 and 3 until k centers have been chosen.

5. Now that the initial centers have been chosen, precede using standard k-means clustering.

This method gives out considerable improvements in the final error of *k*-means. Although the initial selection in the algorithm takes extra time, the *k*-means part itself converges very fast. But lowers the computation time too. The method provides with real and synthetic datasets and obtained typically 2-fold improvements in speed, and for certain datasets close to 1000-fold improvements in error. Additionally, we calculate an approximation ratio for their algorithm. The k-means++ algorithm guarantees an approximation ratio $O(\log(k))$ where k is the number of clusters used. This is in contrast to *k*-means, which can generate clustering arbitrarily worse than the optimum.

B. Database Organization

Visimine: is a system for data mining and statistical analysis of large collections of remotely sensed images. And also provides the infrastructure and methodology required for the analysis of land surveying, aerial photography, and mathematics, geography and satellite images. Visimine uses an SQL-like query language that enables specification of the data mining task, features to be used in an image mining process, and any additional constraints. The query language allows the user to specify the type of knowledge to be discovered, and the set of image data relevant to the image mining process. Based on this an SQL query statement is constructed to retrieve the relevant image data. This new domain requires expertise in image processing, database organization, pattern recognition, content based retrieval and data mining: image processing indicates the understanding and extraction of patterns from a single image; in this system provides users the capability to deal with large collections of images by accessing into large image databases and also to extract and infer knowledge about patterns hidden in the images, so that the set of relevant images is dynamic, subjective and unknown. It enables the communication between heterogeneous source of information and users with diverse interests at high semantic abstraction. The GUI (graphical user interface) enables browsing and manipulation of the images and associated features, creation of data mining queries, and visualization of the results of the data analysis.

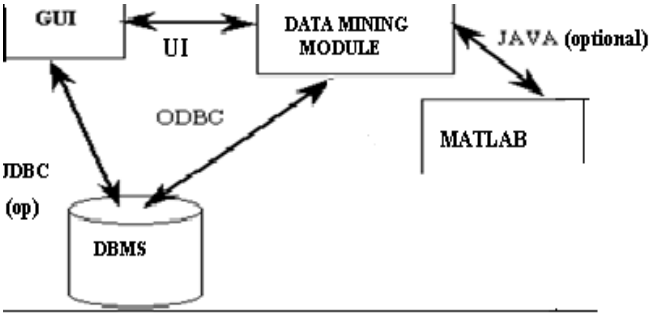


Figure 2. Visimine System

The system is capable of performing similarity searches based on any combination of features. Visimine is based on decision tree models. Visimine data can be accessed from within MatLab by using Java connectivity for images and ODBC connectivity for image and region data .In MatLab, which provides the interactive program for numerical computation and data visualization which along with its

programming capabilities provides a very useful tool for almost all areas of science and engineering, which is used extensively in both academia and industry. The Visimine can also display graphics, which are created using a command line interface and shown within MatLab figure window. The combination of MatLab and Visimine features creates a unique environment for interactive exploration and analysis of remotely sensed image data.

IV. PRELIMINARIES

Problem definition: Modern query languages such as SQL are not sufficient in either performance or sophistication for much of the major development required in a GIS system - but then one would argue that they were not intended for this. One can see why people like SQL; it can give immense power in return for some fairly simple "select" constructs. A problem which has to be addressed is spatial queries within the language, since trying to achieve this with the standard set of predicates provided is extremely difficult and clumsy. If the route adopted is to provide two databases in parallel, a commercial one driven by SQL and a geometry database to hold the graphics, then there is a problem constructing queries that address both databases. Ideally, the query language should be a natural subset of the front end language allowing access to the same seamless environment that the front end language provides. Much work needs to be done in the area of query languages for GIS.

V. EXPERIMENTAL RESULTS

The KVISIMINE++ can also display MatLab graphics, which are created using a command line interface and shown within figure window.

Table 1. Best results

Results over different variations of k-means algorithm using a tree image classified according colors.		
(Total number of records present in dataset = 70)		
Clustering Algorithm	Correctly Classified	Average Accuracy
k-means	68	94.88
k-means++	70	95.83

The combination of K-MEAN++ and Visimine features create a unique graphics environment for image mining and these images are stored in image database in form of vector and MatLab provides the good database connectivity in Visimine. In order to facilitate the analysis of large amounts of image data, we extract features of the images. Large images are partitioned into a number of smaller (segmentation), more manageable image tiles. Partitioning allows fetching of just the relevant tiles when retrieval of only part of the image is requested, and provides faster segmentation of image tiles. Individual image tiles are processed to extract the feature vectors.

PERFORMANCE EVALUATIONS

MatLab connectivity: is an interactive computing Environment for graphics, data analysis, statistics, and mathematical computing. Data was then transferred from MatLab to MS Access using the database connectivity (ODBC) tools as provided by the MatLab Database Toolbox. The data was then transferred from MatLab across the LAN to the SQL Server 7. This process is repeated for matrices with 4 columns per row then 253 columns per row. Each matrix contained 1000 rows. Once the MatLab process was complete, MatLab was closed and MS Access opened. A process was then run that gathered the timestamp information for each row written to the MS Access tables and the SQL Server 7 tables. The SQL Server 7 tables were then emptied, and the row data in MS Access was written to SQL Server 7.It contains a superset of the S object-oriented language and system originally developed at AT&T Bell Laboratories, and it provides an environment for high-interaction graphical analysis of multivariate data, modern statistical methods, data clustering and classification, and mathematical computing. In total, MatLab contains over 3000 functions for scientific data analysis. Visimine data can be accessed from within MatLab by using Java connectivity for images and ODBC connectivity for image and region data. In addition, Visimine has the MatLab command tool, which provides for easy transfer of images, and for data processing. The KVISIMINE++ display MatLab graphics, which are created using a command line interface and shown within figure window. The combination of MatLab and KVISIMINE++ features creates a unique environment for interactive exploration and analysis of remotely sensed image and data. The rich statistical functionality of MatLab, together with the approach user interface and the scalability of its data mining engine, allows for easy and powerful customization of the data analysis process.

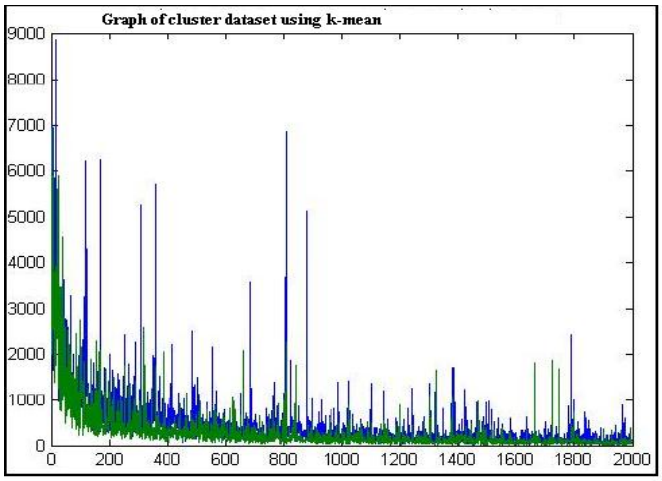


Figure 3.a Graph for k-mean

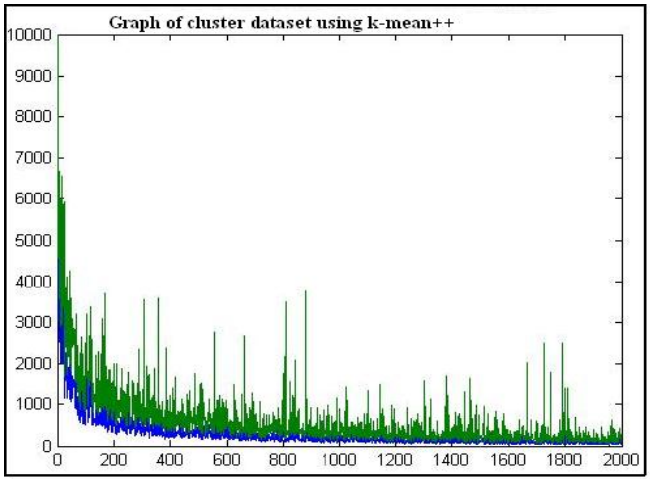


Figure 3.b Graph for k-mean++

According to graph we get better quality of clusters we can use these concepts. And k-means initial selection of cluster centers plays a very important role.

VI. CONCLUSIONS

In this paper we presented a KVISIMINE++ MatLab framework, provides powerful numeric engine and technical programming environment with makes it easy to create and manipulate vectors , MatLab has become the language of technical computing, which explores state-of-the-art data mining and databases technologies to retrieve integrated spectral and spatial information from Geographical information system imagery. A scalable data warehouse containing a huge amount of images may be a better database architecture for fundamentally distributed data management and mining system such as NASA Earth Observing System (EOS). Meanwhile, performance analysis for clustering on and retrieving from large volumes of images is critical for the system to succeed in practical applications. And the results of experiments on the basic of images show that the proposed approach can greatly improve the efficiency and performances of image retrieval, as well as the convergence to user's retrieval concept. Clustering algorithm has been widely used in computer vision such as image segmentation and Visimine is able to distinguish between pixel, region and tile levels of features, providing several feature extraction algorithms for each level. In addition, current implementation provides data and image based search. A segmentation process can be used to segment an image into non-overlapping regions on which we can further apply the texture feature extraction.

REFERENCES

[1] U. M. Fayyad, G. P. Shapiro, P. Smyth, and R. Uthurusamy (Eds.), "Advances in Knowledge Discovery and Data Mining," AAAI/MIT Press, 1996.

[2] M.C. Burl, C. Fowlkes, and J. Roden, "Mining for image content," in Systemic, Cybernetics, and Informatics / Information Systems: Analysis and Synthesis, Orlando, FL, July 1999.

[3] L.K. Soh, C.Tsatsoulis, "Data mining in remotely sensed images: a general model and an application," in Proceedings of IEEE IGARSS 1998, vol. 2, Seattle, Washington, USA, Jul 1998, pp. 798-800.

[4] J. Zhang, H. Wynne, M. L. Lee, "Image mining: issues, frameworks, and techniques," in Proceedings of 2nd International Workshop on Multimedia Data Mining, San Francisco, USA, Aug 2001, pp.13 – 20.

[5] G.B.Marchisio, J.Cornelison,"Content-based search and clustering of remote sensing imagery," in Proceedings of IEEE IGARSS 1999, vol. 1, Hamburg, Germany, Jun 1999, pp. 290 – 292.

[6] A.Vellaikal, C.-C.Kuo, and S. Dao, "Content-based retrieval of remote sensed images uses vector quantization," in Proc. of SPIE Visual Info. Processing IV, vol. 2488, Orlando, USA, Apr 1995, pp.178 – 189.

[7] Ying Liu, Dengsheng Zhang, Guojun Lu, Wei-Ying Ma,"A Survey of content-based image retrieval with high- Level Semantics," Pattern Recognition, Volume 40, Issue 1, January 2007, Pages 262-282.

[8] Muhammad Atif Tahir, Ahmed Bouridane, Fatih Kurugollu, "Imultaneous feature selection and feature weighting Using Hybrid Tabu Search/K-nearest neighbor classifier," Pattern Recognition Letters, Volume 28, Issue 4, 1 March 2007.

[9] Sarbast Rasheed, Daniel Stashuk, Mohamed Kamel, "Adaptive Fuzzy k-NN classifier for EMG signals Decomposition," Medical Engineering & Physics, Volume 28, Issue 7, September 2006, Pages 694-709.

[10] J. Amores, N. SEbE, P. Radeva, "Boosting the distance Estimation: Application to the K-Nearest Neighbor Classifier," Patter Recognition Letters, Volume 27, Issue 3, February 2006, Pages 201-209.

[11] Man Wang, Zheng-Lin Ye, Yue Wang, Shu-Xun Wang," Dominant sets clustering for image retrieval," M. Wang et al. /Signal Processing 88 (2008) 2843–2849., Venables W. N. and Ripley B. D. (2000), S Programming, Springer, New York.

[12] M.C. Burl, C. Fowlkes, and J. Roden, "Mining for image content," in Systemic, Cybernetics, and Informatics / Information Systems: Analysis and Synthesis, Orlando, FL, July 1999.

[13] Zlatanova S,"Large-scale data integration," An Introduction to the Challenges for CAD and GIS Integration, Directions magazine, July 10, 2004.

[14] Van Ostracism P,"Bridging the Worlds of CAD and GIS," Directions magazine, June 17, 2004.

[17] David Arthur, and Sergei Vassilvitskii, "k-means++ the Advantages of careful," seeding, Proceedings of the eighteenth Annual ACM-SIAM symposium on discrete algorithms, pp. 1027—1035, 2007.

[18] Zhang Y, Mao J, and Xiong Z, "An efficient Clustering Algorithm," In Proceedings of Second International Conference on Machine Learning and Cyber netics, Nov 2003.

[19] IEEE Trans. on "Knowledge and Data Engineering, "14, No.5, Sept/Oct 2002.

An Exploratory and Feasibility Study of Implementing Online Based Voting System in Bangladesh

Mohammad Shabbir Hasan[1], Quazi Farhan[2] and Abdullah Al Mahmood[3]

Panacea Research Lab, Dhaka, Bangladesh
[1]shabbir_cse03@yahoo.com, [2]quazifarhan@gmail.com, [3]sajibcseofkuet@gmail.com

***Abstract*:** Modernization in voting system with various technologies has become an attention-grabbing issue in recent years. Currently, internet-based election systems are in the early stages of development and testing even in the developed countries and have been referred to as the ultimate challenge in network security and data encryption. Merging an information system with real-life problem has never been an easy task and to satisfy the zero-tolerance condition, the information systems implementation needs to handle lots of technical details. In Bangladesh, general elections are arranged on entirely paper based ballot system and manual voting procedures are employed. Here, electronic voting system is still in an experimental phase and the possibility of internet-based remote electronic voting in near future is not yet considered as tenable. But recent election problems in Bangladesh have sparked great interest in managing the election process through the use of internet to enhance the voters' scope for participating in the election as well as create scope for more error free rapid tallying of votes and distribution of seats and to enable the election commission to promptly announce the election results within a short time. In this paper, an online based voting system is proposed to eliminate the problems and bottlenecks of the existing voting systems in Bangladesh.

***Keywords*:** Online Voting System, Voting Systems in Bangladesh, Electronic Voting System.

1. Introduction

Election is the way through which people choose their representatives and express their preferences for how they will be governed. Naturally, the integrity of the election process is fundamental to the integrity of democracy itself [1]. Again, elections usually have high media coverage, especially if something goes wrong. Furthermore, voting system seems to have a unique combination of security requirements: voters need to be authenticated as well as results need to be verifiable. Voting systems are hard to make trustworthy because they have strong, conflicting security requirements: Integrity and confidentiality [2]. *Integrity* means election results must be guaranteed so that all voters are certain that votes are counted correctly and *confidentiality* means voters must be assured about the privacy of their votes, avoid selling of votes, and protect voters from coercion. In a word, voters need to be authenticated, results need to be verifiable, but it should not be possible to link a vote to a voter [3]. Although these conflicting requirements are very difficult to satisfy, the adaption of technology in upgrading voting system is practiced worldwide. The first use of computers to count votes — came with the introduction of the punch card system, first used in 1964. DREs (Direct Recording Electronic systems) are the first completely computerized voting systems. They were introduced in the 1970s [4]. Then some European countries started to introduce electronic voting systems. Hence, it can be said that Electronic voting is now a reality but along with many errors and vulnerabilities in commercial electronic voting systems [5, 6, 7, 8]. Many security experts have been skeptical about electronic voting [9, 10, 11, 12, 13], arguing that assurance in electronic voting systems is too hard to obtain and that their deployment creates unacceptable risks. More proofs of the vulnerability of an electronic voting system can be inferred from the fact that the use of similar paperless DREs has been discontinued in California [14], Florida [15], Ireland [16], The Netherlands [17], and Germany [18].

Recently, beside the conventional paper based voting system, Bangladesh Election Commission has also started experiments on electronic voting system by using electronic voting machines known as EVMs in some of the centers for Chittagong City Corporation Election 2010. As a process of building up *Digital Bangladesh*, it is highly likely that EVMs will be extensively used in the next National Election of Bangladesh. Even the possibility of remote electronic voting system, though not tenable in near future, is not out of question as there have been some research works conducted on SMS based secured electronic voting system [19] and internet based secured electronic voting system [20] also. This paper focuses on a comparative exploratory and feasibility study of online voting system with the existing voting system to upgrade a sensitive issue like national election, especially in a country like Bangladesh which is disreputably known for corruption and political mayhem.

The organization of the paper is as follows: section 2 briefly discusses about various voting systems previously implemented in various countries of the world. Section 3 describes the security concern about the existing voting systems in Bangladesh. Section 4 describes the proposed online voting system followed by the verification of the casted vote in Section 5. Features of the proposed system and a comparison between the present voting system in Bangladesh and the proposed system are mentioned in Section 6 and 7 respectively. Section 8 includes some recommendation and conclusion in section 9.

2. Various Types of Voting Systems

Over the years, many innovative changes have occurred to enhance election systems mainly in order to reduce various types of election frauds. According to Coleman and Fischer, currently, five different technologies are in use — paper ballots, lever machines, punch cards, optical scan, and electronic systems (direct recording electronic or DRE) [21]. Many states of USA even use more than one system to strengthen security. Online voting is a reality in many European countries. Multiple casts in online voting became popular by the Estonian's legal binding Local Government Council Election in autumn 2005 [22]. Here in Bangladesh, the scenario is not so advanced in terms of both security and technology. Paper ballot was the only system available in Bangladesh. However, recently the Election Commission started experimenting with EVM (Electronic Voting Machine).

A. Paper Ballots: This is the most common and classical method of voting. In this system, the candidate lists along with their respective parties are placed in a ballot paper. Voters mark their choices on the ballot. Each voter gets one paper. The vote counting system is totally manual. All voting technologies using document ballots use paper or cardstock, but the term paper ballot generally refers to those that are designed to be read by humans rather than machines [23].

B. Lever Machines: There is no document ballot in this technology. A voter enters the voting booth and chooses candidates listed on a posted ballot by pulling a lever for each candidate choice. The votes are recorded by advances in a counting mechanism that are made when the voter leaves the booth [24].

C. Punch Card: This is the first technological approach utilizing computers to count votes. This was first used in 1964. In this system, the voter records choices by punching holes in appropriate locations on a paper computer card that is later fed into a computer reader to record the vote. The computer card serves as the document ballot on which the votes are recorded [24]. Punch cards can be manually recounted and audited.

D. Optical Scan: This technology has been used for decades in scoring standardized tests. It first became available for use in voting in the 1980s. In this system, a voter fills in a box or oval or completes an arrow corresponding to each candidate choice using a paper form and an appropriate writing instrument. The completed ballot is then read by a computerized device that senses and records the marks [24].

E. Electronic Voting Machine: DREs (direct recording electronic systems) are the first completely computerized voting systems. They were introduced in the 1970s [4]. In this system, the voter chooses candidates from a posted ballot which may be printed and posted on the DRE or it may be displayed on a computer screen. Voters make their choices by pushing buttons, touching the screen, or using other devices depending on the equipment used.

DREs can be classified into three basic types. The oldest design mimics the interface of a lever machine. The entire posted ballot is visible at once. The voter pushes a button next to a candidate's name, or pushes on the name itself, triggering an underlying electronic micro switch and turning on a small light next to the choice [4]. In case of the second design, a ballot page is displayed on a computer screen, and the voter uses mechanical devices such as arrow keys and buttons to make choices on a page and to change ballot pages. The third type is similar to the second except that it has a touch screen display. Here, the voter makes a choice by touching the name of the candidate on the computer screen and casts the ballot by pressing a separate button after all choices have been made. In Bangladesh, the oldest type of DRE which is popularly known as EVM is under experiment at present.

Another form of electronic voting currently in development is Internet voting, in which voters make their choices online. This system is far more advanced in terms of technology but poses special challenges for ensuring authentication, secrecy, and security in the voting process.

3. Various Security Vulnerabilities of the Existing Voting Systems in Bangladesh

A. Security Concerns about Paper Based Voting System:

Paper ballot system is the conventional paradigm which is being used in Bangladesh from the beginning of voting. Paper ballots are readily understandable by the voter which is a great advantage as there are a huge number of uneducated voters here in Bangladesh. The security concerns are well understood by the authority as there are no chances of high-tech security breach in this system. The tampering of ballot paper is possible but this requires the involvement of corrupt officials. Ballot box hijacking and coercing the present officials to manipulate the result in a

particular center are some other types of vulnerabilities in this system.

B. Security Concerns about EVMs:

Electronic Voting Machines, known as EVMs are widely used in Indian elections and now under experiment in Bangladesh. These EVMs are said to have less complex code than the previous electronic voting systems like DREs. In spite of this simplicity which makes them less susceptible to some of the threats faced by DREs, it also subjects them to a different set of highly dangerous attacks. An Indian research revealed two major types of possible attacks on Indian EVMs [25]. One of them is named as *Dishonest Display Attack* in which the real display board in the control unit of an EVM is replaced by a dishonest display board developed by the attacker. Another type of attack is called *Clip-on Memory Manipulator Attack* which uses new hardware to alter the internal state of the machine. In "Security Analysis of India's Electronic Voting Machines", the authors demanded that the Indian EVM manufacturers are exporting machines to Bangladesh [25] which is not correct as the EVMs which are currently being used in Bangladesh are developed by Information and Communication Technology (IICT) of Bangladesh University of Engineering and Technology (BUET) and a local manufacturer company named Pi Labs, Bangladesh. However, when asked about the similarities, the developers acknowledged that both share many similar characteristics and design patterns and Bangladeshi EVMs are also vulnerable to both *Dishonest Display Attack* and *Clip-on Memory Manipulator Attack* mentioned above. Bangladeshi EVMs are unique in the way that *Smart Cards* are used for configuring these EVMs. Smart Card is used to configure the Bangladeshi EVM which reduces the cumbersome work at the field level. This configuration procedure is done in two stages. In the first stage Symbols of the candidates and their allotted button number are written on the Smart card from a PC application. And then in the second stage at the field level, authorized person inserts the card to the control unit and presses the specified button and that finishes the configuration process. In contrast to the Indian EVM's configuration procedure the process described above might appear too much user friendly and less cumbersome but Indian EVM manufacturers may claim superior transparency in this aspect as it is done manually. The use of smart cards for configuring EVMs opens up another opportunity of susceptibility. If someone can have the chance of having access to any of the smart cards then he/she can configure it as desired. This *Smart Card Attack* also needs physical access like the previously mentioned two types of attacks. Another major concerning issue of these EVMs is that they do not use any kind of encryption for signal transferring.

4. Proposed Online Based Voting System

The proposed system comprises of several steps. The system is accessible from two sides: (a). Election Commission who is the administrator and (b). The voter. There are some steps which are automated i.e. not accessible from any side. Figure 1 shows the possible input output scenario of the proposed system and total system architecture is presented in detail in Figure 2. Steps of the proposed system are described in the following section.

A. Adding Voter Information

In this proposed system, information of each voter is added according to their National Identity Number. This National Identity Number is unique for each voter and this number is also used to identify the constituency of the voter. After adding information, an auto generated e-mail is sent to the e-mail address of the voter notifying him/her about the information and this e-mail also contains a computer generated password which can be used by the voter for login as well as for changing password and setting verification keys. Here verification keys are used to protect "*Vote Purchase*" and to ensure security. On screen keyboard is used to take the new password from user. The system will not take the password typed in any keyboard other than the onscreen keyboard. Purpose of using on screen keyboard is to prevent capturing password through any software if the voter changes password from any cyber café or in any public computer. Discussion about the use of verification keys is presented later.

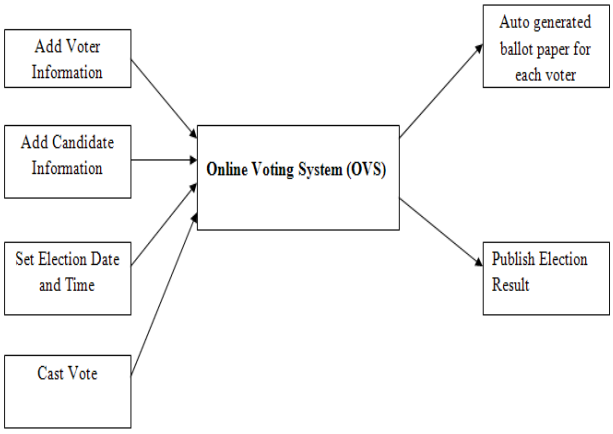


Figure 1. Possible Input-Output scenario of the proposed Online Voting System

B. Adding Candidate Information

Candidate information is added according to the constituency. Here each candidate is assigned an auto generated code to identify uniquely. Party symbol and candidate profile image are also added with other information in this phase.

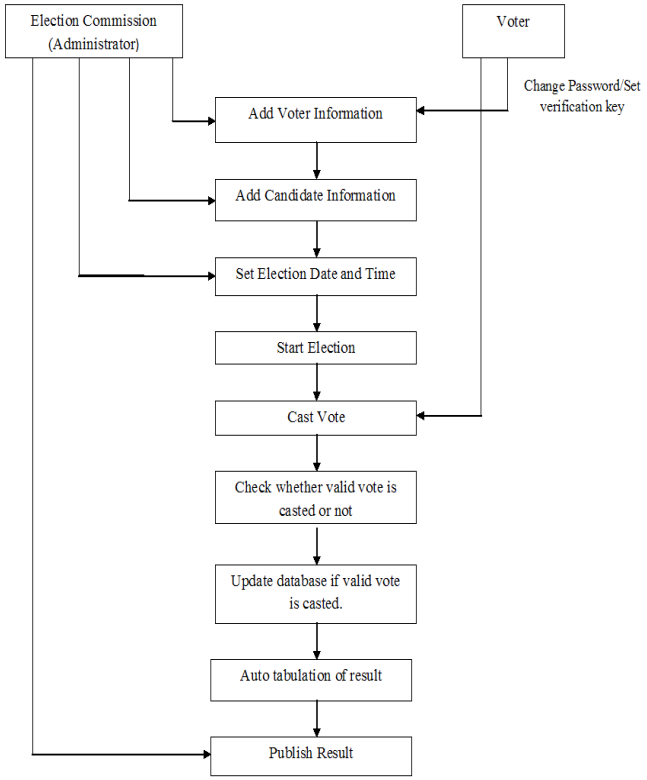


Figure 2. Details architecture of the proposed system.

C. Setting Election Date and Time

In this phase, starting time and ending time of election along with election date are set by the election administrator. Bangladesh is in the "Dhaka" time zone which has +6.00 offset from GMT. The voting server is configured according to this time zone. Voters who reside outside of Bangladesh can cast their vote according to the local time in Bangladesh during Election Day.

D. Start Election

This is an automated phase. During the stipulated date and time, election is started. Voter can cast their vote within this time period. If anyone wants to cast vote before or after the specific time period, an error message is shown.

E. Cast Vote

In this phase, voter has to login first. After logging in with the national id and password, the constituency of the voter is determined from the information stored in the voter database. An E-Ballot paper is created automatically for that constituency from information stored in candidate database. This ballot paper contains the candidate name along with their profile picture, party name with party logo and a radio button to select the candidate for casting vote. There is also a "No" option if the voter is not interested to cast vote to any of the available candidates. Figure 3 shows a sample ballot paper for a constituency.

E-Ballot Paper for the Constitution: Chittagong 1

Candidate ID	Candidate Image	Name	Logo	Party Name	Select
C14		Mr. M		Bangladesh Awami League	○
C15		Mr. N		Bangladesh Nationalist Party	○
C16		Mr. O		Jatiya Party	○
C17		Mr. P		Jatiyo Samajtantrik Dal	○
C18		Mr. Q		Communist Party of Bangladesh	○
C19		Mr. R		Workers Party of Bangladesh	○
No	N/A	N/A	N/A	N/A	◉

Cast Vote

Figure 3. A sample E-Ballot paper for a constituency.

Voter selects radio button of the corresponding candidate and finally press the "Cast Vote" button. A security checking is done to verify whether the vote is casted actually by the voter himself. A discussion about the security checking is presented in the next section.

F. Security Checking

When a voter presses the "Cast Vote" button in the e-ballot paper after selecting the suitable candidate, a security checking is done internally. This security checking is one of the most striking points of this proposed system as it is required to protect "*Vote Purchase*" by any candidate. Figure 4 shows the steps of the security checking process.

Voter can enter verification key before election by logging in his/her account as we mentioned in section A. Number of verification keys to be used are selected by the voter, but among them, only one key is used as "*Real Key*" and rest of them are treated as "*Fake Key*". Among these keys voter will select which one will be used as "*Real Key*". In case of any attempt to "*Purchase Vote*" by any candidate, the voter has the option to hide "*Real Key*" and supply only "*Fake Keys*" to the candidate. If vote is casted by using fake keys, notification will be shown as if vote is casted successfully, but there will be no update in database. So if the voter wants to cast his/her vote, he/she can cast vote using the real key in any time within the voting period. This process will reduce

the tendency of purchasing vote by any candidate and hence make the election process fair.

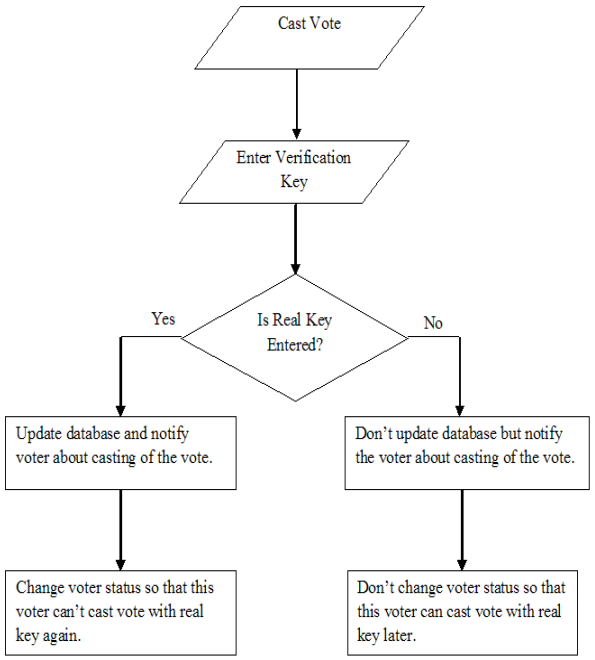


Figure 4. Flow Diagram of the Security Checking Process.

G.Auto Tabulation and Publication of Result

During the election period, result is tabulated automatically and after the election period is over, winner in each constituency is declared automatically. In this proposed system, results can be published immediately which leads to a huge saving of time than the existing methods.

4.1 Security Measurement of the Proposed System

A basic assumption of online voting system is that it does not disregard the underpinning principles of an electoral system. In traditional system voters attend at designated polling centers and nominate their candidate using ballot papers. This way the voter can be assured that his/her confidentiality has been preserved and vote has been casted correctly. But using online system, when voters are casting their vote digitally from their household they cannot be sure that their vote is indeed counted. On the other hand if voters are provided with the opportunity of verifying their vote after casting, this can be used for coercion or vote selling. These contradictory issues are the first hurdle of online voting.

Beside this, there is also a chance of wire tapping or any other means through which voter's identity and vote can be exposed to sniffers. It can be interrupted in the way or may get lost due to bad transmission. To enhance security, in this proposed system, vote is encrypted after casting.

In this proposed system the dilemma between voter integrity and confidentiality is solved using the key that the voter used previously to cast vote. This verification technique is described in detail in the next section.

5. Verification of the Casted Vote

To validate the correctness of the voting, to assure the voter that the vote has been casted exactly the manner he/she intends; multiple, independent communication is required, as employed by the Moguls in India some 500 years ago in the context of combating corruption [26] or mathematically described by Claude Shannon some 50 years ago in the context of combating noise when he introduced his Information Theory [27], a well-known general theory of communication processes.

The problem with the solution is maintaining two parallel schemes is no small feat. The cost, manpower, complexity and newly arisen security holes nullify the advantages of online voting. Worse yet is even with such exaggerated method we cannot rectify an error. Let us consider two ballot of a voter shows two opinions. Online submission chooses one candidate and offline/printed copy supports another. With no way to deduce which was the original choice of the voter, the approach can only ensure that the online version might be wrong but not with certainty, since the offline ballot might tampered with, too.

Now, here comes the most crucial contradiction of using an online system: conflict of integrity and confidentiality. While integrity requires all voters' votes must be counted and voters must be assured that their votes are casted and counted properly. But confidentiality complies that there is no way a voter's ballot and voter's id can be connected, even if a court order comes in or supervisors in charge of election collude. This dilemma can be solved using the key that was used previously by the voter to cast vote. Voter can verify whether the vote has been casted as well as counted correctly by inputting the key that he/she provided during casting his/her vote. In database, information regarding the casted vote and the corresponding key is stored. So voter can retrieve his/her vote information by providing the key. Even if vote is casted using one of the fake keys, it also returns the candidate name that was selected by the unauthorized person during casting vote by using fake keys though this vote has no effect in counting the total vote. As only candidate name with the corresponding key is stored in the table, i.e., no voter information is stored so only voter can check whether vote is correctly casted or not. So it preserves the confidentiality. Again as the system allows voter to become assured of the casted vote, it also ensures integrity of the voter. This verification system is presented in Figure 5.

Another layer of security is client software. Instead of implementing this as a web interface, a java applet or custom software is prepared, which is available to download at the voting site just prior the voting commences and the server always monitors the applet/client for any kind of modification or tampering. In this way we can ensure another layer of security even on the client end.

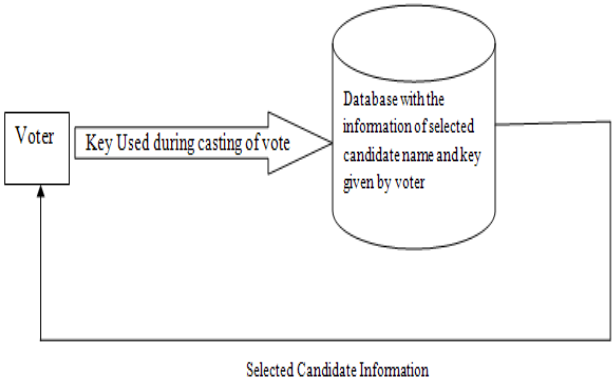


Figure 5. Verification by the voter whether vote is casted and counted correctly.

6. Features of the Proposed System

The features of the proposed system are:

a. No need to stay in the queue for a long time to cast vote, anyone can cast his/her vote from anywhere through internet.

b. Increased voter turnout.

c. The cost for arranging election will be reduced as there is no need to prepare any voting center or no need to manage huge manpower.

d. Impervious security checking to verify whether vote is casted by the voter himself.

e. Verifiability by the voter of the casted vote.

f. Coercion resistant voting system.

g. Result can be published within a very short time as the system automatically tabulate the results and declare winners.

h. Overall, the election will be fair enough and transparent.

7. A Comparative Study between the Proposed System and the Existing System

As paper ballot based voting system is being used in most of the voting centers in Bangladesh, here we compare the proposed system with the paper based voting system. To compare with the existing system, the proposed system was implemented in the trade union election of Technocrats BD on February 2010. The election system was totally online based and the previous elections were paper ballot based like other traditional elections. Comparative statistics between these two election methods are shown in Table 1.

Table 1 shows that the proposed system increases voter turnout than the previous method as online voting system is very much convenient than the traditional voting system. In this proposed system voter cast their vote from any place through internet rather waiting in a queue for long time and this is the main reason for making the voting system interesting which leads to the increase in voter turnout.

From Table 1 it is also clear that there is no way of invalid vote in the proposed system which is a case in the traditional one. As the e-ballot paper contains radio button for each candidate, voter can select only one candidate at a time for the same post. If the voter doesn't select corresponding radio button of any candidate, "NO" vote is casted which means the casted vote doesn't correspond to any candidate. So there is no way of either casting multiple votes in the same ballot paper at a time or cast vote without selecting any radio button. That's the reason of 100% valid vote in the proposed system.

Table 1. Comparative statistics between two election methods.

Criteria	Paper Based Election on January 2006	Online Based Election System on February 2010
Total Voter	1054	1173
Turnout	797 (75.6%)	1053 (89.8%)
Valid Votes	753 (94.47%)	1053 (100%)
Total Cost (in Bangladeshi Taka)	84,320	82,110

Another striking point of the proposed system is the reduction of cost to arrange an election. Lots of manpower and funds are required to arrange the election in the traditional manner. But in this online voting system, there is no need of setup cost for any voting center as voters can cast vote from their personal computer and hence no manpower is required to maintain any voting center. Few people are enough to manage the complete process of the proposed system as many steps are fully automated. Huge reduction of cost is clear from Table 1. In the traditional system, 80 Bangladeshi Taka is required for each voter where as only 70 Bangladeshi Taka is required in the proposed system. With the increase in total number of voters, total cost to arrange election in the proposed system decreases rapidly than the traditional system which is shown in Figure 6.

Finally, after this comparative study, it can be said that better turnout and strong security system with low setup cost have made the proposed voting system better than the traditional system.

8. Recommendation

Before adopting any new voting system there should be at least two official technical evaluations and reviews of those technical evaluations should be made public so that general people can have more faith in new system. These technical evaluations must be performed by an expert committee which should have profound prior knowledge on computer security. One of the technical evaluations for Indian EVMs was conducted by an "expert committee" comprised of C. Rao Kasarbada, P.V. Indiresan, and S. Sampath [28], none of whom appear to have had prior computer security expertise [25]. Unfortunately, Bangladesh also has many previous examples of forming an expert committee without any expert members of the related issue. So, it is to be noted very carefully that a real expert committee must be formed for an official technical evaluation for such a sensitive matter like National Election. Moreover, this step will ensure that people have a reason to put their faith in this system.

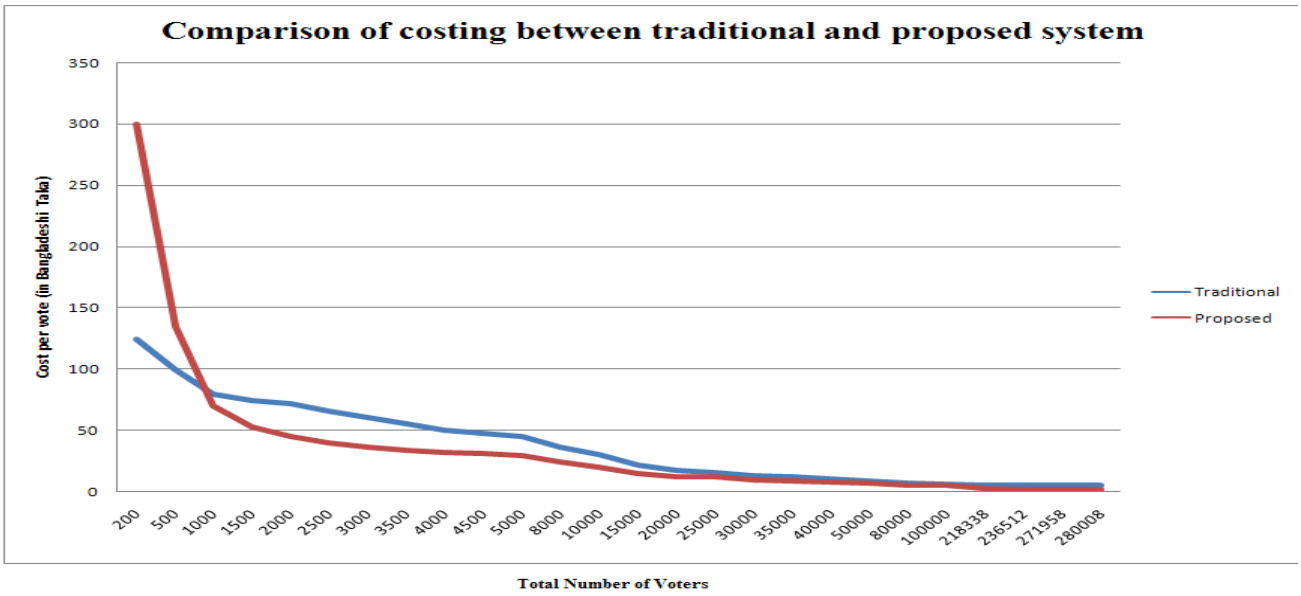


Figure 6. Comparison of costing between traditional and proposed voting system.

9. Conclusion

This paper discusses about various security issues of the existing voting systems in Bangladesh and also describes the design, implementation and evaluation of an online based efficient voting system. To our knowledge, this has not been done before in Bangladesh. Deriving from the previously-known voting scheme, the proposed system is coercion resistant and ensures security and efficiency through technical advances. Experimental results show that cost, tabulation time and security can be practical for real-world elections.

When a system like voting needs to be upgraded from a classical human dependant to a technologically advanced computer dependant system, it is certain that the system authority will have to face difficulty with security as well as lots of adversary of technical, social and also bureaucratic. But perhaps the most important contribution of this work is evidence that secure online voting system could be made possible and we are optimistic about the future of the proposed system which abides by the principals of electoral system.

References

[1] Tadayoshi Kohno, Adam Stubblefield, Aviel D. Rubin, Dan S. Wallach, "Analysis of an Electronic Voting System", Johns Hopkins University Information Security Institute Technical Report TR-2003-19, July 23, 2003.

[2] Michael R. Clarkson Stephen Chong Andrew C. Myers, "Civitas: Toward a Secure Voting System", Computing and Information Science Technical Report TR 2007-2081, May 2007.

[3] Bart Jacobs, Wolter Pieters, "Electronic Voting in the Netherlands: from early Adoption to early Abolishment", "Foundations of Security Analysis and Design", V: FOSAD 2007/2008/2009 Tutorial Lectures, Springer LNCS 5705, 2009, pp. 121-144.

[4] Eric A. Fischer, "Election Reform and Electronic Voting Systems (DREs): Analysis of Security Issues", "CRS Report for Congress", November 4, 2003.

[5] Jonathan Bannet, David W. Price, Algis Rudys, Justin Singer, Dan S. Wallach, "Hack-a-vote: Security issues with electronic voting systems", IEEE Security & Privacy, 2(1):32–37, Jan. 2004.

[6] Brennan Center for Justice, "The machinery of democracy: Voting system security, accessibility, usability and cost", New York University, Oct. 2006.

[7] Tadayoshi Kohno, Adam Stubblefield, Aviel D. Rubin, Dan S. Wallach, "Analysis of an electronic voting system", In Proc. of IEEE Symposium on Security and Privacy, pages 27–42, May 2004.

[8] David Wagner, Matthew Bishop, "Voting systems top-to-bottomreview",http://www.sos.ca.gov/elections/elections_vsr.htm, 2007.

[9] David L. Dill, Bruce Schneier, Barbara Simons, "Voting and technology: Who gets to count your vote?", Communications of the ACM, 46(8):29–31, Aug. 2003.

[10] David Evans, Nathanael Paul, "Election security: Perception and reality", IEEE Security & Privacy, 2(1):24–31, Jan. 2004.

[11] David Jefferson, Aviel D. Rubin, Barbara Simons, David Wagner, "A security analysis of the secure electronic registration and voting experiment (SERVE)", http://www.servesecurityreport.org/paper.pdf, Jan. 2004.

[12] Rebecca Mercuri, "Statement on electronic voting", http://www.notablesoftware.com/RMstatement.html, 2007.

[13] Aviel D. Rubin, "Security considerations for remote electronic voting", Communications of the ACM, 45(12):39–44, Dec. 2002.

[14] D. Bowen et al, "Top-to-Bottom Review of voting machines certified for use in California", Technical report, California Secretary of State, 2007.

[15] A. Goodnough, C. Drew, "Florida to shift voting system with paper trail", The New York Times, Feb. 2007.

[16] Minister Gormley announces Government decision to end electronic voting and counting project, http://www.environ.ie/en/LocalGovernment/Voting/News/MainBody,20056,en.htm, Apr. 2009.

[17] A. U. de Haes, "Dutch government bans electronic voting", IDG News Service, May 2008.

[18] Bundesverfassungsgericht, German Constitutional Court. Judgment 2 BvC 3/07, 2 BvC 4/07, official English translation.
http://www.bverfg.de/entscheidungen/rs200903032bvc000307en.html, Mar. 2009.

[19] Chowdhury M. Rahman, Md. Adnan Khan, "Study of SMS Security as part of an Electronic Voting System", Unpublished Thesis report, BRAC University, Bangladesh, 2006.

[20] Mohammad Shabbir Hasan, Abdullah Al Mahmood, Quazi Farhan, "A Roadmap towards the Implementation of an Efficient Online Voting System in Bangladesh", To be appeared in "International Conference on Computational Intelligence and Software Engineering", 2010, China.

[21] Kevin J. Coleman, Eric A. Fischer, "Elections Reform: Overview and Issues", "CRS Report for Congress", October 30, 2003.

[22] Melanie Volkamer, Rüdiger Grimm, "Multiple Casts in Online Voting: Analyzing Chances", "2nd International Workshop Co-organized by Council of Europe, ESF TED, IFIP WG 8.5 and E-Voting.CC", August 2006.

[23] S.J. Ackerman, "The Vote that Failed", "Smithsonian Magazine", November 1998.

[24] Eric A. Fischer, "Voting Technologies in the United States: Overview and Issues for Congress", "CRS Report for Congress", March 21, 2001.

[25] Hari K. Prasad, J. Alex Halderman, Rop Gonggrijp, Scott Wolchok, Eric Wustrow, Arun Kankipati, Sai Krishna Sakhamuri, Vasavya Yagati, "Security Analysis of India's Electronic Voting Machines", http://indiaevm.org/, April 29, 2010.

[26] Ed Gerck, in an interview by Eva Waskell, "California Internet Voting."The Bell, Vol. 1, No.6, ISSN 1530-048X, October 2000.

[27] Shannon, C., "A Mathematical Theory of Communication." Bell Syst. Tech. J., vol. 27, pp. 379-423, July 1948.

[28] C. R. Kasarbada, P. V. Indiresan, S. Sampath, "Report of the expert committee for technical evaluation of the electronic voting machine", Apr. 1990.

Characterization of QoS Based Routing Algorithms

Dr. Shuchita Upadhyaya[1] , Gaytri Devi[2]

[1] Department Of Computer Science And Applications, Kurukshetra University ,Kurukshetra
shuchita_bhasin@yahoo.com

[2]GVM Institute Of Technology And Management,
M.D. University , Sonipat
gayatri.dhingra1@gmail.Com

Abstract— In recent years, there has been an increasing demand for Internet based multimedia applications. The continuous demand for exchanging multimedia information over the Internet is calling for new networking services that are geared towards providing guarantee of service. This guarantee of service is specified in terms of quality of service (QoS) requirements like desired bandwidth, delay, variation in delay experienced by receiver(jitter),packet loss that can be tolerated, no of hops, cost of links etc. QoS routing is the selection of paths that satisfy the requirements of traffic in the networks. The problem to be solved by QoS routing algorithm is multi constrained path problem. In general, multi constrained path selection problem is NP-complete that cannot be exactly solved in polynomial time. So various types of heuristic and approximate algorithms with polynomial and pseudo polynomial complexities have been presented in literature to solve this problem. This paper presents & discusses the main approaches used to reduce QoS routing algorithm complexity and characterizes them on the basis of their category.

Keywords: Quality Of Service Routing, Heuristic, Approximate, Exact.

I. INTRODUCTION

Quality of Service (QoS) puts some restrictions in the form of certain constraints on the path. These constraints may be desired bandwidth, delay, variation in delay experienced by receiver(jitter),packet loss that can be tolerated, no of hops, cost of links etc.

QoS Constraints are represented in the form of metrics. One metric for each constraint is to be specified like bandwidth metric, jitter (variation in delay) metric, delay metric, no of hops metric, packet loss ratio etc. for one node to all other nodes in the network. Metric for a complete path with respect to each parameter is determined by the composition rules of metrics. The three basic rules are [21].-

i) Additive Metric: The value of that constraint for a path is the addition of all links constituting path. For Example- delay, hop count, cost, and jitter.

It can be represented as

$$D(p_i)=\sum_{e\varepsilon p_i}(d(e))$$

It means delay of path is sum of all its edges.

ii) Multiplicative Metric: Using this metric, the value for the complete path is multiplication of all its edges.

Examples are – reliability (1-lossratio) and error free Transmission (probability)

This can be represented as

$$R(p_i)=\prod_{e\varepsilon p_i}(r(e))$$

The reliability of the path is multiplication of all its edges. Multiplicative metric can be converted into additive by taking logarithm.

iii) Concave Metric: In this metric, either we can take min value or max value among all the edges for a path. For Example- Bandwidth

$$B(p)=\min/\max\,(b(e))$$

For a complete path, the constraints may be required either as a constrained form or in an optimization form. In constrained form, some condition is put on constraint value e.g. Choose that path only which has delay less than or equal to 60 ms. the path obeying the condition is called feasible. On the other hand optimization refers to path having minimum or maximum value for a constraint e.g. Choose the path that has minimum delay among all the paths. This path is called optimal path [1].

Based on these forms QoS routing is broadly classified into two categories .MCP Routing (Multiple constrained path) and MCOP Routing (Multiple constrained optimal path).Where In MCP ,the target is to find the feasible path satisfying multiple constraints, where as MCOP is a special case of MCP problem in which feasible path is found according to one of the constraints. Then from those optimal path is computed according to other constraint .Restricted Shortest Path (RSP) is a type of MCOP problem. Among all the multi constrained path routing problems RSP has received most attention.

A widely studied case of Restricted Shortest Path problem group is DCLC (Delay Constrained least cost) where the goal is to find the least cost path among those that satisfy delay constraint.

In this paper, we discuss various techniques to solve the QoS routing problem. We have characterized the various

QoS routing algorithms on the basis of their category, problem solving strategy, their complexity, number and types of constraints they can handle and their QoS category, We have covered only unicast algorithms.

The layout of paper is as follows: In section II various techniques to solve the QoS routing problem are presented, exact algorithms are discussed in section III, In Section-IV characterization table of algorithms is presented and section V provides summary and conclusion.

II. TECHNIQUES

In general, MCP and MCOP both are NP-complete in nature that can not be exactly solved in polynomial time. Here the idea is to find the solution that will complete in polynomial time .Hence the objective is to find the technique to reduce the computational complexity. To implement these technique, well known shortest path algorithms e.g. Dijkstra, Bellman-Ford algorithms have been used by most of the researchers. Since these algorithms only deal with single weight so these algorithms have been extended or modified to consider multiple constraints for solving QoS routing problem. In general, the techniques to solve NP-complete problem are Parameterization, Restriction, Heuristic, Approximation and Randomization.

A Parameterization

When certain parameters of input are fixed, then the solution can be found. The problem of path selection subject to multiple additive or multiplicative constraints is known to be NP-complete. But if one of constraints is concave and other is additive / multiplicative then problem can be solved in polynomial time. Concave metric is usually dealt with a preprocessing step called topology filtering where all links that do not satisfy constraints are pruned [20].

In general, QoS Constraints are independent and a well known result is that finding a path with (independent) delay & delay-jitter is NP-complete. But in practice these bounds are not independent but the functions of reserved bandwidth. So the problem of finding a path satisfying bandwidth, delay, delay-jitter and buffer-space constraints can be simplified by taking this relationship into consideration [14].

B. Restriction

The problem can be solved in polynomial time, if the structure of input are restricted .If the QoS metrics are real number or unbounded integer then their complexity is NP-Complete, If the metrics take bounded integer then their complexity is polynomial. Chen's algorithm [2] reduced the problem into simpler by converting real weight to integer weight and then applied extended Bellman-Ford and Dijkstra algorithms.

In literature, it has also been suggested that there may exist classes of graphs in which QoS routing is not NP-complete. Also when all the nodes have degree 2, it can be solved in polynomial time, irrespective of link weights [19].

C. Heuristic

A heuristic algorithm does not try to find the perfect solution but an approximate solution where the time or resources are limited. It is free from providing good run times and with provably good or optimal solution quality. Many Researchers have proposed heuristic algorithms which reduces the computational time but do not provide guarantee to find a feasible path even it exist .To find a heuristic, one major method used in literature is metric composition. Metric composition may be-

Linear, Non-linear, lagrange relaxation linear composition.

The combination of additive metrics using Linear composition has been proposed in [3][8].The link weights are computed through linear energy function, where each energy function is weighted sum of the link metrics. This approach is easy to implement but prevents provisioning the guarantee of considering all the constraints.

The second approach is lagrange relaxation linear composition technique. It is a common technique for calculating lower bound & finding good solutions. The basic idea is to first combine the two weights in terms of a parameter α to form an aggregate weight $w=w1+ \alpha w2$, then Dijkstra or Bellman-Ford algorithm is used to find the shortest path [10] [11][34].This approach overcomes the problem of linear composition. These algorithms are having very low time complexity.

The weights can be combined to form a single weight by using non-linear composition [5][13] . This approach can be applied to the metrics that are not correlated. Non – linear length function give higher success rate to find the feasible path than linear function. Korkmaz [13] proposed an algorithm H_MCOP that runs dijkstra algorithm twice: one in reverse direction with a linear cost function and second in forward direction with non linear cost function.

D. Approximation

Approximation algorithms are those heuristic that additionally provide some bounds on error. Ideally, the approximation is optimal up to a small constant factor. An approximation algorithm always returns a solution for a given input whose cost is within some additive or multiplicative factor of the cost of the optimal solution.

The approximate algorithm for MCP problem presented in literature delivers solution with in arbitrarily specified precision ε. An algorithm is said to be ε-optimal if it returns a path whose cost is at most $(1+\varepsilon)$ times the cost of optimal path where $\varepsilon>0$. The complexity of ε-approximate solutions depends on the actual value of link weights ,size of network and $1/\varepsilon$.These solutions are defined by first finding the lower and upper bound values by assuming some initial value and then systematically adjust these bounds using testing procedure. And then rounding and scaling is performed to bind the cost of every link. [7][23][30-32].

E. Randomization

The concept behind randomization is to make random decision during the execution of algorithm. The concept of randomness is used to avoid unforeseen traps when searching for a feasible path. These algorithms are simple and easy to implement but fail with some small probability. Randomized algorithm can balance network load, prevent performance degradation and improve service performance of entire network.[12][15].

III. EXACT ALGORITHM

The Exact solution of multi constrained path problem can be found by systematically examining every path between s & d in brute force manner. But the no of paths grows exponentially with the size of network. Some researches in literature have also proposed exact algorithms instead of defining approximate or heuristic algorithm. The exact algorithm of MCP problem is possible because-

1) NP-complete behavior seems only to occur in specially constructed graph, which are unlikely to occur in realistic communication networks.

2) There exist exact algorithms that are equally complex as heuristic and they do not induce NP-complete behavior.

3) By simply restricting the no of paths explored, the complexity can be decreased at the expose of possibly loosing exactness. [17]

The exact algorithms are constrained Bell-man ford algorithm (CBF), SAMCRA (self-adaptive multiple constraints routing algorithm), TAMCRA (Tunable accuracy multiple constraints routing algorithm) ,A*prune. CBF & A* prune algorithms presents exact solutions but their running time grows exponentially with the network size.

TAMCRA and SAMCRA are based on three fundamental concepts.

i) Non- linear path length measure:

The non linear length functions in more efficient than linear length function, as the curved lines match the constraints boundaries much better than straight lines.

ii) K-Shortest Path Approach:

K-shortest path approaches returns not only shortest path to given destination but also second shortest, third shortest.....Kth shortest path.

iii) Principal of Non Dominance:

A path P2 is said to be dominated by a path P1, if at least one of the weights of path p1 is less than the path p2.Exact algorithms only considers non-dominated paths.

A fourth concept has been added in SAMCRA i.e. Look ahead concept. Look ahead concept proposes to compute the shortest path tree rooted at destination. So the lowest value from destination to a node is stored in the queue of that node n. By using this information the set of possible path can further be limited.

IV. CHARACTERIZATION OF QOS ALGORITHMS

In this section, We have characterized the various QoS routing algorithms on the basis of their category, problem solving strategy, their complexity, number and types of constraints they can handle and their QoS category in a tabular form.

Table1: Characterization table

Algorithm	Problem solving strategy	Time Complexity	Metric	Algorithm strategy	Category
Parameterization					
WQF like scheduling algorithm[14]	The relationship of metrics is used as delay, delay jitter, buffer space are the functions of available bandwidth.	O(m .n)	Delay ,jitter , bandwidth	Bellman ford algorithm	MCP
Bandwidth delay constrained algorithm[22]	It eliminates all links that do not satisfy the bandwidth constraint and find the shortest path w.r.t delay among the remaining paths	$O(n^2)$	Bandwidth & delay	Dijkstra algorithm	MCP

Restriction					
EBF (Extended bellman ford Algorithm)[2]	Reduce the problem into simpler by converting real weight to integer weight then apply extended bellman- ford algorithm	O(xmn) x is an adjustable positive integer.	Delay & cost	Bellman ford algorithm	MCP
EDSP(Extended dijkstra algorithm)[2]	Reduce the problem into simpler by converting real weight to integer weight then apply extended dijkstra.	$O(x^2 n^2)$	Delay & cost	Dijkstra algorithm	MCP
Heuristics algorithms					
Linear composition					
LEFPA (Linear energy function precomputaion algorithm)[3]	Converts two additive weights to a single metric with linear energy functions	O(B (m+n +n logn)) B=No of LEFs	Any two	Dijkstra algorithm	MCP
JAA(jaffe's approx algorthim)[8]	It linearly combines two weights	$O(n^5 b \log nb)$ b=largest weight in the graph	Any two	Dijkstra algorithm	MCP
LPH (limited path heuristic)[33]	Maintains k best paths at each node according to linear combination of weights using linear equation	$O(k^2 v^2)$ K= no of paths	Any two	Dijkstra algorithm	MCP
Lagranges relaxation Linear compositions					
LARAC(lagrange relaxation based aggregate cost)[10]	It uses the concept of aggregated costs and provides an efficient method to find the optimal multiplier based on Lagrange relaxation	$O(m^2 \log^4 m)$	Delay &cost	Dijkstra algorithm	RSP
MCP-IA (Multi constrained path-Iterative algorithm)[34]	Uses iterative procedure to find the appropriate value of α for constructing the mixed weight	$O(k N^2)$ k= No of executions of dijkstra algorithm	Delay & cost	Dijkstra algorithm	MCP
Korkmaz's BSLR algorithm(Binary search for linear relaxation) [11]	It Uses a refined Lagrange relaxation technique to define the weights of metric composition rule. It performs binary search to minimize the linear cost function that is guarantee to terminate with in logarithm no of calls to dijkstra algorithm.	O(log B(m+n log n) B=upper bound	Any two	Dijkstra algorithm	MCP
Non-Linear					

H_MCOP(Heuristic algorithm for multi constrained optimal path)[13]	It runs dijkstra algorithm twice: one in reverse direction with a linear cost function and second in forward direction with non linear cost function.	O(n log n + km log kn + k^2 +1)m) K=no of paths	K	Dijkstra algorithm	MCOP
DCCR(Delay cost constrained routing) [5]	It uses non linear length function and k shortest path algorithm	O(ke log (kn) + k^2 e+ t(A)) K=no of paths	Delay ,cost	-	RSP
SSR+DCCR(search space reduction)[5]	It uses K shortest path algorithm and a new adaptive path weight function together with an additional constraint imposed on path cost to restrict the search space.	O((m(G)+2) e log n)+O(ke log(kn) +k^2 e) G= Total no of iterations of algorithm	Delay & cost	Any shortest path algorithm	RSP

Approximate algorithms					
Hassin's algorithm First FPAS(fully polynomial approximation schemes) Second FPAS[7]	This is a combination of dynamic programming & scaling/rounding & is applicable to acyclic graph Ist algorithm initially starts with LB=1 And UB equals to sum of n-1 largest link costs & systematically adjust these using testing procedure. Second algorithm is a basically extension of ist and uses a technique called Interval partitioning.	O(log logB(m(n/ε)+log log B)) O(m(n^2/ ε)log(n/ ε)	delay & cost Delay & cost	- -	RSP RSP
SEA (Simple Efficient approximation [23]	It computes lower & upper bound using binary search & the Run modified Hassin 's algorithm & it is applicable to general graph	O(mn(loglogn +1/ ε))	delay & cost	-	RSP
FPTAS-DCLC(Fully polynomial time approximation scheme for delay constrained least cost)[31] FPTAS-OMCP(Fully polynomial time approximation scheme for optimization of Multi constrained path) [31]	It is based on the scheme which uses a novel combination of techniques of Hassin & SFPAS. It enforce one constraint and approximate k-1 constraints	O(mn logloglogn+mn/ε) O(mnlogloglog n+m(n/ε)$^{k-1}$)	delay & cost k	Dijkstra algorithm Dijkstra algorithm	RSP MCP
APPROX[32]	It is based on transformation of network for approximation	O((n/ε + logD)1/ε loglogu) u-upper bound d-delay constraint	Delay & cost	-	RSP
K-Approx FPTAS-SMCP [30]	It approximate all k constraint without enforcing any one constraint	O(km+mn) O(m(n/ ε)$^{k-1}$)	k k	Bell man ford Algorithm Dijkstra Algorithm	Optimization version of MCP
ε-OPQR(optimal QoS partition & rounding) [26]	It uses dynamic programming algorithm & presents an approximation technique based on sampling & scaling.	O(m log D + nlogn)loglog b+m logn(log D +n)loglogn +m/ ε log n+ log D + nx) b- ratio of intial bounds D-delay constraint	delay & cost	Dijkstra algorithm	RSP

Randomized algorithms					
Randomized Algorithm[12]	Firstly, the algorithm computes shortest path from every node u to destination and then Randomized BFS discovers those nodes from which there are a chance to go final destination node.	$O(n^3)$	Multiple constraints	Dijkstra algorithm	MCP
Fast QoSRoute Algorithm[15]	Firstly, prunes all the links satisfying one bandwidth and then make the list of candidate paths satisfying delay constraint and then selects one path from computed candidate paths.	O(m+n log n)	Bandwidth & delay	Dijkstra algorithm	MCP
Randomized version of Chen's algorithm[33]	It uses a random method to choose value of x so that algorithm can achieve better performance	O(Txm+Tx nlog(xn))	Any two	Dijkstra algorithm	MCP
Exact algorithms					
Constrained Bell man ford algorithm(CBF)[18]	It maintains a list of paths ordering in increasing cost and decreasing delay using breadth first search & selects paths that satisfies delay constraint and has minimum cost.	O(ΔE) Δ- delay constraint E- edges	Delay & cost	Bell man ford algorithm	RSP
TAMCRA[4]	Concept of non-linear path length, k-shortest path approach, principal of non-dominance	$O(kn \log kn + k^3 + xm)$ X is fixed	Multiple	Dijkstra Algorithm	MCP
SAMCRA[17]	Non-linear path length, k-shortest path, principal of non-dominance, & look-ahead concept	$O(kn \log(kn) + k^2 xm)$ x is fixed	Multiple	Dijkstra Algorithm	MCP
A* prune[28]	It has combined A* search algorithm with a proper pruning technique. The algorithm constructs paths starting at source and going towards destination. But at each iteration, the algorithm get rid of all the paths that are guarantee to violate the constraints there by keeping only those partial paths that have potential to turned into feasible paths from which the optimal paths are drawn.	O(Q N(m+N + log Q)) Q=no of expanded paths	Multiple	-	MCSP(multi constrained shortest path)

m- no of edges

n- no of vertices

ε is approximation factor that reflects how far the solution is from optimal one

V. SUMMARY AND CONCLUSION

In general, searching a route satisfying multiple QoS constraints to support multimedia applications is known to be NP-Complete problem. So mostly heuristics algorithm were proposed for NP-complete problem which close to optimal results and reducing the complexity of path computation problem. Heuristic either imposes relationships among the link metrics to reduce the complexity of the problem which may limit the general applicability of the heuristic or too costly in terms of execution time to be applicable to large networks or too complex in terms of execution time. Heuristic algorithms are fast but are not efficient to provide optimal solution with reasonable probability. The best heuristic algorithm is H_MCOP algorithm. H_MCOP can outperform almost all known heuristic algorithms in terms of success ratio of finding feasible solution. The success ratio of H_MCOP is actually very close to that of an exact algorithm.

Approximate algorithms deliver solution with in arbitrary specified precision. They are very efficient but having very high time complexities thus are very slow unfortunately, in practical cases, the running time of these methods for sufficiently small ε will be worse which makes these results rather theoretical.

Randomized algorithms are useful when networks are having inaccurate or dynamic state. Randomized algorithm can balance network load, prevent performance degradation and improve service performance of entire network but some times fail with small probability.

This Multi constrained path selection problem is not NP-complete in strong sense. The NP-completeness of this problem depends on underlying topology, link weights, value of constraints. So exact algorithm have also been proposed by researchers. Thus the future researches should focus to differentiate the cases for which the complexity is polynomial so that exact algorithms may be refined further.

REFERENCES

[1] S. Chen , K. Nahrstedt, "An overview of quality-of-service routing for the next generation high-speed networks: Problems and solutions", IEEE Network, vol. 12, no. 6, pp. 64–79, Nov-Dec 1998.

[2] S. Chen and K. Nahrstedt, "On finding multi-constrained paths, in Proceedings of the ICC '98 Conference". IEEE, 1998, pp. 874 –879.

[3] Y.Cui,J.Wu ,"Pre computation for multi constrained QoS routing in high speed networks", Proceedings in IEEE infocom

.[4] H. De Neve and P. Van Mieghem, "TAMCRA: A Tunable Accuracy Multiple Constraints Routing Algorithm", Com-puter Communications, 2000, vol. 23, pp. 667-679.

[5] L. Guo and I.Matta, "Search space reduction in QoS routing", in Proceedings of the 19th IEEE International Conference on Distributed ComputingSystems. IEEE, May 1999, pp. 142 – 149.

[6] R. Guerin and A. Orda," QoS routing in networks with inaccurate information: Theory and algorithms", IEEE/ACM Transactions on Networking, vol. 7, no. 3, pp. 350 –364, June 1999.

[7] R. Hassin, "Approximation schemes for the restricted shortest path problem", Mathematics of Operations Research, vol. 17, no. 1, pp. 36–42,1992.

[8] J. M. Jaffe, "Algorithms for finding paths with multiple constraints", Networks, vol. 14, pp. 95–116, 1984.

[9] W..Jianxin,W.Weiping,c.jianer ,"A Randomized Qos Routing algorithm on networks with inaccurate Link state information ", Communication Technology proceedings 2000.

[10] A. Juttner, B. Szviatovszki, I. Mecs, and Z. Rajko, "Lagrange relaxation based method for the QoS routing problem," in Proc. IEEE INFOCOM, vol. 2, Apr. 2001, pp. 859–868.

[11] T. Korkmaz, M. Krunz, and S. Tragoudas, "An efficient algorithm for finding a path subject to two additive constraints", in Proceedings of the ACM SIGMETRICS '00 Conference, June 2000, vol. 1, pp. 318–327.

[12] T. Korkmaz and M. Krunz, "A randomized algorithm for finding a path subject to multiple QoS constraints", in Proceedings of the IEEE GLOBECOM '99 Conference - Symposium on Global Internet: Application andTechnology, Dec. 1999, vol. 3, pp. 1694 – 1698.

[13] T. Korkmaz and M. Krunz, "Multi-Constrained Optimal Path Selection", IEEE INFOCOM 2001.

[14] Q.Ma and P.Steenkiste, "Quality of service routing with performance guarantees". Proc. Of 4th Int. IFIP workshop on QOS May 1997.

[15] H-k.Park, y.Heo , "A Fast randomized Qos routing algorithm"

[16] S. Vutukury and J. J. Garcia-Luna-Aceves,"A simple approximation to minimum-delay routing", in Proc. ACM SIGCOMM'99, Boston, MA, Aug. 1999, pp. 227–238.

[17] P.Van.Miegham and F.A.Kupier, "Concepts of Exact QoS routing algorithms" .IEEE/ACM Transaction on networking , vol.12,no.5 ,oct. 2004

[18] R. Widyono," The design and evaluation of routing algorithms for real time channels"', Tech. Rep. TR-94-024, University of California at Berkeley & International Computer Science Institute, June 1994.

[19] Z. Wang, "On the complexity of quality of service routing", Information Processing Letters, vol. 69, no. 3, pp. 111–114, 1999.

[20] Z. Wang and J. Crowcroft," Quality-of-Service Routing for Supporting Multimedia Applications", IEEE JSAC, vol. 14, no. 7, pp. 1228-1234, September, 1996.

[21] X. Xiao and L. M. Ni," Internet QoS: A big picture", IEEE Network, vol.13, no. 2, pp. 8–18, March-April 1999.

[22] Z.Wang and J. Crowcraft ," bandwidth –dealy based routing algorithms",IEEE Globecom'95.singapur,november 1995.

[23] D. H. Lorenz and D. Raz, "A simple efficient approximation scheme for the restricted shortest path problem," Oper. Res. Lett., vol. 28, pp.

213–219, 2001.

[25] D. Lorenz and A. Orda, "Optimal partition of QoS requirements on unicast paths and multicast trees," in IEEE Infocom, 1999.

[26] D.H. Lorenz, A. Orda, D. Raz, and Y. Shavitt. "Efficient QoS Partition and Routing of Unicast and Multicast." In Proceedings IEEE/IFIP

IWQoS, Pittsburgh, PA, June 2000.

[27] Q. Ma and P. Steenkiste, "On Path Selection for Traffic with Bandwidth Guarantees", Proceedings of the IEEE International Conference on Network Protocols (ICNP .97), Atlanta, Georgia, pp. 191-202, October, 1997.

[28]G. Liu and K.G. Ramakrishnan, "A*Prune: An Algorithm for Finding K Shortest Paths Subject to Multiple Con-straints", IEEE INFOCOM 2001.

[29] A. Orda, "Routing with end-to-end QoS guarantees in broadband networks," IEEE/ACM Trans. Netw., vol. 7, no. 3, pp. 365–374, Jun.

1999.

[30] G.Xue ,W.zhang,,j.Tang "Finding a path subject to many additive constraints" IEEE/ACM Trans. Netw., vol. 15, no. 1, feb 2007.

[31] G.Xue,W.zang,J.Tang "Polynomial Time approximation algorithms for multi constrained QoS routing" IEEE/ACM Trans. Netw., vol.10, no. 4, Jun.2008.

[32] F. Ergun, R. Sinha and L. Zhang, "QoS routing with performance-dependent costs", Proceedings of INFOCOM 2000,vol. 1, pp. 137-146, 2000.

[33] J.Chen "Efficient Heuristic Algorithms for finding Multi-constrained paths" IEEE,vol 0, Feb,2002.

[34] G. Feng, K. Makki , N.Pissinou , C. Douligeris ,"An Efficient approximate Algorithm for Delay-Cost-Constrained QoS Routing"

AUTHORS PROFILE

Dr. Shuchita Upadhyaya
Is an associate professor in the Department of Computer Science and applications in Kurukshetra University, India. She is Ph.D. in Computer Science and application. Her area Of specialization is Computer Network. She has published more than 35 research papers in various national and international Journals. She is having 22 years of teaching and research experience.

Gaytri Devi
Is an assistant professor in GVM Institute of Technology and Management , Sonipat, India.She has done MCA , M.Phil. Now she is pursuing Ph.D. from the department of Computer Science and applications, Kurukshetra University. She is having more than 8 years of teaching experience.

www.sprinterglobal.org/ijcset