Information & Security: An International Journal

Sharma & Ojha, vol.26:2, 2013, 49-60 http://dx.doi.org/10.11610/isij.2605

BIOMETRIC TEMPLATE SECURITY USING CODE BASE CRYPTOSYSTEM

Ajay SHARMA and Deo Brat OJHA

Abstract: This paper presents an enhancement of the accuracy and security of biometric templates based on a code-based cryptosystem—McEliece cryptosystem—which in addition to randomness is also probabilistic, which provides higher susceptibility of templates towards brute force attacks. It is possible to generate many different secure biometric templates for the same system and also unique biometric templates for multiple systems from the same biometric trait; it is just a matter of using a different error vector. It is also easy to cancel a secure template by simply deleting the compromised template and generating a new one by using different error vector.

Keywords: Cryptography, fuzzy commitment scheme, biometric system, template, algorithmic noise, enrolment phase.

Introduction

With the growing use of biometric recognition systems cryptography is considered as one of the fundamental building blocks in the protection of biometric data. Biometric provides a person with a distinct characteristic that is always prevalent. It is a technique of authentication of a person's individuality from one or more behavioral or physiological features.[1] The use of biometrics (e.g., fingerprints, irises, faces) for recognizing individuals is becoming increasingly popular and many applications are already available. Although these applications can be fundamentally different, they can still be grouped into one of two categories: verification and identification.[2,3,4]

A well-known difficulty has been how to cope with the 10 to 20 % of error bits within a biometric data and derive an error-free template. It is fundamentally impossible to avoid noise during biometric data acquisition, because "life means change." For example, faces age and iris patterns are not perfectly invariant to a contraction of a pupil. More noise is introduced by changes in the environmental conditions, which is again an unavoidable circumstance. Finally, noise often finds its way into the sensor during transmission or in the processing of data (the so called "algorithmic noise"). The latter noise sources can be reduced or even removed by improved engineering.

To solve this problem, fuzzy commitment schemes play an important role. The fuzzy commitment scheme is a tool for handling the noise in template of a biometric recognition system. Juels and Wattenberg's fuzzy commitment scheme[5] has been introduced to handle the difference occurring between two captured sets of biometric data using error correcting code.

Various approaches have been proposed to protect the stored template. Some are hardware based and use stand-alone biometric system-on-devices. Others are software based and rely on feature transformation and biometric cryptosystems. Because of interclass variation in the biometric template, common encryption techniques, such as AES (Advance Encryption Standard) or RSA, can not be used in biometric cryptosystems.

This paper defines an application of a fuzzy commitment scheme with the McEliece's cipher.[6] The main idea is to transform the biometric matching problem into an error correcting issue. We carefully studied the error patterns within biometric data, and devised a two-layer error correction technique that combines Hamming code and Goppa code. The error-correcting methods remove noise in the template.[7] Along with accuracy, we suggest an enhancement in the privacy of the biometric cryptosystem, since common encryption techniques such as AES or RSA cannot be used, so that the auxiliary data can be masked by using homomorphic encryption that allows certain arithmetic operation in the encryption domain.[8]

Preliminaries

Biometric System

A generic biometric system consists of five components: Sensor, feature extractor, template database, matcher, and decision module. In general, a biometric based recognition system works in two phases. In the enrolment phase, the biometric template b are processed from a user U and stored or registered in the database. The second phase is the verification phase, when the system captures a new biometric sample b' from U, and then compares it to the registered or reference data via a matching function. Let μ be the biometric measure of U and τ is a recognition threshold, b' will be accepted if $\mu(b,b') \le \tau$, else rejected. Two main kinds of errors are associated to this scheme: False Reject (**FR**), when a matching user, i.e. a legitimate user, is rejected; False Acceptance (**FA**), when a non-matching one, e.g. an impostor, is accepted. Note that, when the threshold increases, the **FR**'s rate (**FRR**) decreases while the **FA**'s rate (**FAR**) grows, and wise versa.[9]

Definitions

A metric space is a set C with a distance function dist: $C \times C \rightarrow R^{+} = [0, \infty)$, which obeys the usual properties (symmetric, triangle inequalities, zero distance between equal points).[10]

Let $C\{0,1\}^{n}$ be a code set which consists of a set of code words c_i of length n. The distance metric between any two code words c_i and c_j in C is defined by

$$dist(c_i, c_j) = \sum_{r=1}^{n} \left| c_{ir} - c_{jr} \right| \qquad c_i, c_j \in C\,.$$

This is known as Hamming distance.

An error correction function f for a code C is defined as $f(c_i) = \{c_j \,/\, dist(c_i, c_j) \text{ is the minimum, over } C - \{c_i\}\}$.

Here, $c_j = f\left(c_i\right)$ is called the nearest neighbour of c_i.[11]

The measurement of nearness between two code words c and c' is defined by $\text{nearness}\,(c, c') = dist(c, c')/n$, it is obvious that $0 \le \text{nearness}\ \ (\mathrm{c}, \mathrm{c}') \le 1$.

The fuzzy membership function for a codeword c' to be equal to a given c is defined as

$$\begin{aligned} FUZZ(c') &= 0 && \text{if nearness}(\mathrm{c},\mathrm{c}') = \mathrm{z} \le \mathrm{z}_0 < 1 \\ &= \mathrm{z} && \text{otherwise} \end{aligned}$$

Fuzzy Commitment Scheme with McEliece scheme

Protocols are essentially a set of rules associated with a process or a scheme defining the process. Commitment protocols were first introduced by Blum.[12] Moreover, in the conventional commitment schemes, opening keys are required to enable the sender to prove the commitment. However, there could be many instances where the transmission involves noise or minor errors arising purely because of the factors over which neither sender nor the receiver have any control, which creates uncertainties. The fuzzy commitment scheme was first introduced by Juels and Martin. The new property “fuzziness” in the open phase was introduced to allow acceptance of the commitment using corrupted opening key that is close to the original one in an appropriate metric or distance. The fuzzy commitment scheme is based on hash function which causes them to share two shortcomings:

1. The hash functions used should be strongly collision free. However, this property can only be empirically checked. It actually turns out that some schemes are inadvertently based on weakly collision-free hash functions.

2. Hash functions alone cannot offer non-repudiability.

Here we use the speed of McEliece and its randomness to enhance the fuzzy commitment scheme by using code base cryptosystem which is based on Goppa Code.

First select secret key W is a random $(k \times k)$ nonsingular matrix over $GF(2)$ called the scrambling matrix, T is a $(k \times n)$ generator matrix of a binary Goppa code T with the capability of correcting n–bit random error vector of weight less than or equal to α, and Q is a random $(n \times n)$ permutation matrix.

Public Key: $V = WTQ$

A tuple $\{P, H, M, f\}$ where $M \subseteq \{0,1\}^k$ is a message set which consider as a code, P is a set of individuals, generally with three elements A as the committing party, B as the party to which commitment is made and TC as the trusted party, f is error correction function and $H = \{t_i, a_i\}$ are called the events occurring at times $t_i, i = 0,1,2$, as per algorithm $a_i, i = 0,1,2$. The scheme always culminates in either acceptance or rejection by A and B.

In the setup phase, the environment is set up initially and public commitment key CK generated, according to the algorithm $setupa\lg(a_0)$ and published to the parties A and B at time t_0. During the commit phase, Alice commits to a message $m \in M$; then she finds $g : m \to mV$.

Encryption: $E_V(m) = mV + e$, where m is the k-bit message, $E_V(m)$ is an n-bit cipher text and e is an n-bit random error vector of weight α.

According to the algorithms $commita\lg(e_1)$ into string c, i.e. her commitment $c = commita\lg(XOR, g(m), E_V(m))$, then after Alice sends c to Bob, which Bob will receive as $t(c)$, where t_f is the transmission function which includes noise.

In the open phase, Alice sends the procedure for revealing the hidden commitment at time t_2 and Bob uses this.

So Alice discloses the procedure $g(m)$ and E to Bob to open the commitment.

$opena\lg(e_2)$: Bob constructs c' using $commita\lg$, message $t(m)$ and opening key, i.e $c' = commita\lg(XOR, t_f(g(m)), t_f(E_V(m)))$, and checks whether the result is same as the received commitment $t(c)$.

Fuzzy decision making

$if\,(nearness(t_f(c), f(c')) \leq Z_0)$

Then A is bound to act as in m

Else he is free not to act as m.

Then after acceptance, Bob calculates $f(c')(WTQ)^{-1}$ and finally gets the message.

Related work

Our work is inspired from a number of authors who combine well known techniques from the area of error correcting code and cryptography to achieve an improved type of cryptographic primitive.[13] Further, numerous works have suggested a combination of biometrics and cryptography and provide details on the related research.[14,15,16]

Proposed System Architecture

In general, the identity theft problem is drastically exacerbated for the biometric systems. The proposed architecture of the biometric system will have enhanced security and accuracy with respect to traditional systems, achieved by combined usage of code base cryptosystem and error correcting code.

In the Enrolment stage (Figure 1) of a typical biometric recognition system, after the biometric acquisition module, some processing is applied in order to obtain the biometric template *b* which is then stored in a database. Here H is called the hamming space of length N, e.g. $\mathrm{H} = \{0,1\}^N = F_2^N$, where $F_2 = \{0,1\}$. However, the biometric data is never stored in the database to prevent it from being stolen. Instead, after the biometric has been acquired and the biometric template has been generated, a cryptographic function will be applied to it. The result of this operation will then be stored in the database; this will be referred to in the remaining text of this paper as the secure biometric template. It should be pointed out that it is impossible to recover any biometric data from this secure template as the cryptographic function is not invertible.

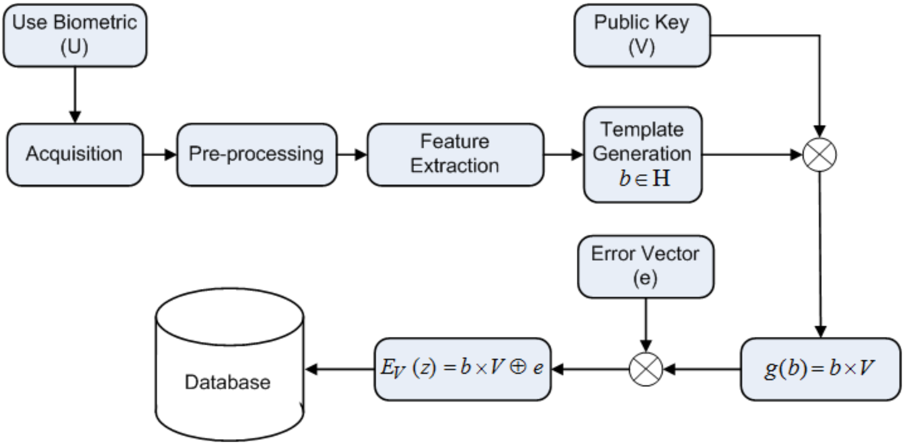

Figure 1: Enrolment Phase.

During the verification stage (Figure 2), the biometric probe is acquired and the corresponding template, b', is generated. The problem here is that b is not stored in the database, but only an encrypted version of it. The original biometric template b is recovered from the database if the user is who he claims, or something completely different if he is not. Therefore, the output of the feature extractor b' needs to be encrypted. Only then is the result compared to the encrypted that is stored in the database. If the $E_V(z')$ and $E_V(z)$ are equal, then the user is validated to be who he claims to be. With this system, the three requirements above are verified. In particular, it is possible to generate many different secure biometric templates from the same biometric trait; it is just a matter of using a different set of error vector (e). It is also easy to cancel a secure template by simply deleting the compromised template and generating a new one by using different error vector (e). Finally, since the biometric data is never stored in a database, this guarantees that this information remains private.

Acquisition

The acquisition module, absolutely necessary in a real biometric verification system, has not been implemented but here, instead of implementation, it is replaced by a large database of iris images, like the one developed by the Chinese Academy of Sciences' Institute of Automation (CASIA)[17] and code from Masek and Kovesi.[18] This database consists of 22051 iris images from more than 700 subjects. All iris images are 8 bit gray-level JPEG files, collected under near infrared illumination.

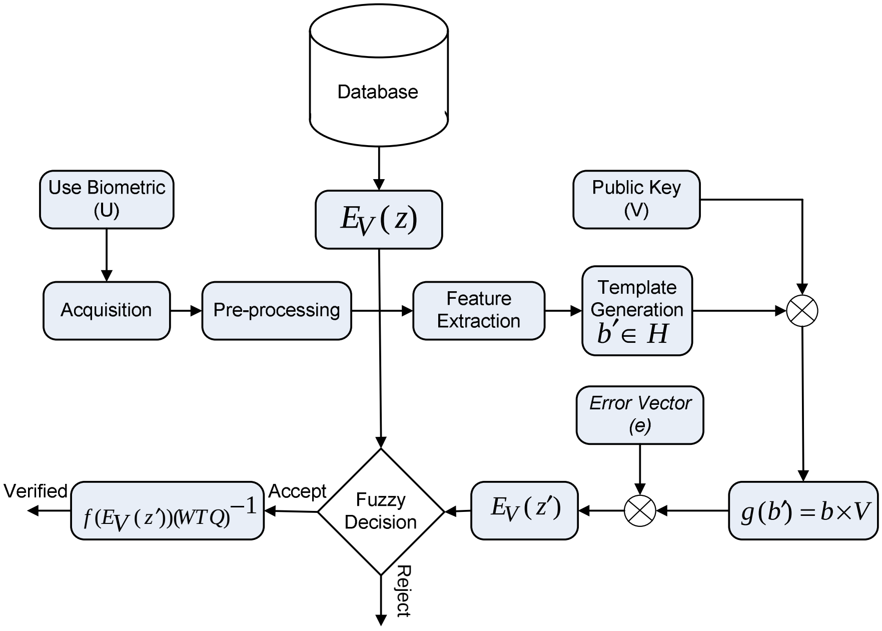

Figure 2: Verification Phase.

Pre-processing

After acquisition this step is to extract the iris from the input eye images. The iris area is considered as a circular crown limited by two circles. The iris inner (pupillary) and outer (scleric) circles are detected by applying the circular Hough transform,[19] relying on edge detection information previously computed using a modified Canny edge detection algorithm.[20] The eyelids often occlude part of the iris, thus being removed using a linear Hough transform.[21] The presence of eyelashes is identified using a simple threshold technique.

Feature Extraction

Once the iris texture is available, features are extracted from it to generate a more compact representation, also called the biometric template. (Readers may refer to Daugman on details of how iris recognition works.[22]) To extract this representation, the two-dimensional normalized iris pattern is convolved with a Log-Gabor wavelet. The resulting phase information is quantized, using two bits per pixel. The resulting iris template is composed of 9600 bits, stored as a 20×480 binary matrix.

Privacy-Protection and Error-Correction

This is the main module of this scheme, using McEliece cryptosystem, which adds some random error at the time of encryption that makes the original template more secure than poorly chosen passwords and other cryptosystems due to its randomness.

In the phase of enrolment, inputs are biometric template (b), error vector (e) which is chosen randomly and public key(V) which has generating a matrix that defines an error correcting code. The output of this phase is an encrypted template which is stored on the system or on a data card (i.e. smart card). Now, it is not easy to gain the template from this data without the knowledge of key and error vector.

In the phase of verification, a similar procedure is used with a newly acquired template b' with same error vector, and key and error correction coding is used to correct biometric templates. In this stage, the probe template of a legitimate user is (error) corrected in order to recover the original template, obtained during enrolment; this should be possible because both templates are fairly similar. However, for a illegitimate user, whose probe template is fairly different from the one originally enrolled by the legitimate user, it should not be possible to recover the original from the probe template.

Therefore, the selected error correcting code should be strong enough to correct templates of legitimate users, but not so strong as to also correct the templates of illegitimate users. Therefore, μ be the biometric measure of U and τ is a recognition threshold, b' will be accepted if $\mu(b,b') \le \tau$, else rejected.

Security Analysis

The accuracy of any biometric system depends on the ability of that system to separate genuine users from impostors. Here we describe a possible attack to the scheme and identify ways of preventing it. It is possible for an attacker to imitate a signer by obtaining a copy of their biometric data, e.g. by some method of duplicating fingerprints.[23] After obtaining a copy of the signer’s biometric data, the attacker can sign a forged message that will appear genuine on verification by the signer. To prevent this attack, genuine messages can be signed in the presence of a trusted witness.

Some issues of security in stored templates under consideration are:

(1) Stored Template should not reveal any data and no close replica made from the stored data.

(2) Multiple systems using the same biometric information should not be able to link templates corresponding to the same individual.

(3) If the stored data is compromised, remove that one and reissue a new one.

The proposed scheme provides the following solutions to these issues:

Explanation of Issue 1

Here we use Goppa code in McEliece by first encrypting a user biometric template and adding at the time of encryption an error vector of fixed weight α. To reveal any template, an attacker should know the solution of decoding problem for unknown weight α of error vector which is very hard to solve. The coding theory based cryptosystem is secure because decoding is hard without the knowledge of a secret.

Explanation of Issue 2

Let’s consider error vector as $e = g(R)$, where g is an invertible function which maps R into an n-bit error vector of weight α, $R = Id_A \square r$ and Id_A is machine identification and r is secret pseudo random vector. Since each system has unique Id_A so same biometric information should not be able to link templates corresponding to the same individual.

Explanation of issue 3

It is possible to generate many different secure biometric templates from the same biometric trait; it is just a matter of using a different set of error vectors. It is also easy to cancel a secure template by simply deleting the compromised template and generating a new one by using different error vector (e).

In this scheme, adequately chosen biometrics have higher entropy than poorly chosen passwords and, therefore, are less susceptible to brute force attacks. Template contains strong correlation, i.e. their bits are not independent from each other. The attacker can, therefore, create a large number of “artificial” templates on a computer.

Further, in this scheme we are using McEliece cryptosystem, which adds some random error at the time of encryption that makes the original template more secure than poorly chosen passwords and other cryptosystem due to its randomness. In addition to randomness, the McEliece cryptosystem is also probabilistic which gives more susceptibility of template towards brute force attacks.

It also provides non-repudiation, i.e. a legitimate user may access the facilities offered by an application and then not to claim that an intruder had circumvented the system. A bank clerk, for example, may modify the financial records of a customer and then cannot deny responsibility by claiming that an intruder could have possibly stolen her or his biometric data. So our proposed scheme enhances the biometric security and accuracy from the previous approaches presented in available literature.

Conclusion

Using a public key cryptosystem to construct a commitment is a way to achieve non-repudiability and authentication, a property which can not be offered by Hash functions alone. By using McEliece in a fuzzy commitment scheme, error vector e used to enhance the security of the function hiding, particularly against matrix factorization attacks. Main enhancement in this approach is randomness of the error vector – we can not obtain any information about the positions in which the error occurs. Thus the information rate is increased and information leakage rate decreased. As soon as identical templates are stored in multiple databases or datasets, it is possible to perform cross matching between them. The randomness property of the error vector is also required to prevent cross-matching of subjects across databases.

References:

[1] Sunil V.K. Gaddam, Manohar Lal, "Efficient Cancellable Biometric Key Generation Scheme for Cryptography," *International Journal of Network Security* 11, no.2 (September 2010): 61–69.

[2] A. K. Jain, S. Pankanti, S. Prabhakar, L. Hong, A. Ross, "Biometrics: A Grand Challenge," in *Proceedings of the International Conference on Pattern Recognition*, vol. 2 (August 2004), 935–942.

[3] James L. Wayman, Anil K. Jain, Davide Maltoni, and Dario Maio, *Biometric Systems: Technology, Design and Performance Evaluation* (London: Springer-Verlag, 2005), http://dx.doi.org/10.1007/b138151.

[4] Davide Maltoni, Dario Maio, Anil K. Jain, and Salil Prabhakar, *Handbook of Fingerprint Recognition* (New York, NY: Springer, 2005).

[5] Ari Juels and Martin Wattenberg, "A Fuzzy Commitment Scheme," in *Proceedings of the 6th ACM Conference on Computer and Communication Security* (November 1999), 28-36. A 2013 update is available at www.arijuels.com/wp-content/uploads/2013/09/JW99.pdf.

[6] Deo Brat Ojha and Ajay Sharma, "A Fuzzy Commitment Scheme with McEliece's Cipher," *Survey in Mathematics and Its Application* 5 (2010): 73-83.

[7] John Daugman, "How Iris Recognition Works," *IEEE Transactions on Circuits and Systems for Video Technology* 14, no. 1 (2004): 23-30.

[8] Julien Bringer and Hervé Chabanne, "An Authentication Protocol with Encrypted Biometric Data," *Progress in Cryptology–AFRICACRYPT 2008*, LNCS 5023 (Berlin: Springer, 2008), 109-124, http://dx.doi.org/10.1007/978-3-540-68164-9_8.

[9] Andrew Burnett, Adam Duffy, and Tom Dowling, "A Biometric Identity Based Signature Scheme," *International Journal of Network Security* 5, no.3 (November 2007): 317–326, http://eprint.iacr.org/2004/176.pdf.

[10] Alawi A. Al-saggaf and H.S. Acharya, "A Fuzzy Commitment Scheme," *IEEE International Conference on Advances in Computer Vision and Information Technology*, 28-30 November 2007, http://arxiv.org/abs/0809.1318.

[11] Vera Pless, *Introduction to the Theory of Error-Correcting Codes*, Third edition (New York, NY: Wiley, 1998).

[12] Manuel Blum, “Coin Flipping by Telephone: A Protocol for Solving Impossible Problems,” in *24th IEEE Computer Society International Conference*, San Francisco, CA, 1982, 133-137.

[13] F.J. MacWilliams and N.J.A. Sloane, *The Theory of Error-Correcting Codes* (Amsterdam: North Holland, 1991).

[14] Feng Hao, Ross Anderson, and John Daugman, “Combining Crypto with Biometrics Effectively,” *IEEE Transactions on Computers* 55, no. 9 (2006): 1081–88, http://dx.doi.org/10.1109/TC.2006.138.

[15] Ann Cavoukian and Alex Stoianov, “Biometric Encryption: A Positive-sum Technology that Achieves Strong Authentication, Security and Privacy,” White Paper (Information and Privacy Commissioner of Ontario, March 2007), www.ipc.on.ca/images/resources/bio-encryp.pdf.

[16] Emine Krichen, Bernadette Dorizzi, Zhenan Sun, and Sonia Garcia-Salicetti, “Iris Recognition,” in *Guide to Biometric Reference Systems and Performance Evaluation*, ed. Dijana Petrovska-Delacrétaz, Gérard Chollet, and Bernadette Dorizzi (London: Springer, 2009), 25–49, http://dx.doi.org/10.1007/978-1-84800-292-0_3.

[17] CASIA website, www.cbsr.ia.ac.cn/IrisDatabase.htm.

[18] Libor Masek and Peter Kovesi, *MATLAB Source Code for a Biometric Identification System Based on Iris Patterns* (School of Computer Science and Software Engineering, University of Western Australia, 2003), http://www.csse.uwa.edu.au/~pk/studentprojects/libor/sourcecode.html.

[19] T. Kawaguchi, D. Hidaka, M. Rizon, “Detection of Eyes from Human Faces by Hough Transform and Separability Filter,” *Proceedings of the IEEE International Conference on Image Processing*, vol. 1 (Vancouver, British Columbia, 2000), 49-52, http://dx.doi.org/10.1109/ICIP.2000.900889.

[20] John Canny, “A Computational Approach to Edge Detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8, no.6 (November 1986): 679-714, http://dx.doi.org/10.1109/TPAMI.1986.4767851.

[21] Richard O. Duda and Peter E. Hart, “Use of Hough Transformation to Detect Lines and Curves in Pictures,” *Communications of the ACM* 15, no. 1 (January 1972): 11-15, http://dx.doi.org/10.1145/361237.361242.

[22] John Daugman, “How Iris Recognition Works,” *IEEE Transactions on Circuits and Systems for Video Technology* 14, no. 1 (January 2004): 21–30, http://dx.doi.org/10.1109/TCSVT.2003.818350.

[23] Ton van der Putte and Jeroen Keuning, “Biometrical Fingerprint Recognition: Don’t Get Your Fingers Burned,” *Proceedings of the Fourth Working Conference on Smart Card Research and Advanced Applications* (Norwell, MA: Kluwer, 2001), 289-303.

About the authors:

AJAY SHARMA holds a Master of Technology (CSE) degree from Guru Jambheswar University of Science and Technology, Hisar (Haryana), India since 2004 and is pursuing a Ph.D. degree in Singhania University, Pacheri Beri, (Rajasthan), India. His major field of study is cryptography and network security. His current research area is on cryptographic protocols, symmetric encryption, asymmetric encryption and biometric template security. He has more than six years of teaching experience. He is working as Associate Professor in the Department of Information Technology, Raj Kumar Goel Institute of Technology, Ghaziabad, (U.P.) India. He is the author or co-author of more than 13 publications in national and international journals and conferences. *E-mail*: ajaypulast@rediffmail.com.

DEO BRAT OJHA holds a Ph.D degree from the Department of Applied Mathematics, Institute of Technology, Banaras Hindu University, Varanasi (U.P.), India. His research is focused on optimization techniques, cryptography and network security. He has more than eight years of teaching experience and more than 11 years of research experience. He is currently a Professor in the Department of Mathematics, Raj Kumar Goel Institute of Technology, Ghaziabad (U.P.), India. He is author or co-author of more than 40 publications in national and international journals and conferences.