ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 2, February 2018

Safeguarding the Data by Matching Two Shares of Captcha

Sathiya.V [1], Sujee.I, Subhashini.M, Uma Maheswari.N

Associate Professor, Department of Computer Science and Engineering, Panimalar Engineering College, Chennai, India[1]

ABSTRACT: A new approach for providing limited information only that is necessary for fund transfer during online shopping thereby safeguarding customer data and increasing customer confidence and preventing identity theft. A cryptographic technique based on visual secret sharing used for image encryption. Using k out of n (k,n) visual secret sharing scheme a secret image is encrypted in shares which are meaningless images that can be transmitted or distributed over an untrusted communication channel. Only combining the k shares or more give the original secret image. Phishing is an attempt by an individual or a group to thieve personal confidential information such as passwords, credit card information etc from unsuspecting victims for identity theft, financial gain and other fraudulent activities The use of images is explored to preserve the privacy of image captcha by decomposing the original image captcha into two shares that are stored in separate database servers such that the original image captcha can be revealed only when both are simultaneously available; the individual sheet images do not reveal the identity of the original image captcha. Once the original image captcha is revealed to the user it can be used as the password. Several solutions have been proposed to tackle phishing.

KEYWORDS: HMAC, halftune, captcha , steganography, phishing

I.INTRODUCTION

OBJECTIVE

The main objective of this project is to safeguard customer data and prevent phishing attack during online shopping by using visual cryptography and steganography.

PROJECT SCOPE

A brief survey of related work in the area of banking security based on steganography and visual cryptography is presented in this section. A customer authentication system using visual cryptography is presented in but it is specifically designed for physical banking. A signature based authentication system for core banking is proposed in but it also requires physical presence of the customer presenting the share. Proposes a combined image based steganography and visual cryptography authentication system for customer authentication in core banking.

OVERVIEW

In this paper, a new method is proposed, that uses text based steganography and visual cryptography, which minimizes information sharing between consumer and online merchant but enable successful fund transfer from consumer's account to merchant's account thereby safeguarding consumer information and preventing misuse of information at merchant side. The method proposed is specifically for E-Commerce but can easily be extended for online as well as physical banking. Identity theft is the stealing of someone's identity in the form of personal information and misuse of that information Online shopping is the retrieval of product information via the Internet and issue of purchase order through electronic purchase request, filling of credit or debit card information and shipping of product by mail order or home delivery by courier. Identity theft and phishing are the common dangers of online shopping for making purchase

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 2, February 2018

and opening of bank accounts or arranging credit cards. In 2012 consumer information was misused for an average of 48 days as a result of identity theft.

Phishing is a criminal mechanism that employs both social engineering and technical subterfuge to steal consumers' personal identity data and financial account credentials. In 2ndquarter of 2013, Payment Service, Financial and Retail Service are the most targeted industrial sectors of phishing attacks. Secure Socket Layer (SSL) encryption prevents the interception of consumer information in transit between the consumer and the online merchant. However, one must still trust merchant and its employees not to use consumer information for their own purchases and not to sell the information to others.

II.RELATED WORKS

[1]K. Bennet Purdue University, "Hiding Information in Document Images". In this paper they proposed how to use spatial information to reliably hide data in document images.The system can be used to discourage unauthorized reproduction and dissemination of either copyrighted documents distributed electronically, or paper copies of confidential executive memoranda.

[2]T. S. Chen and M. W. Cheng, "A New Data Hiding Scheme in Binary Image".In this paper they proposed how to use the data-embedding algorithm that is based on Hamming codes, the proposed scheme embeds authentication information into the cover image with flipping only a small number of pixels. Randomly shuffling the bit-order of the authentication information to be embedded, the information can only be extracted by the designated receiver who has the symmetric key.

[3] Jihui Chen and Xiaoyao Xie, "The security of shopping online".In this paper they proposed about the characteristics of online shopping and the current development of the main safety problems, and make online shopping related security measures and transactions.The online shopping convenient, fast, efficient and economic advantage.

[4] S. Manu and P. Deepa Shenoy, "Secure Authentication using Image Processing and Visual Cryptography for Banking Applications". In this paper proposes a technique of processing the signature of a customer and then dividing it into shares. Total number of shares to be created is depending on the scheme chosen by the bank. The correlation method is used to take the decision on acceptance or rejection of the output and authenticate the customer.

[5]Minal chawla, Siddhartha singh chouhan, **"**A Survey of Phishing Attack Techniques",[2014]**.**In this paper they proposed toward phishing, which is a type of network attack where the attacker creates a replica of an existing web page to fool users in to submitting personal, financial, transaction or password data to what they think is their service provider's website. Here we focus on deceptive phishing using social engineering schemes. In this paper we have reviewed various phishing and anti-phishing methods for detecting and preventing phishing attack. The processing of algorithm it generates multiple rules for the phishing data and suspicious links within short of time. To identify deceptive Phishing measures in modern web browser, and detect phishing websites mostly filtering method and machine learning algorithm.

III.EXISTING SYSTEM

CAPTCHA is a standard security technology to tell humans and computers and the most widely used method is text based scheme. As many text schemes have been broken, 3D CAPTCHAs have emerged as one of the latest one. In this paper, we study the robustness of 3D text-based CAPTCHA adopted by Ku6 which is a leading website providing videos in China and provide the first analysis of 3D hollow CAPTCHA. The security of this CAPTCHA scheme relies on a novel segmentation resistance mechanism, which combines Crowding Character Together (CCT) strategy and side surfaces which form the 3D visual effect of characters and lead to a promising usability even under strong overlapping between characters. However, by exploiting the unique features of the 3D characters in hollow font, i.e. parallel boundaries, the different stroke width of side faces and front faces and relationships between them, we propose a technique that segments connected characters apart and repairs some overlapped apart. The success segmentation rate is 70%. With minor changes, our attack program works well on its two variations, the segmentation rate is 75% and 85% respectively. Captcha is now almost a standard security technology, and has found widespread application in commercial websites. Usability and robustness are two fundamental issues

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 2, February 2018

with captcha, and they often interconnect with each other. This paper discusses usability issues that should be considered and addressed in the design of captchas. Some of these issues are intuitive, but some others have subtle implications for robustness (or security). A simple but novel framework for examining captcha usability is also proposed.

IV.PROPOSED SYSTEM

Proposed System, Visual Cryptography (VC), technique is based on visual secret sharing used for image encryption. Secure Socket Layer (SSL) encryption prevents the interception of consumer information in transit between the consumer and the online merchant. In this paper, a new method is proposed, that uses text based steganography and visual cryptography, which minimizes information sharing between consumer and online merchant. VCS is a cryptographic technique that allows for the encryption of visual information such that decryption can be performed using the human visual system. HMAC Algorithm is used for phishing detection and prevention, we are proposing a new methodology to detect the phishing website. Our methodology is based on the Anti-Phishing Image Captcha validation scheme using visual cryptography.

It prevents password and other confidential information from the phishing websites. Cryptographic technique: (2, 2) - Threshold VCS scheme,(n, n) Threshold VCS scheme, (k, n) Threshold VCS scheme are used in this proposed system.

Advantages: Our methodology is based on the Anti-Phishing Image Captcha validation scheme using visual cryptography. It prevents password and other confidential information from the phishing websites.For phishing detection and prevention, we are proposing a new methodology to detect the phishing website

V.ALGORITHMS

HMAC algorithm: In cryptography, HMAC (Hash-based Message Authentication Code), is a specific construction for calculating a message authentication code (MAC) involving a cryptographic hash function in combination with a secret key. As with any MAC, it may be used to simultaneously verify both the data integrity and the authenticity of a message. Any cryptographic hash function, such as MD5 or SHA-1, may be used in the calculation of an HMAC; the resulting MAC algorithm is termed HMAC-MD5 or HMAC-SHA1 accordingly. The cryptographic strength of the HMAC depends upon the cryptographic strength of the underlying hash function, the size of its hash output length in bits and on the size and quality of the cryptographic key. An iterative hash function breaks up a message into blocks of a fixed size and iterates over them with a compression function. For example, MD5 and SHA-1 operate on 512-bit blocks.

Halftune algorithm: (2,N) Visual Secret Sharing Scheme: This is a simple algorithm for binary (black and white) visual cryptography that creates 2 encrypted images from an original unencrypted image. The algorithm is as follows: First create an image of random pixels the same size and shape as the original image. Next, create a second image the same size and shape as the first, but where a pixel of the original image is the same as the corresponding pixel in the first encrypted image, set the same pixel of the second encrypted image to the opposite color. Where a pixel of the original image is different than the corresponding pixel in the first encrypted image, set the same pixel of the second encrypted image to the same color as the corresponding pixel of the first encrypted image. The two apparently random images can now be combined using an exclusive-or (XOR) to re-create the original image.

VI. SOFTWARE DESCRIPTION

Microsoft Visual Studio

Microsoft Visual Studio is an integrated development environment from Microsoft. It is used to develop computer programs for Microsoft Windows, as well as web sites, web applications and web services. Visual Studio uses such as Windows API, Windows Forms, Windows Presentation Foundation.

Visual Studio includes a code editor supporting IntelliSense (the code completion component).The integrated debugger works both as a source-level debugger and a machine-level debugger. Other built-in tools include a forms

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 2, February 2018

designer for building GUIapplications, web designer, class designer, and database schemadesigner. It accepts plug-ins that enhances the functionality at almost every level—including adding support for source-control systems.

A) C#.NET

C# is one of many .NET programming languages. It is object-oriented and allows you to build reusable components. The Microsoft .NET framework is a software framework developed by Microsoft that runs primarily on Microsoft Windows. C# is a multi-paradigm programming language encompassing strongtyping, imperative, declarative, functional,generic, objectoriented and component-oriented programming disciplines. It was developed by Microsoft within its .NET initiative and later approved as a standard by Ecma and ISO . C# is one of the programming languages designed for the Common Language Infrastructure.

Furthermore, C# has added several major features to accommodate functional-style programming, culminating in the LINQextensions released with C# 3.0 and its supporting framework of lambda expressions, extension methods, and anonymous types. These features enable C# programmers to use functional programming techniques, such as closures, when it is advantageous to their application.

Programs written for .NET Framework execute in a softwareenvironment (as contrasted to hardware environment), known as Common Language Runtime (CLR), an application virtual machine that provides services such as security, memory management, and exception handling. FCL and CLR together constitute .NET Framework. Framework is intended to be used by most new applications created for windows platform. Microsoft also produces an integrated.

B) Features of .NET

Microsoft .NET is a set of Microsoft software technologies for rapidly building and integrating XML Web services, Microsoft Windows-based applications, and Web solutions. The .NET Framework is a language-neutral platform for writing programs that can easily and securely interoperate. There's no language barrier with .NET: there are numerous languages available to the developer including Managed C++, C#, Visual Basic and Java Script. The .NET framework provides the foundation for components to interact seamlessly, whether locally or remotely on different platforms. It standardizes common data types and communications protocols so that components created in different languages can easily interoperate. .NET is also the collective name given to various software components built upon the .NET platform.

C) The .NET Framework

.NET Framework has two main parts:

1. The Common Language Runtime (CLR).
2. A hierarchical set of class libraries.

The CLR is described as the "execution engine" of .NET. It provides the environment within which programs run. The most important features are Conversion from a low-level assembler-style language, called Intermediate Language (IL), into code native to the platform being executed on.

The .NET Framework is a new computing platform that simplifies application development in the highly distributed environment of the Internet.

<table>
<tr><td>ASP.NET
XML web services</td><td>Windows

Forms</td></tr>
<tr><td colspan="2">Base Class Libraries</td></tr>
<tr><td colspan="2">Common Language Runtime</td></tr>
<tr><td colspan="2">Operating System</td></tr>
</table>

Fig 3.1 .NET Framework

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 2, February 2018

D) Managed Code

The code that targets .NET, and which contains certain extra Information - "metadata" - to describe itself. Whilst both managed and unmanaged code can run in the runtime, only managed code contains the information that allows the CLR to guarantee, for instance, safe execution and interoperability.

E) Managed Data

With Managed Code comes Managed Data. CLR provides memory allocation and Deal location facilities, and garbage collection. Some .NET languages use Managed Data by default, such as C#, Visual Basic.NET and JScript.NET, whereas others, namely C++, do not. Targeting CLR can, depending on the language you're using, impose certain constraints on the features available. As with managed and unmanaged code, one can have both managed and unmanaged data in .NET applications - data that doesn't get garbage collected but instead is looked after by unmanaged code.

F) Common Type System

The CLR uses something called the Common Type System (CTS) to strictly enforce type-safety. This ensures that all classes are compatible with each other, by describing types in a common way. CTS define how types work within the runtime, which enables types in one language to interoperate with types in another language, including cross-language exception handling. As well as ensuring that types are only used in appropriate ways, the runtime also ensures that code doesn't attempt to access memory that hasn't been allocated.

G) Common Language Specification

The CLR provides built-in support for language interoperability. To ensure that you can develop managed code that can be fully used by developers using any programming language, a set of language features and rules for using them called the Common Language Specification (CLS) has been defined. Components that follow these rules and expose only CLS features are considered CLS-compliant.

H) Objectives of .NET framework

1) To provide a consistent object-oriented programming environment whether object codes is stored and executed locally on Internet-distributed, or executed remotely.

2) To provide a code-execution environment to minimizes software deployment and guarantees safe execution of code.

3) Eliminates the performance problems.

Introduction To SQL Server

To create a database determines the name of the database, its owner (the user who creates the database), its size and the file groups used to store it. Microsoft SQL Server is a relational database management system developed by Microsoft. As a database, it is a software product whose primary function is to store and retrieve data as requested by other software applications, be it those on the same computer or those running on another computer across a network (including the Internet). There are at least a dozen different editions of Microsoft SQL Server aimed at different audiences and for workloads ranging from small single-machine applications to large Internet-facing applications with many concurrent users. Its primary query languages are T-SQL and ANSI SQL.

Common Language Runtime (CLR) integration was introduced with this version, enabling one to write SQL code as Managed Code by the CLR. For relational data, T-SQL has been augmented with error handling features (try/catch) and support for recursive queries with CTEs (Common Table Expressions). SQL Server 2005 has also been enhanced with new indexing algorithms, syntax and better error recovery systems. Data pages are check summed for better error resiliency, and optimistic concurrency support has been added for better performance. Permissions and access control have been made more granular and the query processor handles concurrent execution of queries in a more efficient way. Partitions on tables and indexes are supported natively, so scaling out a database onto a cluster is easier. SQL CLR was introduced with SQL Server 2005 to let it integrate with the .NET Framework.

Data storage-Data storage is a database, which is a collection of tables with typed columns.SQL Server supports different data types, including primary types as Integer, Float, Decimal, Char (including strings), Varchar (variable length character strings), binary (for unstructured blobs of data), Text (for textual data) among others.

Microsoft SQL Server also allows user-defined composite types (UDTs) to be defined and used. It also makes server statistics available as virtual tables and views (called Dynamic Management Views or DMVs). In addition to

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 2, February 2018

tables, a database can also contain other objects including views, stored procedures, indexes and constraints, along with a transaction log.

SQLCMD-SQLCMD is a command line application that comes with Microsoft SQL Server, and exposes the management features of SQL Server. It allows SQL queries to be written and executed from the command prompt. It can also act as a scripting language to create and run a set of SQL statements as a script. Such scripts are stored as a .sql file, and are used either for management of databases or to create the database schema during the deployment of a database.

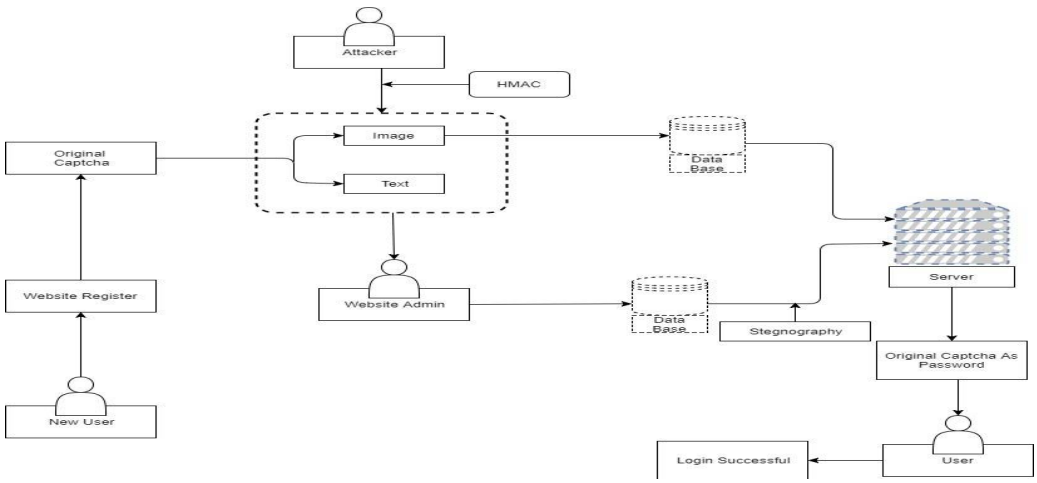
Data retrieval-The main mode of retrieving data from an SQL Server database is querying for it. The query is expressed using a variant of SQL called T-SQL, a dialect Microsoft SQL Server shares with Sybase SQL Server due to its legacy. The query declaratively specifies what is to be retrieved. It is processed by the query processor, which figures out the sequence of steps that will be necessary to retrieve the requested data. The sequence of actions necessary to execute a query is called a query plan. There might be multiple ways to process the same query. For example, for a query that contains a join statement and a select statement, executing join on both the tables and then executing select on the results would give the same result as selecting from each table and then executing the join, but result in different execution plans. In such case, SQL Server chooses the plan that is expected to yield the results in the shortest possible time. This is called query optimization and is performed by the query processor itself.

SQL Server includes a cost-based query optimizer which tries to optimize on the cost, in terms of the resources it will take to execute the query. Given a query, then the query optimizer looks at the database schema, the database statistics and the system load at that time. It then decides which sequence to access the tables referred in the query, which sequence to execute the operations and what access method to be used to access the tables

In the context of a relational database table, a column is a set of data values of a particular simple type, one for each row of the table. The columns provide the structure according to which the rows are composed. The term field is often used interchangeably with column, although many consider it more correct to use field (or field value) to refer specifically to the single item that exists at the intersection between one row and one column. In relational database terminology, column's equivalent is called attribute. For example, a table that represents companies might have the following columns:

i. ID (integer identifier, unique to each row)
ii. Name (text)
iii. Address line 1 (text)
iv. Address line 2 (text)
v. City (integer identifier, drawn from a separate table of cities.

VII. ARCHITECTURE DIAGRAM



ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 2, February 2018

VIII. PROJECT DESCRIPTION

MODULES

1. User and admin authentication
2. Captcha encryption
3. Captcha decryption
4. Customer merchant
5. Amount transaction
6. Transfer to fund

MODULES DESCRIPTION

User and admin authentication
Performs log on for basic form authentication. You can use these login modules to perform authentication with user ID and password. The logic modules used for authentication with client certificates.

Captcha encryption
This module describes the signup form captcha as encrypt and make it as image format to access the user login page for security purpose.

Captcha decryption
This module prescribes the encrypted captcha image to bind the captch in data information in stenography image to protect from attacks. From this technique we can hide the user authentication process from hackers.

Customer merchant
Customer authentication information is sent to the merchant by CA. Upon receiving customer password, bank matches it with its own databases and after verifying, legitimate customer, transfers fund from the customer account to the submitted merchant account. After receiving the fund, merchant's payment system validates receipt of payment using customer authentication information. The problem is that CA does not know to which bank to forward the cover text obtained from combining two shares. It can be solved by appending 9 digit routing or transit number of bank with customer authentication information.

Payment Gateway
Consumer selects item from online shopping portal and then is directed to the payment page. Online merchant may have its own payment system or can take advantage of third party payment system such as PayPal, Pay online web moment and others.

Transfer of fund
During shopping online, after selection of desired item and adding it to the cart, preferred payment system of the merchant direct. The customer to the Certified Authority Portal. In the portal, shopper submits its own share and merchant submit its own account. Consumer selects items from online shopping portal and then is directed to the payment page. Online merchant may have its own payment system or can take advantage of third party payment systems such as PayPal, pay online system, Web Money and others. In the payment portal consumer submit his or her credit. Now one share is kept by the customer and the other share is kept in the database of the certified authority.

IX. CONCLUSION

In this paper, a payment system for online shopping is proposed by combining text based steganography and visual cryptography that provides customer data privacy and prevents misuse of data at merchant's side. The method is concerned only with prevention of identity theft and customer data security. In comparison to other banking application which uses steganography and visual cryptography are basically applied for physical banking, the proposed method can be applied for E-Commerce with focus area on payment during online shopping as well as physical banking.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 2, February 2018

REFERENCES

[1] Jihui Chen, Xiaoyao Xie, and Fengxuan Jing, "The security of shopping online," Proceedings of 2011 International Conference on Electronic and Mechanical Engineering and Information Technology (EMEIT), vol. 9, pp. 4693-4696, 2011.
[2] M. Mohamed, N. Sachdeva, M. Georgescu, S. Gao, N. Saxena, C. Zhang, P. Kumaraguru, P. C. Van Oorschot, and W. B. Chen, "A three-way investigation of a game-captcha: automated attacks, relay attacks and usability," in ACM Symposium on Information, Computer and communication security,pp.195-206,2014
[3] L. Von Ahn, M. Blum, and J. Langford, "Telling humans and computers apart automatically," Communications of the ACM, vol. 47, no. 2, pp. 56–60, 2004.
[4] Jack Brassil, Steven Low, Nicholas Maxemchuk, Larry O'Gorman, "Hiding Information in Document Images," Proceedings of the 1995 Conference on Information Sciences and Systems, pp. 482-489, 1995.
[5] J. Chen, T. S. Chen, M. W. Cheng, "A New Data Hiding Scheme in Binary Image," Proceeding of Fifth International Symposium on Multimedia Software Engineering, pp. 88-93, 2003.
[6] Hu ShengDun, U. KinTak, "A Novel Video Steganography Based on Non-uniform Rectangular Partition," Proceding of 14th International Conference on Computational Science and Engineering, pp. 57-61, 2011.
[7] H. Gao, W. Wang, J. Qi, X. Wang, X. Liu, and J. Yan, "The robustness of hollow captchas," in Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security. ACM, pp. 1075–1086,2013
[8] Jaya, Siddharth Malik, Abhinav Aggarwal, Anjali Sardana, "Novel Authentication System Using Visual Cryptography," Proceedings of 2011 World Congress on Information and Communication Technologies, pp. 1181-1186, 2011.
[9] K. Thamizhchelvy, G. Geetha, "E-Banking Security: Mitigating Online Threats Using Message Authentication Image (MAI) Algorithm," Proceedings of 2012 International Conference on Computing Sciences (ICCS), pp. 276 – 280, 2012.
[10] S. Suryadevara, R. Naaz, Shweta, S. Kapoor, "Visual cryptography improvises the security of tongue as a biometric in banking system," Proceedings of 2011 2nd International Conference on Computer and Communication Technology (ICCCT), pp. 412 – 4,2011