I J C T A, 10(9), 2017, pp. 437-449 ISSN: 0974-5572

Chain Algorithm to Protect Data sets From Hackers

G. Arul Dalton* and Dhumpati Raghu**

ABSTRACT

Computer forensic (sometimes known as Computer forensic science) is a discipline involved in data recovery but with additional guidelines and practices. Are we safe today or at present? But there is no answer from four sides of the corner. Computer's today store lot of information related to banks, ssn, personal details, colleges, schools, flights and lot more. Most countries depend upon databases. Now a day's lots of black professionals involved to cyber crimes, trying to steal, modify or misuse data. These types of professionals are said to be black hats. Sometimes cause damage to systems or data. Day-to-day sophisticated systems, software are changing but no computer system is safe today. White professionals should be very conscious related to their data. Some professional are meant for systems and their safety. These types of professional are called white hats. Chain algorithm is used to safe guard data. Algorithm maintains records in different servers with timestamp. Whenever hacker wants to hack data from the server, hacker should hack all the servers which are not possible. Mean time we can trace hacker involvement with the data records by matching backup records with the records available from the servers. When mismatch it will alert the administrators with a caution that systems are hacked. By checking backup record set with DNS server. There is scope for improvement.

Keywords: computer forensic; chain algorithm; hacker; hack; server, backup

1. INTRODUCTION

Most of the countries in the world depend upon computer systems. Computer Systems can store large volumes of data; Computer Systems are used in areas like banks systems, government organizations, private organizations and lot more places are carried on with it. Lot of data is given to the system daily. Volumes are growing per day. Systems are made to with stand to retrieve and store inform in form of records in the server. Whenever outside world want to retrieve or store inform, it will be easy to have them when there is communication like internet. Everyday lots of people use their systems to retrieve and store information in server. Servers play a key role in today's world. User can use web browser to connect Sserver.

Every coin has two sides. One coin may face safe side and other side can face unsafe. When coming to safe, there is no answer. Because every system in network is unsafe. Both the head and tail face towards unsafe. We are rich in gathering data but poor in saving or protect them daily. People focus on antivirus, but they cannot protect your computer. Lot of viruses programs are coded daily by black hats. Inject virus and then sell antivirus to public. Hackers hack your system or server when you are in network to steal your personal details and information. Hacker can misuse the information which is stolen from your system. Hackers spend lot of time on systems to steal others data. Hacker needs browser, an internet connection and clear mind. Figure 1 focus on hacker. Hackers sometimes referred to as black hats. Hackers focus on government organizations. Recently many countries are targeted by hackers. One among them is India.

* Department of Computer Science & Engineering, MLR Institute of Technology, Hyderabad, India, *Email: arulmlrit@gmail.com*

** Department of Computer Science & Engineering, MLR Institute of Technology, Hyderabad, India, *Email: raghu.dhumpati@gmail.com*

Figure 1: Hacker

1.1. How black hat hackers will hack a web based server

Mostly around the globe end users use systems to buy product, to reserve tickets, to add their personal details to the system, 3-D map to view various locations, search for any details, transfer amount from one account to another account, to book movie tickets, online games and map navigator for various drivers. Today's world has lot of organizations with their own websites. Most websites of the organizations store valuable information such as person details, budget details, email address, bank accounts, credit card numbers, there passwords, etc., Server are free doors for attackers. Attackers have a knowledge related to hacking. They are also called black hats. White hats work against black hats. White hat saves systems from black hats.

1.2. Attacks on Web Servers

Today's internet world saves data on server. A web server is a place where it stores files from different users. Users store, modify, delete and retrieve information from the web servers. The accessibility to the server is by network or internet media. A web server is part of hardware and software. Attackers usually target the software to gain authorized entry to the server. Common attacks that are possible by attackers are as follows:

- Default login user ids and passwords (Fig. 2).
- Executing commands on the server can be dangerous
- Inject bugs or virus into your operating or web server
- Lack of security policy and procedures
- Lack of strong passwords (Mobile code for entry into gmail id Fig.3)

1.3. Types of attacks against Web Servers

An attack exploits bugs in the web server to gain unauthorized access to files and folders that are not in public domain. Once the attacker has gained access, they can download sensitive information execute commands on the server or install malicious software. Top 6 web hacking techniques.

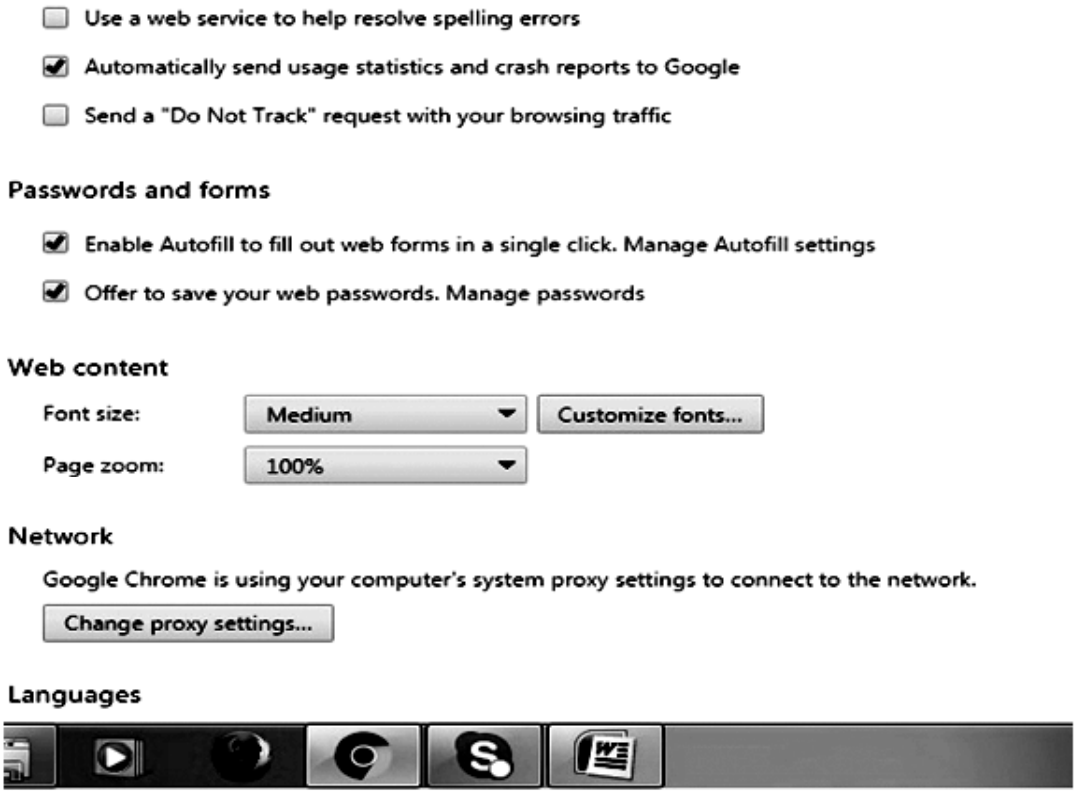

Figure 2: Passwords saving scheme.

1. URL based
2. Directory Browsing
3. Retrieving “non-web” files
4. Reverse Proxying
5. SQL Querying Poisoning
6. Session Hijacking

1.4. Effects of successful attacks

Banks in India will either replace or ask users to change the security codes of as many as 3.2 million debit cards in what’s emerging as one of the biggest ever breaches of financial data in India; several victims have reported unauthorized usage from locations in China and Pakistan.

An organization’s reputation can be ruined if the attacker edits the website content and includes malicious (virus) information or links to the porn website. Websites seem to porn but it is related ready to collapse your system and steals every information available in it. The web server can be used to install malicious software on users who visit the compromised website. The malicious software downloaded on to the visitor’s computer can be a virus, Trojan or botnet software etc. compromised user data may be used for fraudulent activities which may lead to business loss or law suits from the users who entrusted their details with the organization. There are lots of examples in world to tell, speak about sophisticated systems but no answer for security.

2. EXISTING SYSTEM

Websites and web applications are rapidly growing in today’s world. Lot of data in records is stored on databases in server. Now complex business applications are now delivered over the websites (http). HTTP request and HTTP reply from web client to web server. Fig. 3 shows typical web application setup.

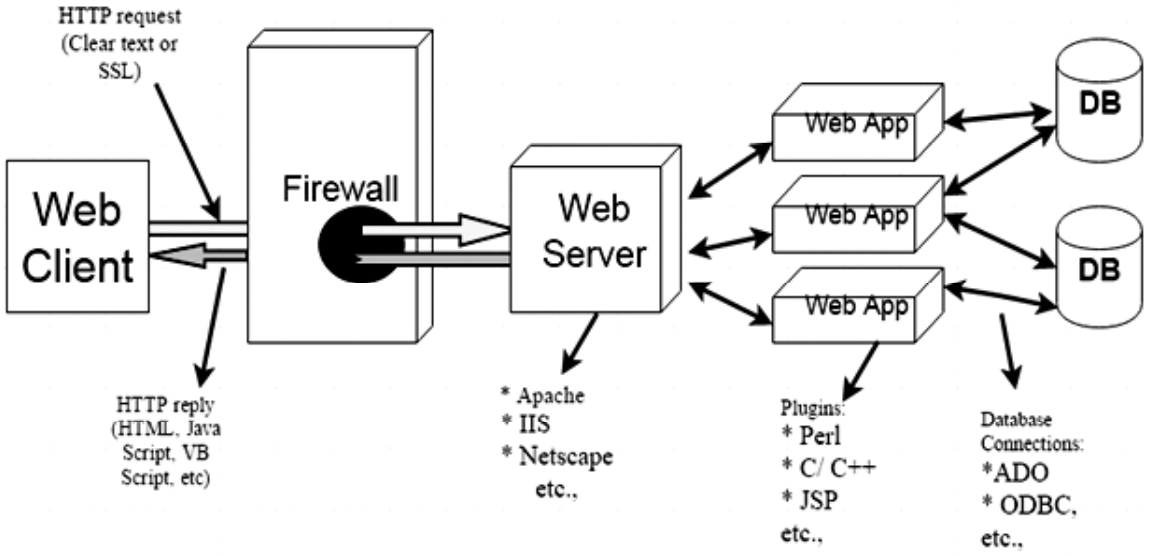

Figure 3: Typical Web Application Setup.

Web hacking activity also started based on increasing volumes. Based on present web application setup cannot save your system from hackers. Server is an open door for hackers. Attackers inject worms into your servers. This will damage you're web server and steal your valuable or privacy information, firewalls play a key role in web applications via firewalls. Web client access your web server. A firewall is a network security system designated to prevent unauthorized access to or from a private network. Firewalls can be implemented in both hardware and software or a combination of both. Hacker from outside can bypass firewall to attack target server. It is very difficult to tell answer.

3. LITERATURE STUDY

People are made to be brought awareness of the computer systems, there security and safe but there are so many stories that are related to unsecure of computers. The major hurdle for safe computers is attackers. Machine learning is process where how a system should work for better usage. Spam detection method is used to detect worms [1]. Identifying hacking attempts and patch working security systems gained by parties. A person who is called attacker causing inconvenience to public or private systems. Protecting this system from hacker is the main theme. Working cycle of ethical hacking gives how user protects the system from different users who use system for different purpose, attackers gain power using internet for attacking user systems. A white hat protects systems from black hats. Ethical hacking is performed with the legal permission like internet to hack system. Internet is main source to attack the system. Paper [3] discuss about different types of attacks on systems. White hats, black hats and grey hats are persons who related to the security. White hats protect the system, black hats are hackers and grey hats are both hacker and cracker. Trojan horse, worm, sniffer, virus and root kit are the tools used by hackers [2][3].

Systems should be protected from hackers who damage system. Paper [4] study gives information about attackers attacking system and reporting it to the owner back again. It gives basic idea about ethical hacking. Protecting systems contains some steps like testing the system thoroughly, risks involved in testing, when the tests takes place and overall time line of testing, way of test process, good knowledge of the system, how to face when major problem arises [2][3][4].

Users are surfing internet but security is very poor. Attackers can hack your system to get valuable information. Suffers are public and private organizations in the world. Attackers can steal valuable information via internet. Need safety from attackers. Attackers find system weakness and hack the system to be benefitted. Impact of the hacking over business and government organizations is more [1][2][3][4][5].

So many users carelessly leave there valuable data in these systems without locks. Hackers have advantage on such careless data; they may destroy or misuse such data. This paper gives suggestions and protection tricks for data [6]. Storing valuable information and transferring valuable information to different places in network is common. Large data is compressed and transferred. Compression algorithms are used. There is possibility of malware. Malware can damage entire systems [7]. Main origin for stealing valuable information source is internet. Based on internet one can steal valuable user passwords and login name, this is called phishing, group of computers grouped together, each computer in a botnet is called a bot. these bots form a network chain of computers, which are controlled by third party and use to transfer malware or spam or launch attacks. To protect systems the novel approach single-linkage hierarchical clustering of the finger prints using normalized compress distances a distance metric [8]. Malicious website are the door ways to steal your valuable information like personnel or in worse-case scenario gain control over your system. After getting control attacker can secretly shoot your activities or record your naked photographs for their misuse. Attackers can install any software or websites in your system without your permission. For safety purpose we have five machine learning algorithms. Based on algorithms we can protect our system [9]. Protecting company, organization and employees from image spam is main theme. From last year onwards global spam volumes have increased five times and research has seen an increase of 75 percent in just the past 3 months. Spam appears on internet to 90% of all email accounts. Image spam is having possibility to introduce virus. Two strategies are followed PCA and SVM. These two provide high accuracy with low computational complexity. We develop a new spam image dataset that cannot be detected using PCA or SVM approach. This new dataset proves to be valuable for improving image spam detection capabilities [10]. The URL as a cruise missile as shown in Fig. 4.

4. PROPOSED SYSTEM

Chain algorithm is used for storing datasets at different servers with timestamp. Based on timestamp we retrieve exact dataset from the server. Maintain dataset on different servers with timestamp can keep our datasets safe. For maintaining servers in network, user should have knowledge of Graph theory, connectivity, servers to be identified in network at various places by using dijkstra algorithm. Dijkstra used for path of servers.

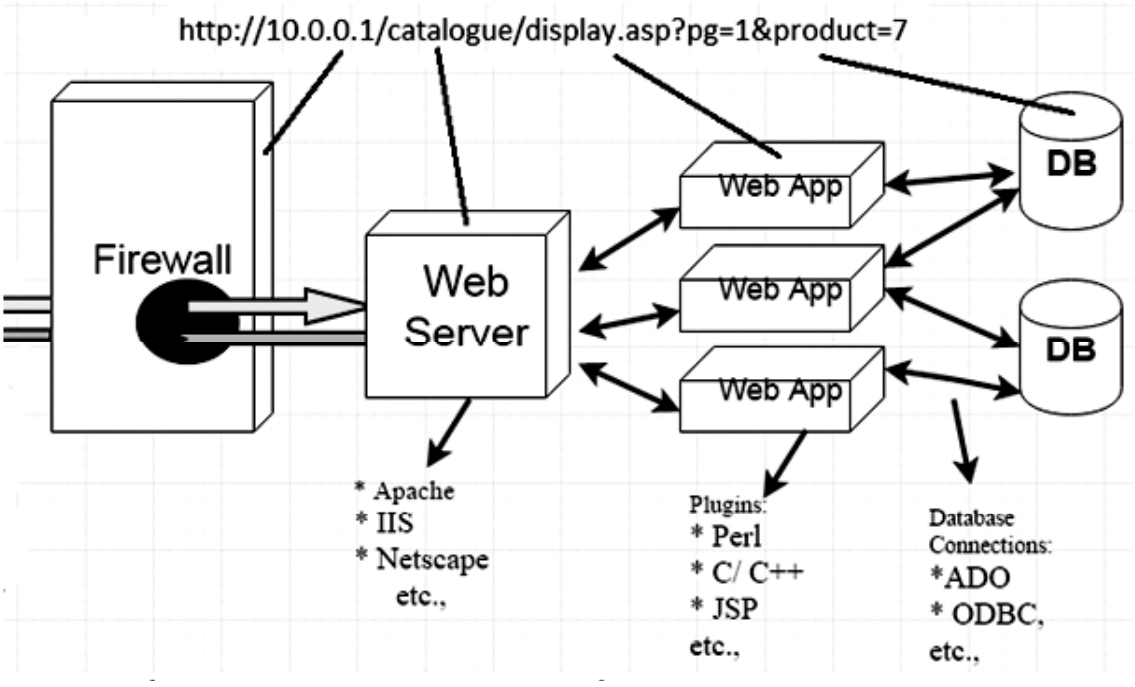

Figure 4: The URL as a cruise missile.

4.1. Graph Connectivity

Graph type is chosen for this problem. There will be servers in the network to place our data record sets. Placing data record sets on different servers that are available in the network is main theme. Servers can be homogeneous or heterogeneous. Two vertices u and v in an undirected graph G are called adjacent (or neighbors) in G if u and v are endpoints of an edge of G. If e is associated with {u, v}, the edge e is called incident with the vertices u and v. The edge e is also said to connect u and v. The vertices u and v are called endpoints of an edge associated with {u, v}. Vertices are also called as node. Node can be server or system. Servers are available in network. Professionals and public interactive with their systems to server for update or delete or provide information. A graph in which each edge connects two different vertices (servers) and where no two edges connect the same pair of vertices is called a simple graph. Figure 5 is a computer network server with diagnostic links.

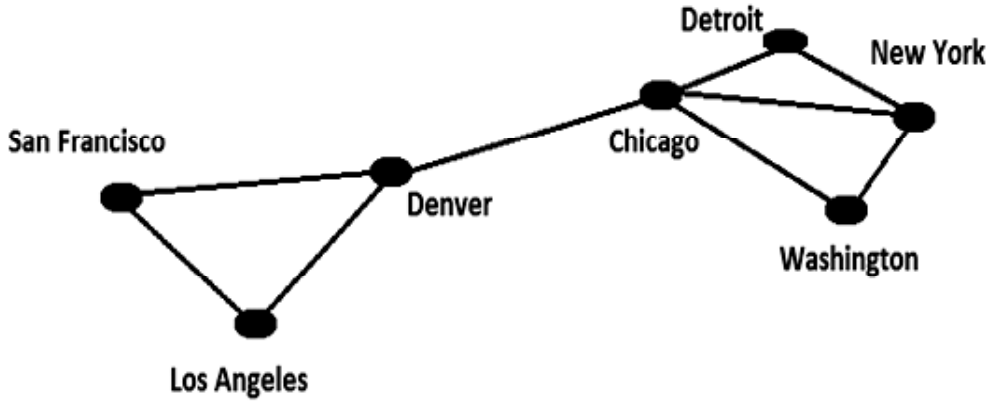

Figure 5: A Computer Network Server with Diagnostic Links.

4.1.1. Connectivity

Many problems can be modeled with paths formed by traveling along the edges of graphs. For instance, the problem of determining whether a message can be studied with a graph model. Problems of efficiently planning routes for mail delivery, garbage pickups, diagnostics in computer network, data record sets and soon can be solved using models that involve paths in graphs.

4.1.2. Paths

Path plays a key role in network. Our concept is related to place our data record sets on different servers with timestamp. Based on timestamp we can get exact updated data record sets. Informally, a path is a sequence of edges that begins at a vertex of a graph and travels from vertex to vertex along edges of the graph. A formal definition of paths and related terminology is given in definition.

a) Algorithm 1

Let n be a nonnegative integer and G an undirected graph. A path of length n from u to v in G is a sequence of n edges e_1,…..e_n of G such that e_1 is associated with $\{x_0, x_1\}$, e_2 is associated with $\{x_1, x_2\}$ and so on, with en associated with $\{x_{n-1}, x_n\}$, where x_0=u and x_n=v. when the graph is simple, we denote this path by its vertex sequence x_0, x_1,…..,x_n (because listing these vertices uniquely determines the path). The path is a circuit is said to pass through the vertices x_1, x_2,….., x_{n-1} or traverse the edges e_1, e_2,…..e_n. A path or circuit is simple if it does not contain the same edge more than once.

b) Algorithm 2

Let n be a nonnegative integer and G a directed graph. A path of length n from u to v in G is a sequence of edges e_1, e_2,…e_n of G such that e1 is associated with (x_0, x_1), e_2 is associated with (x_1, x_2), and so on, with

(x_{n-1}, x_n), where x_0=u and x_n=v. When there are no multiple edges in the directed graph, this path is denoted by its vertex sequence x_0, x_1, x_2,….x_n. A path of length greater than zero that begins and ends at the same vertex is called a circuit or cycle. A path or circuit is called simple if it does not contain the same edge more than once.

4.1.3. Paths in Acquaintanceship Graphs

In an acquaintanceship graph there is a path between two servers if there is a chain of servers linking these servers, where two servers adjacent in the chain know one another. For example, there is a chain of six servers linking DNS server 10.10.10.3 and server site at UK. Chain of servers is linked together like this based on accessibility. This would mean that almost every pair of vertices in the acquaintanceship graph containing all servers in the world is linked by a path of length. Fig. 6 Acquaintanceship graph between few servers.

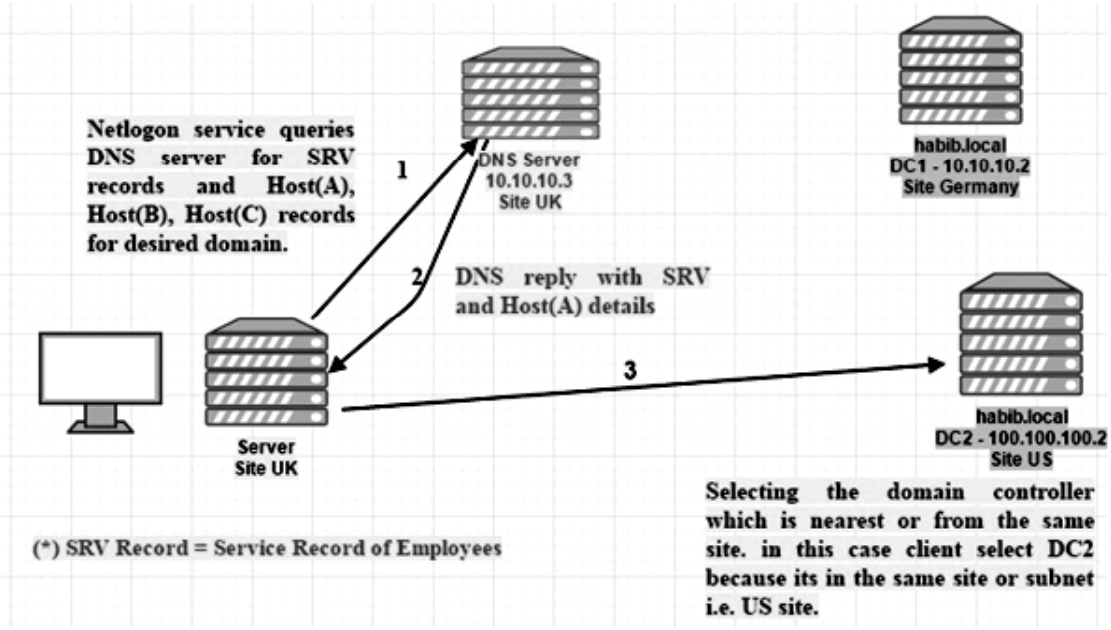

Figure 6: Acquaintanceship between servers.

4.1.4. Menger's theorem

Menger's theorem is a characterization of the connectivity in finite undirected graphs in terms of the minimum number of disjoint paths that can be found between any pair of vertices. It was proved by for edge-connectivity and vertex-connectivity. The edge-connectivity version of menger's theorem was later generalized by the max-flow min-cut theorem.

4.1.5. Edge-Connectivity

The edge-connectivity version of menger's theorem is as follows:

Let G be a finite undirected graph and v_1 and v_2 two distinct vertices. Then the theorem states that the size of the minimum edge cut for v_1 and v_2 (the minimum number of edges whose removal disconnects v_1 and v_2) is equal to the maximum number of pair-wise edge-independent paths from v_1 and v_2.

Extended sub-graphs a maximal sub-graph disconnected by no less than a e-edge cuts is identical to a maximal subgraph with a minimum number e of edge-independent paths between any v1, v2 pairs of nodes in the sub-graph. This theorem let us know about those available servers in networks.

4.1.6. Undirected Graph

An undirected graph is graph that is a set of vertices or nodes that are connected together, where all the edges are bidirectional. An undirected graph is sometimes called an undirected network. In contrast, a graph where the edges point in a direction is called a directed graph. When drawing an undirected graph, the edges are typically drawn as lines between pairs of vertices as shown in Figure 7.

One can formally define an undirected graph as G = (u, v), consisting of the set u of edges and the set v of vertices, which are unordered pairs of elements of u.

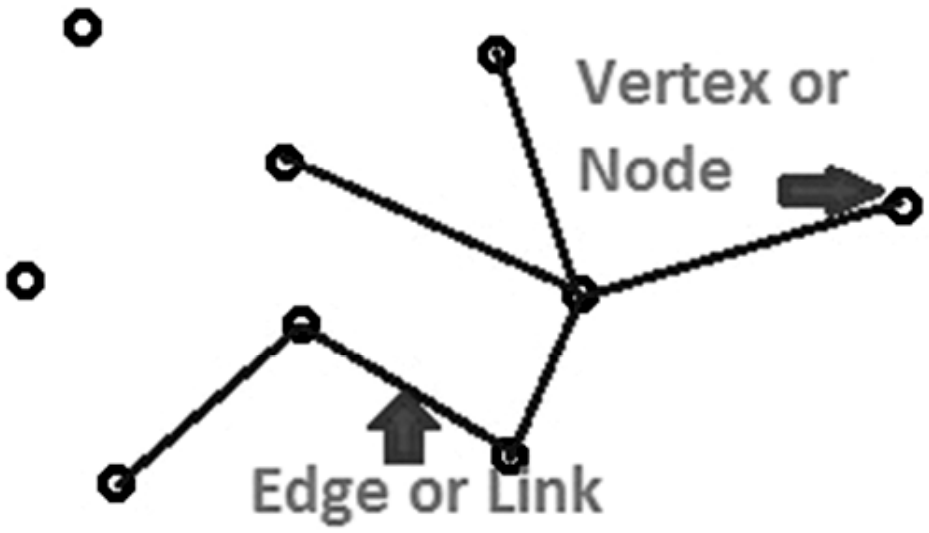

Figure 7: An undirected graph with 6 edges 7 vertices.

4.1.7. Dijkstra's Algorithm

Dijkstra's algorithm plays a key role in this paper. Best example for Dijkstra's Algorithm is road ways. This algorithm is based on adjacency matrix representation of a graph. It finds not only the shortest path from one specified vertex to another, but the shortest path from the specified vertex to all the other vertices. If the graph is weighted, the problem becomes harder, but we can still use the ideas from unweighted case. We keep all of the same information as before. Thus, each vertex is marked as either known or unknown. A tentative distance d_v is kept for each vertex, as before. This distance turns out to be the shortest path length from s to v using only known vertices as intermediates. As before, we record p_v, which is the last vertex to cause a change to d_v.

a) Dijkstra's Algorithm

Procedure Dijkstra(G:weighted connected simple graph, with

all weights positive)

{ G has vertices $a=v_0,v_1,\ldots v_n=z$ *and weights* $w(v_i,v_j)$

Where $w(v_i,v_j)=\infty$ *if* $\{v_i,v_j\}$ *is not an edge in G}*

for i:=1 to n

$L(vi):=\infty$

$L(a):=0$

$S:=\phi$

{the labels are now initialized so that the label of a is 0 and all other labels are ∞, and S is the empty set}

While Z G S

Begin

a:=a vertex not in S with L(u) minimal

S:=S U {u}

For all vertices v not in S

If L(u) + w(u, v) < L(v) then L(v):= L(u) + w(u,v)

{this adds a vertex to S with minimal label and

updates the labels of vertices not in S}

end

Based on these concepts we store our record sets data in servers. Storing of each same type of record sets with timestamp (encrypted) will be based on graph connectivity and Dijkstra's algorithm is used for paths not for shortest paths. My focus is on knowing paths for servers to place data record sets.

Every organization contains records. Each file or sql table contains records, records contain fields and fields contain data. For every entity data will be in form because they are arranged in form of a table like a set. Chain algorithm is used for saving data sets from hackers.

5. IMPLEMENTATION

Maintain set of records on different servers with a timestamp. Replacing will go on in the network or it may be repeated as one cycle. If hacker wants to hack your server he has to hack all servers. This is not possible. By maintaining replicates of the available record sets with timestamps, we can avoid hacking for some extent. Implementation is a transition from above topics.

Chain Algorithm steps involved as follows:

1. Collect record sets. They may be related to bank account, employee records, university database or so on.
2. Copy record sets on different servers with timestamp or encrypted timestamp.
3. Remember path and not to maintain same path between servers in future. Change path.
4. Maintain timestamp for each record sets.
5. This is repeated for n times. Where n is servers list.
6. There should be one server and mobile network system in maintaining all this information.

Servers can be placed in different geographical areas or business client can hire or purchase and maintain different servers as show in below Fig. 8.

Path related to servers should be changed every time for not hacking.

Direction d1 = {Server 1, Server 2, Server 3, ….}

Direction d2 = {Server 1, Server 3, Server 7, …}

Direction d3 = {Server 2, Server 1, Server 15,…}

Direction d_n = { Server 1 to Server i}

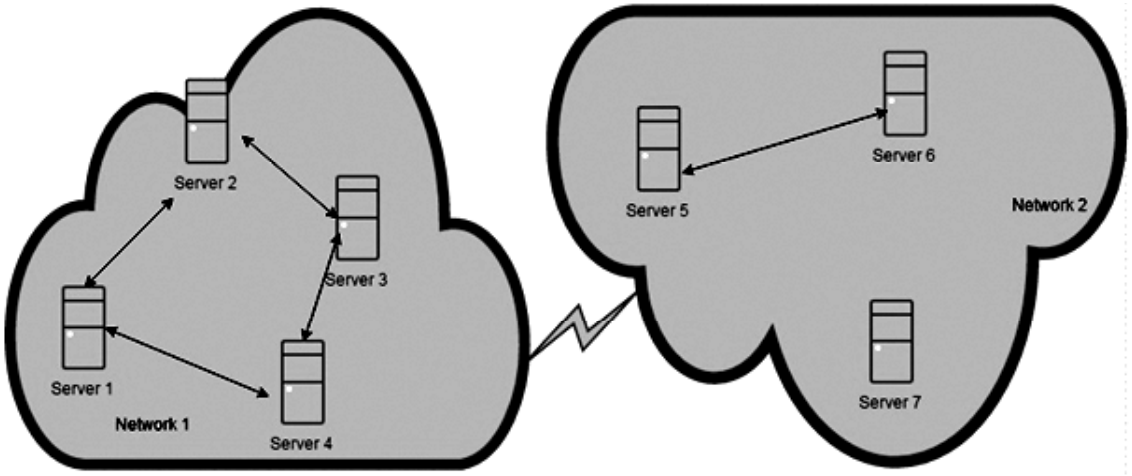

Figure 8: Servers available in various networks.

"i" is any server number from 1 to m, n is n^{th} direction from server 1 to "i". This process is continued and continued.

Initially record set R contains tuples with timestamp T. T timestamp is encrypted to T'. Before replacing in server S1, the record set is R(T').

Server S1 = {R(T')}

Likewise other servers will have the record sets.

Server S_m = {R(T')}

Administrator or Professional should be traced by getting server identity, masking value, and timestamp of last store record set. Because there is possibility of storing record set more than one time. This process can carry on with the help of mobile network or maintaining blackout server. There can be a table system to represent the server information and timestamps.

Server S_m = {R(T)}, record set timestamp is decrypted and match with available data.

Master Server should be there to have monitor directions and check whether the any of server is hacked or not. Blackout Server will have a blackout period. Otherwise hackers make hack this server also. This server will have a backup data with it. Blackout Server is shown in Fig. 9.

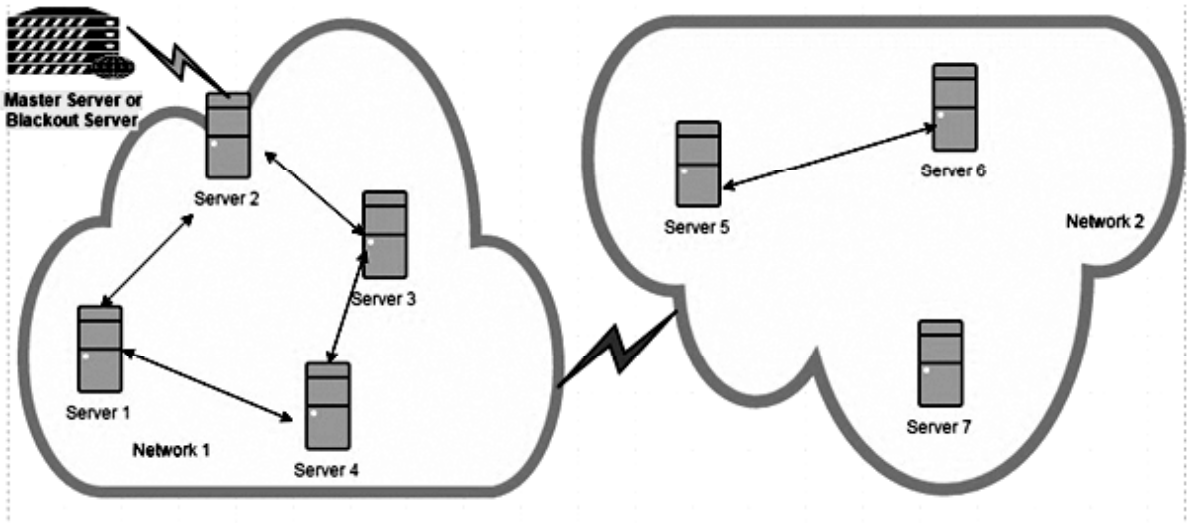

Figure 9: Master Server.

Records set in each server will be in the following way as shown in Fig. 10.

Records set can be any data model or sql table. Table contains a serial number. These serial numbers may be same as they visible or can be encrypted. This encrypted can be generated by system or manually. This may lead to more complex process. At the bottom we have timestamp. Timestamp can be encrypted if possible. Encrypted Timestamp is given in Fig. 11. All this information is maintained and monitored by a blackout server or a master server for every certain period. If it finds any disturbance in any of this data it should inform owner with a SMS request. Every time when records are placed in severs, this information is send to the owner with SMS. Owner can be alert with latest Records set. Because Records set will updated or modified or deleted timely. This information can help in tracking updated records set. By doing this exercise we can eliminate hacking to some extent. This process is also called as chain.

Testing is a process of investigation conducted to provide users with information about the quality of the product under test. Testing provides an objective, scope and view to allow business to appreciate and understand the risks of software implementation.

Testing is not a single day work. Systems should be tested for day-to-day to years. Even sometimes our system may not work properly. It is to be tested for re-installation of the software programs. Testing may give results in form of codes or messages. It depends upon the vendors who supply code.

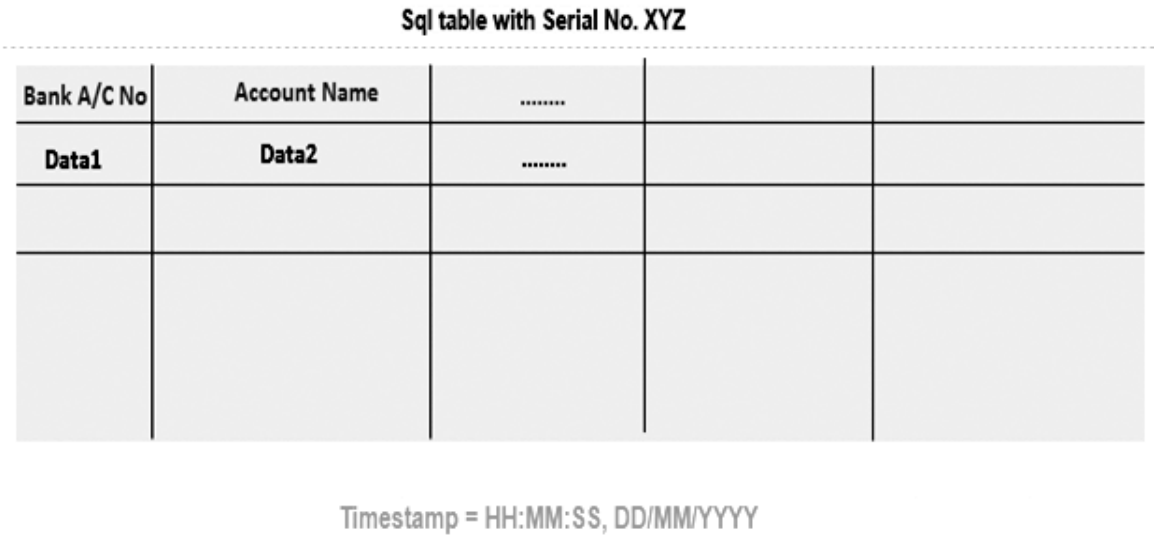

Figure 10: Record Set at each Server.

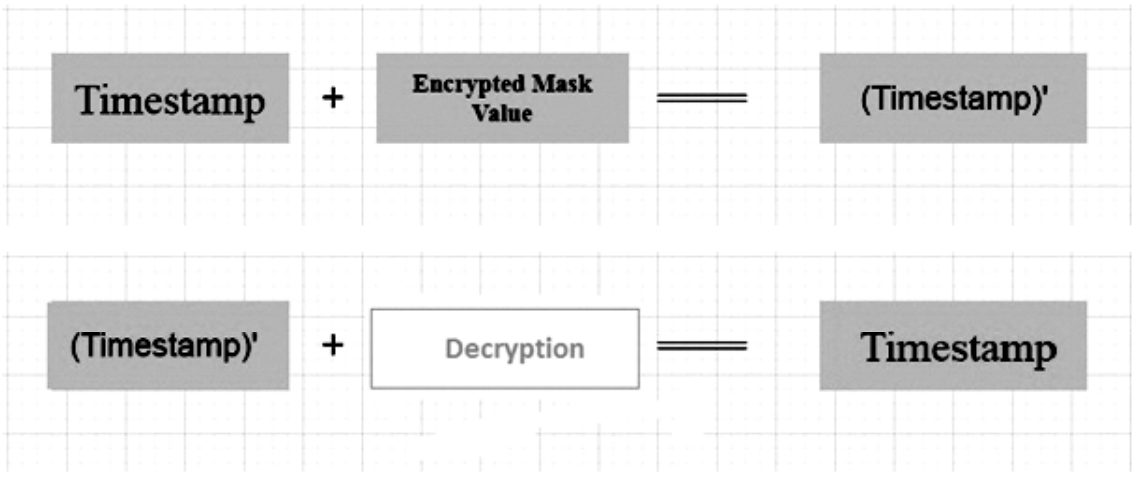

Figure 11: Encrypted and Decrypted Timestamp.

Same way records that are available in servers should be tested with backups or original data. Records can be modified or updated or deleted or create by accessible persons. History about hacking can be known by testing domain information available in the server with backup or original record sets. This type of testing is designed to determine whether there are any basic problems or bugs that will prevent it from malfunction or working. Software itself will go for testing original records with records available in server. If any sense of hacking, it will inform administrator or any IT personnel that system gone for odd. Based on available information professional will act.

All information discussed above should be implemented as programs. Java is best suited for this type of programs. Web server programming can be achieved by setting up and running an Apache server, use of Perl for web CGI programming and support administrative tasks, PHP, Java Servlets programming, Java Server Pages, XML, advanced Services with EJB and finally dot net.

When your PC or server is hacked, the hackers add some PHP code into the header of the homepage redirecting the user to a porn website. Your PC or server will get slow, its behavior will be odd from regular safe days. Denial of service attacks, the web server may crash or become unavailable to the users. Domain name system hijacking, the DNS setting is changed to point to the attacker's web server. Sniffing, unencrypted data sent over the network may be intercepted and used to gain unauthorized access to the web server. Phishing, the attack impersonates the websites and directs traffic to the fake website.

Testing backup records with server record sets, to know whether data is hacked or not. Encrypting timestamp is also known technique used for encrypting timestamp. Decrypting timestamp back to original timestamp is a process of Decryption. These techniques are known to all. Main technology in this paper is chain algorithm. Data record sets are same place in different servers same as images. Each server will maintain same data record sets as if original one that is available in first server. Main concern is testing data record sets in network. Graph related to server hacking can be reduced by following techniques related to avoid hacking as shown in Fig. 12.

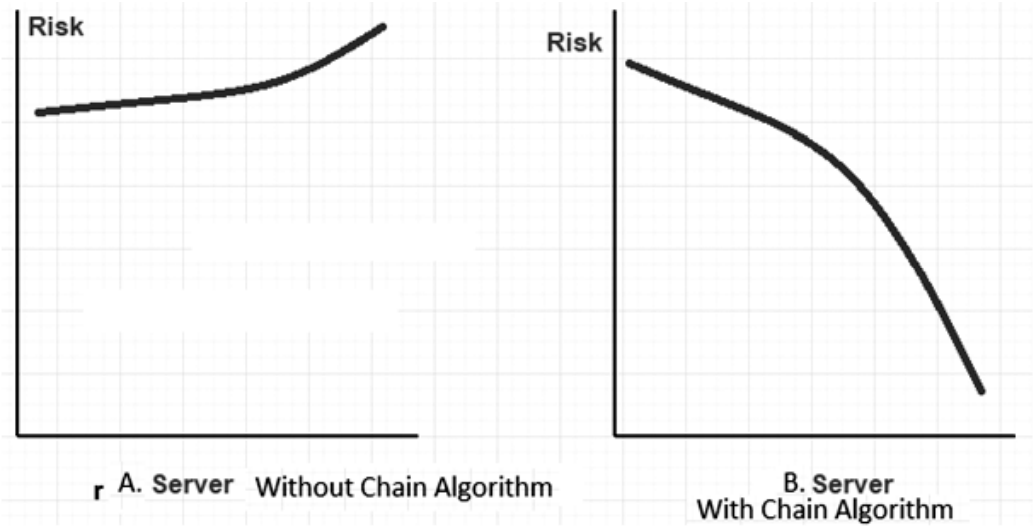

Figure 12: Difference between avoiding hacking.

6. CONCLUSION

There are lots of organizations suffering from different hackers. Not only organizations but lots of individuals are also made to suffer. This paper focuses only on organizations who suffer. To maintain secure data it is mandatory to follow such kind of methods. No data on internet or mobile is safe. Any organization or person should take care of their own servers and mobiles. Paper can guide to some extent for safety. There is scope for improvement in this paper.

ACKNOWLEDGMENT

My sincere thanks to my administrators and other professionals who gave me support and an idea of writing this paper successfully.

REFERENCES

[1] Philip K Chan “Machine Learning for Computer Security”.

[2] Gurpeet K Juneja “Ehtical Hacking: A Technique to enhance information security”.

[3] Sonal Beniwal, Sneha “Ethical Hacking: A Security Technique”.

[4] Susidharthaka Satapathy, Dr. Rasmi Ranjan Patra “Ethical Hacking”.

[5] Bhawana Sahare, Ankit Naik, Shashikala Khandey “Study of Ethical Hacking”.

[6] Kumar Utkarsh “System Security and Ethical Hacking”.

[7] Rebecca Schuller Borbely “On normalized compression distance and large malware”.

[8] Michael Bailey, Jon Oberhiede, Jon Andersen, Z. Morley Mao. Farnam Johanian, Jose Nazario, “Automated Classification and Analysis of Internet Malware”.

[9] Gerardo Canfora, Corrado Aaron Visaggio, “A Set of features to detect web security threats”.

[10] Annapurna Annadatha, Mark Stamp “Image Spam analysis and detection”.

[11] Ayesha Binte Ashraq, Zainab Abaid, Maliha Ismail, Muhammad Umar Aslam, Affan A. Syed, Syed Ali Khayam “Diagnosing bot infection using Bayesian inference”.

[12] Mila Dalla Preda, Federico Maggi “Testing android malware detectors against code Obfuscation: a systematization of knowledge and Unified methodology”.

[13] Yu – Keum Jeong, Roy C Park “Knowledge-based system and security”

[14] Hyung-Jin Mun, Kun-Hee Han “Blackhole attack: user identity and password seize attack using honeypot”.

[15] Robert Luh, Stefan Marschalak, Manfred Kaiser, Helge Janiche, Sebastian Schrittwieser “Semantics-aware detection of targeted attacks: a survey”.

[16] B Smith, William Yurick, D Doss: “Ethical Hacking: The Security justification redux”.

[17] Marilyn Leathers: “A Closer look at Ethical Hacking and Hackers”.

[18] U Murugavel, Dr. Shanthi: “Survey on Ethical Hacking Process in Network Security”.

[19] Kumar Utkarsh: “System Security and Ethical Hacking”.

[20] Chenchu Lakshmi S, P I Basarkod: “Basics of Ethical Hacking”.