IMPLEMENTING ADVANCED INRUSION DETECTION SYSTEM BY MONITORING NETWORK ANAMALIES AND USING ENCRYPTED ACCESS OF DATA

J. Arokia Renjit, Dr. K. L. Shunmuganathan

ABSTRACT

The Telnet, rlogin, rcp, rsh commands have a number of security weakness: all communications are in clear text and no machine authentication takes place. These commands are open to eavesdropping and tcp/ip address spoofing. SSH uses public/private key RSA authentication to check the identity of communicating peer machines, encryption of all data exchanged (with strong algorithms such as blowfish, 3DES, IDEA etc.). In this paper we proposed an IDS for encrypted access with SSH2 protocol to network public servers. Our proposed system detects the intrusions based on transferred data size and timing, which are available without decryption. The results reveal that the proposed system work well for different kinds of intrusions. Pre-operations are not needed and privacy is not violated. The detection is based on anomaly detection, which relies on the frequency of similar accesses and the characteristics of usual HTTP accesses.

Keywords: IDS, SSH, SSH2, MD5,MAC

1. INTRODUCTION

Intrusion detection system is one of the important topic of discussion in network security, that attempts to discover, alert, and possibly respond to an instance of undesired access[1]. The cost of viruses and worms during 2002 was estimated to cross 40 billion dollars [2]. There are lots of techniques that can be combined with IDS, one such attempt is SSH (SECURE SHELL PROTOCOL). SSH is the most reliable protocol for remote network login/access, which protects devices against attacks such as IP spoofing and plain text password interception. Generally SSH uses public key cryptographic techniques which have been broken, so here we employ a new cryptographic algorithm MD5 (Message Digest 5) algorithm into SSH for crypto-conversions. Therefore, the main challenge is to make this intrusion detection system more reliable and more authenticated than the previous IDS's. In recent years, a lot of research activities have been focused on traffic analysis of encrypted data stream [3, 4, 5, 6 and 7]. Widespread use of the SSH protocol greatly reduces the risk of remote computer access by encoding the transmission of clear text usernames and passwords, MD5 (Message-Digest algorithm 5) is a widely-used cryptographic hash function with a 128-bit hash value. Our proposed system is implemented on Snort intrusion detection software [8 and 9]. Our system observes and analyses the encrypted

ingress and outgress network traffic without decryption for anomaly detection. Statistically rare event and an event that differs from typical TCP access behaviours are supposed that have potential to be an intrusion. The proposed system detects an abnormal activity, based on access frequency and traffic properties of related network services. With help of average matrix can find the activity properties such as the volume of input and output traffic.

2. LITERATURE REVIEW

SECURE SHELL PROTOCOL:

SSH (Secure Shell) is a program to log into another computer over a network, to execute commands in a remote machine, and to move files from one machine to another. It provides strong authentication and secure communications over insecure networks. Its features include the following:

- Closes several security holes (e.g., IP, routing, and DNS spoofing). New authentication methods: .rhosts together with RSA based host authentication, and pure RSA authentication.
- All communications are automatically and transparently encrypted. Encryption is also used to protect integrity.
- X11 connection forwarding provides secure X11 sessions.
- Arbitrary TCP/IP ports can be redirected over the encrypted channel in both directions.
- Client RSA-authenticates the server machine in the beginning of every connection to prevent trojan horses (by routing or DNS spoofing) and man-in-the-middle attacks, and the server RSA-authenticates the client machine before accepting .rhosts or /etc/hosts.equiv authentication (to prevent DNS, routing, or IP spoofing).

An authentication agent, running in the user's local workstation or laptop, can be used to hold the user's RSA authentication keys. In the existing system, The Secure Shell Protocol uses public key encryption methods that are breakable with little effort. SSH provides two secure authentication methods that protect the login process from network-based sniffers: standard password and key based; in the latter standard public/private key mechanisms are employed. To use key based Authentication, a user puts in a file on the remote host her public key which is used in conjunction with private key stored on the connecting host during the authentication process. If the keys form a valid pair, the user is logged on. Generally, to get access to the private key, a user would be prompted for the

passphrase that protects the key. It is possible, however, for a user to have a private key without a passphrase. If a user does not have a passphrase on her key, a cracker with access to the private key file could have the user's complete access to the remote system.

SSH2 PROTOCOL

This section introduces some properties of SSH2 protocol, which is used in the proposed system [10, 11 and 12].SSH stands for secure shell and is a secure login program. SSH2 is based on SSH1, although almost totally rewritten and so does not contain the SSH1 protocol flaw[13]. Packet length represents the length of the packet and MAC field. Padding length is the number of octet representing the length of the padding. Packet data is the actual content of the message. Random padding is an arbitrary-length padding appended to packet data, so the payload reaches the block cipher size specified by the protocol. MAC corresponds to the message authentication code, which is computed after negotiate. The table 1 shows the MAC lengths.

MAC length (Bits)	Hash Algorithm
160	SHA-1 HMAC
96	SHA-1HMAC-96
128	MD5 HMAC
96	MD5 HMAC-96

Table 1. Mac's Length

Two end sides of SSH tunnel must agree on what MAC algorithm to use.With observing the mac algorithm-client-to-server and mac-algorithm-serverto-client fields in the packet of type 20, the MAC algorithm is found [14].

3. PROPOSED SYSTEM

The architecture of the proposed system is shown in figure 1. The proposed IDS is placed in the gateway of network and monitors the network traffic behaviours.

Our system observes and analyses the encrypted ingress and outgress network traffic without decryption for anomaly detection. This system has following steps:

A. Separating user activities

To separate user activities, this system deploys IP address. This means that for ingress traffic, the system uses source IP address and for outgress traffic, uses destination IP address. We assume that a client occupies an IP address during the activities and another client does not occupy the address at the same time. Then, the TCP connection for each user is reconstructed and the data packets belong to SSH2 protocol are extracted.

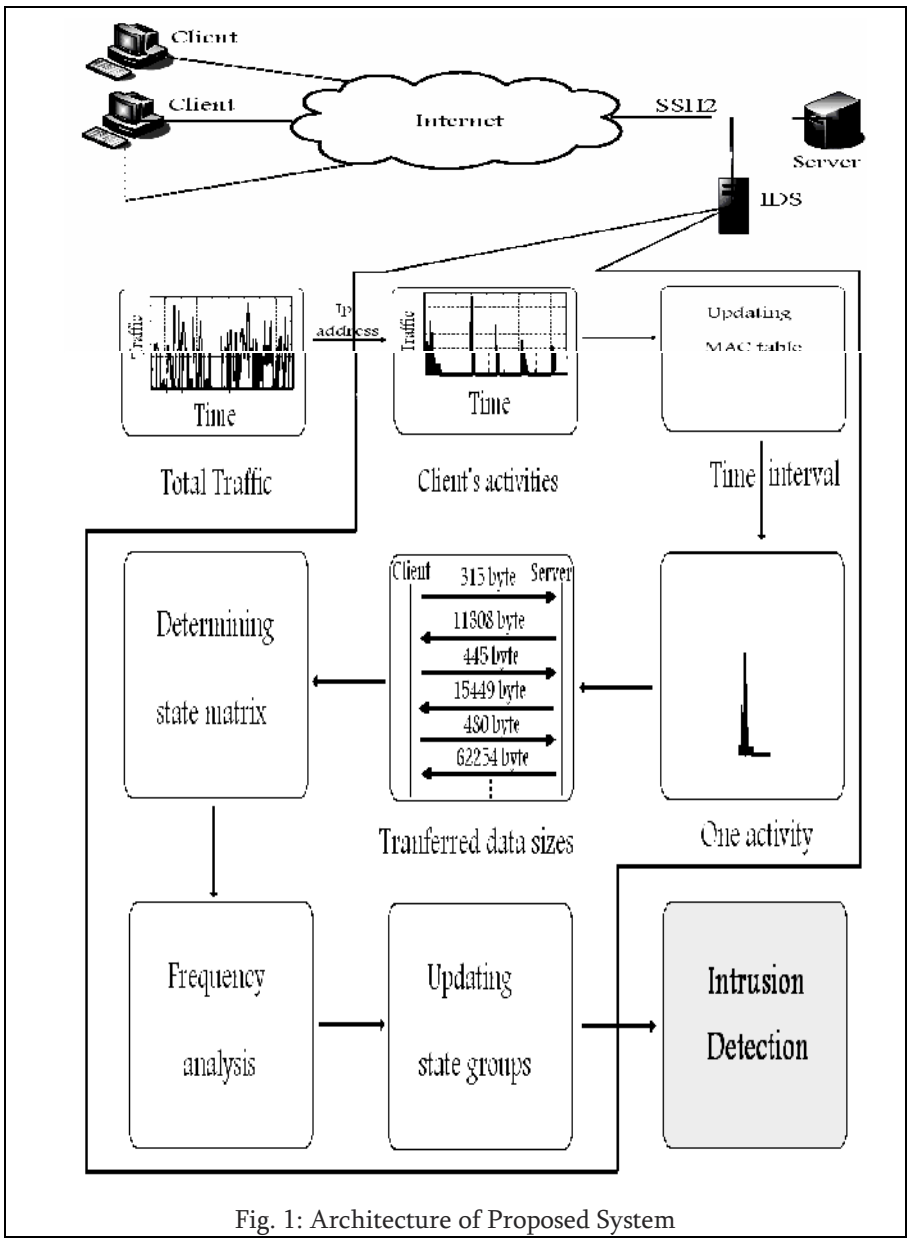


Fig. 1: Architecture of Proposed System

B. Updating MAC table

The MAC table has three columns: user number, user IP address and MAC size and each user has an entry (row) in this table. With this table and formula 1, the data size + padding size of SSH2 packet is computed.

C. Specifying user activities

Activities that belong to a client are divided into each activity based on interval of the packets. For instance, click a hyperlink by a user is an activity.

D. Generating state matrix for each activity

State matrix (Xi) for each activity is generated by the n sample of data size, which transferred between server and client processes. In this research, n is 15.Two n-set samples are selected one for client to server and other for server to client traffic. The samples are collected and stored by data size, then 15 largest samples selected from the two sets. State matrix is a 15 matrix for each activity.

E. Frequency analysis

Formula 1 computes the data size and pad size from SSH2 packet. The purpose of frequency analysis is to reduce the effect of pad size. If an activity carries out repeatedly, the data size is same but the pad size is different. Therefore, it is not possible to determine the number of times that an individual activity is occurred.

F. Updating state groups

In previous step, if an activity is a new activity, then a new group is created and if it is not a new activity then the parameters of group G min are updated .The group table keeps the ID, members number and group average matrix and do not need to keep members data. By assist of group table, the occurrence times of an individual activity can be determined. In addition, the transfer table is also used. This table shows the times of two activities ID m and ID m+1 which happen in sequence.

G. Intrusion detection

Each observed activity compared with the input or output traffic patterns of different services. In normal access, the input traffic (request) is small and the output traffic (reply) is big therefore a big input and small output traffic is rare and as possibly an intrusion. We use two Threq (maximum request size) and Thres (minimum reply size) parameters. For different network servers, these two parameters have different values. An activity is abnormal, if m1 elements of first row from the average matrix are more than Threq or if m2 elements of second row from the average matrix are less than Thres. The accuracy of detection is not high, if only the size of input and output traffic is used for intrusion detection.

To increase the accuracy, we used the threshold parameter Thf that shows the minimum times of a normal event and the value of this parameter is set with regards to the network behavior. If an activity has repeated less than Thf while the size of its traffic is abnormal, then this is a possibly attack. But the system sends a warning message to network manager, if only one of these conditions happens.

To detect the Flooding attacks [15], we are deployed a timer, a recognition table and a counter c. The counter counts the time intervals. The timer produces an interrupt in every t seconds. The table row j is relevant to the activity Gj. The recognition table has three columns a1, a2 and a3. The column a1 presents the frequency of an activity (frequency means the count of the occurrence of an activity) till the last time interval, column a2 presents the frequency of an activity in the new time interval and the column a3 presents the average frequency of an activity in all of time intervals. The initial value of the recognition table is zero. Finally for all G j if (a2 >> a3 or a2 > s × a3) then Flooding Attack is occurred. Where, the "s" is the threshold value of frequency factor. LRU (Least Recently Used) memory management method is used to manage the state group table rows. Proposed System optimizes the used memory with the LRU technique.

4. SYSTEM IMPLEMENTATION

1. Constructing an acyclic Topology and Path Selection:
   - In this module, we construct a acyclic topology.
   - Topology is constructed by getting the names of the nodes and the connections among the nodes as input from the user.
   - While getting each of the nodes, their associated port and IP address is also obtained.
   - For successive nodes, the node to which it should be connected to its parent node.
   - While adding nodes, comparison will be done so that there would be no node duplication.
   - Then we identify the source and the destinations.
   - In this module, we find the possible path for each of the destinations from the source.
   - After finding the possible paths, we find the cost associated with each of those paths.
2. Hop Login

In this module, we Login the Hops for Message Transfer.

Here validate hop name, hop password and port number.

These value are equals means separate server listen for transfer Message.

3. Encryption
   - In this module a file to be encrypted is chosen. The chosen file is given as the message transfer to Destination Hop.
   - A key governs the encryption and the key must be of 8 digits.
   - The Next Hop with in the Network means entered the correct key value after the header information decrypted to view the user.
   - The users can rights to modify the path value. After again encrypted the header information then transfer to next Hop.
   - These it continues up to final Hop identify.
4. Decryption
   - In this module decrypt to view header information.
   - The next Hop users enter the 8 digit key value.
   - The key value same means only the header information decrypt to view the user.
   - After the user change the path value.
5. File View
   - After changing path value identified that is final Hop means check the cost value.
   - The cost value are same means message decrypt to view the destination. Not same cost value means its display Access Denied

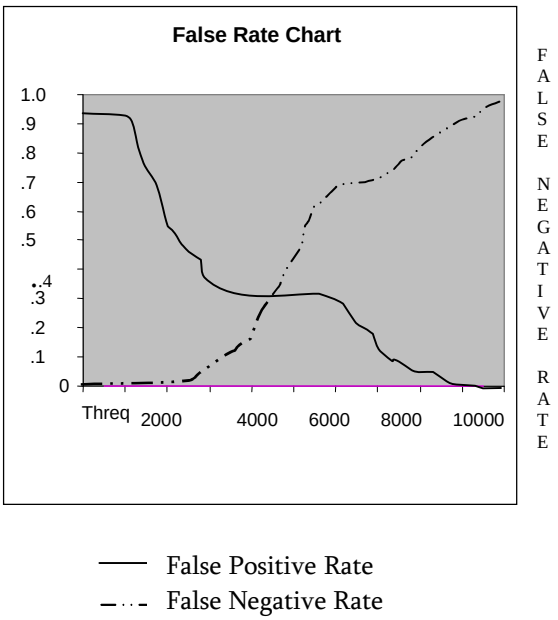
MD5 IMPLEMENTATION

MD5 processes a variable-length message into a fixed-length output of 128 bits. The input message is broken up into chunks of 512-bit blocks (sixteen 32-bit little endian integers); the message is padded so that its length is divisible by 512. The main MD5 algorithm operates on a 128-bit state, divided into four 32-bit words, denoted A, B, C and D. These are initialized to certain fixed constants. The main algorithm then operates on each 512-bit message block in turn, each block modifying the state. The processing of a message block consists of four similar stages, termed rounds; each round is composed of 16 similar operations based on a non-linear function F, modular addition, and left rotation.

5. EVALUATION AND RESULTS

We have implemented our proposed system on Snort intrusion detection software. Figure 2 shows the false rate for different values of Threq and Thres parameters. We assigned 10 to m1 and m2 and 1.3 to s in order to evaluate our system. We achieved 3700 and 600 for Threq and Thres respectively to minimize false rate through this evaluation. The false rate means false positive plus false negative.

The evaluation results show that our proposed system detects intrusions with a low false rate (about 15 percent). The scanning, script and buffer overflow attacks are detected with high accuracy.



6. CONCLUSION

In this paper we proposed IDS for encrypted access with SSH2 protocol to network public servers. Our proposed system detects the intrusions based on transferred data size and timing, which are available without decryption. The results reveal that the proposed system work well for different kinds of intrusions. Pre-operations are not needed and privacy is not violated. The detection is based on anomaly detection, which relies on the frequency of similar accesses and the characteristics of usual HTTP accesses.

Both pseudo-collisions and collisions for the compression function of MD5 have been demonstrated, though collisions for the full MD5

have not yet been achieved. Existing signatures formed using MD5 are not at risk and while MD5 is still suitable for a variety of applications

(namely those which rely on the one-way property of MD5 and on the random appearance of the output) as a precaution it should not be used for future applications that require the hash function to be collision-resistant.

7. ACKNOWLEDGEMENT

We take immense pleasure in thanking our Chairman Dr. Jeppiaar M.A, B.L, Ph.D, the Directors of Jeppiaar Engineering College Mr. Marie Wilson, B.Tech, MBA.,(Ph.D) Mrs. Regeena Wilson, B.Tech, MBA., (Ph.D) and the Principal Dr. Sushil Lal Das M.Sc(Engg.), Ph.D for their continual support and guidance. We would like to extend our thanks to my guide, our friends and family members without whose inspiration and support our efforts would not have come to true. Above all, we would like to thank God for making all our efforts success.

REFERENCES:

1. C. Endorf, E. Schultz and J. Mellander, "Intrusion Detection & Prevention", McGraw-Hill, ISBN: 0072229543, 2004.
2. Reuters.,"Virus damage estimated at $55 billion in 2003". Jan. 2004,
3. A. Hintz, "Fingerprinting websites using traffic analysis", Workshop on Privacy Enhancing Technologies, 2002.
4. G. Bissia, M. Liberatore, D. Jensen, and B. Levine, "Privacy Vulnerabilities in Encrypted HTTPStreams", Workshop on Privacy Enhancing Technologies, 2005.
5. Q. Sun, D. Simon, Y. Wang,W. Russell, V. Padmanabhan and L. Qiu, "Statistical identification of encrypted web browsing traffic", IEEE Symposium on Security and Privacy, 2002.
6. H. Cheng And R. Avnur, "Traffic Analysis of SSL Encrypted Web Browsing", Available at: Http://www.cs.berkeley.edu/?daw/teaching/cs261- f98/projects/final-reports/ronathan-heyning.ps, 1998.
7. S. Mistry and B. Raman, "Traffic Analysis of SSL-Encrypted Web Browsing", Available at: Http://bmrc.berkeley.edu/people/shailen/Classes/SecurityFall98 /paper.ps, 1998.
8. "Snort Intrusion detection system", Available at: www.snort.org/
9. [9] Rafeeq Ur Rehman, "Intrusion Detection Systems with Snort, Advanced IDS Techniques with Snort, Apache, MySQL, PHP,and ACID", Prentice Hall PTR, ISBN 0-13-140733-3, 2003.
10. Himanshu Dwivedi, "Implementing SSH Strategies for Optimizing the Secure Shell", Wiley Publishing, ISBN: 0-471-45880-5, 2004.

11. "The SSHv2 Protocol", available at: Http://cs.wellesley.edu/~cs342/SSH2Protocol.html
12. Daniel J. Barrett and Richard E. Silverman, "SSH, the Secure Shell, The Definitive Guide", O'Reilly & Associates Publishing, ISBN: 0-596-00011-1, 2001.
13. T. Ylonen and C. Lonvick, "The Secure Shell (SSH) ProtocolArchitecture", RFC 4251, 2006
14. G. Lu, B. Krishnamachari, C.S. Raghavendra, "An adaptive energyefficient and lowlatency MAC for data gathering in wireless sensor networks", Proceedings of 18th International Parallel and Distributed Processing Symposium, Pages: 224, 26-30 April 2004
15. J. Chirillo, "Hack Attacks Revealed", Willy computer publishing, ISBN: 0-471-41624-X, 2001.