IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 7, July 2016

Fire Wrap: A cyber defence mechanism against dark web attacks

Feby C Varghese[1], Varghese John[2], Dr. T. A. Ashok Kumar[3]

Student, Institute of Management, Christ University, Bengaluru, India[1, 2]

Assistant Professor, Institute of Management, Christ University, Bengaluru, India [3]

Abstract: Fire Wrap, a network based cyber defence system proposal which detect network anomalies by analysing the pattern of an incoming attack and distinguishes the attacker from the existing networked machines using the Boltzmann machine learning algorithm, then re-routes the incoming signal using double tunnelling approach to a sandbox environment where, the exit node vulnerability of onion routing is exploited to extract the raw data. The attacker will execute the attack in this sandbox environment and we can analyse the behaviour of the attack virtually without affecting the original network system, there by obtaining vital information which will help in forensic studies. The analysis part of Fire Wrap is carried out through Hopfield neural network, which is a simple recurrent network that can work as an efficient associative memory and can store and analysis data in a manner similar to the brain. The system is far advanced than any currently used firewalls and is developed to protect the surface web users from the cybercriminals which use the tor anonymous network to launch cyber-attacks from their hotspots in the dark web, which will steal all the confidential data and causes fatal damage to users including the defence network of a country. We are formulating this system with a provision of planning a counter attack which will make the system future proof.

Keywords: Tunnelling, dark web, neural network, Boltzmann, sandboxing

I. INTRODUCTION

The increased generation of data and the high demand for the security and confidentiality of the data have become a prime concern for all the data handlers. The traditional use of firewall has itself proven to have many drawbacks which are effectively being targeted for cyber-crimes. The most noted attacks are those from the dark web [1] to the middle web in which the use of onion routing [2] have utilized the best of layered tunnelling approach. Our Fire Wrap system model analyses an attack incoming and distinguishes the attacker from the existing networked machines using the Boltzmann machine learning algorithm and re-routes the incoming data using a double tunnelling approach to a sandbox environment where the exit node vulnerability[3] of tunnelling is exploited and allow the attackers encrypted data to execute in the sandbox environment such that the behaviour of the attack could be analysed virtually without the cost of affecting the original network.

The concept of Fire Wrap machine analysis part is carried out through the works of Hopfield neural network, which is a simple recurrent network that can work as an efficient associative memory and can store certain memories in a manner rather similar to the brain.

The inconsistency analysis and optimization is carried out by the network and the connection energy levels are defined from the analysis program. The connection weights are determined by comparing the cost function with energy function of the network.

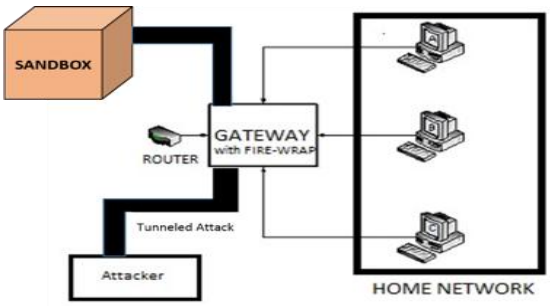

Figure 1: Fire - wrap architecture

The system is developed with the view of protecting the surface web users from the cybercriminals who launch their attacks from the dark web to take out confidential data and causes fatal damage to users including the defence system of a country to the local network of an organization. Onion Routing, which is assumed to be using one of the efficient routing methods to create the so called Tor Anonymous network for creating a tunnelled or a layered approach to initiate an encrypted destructive data. But with the use of double tunnelling concept and that of a sandbox approach, the behaviour of the destructive content is analysed for further cyber forensics since the received data packet is allowed to extract its true nature inside the virtual environment. In this research paper, we put forward a new system which could be much more secure and efficient than the traditional firewall concept which only scans an incoming packet. But, we are not replacing the existing firewall, since most companies

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 7, July 2016

invested heavily in building such capabilities, so Fire Wrap is designed to provide a secondary layer of protection to address the immediate data protection needs. But we are trying to develop the system with a vision for the future, where Fire Wrap will be a stand-alone defence system without the need of a firewall along with the capability of planning a counterattack against the malicious activities.

II. RELATED WORKS

Network security is paramount for any organization, usually the predictive approach is implemented to plan the resources allocation is used. An improved version of this method dealing with time series modelling is put forward in the paper "Predictive Modelling for Intrusions in Communication Systems using GARMA and ARMA models"[5], but the problem is that this kind of statistical analysis will only help if the pattern of attack can be predicted or if the source of attack has a known physical location. There are other systems like the Cyber Panel System[6], where various alert systems are used to sense the incoming attack and evaluate the correlation and look for a strategy to response. These type of defence systems are quite effective but if the attacks are through the tor anonymous network, then it is difficult to sense the attack, due to the complex algorithm used in carrying out the attack. Since they are protected by onion routing, it is difficult to trace back to the source, so it is a humongous task to formulate a strategy based on the cyber panel system. There are lot of security breaches that particularly affect the distributed computer system. To protect these kind of system there are several cyber defence systems, which combine network-based intrusion detection system & anomaly detection system, which uses signature matching and revels network anomalies by internet traffic data mining. These systems are really effective in a surface web, but the scale of complexity the dark web has and the dynamic algorithms they use in their process of attack needs more sophisticated defence system. This paves the way for Fire Wrap, since it is the only systems that can effectively defend the network from anonymous attack and do a forensic study on the nature of attack and the intention behind the attack due to its nature to adapt to the system requirements.

III. PROPOSED SYSTEM

Fire Wrap is network attack protector with enhanced modules for attack analysis in cyber forensics. It helps to understand the type of an attack so that developers can initiate even counter measures to build powerful network shields. The system also opens hope as an efficient alternative for existing network protectors which require efficient shield of their confidential data. System design is carried out mainly in three steps namely:

1. The attack identification
2. Double Tunnelling
3. Routing to sandbox

Each component is itself to be developed as individual modules such that it is scalable to meet the requirements of different sized networks.

The possible attack from an external attacker is identified by the Fire Wrap system using the Boltzmann Machine concept where the change in energy function raises a suspicion to monitor the activity. The change in energy indicates that an external node tries for an active connection to the home network. The next module is the double tunnelling which encapsulates the incoming data through a temporary channel or tunnel which ultimately routes to a virtual environment called the sandbox where the incoming data is allowed to be released and meet a duplicate environment of the internal system in the home network with faked data.

IV. DARK WEB ATTACK

A STUDY

The web we actually know and experience are the ones that can be indexed, this is termed as surface web. The search engines like google, bing, yahoo etc., can search only on the surface web, which itself has a size of about 1 billion. But, it is estimated that these indexed pages are only about 10% of the entire cluster known as the internet. The majority portion, which accounts for 90% of the cluster is the deep web. In the deep web majority sites are academic, scientific research, military databases. These are meant to be anonymously kept so as to protect the data from unauthorized access, which may result in catastrophes considering the gravity of the data stored. The layer below the deep web is called the dark web (sometimes it's not even be termed as a layer because of its integrity with the deep web, so it is often mistakenly referred as deep web) which the cybercriminals use as a gangland for their criminal activities[1]. The dark web is hidden inside the tor anonymous network[2], which acts as a protection layer from outside snooping combined with onion routing, provided a layered encryption to the data transferred from and to the dark web, which itself made the dark web to be called as onion land[3]. In the dark web you can hire hit men, buy drugs, weapons, see snuff films, human & animal tortures etc., just naming a few out of the estimated 45,000 service provided. There are lot of market places which act as a centralized hub for all the illegal trades. They uses crypto currencies like bit coin for their transaction, so it can't be traced back to the user. The only successful operation against the dark web was the Operation Ominous by FBI, where the original dark web market called the Silk Road was shut down and the founder Ross Ulbricht was arrested. But in a counter twist, admin of the site Dread Pirate Roberts (DPR) was not a single person but multiple person, so a new version of the site kept up and running in a short span of time. Also it made the rise of other marketplaces like Agora, which act as an alternative to Silk Road. All these points out to the complexity of the tor anonymous network and the challenges the security agencies face to tackle these cybercriminals[4].

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 7, July 2016

V. BOLTZMANN ANALYSIS

Every Boltzmann machine is often referred to as a stochastic neural network which is built up of one layer of visible units and one layer of hidden units. The visible neurons provide a platform between the network and the network's surroundings. On the other hand, the connections between the neurons or host are two ways and it is a symmetric network where the information can flows in both directions and the weights are the similar in both ways. In order to use the Boltzmann machine concept to analyse the networked hosts, we need to modify Hopfield's updating[8] rule. Since, the Boltzmann machine[9] operates by picking a unit at random, let say unit i, and twisting the state of unit i from S_i to $-S_i$ at temperature T with a probability of:

$$p_i = \frac{1}{1+e^{\left(\frac{\Delta E_i}{T}\right)}}$$

Where, ΔE_i can be formulated from the expression of the energy calculation of E , given by:

$$E = -\left(\sum_{i<j} w_{ij}\, s_i s_j + \sum_i \theta_i\, s_i\right)$$

Where:

- w_{ij} is the connection strength between host units j and i.
- s_i is the state of the machine, $s_i \in \{0,1\}$ of unit i.
- θ_i is the bias of unit in the global energy function.

Now, ΔE_i can be expressed as the difference of energies of two states.
i.e.

$$\Delta E_i = E_{i=off} - E_{i=on}$$

Here, ΔE_i denotes the global that results from a single unit being 0 (off) versus 1 (on). The temperature parameter T is necessary for the simulated annealing method of selecting the node i. If the spinning procedure is used constantly to the units, the units will change state and the relative probability of two global states is determined directly by their energy difference when the system rested to the thermal equilibrium and follows a Boltzmann

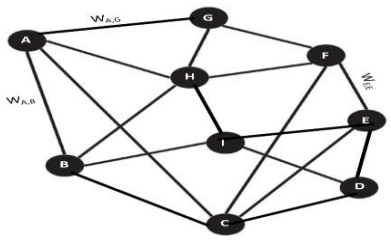

Figure 2: A network with few weights marked showing the node dependency

Even though the Boltzmann machine learning is a widely discussed algorithm for pattern matching and image processing[10], with the study on Fire Wrap, we propose implementing the basic concept of the same in identifying an energy function for a network so that any deviation from the standard value of the energy constant E, invokes a suspicion. This makes the system to proceed to the next phase of handling a possible attack.

A. Algorithm suspicion process (energy change)
1. Calculate energy constant E of the network.
2. Store each connection energy w_{ij} between two nodes, to variables.
3. For an incoming request message from an external node, if(E.change=0)

```
{
Packet.scan=no suspicion;
Accept();
}
Else
{
Create(double_tunnel)
Create(sandbox.process)
Exit();
}

Double_tunnel()
{
Encapsulate_rcvd(tunnelled_packet);
Packet.destination = sandbox(address);
Exit()
}
```

4. Prevent further connection from same address.

VI. DOUBLE TUNNELLING

The concept of double tunnelling very similar to the conventional example of covering a pipe with another protective pipe such that the ultimate destination of the inner content is determined by the outer pipe.
The module is initiated only when E.change=1 and such that there is initial suspicion been found with the primary investigation of the incoming packet. It routes the incoming to a virtual destination called the sandbox. The advantage of this system is that it can even route a tunnelled data to the required target and eliminates the traditional drawback of difficulty in extracting tunnelled data, especially the onion routed messages.
Here, we propose the tunnel mode IP sec as the basic encapsulation mechanism and the security guarantee of the tunnel, make use of the Internet key exchange protocol as the tunnel configuration. The tunnel protocols usually support protocols such as IP, IPX, Apple talk and so on.
Originally[11] IPSec was designed to transport IP packet, so it cannot support multi-protocol . But we can extend it and make suitable for multi-protocol environment. An extending method is encapsulating the non-IP protocol (protocol X) in the IP protocol using another tunnelling protocol such as GRE before IPSec encapsulation. Then

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 7, July 2016

IPSec encapsulation is adopted. In this way the encapsulation form becomes:

$$(IP(IPSec\ (IP(GRE(Protocol\ X))))$$

VII. SANDBOXING

A software sandbox is no different from the one built for a child to play. By providing a sandbox to a child we simulate the environment of real playground (in other words an isolated environment) but with restrictions on what a child can do. This I done in order to ensure that the child doesn't to get infected or we don't want him to cause trouble to others. What so ever the reason is, we just want to put restrictions on what child can do for Security Reasons.

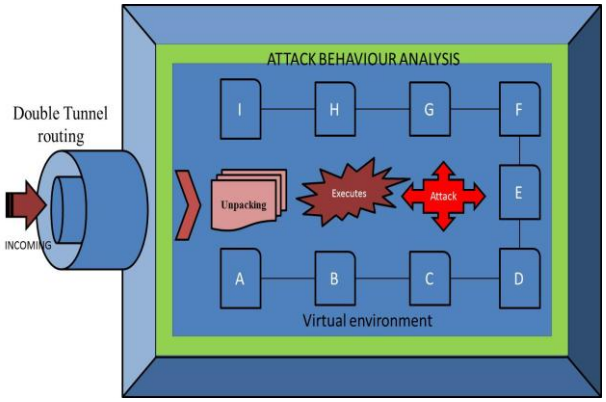

Figure 3: Sandbox Architecture using dynamic malware analysis mechanism

Execution of software sandbox is also of similar concept. We let any software(child) to play (execute) but with some restrictions over what it (he) can do. We can feel safe & secure about what the executing software can do. If you have a proper sandbox you can even run a virus infected file and stop all the malicious activity of the virus and see for yourself what it is trying to do. In fact, this will be first step of Antivirus researcher.

Fire Wrap uses this concept by exploiting the end node vulnerability[12] to get the actual harmful data onto a sandbox such that the original data is left unharmed. The conventional static malware analysis technique suffers limitations such that generally, the source code of malware samples are not readily available. Additionally, attacks relying on values that cannot be statically determined (e.g., current system date, indirect jump instructions) exacerbate the application of static analysis techniques. The other is that the malware authors know of the limitations of static analysis methods and thus, will likely create malware instances that employ these techniques to thwart static analysis. Therefore, it is necessary to develop analysis techniques that are resilient to such modifications, and are able to reliably analyse attacks that can even withstand a dark web intrusion.

The dynamic malware analysis technique executes a given malware sample within a controlled environment and monitor its actions in order to analyse the malicious behaviour. Since Dynamic Malware Analysis is performed during runtime and malware unpacks itself, dynamic malware analysis evades the restrictions of static analysis (i.e., unpacking and obfuscation issues). Thereby it is easy to see the actual behaviour of a program. Another major advantage is that it can be automated thus enabling analysis at a large scale basis.

However, the main drawback is so-called dormant code: That is, unlike static analysis, dynamic analysis usually monitors only one execution path and thus suffers from incomplete code coverage. In addition, there is the danger of harming third party systems, if the analysis environment is not properly isolated or restricted respectively. Furthermore, malware samples may alter their behaviour or stop executing at all once they detect to be executed within a controlled analysis environment.

Mainly two basic approaches for dynamic malware analysis can be distinguished:

.

- Analysing the difference between defined points:
  A given malware sample is executed for a certain period of time and afterwards the modifications made to the system are analysed by comparison to the initial system state. In this approach, Comparison report states behaviour of malware.
- Observing runtime-behaviour:
  In this approach, malicious activities launched by the malicious application are monitored during runtime using a specialized tool.

VIII. FUTURE SCOPE

The Fire Wrap system proposal also opens the scope for return attack method. This can be made possible through the fact that the double tunnelling ensures that the originality of the incoming data, whether it is tunnelled or even encrypted, is routed to a virtual environment and the same path can be used to retrace the location of the source of the attack. However, this involves effective and deeper analysis on the methods to implement because in case of an encrypted packet attack, the return attack should be such that it should retrace the packet source address and thus modify the packet header content. Further, the current system proposes as an added tool for the existing firewall architecture because installing a new infrastructure at once could incur more expense. But it could also be made as a standalone system.

REFERENCES

[1] Hsinchun Chen "Dark Web: Exploring and Mining the Dark Side of the Web" European Intelligence and Security Informatics Conference, 2011
[2] Roger Dingledine, Nick Mathewson, Steven Murdoch, Paul Syverson "Tor: The Second-Generation Onion Router" IEEE, 2014
[3] Jan Camenisch, Anna Lysyanskaya "A Formal Treatment of Onion Routing" IEEE

[4] Abhishek Sachan "Countering Terrorism through Dark Web Analysis" IEEE-20180, 2012
[5] Thulasy Ramiah Pillai, Azween Abdullah, Sellappan Palaniappan, Hafiz Muhammad Imran, "Predictive Modeling for Intrusions in Communication Systems using GARMA and ARMA models" IEEE 2015
[6] Laura S. Tinnel, O. Sami Saydjari, Joshua W. Haines "An Integrated Cyber Panel System" DARPA Information Survivability Conference and Exposition 2003
[7] Kai Hwang, Ying Chen, Hua Liu "Defending Distributed Systems Against Malicious Intrusions and Network Anomalies" IEEE International Workshop on Security in Systems and Network 2005
[8] Inderjeet Singh Behl, Ankush Saini, Jaideep Verma "Hopefield Network" International Journal Of Scientific Research And Education, 2013
[9] Ruslan Salakhutdinov , Geoffrey Hinton "Deep Boltzmann Machines" Proceedings of the12th International Conference on Artificial Intelligence and Statistics (AISTATS), 2009
[10] Nitish Srivastava, Ruslan Salakhutdinov "Multimodal Learning with Deep Boltzmann Machines" Journal of Machine Learning Research 15, 2014
[11] Zho Aqun, Yuan Yuan , Ji Yi aand Gu Guanqun "Research on Tunnelling Techniques in Virtual Private Network" IEEE, 2000
[12] Meenskashi Sharma , Supriya "Deep Web Data Extraction Using Query String Formation" International Conference on Reliability, Optimization and Information Technology, 2014.

BIOGRAPHIES

Dr. T. A. Ashok Kumar, M.C.A, M.B.A., M. Phil., Ph.D is working as Assistant Professor in Institute of Management, Christ University, Bengaluru, India since 2015. He has completed Ph.D in Computer Science MS University, Tirunelveli, Tamilnadu, India. Also, he completed Post Graduation in MCA from SNR Sons College, Coimbatore, MBA from Bharathiar University, Coimbatore and Master of Philosophy (M.Phil) from Bharathidasan University, Tiruchi, Tamilnadu. With 17 years of teaching experience he had conducted course curriculum for various universities and autonomous colleges in Tamilnadu. He also served as Chairman & Member of the Board of Studies in CMS College of Science & Commerce (Autonomous), Coimbatore, Bharathiar University and other various universities in Tamilnadu. He also member and editor for various International Journals in Computer Science like Inderscience Publications, SCI, AIRCC etc., His research interests are in Data Mining, focusing on Computer Networking, Distributed Computing and Human Resource Management. In addition, he has made numerous contributions to Data Mining and has examined the impact of Software Engineering Techniques and Applications on the design of various statistical models. Also, he has presented various papers in International and National Conferences.

Feby C Varghese, B.Tech has done his B.Tech in Electrical and Electronics Engineering from the University of Kerala, Kerala, India. He is currently pursuing his MBA in Lean Operations and System from the Institute of Management, Christ University, Bengaluru, India. He is a has a Six Sigma Green Belt certification from KPMG and an advanced diploma in Industrial Automation. He is a winner of paper presentation competitions at national level.

Varghese John, B.Tech has done his B.Tech in Computer Science Engineering from the University of Calicut, Kerala, India. He is currently pursuing his MBA in Lean Operations and System from the Institute of Management, Christ University, Bengaluru, India. He is a has a Six Sigma Green Belt certification from KPMG. He is a winner of paper presentation competitions at national level.