Malicious Code Detection Method over IPv4/IPv6 Tunneling Using Naive Bayesian Classifier

KYU-CHEOL OH, KI-HO LEE, YOU-JAE WON
Korea Information Security Agency
78,Garak-Dong,Songpa-Gu,Seoul,Korea 138-805
Korea

Abstract: - As IPv4/IPv6 transition mechanism, there are a Dual Stack, Tunneling, and Translation. Among them, tunneling may be misused in order for an malicious code to avoid a firewall system or intrusion detection system. In this study, the methodology which classifies the normal traffic and malicious traffic in IPv4/IPv6 tunneling environment, by using 'Naive Bayes Classifier' which shows an excellent performance for a text categorization is discussed.

In general, Internet worms or remote attack scripts include the certain features, that is, signature or machine instructions. Network packet can be supposed as one of general document. Accordingly, 'Naive Bayes Classifier' can be utilized for the network traffic analysis.

This study indicates the method which can detects the new form's malicious code in IPv4/IPv6 transition environment by applying 'Naive Bayes Classifier', and it can be effectively applied to an encapsulated packet.

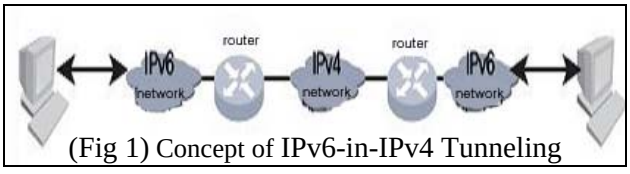
Key-Words: - Naive Bayesian Classifier, Malicious Code Detection, IPv4/IPv6 Transition, Tunneling, IDS(Intrusion Detection System)

1 Introduction

1.1 IPv4/IPv6 Tunneling and Security Vulnerability

As informationization has rapidly developed, inter-information device communication has been considered as a significant factor. Also, ALL-IP process has been quickly made in wire and wireless network environment.

Because currently-used IPv4 has 32bit address space, it can support 4.3billion addresses by arithmetic calculation. However, the Internet address is in danger of running out due to an inefficient address system which classifies the address in the class unit and rapidly-increasing number of Internet connection host. Also, as the limit of IPv4, the actual solution for a quality control(QoS) and security is insufficient. Hence, a demand for IPv6 infra expansion has gradually increased.



(Fig 1) Concept of IPv6-in-IPv4 Tunneling

But, it is actually difficult to alternate IPv4-based network with IPv6 once. Accordingly, the confounded network of IPv4 and IPv6 is expected to be used for some time. Tunneling as the actual traffic transition mechanism is available to support compatibility of IPv4 and IPv6.

However, this tunneling can be misused to avoid a firewall system or intrusion detection system. For example, when the firewall system is located only in IPv4 section due to costs in IPv4/IPv6 confounded network like (Fig 1), the firewall system which is set by IPv4 environment doesn't recognize IPv6 packet and so it cannot prevent the malicious attack.

1.2 Intrusion Detection System

One of representative methods which recognize the network attack is an operation of intrusion detection system. This system can be largely classified into Misuse Intrusion Detection and Anomalous Intrusion Detection.

Because Misuse Intrusion Detection patterns an existence of if-then form's rule or certain data and then inspects whether it is corresponded by using information on known attacks, there is low False

Positive Error. Because it searches only the attack pattern, it has an economical advantage but on the other hand, it cannot cope with unknown and new danger. On the other hand, because Anomalous Intrusion Detection considers an behavior which is against the modeled abnormal behavior as attack, there is low False Negative Error. But, because it must analyze the large capacity data for normal behavior modeling, there is a high implementation cost and high False Positive Error as disadvantages.

Also, For newly-appeared attack pattern, the detection methods which apply the methods of clustering, neural networks, associations rules, etc. which has been studied in the Anomalous Intrusion Detection have all the limitations of detection.

For the clustering, the learning speed is fast and it is easy to implement. But, the valuation function is expressed in a multi-dimension's euclidian distance. At this time, the result is different as the features and euclidian distance is set and the result is not good[1]. Association Rules find the irregular behaviors by defining the normal patter and by detecting irregularity among records, and is the most widely used[2]. Neural Network generates the data record's category and predicts the category which is included in the certain record. However, the above two cases are not proper to determine all the new worms or remote attack scripts because it is difficult to find what the most significant and important attribute of related pattern with only one data packet information.

There is difficulty in detecting the Internet worm and remote attack scripts which appear as the new danger, by applying the data mining in the intrusion detection system. But, according to the latest result of study, there was the good performance in the virus and normal programming classification by using the Naive Bayesian Classifier[3][8].

Naive Bayesian Classifier as one of algorithm which shows the most excellent performance in the text categorization, modified all the malicious codes and abnormal codes to a hexa value and classified by treating like the general text in [3].

This study is intended to present the method which can classify the normal data and malicious data (worm, attack script, etc.) for the traffic which is tunneled in IPv4/IPv6 confounded network, by using Naive Bayesian Classifier, as the network packet is considered as one of text.

2 Related Studies

2.1 Bayesian Learning

Bayesian Learning method is one of methods which are largely used for the problem which is learned with an algorithm to calculate the probably explicitly on the hypothesis.

Decision-making using the Bayesian Theorem means that the classification of the highest possibility is selected in the state that the feature value is given. It is hypothesized that the feature value is x, the classification is C, the probability distribution to feature value x in whole mother group is P(x), the prior probability that a random sample is included in Classification C is P(C), and the conditional probability that the variable value x will be gained from Classification C is P(x|C). When P(x|C), P(C), and P(x) are given, Probability P(C|x) that the feature value x will be included in Classification C in the given situation is calculated. The above probability is described in the following formula.

$$P(C \text{ and } x) = P(C)P(x|C)$$
$$= P(x)P(C|x) = \frac{P(x|C)P(C)}{P(x)} \qquad (2.1)$$

At this time, the closest hypothesis among many classifications means the Maximum A Posterior hypothesis (MAP). It is formulated with Bayesian theory again as follows.
Because P(C|x)map = : P(x) is independent for C, P(x) can be omitted.

$$= \max_{c} (P(x|C)P(C)) \qquad (2.2)$$

2.2 Naive Bayes

Being the most widely used among Bayesian learning methods, Naive Bayes classifier is a statistical algorithm which is the most largely used for the text categorization. Naive Bayesian Classifier can classify the text by using the statistical information from the training document, when the new document is given.

Naive Bayesian Classifier is formed as each Instance x is combined with the attribute values, and is applied to the learning process that the target function f(x) can take the value from the finite set V. When the target function's learning set is offered and the new instances which are expressed in the line of attribute

values $<a_1,a_2,\ldots a_n>$ are shown, Naive Bayesian Classifier predicts the target value or classification process for the new instance.

Bayesian approach to classify the new instance, allots the highest probability value , when the attribute value $a_1,a_2,\ldots a_n$ which explains the instance is given.

$$v_{MAP} = \arg\max_{v_j \in V} P(v_j | a_1, a_2, \cdots, a_n)$$

In here, the following formula can be expressed again by using the Bayesian Theory.

$$v_{MAP} = \arg\max_{v_j \in V} \frac{P(a_1, a_2, \cdots, a_n | v_j) P(v_j)}{P(a_1, a_2, \cdots, a_n)}$$

$$= \arg\max_{v_j \in V} P(a_1, a_2, \cdots, a_n | v_j) P(v_j) \quad (2.3)$$

In here, $P(a_1,a_2,\ldots a_n|v_j)$ and $P(v_j)$ can be calculated based on the training set. As each target value calculates the frequency which occurs from the learning data set, $P(v_j)$ can be easily estimated. But, it is not easy to estimate $P(a_1,a_2,\ldots a_n|v_j)$,unless we have a huge learning data set.

To calculate the reliable estimate, the number that multiplies the number of instances by possible target values must be estimated. Because NBC is based on the hypothesis that the attribute values are conditionally independent to the given target value, the given instance's target value can be indicated by multiplying the joint probability of $a_1,a_2,\ldots$ an by each attribute's probability.

That is, the formula is $P(a_1,a_2,\ldots a_n|v_j) = \prod_i P(a_i|v_j)$ When it is substituted to (2.3), Naive Bayes Classifier can be calculated.

Naive Bayes Classifier :

$$v_{NB} = \arg\max_{v_j \in V} P(v_j) \prod_i P(a_i | v_j) \quad (2.4)$$

In here, V_{NB} indicates the target value and result value by Naive Bayes Classifier. $P(a_i|v_j)$ is estimated from the learning set in Naive Bayes Classifier is the number which multiplies the number of attribute value by the number of target values. Because it is the probability mode, the bad performance may be shown if there is no sufficiently suitable data.

3 Example of Detection Applying Naive Bayesian Classifier

In the certain data, the following “3fbf 1e04“, ”fefc“, ”9a57“ sentences were found. When it is hypothesized that the probability from each classification is

P(“3fbf 1e04“|malicious)=2/9,
P(”fefc“|malicious)=3/9, P(”9a57“|malicious)=3/9,
P(malicious)=9/14, P(Normal)=5/14
P(“3fbf 1e04“|Normal)=3/5,
P(”fefc“|Normal)=1/5, P(”9a57“|Normal)=4/5,

the probability can be estimated as follows.

P(Malicious | 3fbf 1e04, fefc, 9a57)

$$= \frac{2/9 \times 3/9 \times 3/9 \times 9/14}{P('3fbf\ 1e04', 'fefc', '9a57')}$$

$$= \frac{0.0158729}{P(\text{"3fbf 1e04"}, \text{"fefc"}, \text{"9a57"})}$$

P(Normal | 3fbf 1e04, fefc, 9a57)

$$= \frac{3/5 \times 1/5 \times 4/5 \times 5/14}{P(\text{"3fbf 1e04"}, \text{"fefc"}, \text{"9a57"})}$$

$$= \frac{0.0342857}{P(\text{"3fbf 1e04"}, \text{"fefc"}, \text{"9a57"})}$$

When the probability value is formulated, the probability of being normal is high as below.

$$Malicious \quad \frac{0.0158729}{(0.0158729+0.0342857)} = 0.3164536$$

$$Normal \quad \frac{0.0342857}{(0.0158729+0.0342857)} = 0.6835463$$

However, if the probability of “ffff ffff” among the learning data that we obtained in here is followed,

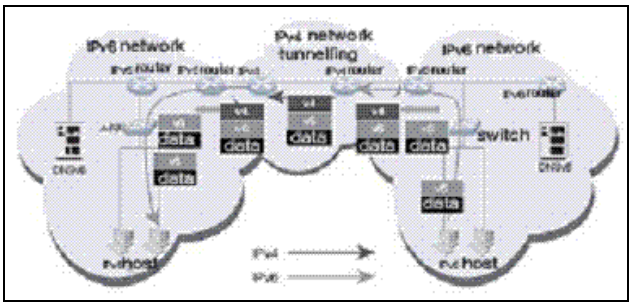
P("ffff ffff"|Normal) = 0/100
P("ffff ffff"|Malicious) = 87/100

Whatever the other attributes do, if there is “ffff ffff, it will be classified as 'Malicious'. Accordingly,

we used the smoothing method to prevent such the error.

For reference, there is the high false positive error in the signature (a special string which exists only in the certain malicious code)-based detection method which is used in the intrusion detection system or virus vaccine due to the above reason.

(Fig 2) is the example which applies IPv6-in-IPv4 tunneling for the network section which recognizes only IPv4 in a transmission route of IPv6 data. In the tunneling entrance, IPv6 data is encapsulated and transmitted to IPv4 packet. In order to describe that the encapsulated data is IPv6, IPv4 packet header's protocol field value is set to 41.



(Fig 2) Example of IPv6-in-IPv4 Tunneling

When IPv6-in-IPv4 tunnel is set, it is operated like a point to point link. That is, The whole tunnel is operated like one hop regardless of IPv4 tunnel's internal network configuration. If any problem occurs to the tunnel's internal link, the operator can hardly identify which link the problem occurs to.

Accordingly, IPv6-based firewall system which can inspect and filter the inside-flowing IPv6 traffic properly must be installed in the end of tunnel in order to solve the security vulnerability which avoids IPv4 firewall system by using IPv6-in-IPv4 tunneling. Also, the firewall system must convert ACL rule used to IPv4 in the site border for IPv6 to IPv6 rule and then must apply it. In other word, there needs the separate equipment and security policy for IPv4 traffic and IPv6 traffic.

4 Conclusion and Future Study

When the detection method using Naive Bayesian Classifier described in this study is applied, the normal traffic and malicious traffic can be classified in any position of data transmission route.

Because Naive Bayesian Classifier classifies by hypothesizing that all texts are the general text without classifying the general traffic or encapsulated traffic, Naive Bayesian Classifier can analyze IPv6 traffic which is encapsulated in IPv4 packet in spite of the intrusion detection system which is positioned in IPv4 section.

We will plan to precisely measure the detection's False Positive Error and False Negative Error, and we will execute the study which corrects the error and will compare the performance with the existing detection method.

References:

[1] Hyoseung Lee, Cheoljun Shim, Ilyong Won, Changhun Lee, “Comparative Analysis of unsupervised learning algorithm for network-based abnormal behavior detection model generation“, 2002 Korea Information Processing Society (KIPS) Autumn Conference

[2] W. Lee and S. J. Stolfo. "Data mining approaches for intrusion detection." In in Proceedings of the 1998 USENIX Security Symposium, 1998.

[3] Matthew G. Schultz, eleazar Eskin, Erez Zadok, and Salvatore J. Stolfo. "Data Mining Methods for Detection of New Malicious Executables." To appear in IEEE Symposium on Security and Privacy, May 2001.

[4] francisco Fernandez, "Heuristic engines," 11th International Virus Bulletin Conference, 2001

[5] Baudouin Le Charlier, Morton Swimmer, Addelaziz Mounji, "Dynamic detection and classification of computer viruses using general behaviour patterns," fifth International Virus Bulletin Conference, Boston, september 20-22, 1995.

[6] http://packetstormsecurity.nl/assessment.html

[7] Network intrusion accident investigation team, Network intrusion accident investigation result, 2003

[8] Study on new malicious code detection mechanism using 'Naive Bayes Classifier', 2003