走査活動観測に基づくネットワーク攻撃意図の推定

ゴ　キム　クォン†　　中村　康弘†

†防衛大学校
239-8686 神奈川県横須賀市走水 1-10-20
em53039@nda.ac.jp, yas@nda.ac.jp

あらまし 意図推定は攻撃行動とネットワーク環境によって攻撃者の目標を判断し、予測する能力である。攻撃者の意図を理解することは、管理者がネットワークリソースを保護するために正しい判断を下すことに必要である。攻撃目的の有用な証拠を再組み立てすることは、原因スキャン処理の不確実性のために困難であり、これらの活動の分析は攻撃の意図と戦略を推定するために、非常に有用である。この論文で、受信ネットワークトラフィックの攻撃の意図の分類は紹介され、その後のスキャン動作に基づいて、攻撃の意図モデルが提案され、攻撃者の意図を推定するために使用される。

Estimation of Attackers' Intentions Based on Observation of Scanning Activities

Ngo Kim Cuong†　　Yasuhiro Nakamura†

†National Defense Academy
1-10-20, Hashirimizu, Yokosuka, Kanagawa 239-8686, JAPAN
em53039@nda.ac.jp, yas@nda.ac.jp

Abstract

Intention estimation is the ability to judge and predict the attackers' goal according to attack behavior and network environment. Understanding an attacker 's intention can support the administrators make the right decision to protect the network resources. Reassembling the useful evidence of an attack purpose is difficult due to the uncertainty of the scanning processes, analyzing these activities are very helpful in estimating the intention and strategy of the attack. In this paper, a taxonomy of attack intentions in inbound network traffic is introduced, then the attack intention models based on scanning activities is proposed and used to estimate attackers' intentions.

1 Introduction

Intention estimation is the ability to judge and predict the goal of attackers according to attack behavior and network environment. Additionally, intent analysis plays an important role in the calculation of the essential threat value and it has become a hot research topics in network security area recently. There are different types of attack 's intentions that affect the network resources, and each of these intention has its aim or objectives for attacking the network. Some attack 's intention alters system resources or affect their operation thereby compromising integrity or availability while others attempts to learn or make use of information from the system but does not affect system resources thereby compromising confidentiality. These network attacker 's intentions basically can be separated into 4 main types: Information gathering for intrusion, Malware infection, Denial of Service (DoS) attack,

and others (unknown/unclear - intention, etc).

Network scanning is a reconnaissance operation aimed at finding services and potential vulnerabilities on a network, for the purpose of attacking them or deploying other intentions. Understanding an attacker 's intention can support the administrators make the right decision to protect the network resources. Reassembling the useful evidence of an attack purpose is difficult due to the uncertainty of the scanning processes, analyzing these activities are very helpful in estimating the intention and strategy of the attack.

In this paper, a taxonomy of attack intentions in inbound network traffic is introduced, then the attack intention models based on network scanning activities is proposed and used to estimate these attackers' intentions.

2 Related Work

There are many research papers proposed some different methods and various classification algorithms to detect Network Attack intention. Jordan Kiprop Koskei [1], used known information about an attacker 's behavior to create Hidden Markov Models, then decoded the alerts from an IDS (Snort) to discover the intruders high level intentions for the given alerts and predict the future intention. Qiu Hui and Wang Kun [2], proposed a dynamic real-time network attack intention recognition algorithm. By correlating real-time security alerts and vulnerabilities, they found the spread route and stage of attacks based on graph theory and probability theory, then identified the attack intention and predicted the possible transition of attacks. Xinzhou Qin and Wenke Lee [3], presented an approach to identify attack plans and predict upcoming attacks. They developed a graph-based technique to correlate isolated attack scenarios derived from low-level alert correlation based on their relationship in attack plans.

All of these three research papers have one thing in common: only focus on detecting the attack activity, but not observing the scanning which could early warn of the next step in intrusion. Sometimes, these proposed systems still have false discoveries.

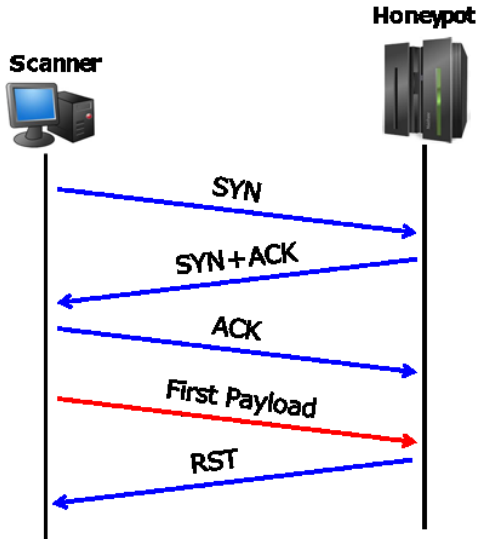

Figure 1: Network Observation System

3 Proposed Method

3.1 Network Observation System

To collect the network attack traffic data, this research setup an observation system on a gateway of the network that allowed to capture all packets that targeted to unused IP address range, but severely limited outgoing connections to minimize damage by the attacker. This reactive observation system responds SYN+ACK packet to SYN request connection via TCP/IP, then this system will get the first payload which is sent to the target (as show in Figure 1). By observing the scanning activities on these inbound network traffic might reveal the purpose of attackers.

3.2 Features for Observation

In order to observe the network scanning activity in the inbound network traffic datasets, it needs to extract certain features that characterize the current trends in malicious traffic. This research targets to analyze the first payload from attacker, which includes the network scanning activities, malwares infection activities, and so on. According to the results of preliminary investigation of scanning packets,

the arrival payload activities are investigated, then the following several facts are revealed:

- This observe system can only get the first payload for each connection request.
- For each destination IP address can be received many payloads for any time.
- The same source IP address usually continuously sends the same payloads.
- Sometime, the same payloads are received from many different source IP addresses.

From these above conditions, six features are selected for each source IP address as shown below:

1. **dstip**: Number of destination IP address
2. **sport**: Number of source port
3. **dport**: Number of destination port
4. **hash** : Number of Payload's Fuzzy hash value
5. **span** : Average time for each arrival time of Payload, [s]
6. **sdev** : Standard deviation of the **span**, [s]

Six kinds of features each has three or four ranges are indicated in Table 1, then total number of classes are: 4x3x3x4x4x3 = 1728 classes.

Table 1: Feature Value Range

Name	Range of Value
`dstip`	1, 2-9, 10-99, 100-
`sport`	1, 2-99, 100-
`dport`	1, 2-99, 100-
`hash`	1, 2-9, 10-99, 100-
`span`	<0.1, 0.1-0.9, 1.0-10.0, 10.0<
`sdev`	<1, 1-10, 10<

All of these classes are generated as a decision tree based on C4.5 algorithm. To each source IP address, each packet will be extracted 6 features then be classified into one of attack type class based on prepared decision tree. This process repeats every 10 minutes to detect the change of scanning behavior for each source IP address, thus the scanning activities can be observed.

3.3 Taxonomy of Attack Intentions using Latent Dirichlet Allocation Algorithm

The purpose of this research is to estimate the attack 's intention that are contain in the inbound traffic data, but it is still not clear what type of attack exactly are exist in captured traffic data. Based on these 6 features above to observe the scanning behaviors, it can easily classify into 2 main attack 's intentions: Network Probing attack, Denial of Service attack, and all others unknown intentions. In order to cluster these scanning activities result to 3 attack 's intention groups, this research uses Latent Dirichlet Allocation algorithm (LDA) to partition scanning class items in a data set into subgroups. LDA is unsupervised,so it does not need labelled samples. After running LDA it would end up with a number of unnamed groups, each containing tokens related to that groups.

For example: source IPs have the following set of scanning classes:

- Source IP 1 (113.200.235.38): c212222; c211232, c212221, c211211.
- Source IP 2 (180.97.161.225): c111121, c111221.
- Source IP 3 (185.94.111.1): c431432, c321332, c431132, c4321132.
- Source IP 4 (177.23.85.195): c421112.
- Source IP 5 (119.187.191.180): c111132.

Latent Dirichlet allocation is a way of automatically discovering groups that these scanning classes contain. For example, given these set of scanning classes and asked for 2 groups, LDA might produce something like

- Source IP 1 and 2: 100% Group A
- Source IP 3 and 4: 100% Group B
- Source IP 5: 60% Group A, 40% Group B

Group A: 17% c212xxx, 17% c111xxx, ⋯ (at which point, it could interpret group A to be about DoS Attack)

Group B: 25% c43xxx32, 17% cxxx132 ⋯ (at which point, it could interpret group B to be about Probe Attack)

In more detail, LDA represents source IPs as mixtures of groups that spit out scanning class name with certain probabilities (attack

intentions). It assumes that source IPs are produced in a group by the following fashion:

- Decide on the number of scanning class name or the type of scanning the source IP will have.
- Choose a group mixture for the source IP (according to a Dirichlet distribution over a fixed set of K groups, in this case K=3). For example, assuming that it has the two DoS attack and Probe attack groups above, it might choose the source IP that consist of 1/4 c43xxx32 and 1/6 cxxx132.
- Generate each scanning class name w_i in source IP by:
  - First picking a group (according to the multinomial distribution that have sampled above; for example, it might pick the group A with 1/3 probability and the group B with 2/3 probability).
  - Using the group to generate the scanning class itself (according to the group s multinomial distribution). For example, if selected the group A, it might generate the scanning class“ c212xxx ” with 17% probability, “ c111xxx ” with 17% probability, and so on.

Assuming this generative model for a collection of source IP addresses, LDA then tries to backtrack from the source IPs to find a set of groups that are likely to have generated the collection.

3.4 Attack Intentions

Many Internet attacks involve various network activities that enable remote attackers to locate a open services or open ports or even a vulnerable system, to gain unauthorized access to a target host, or to aim at disrupting the normal service of a specific target system, etc. In this section, this research surveys three common malicious activities: Network Probing attack, DoS-DDoS attack, and others. These three malicious activities are only a subset of many Internet attacks occurring lately, but their prevalence has drawn much attention from this research based on the results of preliminary investigation of inbound traffic data.

1. Network Probing attack: The purpose of which is to identify specifics about network resources. Sometime attacker is undoubtedly check for open ports on routers, firewalls, even on victim computer and identify what system services are available for exploitation.
2. Denial-of-Service(DoS) attack: An attack in which an attacker floods a target system with malicious traffic in order to prevent legitimate access. DoS attacks can both overload the victim host or its Internet link and causes a partial or complete service disruption.
3. Other intentions: all other unknown or unclear intentions.

, 4 Experiments and Results

By calculating the change in 6 determined features to each source IP address, it could define the scanning class for every 10 minutes. The Table 2 below shows the scanning class name for each source IP address by every time nterval in the day 2016/04/11.

Table 2: Scanning Class for each source IP

Source IP	Scanning Class Name
99.92.91.85	c111132
99.75.190.61	c111132
99.70.223.89	c111132
99.66.213.233	c111132
99.244.209.71	c111132
. . . . .	. . . . .
110.4.52.12	c111143
110.248.30.109	c111132
109.65.28.236	c111131
104.224.41.53	c211232

Finally, apply LDA algorithm to the result above,it would end up with 3 unnamed groups, each containing tokens related to that groups.

Here are some of the typically scanning classes that are classified to each group:

1. Group 1: c122321, c411141, c221342, c132312, c431313, c312221, c232241, c221331, c221343 ...
2. Group 2: c411141, c311221, c332442, c311223, c131223, c331211, c411412, c311222, c132312 ...
3. Group 3: c122321, c111342, c411141, c232243, c312311, c132232, c413113, c413441, c411412 ...

According to these 9 typically scanning class names in each group and based on the results of preliminary investigation of inbound traffic data, each group's name can be determined as:

1. Group 1 is Unknown attack intentions;
2. Group 2 is DoS, DDoS attack;
3. Group 3 is Probe attack.

5 Conclusion

This research shows that by using only 6 determined features, the inbound traffic data can be investigated and analyzed scanning activities for each source IP address. Then applied the LDA algorithm to these scanning classes it can classify into 3 groups of attacks intention, that are Probe Attack, Denial of Services Attack and all other unknown Attack Intentions.

It could be the initial idea to classify and estimate the attack intention that are contained in inbound traffic data, and it might not be complete, because there are still many unknown intentions and unclear intentions are not be recognized in the 3nd group.

The next research will improved the partion algorithm to get the better results in classifying attack intentions. Futhermore, it would generate an weighted average value to calculate the average of possibilities value of an attack intention type in each group. (To estimate the percentages of possibilities of an attack intention taking place in each group)

References

[1] Jordan Kiprop Koskei, "An Attacker Intention Discovery Layer for Intrusion Detection System using Hidden Markov Models", Master Thesis, 70 pages, Publisher Biblio Bazzar, 2012.

[2] Qiu Hiu and Wang Kun, "Real-time Network Attack Intention Recognition Algorithm" International Journal of Security and Its Applications, Vol.10, No.4, pp.51-62, Oct. 2016.

[3] Xinzhou Qin and Wenke Lee, "Attack Plan Recognition and Prediction Using Casual Networks", The 20th Annual Computer Security Applications Conference (ACSAC), 1997.