Exploitation of NetEm Utility for Non-payload-based Obfuscation Techniques Improving Network Anomaly Detection

Ivan Homoliak, Martin Teknos, Maros Barabas, and Petr Hanacek

Faculty of Information Technology, BUT,
Bozetechova 1/2, 612 66 Brno, Czech Republic
ihomoliak@fit.vutbr.cz
http://www.fit.vutbr.cz/~ihomoliak/.en

Key words: NetEm • Network anomaly detection • Intrusion detection • Obfuscation • Evasion • ADS • NBAD • Naive Bayes

1 Introduction

The impact of a successfully performed intrusion can be very crucial. There exists a lot of space which needs research in order to improve detection capabilities of various types of intrusions. Therefore, many researchers and developers are encouraged to design new methods and approaches for detection of known and unknown (zero-day) network attacks. These facts are the most important reasons why Anomaly Detection Systems (ADS) intended for intrusion detection arose. Network ADS (further ADS) approaches attack detection by utilizing packets' headers and communication behavior, not the content of the packets. Thus, basic principles of ADS open possibilities of an attacker to evade ADS detection by obfuscation techniques.

The goal of our work is to train the ADS detection engine to be aware of the behavior of obfuscated attacks, and thus correctly predict other similar obfuscated attacks. The obfuscation techniques leveraged in our current research are based on non-payload-based modifications of connection-oriented communications. Our work instantiates ADS features by *Advanced Security Network Metrics* (ASNM) [4], which are aimed at offline intrusion detection. In our previous work, experiments showed interesting intrusion detection capabilities on CDX 2009 dataset. But the possibility of evading an intrusion detection employing such features still exists, which is the subject of our current research.

2 Related Work

Although non-payload-based evasions of network attacks in the area of intrusion detection were considered as an actual research subject of more than one and a half decades ago [2, 5, 6], it revealed to be actual a few years ago as well [1]. There exist several related works considering non-payload-based evasions of network

attacks for payload-based intrusion detection, however, there is a lack of works performing investigations into non-payload-based network behavior anomaly detection and this kind of evasion.

3 Obfuscation Tool

We designed and implemented a tool for automatic exploitation of network services which is able to perform various obfuscation techniques based on NetEm utility and *ifconfig* Linux command. Execution of direct attacks (non-obfuscated ones) is also supported by the tool as well as capturing network traffic.

The most suggested obfuscations are performed by *tc* utility and its extension NetEm [3], respectively. NetEm enables us to add latency of packets, loss of packets, duplication of packets, reordering of packets and other outgoing traffic characteristics of the selected network interface. The modification of MTU is performed by the linux utility *ifconfig*. Table 1 presents instances of these techniques and contains appropriate empirically recognized parameters.

4 Data Mining Experiments

All experiments were performed in Rapid Miner Studio [7] using a 5-fold cross validation and conditional probability based Naive Bayes classifier.

Forward Feature Selection Experiment

For the purpose of finding the best subset of ASNM features [4], we performed the forward feature selection (FFS) method. The experiment considered two class prediction – the first for legitimate traffic and the second for intrusive traffic.

The experiment consisted of two executions of the FFS. The first took as input just legitimate traffic and direct attack entries (denoted as FFS DL), and represented the case where ADS was trained without knowledge about obfuscated attacks. The second execution took as input the whole dataset of network traffic – consisting of legitimate traffic, direct attacks as well as obfuscated ones (denoted as FFS DOL), and thus, represented the case where ADS was aware of obfuscated attacks.

Binominal Classification Experiment

A 5-fold cross validation was performed using direct attacks with legitimate traffic considering FFS DOL features. The classifier achieved average recall of 99.35%, while it correctly predicted 98.71% of direct attacks. The classifier trained on all direct attacks and legitimate traffic instances was then applied in the prediction of the whole dataset (including obfuscated attacks) and it correctly predicted 71.25% of obfuscated attacks and 78.26% of all attacks respectively. The achieved result proclaimed the existence of some successful obfuscations of attacks which were predicted as legitimate traffic.

Table 1: Experimental obfuscation techniques with parameters

Technique	**Instance**	**ID**
Spread out packets in time	• constant delay: 1s	(a)
	• constant delay: 8s	(b)
	• normal distribution of delay with 5s mean 2.5s standard deviation (25% correlation)	(c)
Packets' loss	• 25% of packets	(d)
Unreliable network channel simulation	• 25% of packets damaged	(e)
	• 35% of packets damaged	(f)
	• 35% of packets damaged with 25% correlation	(g)
Packets' duplication	• 5% of packets	(h)
Packets' order modification	• reordering of 25% packets; reordered packets are sent with 10ms delay and 50% correlation	(i)
	• reordering of 50% packets; reordered packets are sent with 10ms delay and 50% correlation	(j)
Fragmentation	• MTU 1000	(k)
	• MTU 750	(l)
	• MTU 500	(m)
	• MTU 250	(n)
Combinations	• normal distribution delay ($\mu = 10ms$, $\sigma = 20ms$) and 25% correlation; loss: 23% of packets; corrupt: 23% of packets; reorder: 23% of packets	(o)
	• normal distribution delay ($\mu = 7750ms$, $\sigma = 150ms$) and 25% correlation; loss: 0.1% of packets; corrupt: 0.1% of packets; duplication: 0.1% of packets; reorder: 0.1% of packets	(p)
	• normal distribution delay ($\mu = 6800ms$, $\sigma = 150ms$) and 25% correlation; loss: 1% of packets; corrupt: 1% of packets; duplication: 1% of packets; reorder 1% of packets	(q)

In the next part of the current binominal classification experiment, we performed 5-fold cross validation of the whole dataset including obfuscated attacks. The classifier achieved average recall of 99.63%, while it correctly predicted 99.37% of all attacks. Therefore, we confirmed the assumption that a classifier trained with knowledge about some obfuscated attacks is able to detect the same or similar obfuscated attacks later.

Comparing the results of the current experiment reproduced with the FFS DL feature set, we concluded that the model using FFS DL features had achieved slightly better results in learning direct attacks and legitimate traffic characteristics than the case of the first model (using DOL features), but on the other hand, it resulted in more misclassified cases of obfuscated attacks than the first one (i.e. 155:138) as well as it achieved worse results in cross validation of the whole dataset.

5 Summary of the Obfuscation Techniques

The results presented in the section originate from a binominal classification experiment in which the classifier is trained without obfuscated attack knowledge and validated on the whole dataset. The obfuscations are considered successful if they are predicted as legitimate traffic, and therefore the situation represents the ADS evasion case. The most successful obfuscations use combinations of more techniques (i.e. o, q, p), damaging of packets (i.e. f, e) and spreading out packets in time with delays specified by normal distribution (i.e. c). From the MTU modification techniques, (n) appear to be the most successful.

(Non) Exigency of a Network Normalizer

If we would assume existence of an optimal network normalizer for ADS which would be able to completely eliminate the impact of proposed non-payload based obfuscation techniques, then these obfuscation techniques would be useless. If such optimal network normalizer would exists, then it would still be prone to state holding and CPU overload attacks.

Contrary, if we would not assume network normalizer as part of ADS system, then non-payload-based obfuscation techniques might be employed as training data driven approximation of network normalizer, which would not be prone to previously mentioned issues and attacks.

Acknowledgment

This article was created within the project Reliability and Security in IT (FIT-S-14-2486) and supported by The Ministry of Education, Youth and Sports from the National Program of Sustainability (NPU II); project IT4Innovations excellence in science – LQ1602; and Aggregated Quality Assurance for Systems (AQUAS) – 8A17001.

References

1. Boltz, M., Jalava, M., Walsh, J.: New Methods and Combinatorics for Bypassing Intrusion Prevention Technologies. Tech. rep., Stonesoft (2010)
2. Handley, M., Paxson, V., Kreibich, C.: Network Intrusion Detection: Evasion, Traffic Normalization, and End-to-End Protocol Semantics. In: 10th USENIX Security Symposium. pp. 115–131 (2001)
3. Hemminger, S., et al.: Network Emulation with NetEm. In: Australia's 6th National Linux Conference. pp. 18–23. Citeseer (2005)
4. Homoliak, I., Barabas, M., Chmelar, P., Drozd, M., Hanacek, P.: ASNM: Advanced Security Network Metrics for Attack Vector Description. In: Proceedings of the International Conference on Security & Management (SAM). pp. 350–358 (2013)
5. Ptacek, T.H., Newsham, T.N.: Insertion, Evasion, and Denial of Service: Eluding Network Intrusion Detection. Tech. rep., DTIC Document (1998)
6. Puppy, R.F.: A look at Whisker's Anti-IDS Tactics (1999), `http://www.ussrback.com/docs/papers/IDS/whiskerids.html`
7. RapidMiner: RapidMiner Studio, `https://rapidminer.com/products/studio/`