Research Paper

Volume : 4 | Issue : 6 | June 2015 • ISSN No 2277 - 8179

Alert Aggregation Based Pattern Classifier for Effective Network Intrusion Detection System

Engineering

KEYWORDS : Alert aggregation, intrusion detection, Pattern Classification, Spam Filters, Multimodal Biometric Methods, Biometric Authentication Systems, Good Word Attacks

Fabi A.K M.Tech CSE AWH Engineering College Calicut, India

ABSTRACT

Alert aggregation is an important subtask of intrusion detection. The goal is to identify and to cluster different alerts—produced by low-level intrusion detection systems.—belonging to a specific attack instance which has been initiated by an attacker at a certain point in time. A novel technique for online alert aggregation which is based on a dynamic, probabilistic model of the current attack situation was produced. Basically, it can be regarded as a data stream version of a maximum likelihood approach for the estimation of the model parameters.

Detecting intrusions in networks has become one of the most critical tasks to prevent their misuse by attackers. The rapid increase in network traffic and attacks made the Intrusion Detection Systems to fail in terms of accuracy and efficiency in many situations. Present networks and enterprises follow a layered defence approach to ensure security at different access levels by using a variety of tools such as network surveillance, perimeter access control, firewalls, network, host and application intrusion detection systems, data encryption and others. Given this traditional layered defence approach, only a single system is employed at every layer which is expected to detect attacks at that particular location. The main goal of this approach in Intrusion Detection System is to achieve high accuracy and efficiency.

I. Introduction

In Pattern classification systems machine learning algorithms are used to perform security-related applications like biometric authentication, network intrusion detection, and spam filtering, to distinguish between a "legitimate" and a "malicious" pattern class. The input data can be purposely manipulated by an adversary to make classifiers to produce false negative. This often gives rise to an arms race between the adversary and the classifier designer. Well known examples of attacks are: Spoofing attacks where one person or program purposely falsifying data and thereby gaining an illegitimate advantage [1][2],modifying network packets belonging to intrusive traffic manipulating contents of emails [3],modifying network packets belonging to intrusive traffic. Mainly three main open issues are identified: (i) analyzing the vulnerabilities of classification algorithms, and the corresponding attacks (ii) developing novel methods to assess classifier security against these attacks (iii) developing novel design methods to guarantee classifier security in adversarial environments. Machine learning is used to prevent illegal or unsanctioned activity which is created from adversary. Machine learning is used in security related tasks involving classification, such as intrusion detection systems, spam filters, biometric authentication. Measuring the security performance of these classifiers is an essential part for facilitating decision making.

II. Related works

It is highly desirable to have an efficient security strategy in the context of spoofing the biometric systems used to perform automatic personal authentication [1]. The basic necessity of universality, distinctiveness, and permanence make the biometric traits like finger print, face, voice, hand geometry, iris pattern etc. very suitable for personal authentication process. But the spoofing of biometric system is a serious threat faced by secure personal authentication. Some fusion schemes that can increase the security of multimodal biometric systems seem to be highly efficient when one of the modes is successfully spoofed. When compared with traditional fusion methods these fusion schemes appear to have a promising potential against spoof attacks since these methods take into account the intrinsic security of each biometric system being fused, besides the matching score and sample quality score.

The use of multimodal biometric systems became a trend in the context of spoofing where an imposter will fake a biometric trait [2]. It is assumed that an imposter must spoof all modalities in the system to be accepted. A method for assessment of multimodal systems and underlying fusion algorithms deserves significance when an imposter spoofs only a subset of modalities used in the system and leaves the remainder as zero effort attempts. In this method many traits are analyzed and made them into a normalized score. All these normalized scores are fused to form a single score thereafter. If this score exceeds a predefined threshold it is assumed that the operator is genuine and if it is otherwise the operator is an imposter.

Statistical spam filters have been useful to overcome the problem of spam emails [3]. But the senders of spam often bypass these filters by systematically modifying their email messages by inserting or appending words that looks like as that in legitimate email. In the active good word attacks two pairs of messages are taken; in the first pair one is legal and the other one is spam where as in the next pair the messages differ only by one word.

The automatic signature generators (ASGs) are used as a defence against fast propagating worms [4]. Many attacks like allergy attack aim these systems. Allergy attack which is a much less recognized problem aims to introduce false positives. Type II Allergy Attack employs the date encoded in URLs. The Type II Allergy Attack against an old corpus is found more powerful than that against a fresh one.

In the fingerprint spoofing and face spoofing several methods are adopted to mock the traits of the users [5]. This may be as follows, In finger print spoofing ,1.The user presses his finger on a soft material such as wax, play doh, dental impression material, or plaster, to create the negative impression of the fingerprint as a mold; 2. a casting material such as liquid silicon, wax, gelatin, moldable plastic, plaster or clay, is poured in the mould; 3. when the liquid is hardened, the fake fingerprint is formed.

SpamAssassin is one of the most well-known and widespread open-source spam filters [6]. It is made up basically of a set of 'if-then' rules, each one of which is dedicated to a different characteristic of an email which can be useful to determine either its spamminess or its hamminess. A score is associated with each rule: higher scores denote a higher degree of spamminess.

Pattern recognition systems have been widely deployed in security-sensitive applications like spam filtering, malware detection, and biometric authentication [7]. A face verification system based on Principal Component Analysis (PCA), where each client is authenticated by comparing the submitted face image with the stored template belonging to the claimed identity, in the feature space induced by PCA is supposed to have a desirable advantage.

Security applications like malware detection, intrusion de-

tection, and spam filtering employ Support Vector Machines (SVMs) as a most popular classification technique. The following steps can be adopted to correctly use the framework in specific application scenarios [8]: 1. hypothesize an attack scenario by identifying a proper adversary's goal, and according to the taxonomy in 2. define the adversary's knowledge, and capabilities; 3. formulate the corresponding optimization problem and devise the corresponding attack strategy; 4. resample the collected (training and test) data accordingly; 5. evaluate classifier's security on the resampled data (including attack samples); 6. repeat the evaluation for different levels of adversary's knowledge and/or capabilities, if necessary; or hypothesize a different attack scenario .

Inexact string matching is a possible solution to the problem of filtering obscured text [9]. Spam filters that rely on traditional word-based features may be defeated by the spam attack of obfuscation, in which words are intentionally obscured through the use of misspellings, character substitutions, insertions, and deletions. The use of character level k-mers is one basic method of inexact string matching.

The commonly used statistical spam filters seem to be vulnerable to adversarial attacks. One of the more common adversarial attacks, known as the *good word attack*, spoils spam filters by appending to spam messages sets of 'good' words, which are words that are common in legitimate email but rare in spam [10]. An email is classified as spam if at least one instance in the corresponding bag is spam, and as legitimate if all the instances in it are legitimate.

III. Existing System

Pattern classification systems based on classical theory and design methods do not take into account adversarial settings; they exhibit vulnerabilities to several potential attacks, allowing adversaries to undermine their effectiveness. A systematic and unified treatment of this issue is thus needed to allow the trusted adoption of pattern classifiers in adversarial environments, starting from the theoretical foundations up to novel design methods, extending the classical design cycle of. In particular, three main open issues can be identified: (i) analyze the vulnerabilities of classification algorithms, and the corresponding attacks. (ii) Developing novel methods to assess classifier security against these attacks, which are not possible using classical performance evaluation methods. (iii) Developing novel design methods to guarantee classifier security in adversarial environments.

A. Disadvantages

1. Poor analyzing the vulnerabilities of classification algorithms, and the corresponding attacks.
2. A malicious webmaster may manipulate search engine rankings to artificially promote website

IV. Proposed System

A framework was developed for the empirical evaluation of classifier security at design phase that extends the model selection and performance evaluation steps of the classical design cycle. Three main ideas emerged. First, to pursue security in the context of an arms race it is not sufficient to react to observed attacks, but it is also necessary to proactively anticipate the adversary by predicting the most relevant, potential attacks through a what-if analysis; this allows one to develop suitable countermeasures before the attack actually occurs, according to the principle of security by design. Second, to provide practical guidelines for simulating realistic attack scenarios, define a general model of the adversary is defined, in terms of her goal, knowledge, and capability, which encompass and generalize models proposed in previous work. Third, since the presence of carefully targeted attacks may affect the distribution of training and testing data separately, a model of the data distribution that can formally characterize this behaviour, and that allows us to take into account a large number of potential attacks is found desirable. It is advantageous to have an algorithm for the generation of training and testing sets to be used for security evaluation, which can naturally accommodate application-specific and heuristic techniques for simulating attacks.

A. Advantages

1. New system prevents developing novel methods to assess classifier security against these attacks.
2. The presence of an intelligent and adaptive adversary makes the classification problem highly non-stationary.

V. Implementation

A. Attack Scenario and Model of the Adversary

Although the definition of attack scenarios is ultimately an application-specific issue, it is possible to give general guidelines that can help the designer of a pattern recognition system. This model is based on the assumption that the adversary acts rationally to attain a given goal, according to her knowledge of the classifier, and her capability of manipulating data. This allows one to derive the corresponding optimal attack strategy.

B. Pattern Classification

Multimodal biometric systems for personal identity recognition have received great interest in the past few years. It has been shown that combining information coming from different biometric traits can overcome the limits and the weaknesses inherent in every individual biometric, resulting in a higher accuracy. Moreover, it is commonly believed that multimodal systems also improve security against Spoofing attacks, which consist of claiming a false identity and submitting at least one fake biometric trait to the system (e.g., a "gummy" fingerprint or a photograph of a user's face). The reason is that, to evade multimodal system, one expects that the adversary should spoof all the corresponding biometric traits.

C. Adversarial Classification

Assume that a classifier has to discriminate between legitimate and spam emails on the basis of their textual content, and that the bag-of-words feature representation has been chosen, with binary features denoting the occurrence of a given set of words

D. Security Modules

Intrusion detection systems analyze network traffic to prevent and detect malicious activities like intrusion attempts, ROC curves of the considered multimodal biometric system under a simulated spoof attack against the fingerprint or the face matcher. Port scans, and denial-of-service attacks. When suspected malicious traffic is detected, an alarm is raised by the IDS and subsequently handled by the system administrator. Two main kinds of IDSs exist: misuse detectors and anomaly-based ones. Misuse detectors match the analyzed network traffic against a database of signatures of known malicious activities. Their training set is constructed, and periodically updated to follow the changes of normal traffic, by collecting unsupervised network traffic during operation, assuming that it s normal.

VI. Conclusion

During the overall development of the project the well-known fact that there is no one single classifier, which can outperform all the other classifiers in all the datasets was consistently proved. The combination of multiple classifiers certainly improved the accuracy, but at the cost of heavy increase in computation time. This intrinsic method is application-specific.

Volume : 4 | Issue : 6 | June 2015 • ISSN No 2277 - 8179

REFERENCE

[1] Robustness Of Multimodal Biometric Fusion Methods Against Spoof Attacks, R.N. Rodrigues, L.L. Ling, and V. Govindaraju | [2] Multimodal Fusion Vulnerability to Non-Zero Effort (Spoof) Imposters ,P. A. Johnson 1, B. Tan 2, S. Schuckers 3 ECE Department, Clarkson University Potsdam, NY 13699, USA 1 johnsopa@clarkson.edu 2 tanb@clarkson.edu 3 sschuckers@clarkson.edu. | [3] Good Word Attacks On Statistical Spam Filters, Daniel Lowd Dept. of Computer Science and Engineering University of Washington Seattle, WA 98195-2350 lowd@cs.washington.edu.,Christopher Meek Microsoft Research Redmond, WA 98052 meek@microsoft.com | [4] Advanced Allergy Attacks: Does A Corpus Really Help?, Simon P. Chung and Aloysius K. MokDepartment of Computer Sciences, University of Texas at Austin, Austin TX 78712, USA {phchung,mok}@cs.utexas.edu | [5] Security Evaluation Of Biometric Authentication Systems Under Real Spoofing Attacks, Battista Biggio, Zahid Akhtar, Giorgio Fumera, Gian Luca Marcialis, and Fabio Roli Department of Electrical and Electronic Engineering, University of Cagliari Piazza d'Armi, 09123 Cagliari, Italy | [6] Evading Spamassassin With Obfuscated Text Images, Battista Biggio, Giorgio Fumera, Ignazio Pillai, Fabio Roli and Riccardo Satta University of Cagliari, Italy | [7] Pattern Recognition Systems Under Attack, Fabio Roli, Battista Biggio, and Giorgio Fumera | Dept. of Electrical and Electronic Engineering, University of Cagliari, Piazza d'Armi, 09123 Cagliari, Italy {roli, battista.biggio, fumera}@diee.unica.it, WWW home page: http://pralab.diee.unica.it | [8] Security Evaluation Of Support Vector Machines In Adversarial Environments, Battista Biggio, Igino Corona, Blaine Nelson, Benjamin I. P. Rubinstein, Davide Maiorca, Giorgio Fumera, Giorgio Giacinto, and Fabio Roli. | [9] Spam Filtering Using Inexact String Matching In Explicit Feature Space With On-Line Linear Classifiers. D. Sculley, Gabriel M. Wachman, and Carla E. Brodley Department of Computer Science, Tufts University Medford, MA 02155, USA {dsculley, gwachm01, brodley}@cs.tufts.edu | [10] A Multiple Instance Learning Strategy For Combating Good Word Attacks On Spam Filters, Zach Jorgensen ZDJORGEN@JAGUAR1.USOUTHAL.EDU Yan Zhou ZHOU@CIS.USOUTHAL.EDU Meador Inge WMI601@JAGUAR1.USOUTHAL.EDU School of Computer and Information Sciences University of South Alabama Mobile, AL 36688, USA