Game Theory and Intrusion Detection Systems

By Anis Alazzawe, Asad Nawaz, Murad Mehmet Bayraktar

ISA 767-Secure E-Commerce
Spring 2006
Prof. Cervesato

Abstract

Intrusion detection systems have long been utilized in detection and response strategy to potential attacks. However, effective policing and finding a right balance between tradeoffs has always been an issue. In fact, it has been the focus of very extensive research. Nevertheless, it remains an inexact science as no solution has yet been found. The balancing of tradeoffs must still be accomplished according to individual network necessities. Tradeoffs could include IDS sensitivity versus false alarm rates. Even after obtaining a satisfactory solution, the system may wish to have a response strategy. Traditional intrusion detection schemes are not evolved enough for complex detection and response strategy. Therefore there is a need for logically formalizing not only effective detection policies but also effective response policies. In an attempt to overcome the shortcoming of traditional intrusion detection systems, we have described a model for a game theoretic approach to intrusion detection systems. Game theory has been widely studied and implemented in disciplines like economics, war strategy, etc. Research in the realm of networking and intrusion detection has been limited, though. Presented is an introduction to traditional IDS and challenges associated therewith, introduction to game theory, a proposition of foundations for the theoretical limits of a game theoretic approach to IDS, current available prototypes, common attacks and countermeasures thereof, and some limitations of the game theoretic approach.

Introduction

As technology evolves, so do the potential threats. To defend against these threats, corporations, companies, and even individuals have developed numerous countermeasures. Specifically in the area of networking, to prevent or detect an increasing amount of intrusions, intrusion detection systems (IDS) have been utilized. As time has passed, IDSs have also evolved and become increasingly popular. Although one cannot be naive and expect a perfect solution, an optimal solution is always desired and sought. The challenges with modern IDSs are many. An IDS has to make quick decisions, decisions on-the-fly. Therefore, somehow, an IDS has to not only gather and collect various information, but also process it and apply actions based upon the given policies. However, there are also many challenges to IDS policies. First, the policy can only be as good as the administrator that writes them. Second, effective policing has been the focus of quite extensive research. Tradeoffs must be adequately balanced such as what can be the acceptable rate of false positives, false negatives, etc. Therefore, we introduce the concept of game theory into modern IDS solutions. Game theory provides a way of mathematically formulizing the decision making process of policy establishment and execution. Thus, a logical methodology is applied to the selection of certain threshold values. Game theory has been studied and applied in many fields and applications. However, there has been limited and minimal research in the area of networking and more specifically in the realm of intrusion detection systems. Therefore, it is the object of our research to analyze the available game theoretic approaches for intrusion detection systems. We will also attempt to provide a concise problem definition and identifications of interactions between players of the game theoretic systems. Furthermore, we will describe a model for the theoretical limits of a game theoretic approach to IDS. This paper will further contain an in-depth investigation of data analysis based decision and intrusion detection processes through game theoretic models. We also will introduce an ever-evolving decision and control framework for intrusion detection systems to address areas in attack modeling, analysis of possible threats, and decision on response actions. Our main objective is to conduct a deep research and analyze the future of IDS in the game theoretic approach which lies in data classification and correlation. The IDS of tomorrow will produce results by examining input from several different sources. The way to solve this challenge lies in classification analysis and predictive artificial intelligence performed on strange data sets. Intrusion detection systems face several daunting, but exciting challenges in the future and are sure to remain one of our best weapons in the arena of network security. But in order to fully comprehend and appreciate the advances made by a game theoretic approach to intrusion detection systems, one must fully understand the current intrusion detection techniques and drawbacks thereof. Thus, a brief overview of modern intrusion detection systems is necessary.

Traditional IDS Systems

Intrusion detection systems have become increasingly popular in detecting and defending against intrusions into a given network. But how does one define an attempted intrusion into a system. In order to get a formal definition of an intrusion attempt, we look at the early works of the famous J. P. Anderson. In 1980, J. P. Anderson defined an intrusion attempt [1] as the potential possibility of a deliberate unauthorized attempt to

1. Access information
2. Manipulate information
3. Render a system unreliable

Thus the study of such intrusions and countermeasures thereof is immensely important. In fact a study reported by Purdue University [3] shows an increasingly dangerous trend in network intrusions. Specifically, the study shows that information theft is up over 250% in the span of five years most recent years. The study further points out that 99% of all major companies have reported at least one major incident. Finally, the cost of fraud has increased dramatically in the areas of telecom and computer totaling to $10 billion in the US alone. These are staggering statistics but it should be noted here that the intrusions could take one of many forms [4]. These intrusions can include but are not limited to

1. Attempted compromises to system
2. Masquerade attacks
3. Denial of Service
4. Any type of malicious use

To deal with the above-mentioned attacks and all intrusions in general, many intrusion detection systems have been formalized. It is worthwhile to explore some of these systems in a little depth.

Types of IDS

To counter the increasing number of intrusions and intrusion attempts into a given system, companies are deploying intrusion detection systems. Each solution has its own advantages and also its distinct disadvantages. Generally, the primitive intrusion detection systems simply detect all traffic and discern malicious traffic from legitimate use. Modern and more evolved intrusion detection systems now have the ability to take responsive or pre-emptive actions. Furthermore, some intrusion detection systems can be classified as passive and reactionary to others that are aggressive and the most famous and widely deployed network intrusion detection systems are misuse detection and anomaly detection.

Misuse Detection

Misuse detection systems are not unlike virus detection methods. They simply rely on pattern or signature recognition wherein each pattern or signature represents a malice activity (intrusion). Therefore, each attack is also represented in a form of a pattern or signature. Depicted in figure 1 below is a typical misuse detection system [5].

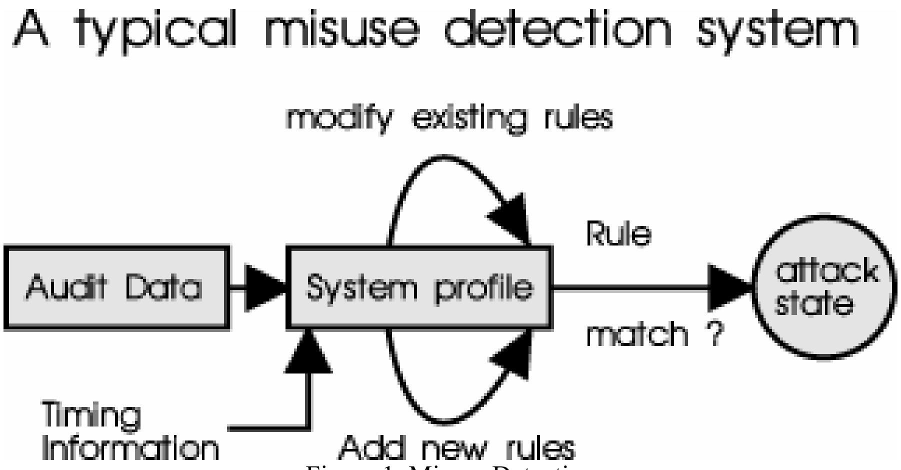


Figure 1. Misuse Detection
source: www.cs.utk.edu/~abdulrah/ netsecurity/paper.html

The misuse detection systems do, however, have many challenges. One such challenge would be developing a signature that would represent all variations of a pertinent attack. Surely a signature or pattern of a known attack can be slightly modified to fool a misuse detection system. Another challenge would be controlling false positives. False positives are activities flagged by a system as intrusive when in fact they are legitimate uses. A false positive is usually attained when a signature representing an intrusive activity also matches a non-intrusive activity. A high false positive rate can render the entire intrusion detection system useless as an effect of repeatedly "crying wolf".

Anomaly Detection

Anomaly detection systems usually do not present a significant amount of false positives. They detect intrusive behavior by defining all intrusive activities as anomalous. Therefore, all anomalies are by default intrusions. In order to determine what an anomaly is, the system consults a profile engine. The profile engine typically contains a normal activity profile among other things. This normal activity profile, as its name suggests, contains all legitimate activities the given system can perform or can be preformed on. Thus, all activities are checked with the profile engine to determine whether they match a certain activity in the normal profile. If they do match an activity within the profile, they are considered to be normal activities and subsequently legitimate use. However, if an activity is detected that does not match an activity within the normal activities profile, it is flagged as an intrusive activity and appropriate measures are taken as specified by the administrator.

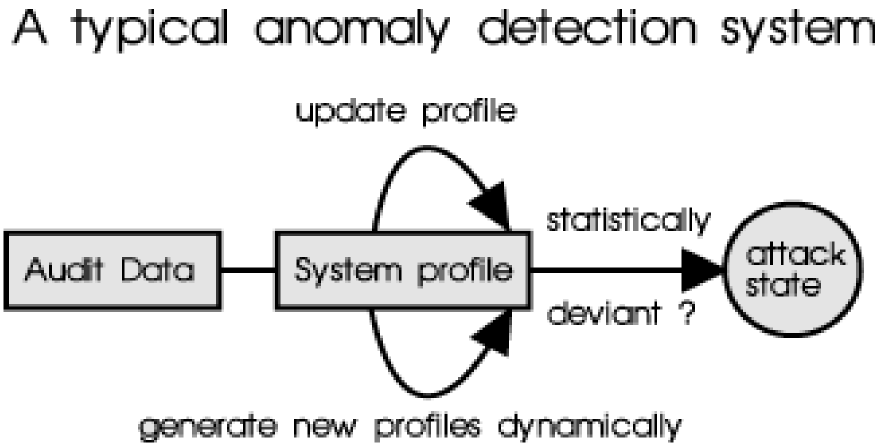


Figure 2. Anomaly Detection
source: www.cs.utk.edu/~abdulrah/ netsecurity/paper.html

There are challenges to the anomaly detection systems. As one can imagine, developing, maintaining, updating, and storing profile metrics can be quite tedious and the overhead can be enormous. Furthermore, in an area where on-the-fly decisions are valued, anomaly detection systems present a rather computationally expensive alternative. But perhaps the most significant challenge is controlling and limiting the false negatives. Activities flagged as false negatives are much more dangerous than those flagged as false positive. This is because a false negative is an intrusive activity that matches an activity in the normal profile and thus considered non-intrusive. Limiting this rate is essential to maintaining a functional anomaly detection system.

Trade-offs and Drawbacks

Whether one has deployed a misuse detection system or an anomaly detection system, effective and well-balanced policing is paramount to the success of the intrusion detection technique. One of the main trade-offs is that of convenience. An administrator must balance the security risks with the ease of accessing the system. The harder it is for a legitimate user to access the system, the harder it is for an attacker to access the system resulting in an increased security of the overall system. However, the administrator must realize that a legitimate user might "give up" trying to access the system and thus losing usability functions of the system. Another area that is of concern is the amount of overhead an administrator is willing to handle. Decisions in any intrusion detection technology must be made as quickly as possible. The overhead increases the computational efficiency and thus slowing the system's response/detection time. Another area of concern, which has been the focus of extensive research, is the sensitivity of the system compared to false alarms. This area of effective policing has been depicted in figures 3-5. As was previously discussed, an increased amount of false positives can render the system useless in that it flags an increased amount of legitimate activity. However, if one was to desensitize the system, it would be easier to focus one's resources in diagnosing and treating the given flagged activity as it would be most likely an intrusive action. On the other hand, intrusive activities have a higher likelihood of not being flagged as intrusive. It is a similar case with false negatives. If the system is sensitized, more activities will be flagged including legitimate activities resulting in an increased false positive rate. Finally, in an intrusion detection system that is distributed by means of software agents, there are trade-offs between scalability and robustness versus communication overhead. Also, an administrator has

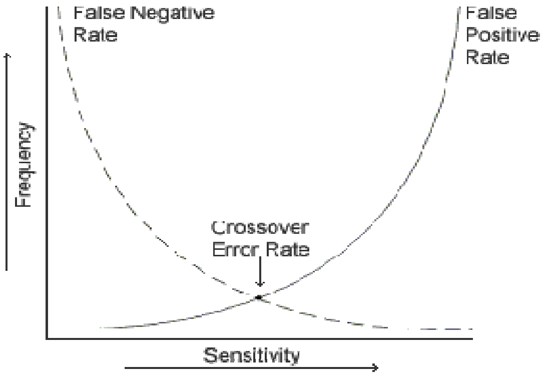
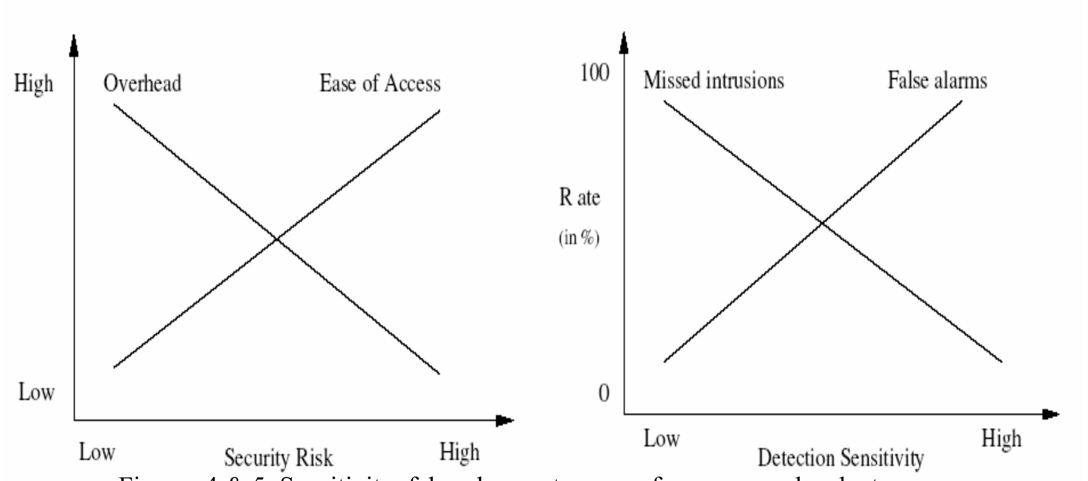


Figure 3: The fragile balance of tradeoffs



Figures 4 & 5: Sensitivity, false alarm rate, ease of access, overhead, etc.

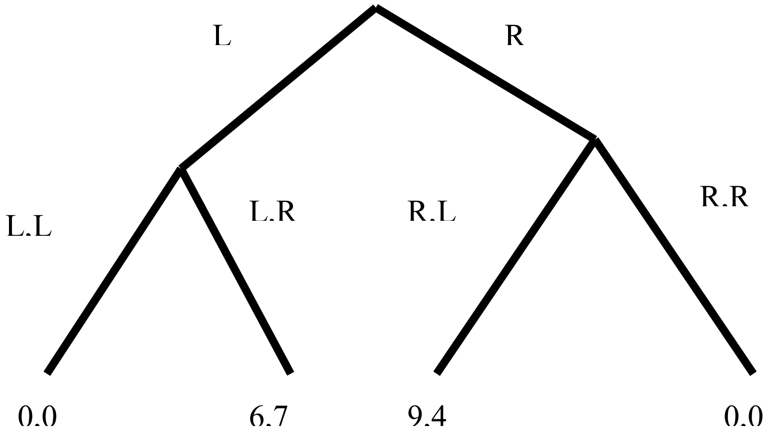
the added responsibility of allocating the level of autonomy for a given agent. No matter the technique utilized for detecting intrusions to a system, it is evident that numerous factors must be balanced in constructing a proper policy. An optimal policy is often never achieved. Thus there is need for improvements to existing techniques in not only detecting but also responding to activities deemed intrusive for a given system.

History of Game Theory

A branch of applied mathematics, game theory is similar to its well-known sibling decision theory. The study of game theory dates back to 1713 but was not a serious discipline until the publication of *Researches into the Mathematical Principles of the Theory of Wealth* in 1838. The field of game theory gained legitimacy in 1928when John von Neumann, known as the father and inventor of game theory, published a series of papers. The studies were, however, based upon economic behavior. In the 1050's experiments using game theory were being conducted to solve the famous prisoner's dilemma. John Nash later developed what is now termed the Nash equilibrium. This theory facilitated the analysis of non-cooperative and cooperative games. Subsequent prize winning research was conducted but mostly in the area of economic sciences. In the 1970s, game theory was applied heavily in the field of biology. Just recently has the idea of game theory been introduced into complex computing problems.

Game Theory

Game theory provides a way of mathematically formulizing the decision making process of policy establishment and execution. Thus, a logical methodology is applied to the selection of certain threshold values applied to the overall system policy. Game theory has been studied and applied in many fields and applications. However, there has been limited and minimal research in the area of networking and more specifically in the realm of intrusion detection systems. In fact, it was first developed as a tool for understanding economic behavior. Thereafter, it was utilized in defining nuclear strategies for the RAND Corporation. Today, some areas of interest have been game and adversarial situations. Also, there have been well-researched applications of game theory in the fields of economics, political strategy, biology, psychology, sociology, philosophy, and war strategy [11]. In game theory, players are pinned against each other taking turns sequentially to maximize their gain in an attempt to achieve their ultimate goal.



As is illustrated from the tree representation in extensive form in figure 6, party A may choose to move left (L), party B would choose a subsequent move to the right (LR) in order to achieve maximum gain

Figure 6: Extensive representation of a 2-person game

and victory over adversary A. In a pre-emptive strategy, party A might decide to make the first move an R. In this case, it will never lose but can win. There are several models of game theoretic approaches [10] available including but not limited to

- 2−Person vs. N−person
- Cooperative vs. Non-cooperative
- Zero-sum vs. Non-zero sum
- Finite vs. Infinite
- Min-max strategy

The different models of game theoretic approaches are categorized according to certain features present. For the first example in the list above, it is obvious that the feature of interest is the number of players involved in the game. One-person games are of little or no interest and often dismissed because of their simplicity and lack of adversarial conflict. An example of a one-person game would be a person deciding whether to buy an item. A person would simply list the risks and possible outcomes and make a decision. Before talking about 2-person games, it is important to introduce the notion of zero-sum and nonzero-sum games. Zero-sum games are games in which there is a balance of power. This means that if one party accumulates a gain, the other party subsequently loses some. Nonzero-sum games, on the other hand, can have multiple winners and losers. There need not be a clear shift of powers. An example of this would be a negotiation. Two adversaries can negotiate to a mutually acceptable solution and thus both be winners. Nonzero-sum games can be further categorized into cooperative or non-cooperative games. In a cooperative game, the two adversaries may cooperate in advance. While in non-cooperative games, the adversaries work independently to achieve maximum gain. Another subcategory of 2-person games is finite and infinite games. Finite games are like most scenarios in which there are a finite amount of decisions or outcomes possible. Infinite games are more complex and subtle in that there are an infinite number of alternatives each adversary can employ. The minmax theorem, proven by von Neumann in 1928, states that “every finite, two-person zero-sum game has a solution in mixed strategies” [10]. Furthermore, it states that there is a value, V, that will yield to either party’s favorable outcome and that both parties have the motivation and power to enforce that outcome V. N-person games in which the players do not cooperate are not much different from 2-person games. There is the added factor of players joining powers and forming alliances to achieve a common goal. This area of game theory is still being researched, as there is no concrete way of defining the best strategy for a given player.

Game Theoretic solutions in IDS

One can utilize the notion of game theory in current intrusion detection systems in assisting in defining and reconfiguring security policies given the severity of attacks dynamically. Furthermore, a game theoretic approach to intrusion detection systems assist in the decision process involved with allocation or reallocating limited resources for detecting significant threats to vital subsystems of a large networked system in near real time. Lastly, the aim of a game theoretic approach to intrusion detection systems should allow for the quantification of appropriate responses that would match threat levels.

Model

In an attempt to formalize a game theoretic approach to IDS, we have defined the following model. For any given IDS, there must be a set of sensors, S, within a network N such that

$$S = \{s1, s2, s3,\ldots,sn\}.$$

The intrusion detection system utilizing a game theoretic approach can be distributed and thus represented as a set of subsystems T wherein

$$T= \{t1, t2, t3,\ldots,tn\}.$$

There also exists a set of documented threats and detectable anomalies, I, such that it may indicate a potential intrusion wherein

$$I = \{I1, I2, I3,\ldots,In\}.$$

As with traditional intrusion detection systems, the given sensors within set S are able to detect more than one anomaly, threat, or possible intrusion. By utilizing a one-to-many mapping from the set S to the set $I \cup \{0\}$, an output vector of the network of sensors d is yielded wherein

$$d := \{d1, d2, d3,\ldots,dn\}.$$

Furthermore, there exists a matrix A such that it contains a description of the relationship between the sensor output vector and the subsystems wherein

$$A\{I, J\} = \begin{cases} 1; \text{ if the sensor } j \text{ monitors subsystem } i \\ 0; \text{ if sensor } j \text{ does not monitor subsystem } i \end{cases}$$

$$\text{where } i \subset I \text{ and } J \subset d(s).$$

There also exists a fixed number of security levels with predefined thresholds. Each known intrusion or anomaly from the set I is associated with a security risk value quantified with a positive real number, denoted f(I). The given intrusion detection system switches automatically between the different levels. This switching is done responsive to the sum of the risk values of the detect intrusions from set I. This method has several advantages in that the security warning system enables the intrusion detection system to operate in different modes at each security level. It also provides the administrator an intuitive overview of the current security situation in the network. So one can see that the security level thresholds, denoted m, are different than the IDS security level, L. In fact, m determines if L should be adjusted or not. The IDS security level L can be determined by

$$\mathcal{L} = \begin{cases} l_1, & \text{if } \sum_{i=1}^{N} f(d_i) < m_1 \\ l_j, & \text{if } m_{j-1} \leq \sum_{i=1}^{N} f(d_i) < m_j \\ l_L, & \text{if } \sum_{i=1}^{N} f(d_i) \geq m_L. \end{cases}$$

Available Approaches

There are less than a handful of prototypes of intrusion detection systems with a game theoretic approach that have been implemented. One of them is Bro. Bro is a stateful, event based analysis system. It is limited in detecting events such as connection establishment, http requests, etc. One of the foremost and well-known IDS systems, Snort, has a prototype termed Adaptive Snort. This is a packet based analysis method with a rule tree structure used for detection. It monitors performance of packets received, dropped, average inter-packet arrival time, rule hits, rule misses, and types of packets received. Figures 7-8 denote the traditional snort on the right and the adaptive snort on the left with respect to number of dropped packets.

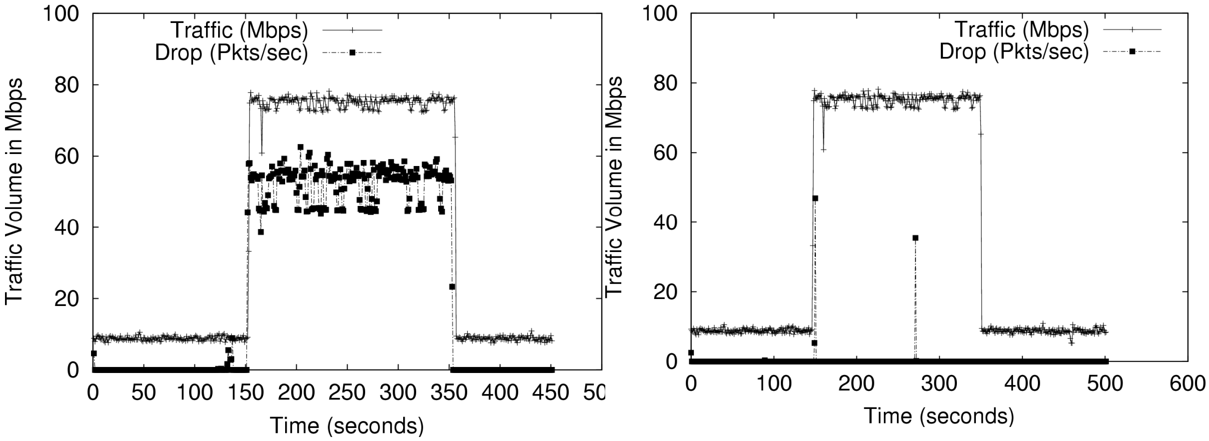


Figure 7: Traditional Snort vs Adaptive Snort
Source: www-static.cc.gatech.edu/classes/ AY2003/cs8803k_spring/selected_projects.ppt -

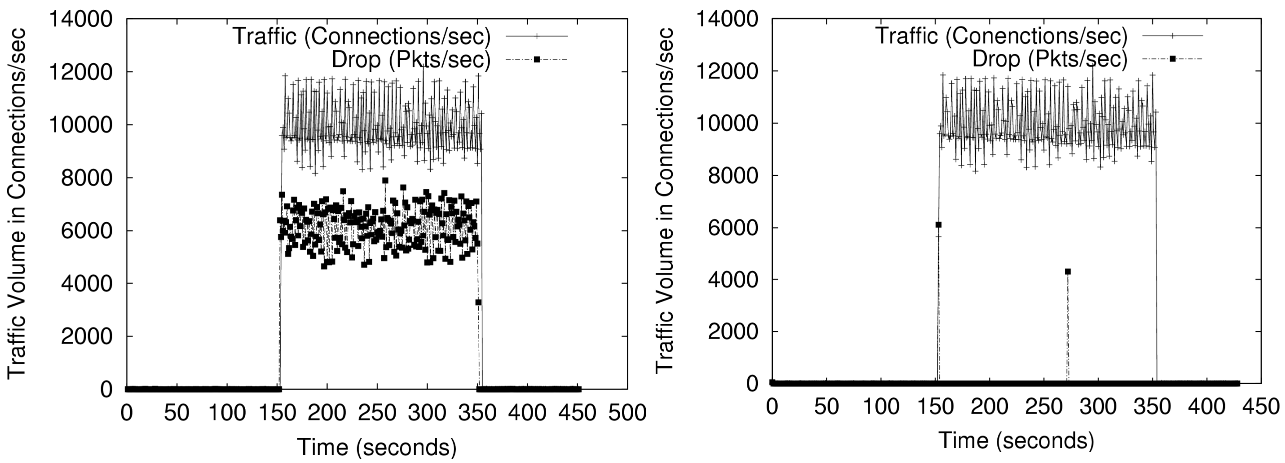


Figure 8: Traditional Snort vs Adaptive Snort
Source: www-static.cc.gatech.edu/classes/ AY2003/cs8803k_spring/selected_projects.ppt

As one can clearly see, there is a vast improvement with less dropped packets. No matter which prototype one is trying to incorporate, providing performance assurance must be the basic requirement for intrusion detection systems.

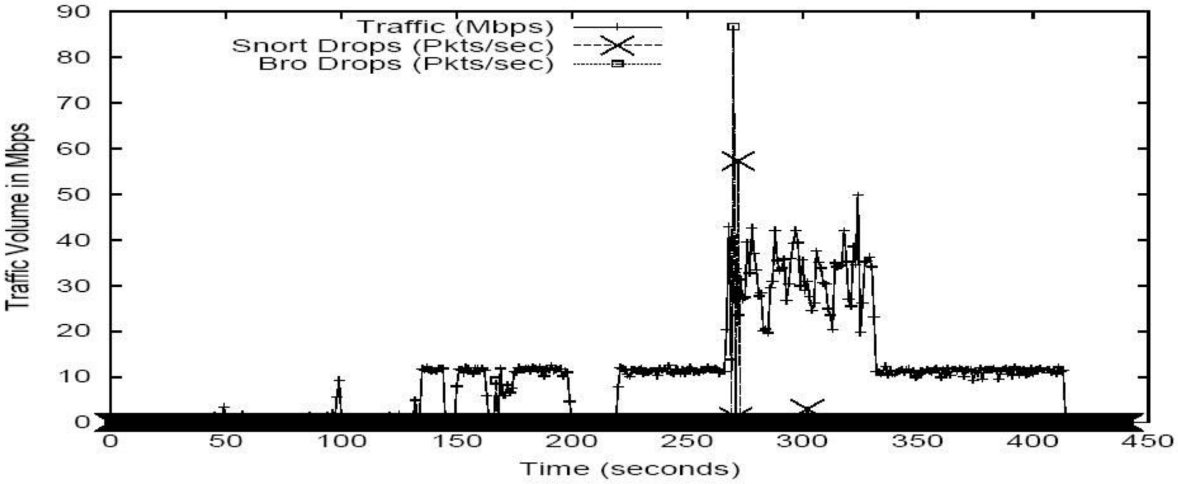


Figure 9: Comparison of Adaptive Snort and Bro
Source: www-static.cc.gatech.edu/classes/ AY2003/cs8803k_spring/selected_projects.ppt

Through our research, we provide an analysis of metrics and limitations of intrusion detection systems. Having the operational limitation established in the IDS, the real values to be adapted must not be deviated from. The original purpose was to ensure that an IDS must achieve the basic goal of performance adaptation given the different environment and operation of the system. This can only be accomplished with monitoring and reconfiguration mechanisms. Future work needs to focus on efforts to lower the bound and overhead by allowing them to be updated/edited dynamically. This essentially is an optimization problem..

Attack Types and Countermeasures

Once one has an understanding of the components of the intrusion detection system with game theory, one can begin to study the different types and components of the attacks. Each type of attack consists of three components: type of influence, specificity, and security violation. The type of influence can be either a causative where the attacker wants to affect the system's behavior or merely exploratory in nature. The specificity component describes whether the attack is targeted or indiscriminate. Finally, the security violation describes the type of security breach the attack would like to effect such as availability or integrity.

Combining these components together yields a type of attack. For example a CI attack or a Causative integrity attack is an attack in which an attacker attempts to fool an IDS into not flagging a known exploit as an intrusion. The specificity parameter would define whether one particular exploit is chosen or is it indiscriminate. This type of attack has an ultimate goal of forcing the administrator to disable the IDS by causing severe degradation of the engine's performance. The engine is the component within the IDS that calculates and stores the game theory strategy. This component is usually a learning engine and adapts its thinking dynamically, evolving over time. An EI attack, on the other hand, is an exploratory integrity attack that attempts to discover information about the state of the engine. It is important to note that the attacker does not attempt to influence the engine in any way. Rather the attacker simply attempts to find intrusions that are not recognized by the engine.

More complex attacks try to influence the learning of the engine. As discussed before, the engine is a learning component and thus must be regularized. It must go through a learning phase often described as a penalty term. The more constraints the engine is presented with during the penalty term, the more independent the engine will be leaving less of an opportunity for the adversary to exert influence over the engine. A causative attack, however, can be detected my creating a special test set containing intrusions and activities similar to intrusions. A subsequent analysis of the classification of the activities would give a strong indication of the compromises within the trained engine. Exploratory attacks can also be detected by running a separate clustering algorithm. A large cluster near the decision boundary indicates a systematic probing.

One can also confuse the attacker's estimate of the engine's state, thereby preventing the attacker from learning the decision boundary. Here the roles of the attacker and the engine are reversed. The engine can also set up honeypots or honeynets by indicating that some intrusions are not included in the training set. Thus, the attacker is baited in exploiting a vulnerability that the intrusion detector is closely watching.

Limitations of Game Theory

Although the introduction of game theory into the realm of network-based intrusion detection systems presents a vast improvement over existing methods, there is no perfect solution as of yet. A considerable amount of research still needs to be done in order to solve some limitations of our approach and game theory in general. For example, currently, not all types of intrusions can be modeled as the game theory tree must remain finite. This problem is further highlight by attacks that span a long time frame. These attacks may not be recognized or simply overlooked. The adversary may also try to train the learning engine, as has been discussed previously. Perhaps the most important flaw of game theory is that it assumes the adversary's rationality. It assumes that it would take steps to obtain a maximum gain (or near maximum) for itself or its interests. However, in reality, this may not always be the case as human behavior is still unpredictable. So it is hard to discern whether the user's attack is rational or whether he/she is simply trying to confuse the IDS by employing another attack vector. Another area of weakness of the game theoretic approach to intrusion detection is the unsolved approach on how to detect and handle simultaneous attacks. While most other limitations seem to be resolvable, this area will be the test of whether game theoretic solutions will have a future in the realm of intrusion detection systems.

Conclusion

Intrusion detection systems have proven their value to network and system security. Researchers and administrators alike are constantly searching for methods to improve existing technologies. An experimental approach of introducing game theory into the realm of intrusion detection has yielded better than expected results. However, there are still much to research and improve before this technology becomes truly deployable. In an attempt to further the research in this area, we have introduced a foundation for theoretical limits in applying a game theoretic approach to intrusion detection systems. There are limitations to our proposed method but with continued research, we are confident that they are viable solutions to these limitations. We further plan on expanding our research to include n-person games as well as incorporation of different technologies. We also hope to formulize a variety of attack vectors and construct a conducive test environment in which we can perform experiments.

References Cited

1. J.P Anderson. Computer Security Threat Monitoring and Surveillance. Technical report, James P Anderson Co., Fort Washington, Pennsylvania, April 1980.
2. An Introduction to Intrusion Detection by Sundaram, Aurobindo. "An Introduction to Intrusion Detection". ACM Crossroads, Issue 2.4, April 1996.
   http://www.acm.org/crossroads/xrds2-4/intrus.html
3. Eugene H Spafford. Security Seminar, Department of Computer Sciences, Purdue University, Jan 1996.
4. Steven E Smaha. Haystack: An Intrusion Detection System. In *Fourth Aerospace Computer Security Applications Conference*, pages 37-44, Tracor Applied Science Inc., Austin, Texas, December 1988.
5. www.cs.utk.edu/~abdulrah/ netsecurity/paper.html
6. A Game Theoretic Approach to Decision and Analysis in Network Intrusion Detection Tansu Alpcan and Tamer Basar - Decision and Control Laboratory, CSL,, University of Illinois at Urbana-Champaign
7. An Intrusion Detection Game with Limited Observations Tansu Alpcan and Tamer Basar
8. Can Machine Learning Be Secure? Barreno et al, Computer Science Division, UC Berkeley
9. Martin J. Osborne and Ariel Rubiunstein. A Course in Game Theory., MIT Press, 1994
10. Classification of Games, Encyclopedia Britanica, Accessed May 2006., http://wwwa.britannica.com/eb/article-22612?tocID=22612
11. Game Theory, http://en.wikipedia.org/wiki/Game_theory