CYBER THREAT ANALYSIS – A KEY ENABLING TECHNOLOGY FOR THE OBJECTIVE FORCE (A CASE STUDY IN NETWORK INTRUSION DETECTION)

Aleksandar Lazarevic, Jaideep Srivastava, Vipin Kumar
Army High Performance Computing Reearch Center, Computer Science Department, University of Minnesota
aleks@cs.umn.edu, srivasta@cs.umn.edu, kumar@cs.umn.edu

Introduction

Effective use of state-of-the-art communication has always been key to societal progress, be it smoke signals of yesteryear, or network-based computer systems of today. Nowadays computers control power, oil and gas delivery, communication systems, transportation networks, banking and financial services, and various other infrastructure services critical to the functioning of our society.

Since the Second World War, science and technology have been a key enabler of the US military's global leadership. Tremendous progress in information technology is the critical to the ongoing transformation and eventual fielding of the Objective Force, as spelled out by many of the service's leaders at the Association of the U.S. Army's 2002 Winter Symposium. According to the US Army White Paper "Concepts for the Objective Force" [1], soldiers and leaders enabled by advanced technologies will provide revolutionary increases in operational capability. In addition, information systems will provide dominant situational understanding, enabling combined arms units to conduct simultaneous, non-contiguous, distributed operations. Platform designs in an arrangement of system-of-systems technologies will enable decisive maneuver, both horizontal and vertical, during day and night, and in all terrain and weather conditions [1]. These breakthroughs will give Objective Force units the lethality and survivability needed to deliver full spectrum dominance, the versatility to change patterns of operation faster than the enemy can respond, and the agility to adjust to enemy changes of operation faster than he can exploit them [1].

Notwithstanding the tremendous benefits that information technology brings, there is inevitably an escalation of the "dark side of the force" in the form of cyber terrorism, which is the use of informational technology capabilities to launch an attack on an organization's information resources. Today our real assets are stored electronically, not in Fort Knox, and the targets are increasingly not only government and military installations, but public and private computer network systems as well. Information warfare extends the battlefield to incorporate all aspects of society. The belief exists that the United States has not been invaded since 1812, but invasion through the cyberspace is now a daily occurrence. We can no longer afford to rely on the two oceans that have historically protected our country; and instead must develop the means to mitigate risk in an electronic environment that knows no boundaries. To compound the problem, military and law enforcement authorities report that every month, assailants make thousands of unauthorized attempts to gain access to these systems, amounting to a nearly continuous assault [2].

Cyber Threat Analysis

According to a recent research survey, cyber attacks have increased by almost 80 percent over the last 6 months [3]. This indicates that there is an urgent need to expand efforts in the battle against cyber terrorism. The key question is whether contemporary information technologies such as artificial intelligence and data mining can contribute to this battle and further enhance defense mechanisms. This paper addresses some possible directions in this battle.

With an eye towards the future, the Army is undergoing a transformation from a forward deployed 'Cold War' army to a power projection force. This transition will eventually result in a fully digitized, more configurable, rapidly expandable, strategically deployable, and effectively employable organization. It is clearly evident that the advanced information technologies will play an important role in this transition. Cyber Threat Analysis, as one of the most emerging advanced technologies, has many different components including information assurance, methods to identify the most critical infrastructures, methods to detect cyber terrorist attacks and protect against cyber terrorism, intrusion detection and recovery from intrusions. All these components inevitably cause the changes to army doctrines, tactics, techniques, and procedures on how we integrate digitized and non- digitized systems and organizations into the fight.

Intrusion Detection

This paper presents the scope and status of our efforts in detecting cyber attacks on the real network at University of Minnesota. Intrusion Detection includes detecting and defending against malicious actions that compromise the integrity, confidentiality, or availability of information resources. The most widely deployed methods for intrusion detection employ signature-based detection techniques. These methods extract features from various audit streams, and detect intrusions by comparing the feature values to a set of attack signatures provided by human experts. Such methods can only detect previously known intrusions that have a corresponding signature. The signature database has to be manually revised for each new type of attack that is discovered. Limitations of signature-based methods have led to an increasing interest in intrusion detection techniques base upon data mining.

Data mining based intrusion detection techniques generally fall into one of two categories; namely misuse detection and anomaly detection. In misuse detection approaches, each instance in a data set is labeled as normal or intrusion (attack) and a learning algorithm is trained over the labeled data. These approaches are able to automatically retrain intrusion detection models on different input data that include new types of attacks as long as they have been labeled appropriately. The main advantage of misuse detection is that it can accurately detect known attacks, while its drawback is its inability to detect novel, previously unseen attacks.

Traditional anomaly detection approaches, on the other hand, build models of normal data and detect deviations from the normal model in observed data. Anomaly detection

applied to intrusion detection and computer security has been an active area of research since it was originally proposed by Denning [4]. Anomaly detection algorithms have the advantage that they can detect new types of intrusions as deviations from normal usage. However, they suffer from a high rate of false alarms primarily because previously unseen (yet legitimate) system behaviors may also be recognized as anomalies, and hence flagged as potential intrusions.

This paper summarizes our current research in the area of network intrusion detection. It addresses both misuse detection and anomaly detection. In misuse detection related problems, standard data mining techniques are not applicable due to several specific details that include dealing with skewed class distribution (intrusion as a class of interest is much smaller i.e. rarer than the class representing normal network behavior) and learning from data streams (intrusions very often represent sequence of events). We have developed several novel classification algorithms designed specially for learning from the rare classes. PN rule [5] is a two-stage learning algorithm based on computing the rules. The first stage is aimed at discovering P-rules that cover most of the intrusive examples, while the second phase discovers N-rules and eliminates false alarms generated in the first phase. Another technique includes generating artificial examples from the minority (intrusion) class, then learning classifiers from different newly generated data and finally combining these classifiers [6]. Experimental results on publicly available intrusion detection data sets [7] for both approaches have demonstrated that the proposed techniques are much more efficient in detection intrusive behavior that the standard data mining techniques.

For the case where the nature of the attack is unknown, we have developed outlier detection schemes to detect novel attacks/anomalies. The implemented outlier detection techniques are based on computing the distances between pairs of points and densities of specific regions [8]. In distance-based approaches, outliers (anomalies) are points that do not have enough neighbors, while in density-based approaches outliers (anomalies) are the points that belong to the highly sparse regions of the data space. Several existing outlier detection schemes and their variations are evaluated both on the publicly available data set of network connections [7] as well as on the data collected at our center. Our experimental results indicate that some outlier detection schemes appear more promising than others when detecting novel intrusions.

Finally, encouraged with excellent initial results we are currently developing an on-line anomaly detection system for network connections that is based on publicly available intrusion detection system (IDS) called SNORT [9]. SNORT is an open source signature-based IDS that allows us to integrate our outlier detection schemes. In order to incorporate these schemes, the SNORT database is modified to include extra features that are derived using the network statistics and later used in the outlier detection algorithms. In addition, we have also implemented an visualization tool for reporting anomalous/suspicious behavior detected using our schemes. Although SNORT already has its own anomaly detection engine SPADE, it achieves poor performance due to very high false alarm rate. In contrast, preliminary results have shown that our integrated outlier detection schemes improve the detection rate achieved by SNORT while keeping the false alarm rate quite low.

Conclusion

Leading edge information technology is a critical underpinning for the Objective Force, which is the main thrust of the Army of the future. Unfortunately, the very same information technology advances that provide new capabilities to our forces, also provide the enemy with new and powerful tools to launch attacks on our critical information resources. A specific example of this trend is the rapidly increasing rate of cyber attacks against our computers in the past few years. It is crucial that we pay sufficient attention to making our information systems – especially those used for critical functions in the military and commercial sectors – resistant to and tolerant of such attacks. Research at the Army High Performance Computing Research Center (AHPCRC) is focusing on applying data mining to develop techniques that can be used to detect, thwart, and recover from known as well as unknown cyber threats.

Acknowledgments. This work was partially supported by Army High Performance Computing Research Center contract number DAAD19-01-2-0014. The content of the work does not necessarily reflect the position or policy of the government and no official endorsement should be inferred. Access to computing facilities was provided by AHPCRC and the Minnesota Supercomputing Institute.

References

[1] United States White Paper: Concepts for the Objective Force, 2001.

[2] Frank Vizard, Waging War.com: A Hacker Attack Against NATO Spawns a War in Cyberspace, Popular Science, p. 80, July 1999.

[3] Successful Real-Time Security Monitoring, Riptech Inc. white paper, September 2001.

[4] D.E. Denning, An Intrusion Detection Model, IEEE Transactions on Software Engineering, SE-13:222-232, 1987.

[5] R. Agarwal, M. Joshi, Pnrule: A New Framework for Learning Classifier Models in Data Mining (A Case-study in Network Intrusion Detection), Proceedings of First SIAM International Conference on Data Mining, April 2001.

[6] A. Lazarevic, N. Chawla, L. Hall, K. Bowyer, SMOTEBoost: Improving the Prediction of Minority Class in Boosting, in review.

[7] R. P. Lippmann, D. J. Fried, I. Graf, J. W. Haines, K. P. Kendall, D. McClung, D. Weber, S. E. Webster, D. Wyschogrod, R. K. Cunningham, and M. A. Zissman, Evaluating Intrusion Detection Systems: The 1998 DARPA Off-line Intrusion Detection Evaluation, Proceedings DARPA Information Survivability Conference and Exposition (DISCEX) 2000, Vol 2, pp. 12-26, IEEE Computer Society Press, Los Alamitos, CA, 2000

[8] A. Lazarevic, H. Ramnani, L. Ertoz, J. Srivastava, V. Kumar: Evaluation of Outlier Detection Schemes for Detecting Network Intrusions, in review.

[9] SNORT Intrusion Detection System, www.snort.org