International Journal of Science and Technology (STECH)
Bahir Dar- Ethiopia

Vol. 3(3), S/No 8, September, 2014: 106-112
ISSN: 2225-8590 (Print) ISSN 2227-5452 (Online)
DOI: http://dx.doi.org/10.4314/stech.v3i3.10

Data Mining for Cyber Security

Ikemelu, Chinelo Rose_Keziah
Computer Science Department
Nwafor Orizu college of Education, Nsugbe
Anambra State, Nigeria
E-mail: Ikemeluchinelo@yahoo.com

Abstract

Cyber security is concerned with protecting computer and network system from corruption due to malicious software including Trojan horses and virus. Security of our network system is becoming imperative as massive sensitive information is transmitted across the network. In this research paper, data mining application for cyber security is highly explored. We discussed various cyber-terrorism or attack committed across the network such as malicious intrusion, credit card fraud, identity thefts, and infrastructure attack. Data mining techniques such as classification, anomaly, link analysis and so on are being applied to detect or prevent the aforementioned cyber-terrorism or attack. Recommendations were made and suggestion for further study was indicated.

Introduction

Cyber security consists of security mechanism that attempt to provide solutions to cyber attacks or cyber-terrorism. Security of network systems is becoming increasing important as more and more sensitive information is being stored and manipulated online. As a result, the integrity of computer networks, both in relation to security and with regard to institutional life of the nation in general is a growing concern. Security intelligence reports that cybercrime wave around the world is worrisome. According to the U.S Federal Trade commission (FTC)credit card fraud cost card holders and issuers hundreds of millions of dollars each year which create a tremendous damage to the economy. Also the biennial department of trade and industry (DTI) security breaches survey reports stipulated that 68% of UK business had a computer security crime incident in 2012. Furthermore, security and defense networks, proprietary research, intellectual property, and data based market mechanisms that depend on unimpeded and undistorted access, can all be severally compromised by malicious intrusions. There is a need for the best way to protect these systems. That is an effective technique is needed to detect security breaches.

Data mining has many applications in security includes in national security (e.g. surveillance) botnet detection, as well as in cyber security (like virus detecting). Data mining or knowledge discovery (KDD) according to Bradfond (2012) is a method used to analyze data from a target source and compose that feedback into useful information. Data mining techniques are being used to identify suspicious individuals and groups, and to discover which individuals and groups are capable of carrying out terrorist activities. Data mining is also being applied to proffer solutions such as intrusion detection and auditing.

In this study, we will focus mainly on Data mining application for cyber security. To comprehend the mechanism to be adopted in order to safeguard the nation’s computers and network, it is imperative to understand the types of threats that endanger the cyber network. Against this backdrop, this paper discusses cyber terrorism, threats and external attacks, malicious intrusions as well as credit card identify theft. Attacks on critical infrastructures were also discussed. Finally, data mining application for cyber security was discussed and suggestions for further study was made

Cyber Terrorism, Threats and External Attacks

Cyber-terrorism, according to the O’ Leary (2010) is committed through the use of cyberspace or computer resources. Cyber-terrorism is one of the major terrorist threats posed to our nation today. This threat escalated as a result of vast information transmitted electronically across the web. Attacks on our computers, networks, internet intra-structure and databases could be devastating to the business. It is estimated that cyber – terrorism could cause billions of dollars to business. A classic example is that of a banking information system. According the Federal Trades Commission (FTC), for all internets related complaints received 20% of cases involved a bank account debit and misuse of account numbers. If terrorist attack such a system and deplete accounts of funds, then the bank could lose millions or billions of dollars. Crippling the computer system millions of hours of productivity could be a lost, which is ultimately equivalent to money loss. Even a simple power failure at work through some accident could cause several hours of productivity loss which leads to financial loss. Therefore, it is imperative that our information system could be secured.

Threats can occur from inside or outside the organization. Outside attacks are attacks on computer system from someone outside the organization. Crackers can break into the computer system and cause a lot if havoc within the organization. Some crackers spread virus that damage files in various computer systems. But the more devastating one is that of the inside threat. These people are often under locked, but they are the people who understood the vital information in the organization. People inside an organization who have studied the business practices and procedures have an enormous advantage when developing scheme to cripple the organization's information assets. These people may be regular employees or even those working at the computer centres. The problem is very devastating as someone masquerades as someone else, and causes all kinds of damage such as stealing important technical information or introducing what is called a "bomb" that is a destructive computer programme into the system. The rest of the paper will be examining how data mining could be used to prevent such attacks.

Malicious Software and Malicious Intrusion

Malicious software is programs specifically designed to damage or disrupt a computer system (O' Leary 2010). The most common types of these programs are viruses, worms and Trojan horses and are normally propagated to the computer network through Hackers and crackers. The target of malicious intrusions includes networks, web clients and servers, operating system and database. Many cyber-terrorism are due to malicious intrusions. We hear much about of network intrusions. What happens here is that the intruders try to tap into the networks and get vital information that is being transmitted. These intruders may be human intruders or automated malicious software set by humans. Intrusion can also target files inside of network communications. For instance, an attacker can masquerade as a legitimate user and use their credentials to log in and access restricted files. Malicious intrusion can also occur on databases, in some cases, stolen credentials enable the attacker to pose queries such as SQL queries and access restricted data.

In discussing malicious intrusion or cyber attacks it is often helpful to draw analogies from the non cyber world (that is non-information related terrorism) and then translate those attacks to attacks on computers and networks. For instance a thief could enter a building through a door. In the same way, a computer intruder could enter the computer or network through some sort of a trap door that has intentionally build by a malicious insider and left unattended perhaps through careless design.

In the case of credit card fraud, which is a more serious problem too, here an attacker obtains a person's credits card and uses it to make unauthorized purchases. By the time the owner of the card becomes aware of the fraud, it may be too late to reverse the damage or apprehend the culprit. A similar problem occurs when using ATM card on Automated Teller Machine (ATM) in withdrawing money from the bank. I have become a victim of such scenario. In fact, these type of fraud has happened to me personally where someone stole my ATM card and made away with substantial amount of money (70,000) seventy thousand naira from my account and by the time I got the information from my phone it was too late. When such information alerted to my bank that, i am not the person that withdrawal the money, they could not do anything to recover the money and up till today nothing has been done. Such fraudulent can cause a disaster to the health of the victim. There is a need to explore the use of data mining for both credit card and detection and as well as identity theft.

Critical Infrastructures Attack

Attack on critical infrastructure is very dangerous because this could cripple a nation and its economy. Infrastructure attacks consist of attacking telecommunication lines, power, electricity, gas, reservoirs and water supplies, food supplies and other basic amenities that are paramount for the operation of the nation.

Critical infrastructure attacks could occur during any type of attack whether they are information related, non- information related or bio terrorism attack. For instance one could attack the software that runs the telecommunications industry and close down all the telecommunication lines. Telecommunication lines could also attack physically through bombs and explosive chemicals. Similarly, software that runs the power and gas supplies could be attacked. Attack on infrastructure could also occur in transportation line such as attacking railways and high ways. All these could cause a lot of damages if it is not protected.

Furthermore, infrastructures could also be attacked by natural disaster such as hurricanes and earthquakes. But our main interest here is the attacks on infrastructures through malicious attacks, both information and non-information related. The main goal of the researchers is to apply data mining to prevent or detect such infrastructure attacks.

Data Mining Application for Cyber Security

Data mining According to Silltow (2012) automates the detection of relevant patterns in a database, using defined approaches and algorithms to look into current and historical data that can then be analyzed to predict future trends. Because data mining tools predict future trends and behaviors by reading through database for hidden patterns, they allow organizations to make proactive, knowledge driven devious and answer questions that were previously too time-consuming to resolve.

Data mining application for cyber security is the use of data mining techniques to detect cyber security. Data mining is being applied to problems areas such as intrusion detection and auditing. For instance, in anomaly detection techniques, it could be used to detect usual patterns and behaviors. According to Thuraisingham, Khan and masud (2013) data mining classification technique may be used to group various cyber attacks when it occurs. Link analysis may be used to trace self-propagating malicious code to its authors. Prediction data mining technique may be use to determine potential future attacks depending in a way on information learnt about terrorists through our e phone conversations and email. Moreover, in the application of data mining for cyber security, non-real data mining may be elicit or apply to some threats while for certain other threat real-time data mining may be used such as in Network intrusion where real time data mining may be applied to detect fraud. According to Thuraisingham, and etal (2013) they stipulated again that real time data mining may be used for credit card fraud detection as it is inform of real time processing, that it is imperative that the result and the models build should be generated in real time. However, the models are usually built ahead of time. Although building model in real-time may be challenging. Data mining can also be used for analyzing web logs, here base on the results of data mining tool one can determine whether any unauthorized intrusions have occurred.

Data mining application for cyber security other areas include in the analysis of audit data. Here one could build a warehouse or repository containing the audit data and then conduct an analysis using various data mining tools to find out if there is a potential anomaly. For example, there could be a situation when a person has been accessing the data base between the hours of 3-4am but for the last 2 days he has accessed database between the hours of 3-4pm. This could then be tagged as an unusual pattern that could be need proper investigation to know whether if an unauthorized query has been posed on the database. Furthermore, insider threat analysis is also a problem from a cyber security. Insider threat are those employee working in a corporation who are considered to be trusted could commit espionage. Also those who have proper access to the computer system could plant virus and Trojan horses and to gets such terrorist will be very difficult than catching terrorists outside the organization. As a result of this it is imperative that one should monitor the access patterns of all the individuals of a corporation even if they are system administrators to see whether they are carrying out cyber terrorism activities. Data mining technique could be used to introduce sensitive associations from the legitimate responses.

Conclusion

This paper had explored the various cyber terrorism or attacks such as malicious intrusion, credit card fraud, identity theft, and attack on critical infrastructure committed on the network that has deterred the integrity and effectiveness of computer network. These cyber-crimes should be detected or prevented using data mining techniques.

Recommendations and Suggestions for Further Study

Virtually every sector is being computerized on daily basis. There is a tendency for escalation of sensitive, information, transmitted across the network. Therefore, the security of our network and computer systems must be assured. Data mining techniques should be explored as the resent mechanism to combat cyber-terrorism or attack. As a result, computer or information technology professional and non-professional should be effectively trained on the application of data mining techniques base model as decision system to successfully ensure the security of our network which would enhance the profit potentials of the entire world economy.

The researcher suggests further research on the application of data mining techniques for botnet detection for an interested scholar who came across this work. The term "bot" comes from a word robot. A bot is typically autonomous software capable of performing certain functions. A botnet is a network of bots that are used by a human operator or bot master to carryout malicious actions. Botnets are one of the most powerful tools used in cyber-crime today as it is capable of affecting Dos (Denial – of – service attack), spamming, phishing and eavesdropping on remote computers. However, with the help of bonet, individuals, business and Government are facing a lot of millions of dollars damages. It is paramount that the cyber-security community or any interested research scholar should proffer a data mining technique to combat this challenge threat.

References

Contel, Bradford (2012). Different types of data mining technique. Follow e wise.com. Retrieved 11th Sept, 2012.

Daniel, B. & Sushil, S. (2012). Modern intrusion dictation, Data mining and degrees of attack guilt. Retrieved 28th January, 2014.

O'Leary, T. (2010). *Computer essential, introductory edition*. Arizona: Italy Retrieved, 14th July, 2005.

Silltow, J. (2012). Pattern recognition in data mining. Maryland, U.S.A: University of Maryland College Park.

Thuraisingham, B., Khan, L., & Masud, M. (2013). Data mining for security Applications. Retrieved 27th January, 2014.

U.S. Federal Trade Commission (F.T.C) (2012). Survey released. Retrieved March 13, 2012.