Int. J. Advanced Networking and Applications
Volume: 04, Issue: 04, Pages:54-60, Jan – Feb (2013)
Special Issue on "Computational Intelligence- A Research Perspective" held on "21st – 22nd February 2013"

Application of Machine learning in Detecting Insider Threat-State of Art and Survey

R.Rajamenakshi
Research Scholar,
Department of Computer Science,
Avinashilingam Institute for Home Science and Higher
Education for Women, Coimbatore
Email: rajamenakshi@gmail.com
Dr.G.Padmavathi
Professor & Head
Department of Computer Science,
Avinashilingam Institute for Home Science and Higher
Education for Women, Coimbatore
Email: ganapathi.padmavathi@gmail.com

ABSTRACT

Research in computer security is more focused to prevent unauthorized and illegitimate access to systems and information. But, many times, the most damaging malicious activity is the result of internal misuse within an organization, which has not drawn much attention. Data Exfiltration refers to illegitimate transfer of data out of a given organization or network. Organizations employ security solutions like IDS, IPS and firewalls at the perimeter level to safe guard their network from external attacks. Insider attacks in the recent decade poses serious impact on the organization in terms of confidentiality and reputation. Machine learning algorithms and techniques has provided solutions to many of the complex real time problems in diversified fields and of great help in decision making and taking preventive and corrective measures. Many of soft computing techniques have been applied in intrusion detection in the recent years to detect and to prevent network intrusions by both external and internal attackers. This paper presents overview of insider threat, its current state of art in research, research challenges, data exfiltration steps, detailed on the machine learning approaches applied to address this problem.

Keywords – **Insider threat, Data exfiltration, detection models, machine learning approaches, challenges**

1 Introduction

In this digital era, there is a persistent threat to confidential data that leads to huge data and financial loss. Organizations protect their information from external attackers by deploying perimeter security solutions like IDS, IPS, UTMs etc. to safe guard their network attacks. Insider cyber attacks of the recent times have gained attention where the data that is internal to the organization is compromised and ex-filtered by the insiders. Insider threat affects relatively many critical infrastructure sectors ranging from defense, finance, insurance, credit-card fraud, hospital records, power sector etc that results in the leakage of sensitive information.

[15] discusses the problem of insider threats in handling records in the health care systems. [1]has briefed about the study conducted by CERT/CMU since 2001 and their analysis projected that 52% of organizations are concerned about the magnitude of the damage caused by insider attacks. [9] presents summary of the insider attack cases and [10] gives the statistics of the demographics of the insiders who performed the crime. The studies indicate that 59% of the insiders were ex-employees who were fired, 77% were full time employees, 86% were technical positions, 96% of were male and 18% had history of violations. With the researches in the recent times different approaches have evolved over the time. These techniques have not gained popularity and are relatively immature. Lot of research is happening to detect, prevent and mitigate insider attacks. Different approaches have evolved over time, but these techniques are relatively immature and have not gained popularity.

1.1 Insiders and Insider Attack

[3] defines insider attack as intentional misuse of computer systems by users who are authorized to access those systems and networks. Insiders are those who perform those attacks who are eventually the employees, ex-employees, contractors and business partners who are associated with the organization either directly or indirectly and have privileges to access the organizations data. Bishop [16] defines insiders as entities who misuse organizations data either by violating their legitimate access or by obtaining unauthorized access for malicious activities. [23] discusses different types of insiders as inadvertent, intentional and malicious.

Insider attacks are any malicious attack on the organizations system or network leading to leakage of sensitive information, modification or deletion of records by an insider. These attacks include efforts to retrieve, tamper or change the information. With outsourcing in businesses, it becomes difficult to draw a strict line

Department. of Computer Science,
Avinashilingam Institute for Home Science and Higher Education for Women University,
Coimbatore

between insiders and outsiders. Since these attacks are carried out during business hours it is difficult to differentiate between malicious intent and non-malicious intent of the insiders performing these attacks.

1.2 Data Exfiltration

Data exfiltration on the other hand is defined as an unauthorized transfer of data over a network. [26] defines exfiltration as a mechanism by which a malicious insider moves the data out of the network from a compromised system within a confidential network. With a good knowledge about the internals of the organization and privileges to access data it becomes easy for the insider to exfiltrate data using different means. It is very difficult to differentiate between a normal activity and a malicious activity by an insider. The present day security solutions are not equipped to capture and classify the data movement within an organization.

Data exfiltration can happen over physical or networked means. Exfiltration over physical means include using external devices such as printers, copiers, FAX, USB devices, scanners and laptops. Exfiltration by physical means can be controlled by disabling or removing access to USB and providing controlled access to copiers and scanners. There are a good number of data leakage prevention (DLP) solutions that can be deployed at the network, end point and file level that would detect potential data breaches and prevent the same. Unfortunately, data can easily bypass DLPs (1) either by encoding or encrypting which is very easy to carry out using SSL or by using file based encryptions like WinZip, RAR etc or (2) by leveraging VoIP for exfiltration. It is difficult to mitigate data exfiltration unless we have a good knowledge on the nature of data movement and accesses in the given network. Windows Network share, Anonymous FTP, malware, encrypted backdoor, physical access and SQL injection are some of the exfiltration methods.

This paper is organized as follows. Section 2 presents the current state of art in insider threat and data exfiltration and discusses research gaps and challenges. Section3 presents system dynamics of the Insider threat. Section 4 presents the Existing detection models. Section 5 presents Conclusion and future work.

2 Current State of Art & Survey

2.1 Insider threat: Problem Perspective and Taxonomy

Anderson [14] discusses this problem as a complex problem as a factor of behavioral, technical and organizational issues. Most of the organizations take care of physical and well defined authentication mechanisms to protect their data from internal and external attackers. However, the insiders can always override all these policies to perform their day to day activities. Hence this problem is approached in a combined fashion by addressing organizational and behavioral issues, in addition to technology perspective

[20] A.H.Phyo has provided a detailed taxonomy of insider misuse. Cheswick Bellovin classifies insider attacks into seven generic categories based on their work on firewalls, whereas Neumann-Parker classifies the attacks into nine heads based on the insider attacks. Lindqvist–Jonsson classifies Neumann's classification further based on the System perspective. [13][20]Anderson classified the malicious insiders into three groups viz, masqueraders, misfeasors and clandestine. Tuglular provides taxonomy based on the insider incidents that had three dimensions viz incident, response and consequences. Magklaras-Furnell provided a human centric based prediction model based on system role, misuse factors and system consequences.

According to Gafny[10], the most commonly reported attacks are unintentional exposure of sensitive data, theft of IP and theft of other data such as financial information, customer records etc. Yoohwan [16] includes manipulation of data, using unauthorized security mechanisms and network connections, covert channels, physical damage and destruction including information extrusion/ex-filtration as a major threat arising out of insiders. The common insider attacks are IP theft, IT sabotage and Insider frauds resulting in data leakage using unauthorized access and resource misuse for a variety of reasons. The characteristics of most of these attacks are multi-step, multi-stage and stealthy in nature with a well planned contingency. The problem is more complex as it is very difficult to identify the motive and the technique behind the attack. [1] it is very complicated to distinguish the normal activity from the malicious activity of a user.

2.2 Data Exfiltration- State of Art

[28] [29] presents a detailed study on data exfiltration and [28] also presents taxonomy of the exfiltration based on the medium used for transferring of files viz physical, network and cognitive. These attacks can further be malicious or non-malicious in nature.

Further [29] details most of the data exfiltration techniques uses network protocols such s HTTP, FTP and SMTP and proposes a method to detect data exfiltration happening in databases. Data exfiltration can happen over encrypted or unencrypted covert channels using VoIP. Penetration testing done from multiple vectors such as external and internal network, wireless and application is used to detect exfiltration that has happened. It is also evident that the bandwidth utilization is proportional to the degree of covertness. Whenever there is a huge bandwidth is utilized it is easier to detect the exfiltration if any that happens. But if a low bandwidth is utilized it is very difficult to detect such exfiltration happening. Here most of the covert channels utilize only a smaller proportion of the bandwidth

2.2.1 Steps in Data Exfiltration

Data exfiltration does not happen in a single day and it's a slow and stealthy process that is done in a phased manner. [30][19]define these attacks as multi-step, multi-stage. [34] elaborates on the different attacks on multistage/steps that lead to cyber exploitation and exfiltration of data and identity theft. With the first exfiltration attack reported in

2004, there have been reports of this kind since then. Tracing back across different boundaries through different networks involving several stages across multiple hosts, multi-stage attacks possesses a serious challenge to the network researchers. There exist a number of steps before the actual exfiltration takes place. People have presented different steps and stages that are involved before the actual exfiltration. Sean Coyne[31] defines Data preparation, Staging, Data exfiltration, Persistence, Continue Data Exfiltration as the sequence of steps involved in data exfiltration. Amit [32] presents Infiltration, Data targeting and acquisition, command and control and exfiltration as the major steps in exfiltration. Here the infiltration happens both by human and technical factors. In all it is evident that data that is targeted is acquired and moved to a location, from where the actual extrusion takes place. Table 1 lists the different methods used at different stages leading to exfiltration.

TABLE 1- Methods Used in Exfiltration at different Stages

Initial Entry	Remote Access Application,3rd party Connections, SQL Injection, Exposed Services, Remote File inclusion, Email Trojan, Physical Access
Staging Area	Common Loading points, Workstations
Data Harvesting	Malware, Memory Parsers, Keystroke Loggers and Network Sniffers
Exfiltration	Using DNS/HTTP, Covert channels and VoIP

2.3 Research Gaps and challenges

Sandip et.al [5] presents a brief account of the data exfiltration detection challenges. Based on the survey, it is evident that, there exist a number of models and techniques to detect, prevent insider attacks from happening. There exist a number of DLP solutions that attempts to prevent data exfiltration happening. But there exist no comprehensive solution and gap still exists. We need a comprehensive solution that addresses these gaps.

(i) Current day attacks are performed either directly by humans or by using different tools like Malware, backdoor Trojans etc. It is difficult to differentiate between the actions and events initiated by either. Hence we need mechanisms to differentiate between program-initiated and human initiated events

(ii) It is very difficult to detect exfiltration happening over an encrypted channel. While the encryption of data is done mainly to ensure secure transfer and maintaining CIA, it is difficult to decrypt the encrypted one. Even the DLP solutions and deep inspection of the packet at the network level fails. We need mechanisms to detection and prevention of exfiltration happening over an encrypted channels

3 System Dynamics of Insider threat

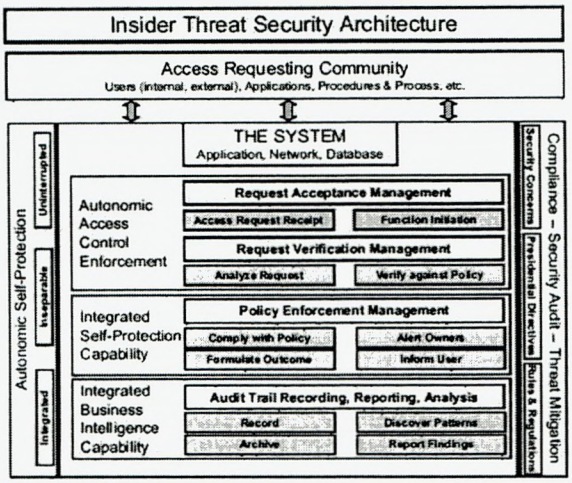

Figure 1- Insider Threat Security (ITSA) Framework

There exist different approaches in defining the system dynamics of the insider threat. [17] describes a four state model that describes the process by which an insider becomes an attacker. The different state transits are given as Disaffected → Contacted →Planning and Rehearsal → Attack. [8] suggests systems dynamics technique for detecting and mitigating insider threat. Neumann suggested that separation of organizational duties, two person control, split-key encryption and enlightened management would reduce the number of insider attacks happening. [14] presents system dynamics model of the insider cyber attack that is based on the dynamic trigger hypothesis. Later [2] enhances this model by adding judgments, decisions, outcomes and learning process. Martinez says that signal detection theory will well suit insider problem for selection and detection of insiders.[15] gives a framework that gives a score based on the user behaviors that uses psychology, anomaly detection theory and item response theory. This framework computes privacy risk scores using the item response theory and controls the information accesses with the help of these scores. Users who reveal sensitive information are supposed to have high privacy risk score and they are sent for awareness training for privacy protection. The system uses a DLP system in place that helps in setting this score.

4 Overview of Detection Techniques

The detection or prevention models starts with the collection of data, applying different techniques and detection, detecting preventing or predicting based on the outcome. Data abstracted at different levels – host, application and the network are used in different ways. Over the time, many approaches have evolved based on

(a) anomaly and signature based approaches on the network , (b) user behavior models based on user profiling on various user specific indicators and (c) Graph based approaches. Anomaly based approaches uses a number of machine learning algorithms for training and for detection. On the other hand, signature based approaches devise signatures that capture the specific kind of attack.. Stream mining is relatively new area that analyses the stream data for anomalies by applying machine learning techniques. There exists hybrid approaches such as SIDD multilevel framework as proposed by Liu [24] consisting of (1) network level application identification,(2) Signature generation and detection and (3) covert channel detection. The framework uses network flow data to analyze and generates signatures for the content and compares the signatures for detecting data exfiltration. Lance Spitzner[25] showed a detection mechanism that proved that honeypots can not only be used to detect external threats but can also be used for detecting insider threat. Using honeypots we will be able to collect the data, intercept the attack and analyze attacker's behavioral patterns.

Myers [1] views this problem as a subset of IDS, in a way that, data can be collected and analyzed from many different sources to identify potentially harmful events or event sequences. Myers mainly focuses on the threat sources from where the threats are generated from within a logical boundary. Number of sensors, their type and placing, the detection algorithm impacts the degree of detecting using IDS. Here the key challenge is the selection and placing of the sensors

Myers[2] indicates that correlating multiple activities of multiple devices across the network provides a mechanism to understand and identify the malicious insider.[4] talks about misuse in databases and assigns a misuse weight score that is calculated based on the sensitivity of the data the user accesses.[12] discusses how combining of monitoring of access in the directory services along with the policies help to detect insider misuse if any. This also discusses the anatomy of the gaining unauthorized access for misuse.

4.1 Behavior based models

Since this problem revolves around the user and the attacks they perform, number of models and solutions have evolved that analyses the behavior of the user, predicting the users intent, deriving attack patterns etc. User profiling is a very broad, where the context of the user and the associated indicators vary depending on the profile we are looking for. Here user context refers to the user profile that revolves around, user details, system and the network usage and the command profile, log profiles and temporal details etc. The list includes monitoring of file/database accesses, system calls and command line calls by the user, organization rules and policy violations etc. The main focus on the behavior analysis is to accurately detect the unseen command/usage in the users' behavior.

Deris Stiawan[6] discusses about using the outbound traffic profile of the user to detect the exfiltration . [36]proposed a learning based behavior model that uses kernel density estimation technique and then correlating the system and the network profiles to detect the exfiltration happening in a given network. Qaio et.al [37] proposes a behavior based learning model that uses cluster based outlier detection that detects anomaly using the user system and the command profiles. Jones presents taxonomy of anomaly detection technique in network-based intrusion. Koch proposed neural network based pre-processor component with fast-learning modular. Galassi et al presents automatic construction of user profile from the users' activity log. [9] also considered the user commands as bag of words and applied text classification approach using n-gram technique.[9] presents a range of detection mechanisms based on web log, windows profiling and program profiling.

Stalfo[9] presents high order Markov chain model to identify a signature pattern based on the user commands. Schonlau et .al applied a variety of techniques on the user command sequences and the detection rates are given in the Table 2. Stolfo experimented with one-class training method using Naive Bayes and SVM of user commands. Results show that one-class SVM is as efficient as two class training algorithms.

Table 2: Two -Class Based Anomaly detection using Schonlau data set

Method	False Alarms (%)	Missing alarms (%)
Uniqueness	1.4	60.6
Bayes, one-step Markov	6.7	30.7
Hybrid multi-step Markov	3.2	50.7
Compression	5	65.8
Sequence Match	3.7	63.2
IPAM	2.7	58.9
Naïve Bayes (Updating)	1.3	38.5
Semi-Global Alignment	7.7	24.2
Eigen Co-occurrence Matrix	2.5	28
Naïve Bayes + EM	1.3	25

Covert channels both storage and timing channels are general purpose transmission mediums that can be used for good or bad. An entropy-Based Approach has been applied to network data in detecting Covert Timing Channels. Steganalysis of compressed speech and retransmission steganography are is applied to detect covert VoIP channels.[43][9]Davidson and Hirsch presents Incremental Probabilistic Action Modeling (IPAM) on the stream data, based on one-step command transition probabilities estimated from the training data. Pallabi[43] has applied both supervised and unsupervised ensemble based stream mining techniques on the stream data to detect exfiltration. The comparative shows that the ensemble based mining is effective than incremental technique. Gonzalo applied a specialized random projection matrices approach on the compressed and

uncompressed streams of covert network traffic and analyzed the performance.

4.2 Graph Based Approaches

Liu[40] presented a algorithm that studies the multi-step attack patterns, correlates the activities and then constructs attack scenario. This uses attack weighted cost and degree to the events in the attack graphs. Lawrence has applied Graph based anomaly detection methods on email, cell-phone traffic, business processes and cyber crime datasets. [23] proposes an Intent Driven Insider threat Detection model that uses a user model and the metrices to detect the insiders. In this model, the main focus is on grasping the intention of the users. A document graph is evolved to understand the user knowledge context. [17] proposes a method that uses statistical and graph based anomaly detection that analyses the short-term and the long term changes in the behavior of the employees. William Eberle [7] used graph based approach to discover anomalous structural patterns n this approach, the authors search for the activities or transactions that appear near normal. Here the task is to detect the anomalous substructure from other sub structures.

Panda [21] proposes a domain oriented approach for predicting and mitigating an insider threat. In this approach, a knowledge graph of the insider and the internal resource is generated. More the number of resources, the insider access, and the system assign a threat score and the risk is computed. The possibility of the insider threat is computed by setting predefined threshold values ranging from 1 for a very low important and 5 for very high important. The risk is calculated by including these predefined threshold values. Hui Wang et.al [22] presents framework for detecting and predicting attacks by forming attack trees that combines users' intention with their actions. Nugent proposed an intent driven framework that comprises of a user model that holds the users intention and a insider detection metrics. This uses a document graph approach

[4] presents a detailed work on profile based approaches in detecting data exfiltration. Researchers have worked on a number of techniques by profiling users commands, profiling in windows environment of their applications and processes, web accesses, program profiling etc. The table shows different techniques that have applied on the profiled data using Schonlau data set and their estimated accuracy levels.

[26] proposes a unified model that monitors the network traffic for signs of data exfiltration in addition to monitoring file transfer using FTP. This uses anomaly detection approach that learns the normal behavior of the outbound traffic and identifies the anomaly. [35] uses a anomaly detection approach using Kernel density estimation technique for detecting the data exfiltration using the network and the system parameters by correlating the system and the network parameter utilization. Since there is a correlation between the system utilization and the network bandwidth utilization.

5 Conclusion and Future work

This paper presents a detailed survey of the insider threat and techniques applied for detection , prevention and prediction of insider threats. It is interesting to note that most of these techniques uses data mining and machine learning approaches such as clustering, supervised and unsupervised learning and stream and text mining algorithms. The paper also discusses the research gaps that exist. There exists no comprehensive solution that is capable of detecting and preventing of data exfiltration happening over a network. It is very much evident that the right choice of attack indicators along with the machine intelligence techniques we would be able to address the existing gaps.

REFERENCES

1. Justin Myers, Michael R Grimaila, Robert F Mills, Towards Insider threat Detection using Web Server logs, ACM CSIIRW Journal,Apr'09
2. Ignacio J. Martinez- Moyanoa, Eliot Rich, Stephen Conrad et.al, Behavioral Theory of Insider-Threat Risks: A System Dynamics Approach, ACM Transactions on Modeling and Computer Simulation, Vol. 18, No. 2, Article 7, Publication date: April 2008
3. E. Eugene Schultz. A framework for understanding and predicting insider attacks Computers & Security, 21(6):526 – 531, 2002.
4. Jaeseung Hong, Jongwung Kim, Jeonghun Cho,The Trend of the Security Research for the Insider Cyber Threat, International Journal of Future Generation Communication and Networking, Vol. 3, No. 2, June, 2010
5. Sandip A.Kale C,Prof.S.V. Kulkarni ,Data Leakage Detection: A survey, IOSR Journal of Computer Engineering (IOSRJCE) ISSN : 2278-0661 Volume 1, Issue 6 (July-Aug 2012), PP 32-35
6. Deris Stiawan*, Mohd. Yazid Idris, Md. Sah Hj Salam and Abdul Hanan Abdullah, Intrusion threat detection from insider attack using learning behavior-based International Journal of the Physical Sciences Vol. 7(4), pp. 624 - 637, 23 January, 2012
7. WILLIAM EBERLE, and JEFFREY GRAVES, Lawrence Holder, Insider threat detection using graph based approach ,Journal of Applied Security Research, 6:32–81, 2011
8. Insider threat Analysis of Case Based System Dynamics, Advanced Computing An international Journal (ACIJ) Vol2, No.2 March 2011
9. Salvatore J.Stolfo, Insider attack and Cyber Security- Beyond the Hacker , Advances in Information security series, Volume 39,Springer ISBN-13: 978-0-387-77321-6
10. Ma'ayan Gafny, Asaf Shabtai, Lior Rokach et. Al, Detecting Data Misuse by Applying

Context-Based Data Linkage, ACM InsiderThreats'10, October 2010
11. Amir Harel, Asaf Shabtai, Lior Rokach, Yuval Elovic, M-Score: Estimating the Potential Damage of Data Leakage Incident by Assigning Misuseability Weight, ACM InsiderThreats'10, October 2010
12. William R Claycomb, Dongwan Shin, Detecting Insider Activity Using Enhanced Directory Virtualization, ACM InsiderThreats'10, October 2010
13. Suraj Nellikar, David M. Nicol, Jai J Choi, Role-based Differentiation for Insider Detection Algorithms,ACM InsiderThreats'10, October 2010
14. Andersen, D. F., Cappelli, D., Gonzalez, J. J., Mojtahedzadeh, M, Moore, A. P., Rich, E., Sarriegui,J. M., Shimeall, T. J., Stanton, J. M., Weaver, E., and Zagonel, a. 2004. Preliminary system dynamics maps of the insider cyber-threat problem. In Proceedings of the 22nd International Conference of the System Dynamics Society (Oxford, U.K., July)
15. Ahmed AL Faresi, Duminda Wijesekera, Preemptive Mechanism to Prevent Health Data Privacy Leakage, ACM MEDES 2011
16. Bishop, M. and Gates, C. 2008. Defining the insider threat. Proceedings of the 4th Annual Workshop on Cyber Security and Information Intelligence Research (Oak Ridge, Tennessee, May 12 - 14, 2008). F. Sheldon, A. Krings, R. Abercrombie, and A. Mili, Eds. CSIIRW '08, vol. 288. ACM, New York, NY, 1-3.
17. Yoohwan Kim, Frederick T. Sheldon, Lee M. Hively, Anomaly Detection in Multiple Scale for Insider Threat Analysis, ACM CSIIRW '11
18. Elisa Bertino, Gabriel Ghinita, Towards Mechanisms for Detection and Prevention of Data Exfiltration by Insiders, ASIACCS '11
19. Anderson, J.P. (1980), 'Computer Security Threat Monitoring and Surve illance',Technical Report, James P Anderson Co., Fort Washington, April 1980
20. A.H.Phyo and S.M.Furnell, A Detection-Oriented Classification of Insider IT Misuse
21. Qutaibah Althebyan, B. Panda.: Performance analysis of an insider threat mitigation model. ICDIM 2008: 703-709
22. Hui Wang, Shufen Liu, injia Zhang.: A Prediction Model of Insider Threat Based on Multi-agent. 2006 1st International Symposium on Pervasive Computing and Applications, 2006
23. Eugene Santos, Jr, Hien Nguyen, Fei Yu, Keumjoo Kim, Deqing Li, John T. Wilkinson, Adam Olson, Russell Jacob.: Intent-driven Insider Threat Detection in IntelligenceAnalyses. 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2008.
24. Yali Liu, Cherita Corbett, Rennie Archibald and Biswanath.: SIDD: A Framework for Detecting Sensitive Data Exfiltration by an Insider Attack. Proceedings of the 42nd Hawaii International Conference on System Sciences, 2009
25. Lance Spitzner: Honeypots: Catching the Insider Threat, 19th Annual Computer Security Applications Conference (ACSAC '03)
26. Nitha Rachel Suresh, Nikhil Malhotra, Rohit Kumar, B.Thanudas, An Integrated Data Exfiltration monitoring tool for a large organization with highly confidential data source, 2012 4th Computer Science and Electronic Engineering Conference (CEEC)
27. P. Garcıa-Teodoroa,*, J. Dıaz-Verdejoa, G. Macia-Fernandeza, E. Vazquezb, Anomaly-based network intrusion detection: Techniques, systems and challenges, Computers & security (2009) Pg.18–28
28. Annarita Giani and Vincent H. Berk and George V. Cybenko, Data Exfiltration and Covert Channels , Book Title "Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense V, 2006
29. Elisa Bertino,Gabriel Ghinita, Towards Mechanisms for Detection and Prevention of Data Exfiltration by Insiders, Key note paper, ASIACCS '11, March 22–24, 2011
30. David D. Clark, Susan Landau, The Problem isn't Attribution; It's Multi-Stage Attacks, ReARCH 2010
31. Sean Coyne, Ryan Kazanciyan, The Getaway: Methods and defenses for data exfiltration Black Hat DC 2011
32. Iftach Ian Amit, Advanced Data Exfiltration-The way Q would have done it- Going above and Beyond DNS and HTTP tunnels, Septemeber 2011
33. YI Hu "Profiling File Repository Access Patterns for Identifying Data Exfiltration Activities" IEEE 978-1-4244-906-9, 2011
34. Tyrell William Fawcett, Exfild: a tool for the detection of data exfiltration using entropy and encryption characteristics of network traffic.
35. Rajamenakshi Ramachandran, N.Subramanian, Ajay Shankar Bidyarthy, "Behavior model for detecting data exfiltration in Network Environment" IEEE 978-1-4577-1328-6, 2011.
36. David D Clark, Susan Landau, Untangling Attribution, Proceedings of a Workshop on Deterring Cyber attacks: Informing Strategies and Developing Options for U.S. Policy , National Research Council Spetember 2010 pg 25-40
37. Qiao H, Peng J, Feng C, Rozenblit JW Behavior Analysis-Based Learning Framework for Host Level Intrusion Detection. Proceedings of the 14th Annual IEEE International Conference and Workshops on the Engineering of Computer-Based Systems (ECBS07), 2007
38. Oh SH, Lee WK An anomaly intrusion detection method by clustering normal user

behavior. Computers. Security, 2003, 22(7): 596-612.

39. Lim SY, Jones A (2008). Network Anomaly Detection System: The State of Art of Network Behavior Analysis British Telecommunications plc, Security Research Centre, Ipswich, United Kingdom. Int. Confer. Converging. Hybrid. Inform. Technol., pp. 459-465. doi:10.1109/ICHIT.96
40. Liu Z, Wang C, Chen S (2008). Correlating Multi-Step Attack and Constructing Attack Scenarios Based on Attack Pattern Modeling.2008 International Conference on Information Security and Assurance, pp. 214-219 doi:10.1109/ISA.2008.11
41. Gonzalo Garateguy Gonzalo R. Arce Juan Pelaez ,Covert Channel detection in VoIP streams , IEEE CISS 2011 pg 1-6
42. Pallabi Parveen, Jonathan Evans, Bhavani Thuraisingham, Kevin W. Hamlen, and Latifur Khan, Insider Threat Detection using Stream Mining and Graph Mining, IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Social Computing, 2011 , pg 1102-1110