Research Journal of Applied Sciences 12 (3-4): 304-310, 2017
ISSN: 1815-932X

Survey Paper on Various Hybrid and Anomaly based Network Intrusion Detection System

J. Josemila Baby and J.R. Jeba
Department of Computer Applications,
Noorul Islam Centre for Higher Education, 629180 Kumaracoil, India

Abstract: With the dramatically development of computer network technology in our current society, the threat of cyber intrusion also highly increases. With the increase of usage in computers, criminal activity has also shifted from physical intrusion to cyber intrusion. Intrusion Detection System (IDSs) plays a significant role in monitoring and analyzing daily activities occurring in computer systems to detect occurrences of security threats. Develop security measures to prevent unapproved access to system resources and data become an urgent problem in the network security field. It is so necessary to discover the intrusion as soon as possible to take effective measures to identify the loopholes and repair the system is called as intrusion detection research. With the incredible expansion of network-based services, network protection and security is more and more significant than ever. IDSs constitute a serious security risk in networking surroundings. Data mining techniques are used to monitor and analyze large amount of network data and classify these network data into anomalous and normal data. Among the different data mining techniques, classification and clustering are the commonly used techniques to build IDS. An effective IDS requires high detection rate, low false alarm rate as well as high accuracy. Our current study presents the review and useful insights into the recent IDS techniques applied for the effective detection of normal and malicious activities in the network.

Key words: Review, intrusion detection, IDS, data mining, anomaly detection, alarm rate

INTRODUCTION

Life today is highly dependent on computers and internet. Internet is a global network connecting millions of computers linked into exchanges of data. The information communication technology has highly improved the lives of modern society. Currently, network based computer systems play a leading role in our society and its economy. The rapid increase in the number of networked computers and the widespread use of the internet in organizations have pave the way to an increase in the number of unauthorized activities, not only by external attackers but also by internal sources such as fraudulent employees or people violate their privileges for personal gain. It is very important to maintain a high level security to ensure safe and trusted communication of information between various organizations.

These immense developments in the internet world have become the targets of wide array of malicious threats that invariably turn into real intrusions. Intrusions cause disaster inside the Local Area Networks (LANs) and the time and cost to remove these damages increased in greater proportions. As a result of these, Intrusion Detection Systems (IDSs) have received tremendous attention in recent years. Intrusion detection is the process of identifying malicious activity targeted at computing and networking resources. An IDS monitors computers and/or networks to identify suspicious activity. When such an unauthorized event is detected, the IDS typically raise an alert. Identification of intrusion in a network is quite tedious. IDSs are first introduced by Anderson (1980) and later formalized by Denning (1987). Since then, intrusion detection techniques are considered as the second gate for providing network security behind firewalls. IDSs are aim to provide a layer of defense against malicious use, misuse and abuse of computing systems by sensing attacks and alerting users. In general, intrusion detection is a mechanism of gathering intrusion related knowledge occurring during the system monitoring and then analyzing collected data to draw a conclusion whether the system is intrusive or nor according the user activity, system logs, etc. After detecting some possible intrusion behaviors, the IDS raise the alarm to the network administrator and start to do the protection processing.

Corresponding Author: J.R. Jeba, Department of Computer Applications, Noorul Islam Centre For Higher Education, 629180 Kumaracoil, India

MATERIALS AND METHODS

Classification of intrusion detection system: IDSs are commonly classified based on the type of data source involved and also the detection mechanism.

IDS are the area where data mining concept is used in large measure. There are two reasons for this namely, an IDS is a very common, popular and extremely critical activity for malicious detection and prevention, large volume of data on the network is dealing so this is an ideal condition for the data mining to use it. The data mining technology has tremendous benefits in the data extracting attributes and the rule so it is significant to use the data mining techniques in the intrusion detection (Panda and Patra, 2009).

Types of networking attacks: In Computer and networking community an attack is any attempt to destroy, expose, alter, disable, steal or gain unauthorized use of a source. There are four major categories of networking attacks available as described.

Denial of Service (DoS): In DoS attack, legitimate networking requests are not served because attacker makes the resources either too busy or full to serve the request. Hence, the legitimate user cannot access the services of a machine or network resources. Example: apache, mail bomb, back, etc.

Probing (probe): In probing, attacker scans a machine or a network device for gathering the information about weaknesses or vulnerabilities that can be exploited later to compromise the target system. Example: saint, mscan, nmap, etc.

User to Root (U2R): In U2R attacks, an authorized user attempt to abuse the vulnerabilities of the system in order to gain privilege of root user for which they are not authorized. Example: perl, xterm, Fd-format, etc.

Remote to Local (R2L): In this type of attacks, a remote user tries to gain access as a local user to a local machine by sending packets to a machine over the internet. An external intruder exploits vulnerabilities of the system to access the privileges of a local user. Example: xlock, phf, guest, etc.

Literature review: In this particular study, we present a complete survey of data mining techniques that have been applied to different IDSs by several research groups is presented. A number of research articles regarding to intrusion detection are discussed below. This review is widely classified into articles related to hybrid IDS and anomaly based IDS.

Various intrusion detection system

Study on hybrid IDS: Intrusion detection systems are try to detect the attacks. It monitors and analyzes the events occur in computer or network device. In IDS technologies have categorized as misused based detection, anomaly based detection and developed version of host based (for individual computers) and network based detection (for network level). Combination of two IDS based detection, the researcher introduce the hybrid intrusion detection systems.

Panda *et al.* (2012) proposed hybrid intelligent approaches using combination of classifier to enhance the performance. First, set of data is filtering with classifier thro whole training data. Output is applied to another classifier to classify the data, experiment using KDDCUP1999 dataset. Result shows high rate detection data and low false rate.

Dhakar and Tiwari (2014), this study represent novel hybrid model for intrusion detection. Here, the proposed frame work makes use of the crucial data mining classification algorithms profitable for intrusion detection. It leads to effective, adaptive and intelligent intrusion detection. As the frame work uses TAN and REP techniques of datamining. This framework is able to detect U2R and R2L.

Tesfahun and Bhaskari (2015) have focused two layer systems that combine misuse and anomaly IDs. In first layer they are using misuse detector based on "random forest classifier". In second layer using anomaly detector based on "one class support vector machine classifier". Enhanced result detection is previously know and zero day attacks.

Moorthy and Sathiyabama (2012) provide the technique of hybrid intrusion detection system by connecting a misuse detection module to the anomaly detection module. Here anomaly detection is performed using the Bayesian network technique and misuse detection is performed using the Support Vector Machine (SVM) technique. Full system is performed by the fuzzy logic. Improvement of hybrid detection system is performance by merging the advantages of the misuse and anomaly detection.

Aydin *et al.* (2009) describe the hybrid intrusion detection system consecutively combining Packet Header Anomaly Detection (PHAD) and Network Traffic Anomaly Detection (NETAD) with snort. Here the combinations of Anomaly Detectors (PHAD and NETAD) as act like a preprocessor for snort.

Priya and Vasantha (2014) presented a new intrusion detection system called Heuristic based Hybrid Intrusion Detection System (HH-NIDS). Their approach is to identify the known and unknown attacks and the unknown intrusions are classified from the known intrusions by a series of emulations technique in a dedicated virtual machine. This proposed system consists of two clusters namely anomaly and heuristic detection clusters in which the heuristic cluster is invoked on demand.

Koshal and Bag (2012) represented two hybrid algorithms for developing the intrusion detection system. The researchers combined the C4.5 decision tree algorithm and Support Vector Machine (SVM) algorithm. The results presented after combining these two algorithms indicates that improvement in the accuracy and detection rate. This research group concluded that cascading of different algorithms shows better results that that single algorithm in the detection mechanism of intrusions.

Study on anomaly based IDS: Karami and Guerrero-Zapata (2015) presented a fuzzy anomaly detection method on hybrid Particle Swarm Optimization (PSO) algorithm in content-centric networks. They delivered the work in two different phases. The first phase is a training phase which gives the hybridization of PSO and k means algorithm with separated clusters determine and local optimization to determine the optimal number of clusters. In the second phase called detection phase, Karami and his co-worker employ a fuzzy approach by the combination of two distance-based methods as classification and outlier to detect anomalies in new monitoring data.

Feng *et al.* (2014) introduced a new machine-learning-based data classification algorithm which is applicable to intrusion detection. The basic task in this research is to classify the normal or abnormal network activities while minimizing misclassification. These researchers developed a new approach that combines the Support Vector Machine (SVM) method and the Clustering based on Self-Organized Ant Colony Network (CSOACN). It takes the advantages of both the methods and avoids the weaknesses. Their experiment result shows that the combined method outperforms SVM alone or CSOACN alone in terms of both classification rate and run-time efficiency.

Muda *et al.* (2011) developed a two learning approaches called K-means clustering and Naive Bayes classifier (KMNB) to perform intrusion detection. In this method the K-means is used in the first stage to identify groups of samples that behave similarly and dissimilarly. The second stage is used to classify all data into correct class category. Their experimental results show that KMNB method used in the current study is significantly improve and increase the accuracy, detection rate and false alarm rate.

Macia-Perez *et al.* (2011) proposes a Network Intrusion Detection System (NIDS) embedded in a smart-sensor-inspired device under a Service-Oriented Architecture (SOA) approach and this is able to operate independently as an anomaly-based NIDS or integrated transparently in a Distributed Intrusion Detection Systems (DIDS). This proposal is innovative because it combines the advantages of the smart sensor approach and the subsequent offering of the NIDS functionality. The main goal in this study is to reduce the huge volume of management tasks inherent to the network services. This research also addresses the construction of physical sensor prototype. This prototype was used to carry out the test that has demonstrated the proposals validity, providing detection and performance ratios.

Tan *et al.* (2015) presented a system that treats traffic records as images and detection of Denial-of-Service (DoS) attacks as a computer vision problem. In this study the researchers introduce a multivariate correlation analysis which accurately depict network traffic records and to convert the records into their respective images. The proposed DoS attack system is developed based on a widely used dissimilarity measure, namely Earth Movers Distance (EMD). EMD takes cross-bin matching into account and provides a more accurate evaluation on the dissimilarity between distributions. The results presented in the present study illustrate that their detection system can detect unknown DoS attacks and achieve 99.95% detection accuracy on KDD CUP 99 dataset.

Casas *et al.* (2012) explained Unsupervised Network Intrusion Detection System (UNIDS) capable of detecting unknown network attacks without using any kind of signatures, labeled traffic or training. UNIDS uses a novel unsupervised outliers detection approach based on sub-space clustering and multiple evidence accumulation techniques to pin-point different kinds of network intrusions and attacks such as Denial of Service attacks (DoS), probing attacks, propagation of worms, buffer overflows, illegal access to network resources, etc. In this study, the researchers particularly show the ability of UNIDS to detect unknown attacks, comparing its performance against traditional misuse-detection-based NIDSs. They show their algorithm used by UNIDS is highly adapted for parallel computation which permits to drastically reduce the overall analysis time of the system.

Narayana *et al.* (2011) developed their research to identify the abnormal instances in knowledge discovery and data mining. These researchers designed a novel,

unsupervised, domain independent framework that utilizes the information provided by different types of links to identify abnormal nodes. Their approach measures the dependencies between nodes and paths in the network and then applies a distance-based outlier detection method to find abnormal nodes that are significantly different from their closest neighbors. To facilitate the validation, the group designed a novel explanation mechanism that can generate meaningful and human-understandable explanations for abnormal nodes. Their research leads to potential applications in a variety of areas such as scientific discovery, data analysis and data cleaning.

Abd-Eldayem (2014) proposed the IDS based on Naive bayes classifier used to analysis the HTTP service based on traffic and identifies the HTTP normal connection and attacks. This experiment measured by NSL-KDD data set. This proposed IDS holds the high detection rate and lowest false alarm rate compared with other leading IDS.

Nadiammai and Hemalatha (2014) present datamining concept integrated with an IDS to identify relevant hidden data with less execution time. Four issues such as classification of data, high level of human interaction, lack of labeled data and effectiveness of distributed denial of service attack are being solved using the proposed algorithms like EDADT algorithm, hybrid IDS Model, Semi supervised approach and varying, HOPERAA algorithm, respectively. The above said method was tested using KDD CUP set. It shows better accuracy and reduce false alarm rate.

Mafra *et al.* (2014) proposed a new IDS Model focusing on the distributed algorithm and their computational costs. They used fault tolerance technique and cryptographic mechanism to detect the malicious. Set of distributed algorithms that support an intrusion detection model for mobile ad hoc networks.

Feng *et al.* (2014) propose new machine learning based data classification algorithm applied on IDs. Support Vector Mechanism (SVM) and the clustering based on self-organized Ant colony networks. Combining the SVM method with CSOACN to take the advantages of both while avoiding their weakness. It is evaluated using KDD99 data set alone in terms of classification rate and run time efficiency.

Ramsey *et al.* (2015) demonstrate that the precise manipulation of the physical layer header prevents a subset transceiver types from receiving the manipulated packet by soliciting acknowledgement from wireless device using a small number of packets with manipulated preambles and frame lengths, a classes >99% accuracy, irrespective environment.

Ponomarev and Atkison (2016) propose an approach to detect the intrusion into network attached industrial control system with internet benefits company and engineer who used by measuring and verifying data that is transmitted through the network. But it is not inherently the data used by transmission protocol-network telemetry. Using simulated PLC units, the developed IDS was able to achieve 94.3% accuracy when differentiating between attacker and engineer on the same network. And 99.5% accuracy when differentiating between attacker and engineer on the internet.

Eesa *et al.* (2015) proposed model uses the Cuttlefish Algorithm (CFA) as a search strategy to ascertain the optimal subset of features and the Decision Tree (DT) classifier as a judgement on the selected features that are produced by the CFA. The KDD CUP 99 data set is used to evaluate the proposed model. Result using CFA gives a higher detection rate and accuracy rate with a lower false alarm rate, when compared with the obtained results using all features.

Patel and Panchal (2015) presented the hybrid model; integrates anomaly based intrusion detection system with signature based intrusion detection is divided into two stages. In first stage, signature based IDS SNORT is used to generate alerts for anomaly data in second stage, data mining techniques "k-means+CART" is used to cascade k-means clustering and CART cup dataset. The proposed assemblage is introduced is maximize the effectiveness of identifying attacks and achieve high accuracy rate as well as false alarm rate.

Arthur and Kannan (2016) reported the cross-layer based multiclass intrusion detection system for secure multicast communication of mobile ad hoc network in military networks. Here, the proposed method delivers common content to more than one receiver at a time. Using this multicast communication system it easy to connect the front-end war field, rescue troops to carry out their missions. The researchers used efficient multilayer features to improve the accuracy of intrusion detection system in terms of detection of direct and indirect internal stealthy attacks.

Joseph *et al.* (2011) presented an autonomous host-based intrusion detection system for detecting malicious sinking behavior. The proposed detection system maximizes the detection accuracy by using cross-layer features to define a routing behavior. To study the new attack scenarios and network environments they utilize two machine learning techniques. Support Vector Machine (SVM) and Fisher Discriminant Analysis (FDA) are used together to exploit the better accuracy of SVM and faster speed of FDA. The researchers have conducted various experiments with varying network conditions and malicious node behavior.

RESULTS AND DISCUSSION

Shakshuki *et al.* (2013) proposed a new and secure intrusion detection system for MANETS. With the help of current improvements in the software technology and hardware reduction costs, the researchers expanded the MANETs into industrial applications. They developed and implemented an intrusion detection system called Enhanced Adaptive Acknowledgment (EAACK) especially for MANET. Compared to the other present approaches EAACK demonstrates higher malicious-behavior-detection rates in certain circumstances without affecting the network performances.

The different attacks and anomalies are deliberate actions against data, contents, software or hardware that can destroy, degrade, disrupt or deny access to a computer network. An efficient and effective security mechanism is required to secure content and defense against unknown and new forms of attacks and anomalies. In order to disarm new kinds of attacks, anomalous traffics and any deviation, not only the detection of the malevolent behavior must be achieved but also the network traffic belonging to the attackers should be also blocked. In an attempt to tackle with the new kinds of anomalies and the threat of future unknown attacks, many researchers have been developing Intrusion Detection Systems (IDS) to help filter out known malware, exploits and vulnerabilities.

The anomaly detection systems have two major advantages over signature based intrusion detection systems. The first advantage that differentiates anomaly detection systems from signature detection systems is their ability to detect unknown attacks as well as "zero day" attacks. This advantage is because of the ability of anomaly detection systems to model the normal operation of a system/network and detect deviations from them. A second advantage of anomaly detection systems is that the aforementioned profiles of normal activity are customized for every system, application and/or network and therefore making it very difficult for an attacker to know with certainty what activities it can carry out without getting detected. However, it is essential to address some technical challenges like the intrinsic complexity of the system, the high percentage of false alarms and the associated difficulty of determining which specific event triggered those alarms before anomaly detection systems can be adopted. So that, data mining technique is generally used to find the interesting rules from a large database depending upon the user defined support and confidence. An idea of behind using data mining is to help the decision maker to differentiate between data as useful or irrelevant.

Although, many kinds of clustering methods such as Fuzzy C-Means (FCM), K-means are widely used in intrusion detection. Usually, clustering algorithms would fit the requirements for building a good anomaly detection system. K-means is a popular anomaly detection method to classify data into different categories. However, it suffers from the local convergence and sensitivity to selection of the cluster centroids. But unfortunately from all of the above systems any of the systems so far is not completely free from defects. In order to overcome the above mentioned issues it is essential to develop a multidimensional hierarchical intrusion detection mechanism in CCN. Block diagram for proposed work is given in Fig. 1. Figure 1 explains the exact flow chart of the proposed system in our present study. Based on the above survey we would present that combining two algorithms give better results than single algorithm. Our proposed work consists of multidimensional hierarchical K-means algorithm and cuckoo search optimization algorithm. This novel hybrid method improves the accuracy and detection rate also minimizes the false alarm rate.

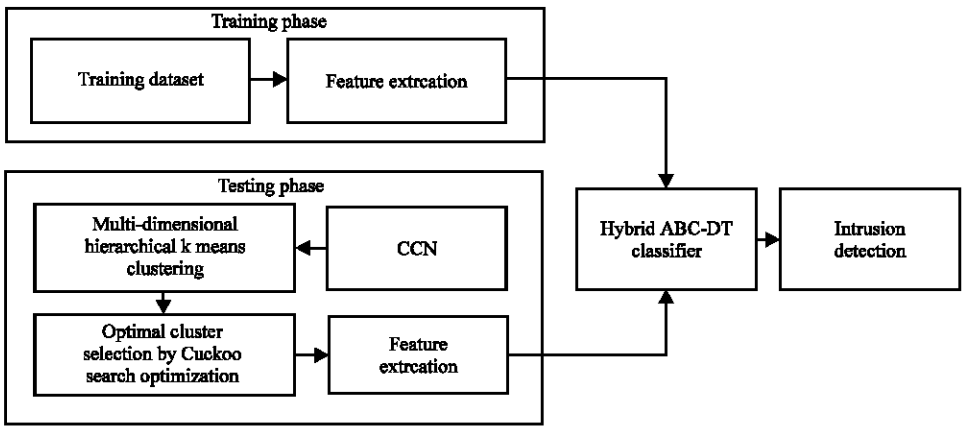

Fig. 1: Block diagram of the proposed work

CONCLUSION

Intrusion detection is attracting many researchers and networking companies due to vulnerability in the networks. Most systems today classify data either by misuse detection or anomaly detection: both approaches have its own merits and a set of drawbacks. In this study, we presented a detailed review of some important anomaly and hybrid based intrusion detection techniques and their properties together with some useful related work. A good intrusion detection system promises to allow greater confidence in the results. The proposed method will be implemented on MATLAB working platform and tested on most widely used dataset KDD CUP 99 and the results will be compared with existing methods.

REFERENCES

Abd-Eldayem, M.M., 2014. A proposed HTTP service based IDS. Egypt. Inf. J., 15: 13-24.

Anderson, J.P., 1980. Computer security threat monitoring and surveillance. James. P. Anderson Co., Washington, Pennsylvania.

Arthur, M.P. and K. Kannan, 2016. Cross-layer based multiclass intrusion detection system for secure multicast communication of MANET in military networks. Wirel. Netw., 22: 1035-1059.

Aydin, M.A., A.H. Zaim and K.G. Ceylan, 2009. A hybrid intrusion detection system design for computer network security. Comput. Electr. Eng., 35: 517-526.

Casas, P., J. Mazel and P. Owezarski, 2012. Unsupervised network intrusion detection systems: Detecting the unknown without knowledge. Comput. Commun., 35: 772-783.

Denning, D.E., 1987. An intrusion-detection model. IEEE Trans. Software Eng., SE-13: 222-232.

Dhakar, M. and A. Tiwari, 2014. A novel data mining based hybrid intrusion detection framework. J. Inf. Comput. Sci., 9: 037-048.

Eesa, A.S., Z. Orman and A.M.A. Brifcani, 2015. A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems. Expert Syst. Applic., 42: 2670-2679.

Feng, W., Q. Zhang, G. Hu and J.X. Huang, 2014. Mining network data for intrusion detection through combining SVMs with ant colony networks. Future Gener. Comput. Syst., 37: 127-140.

Joseph, J.F.C., B.S. Lee, A. Das and B.C. Seet, 2011. Cross-layer detection of sinking behavior in wireless ad hoc networks using SVM and FDA. IEEE Trans. Dependable Secure Comput., 8: 233-245.

Karami, A. and M. Guerrero-Zapata, 2015. A fuzzy anomaly detection system based on hybrid PSO-kmeans algorithm in content-centric networks. Neurocomputing, 149: 1253-1269.

Koshal, J. and M. Bag, 2012. Cascading of C4.5 decision tree and support vector machine for rule based intrusion detection system. Intl. J. Comput. Netw. Inf. Secur., 4: 8-20.

Macia-Perez, F., F. Mora-Gimeno, D. Marcos-Jorquera, J.A. Gil-Martinez-Abarca, H. Ramos-Morillo and I. Lorenzo-Fonseca, 2011. Network intrusion detection system embedded on a smart sensor. IEEE Trans. Ind. Electron., 58: 722-732.

Mafra, P.M., J.S. Fraga and A.O. Santin, 2014. Algorithms for a distributed IDS in MANETs. J. Comput. Syst. Sci., 80: 554-570.

Moorthy, M. and S. Sathiyabama, 2012. A hybrid data mining based intrusion detection system for wireless local area networks. Intl. J. Comput. Appl., 49: 19-28.

Muda, Z., W. Yassin, M.N. Sulaiman and N.I. Udzir, 2011. A K-means and naive bayes learning approach for better intrusion detection. Inform. Technol. J., 10: 648-655.

Nadiammai, G.V. and M. Hemalatha, 2014. Effective approach toward intrusion detection system using data mining techniques. Egypt. Inf. J., 15: 37-50.

Narayana, M.S., B.V.V.S. Prasad, A. Srividhya and K.P.R. Reddy, 2011. Data mining machine learning techniques: A study on abnormal anomaly detection system. Intl. J. Comput. Sci. Telecommun., 2: 8-14.

Panda, M. and M.R. Patra, 2009. Ensembling rule based classifiers for detecting network intrusions. Proceedings of the 2009 International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom'09), October 27-28, 2009, IEEE, Kottayam, India, ISBN:978-1-4244-5104-3, pp: 19-22.

Panda, M., A. Abraham and M.R. Patra, 2012. A hybrid intelligent approach for network intrusion detection. Procedia Eng., 30: 1-9.

Patel, J. and K. Panchal, 2015. Effective intrusion detection system using data mining technique. J. Emerging Technol. Innovative Res., 2: 1869-1876.

Ponomarev, S. and T. Atkison, 2016. Industrial control system network intrusion detection by telemetry analysis. IEEE. Trans. Dependable Secure Comput., 13: 252-260.

Priya, N. and S. Vasantha, 2014. Heuristic based hybrid network intrusion detection system: A novel approach. Asian J. Inf. Technol., 13: 733-738.

Ramsey, B.W., B.E. Mullins, M.A. Temple and M.R. Grimaila, 2015. Wireless intrusion detection and device fingerprinting through preamble manipulation. IEEE. Trans. Dependable Secure Comput., 12: 585-596.

Shakshuki, E.M., N. Kang and T.R. Sheltami, 2013. Eaack: A secure intrusion-detection system for Manets. IEEE. Trans. Ind. Electron., 60: 1089-1089.

Tan, Z., A. Jamdagni, X. He, P. Nanda and R.P. Liu *et al*., 2015. Detection of denial-of-service attacks based on computer vision techniques. IEEE. Trans. Comput., 64: 2519-2533.

Tesfahun, A. and D.L. Bhaskari, 2015. Effective hybrid intrusion detection system: A layered approach. Intl. J. Comput. Netw. Inf. Secur., 7: 35-41.