Hybrid Approach for Optimised Intrusion Detection System

Mr.M.Mangaleswaran

Assistant Professor
Department of Computer Science and Engineering
Jansons Institute of Technology
Coimbatore- 641659, India

Abstract
Intrusion detection is an important way to ensure the protection of computers and networks. In this paper, a new intrusion detection system is proposed based on Hidden Conditional Random Fields. In order to improve the performance of HCRFs, we bring forward the Two-stage Feature Selection method, which contains Manual Feature Selection method and Backward Feature Elimination Wrapper method. The BFEW is an aspect selection method which is introduced based on wrapper approach. Experimental results on KDD99 dataset show that the proposed IDS not only have a great advantage in detection efficiency but also have a higher accuracy when compared with other well-known methods. With the ever increasing number and diverse type of attacks, including novel and previously invisible attacks, the success of an Intrusion Detection System is very essential. Hence there is high demand to reduce the threat level in networks to ensure the information and services accessible by them to be extra secure. In this paper we developed a practical test suite for getting better the competence and precision of an intrusion detection system use the layered CRFs. We set up altered types of checks at several levels in each layer .Our framework examine various attribute at every layer in order to effectively classify any infringe of security. Once the attack is detected, it is intimated through mobile phone to the system administrator for preservation the server system. We established experimentally that the layered CRFs can thus be more professional in detecting intrusions when compared with the other previously known techniques.

Keywords
Intrusion detection system; Hidden conditional Random Fields; Conditional random fields; Anomalous Activity; Layer-based Intrusion Detection System.

I. INTRODUCTION

The emergence and development of the network has altered the life of human. Regrettably, we have to tackle a variety of risks of network intrusion while enjoying the ease brought by the network. In order to shrink and evade the risks, a series of intrusion prevention techniques like firewalls have been presented. The IDS monitors the events happening in a system energetically, and decides whether these actions are intrusions or just standard behaviours. IDSs can be classified as anomaly based or signature based according to the attack detection method. Another method for intrusion detection is to make use of both the normal and the abnormal behaviour for training IDS. This combines the advantages of both the anomaly-based and the signature-based methods.

Intrusion Detection Systems are based on two concepts, matching of the formerly seen and hence known anomalous patterns from an internal database of signatures or building profiles based on normal data and detecting deviations from the expected behaviour. Based on the type of deployment, the Intrusion Detection Systems are classifying as Network based and Host based. Network based systems make an assessment by Analysing the network logs and packet headers from the inward and outgoing packets. Host based systems monitor's individual systems and uses system logs broadly to make any decision. Intrusion Detection Systems are moreover Signature based or Behaviour based. The main purpose of intrusion detection is that finding the attack which are precise in the signatures and as well as finding the new or invisible attacks efficiently and also come up with large amount of network traffics. Network Intrusion Detection Systems are placed at a considered point or points within the network to monitor traffic to and beginning all devices on the network. Ideally one would scan all inbound and outbound traffic. HIDS Host Intrusion Detection Systems are run on individual hosts or devices on the network. A HIDS monitors the inbound and outbound packets from the device only and resolve aware the user or administrator if suspicious activity is detected. It will analyze network traffic and system-specific settings for effectively decision the attacks survive in the current network environment. A signature based IDS will monitor packets on the network and compare them against a database of signatures or patterns of known malicious threats. An IDS which is anomaly based will monitor network traffic and compare it against an established baseline. The Generic representation of this system is shown in Fig.1. Intrusion Detection Systems refers to a program used to identify an intrusion once it happens and to prevent a system from being compromised. An intrusion detection system monitors the activities of a given environment and detects inaccurate and Inappropriate and anomalous activity as defined by the Sysadmin, Audit, Networking, and Security institute.

Manuscript received February 5, 2025
Manuscript revised February 20, 2025
https://doi.org/10.22937/IJCSNS.2025.25.2.13

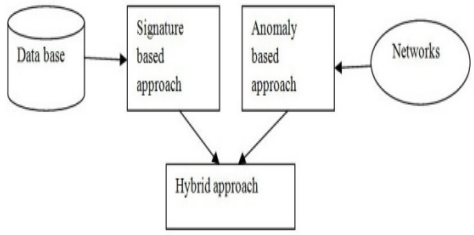

Fig.1. Generic representation of the system

Hybrid approach is another method for intrusion detection which is trained with both the common and the recognized anomalous patterns. Hybrid systems are efficient and perform classification on test data. An IDS is a security counter measure. It monitors things looking for signs of intruders. An intrusion detection system is a software and/or hardware designed to detect unwanted attempts at accessing, manipulating or disabling of computer systems mainly through a network. In the present system, Conditional Random Fields with multilayer approach used to build Intrusion Detection System that is efficient in detecting a broad variety of attacks. In the present work CRF approach is efficiently used for attack pattern recognition. Layered Framework for building intrusion detection systems has been introduced which can detect a broad variety of attacks constantly and efficiently when compared to the traditional network intrusion detection systems. In the present layered framework, a number of subsystems individually trained with KDD'99 training data and consecutively arranged sub systems in order to reduce the number of false alarms and increase the attack detection coverage. An intrusion detection system monitors the performance of a given environment and decides whether these activities are malicious or legitimate based on system reliability, secrecy and the availability of information resources. Intrusion detection as distinct by the Sysadmin, check, Networking, and safety institute is the act of detecting actions that attempt to compromise the privacy, reliability or accessibility of a resource. Detecting intrusions in networks and applications has become one of the most crucial tasks to prevent their abuse by attackers. The cost involved in protecting these expensive resources is often insignificant when compared with the actual cost of a successful intrusion, which strengthens the need to extend more influential intrusion detection systems. There are two types of IDS depending on their method of deployment and data used for analysis. Network Intrusion Detection Systems and the other is Host Intrusion Detection Systems .NIDS monitors the packets from the network and it is an free platform that identifies intrusion by examining the network traffic and multiple hosts .Hybrid systems can be very efficient, subject to the classification process used, and can also be used to label unobserved or new instances as they assign one of the identified classes to each test instance. This is feasible because during training the system learns features from all the classes. The only unease with the hybrid method is the accessibility of labelled data.

II. CONDITIONAL RANDOM FIELDS

Conditional random fields are a group of numerical modelling mode often applied in pattern recognition and machine learning, where they are used for structured prediction. Whereas an common classifier predict a label for a single sample devoid of regard to "neighbouring" samples, a CRF can take context into account; e.g., the linear chain CRF popular in accepted language processing predicts sequences of labels for sequences of input samples. CRFs are a category of discriminative undirected probabilistic graphical form. It is used to encode known relationships between observations and create constant Interpretations. It is frequently used for labelling or parsing of chronological data, such as natural language text or biological Sequences and in computer idea. Specifically, CRFs find applications in low parsing, named entity recognition and gene finding, among other tasks, being an alternative to the related hidden Markov models. In computer idea, CRFs are often used for object recognition and image segmentation.

III. RELATED WORKS

As a vital barrier to protect computers and networks from intrusions, intrusion detection system must be skilled of detecting network intrusions perfectly and processing huge amounts of network traffic powerfully. In order to handle these two issues, researchers have brought forward a variety of methods, and have made unremitting efforts the authors applied Support Vector Machines to intrusion detection. SVMs are proficient of handling these datasets which have high dimension feature space. The Naive Bayes method is also applied to intrusion detection. However, there is a conditional independence assumption imposed on this method. Manual Feature Selection on each type of attacks, and their experimental results show that the MFS performs improved than some regularly used automatic characteristic selection methods such as Neural Network and PCA. There are a number of methods and frameworks been projected and many systems have been built to detect intrusions. Various techniques such as association rules, clustering, naive Bayes classifier, artificial neural networks, support vector machines, genetic algorithms, and others have been useful to detect intrusions.

A. Data mining Approach

Data mining approaches for intrusion detection include association rules, common episodes and outlier detection, which are based discovering significant patterns of program and user behaviour for building up the Classifiers.

B. Data clustering methods

Data clustering methods in interruption detection includes the k-means and the fuzzy c-means Clustering methods. The main drawbacks in these clustering technique is that it is based on determining the numeric distance among the observations resulting that, the observations must be numeric .Hence observations with symbolic features cannot be easily identified and used for clustering, resulting in inaccuracy for finding the attacks.

C. Naive Baye's classifiers

The next approach discussed here in intrusion detection is Naive Bayes classifiers. Those approaches will make strict free assumption between the features in an observation records which resulting in very low detection accuracy when the features in the observation are having correlation between them

D. Decision trees

The intrusion detection also performed using Decision trees approach this approach presents decision tree techniques that are used to automatically learn intrusion signatures and perform the arrangement activities in computer network systems as ordinary or intrusive.

E. Neural Networks

The neural network components also used for finding the intrusive events in the network .The neural network in intrusion will work well with correlated kind of data in the observation.

IV. CHALLENGES AND REQUIREMENT

It is important intrusion detection must notice attacks at a premature stage in order to minimize their collision. The major challenges and requirements for building intrusion detection systems are:

- The system must be able to detect as many attacks as possible lacking giving false alarms.
- The system must be able to handle large amount of data without disturbing performance and without sinking data.
- A system must not only detect an attack, but also able to classify the type of attack.
- A system must be resistant to attacks since, a system that can be broken during an attack may not be able to detect attacks constantly.
- The challenge is to build a system which is scalable and can be easily adapted as per the specific requirements of the environment where it is deployed.

V. ANOMALY DETECTION

Intrusion Detection System plays key role of detecting various kinds of attacks and secures the applications and networks in the pervasively connected network environment. Intrusion detection is the method of monitoring computers or networks for unauthorized access, activity, or file modification. Anomaly-based IDS establish a baseline of normal usage patterns, and anything that commonly deviates from it gets flagged as a possible intrusion. Anomaly detection method can investigate user patterns, such as profiling the programs executed daily or the confidential processes executed with access to resources that are unreachable to ordinary user. The major advantage of anomaly-based IDS is their ability to detect attempts to exploit new and unpredicted vulnerabilities. Their difficulty, due to the inherently dynamic nature of computer networks, is their major disadvantage, as well as their high false alarm rate because the entire scale of the behaviour of an information system may not be roofed during the learning phase. To overcome this drawback we train our system using conditional random fields as they are more accurate and due to this false alarm rate goes on decreasing.

A. Analysis of Activities

The analysis is the heart of the anomaly intrusion detection system. In this system we investigate user patterns, such as profiling the programs executed on a daily basis or the privileged processes executed with admittance to resources that are inaccessible to ordinary user. For this we collect the volatile data from the system. To collect this data we use system log file which gives us the number of processes which are running on the system, the number of resources which are assigned to the user, and the system privileged.

B. Prevention of Anomalous Activity

Once the anomalous activity occurs, we can prevent it. The admin may log on to the system locally or remotely. If the admin is at local level then he/she can view the running activities, or he/she can stop the anomalous activity and if the admin is at remote level then he/she can log on to the system using Internate or GPRS using cell phone. After that the user can stop anomalous activity, or start new activity. But if the controlling of the anomalous activity is not possible then admin may shutdown or reboot that system.

VI. INTRUSION DETECTION

Intrusion detection is the method of monitoring computers or networks for illegal entrance, activity, or file alteration. An IDS is a device or application used to scrutinize all network traffic, thereby detecting if a system is being under attack by a network such as a denial of service attack. In some cases the IDS may also respond to uncharacteristic or malicious traffic by taking action such as jamming the user or source IP address from accessing the network. IDS protect a network and attempt to prevent intrusions.

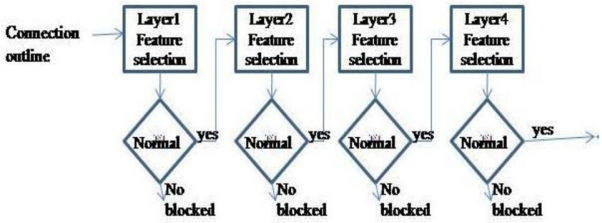

Fig 2: Layered security model

They don't fully assurance security, but when used with security procedure, susceptibility assessments, data encryption, user authentication, access control, and firewalls, they can greatly develop network safety .Intrusion detection systems provide three essential security functions: they monitor, detect, and respond to unauthorized activity by company insiders and outsiders. Intrusion detection systems use policies to term certain events that, if detected will issue an alert.

VII. PIPELINING OF LAYERS

A. Layered security model

Layered security model is a sequential model in which number of security checks are performed one after other in sequence. Sequential Layered Approach and is based on ensuring accessibility, secrecy, and reliability of data and services over a network. The objective of using a layered security model is to condense computation and the generally time required detecting abnormal events. The time required to detect an intrusive occurrence is considerable and can be reduced by eliminating the communication overhead among different layers. This can be achieved by making the layers independent to block an attack without the need of a central decision maker. In model every layer is trained separately and then deployed successively. We define four layers that correspond to the four attack groups. They are Probe layer, DoS layer, R2L layer, and U2R layer.

B. Pipelining

Pipelining is used by virtually all modern multicore processors to enhance performance by overlapping the execution of instructions. The pipeline is divided into segments and each segment can implement it operation simultaneously with the other segments. Once a fragment completes an operation, it passes the result to the next fragment in the pipeline and fetches the next operations from the preceding segment.

C. Proposed Pipelined Layered Security Model

The layered security model improves the intrusion detection system performance by finding attacks in four layers. If the attack is found in initial layer then it is blocked otherwise test instance is passed to next layer. If the test instance is passed through all the four layers then it indicates there is no attack. In this technique the test instance is passed sequentially through all the layers. Only one test instance is checked at a time. Next instance has to wait until the previous instance completes its checking through all layers.

VIII. FEATURE SELECTION FOR EACH LAYER

We first select four layers corresponding to the four attack groups those are Probe, DoS, R2L, and U2R; and perform feature selection for each layer. We illustrate our approach for selecting features for every layer and why some features were chosen over others.

A. Denial of Service Layer

Denial of Service attack (DoS) is an attack in which the attacker makes some computing or reminiscence source too busy or too full to hold genuine requests, or denies legitimate users access to a machine. Therefore, for the DoS layer, traffic features such as the "component of links having similar object host and related service" and packet level features such as the "source bytes" and "percentage of packets with errors" is considerable. To spot DoS attacks, it may not be vital to know whether a client is "logged in or not."

B. Probe Layer

Probing attack is an attempt to gather information about a network of computers for the apparent purpose of circumventing its security controls. Therefore, essential connection level features such as the "duration of connection" and "source bytes" are important while features like "sum of files creations" and "sum of files accessed" are not ordinary to provide information for detecting probes.

C. User to Remote Layer

User to Remote attack is a class of utilize in which the attacker starts out with access to a regular user account on the system and is able to exploit some vulnerability to gain root access to the system.

D. Remote to Local Layer

Remote to Local attacks occurs when an attacker who has the ability to send packets to a machine over a network but who does not have an account on that machine exploits some susceptibility to gain restricted access as a user of that machine.

IX. PROPOSED System

In our proposed system we depict the Layer-based Intrusion Detection System. The LIDS draws its inspiration from what we call as the Airport Security model, where a quantity of security checks are performed one behind the other in a series. Similar to this model, the LIDS represents a sequential Layered Approach and is based on ensuring accessibility, secrecy, and integrity of data and (or) services over a network. The purpose of using a layered model is to trim down computation and the generally time required detecting Anomalous events. The time required to detect an interfering event is important and can be reduced by eliminating the communication slide among different layers. We classify four layers they are Probe layer, DoS layer, R2L layer, and U2R layer. Every layer is separately trained with a small set of characteristics. The layers basically act as filters that chunk any anomalous connection, thereby eliminating the need of advance processing at subsequent layers enabling quick response to intrusion. The effect of such a sequence of layers is that the anomalous events are known and blocked as soon as they are detected. just the once the attack is detected, it is intimated through mobile phone to the system administrator for secure guarding the server system. We realize the LIDS and select four set of features which reduces the computational period. Methods such as naive Bayes assume independence among the observed data.

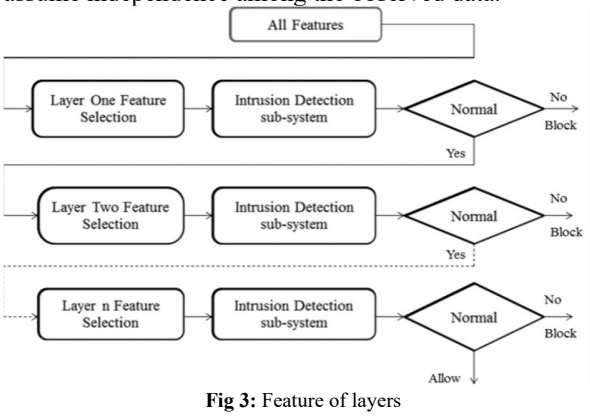

Fig 3: Feature of layers

To balance this trade-off, we use the CRFs that are more perfect, though pricey; however we execute the Layered Approach to improve overall system performance. Our proposed system, Layered CRFs, performs drastically superior than last systems.

Proposed Algorithm:

Step 1: Choose the number of layers, n, for the whole system.

Step 2: Separately execute nature selection for each layer.

Step 3: Plug in the layers sequentially such that only the connections categorized as regular are passed to the next layer

Step 4: For every (next) test instance perform Steps 5 through 8.

Step 5: Test the case and label it whichever as attack or normal.

Step 6: If the instance is labelled as an attack, block it and then identify it as an attack with the corresponding Layer name at which it is detected and go to step 4. Pass the sequence to next layer.

Step 7: If the present layer is not the last layer in the system, test the occurrence and go to step 6.Else go to step 8.

Step 8: Test the instance and label it any as normal or as an attack. If the instance is considered as an attack, block it and identify it as an attack related to the layer name.

Step 9: If the instance is considered as an attack at any layer then near it to system admin's mobile with a corresponding appropriate message of attack.

X. CONCLUSION AND FUTURE WORK

Hybrid intrusion detection is a novel kind of model combining the advantages of anomaly based intrusion detection and signature based intrusion detection. Intrusion and anomaly are two different kinds of uncharacteristic traffic events in an open network environment. An intrusion takes place when an unlawful access of a host computer system is attempted. A glitch is observed at the network connection level. Both attack types may compromise important hosts, disclose sensitive data, deny services to genuine users, and pull down network based computing resources. The intrusion detection system offers intellectual protection of networked computers or distributed resources much better than using fixed-rule firewalls. Existing IDSs are built with any signature-based or anomaly-based systems. Signature matching is based on a misuse model, whereas anomaly detection is based on a normal use model. The

design philosophies of these two models are quite poles apart, and they were not often mixed up in existing IDS products from the security industry. The signatures are manually constructed by security experts analyzing earlier attacks. The collected signatures are used to match with arriving traffic to detect intrusions. These are knowable systems that detect identified attacks with low false alarms. However, the signature-based IDS cannot detect mysterious attacks without any recollected signatures or lack of hit classifiers.

The Hybrid Intrusion Detection System integrates the litheness of Anomalous Detection System with the exactness of a signature-based Intrusion Detection System. Anomalous Detection System is planned by acquiring the unpredictable data when there is no any anomalous activity. Here we train our system using conditional random field for the anomalous activity. This new approach mechanically enables HIDS to detect related anomalous attacks in the future.

REFERENCES

[1] KK Gupta, B Nath and K Ramamohanarao, “Layered approach using conditional random fields for intrusion detection”, IEEE Transactions on Dependable and Secure Computing, vol. 7, no. 1, (2010), pp. 35-49.
[2] S Mukherjeea and N Sharma, “Intrusion Detection using Naive Bayes Classifier with Feature Reduction”, Procedia Technology, vol. 4, (2012), pp. 119-128.
[3] SJ Horng, MY Su, YH Chen, TW Kao, RJ Chen, JL Lai and CD Perkasa, “A novel intrusion detection system based on hierarchical clustering and support vector machines”, Expert Systems with Applications, vol. 38, (2011), pp. 306-313.
[4] J Vlcek and L Luksan, “Generalizations of the limited-memory BFGS method based on the quasi-product form of update”, Journal of Computational and Applied Mathematics, vol. 241, (2013), pp. 116-129.
[5] Li, J Xia, S Zhang, J Yan, X Ai and K Dai, “An efficient intrusion detection system based on support vector machines and gradually feature removal method”, Expert Systems with Applications, vol. 39, no. 1, (2012), pp. 24-430.
[6] V Bolón-Canedo, N Sánchez-Maroño and A Alonso-Betanzos, “Feature selection and classification in multiple class datasets: an application to KDD Cup 99 dataset”, Expert Systems with Applications, vol. 38, no. 5, (2011), pp. 5947-5957.
[7] W Alsharafat, “Applying Artificial Neural Network and eXtended Classifier System for Network Intrusion Detection”, The International Arab Journal of Information Technology, vol. 10, no. 3, (2013), pp. 230-238.
[8] L Zhang, LG Meng and CJ Hou, “Intrusion Detection Based on Immune Principles and Fuzzy Association Rules”, Intelligence Computation and Evolutionary Computation, vol. 180, (2013), pp. 31-35.
[9] C Guo, YJ Zhou, Y Ping, ZK Zhang, GL Liu and YX Yang, “A distance sum-based hybrid method for intrusion detection”. Appl Intell, vol. 40, (2014), pp. 178-188.
[10] SANS Institute, (2012) Intrusion Detection FAQ.http://www.sans.org/resources/idfaq/
[11] Autonomous Agents for Intrusion Detection, http://www.cerias. purdue.edu/research/aafid/, 2010.
[12] Kapil Kumar Gupta, Baikunth Nath, Senior Member, IEEE, and Ramamohanarao Kotagiri, Member, IEEE, ―Layered Approach Using Conditional Random Fields for Intrusion Detection‖, ieee transactions on dependable and secure computing, vol. 7, no. 1, January -march 2010
[13] CRF++: Yet Another CRF Toolkit, http://crfpp.sourceforge.net/, 2010.
[14] KDD Cup 1999 Intrusion Detection Data, http://kdd.ics.uci.edu/ databases/ kddcup99/ kddcup99.html, 2010.
[15] J.P. Anderson, Computer Security Threat Monitoring and Surveillance,http://csrc.nist.gov/publications/history/ande80.pdf, 2010