A Research Paper on Hybrid Intrusion DetectionSystem

[1]NILAMADHABA MISHRA, *Gandhi Institute of Excellent Technocrats, Bhubaneswar,India*

[2]RAMYA RACHITA ROUL, *KMBB College of Engineering and Technology, Khordha, Odisha, India*

Absrtact-_An intrusion detection system (IDS) is a device orsoftware application that monitors network or systemactivitiesfor malicious activities or policy violations and produces reportsto a Management Station. Some systems may attempt to stop anintrusion attempt but this is neither required nor expected of amonitoring system. Intrusion detection and prevention systems(IDPS) are primarily focused on identifying possible incidents,logginginformationaboutthem,andreportingattempts.In addition, organizations use IDPSes for other purposes, such asidentifying problems with security policies, documenting existingthreats and deterring individuals from violating security policies.IDPSeshavebecomeanecessaryadditiontothesecurityinfra structure of nearlyeveryorganization.Different methodscanbeusedtodetectintrusionswhichmakeanumberofassumptionsthatarespecificonlytotheparticularmethod.Hence, in addition to the definition of the security policy and theaccesspatternswhichareusedinthelearningphaseofthedetector, the attack detection capability of an intrusion detectionsystem also depends upon the assumptions made by individualmethodsforintrusiondetection.Thepurposeofanintrusiondetectionsystemistodetectattacks.However,itisequallyimportant to detect attacks at an early stage in order to minimizetheirimpact.IhaveusedDatasetandClassifiertorefineIntrudersinNetworks.

Keywords-(IDS),(IDPS),IDPSes.

I. INTRODUCTION

In the 21st century the development of telecommunicationsnetworkshastakengiantleapsfromcircuitandpacketswitchednetworkstowardsall-IPbasednetworks.Thisdevelopmenthascreatedaunifiedenvironmentwherecommunication of applications and services (data and voice)arebeingtransferredontopofthe IP-protocol. Although the development of communication networks hasbeen towards a better sustainability of technologies it hasalso raised new unwanted possibilities. Threats that wereapplicable only in the fixed networks are now feasible in theradio access networks. When taken into account that threatsarebecomingmoreandmoresophisticateditalsomeansthat the security systems have to become more intelligent.Thebasicsecuritymeasurementssuchasfirewallsandantivirusscannersareintheirlimitstocopewiththeovergrowing number of intelligent attacks from the Internet.A solution to enhance the overall security of the networks istoincreasethesecuritylayerswithintrusiondetectionsystems. Tounderstandwhatroleintrusiondetectionhasintelecommunications networks it can be thought through asimpleexample. Thinkofintrusiondetectionasasecurityguardthatisguardingthe frontgate ofafactorypremises.

Thepremisesofthefactoryrepresentthenetworkofamobile operator and the fence surrounding the factory is theoperator’s firewall. Employees of the factory represent thetraffic in the operator’s network. It is know that factories arewell protected and they do not want to let people inside thepremises that do not have the required clearances. The fenceor firewall in this case, is in charge to keep all unwantedvisitors outside the factory premises. Just like in a firewall, afencehasholes (gates) in it to let employeesmove in andout of the factory premises. These holes in the fence thoughleave the factory vulnerable to the unwanted visitors and thisis why the factory has a security guard guarding the gate.Depending on the role that the security guard is in, while heis monitoring the people going in and out of the factorypremises, he either notifies the head of security when hedetectsasuspiciouslookingpersonwalkingthroughthegate. Or he steps in and prevents this person from enteringthe factory premises. The basic functionality of an intrusiondetection system is the first example of the security guard.IDS generate an alarm when it detects something suspiciousandthenthesecuritypersonnelofthenetworkoperator furtherinvestigatethecauseofthealarm.Anintrusiondetection system (IDS) is a device, typically a designatedcomputersystem,whichmonitorsactivitytoidentify maliciousorsuspiciousalerts.Itisplacedinsideanorganisation to monitor what occurs within the network oftheorganisation.Thegoalofanintrusiondetectionsystemistoaccuratelydetectcomputersecurityincidents,andnotifynetwork administrators.Adistinctionismadebetweenalertsandincidents byanintrusiondetectionsystem.Alerts are definedasall the observable actions onthe computer network that are picked up by the sensors of anintrusiondetectionsystem.Incidentsaremaliciousorsuspiciousalertsthathaveahighenoughvaluetobeconsideredasecurity-relevantsystem eventinwhich thesystem’s security policy is disobeyed or otherwise breached.AnIDSconsistsoffourcomponents,accordingtotheCommonIntrusionDetectionFramework(CIDF);eventgenerators, analysers, event databases and response units. Inthe research of this thesis, Dataset is used to provide attacksandnormaldatatoanalyzer.Aneffortwillbemadetochoose a machine learning method that can be used as ananalyser,whichimprovesthedetectionratealertsfromincidents.Aneventdatabasewillbeusedtotraintheanalyser, and to evaluate its predictions. The response unitswillnotbewithinthescopeofthisthesis,butcanbecontrolled bythe decisionsofthe analyser.

II. INTRUSIONDETECTIONANDINTRUSION DETECTIONSYSTEM

The intrusion detection systems are a critical component inthenetworksecurityarsenal.

2.1 PrinciplesandAssumptionsinIntrusionDetection

Denning defines the principle for characterizing a systemunder attack. The principle states that for a system which isnotunder attack,thefollowingthreeconditions hold true:

1. Actionsofusersconformtostatisticallypredictablepatterns
2. Actions of users do not include sequences which violatethesecuritypolicy.
3. Actionsofeveryprocesscorrespondtoasetofspecifications whichdescribewhattheprocessisallowed todo.

Systems under attack do not meet at least one of the threeconditions. Further, intrusion detection is based upon someassumptionswhicharetrueregardlessoftheapproachadoptedbytheintrusiondetectionsystem.Theseassumptionsare:

1. There exists a security policy which defines the normaland (or)theabnormalusageofeveryresource.
2. Thepatternsgeneratedduringtheabnormalsystemusage are different from the patterns generated duringthe normal usage of the system; i.e., the abnormal andnormal usage of a system results in different systembehavior. This difference in behavior can be used todetectintrusions.

As we shall discuss later, different methods can be used todetect intrusions which make a number of assumptions thatare specific only to the particular method. Hence, in additiontothedefinitionofthesecuritypolicyandtheaccesspatterns which are used in the learning phase of the detector,theattackdetectioncapabilityofanintrusiondetectionsystemalsodependsupontheassumptionsmadebyindividualmethodsforintrusiondetection.

2.2 ComponentsofIntrusion DetectionSystems

An intrusion detection system typically consists of three subsystemsor components:

1. **Data Preprocessor** – Data preprocessor is responsibleforcollectingandprovidingtheauditdata(inaspecified form) that will be used by the next component(analyzer)tomakeadecision.Datapreprocessor is,thus,concernedwithcollectingthedatafromthedesired source and converting it into a format that iscomprehensible by the analyzer.Data used for detectingintrusions range from user access patterns (for example,thesequenceofcommands issuedat the terminal andtheresourcesrequested)tonetworkpacketlevelfeatures (suchasthesourceanddestinationIPaddresses,typeof packets andrateof occurrenceofpackets) to application and system level behavior (suchas the sequence of system calls generated by a process.)Wereferto thisdataasthe auditpatterns.
2. **Analyzer (Intrusion Detector)** – The analyzer or theintrusion detector is the core component which analyzesthe audit patterns to detect attacks. This is a criticalcomponentandoneof themostresearched.Variouspattern matching,machine learning, datamining andstatistical techniques can be used as intrusion detectors.The capability of the analyzer to detect an attack oftendeterminesthestrengthoftheoverallsystem.
3. **Response Engine** – The response engine controls thereactionmechanismanddetermineshowtorespondwhen the analyzer detects an attack. The system maydecide either to raise an alert without taking any actionagainstthesourceormaydecidetoblockthesourcefor a predefined period of time. Such an action dependsuponthepredefined securitypolicyofthenetwork

TheauthorsdefinetheCommonIntrusionDetectionFramework(CIDF)whichrecognizesacommonarchitectureforintrusiondetectionsystems.TheCIDFdefines four components that are common to any intrusiondetectionsystem.Thefourcomponentsare;Eventgenerators(E-boxes),eventAnalyzers(A-boxes),eventDatabases (D-boxes) and the Response units (R-boxes). Theadditional component, called the D-boxes, is optional andcanbe usedfor later analysis.

III. PROPOSEDWORK

Weusetwoclassificationtechniquesforourproposedarchitecture,inacombinedmanner.Consequently,anincreasing number of approaches have been developed foraccomplishing such purpose, including k-nearest-neighbor(KNN)classification,NaïveBayesclassification,supportvector machines (SVM), decision tree (DT), neural network(NN),andmaximumentropy.Ourchoiceamongallavailableclassificationtechniquesisdependsuponourstudies about all classifier. We put our motivations for theseclassifiersinbelowtopic at aglance.

3.1 Bayes'Theorem

Let $\boldsymbol{X}$ be a data tuple. In Bayesian terms, $\boldsymbol{X}$ is considered"evidence." As usual, it is described by measurements madeon a set of n attributes. Let H be some hypothesis, such asthat the data tuple $\boldsymbol{X}$ belongs to a specified class C. Forclassification problems, we want to determine $P(H|\boldsymbol{X})$, theprobability that the hypothesis H holds given the "evidence"or observed data tuple $\boldsymbol{X}$. In other words, we are looking forthe probability that tuple $\boldsymbol{X}$ belongs to class C, given that weknow the attribute description of $\boldsymbol{X}$. $P(H|\boldsymbol{X})$ is the posteriorprobability, or *a posteriori probability*, of H conditioned on$\boldsymbol{X}$.

Inthisway, Bayes' theoremadjuststheprobabilitiesasnewinformationonevidencesappears.

AccordingtoitsclassicalformulationgiventwoeventsAand B, the conditional probability

- $P(A|B)$thatAoccursifBoccurscanbeobtainedifweknow P(A),theprobabilitythat Aoccurs
- $P(B)$,theprobabilitythatBoccurs,
- $P(B|A)$ the conditional probability of B given

A,(Asshowninequation):

$$P(A|B)=\frac{P(B|A)P(A)}{P(B)}$$

3.1.2NaïveBayesClassifierforintrusionDetection

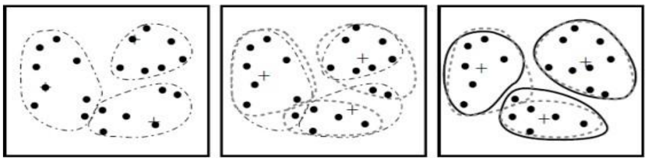
In Bayesian classification, we have a hypothesis that thegiven databelongs toa particularclass.We then calculatethe probability for the hypothesis to be true. This is amongthe most practical approaches for certain types of problems.Theapproachrequiresonlyonescanofthewholedata. Also, if at some stage there are additional training data, theneach training example can incrementally increase/decreasethe probability that a hypothesis is correct. Thus, a Bayesiannetworkisusedtomodeladomaincontaininguncertainty.

3.2 K-meansClustering

The k-means algorithm takes the input parameter,k, andpartitionsasetof*n*objectsinto*k*clusterssothattheresulting intra-cluster similarity is high but the inter-clustersimilarityislow.Clustersimilarityismeasuredinregardtt he*mean* value of the objects in a cluster, which can beviewed asthe cluster's *centroid*or *centerofgravity*.

Network intrusion class labels are divided into four mainclasses, which are DoS, Probe, U2R, and R2L. Fig. 1(a) toFig. 1(c) shows the steps involved in K-Means clusteringprocess. Fig.2 will later show the final overall result withapplication of the classification approach. The main goal toutilize K-Means clustering approach is to split and to groupdata intonormal and attack instances.K-Means clusteringmethods partition the input dataset into k- clusters accordingtoaninitialvalueknownastheseed-pointsintoeachcluster's centroids orclustercenters.Themeanvalueofnumericaldatacontainedwit hineachclusteriscalledcentroids. In our case, we choose k = 3 in order to cluster thedata into three clusters (C1, C2, C3). Since *U2R* and R2Lattackpatternsarenaturallyquitesimilarwithnormalinstanc es, one extra cluster is used to group U2R and R2Lattacks.

Back to Fig. 1(b), each input will be assigned to the closestCentroid by squared distances between the input data pointsand the centroids. New centroids will then be generated foreach cluster by calculating the mean values of the input setassigned to eachcluster asshown inFig.1(c).



(a) (b) (c)

Algorithm:*k*-means.The*k*-meansalgorithmforpartitioning,where each cluster's center is represented bythemeanvalueofthe objectsinthecluster.

Input:

k:thenumberofclusters,

D:adatasetcontaining*n* objects.

Output: Asetof*k*clusters.

Method:

(1) Arbitrarilychoose*k*objectsfrom*D*astheinitialclustercenters;
(2) Repeat
(3) (Re)assigneachobjecttotheclustertowhichtheobject is the most similar, based on the mean value oftheobjectsinthe cluster;
(4) Update the cluster means, i.e., calculate the mean valueofthe objectsfor eachcluster;
(5) Untilnochange;

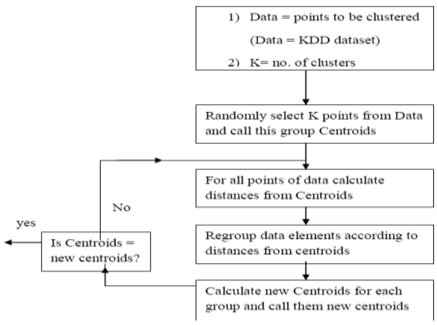


Figure3.2:blockdiagramforK-meansclustering

FUTURESCOPE

Inthefuture:

- WerecommendconsideringtheHybridIntrusionDetection System which is better at detecting R2L andU2Rattacks.
- The misuse detection approach better at detecting R2Land U2R attacks more efficiently as well as anomalydetectionapproach.
- Work for approach which isbetter at detecting attacksat the absence of match signatures as provided in themisuserule files.

Thecriticalnatureofthetaskofdetectingintrusionsinnetworks and applications leaves no margin for errors. Theeffective cost of a successful intrusion overshadows the costofdevelopingintrusiondetectionsystemsandhence,itbeco mes critical to identify the best possible approach fordevelopingbetterintrusiondetectionsystems.

Every network and application is custom designed and itbecomesextremelydifficulttodevelopasinglesolutionwhich can work for every network and application. In thisthesis,weproposednovelframeworksanddevelopedmethod swhichperformbetter.However,inordertoimprove the overall performance of our system we used thedomain knowledge for selecting better features for trainingour models. This is justified because of the critical nature ofthe task of intrusion detection. Using domain knowledge todevelopbettersystemsisnotasignificantdisadvantage;howev er, developing completely automatic systems presentsaninterestingdirection for futureresearch.

Thefieldofintrusiondetectionhasbeenaroundsince1980's and a lot of advancement has been made in the same.However, to keep pace with the rapid and ever changingnetworksandapplications,theresearchinintrusiondet ectionmustsynchronizewiththepresentnetworks.Present networks increasingly support wireless technologies,removable and mobile devices. Intrusion detection systemsmust integrate with such networks and devices and providesupportforadvancesinacomprehensiblemanner.

REFERENCES

[1] StefanAxelsson.ResearchinIntrusion-DetectionSystems:ASurvey.TechnicalReport98-17,DepartmentofComputerEngineering,ChalmersUniversityofTechnology, 1998.

[2] SANSInstitute - Intrusion Detection FAQ. Last accessed: August30,2012.http://www.sans.org/resources/idfaq/.

[3] KotagiriRamamohanarao,KapilKumarGupta,TaoPeng,andChristopher Leckie. The Curse of Ease of Access to the Internet. InProceedings of the 3 rd International Conference on InformationSystemsSecurity(ICISS),pages234–249.LectureNotesinComputerScience, SpringerVerlag, Vol(4812),2008.

[4] Overview ofAttackTrends, 2002. Last accessed: November 30,2008.http://www.cert.org/archive/pdf/attack_trends.pdf.

[5] Kapil Kumar Gupta, BaikunthNath, KotagiriRamamohanarao, andAshraf Kazi. Attacking Confidentiality: An Agent Based Approach.In Proceedings of IEEE International Conference on Intelligence andSecurity Informatics, pages 285–296. Lecture Notes in ComputerScience, SpringerVerlag, Vol(3975),2006.

[6] TheISCDomainSurvey.Lastaccessed:Novmeber30,2008.https://www.isc. org/solutions/survey/.

[7] PeterLyman,HalR.Varian,PeterCharles,NathanGood,LaheemLamarJordan,JoyojeetPal,andKirstenSwearingen.HowmuchInformation. Last accessed: Novmeber 30, 2008.http://www2.sims.berkeley.edu/research/projects/how-much-info-2003.

[8] AnimeshPatchaandJung-MinPark.AnOverviewofAnomalyDetection Techniques: Existing Solutions and Latest TechnologicalTrends.ComputerNetworks, 51(12):3448–3470,2007.

[9] CERT/CCStatistics.Lastaccessed:Novmeber30,2008.http://www.cert.org/stats/.

[10] Thomas A. Longstaff, James T. Ellis, Shawn V. Hernan, Howard F.Lipson,RobertD.Mcmillan,LindaHutzPesante,andDerekSimmel.SecurityoftheInternet.TechnicalReportTheFroehlich/KentEncyclopediaofTelecommunicationsVol(15),CERTCoordinationCenter1997.Lastaccessed:Novmeber30,2008.http://www.cert.org/encyc_article/tocencyc.html.

[11] KDD Cup 1999 Intrusion Detection Data. Last accessed: Novmeber30,2008.http: //kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.

[12] Kapil Kumar Gupta, BaikunthNath, and KotagiriRamamohanarao.ApplicatiobaseIntrusionDetectionDataset. Lastaccessed:Novmeber30,2008.http://www.csseunimelb.edu.au/˜kgupta.

[13] Stefan Axelsson. Intrusion Detection Systems: A Taxomomy andSurvey.TechnicalReport99-15,DepartmentofComputerEngineering,ChalmersUniversityofTechnology, 2000.

[14] Anita K. Jones and Robert S. Sielken. Computer System IntrusionDetection: A Survey.Technical report,Departmentof ComputerScience, University of Virginia, 1999. Last accessed:Novmeber 30,2008. http://www.cs.virginia.edu/˜jones/IDS-research/Documents/jones-sielken-survey-v11.pdf.

[15] PeymanKabiriandAliA.Ghorbani.ResearchonIntrusionDetectionand Response:ASurvey.*InternationalJournalofNetworkSecurity*,1(2):84–102,2005.

[16] Joseph S. Sherif and Tommy G. Dearmond. Intrusion Detection:SystemsandModels.In*ProceedingsoftheEleventhIEEEInternational Workshops on Enabling Technologies: InfrastructureforCollaborativeEnterprises.WETICE*,pages115–133.IEEE,2002.

[17] MikkoT.SiponenandHarriOinas-Kukkonen.AReviewofInformation Security Issuesand Respective Research Contributions.*SIGMISDatabase*,38(1):60–80,2007. ACM.

[18] TeresaF.Lunt.Asurveyofintrusiondetectiontechniques.*Computers and Security*, 12(4):405–418, 1993. Elsevier AdvancedTechnologyPublications.

[19] Emilie Lundin and ErlandJonsson. Survey of Intrusion DetectionResearch.TechnicalReport02-04,DepartmentofComputerEngineering,ChalmersUniversityofTechnology, 2002.

[20] JamesP.Anderson.ComputerSecurityThreatMonitoringandSurveillance,1980.Lastaccessed:Novmeber30,2011.http://csrc.nist.gov/publications/history/ande80.pdf.

[21] DorothyE.Denning.AnIntrusion-DetectionModel.*IEEETransactionsonSoftwareEngineering*,13(2):222–232,1987.IEEE.

[22] H.S.JavitzandA.Valdes.TheSRI IDESStatisticalAnomalyDetector.In*Proceedingsof*

[23] *theIEEESymposiumonSecurityandPrivacy*,pages316–326.IEEE, 1991.

[24] S.E.Smaha.Haystack:AnIntrusionDetectionSystem.In*Proceedings of the 4th Aerospace Computer Security ApplicationsConference*,pages37–44.IEEE, 1988.

[25] Paul Innella. The Evolution of Intrusion Detection Systems, 2001.Last accessed: Novmeber 30, 2008. http://www.securityfocus.com/infocus/1514.