A Machine Learning Intrusion Prevention and Detection System using Securing Smart Grid

N. Raghavendra Sai, Akkili Guru Raghavendra, Nerella Chandra Mouli Deepak, M.Poojitha

Abstract: ***Ordinary control arrange abilities have improved and expanded by the insightful framework however because of indistinguishable time making it further inclined to contrasting kinds of assaults. These vulnerabilities permit partner degree transgressor to breakdown trustworthiness and classification and enable access to the system. insignificant decisions in information have caused semi changeless downside in arrange traffic Classification. These decisions obstruct the technique for grouping and hinder classifier from settling on right choice, especially once tending to enormous data. An IDS ,named Least sq. Bolster Vector Machine based IDS,Is fabricated example alternatives} picked by our arranged Feature decision recipe. The exhibition of Least sq. Bolster Vector Machine assessed abuse a couple of interruption identification examination datasets, explicitly KDD Cup ninety ninemsn-KDD dataset. The examination Results show that our element choice recipe contributes further important decisions for Least sq. Bolster Vector Machine to comprehend higher precision and lower calculation worth contrasted and dynamic ways that.***

Key Words: Intrusion Detection System (IDS), Least Square Support Vector Machine(LSSVM)

I. INTRODUCTION

System security has slash hack become one among the first squeezing issues and of worry for web clients and fix providers with a proceeding with development in net exercise (Medaglia and Serbanati, 2010). A protected system are frequently described regarding its equipment and PC code insusceptibility against very surprising interruptions. A system are regularly verified by fusing solid keen, assessment and protect methods. System Intrusion Detection System (NIDS) consolidates these strategies to guard against organize interruptions (Debar, Dacier and Wespi, 1999). These barrier frameworks perform consistent watching of system traffic, dissect and report any interruptions. the key components of this method epitomize traffic gatherer, investigation motor, signature data and caution stockpiling, as appeared in Figure one.

Manuscript published on January 30, 2020.
* Correspondence Author

Dr.N. Raghavendra Sai*, Assoc.Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India.
Email: nallagatlaraghavendra@kluniversity.in.

Akkili GuruRaghavendra, B.Tech IV Year CSE Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India.akkiliraghu007@gmail.com

Nerella Chandra Mouli Deepak, B.Tech IV Year CSE Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India. deepaknerella99@gmail.com

M.Poojitha, B.Tech IV Year CSE Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India.

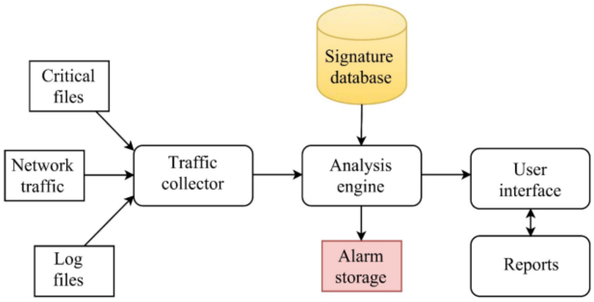


Figure 1: Components of IDS

Every component assumes a significant job in interruption location. System traffic is caught by the traffic authority, that is, parcel follows, examination motor leads a profound investigation of the caught traffic information and sends caution sign to alert stockpiling once interruption is recognized. The mark data stores the marks or examples of known gatecrashers, and these marks zone unit utilized for coordinating reason. A commonplace NIDS is outlined in Figure 2.

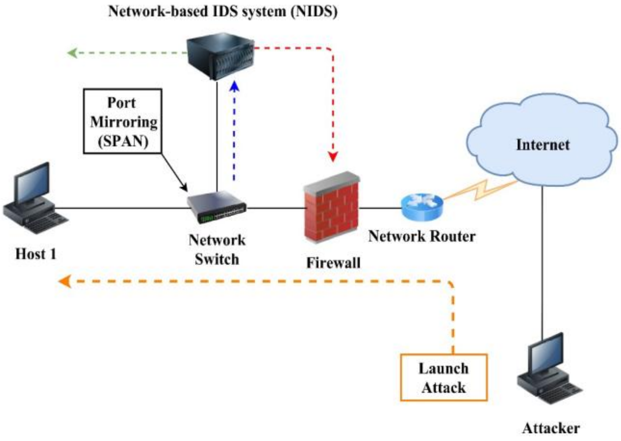


Figure 2: ILLUSTRATION OF IDS

An Intrusion Detection System(IDS) screens framework traffic and screens for suspicious activity and cautions the structure or systemoverseer .Now but again the IDS would potentially moreover respond to odd or vindictive traffic by taking activities. for example, discouraging the benefactor or offer logical order address from going into the framework. IDS is that the cutting edge like Associate in Nursing caution sorted out to see passageways, threatening exercises, and known interlopers. DS likely could be a location framework that is aware of how to peruse and translate the substance of log records or libraries from switches ,firewalls ,servers and diverse system gadgets.

Partner in Nursing IDS for the most part stores a data of known assault marks and will think about examples of activity,trafficor conduct it sees inside the logs or libraries it's keen against those imprints to get a handle on once a near to coordinate between a mark and blessing or late lead happens..At that point, the IDS can give cautions or alarms, make changed sorts of programmed move ragging from detaching of net connections or explicit servers to propelling back follows ,and fabricate diverse dynamic makes a shot to recognize assailants and action gather evidence of their exercises .Often, IDS bundle runs on consistent framework or servers wherefirewalls, intermediaries or distinctive limit administration work; Associate in Nursing IDS not running on steady gadget or server where the firewall or various administrations square measure place in screens those techniques intently and carefully. albeit such gadgets will in general oversee at organize outskirts, IDS's can sight and impact inside the framework assaults furthermore as outer assaults.

II. METHODOLOGY

In the Methodology we have done a half and half component elective calculation (HFSA) design Least sq. Bolster Vector Machine's comprises of two segments: the upper part directs a fundamental inquiry to dispose of digressive and repetition decisions from the underlying date. This helps the wrapper method(lower stage) to diminish the needing change from the whole unique element space to the pre-chosen features(the yield of the upper phase).This work proposes a novel channel based element elective system, all through that thermotical investigation of shared information is acquainted with guage the reliance among decisions and yield classes. we will in general will in general lead total examinations on two standard datasets moreover to the dataset utilized. this will be significant in assessing the exhibition of IDS since KDD dataset is dead and doesn't contain most novel assault designs in it. These datasets oft utilized inside the writing to guage the presentation of Intrusion Detection System(IDS).Moreover, these datasets have tests in changed sizes and completely entirely unexpected quantities of featurisms they supply inexhaustible further challenges for completely testing feature assurance computations. about a comparative because of the ID structure guided that plans only for twofold portrayal, we will in general will in general improve our anticipated framework to contemplate multiclass request gives this will be to implies the adequacy and moreover the utility of the anticipated strategy.

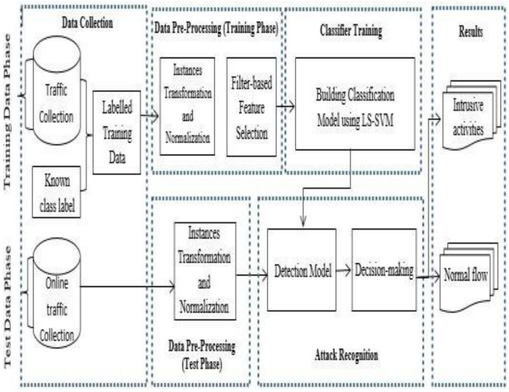

Training Data Phase
Test Data Phase
Data Collection
Traffic Collection
Known class label
Labelled Training Data
Online traffic Collection
Data Pre-Processing (Training Phase)
Instances Transformation and Normalization
Filter-based Feature Selection
Classifier Training
Building Classification Model using LS-SVM
Results
Intrusive activities
Instances Transformation and Normalization
Data Pre-Processing (Test Phase)
Detection Model
Decision-making
Attack Recognition
Normal flow

Figure 3 : Methodology.

We will isolate the whole information into train information segment and take a look at information part. There square measure completely very surprising modules all through this style. They are:

a) information pre-handling

b) Filter based thoroughly Feature elective

c) Attack order and Recognition. U2R(User To Root Attack) R2L(Remote To local Attack) inquisitory

d) Performance examination

a) information pre-handling: the information got all through the area of information combination square measure initially prepared to return up with the essential decisions like those inThe KDD Cup 9ty nine dataset is one in every one of the chief overflowing and complete interference disclosure datasets and is for the most part applied to evaluate the presentation of interference recognizable proof frameworks..Thus, each representative decisions grasp the kind of convention, administration kind and convention standing banner. the strategy just replaces the estimations of the exact qualities with numeric qualities. A basic advance of information pre-handling once moving every single representative property into numerical esteems in institutionalization. information institutionalization can be a method for scaling the estimation of each trait into a Junoesque shift, that the predisposition for decisions with bigger qualities is disposed of from the dataset.

b) Filterbased Feature Selection: If one considers relationships between's system traffic records to be straight affiliations, at that point a live of dependencesuch as direct measurement could likewise be acclimated live the reliance between two irregular factors. Notwithstanding, considering the $64000 world correspondence, the connection between's factors could likewise be nonlinear moreover. Obviously, a live can't uncover the connection between two nonlinearly subordinate factors. on these lines, we'd kind of a live prepared for inspecting the alliance between two factors regardless of whether or not they're straightforwardly or nonlinearly needy.. Hence, this work means to investigate how of choosing ideal element space disregarding the kind of connection between's them. we tend to create two calculations for highlight elective procedure. These are: adaptable common information based absolutely include elective and have elective upheld straight measurement.

c)Attack Classification and Recognition;: by and large, it's simpler to make a classifier to separate between two classes than considering multi classifications in partner amazingly move back. is this is normally this could be} as a consequences of the choice limits at interims the main case are regularly simpler. the initial an area of the examination all through this paper utilizes two classes, where records coordinating to traditional|the standard} class square measure reputed as ordinary information, than two classes, there square measure two popular systems: One-Vs-One(OVO),One-Vs-All(OVA). Subsequent to completing all the said advances and moreover the classifier is prepared practice the ideal arrangement of decisions which incorporates the principal related with and significant choices, the standard and interruption deals could likewise be well known by training the spared prepared model to locate interruptions. Records coordinating to traditional|the standard} class square measure pondered as ordinary information, and furthermore the decision records square measure reputed as assaults. In the event that the classifier model affirms that the record is strange, the organic gathering of the

unusual record(type of assaults) could likewise be acclimated check the record kind.

c.1)U2R(User To Root Attack):U2R likely could be a class of movement that the awful individual beginnings with access to a run of the mill client account on the system(perhaps picked up by sniffing passwords, a book of certainties assault, or social building) and is in a very position to require bit of leeway of some defenselessness to accomplish root access to the framework.

c.2) as a client of that machine.

c.3)Probing: R2L(Remote To local Attack):R2L happens once partner terrible individual authoritative unit can send bundles to a machine exercises some defenselessness to accomplish local access inquisitor is choose to assemble data many system of PCs for the obvious reason for staying away from its security controls Attacks include fundamental three classifications:

U2R: unapproved access from an inaccessible machine, e.g., estimation mystery.

R2L:unauthorized access to local super client (root) benefits, E.g., different "cradle flood" assaults.

Examining: exploring and different inquisitor. E.g., port filtering.

d)Performance Evaluation: most of the IDS tests were performed on third Cup 9ty nine datasets. furthermore, these data sets have absolutely totally unique data sizes and differing numbers field-grade official decisions which give far reaching tests in affirming highlight elective manners that.

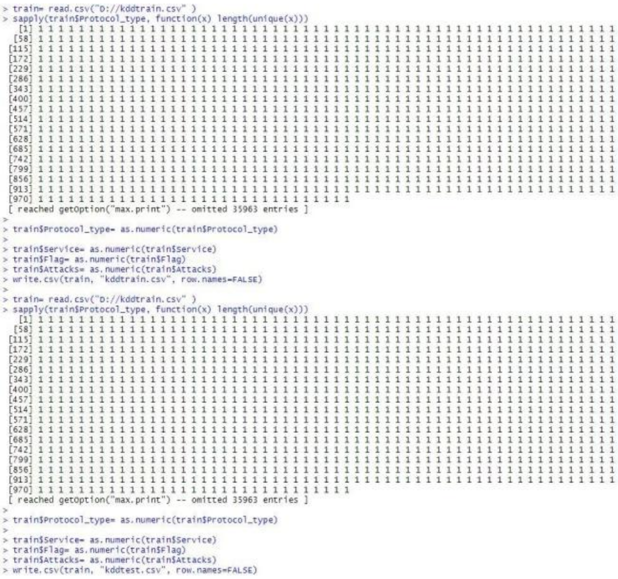
The KDD Cup 9ty nine dataset is one in every one of the premier popular and far reaching interruption recognition datasets and is wide applied to assess the presentation of interruption location frameworks. It comprises of five absolutely totally various classes that ar old and three styles of assaults (I . e . ., U2R , R2L, Probing).It contains business data with extra or less hardly any million connection records and take a look at data with concerning 2,000,000 association records. each record in these datasets is labeled as either old or partner assault, and it's forty totally

unique quantitative and subjective decisions. numerous tests ar semiconductor diode to evaluate the presentation and sufficiency of the anticipated LSSVM IDS ..For this reason, the precision rate, bogus positive rate and F-estimated measurements applied.

The significant note that the take a look at data isn't from consistent probability of dispersion because of the business data, it grasp explicit assault kind not among the work information. This makes the assignments extra right. Some interruption authorities accept that pretty much all novel assaults ar variation of known assaults and conjointly the mark of better-realized assaults could likewise be cozy to get novel variations. The datasets contains aggregate of twenty four distinct kinds of assaults, with an extra fourteen assortments among the take a look at information in particular.

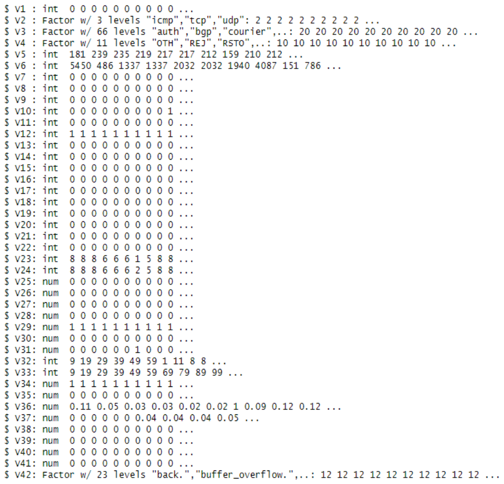
III. RESULTS

1)Creating Training & Testing Data:



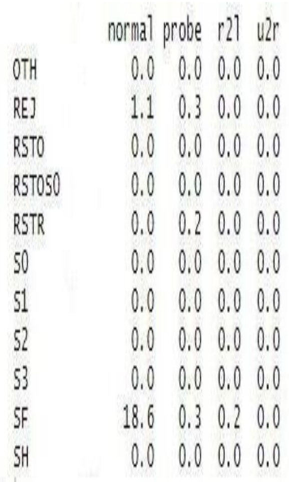
```
> train= read.csv("D://kddtrain.csv" )
> sapply(train$Protocol_type, function(x) length(unique(x)))
   [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
  [58] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [115] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [172] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [229] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [286] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [343] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [400] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [457] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [514] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [571] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [628] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [685] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [742] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [799] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [856] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [913] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [970] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [ reached getOption("max.print") -- omitted 35963 entries ]
>
> train$Protocol_type= as.numeric(train$Protocol_type)
>
> train$Service= as.numeric(train$Service)
> train$Flag= as.numeric(train$Flag)
> train$Attacks= as.numeric(train$Attacks)
> write.csv(train, "kddtrain.csv", row.names=FALSE)
>
> train= read.csv("D://kddtrain.csv" )
> sapply(train$Protocol_type, function(x) length(unique(x)))
   [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
  [58] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [115] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [172] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [229] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [286] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [343] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [400] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [457] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [514] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [571] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [628] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [685] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [742] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [799] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [856] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [913] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [970] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [ reached getOption("max.print") -- omitted 35963 entries ]
>
> train$Protocol_type= as.numeric(train$Protocol_type)
>
> train$Service= as.numeric(train$Service)
> train$Flag= as.numeric(train$Flag)
> train$Attacks= as.numeric(train$Attacks)
> write.csv(train, "kddtest.csv", row.names=FALSE)
```

2) Pre-Processing of Data :



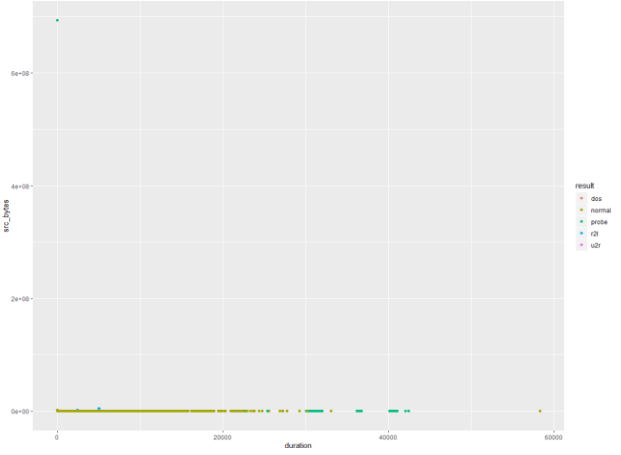
```
$ V1 : int  0 0 0 0 0 0 0 0 0 0 ...
$ V2 : Factor w/ 3 levels "icmp","tcp","udp": 2 2 2 2 2 2 2 2 2 2 ...
$ V3 : Factor w/ 66 levels "auth","bgp","courier",..: 20 20 20 20 20 20 20 20 20 20 ...
$ V4 : Factor w/ 11 levels "OTH","REJ","RSTO",..: 10 10 10 10 10 10 10 10 10 10 ...
$ V5 : int  181 239 235 219 217 217 212 159 210 212 ...
$ V6 : int  5450 486 1337 1337 2032 2032 1940 4087 151 786 ...
$ V7 : int  0 0 0 0 0 0 0 0 0 0 ...
$ V8 : int  0 0 0 0 0 0 0 0 0 0 ...
$ V9 : int  0 0 0 0 0 0 0 0 0 0 ...
$ V10: int  0 0 0 0 0 0 0 0 0 1 ...
$ V11: int  0 0 0 0 0 0 0 0 0 0 ...
$ V12: int  1 1 1 1 1 1 1 1 1 1 ...
$ V13: int  0 0 0 0 0 0 0 0 0 0 ...
$ V14: int  0 0 0 0 0 0 0 0 0 0 ...
$ V15: int  0 0 0 0 0 0 0 0 0 0 ...
$ V16: int  0 0 0 0 0 0 0 0 0 0 ...
$ V17: int  0 0 0 0 0 0 0 0 0 0 ...
$ V18: int  0 0 0 0 0 0 0 0 0 0 ...
$ V19: int  0 0 0 0 0 0 0 0 0 0 ...
$ V20: int  0 0 0 0 0 0 0 0 0 0 ...
$ V21: int  0 0 0 0 0 0 0 0 0 0 ...
$ V22: int  0 0 0 0 0 0 0 0 0 0 ...
$ V23: int  8 8 8 6 6 6 1 5 8 8 ...
$ V24: int  8 8 8 6 6 6 2 5 8 8 ...
$ V25: num  0 0 0 0 0 0 0 0 0 0 ...
$ V26: num  0 0 0 0 0 0 0 0 0 0 ...
$ V27: num  0 0 0 0 0 0 0 0 0 0 ...
$ V28: num  0 0 0 0 0 0 0 0 0 0 ...
$ V29: num  1 1 1 1 1 1 1 1 1 1 ...
$ V30: num  0 0 0 0 0 0 0 0 0 0 ...
$ V31: num  0 0 0 0 0 0 1 0 0 0 ...
$ V32: int  9 19 29 39 49 59 1 11 8 8 ...
$ V33: int  9 19 29 39 49 59 69 79 89 99 ...
$ V34: num  1 1 1 1 1 1 1 1 1 1 ...
$ V35: num  0 0 0 0 0 0 0 0 0 0 ...
$ V36: num  0.11 0.05 0.03 0.03 0.02 0.02 1 0.09 0.12 0.12 ...
$ V37: num  0 0 0 0 0 0 0.04 0.04 0.04 0.05 ...
$ V38: num  0 0 0 0 0 0 0 0 0 0 ...
$ V39: num  0 0 0 0 0 0 0 0 0 0 ...
$ V40: num  0 0 0 0 0 0 0 0 0 0 ...
$ V41: num  0 0 0 0 0 0 0 0 0 0 ...
$ V42: Factor w/ 23 levels "back.","buffer_overflow.",..: 12 12 12 12 12 12 12 12 12 12 ...
```

3)Attack classification and recognition:



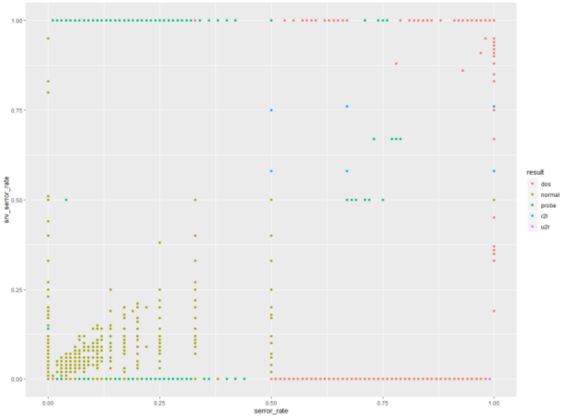
	normal	probe	r2l	u2r
OTH	0.0	0.0	0.0	0.0
REJ	1.1	0.3	0.0	0.0
RSTO	0.0	0.0	0.0	0.0
RSTOS0	0.0	0.0	0.0	0.0
RSTR	0.0	0.2	0.0	0.0
S0	0.0	0.0	0.0	0.0
S1	0.0	0.0	0.0	0.0
S2	0.0	0.0	0.0	0.0
S3	0.0	0.0	0.0	0.0
SF	18.6	0.3	0.2	0.0
SH	0.0	0.0	0.0	0.0

4)Source port rate has some has slight effect on Intrusion Type.



As the duration increases the data from source to destination doesn't change but here we can find attacks when time changes. Initially at higher data we can see remote to local (R2L) attack. Simultaneously as the time or duration increase the data is getting intruded mostly by remote to local attack. In this type of attack the attacker takes advantages over the bugs or weakness of the system using the network.

5)Server rate has some slight information of intrusion type.



In this graph we can observe synchronization errors in the graph. There is probe attack at the peak stage of service and as the host error rate of increased it changes to DoS.It is normal at lower stages of the service error rate and increases it changes and changes as the host error rate increase the data seems to be attacked at lower service error by probe and normal when service error rate is increasing. As the host error rate alone increases the type of attack changes from probe to Dos and at the peak stages of both service error rate and host error rate.

IV. CONCLUSION AND FUTURE SCOPE

Intrusion Detection System(IDS) is used to find the different types of attacks that takes place in the network or computer applications. Now-a-days networks are easily vulnerable to the attackers and finding the loop holes to access data. If we use better detection system we can easily detect the loop holes and can be aware of these attackers. In this we have used LSSVM for detecting the intrusions and it gives output more accurately than SVM and within a short period of time. Now-a-days all the processes are being done online and the usage of internet is being increased in coming generations and attackers will try for other loop holes in order to access the private data, this detection system will be helpful in order to secure your data from the attackers by finding the intrusion that are being on your network and the system.

REFERENCES

1. Y.Li,C.Zhang, L.Yang,"The Research of AMI Intrusion Detection Method using ELM in Smart Grid," in International Journal of security and its applications,Vol 10,No.5 2016,pp.283-296.
2. A.Arvani and V.S Rao,"Detection and Protection Against Intrusions on Smart Grid Systems," in International Journal of Cyber-Security and Digital Forensics Vol 3,No 1,2016,pp,. 38-48
3. F.Aloul,A.R. Al-Ali,R.Al-Dalky,M.Al-Mardin,W.El-Haji"Smart Grid Security: Threats, Vulnerabilities and Solutions" International Journal of Smart Grid and clean Energy Vol1,no. 1,2012

4. R.K Sharma,H.K.Kalita P.Borah "Analysis of Machine Learning Techniques Based Intrusion Detection Systems" Proceedingd of 3rd International Conference on Advanced Computing,Networking and Informatics,2015,pp.485-493
5. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. ACM SIGKDD Explorations Newsletter, 11(1), 10-18.
6. Hand, D. J. (2009). Measuring classifier performance: A coherent alternative to the area under the ROC curve. Machine Learning, 77(1), 103-123.
7. Hasan, M. A. M., Nasser, M., Ahmad, S., & Molla, K. I. (2016). Feature selection for intrusion detection using random forest. Journal of Information Security, 7(03), 129-140.

AUTHORS PROFILE

Dr. N. Raghavendra Sai currently working as Assoc.Professor in the Department of Computer Science and Engineering at KL University. He has 12 years of Teaching Experience. He did his Ph.D from Bharathiar University. He has participated in various International Conferences and workshops held at different places. His area of interest includes web mining,Netwrok Security,Information retrieval and social network analysis, semantic analysis.

A. Guru Raghavendra Studying B.Tech (CSE) at KL University and his basic Interset of research area is Network Security and his hobbies are playing cricket and reading books.

Nerella Chandra Mouli Deepak Studying B.Tech (CSE) at KL University and his basic Interset of research area is Network Security and his hobbies are playing cricket and listening music.

M. Poojitha Studying B.Tech (CSE) at KL University and his basic Interset of research area is Network Security and his hobbies are playing chess and reading books.