8th Sem Project: Intrusion Detection System

Pratyay Amrit <p.amrit@live.com> Thu, Jan 11, 2018 at 11:42 AM
To: "Ipsita Upasana [MAHE-MIT]" <ipsita.upasana@manipal.edu>

Ma'am,

This is my **weekly report for tomorrow, Friday, 12th January, 2018.**

We have not received a format for this report yet, so I will be writing everything in this mail directly.

I have narrowed down my search to 3 papers:

[1] "Using Data Mining Algorithms for Developing a Model for Intrusion Detection System (IDS)" (2015)
[2] "IDS Using Bagging with Partial Decision Tree Based Classifier" (2015)
[3] "Multilayer Perceptron with Binary Weights and Activations for Intrusion Detection of Cyber-Physical Systems" (2017)

Each of these are based on different machine learning and data mining methods of building an IDS.

[1] and [2] focus on a generic device, while [3] focuses more on low powered devices. I believe, with how development is going in the field of WSN and IoT, focusing on low powered devices is important.

I have gone through the papers in some detail, and on the basis of knowledge that I already have, I believe the difficulty of implementing will be in the order [1] < [3] < [2].

Each paper has a few drawbacks associated with them:
[1] has the classic problem of selecting the right number of clusters in k-means.
[2] has the problem that it requires a lot of time to build the model.
[3] has the problem that the error rate increases by 4 times when the weights are made binary to make the model implementable in low powered devices.

We have two options,
1. We can start implementing and comparing these models one by one.
2. We can try to fix an issue associated with any one of these papers.

As of now, I have no idea how a problem may be fixed. It's hard to come up with something new unless I have implemented these at least once. I would like to change my topic, if possible, in case I figure out a way to fix a problem during the implementation.

I do believe that all three of these papers can be implemented within 4 month time.

While implementing these, I would also like to try to scale the algorithms down for low powered devices as done in [3] for methods used in [1] and [2].

Acquiring data for training and testing the machine learning models will not be a problem during implementation as most papers on IDS use the same data set which is publicly available at http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.

About the report,

Sanjay Singh sir has not sent the format for the weekly report, so I believe getting a physical signature is not mandatory yet. I believe this is because the majority of students have not chosen their guides.

I request you to go through this report when you find the time and let me know if the meeting due tomorrow is necessary.

I will start writing the synopsis for the implementation and comparative study of these methods and submit a soft copy to you before the actual submission as soon as possible.

Thanks,
Pratyay Amrit
140953430
ICT Dept., MIT
Manipal