DETECTING ANOMALOUS INSIDERS IN COLLABORATIVE INFORMATION SYSTEMS

MINI PROJECT

The company providing us the opportunity to work with them to complete the mini project!

Team Members:

- A.S.S.BHARADWAZA(11T81A0525)
- T.NAVEEN KUMAR(11T81A0517)
- P.SAI ARUN KUMAR(11T81A0528)

Under the guidance of

Mr. Ravinder Reddy, Asst.professor, Dept.CSE

Department of Computer Science and Engineering

ANNAMACHARYA INSTITUTE OF TECHNOLOGY AND SCIENCES

(Affiliated to JNT University, Hyderabad)

Piglipur, Batasingaram (V), Hayathnagar (M), R.R.Dist, Hyderabad-501512.

2014-2015.

21st century software solutions here mentioned, Detecting Anomalous Insiders in Collaborative Information Systems, project in detail as an abstract, which includes what is actually involved in the entire mini project to be done by A.S.S.BHARADWAZA and his team members under guidance of Raghava Rao (Associate Developer) in the company premises

We are taking responsibility to train them in all basic requirements in general use to develop the project

The project proposal is as follows

Initially we will record (EHR) system in a large medical center. The results illustrate our models exhibit significant performance gains over state-of-the-art competitors. When the number of Collaborative information systems (CISs) are deployed within a diverse array of environments that manage sensitive information. Current security mechanisms detect insider threats, but they are ill-suited to monitor systems in which users function in dynamic teams. In this paper, we introduce the community anomaly detection system (CADS), an unsupervised learning framework to detect insider threats based on the access logs of collaborative environments. The framework is based on the observation that typical CIS users tend to form community structures based on the subjects accessed (e.g., patients' records viewed by healthcare providers).
CADS consists of two components: 1) relational pattern extraction, which derives community structures and 2) anomaly prediction, which leverages a statistical model to determine when users have sufficiently deviated from communities? We further extend CADS into MetaCADS to account for the semantics of subjects (e.g., patients' diagnoses). To empirically evaluate the framework, we perform an assessment with three months of access logs from a real electronic health illicit users is low, MetaCADS is the best model, but as the number grows, commonly accessed semantics lead to hiding in a crowd, such that CADS is more prudent.

Existing System:

It can be seen that the performance of the supervised classification models is significantly worse than the unsupervised models. The supervised models consistently have a lower true positive rate at all operating points. Second, unlike the previous experiment, HVU achieves comparable results to the supervised classification models. This is due to the fact that this model is correctly characterizing the intruders that access a larger number of records. Third, with respect to AUC, we observe the same trend as earlier regarding the dominance of the unsupervised models as a function of the mix rate. Specifically,

How our proposal overcomes the limitations of existing

MetaCADS dominates when the mix rate is low, but CADS dominates when the mix rate is high. Notably the disparity between MetaCADS and CADS is more pronounced at the low mix rate (0.91 versus 0.88) in this setting than in the previous setting. However, at lower false positive operating points, CADS appears to dominate MetaCADS.

Proposed system

Several notable approaches have been proposed to address this type of intruder. The first is nearest neighbor anomaly detection techniques, which are designed to measure the distances between instances by assessing their relationship to “close”
Instances. If the instance is not sufficiently close, then it may be classified as an anomaly. However, social structures in a CIS are not explicitly defined and need to be inferred from the utilization of system resources. If distance measurement procedures are not tuned to the way in which social structures have been constructed, the distances will not represent the structures well. Our experimental results confirm this notion.

Modules:

- Pattern Extraction
- Anomaly Detection
- Detection Performance Metrics
- Varying Number of Accessed Subjects

Software Requirements:

- OS : Windows XP with SP2
- Database : MS-SQL server 2008
- Language : C#.NET
- IDE : Visual Studio .Net 2010
- Browser : IE

Hardware requirements

- Processor :Pentium
- HDD :20 GB Min 40 GB Recommended
- RAM :1 GB Min 2 GB Recommended