International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

Fraud Analytics for Wireless Sensor Networks

V. N. Niranjan[1], K. V. D. Kiran[2]

[1]M.Tech (CNS) Student, KL University, Guntur, Andhra Pradesh, India

[2]Assistant Professor, KL University, Guntur, Andhra Pradesh, India

Abstract: ***Wireless Sensor Networks (WSN) are deployed at locations where continuous monitoring is required. Physical protection for these sensor devices is hectic task and this is an opportunity for establishment of fraudulent connections by fraudsters and delivering fault responses. This badly affects the monitoring system and as a result the chances of false alarm are increased which badly degrades the performance of system. To avoid these fault alarms, fraud analytics is applied in this paper to differentiate actual monitored values and fault values. The main intention of this paper is to develop a Hadoop based fraud analytics mechanism to detect faults in Wireless Sensor Network reporting system and which helps in improving the performance of the monitoring system.***

Keywords: Wireless Sensor Networks (WSN), Fraud Analytics, Hadoop, Map-Reduce, HDFS.

1. Introduction

Wireless Sensor Networks are more vulnerable for attacks from fraudsters. This paper deals with the introduction of Hadoop based fraud analytics mechanism to detect faults in wireless sensor networks and improve performance of the monitoring system. Fraud Analytics is used to achieve mainly three goals. Firstly, Increasing Operational Efficiency and Productivity is achieved through Data Pre-Processing, Data Aggregation and Infrastructure Management. Secondly, Decreasing Risk through detection, prevention and prediction of fraud using analytic models from previous data and trend of fault reports for a certain period of time. Finally, the main aim of any business, Increasing Income and Profits is achieved by providing attractive representation of results through effective data visualization techniques and Automation and Integration of different types of analytic models into a single mechanism. The overall advantages and improvements in a system achieved by fraud analytics can be observed in detail as shown in below figure:

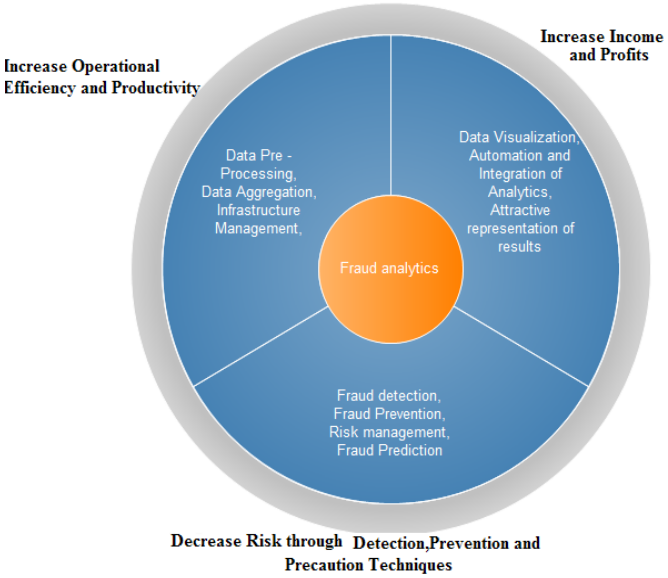

Figure 1: Overview of Fraud Analytics

2. Survey on Scope of Fraud Analytics

As per the article published in "Markets And Markets" in January 2014, interesting facts have been revealed regarding Fraud Analytics. The article explains that , fraud detection and prevention market with an in depth analysis will grow to worth of $7.5 Billion by 2018.Fraud Analytics is in great demand for defense, Banking and Financial Services, Insurance Companies, Telecommunication, public sector and many more industries which are frequently targeted by fraudsters and hackers.

3. Purpose

The purpose of this paper is to investigate the feasibility of fraud analytics with Wireless Sensor Networks and hence improving the performance and efficiency of the wireless sensor monitoring system .Future scope of research is also explained with clear methodology in this paper.

4. Related Work

There are several analytic models and most frequently used analytics techniques are Real-Time analytics and fraud analytics. As the data collected from huge number of wireless Sensor Nodes is heavy it is not easy to analyze total data and finding out the fault values and malfunctioning nodes is almost a very difficult task. To reduce the complexity of fraud detection in Big-data, fraud analytics is performed using Hadoop. The data collected from all nodes in wireless Sensor Networks is processed in Hadoop mechanism in which several Transformations of data is carried out with the help of Hadoop Distributed File System (HDFS) and Map-Reduce code written in java or Hive or Pig.

5. Development Environment

The development environment consist a Hadoop Cluster with many nodes working in a distributed manner to process the huge data that is coming from numerous sensor nodes. Flume Collector is used to collect the data from wire-less sensor networks and processed to Hadoop Cluster. Hadoop

Distributed File System (HDFS) which is on top of Hadoop distributes the data among various nodes. HDFS is a distributed file system that runs on huge clusters which are made up of many individual machines. Map- Reduce tasks can be carried out by writing map-reduce code in java or hive or pig. Data transformations are carried out in a rapid speed in Hadoop Clusters and results can be stored in HBase. HBase is a column-oriented database. HBase uses HDFS for its storing and supports computations using Map-Reduce and queries. Data can be accessed using HiveQL or Pig scripts.

Several analytical procedures are performed to achieve the task of obtaining results of fraudulent data and these results can be visualized using Data visualization tools. The data processing from Wire-less sensor networks to Data visualization is shown in figure below:

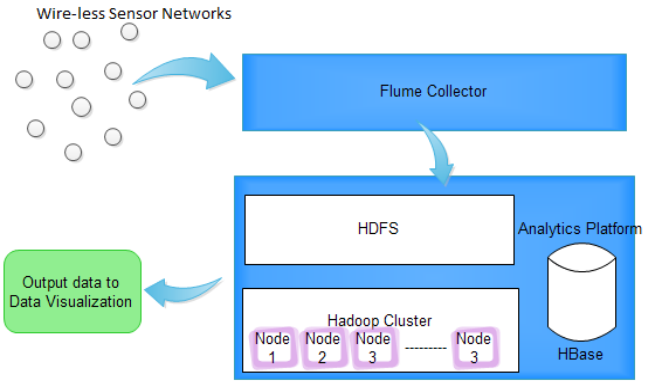

Figure 2: Data Processing from Collection to Visualization

6. Conclusion

This paper discusses about the process involved in implementing fraud analytics using Hadoop for huge data that is regularly collected from Wire-less sensor Networks. More over the importance of fraud analytics and its advantages are discussed in brief. Spark which is 100 times faster than Hadoop Map-Reduce can be used in future for increasing speed of computation.

7. Future Scope

Further developments in this paper include many areas. Two major developments can be automating system with Real-Time Analytics which can perform fraud analytics mechanism and a Smart Phone Application for Analytics which can generate graphs for data visualization.

7.1 Automating Monitoring System with Analytics

Upgrading the monitoring system with Analytics is major research area which includes an automation system with analytics engine which can perform multiple analytics like fraud analytics and real-time analytics. The system must be capable of deciding the actions needed to be performed when a fraudster attacks one or many nodes in Wire-less Sensor Network and eliminate fault alarms that are caused by system internal issues. This can be achieved by applying multiple Machine Learning algorithms.

7.2 Smart Phone Application for Fraud Analytics

Upgrading the monitoring system compatible with smart phone applications and providing data visualization in smarter way would be an interesting area of research which facilitates the operator to view, visualize, analyze and even control the system more effectively.

References

[1] http://en.wikipedia.org/wiki/Apache_Hadoop
[2] http://www.marketsandmarkets.com/Market-Reports/fraud-detection-prevention-market-1312.html
[3] Hadoop: The Definitive Guide, Second Edition by Tom White, Published by O'Reilly Media,Inc.
[4] http://hadoop.apache.org
[5] http://en.wikipedia.org/wiki/Apache_Spark

Author Profile

V. N. Niranjan is studying M.Tech (Computer Networks and Security) in KL University. He is working as Project Trainee at Honeywell Technology Solutions.

Mr. Venkata Durga Kiran.Kasula did M.Tech(CSE) from Archarya Nagarjuna University and pursuing Ph.D in risk assessment and security in Distributed systems from the same university. He has published thirteen international journal papers and four international conferences. He is presently working as a Assistant Professor at KL University, Vijayawada