

CORPORATE EMPLOYEE ATTRITION

THIRISA T.V , PRIYA .C , SARUMATHY .M , PRIYADHARSHINI.V

Abstract: The airport codes may refer to either the IATA airport code, a three-letter code that is used in passenger reservation, ticketing and baggage-handling systems, or the ICAO airport code which is a four-letter code used by ATC systems and for airports that do not have an IATA airport code. To provide better Airline and AirPort services and to avoid delays in Air Travel across different locations at Municipality level. The aim is to provide airports, airlines, and the travelling public with a neutral, third-party view of which airlines are delivering on their promise to get passengers from Point A to Point B on-time. For that purpose we use the method of data analytics and internet of things to store the data which is used for the needed. Data set of various countries , Number of flights based on continent , weather , region , Air station , Number of airports that are suitable for the running of the flights are being collected in the form of the data sets and penned down in charts . In IBM Cognos cloud area the collected information is stored for further process. On following the collection of the data set those processed information are dropped in the spreadsheets.

Using the given dataset, we plan to create various graphs and charts to highlight the insights and visualizations. On doing this a perfect information of the flights are obtained and many overhead delay of the arrival , departure of the planes can be avoided.

Keywords: Data analytics , Aviation , Data Set

INTRODUCTION

There is an interest in large scale data analytics in the aviation domain for analysis and prediction of capacity and flow in the US NAS. This is facilitated by data analytics systems in addition to availability of massive amounts of surveillance data. In order for companies in the aviation domain to better make sense of current and historical data, and make predictions using descriptive behavior, a scalable analytics service is needed. In response to BCA's need in this nature, we built a data warehouse that will aid making critical decisions. In 1992 the Federal Aviation Administration (FAA) started a program to provide flight plan and track information from the NAS to airlines and other organizations in real-time or near real-time depending on the need. The feed, called the ASDI, is a product of the Enhanced Traffic Management System (ETMS). The feed originates from the Traffic Flow Management (TFM) Production Center located at the William J. Hughes Technical Center (WJHTC) in Atlantic City, New Jersey. The FAA distributes the ASDI data stream to a tier of vendors who in turn distribute it to their clients

In addition to processing of previously stored historical ASDI data in flat text files, we also needed to handle processing of continuous streams of ASDI data in near real-time which would then enable users to perform analytics operations once the data has been pushed into the data warehouse.

In response to these needs, we decided building a data warehouse that contains summary, historical, and detail data to support tactical and strategic decision making. Data would be extracted from operational data sources, transformed, cleaned, reconciled, aggregated, and summarized in preparation for warehouse processing

The data warehouse is implemented by a backend DB2 database . For near real-time ASDI feed consumption, message routing, transformation, and publish/subscribe capabilities, we used WebSphere Message Broker and WebSphereMQ . SPSS Modeler by IBM has been our choice of predictive data analytics tool that provides data collection, statistics, predictive modeling, and deployment capabilities Front-end visualization for business intelligence support was provided by IBM's Cognos BI tool .

LITERATURE SURVEY

Big data means data that cannot be handled and processed in a traditional manner. It will be so large as to not fit on a single hard drive, as a result, it will be processed on a number of cores [8].

There are number of articles and books on big data, analytics, data warehousing and OLAP technology and related research issues. While some of these research focus on physical and conceptual design, others target maintenance issues and stream processing. However, to the best of our knowledge, there is no work done similar to ours in the aviation domain where operational real-time or near real-time surveillance data is turned into a warehouse enabling critical decision making and predictive analytics in the literature

Many cybercrime is being casually executed for example spam, fraud, identity theft cyber terrorisms and phishing. Among this phishing is known as the most common cybercrime today. Phishing has become one amongst the top three most current methods of law breaking in line with recent reports, and both frequency of events and user weakness has increased in recent years, more combination of all these methods result in greater danger of economic damage and issues.

Phishing is a social engineering attack that targets and exploiting the weakness found in the system at the user's end. This paper proposes the Agile Unified Process (AUP) to detect duplicate websites that can potentially collect sensitive information about the user. The system checks the blacklisted sites in dataset and learns the patterns followed by the phishing websites and applies it to further given inputs. The system sends a pop-up and an e-mail notification to the user, if the user clicks on a phishing link and redirects to the site if it is a safe website. This

system does not support real time detection of phishing sites; user has to supply the website link to the system developed with Microsoft Visual Studio 2010 Ultimate and MySQL stocks up data and to implement database in this system.

The recent approaches to prevent the attacks like heuristics approach, blacklist approach, fuzzy rule-based approach, machine learning approach etc. and finally filtering all detection techniques based on accuracy and performance proposed a framework to detect and prevent phishing attacks. A combination of supervised and unsupervised machine learning techniques is used to detect malicious attacks.

CONCLUSION

This survey presented various algorithms and approaches to detect phishing websites by several researchers in Machine Learning. On reviewing the papers, we came to a conclusion that most of the work done by using familiar machine learning algorithms like Naïve Bayesian, SVM, Decision Tree and Random Forest. Some authors proposed a new system like for detection. The combinations of features with regards to accuracy, precision, recall etc. were used. As phishing websites increase day by day, some features may be included or replaced with new ones to detect them.

REFERENCE

1. Machine Learning-Based Phishing Attack Detection by Sohrab Hossain, Dhiman Sarma, Rana Joythi Chakma published in the year 2020.
2. Mahmoud Khonji, Youssef Iraqi, "Phishing Detection: A Literature Survey IEEE, and Andrew Jones, 2013