Traveller Usage Prediction And Renderingsolutions For Railway Infrastructure

Muthulakshmi.K[1], Madhuvanthy.R[2], Anusuya.S[3], B.Monisha[4]
[1]Assistant Professor, Dept of Information Technology
[2, 3, 4]Dept of Information Technology
[1, 2, 3, 4] Panimalar Engineering College, Tamil Nadu, India

Abstract- This project presents the main foundations of Big Data applied to Smart Cities. A general Internet of Things based architecture is proposed to be applied to different smart cities applications. In this project we are trying to predict and provide solution to improvise the railway infrastructure and train services. Indian Railways is a mode of transport service where thousands of transit-card transactions should be processed every minute. Thus our proposed system involves smart transit cards which can monitor the user behavior. The smart transit cards are also used for payment. Thus these transit cards would produce great amount of data which can be used for analytics and prediction. Predicting the consumers count and behavior who uses the railway services are solved through the application of big data technique .From the transit card, passenger source and destination are been analyzed for providing solution for improvisation. Also to make the collected data available to the railway department we encrypt the data using encryption algorithm and store in cloud

Keywords- Railway, Ticketing,Smart transit card.

I. INTRODUCTION

The main aim is to introduce a new ticketing system for theTrain system is in order to authenticate and authorize valid commuters to suit one's comfort, purpose or needs while travelling. Due to very high job prospects in city, the population which started travelling through this train increased immensely, which resulted in various issues such as never ending queues, wastage of paper, lots of resources and staff utilization. The alternatives techniques introduced to resolve these issues failed drastically. Moreover in spite of having such a massive amount of data generated of the commuters, no analysis was ever done ever to improve the railway service and the commuter's experiences. To overcome the above pitfalls this paper proposes a smart ticketing system architecture for railways which completely scraps the idea of paper tickets and completely harness the amount of money commuters has invested for their travelling. This model also enables the authorities to detect those commuters who never pay and buy or fail to carry their tickets or pass while travelling. The proposed system will also have features like crowd analysis and suggesting cost effective offers to the railway commuters

The aim of this project is to **avoid crowd, queue in stations to get tickets and to avoid paper tickets**. This system will make use of Radio Frequency Identification technology along with the **mobile phone and smart card** to carry out the ticketing transactions. The user prediction is not at all taken into account for improvisation. Railway facilities have not improved very substantially over the past few decades which leads to congestion and infrastructure damage. Also there is no initiative to improve the quality of service. There are many serious problems, some are to book a general class ticket, and one has to go to the railway station or book from the authorized agents. The process is quite time consuming and people don't like to stand in queues. The people are afraid of losing the train, so some people travel without ticket.

It is a 1st initiative to provide improvisation solution for our railway system using real time data. The real time data are collected using smart transit cards. In this project we are trying to predict and provide solution to improvise the railway infrastructure and train services. Predicting the consumer's behavior who uses the railway services are solved through analysis. From the smart transit card, passenger name, ticket amount and destination are been analyzed for providing solution for improvisation.The real time data collected are stored in real time cloud which can be accessed and viewed anywhere. Each user's travel history is stored in the database. A login is provided to user by which they can view their travel history. Analysis is made from data in real time cloud. All the data and analysis can be viewed by a railway admin, TTR and user from the cloud.

An android application is also provided for a user to view their travel details in an easy way.

II. FEATURES

a) **JAVA**

Java architecture provides a portable, robust, high performing environment for development.

Java provides portability by compiling the byte codes for the Java Virtual Machine, which is then interpreted on each platform by the run-time environment. Java is a dynamic system, able to load code when needed from a machine in the same room or across the planet.

b) IDE

An **integrated development environment** (**IDE**) or **interactive development environment** is a software application that provides comprehensive facilities to computer programmers for software development. An IDE normally consists of a source code editor, build automation tools and a debugger.

c) NET BEANS

Net Beans IDE is the official IDE for Java 8. With its editors, code analyzers, and converters, you can quickly and smoothly upgrade your applications to use new Java 8 language constructs, such as lambdas, functional operations, and method references.

d) ECLIPSE:

In computer programming, **Eclipse** is an integrated development environment (IE). It contains a base workspace and an extensible plug-in system for customizing the environment. Written mostly in Java, Eclipse can be used to develop applications.

e) MY SQL

MySQL is (as of March 2014) the world's second most widely used open-source relational database management system (RDBMS). Free-software-open source projects that require a full-featured database management system often use MySQL.

f) DROP BOX:

Drop box is a cloud storage provider (sometimes referred to as an online backup service) that is frequently used as a file-sharing service. Drop box is a file hosting service.

III. HARDWARE COMPONENTS

1. RFID

Radio-frequencyidentification (RFID) uses **electromagnetic fields** to automatically identify and track tags attached to objects. The tags contain electronically stored information. Passive tags collect energy from a nearby RFID reader's interrogating **radio waves**. Active tags have a local power source (such as a battery) and may operate hundreds of meters from the RFID reader. Unlike a **barcode**, the tag need not be within the line of sight of the reader, so it may be embedded in the tracked object. RFID is one method for **Automatic Identification and Data Capture**

RFID tags contain at least three parts:

- an **integrated circuit** for storing and processing information
  that **modulates** and **demodulates** a **radio-frequency** (RF) signals;
- a means of collecting DC power from the incident reader signal; and
- An **antenna** for receiving and transmitting the signal.

The tag information is stored in a non-volatile memory. The RFID tag includes either fixed or programmable logic for processing the transmission and sensor data, respectively.

2. EM18 MODULE

This module directly connects to any microcontroller UART or through a RS232 converter to PC. It gives UART/Wiegand26 output.

3. ARDUINO UNO

Arduino is an open source computer hardware and software company, project, and user community that designs and manufactures single board microcontrollers and microcontroller kits for building digital devices and interactive objects that can sense and control objects in the physical world.License, permitting the manufacture of Arduino boards and software distribution by anyone. Arduino board designs use a variety of microprocessors and controllers. The boards are equipped with sets of digital and analog input/output (I/O) pins that may be interfaced to various expansion boards (*shields*) and other circuits. A program for Arduino may be written in any programming language with compilers that produce binary machine code for the target processor. Atmel provides a development environment for their microcontrollers, AVR Studio and the newer Atmel Studio

4. ZIGBEE

Zigbee is IEEE 802.15.4-based specification for a suite of high-level communication protocols used to create with small, low-power digital radios, such as for home automation, medical device data collection, and other low-power low-bandwidth needs, designed for small scale projects which need wireless connection. Hence, Zigbee is a low-power, low data rate, and close proximity (i.e., personal area) wireless ad hoc network.

The technology defined by the Zigbee specification is intended to be simpler and less expensive than other wireless personal area networks (WPANs), such as Bluetooth or more general wireless networking such as Wi-Fi. Applications include wireless light switches, home energy monitors, traffic management systems, and other consumer and industrial equipment that requires short-range low-rate wireless data transfer.

IV. SYSTEM DESIGN

1. IDENTITY MANAGEMENT

Identity management (IdM) is the task of controlling information about users on computers. Such information includes information that authenticates the identity of a user, and information that describes information and actions they are authorized to access and/or perform. It also includes the management of descriptive information about the user and how and by whom that information can be accessed and modified. Managed entities typically include users, hardware and network resources and even applications.

In the real-world context of engineering online systems, identity management can involve four basic functions:

1. The pure identity function: Creation, management and deletion of identities without regard to access or entitlements;
2. The user access (log-on) function: For example: a smart card and its associated data used by a customer to log on to a service or services (a traditional view);
3. **The service function**: A system that delivers personalized role-based, online, on-demand, multimedia (content), presence-based services to users and their devices.
4. **Identity Federation**: A system that relies on federated identity to authenticate a user without knowing his or her password.

2. SECURITY

2.1 SQL INJECTION

SQL injection is a code injection technique, used to attack data-driven applications, in which nefarious SQL statements are inserted into an entry field for execution (e.g. to dump the database contents to the attacker).SQL injection must exploit a security vulnerability in an application's software, for example, when user input is either incorrectly filtered for string literal escape characters embedded in SQL statements or user input is not strongly typed and unexpectedly executed. SQL injection is mostly known as an attack vector for websites but can be used to attack any type of SQL database.

SQL injection attacks allow attackers, to spoof identity, tamper with existing data, cause repudiation issues such as voiding transactions or changing balances, allow the complete disclosure of all data on the system, destroy the data or make it otherwise unavailable, and become administrators of the database server.

2.2 BRUTE FORCE ATTACK

A brute-force attack is a crypt analytic attack that can, in theory, be used to attempt to decrypt any encrypted data (except for data encrypted in an information-theoretically secure manner). Such an attack might be used when it is not possible to take advantage of other weaknesses in an encryption system (if any exist) that would make the task easier.

When password-guessing, this method is very fast when used to check all short passwords, but for longer passwords other methods such as the dictionary attack are used because a brute-force search takes too long. Longer passwords, passphrases and keys have more possible values, making them exponentially more difficult to crack than shorter ones.

Brute-force attacks can be made less effective by obfuscating the data to be encoded making it more difficult for an attacker to recognize when the code has been cracked or by making the attacker do more work to test each guess. One of the measures of the strength of an encryption system is how long it would theoretically take an attacker to mount a successful brute-force attack against it.

Brute-force attacks are an application of brute-force search, the general problem-solving technique of enumerating all candidates and checking each one.

3. PAYMENT

This card looks like any plastic payment card but it has a microchip embedded on its face. This can hold more information than ordinary credit cards with magnetic strips.

Rather than holding only card's information, it can also hold information for such as transportation, identification and others.

This enables information for different purposes to bestored in one location. The smart card can be used to make purchases over the internet with the use of a card reader to read the card details necessary for payment and secure sending of data over the Internet

ADVANTAGES
- Portability
- Increasing data storage capacity
- Reliability that is virtually unaffected by electrical and magnetic fields.

IV. ANALYSIS

4.1 CLUSTERING

Clustering is the grouping of a particular set of objects based on their characteristics, aggregating them according to their similarities. Regarding to data mining, this methodology partitions the data implementing a specific join algorithm, most suitable for the desired information analysis.
This **k means clustering** analysis allows an object not to be part of a cluster, or strictly belong to it, calling this type of grouping hard partitioning. In the other hand, soft partitioning states that every object belongs to a cluster in a determined degree. More specific divisions can be possible to create like objects belonging to multiple clusters, to force an object to participate in only one cluster or even construct hierarchical trees on group relationships (based on check in timing of passengers).

4.2 PREDICTION AND SOLUTION

One copy of data is taken to the cloud and other copy is kept in an excel sheet. The data in the excel sheet is analyzed by using k-means clustering algorithm. The output can be viewed by the **user, TTR and a railway admin**. The output is included with the details of the destination railway infrastructure, the count of people who have already checked in. Based on this count the user can plan their travel schedule and the admin can record the user travel pattern and can provide different or further trains if necessary. All these information can also be viewed through an android application and by using a real time cloud URL.

V.CONCLUSION

This paper displays the benefits of applying big data techniques over data originated by IoT-based devices deployed in smart cities. Especially for the public transport service of a city. Thus this project collects the real time data and process using big data analytics tool. The predicted result would provide an efficient solution for railway department to improvise the service and infrastructure.

V1. TABLE OF FIGURES
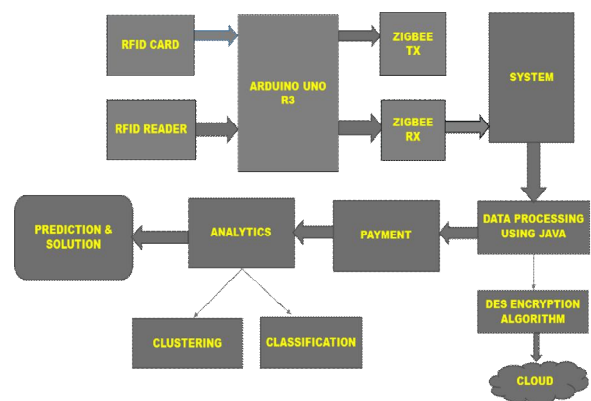

Chart-1: Hardware input

Chart-1: Architecture diagram

REFERENCES

[1] N. Komninos, "Intelligent cities: variable geometries of spatial intelligence," Intelligent Buildings International, vol. 3, no. 3, pp. 172–188, 2011.

[2] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," Computer networks, vol. 54, no. 15, pp. 2787–2805, 2010.

[3] L. Da Xu, W. He, and S. Li, "Internet of things in industries: a survey," Industrial Informatics, IEEE Transactions on, vol. 10, no. 4, pp. 2233– 2243, 2014.

[4] R. Iqbal, F. Doctor, B. More, S. Mahmud, and U. Yousuf, "Big data analytics: Computational intelligence techniques and application areas," Int. J. Inf. Manage, pp. 10–15, 2016.

[5] Z. Yan and D. Chakraborty, "Semantics in mobile sensing," Synthesis Lectures on the Semantic Web: Theory and Technology, vol. 4, no. 1, pp. 1–143, 2014.

[6] A. Carroll and G. Heiser, "An analysis of power consumption in a smartphone." in USENIX annual technical conference, 2010, pp. 1–14.

[7] M. Kantardzic, Data mining: concepts, models, methods, and algorithms. John Wiley & Sons, 2011.

[8] L. Daniele, F. den Hartog, and J. Roes, "Created in close interaction with the industry: The smart appliances reference (saref) ontology," in Formal Ontologies Meet Industry. Springer, 2015, pp. 100–112.

[9] K. Janowicz and M. Compton, "The stimulus-sensor-observation ontology design pattern and its integration into the semantic sensor network ontology," in Proceedings of the 3rd International Conference on Semantic Sensor Networks-Volume 668. CEUR-WS. Org, 2010, pp. 64–78.

[10] M. Compton, P. Barnaghi, L. Bermudez, R. GarcíA-Castro, O. Corcho, S. Cox, J. Graybeal, M. Hauswirth, C. Henson, A. Herzog et al., "The ssn ontology of the w3c semantic sensor network incubator group," Web Semantics: Science, Services and Agents on the World Wide Web, vol. 17, pp. 25–32, 2012.