 ISBN : 978-93-5254-640-5

RFID Implementation hurdles for Public Transportation

[1]Amit K. Patel, [2]Jitendra P. Dave

[1]Assitant Professor, Sri C.U. Shah College of Master of Computer Application Wadhwan City, Surendranagar, Gujarat
[2]Assitant Professor, Sri C.U. Shah College of Master of Computer Application Wadhwan City, Surendranagar, Gujarat
[1]amitpatel_mca10@yahoo.co.in, [2]davejitendra1982@gmail.com

Abstract: **Radio Frequency Identification (RFID) tags has been proposed for use in novel ways for hundreds of applications. RFIDs hold the promise of revolutionizing business processes. This paper focuses on how RFID Technology can be used to solve problems faced by public transport in metropolitan cities of the country. Automated tracking of buses can be used to provide useful estimates of arrival times and enhance commuter convenience. There are, however, formidable hurdles in the way of widespread RFID deployment. From a systems perspective, we highlight and explore the problem of data capturing, storage and retrieval and how Event, Condition and Action (ECA) rules developed for active databases can help us in managing the huge number of events generated each day. We also discuss how the collected data can be used to predict bus movement timings in order to provide better service.**

*Keywords—**RFID, ECA, Data Management, Forecasting, Public Transportation.***

[1] INTRODUCTION

Radio Frequency Identification (RFID) tags have emerged as a key technology for real-time asset tracking. It is an automated identification technology that allows for non-contact reading [1] of data making it attractive in verticals such as manufacturing, warehousing, retail [2, 3], logistics, pharmaceuticals [4], health care [5] and security. RFID systems are foreseen as replacement to the legacy bar code system of identifying an item. One of the major advantages of RFIDs over bar codes is that it is a non-line-of-sight technology - thus every item need not be handled manually for reading. In addition, RFID readers can read tags even when they are hidden.

The bus system is one of the largest in transportation medium all around the India. Often the buses are overcrowded. As a result commuters usually spend long hours at bus stops waiting. The bus arrivals at a particular stop are stochastic variables thanks to traffic congestion. This unpredictability can be partly alleviated by deploying a bus tracking and reporting system. There are a couple of ways to address this problem; one approach is to use the Global Positioning System (GPS) and another is through the use of RFIDs.

In this paper, we propose a solution using RFID technology and present issues related to its deployment.

In section 2 we briefly introduce RFID technology and its components. Section 3 explains the solution for the proposed problem using RFID Technology. In sub-section 4.1 & 4.2 we pose the challenges of handling huge number of events and discuss how ECA rules [16] can be used in a distributed manner to handle event explosion. Section 5 provides a framework for using the collected data in predicting arrival times for buses at different stops.

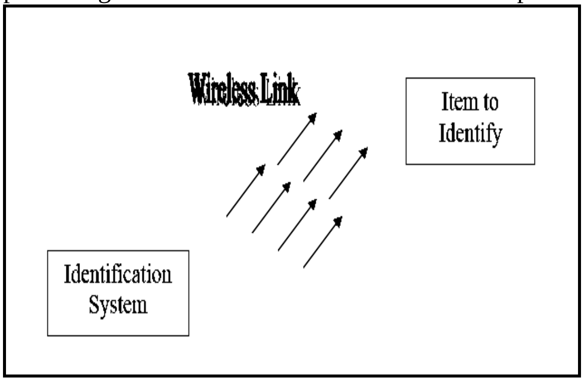

Figure 1. Radio Frequency Identification

[2] COMPONENTS of RFID System

RFID system comprises of RFID tag, RFID transceiver, servers, and middleware and application software. The RFID tag is a low functionality microchip with an antenna connected to the item to be tracked, or identified, and stores the unique identification number of the item. These chips transform the electromagnetic energy of radio-frequency signals/queries from a RFID reader/transceiver to respond by sending back information they enclose. The readers communicate with the tags for reading/writing the information stored on them as well as update the servers which may be standalone or networked. Readers may be fixed or mobile. Finally, a computer hosting a specific RFID application pilots the reader and processes the data it sends.

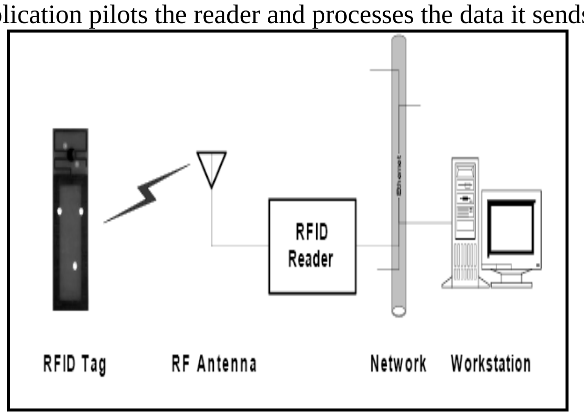

Figure 2. RFID System Components

ISBN : 978-93-5254-640-5

RFID Tags can be Active or Passive. Active RFID tags (beacons) are powered by an internal battery which is used to power ICs and generate the outgoing signal. They are typically read/write type and the size of memory used varies according to application requirements. The battery supplied power of an active tag gives it a longer read range, but such tags have large size, higher cost, and a limited operational life.

Passive RFID tags operate without an external power source. They use the operating power generated from the reader. Electrical current induced in the antenna by the incoming radio frequency signal provides enough power for the CMOS integrated circuit in the tag to power up and transmit a response. The absence of battery makes them lighter than active tags, less expensive, and offers a virtually unlimited operational lifetime. But they have shorter read range than active tags and require a higher-powered reader. Passive tags are typically Read-only and are programmed with a unique set of data (usually 32 to 128 bits) that cannot be modified.

The data transmitted by the tag may provide identification or location information, or specifics about the product tagged, such as price, color, date of purchase, etc. The interrogator, an antenna packaged with a transceiver and decoder, emits a signal activating the RFID tag so it can read and write data to it. When an RFID tag passes through the electromagnetic zone, it detects the reader's activation signal. The reader decodes the data encoded in the tag's integrated circuit (silicon chip) and the data is passed to the host computer.

RFID system can also be distinguished by frequency range. Low-frequency (30 KHz to 500 KHz) systems have short reading range and lower system costs. They are most commonly used in security access, asset tracking, and animal identification applications. High-frequency (850 MHz to 950 MHz and 2.4 GHz to 2.5 GHz) systems, offering long read range (greater than 90 feet) and high reading speeds, are used for applications such as tracking fast moving vehicles and automated toll collection. However, high-frequency RFID systems incur higher system costs.

[3] Proposed RFID based Solution

We first propose a solution based on RFIDs followed by an introduction to the GPS based solution.

[3.1] RFID BASED APPROACH

Each bus could have an RFID tag affixed to it while the readers are conveniently mounted at intersections, lamp posts or bus stops. The crucial information associated with a tag is the specific bus number, the capacity of the bus, the route number currently plying and the termination point (for example, during non-peak hours a bus may terminate at a depot before its usual terminating point). Tag readers continually monitor passing buses and transfer this information in real-time to a central computer.

A commuter with access to a cell phone could subscribe to the following service from the mobile network provider. The subscriber may enter his destination stop, D, (and optionally location of nearest bus stop) on his cell phone in the comfort of his home. The system will inform him of the relevant buses closest to him and expected arrival times of these buses.

The above service can be provided by the mobile network operator. The provider contacts a central computer to obtain the set of buses traveling to D through the closest bus stop to the customer. This list is obtained in sorted order and could possibly be filtered or enhanced in some way depending on the preferences of the customer. For example, if there are several bus stops in proximity to the commuter, information on relevant arrivals at all these bus stops can be provided. The provider can provide customized service to each subscribing commuter for a small fee.

This service can be used for multiple purposes to locate and control bus movement in the metro city. For example, in the event of an accident causing traffic congestion on a particular road, the buses leading to the road can be informed. In some cases, the routes of the bus can be changed temporarily and accordingly bus driver can be informed via wireless network. Or if it is found that a particular bus was stuck in traffic and that has led to a smaller gap with the next bus, the bus driver of the next bus can be informed to slow down to increase the gap. Many such applications can be thought of based on such an RFID application.

Before describing some of the research challenges in deploying such a system, we first describe an alternate technology that can also achieve similar objectives.

[3.2] GPS – BASED APPROACH

A GPS tracking system uses GPS (Global Positioning System) [14] to determine the location of a vehicle, person, or pet and to record the position at regular intervals in order to create a track file or log of activities. The recorded data can be stored within the tracking unit, or it may be transmitted to a central location, or Internet-connected computer, using a cellular modem, 2-way radio, or satellite. This allows the data to be reported in real-time; using either web browser based tools or customized software’s. More often, GPS receivers are used for navigation, positioning, time dissemination, and other research. Research projects [15] include using GPS signals to measure atmospheric parameters.

Though GPS based systems are widely used in the developed countries, there exists some serious limitations of this technology in developing countries like India. Firstly the coverage of GPS system in developing countries is not as wide. Secondly, effective implementation of a GPS system will require mapping the roads to the GPS system. Such mapping so far does not exist for metro cities in India. In the developed world, road infrastructure is almost static. However in the developing world metros (e.g. Mumbai, Delhi, Chennai, Hyderabad), new roads are being constantly built and layout of old roads is frequently changed. This will require remapping of roads at regular intervals. On the other hand with RFID systems new roads and change of old roads will require just reinstalling few RFID scanners or changes in the positions of these scanners.

ISBN : 978-93-5254-640-5

[4] RESEARCH CHALLENGES OF RFID BASED SOLUTIONS

There are many technical challenges associated with deployment of RFIDs. For example, there are problems with false or missing reads as a result of radio waves being easily distorted, detected, absorbed, or interfered with. There are a number of system-level challenges such as determining the number, type and placement of readers. In this paper we primarily focus on the challenges related to data management which deals with capturing, storing and querying RFID data.

[4.1] Data management problem

BEST runs over 335 bus routes in Mumbai [13]. The average time between two buses on a route is about 15 minutes. But, due to traffic congestion and peak crowds, the maximum time may exceed 30 minutes. Overall, there are around 3380 buses in B.E.S.T. which carry around 45,00,000 passengers everyday.

BEST buses, on many routes run for 21 hours (from 4:00AM to 1:00AM) a day. So the number of trips along a route will be 84 trips (21 x 60 /15), on an average. Thus, with 84 trips per bus route, and an estimated average number of bus stops per route = 17, we could estimate the number of events that will be generated in this scenario as below:

84 trips x 335 routes x 17 stops = 4,78,380 events

Processing and relating so many events to derive a meaningful real-time decision is a challenging task. The above estimates occur in the case when readers are placed at bus stops and depots and when only BEST buses are taken into consideration. If the data is captured not only from bus stops but also from several traffic lights to get intermediate information between two bus stops the number of events will further increase. This situation will be exacerbated if other kinds of traffic movements such as taxis, trucks are also monitored.

Managing such high volume of events and generated data poses the challenges to applications as well as back-end databases. This data is often redundant and needs to be filtered/cleaned and consolidated in order to occupy less space in database. In doing so, care must be taken that no useful information is lost.

Researchers in the database community have presented techniques and models for warehousing as well as cleaning/filtering RFID data. EPC-IS [9] and PML Core [10] are the RFID system standardization efforts by auto-ID center. [9] Summarizes the data characteristics, models data as events and provides some reference relation to represent data. Dynamic Relationship ER (DRER) presented in [17] is an expressive temporal data model which enables support for basic queries for tracking and monitoring RFID tagged objects. A simple observation that objects move together in initial stages bring a couple of more proposals. Hu et al. [12] used bitmap data type to compress the information corresponding to objects that move together. RFID-Cuboids [11] are a new warehousing model that preserves object transitions while providing significant compression and path-dependent aggregates. FlowCube [11] is a method to construct a warehouse of commodity flows. Some of these are simple representations of various relationships in the Relational DBMS.

[4.2] Real-time Decision Making

In this sub-section we describe how the application of Event-Condition-Action (ECA) framework can address some of the real-time event management issues.

Assume that we record each instance of reader-tag interaction with the help of a tuple: {object_epc, location, timestamp}. Here, object_epc is Electronic Product Code used to uniquely identify an object (the bus and the route), location denotes the place where the interaction took place (say in some bus stop), and timestamp denotes the time at which the interaction took place.

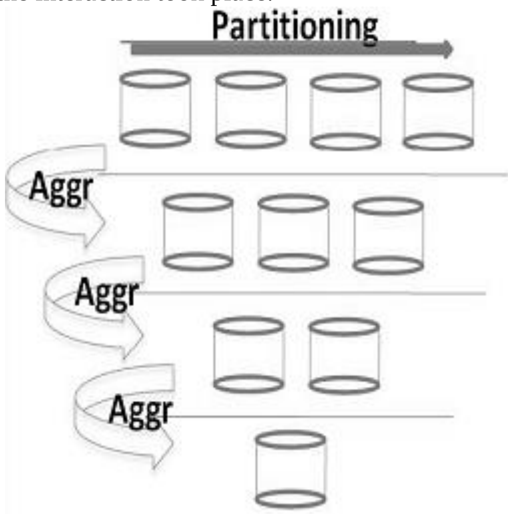

Figure 3: Data Aggregation & Partitioning in Databases

Each event is characterized by certain dimensions like time of scan, location of the reader, etc. Similarly, the conditions and actions also have some dimensions. Consider an example event that the distance between two consecutive buses is below a certain threshold. This can be expressed in the ECA form as:

EVENT e1 = {location = l1, timestamp = t1, epc = {route1, bus1}}

EVENT e2 = {location = l2, timestamp = t2, epc = {route2, bus2}}

EVENT e3 = {e1 AND e2}

CONDITION = {e2.l2 = e1.l1 + 1 AND e2.t2 < e1.t1 + threshold AND e1.route1 = e2.route2}

ACTION = {Notify bus driver of e2.bus1 to slow}

It is possible to dynamically add, delete and modify such rules without interrupting the system. Moreover such ECA rules will be invoked in real-time by a distributed ECA framework. These ECA rules can also be used to do several data management tasks – data cleaning, data aggregation and prediction.

[4.3] SCALABILITY

As the amount of data generated by RFID system is enormous, scalability of the proposed system is an important

ISBN : 978-93-5254-640-5

aspect. A couple of techniques can be employed in order to make our system scalable.

[4.3.1]Data Aggregation:

In many scenarios, granularity of recently generated data is more important as opposed to old data. So we can have multi-layered data architecture with high granularity recent data at the top and consolidated, less granular data towards the bottom. If for example, the data on individual bus movement is required for immediate action or for analysis over a day. However beyond few days, the historical data of individual buses will be of much less interest. Rather it will be more interesting to know the aggregated overall bus movement data. Following this, in our proposed RFID system, the historical data will be aggregated and stored.
This is a very good example for layered architecture for data storage on basis of aggregation. But this assumption may not always hold true. Sometimes, special events might take place making it necessary to violate this generalized structure. E.g., for future application it will be logical to study the movement pattern of individual buses on the day of the Mumbai bomb-blast (7/11). Such analysis can be used to effectively plan for disbursement of traffic in case of a future terrorist attack. So, individual bus movement data needs to be kept for that particular day. It is possible to describe such data aggregation rules based on the proposed ECA framework.

[4.3.2] Event Explosion:

Each time the tag is scanned a tuple is produced indicating location of object and time of scan. This tuple directly or indirectly leads to an event (assume filtering of tuples is not done). So, the number of events generated will be proportional to the number of objects which need to be tracked or monitored. This on an average is large. We can have a centralized event manager which identifies all the events in the systems, checks for conditions associated with it, evaluates these conditions and finally fires events. This event manager will also be overloaded and will become the bottleneck for our system.
The solution is to distribute the job of the event manager. The whole purpose behind this distribution is to capture events as close as possible to its generation point and execute corresponding action if the conditions defined are true. As an example, consider the ECA "when a bus arrives at bus stop A, if a previous bus of the same route has left A, beyond a threshold time then inform the bus driver to increase the speed". This ECA can be executed locally within the RFID scanner or a nearby computer to which the RFID scanner is connected. The event need not be propagated to the centralized system. Such a system will reduce the number of events to be handled at any single location. To achieve the optimal execution of RFID ECA rules, the system should be able to specify the location of event detection, condition checking and action handling.

[5] TURNING DATA INTO VALUABLE INFORMATION

Consider that a customer has queried the system at time t to obtain an estimate of the earliest arrival time of Bus # 252 at stop S3. Figure 2 shows the locations of two instances of Bus # 25 closest to the customer. The closer of the two busses is between stops S1 and S2. To estimate the arrival time of this bus at S3, we need to forecast the following times:

- Time to complete hop h1 and
- Times to traverse hops h2 and h3.

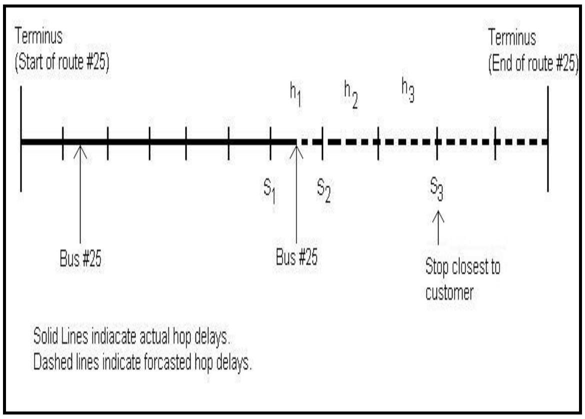

Figure 4: Estimating the earliest bus arrival time

The estimated arrival time of this bus at S3 is t plus the sum of the above delays. We next attempt to answer the customer's query by formulating it as a problem in the domain of forecasting theory.
Let $\tau_{r.h}(t_i)$ be the time taken to traverse the hth hop on bus route r given that the bus starts from its depot at time = t_i. The variable i is the bus departure index – it corresponds to the ith departure of Bus # r since bus arrival records were maintained. τˆ denotes an estimate of delay. Specifically $\hat{\tau}_{r,h}(t_i, t)$ denotes the time, estimated at time t, to traverse the hth hop on bus route r given that the bus starts from its depot at time = t_i.
Estimated times can best be analyzed by considering two categories of days – normal or regular days (these are typically working days in the week) and "other" days (these include week-ends and other holidays.)
Let $T_{r,h}$ denote the time series:
$\tau_{r,h}(t_1)$, $\tau_{r,h}(t_2)$, . . .
Note that t_1, t_2, t_3 . . . are the departure times from the depot of the 1st, 2nd, 3rd buses, etc. plying on route r. We clarify that, if there are n departures per day, then the first n terms of the series, $T_{r,h}$ correspond to the departures on Day 1, the next n departures are those on Day 2 and so on. For simplicity, we include only regular days in the series. We expect to mine certain patterns from such series. In particular for a given series (and hence a particular hop on a specific route, r).

- Values of successive terms in the series would tend to be close.
- The hop delay at 5:00 am (early morning, low-traffic load) would not be a good indicator of hop delay during evening rush hour (say at 6:00 pm). On the other hand, hop delay at

6:00 pm yesterday may be a better predictor of today's hop delay at 6:00 pm.

Observation 2 suggests a form of seasonality with period n in the time series (where n is the number of departures of Bus # r on any given regular day).

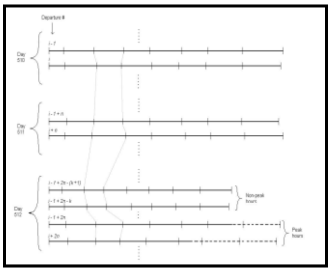

Figure 5: Relating delay in traversing the 3rd hop to corresponding delays in previous trips.

We thus see that the traversal time for hop h on route r may be estimated with the aid of previous terms in the series, $T_{r,h}$. This, however, may not always be the case. Consider, for example, a road accident or other such event that creates congestion at a certain point in the road network. All routes that converge at that point may also be affected. This means that hop delay estimates should also factor fresh information concerning hop delays encountered on other routes that intersect the route of interest to our customer.

6. CONCLUSION

Practical RFID systems are involved in real time tracking and monitoring of events. The system performs appropriate actions in response to events based on certain conditions. It is natural to consider the use of the Event, Condition and Action (ECA) framework to address event management issues. Since the number of events captured by many RFID systems is very large, clever filtering and aggregation techniques should be employed. We are currently engaged in the development of a RFID rule based management system using existing RFID middleware from Sun.

Transportation is a fertile area for deployment of RFID-based systems. Tracking of buses and other vehicles in crowded metros could greatly benefit

Comments who could plan their trips to avoid long delays at bus stops. In this paper, we use of RFIDs for bus tracking using readers strategic location. Event such as arrival of buses can be used to generate useful information such as earliest arrival time of a bus on a given stop. By informing a commuter about bus arrival times the commuter can save valuable waiting time. We

Formulated the goal of estimating bus arrival times as a technique developed for seasonal time series as well as regression analysis.

7. REFERENCES

[1] Walmart. http://www.walmart.com

[2] Albertsons Inc. https://www.albertsons.com

[3] RFIDs in Pharmaceuticals. http://www.rfidjournal.com/article/articleview/840/1/1

[4] RFIDs in Healthcare. http://www.rfidgazette.org/healthcare/index.html

[5] F Wang and P Liu. Temporal management of RFID data. In Proceedings of the 31st International Conference on Very Large Data Bases, 2005.

[6] Sun Microsystems, Inc. The Sun Java System RFID Software Architecture, March 2005.

[7] I-konnect. The Sun Java System RFID Software Architecture, May 2005.

[8] M. Harrison. EPC information service - data model and queries. Technical report, Auto-ID Center, October 2003.

[9] C. Floerkemeier, D. Anarkat, T. Osinski, and M. Harrison. PML Core Specification 1.0. Technical report, Auto-ID Center, September 2003.

[10] H Gonzalez, J Han, X Li, and D Klabjan. Warehousing and analyzing massive RFID data sets. In The 22nd International Conference on Data Engineering, 2006.

[11] Y Hu, S Sundara, T Chorma, and J Srinivasan. Supporting RFID-based item tracking applications in oracle DBMS using a bitmap datatype. In Proceedings of the 31st International Conference on Very Large Data Bases, 2005.

[12] BEST Undertaking. About the Organization, http://www.bestundertaking.com/org_info.asp

[13] Wikipedia: Global Positioning System, http://en.wikipedia.org/wiki/GPS

[14] Peter H. Dana. The Global Positioning System, http://www.colorado.edu/geography/gcraft/notes/gps/gps_f.html

[15] Norman W. Paton and Oscar Díaz, Active Database Systems, ACM Computing Surveys, 1999

[16] Fusheng Wang and Peiya Liu. Temporal management of RFID data. In Proceedings of the 31st International Conference on Very Large Data Bases, Trondheim, Norway, 2005.

[17] Sun Java System RFID Software 3.0 (Developer's Guide, Installation Guide and Administration Guide), http://docs.sun.com/app/docs/coll/434.2