Research Tribulation in RFID Data Warehouse

Barjesh Kochar\(^1\) and Rajender Singh Chhillar\(^2\)
\(^1\)GNIM, \(^2\)Department of Computer Science, M.D.U
barjeshkochar@gmail.com and chhillar01@rediffmail.com

ABSTRACT
RFID (Radio Frequency Identification) warehousing is the emerging field where we can store data capture by RFID technology. This technology has a lot of advantages and it is being used in various application areas like supply chain management, Retail stores and many more. In spite of lot of advantages RFID data has a number of challenges in the data management. In this research paper we are providing detail of RFID technology its working and challenges so that in future it can be used more effectively.

KEYWORDS
RFID, Data Warehousing, Data Management

1. INTRODUCTION
1.1 DATA WAREHOUSE
Data Warehouse is an information repository of an organization which stores data and is designed to facilitate reports and analysis. It is an informational environment that provides various facilities such as integrated and total view of organization or enterprise, makes current and historical data or information available at a particular time for decision making, and provides an interactive source of strategic information.

A Data Warehouse is neither hardware nor software but it is a BLEND OF MANY TECHNOLOGIES (Data management, Store Management, Data Acquisition). Data Warehouse collects all its data from operational systems, transform data by performing various operations (Extraction, Cleansing, and Aggregation) and then stores the data in suitable formats for easy decision making.

Apart from Operational System, there are many other sources from where data maybe collected. These sources are External Sources, Production Sources, Internal Sources and Archived Sources.

**External Sources:** - External data from outside environment of an organization contribute in External Source. For example, the Data Warehouse of any Car Rental Company contains data on current production schedules. The External Data in the Warehouse helps the company plan their Fleet Management.

**Production Sources:** - Data is collected from various operational system in different formats and from collected data, we choose segments of data.

**Internal Data:** - An organizations private data that is not to be disclosed such as spreadsheets, financial reports, customer profiles etc. contribute in the Internal Data.

**Archived Data:** - All the historical data, stored as back storage files or as snapshots contribute to Archived Data for Warehouse.

1.2 RADIO FREQUENCY IDENTIFICATION TECHNOLOGY
RFID technology involves embedding or applying an object to a person or product for the purpose of identification and tracking via the use of radio waves. It’s bifurcated into two main components –
1. Interrogators (also called readers)
2. Tags (also called labels)
Before delving further into the realm of RFID, it would be really advantageous if we knew the applications.
Some of the recent venues where RFID played are listed as follows
1. RFID was used to track participants at the “world’s largest 5K event”.
2. Baptist Memorial Healthcare installed RFID system to protect babies.
3. Bibliotheca is going to provide RFID system to 54-Library consortium in New York.

2. TECHNOLOGY BEHIND RFID
With RFID, the signals are transmitted using the electromagnetic or electrostatic coupling in RF(radiofrequency)portion of the electromagnetic spectrum. Any RFID system consists of two main parts-
1. An Antenna
2. A transceiver
The RF is read by the transceiver and the information is then transferred to a processing device [i.e. reader] and a transponder or a RF tag, which contains RF circuitry and information to be transmitted.
The information is transmitted by the antenna for the integrated circuit which then converts the radio wave reflected back from the RFID tag into digital information that is then headed to computers which will analyze it accordingly.
The bifurcation of the tags holding data is done as follows:
- Passive Tags:
The radiofrequency form the reader is used to transmit the signal by these tags. Herein the data is burned into the tag during its manufacturing, although some data can be rewritten.
- Active Tags:
Herein, the data signal is transmitted by an on board battery as the power and hence cover a greater radius and power random access memory[RAM], thereby they are dubbed as the more sophisticated of the instruments and have an ability to store upto 32000 bytes of data.

3. RFID frequencies
It is quintessential that the RFID tags and readers be tuned into the same frequencies so as to enable communications. A variety of frequencies can be used by the RFID systems to communicate but owing to the different ways any radio wave works and acts at different frequencies, most RFID systems are frequency specific.
Transmission range of more than 90 feet is offered by High Frequency RFID Systems[850 MHz to 950 MHz and 2.4 GHz to 2.5 GHz]. Though it has a bottleneck, the wavelengths in the 2.4 GHz range are absorbed by water, including the human body, hence reducing the range of application.

4. Architecture Of Managing RFID Data: -
The architecture presented above is divided into three categories. At the lowest layer we have various RFID tags (containing information). The next layer we have RFID readers. The layer that is responsible for capturing data and deals with a stream of tuples in the form (EPC, Location, time). Tuples contains information where EPC stands for a unique identifier read by RFID reader, Location denotes the place where RFID tag is read or scanned by RFID reader, and Time means the time of the reading of RFID tag.

The process of how data is captured from RFID tags and stored in Data Warehouse is explained below:

When a tag is scanned on a RFID reader, the Data capture layer responsible for storing an event stores the reading as (EPC, Location, and Time). The data at low level data is also called Dirty Data. Due to inconsistencies in low level data, Data capture layer performs various processes to make data in consistent state, these processes are Data Cleaning and Data Aggregation and the area were these processes are performed is called Data Staging Area.

At Data capture layer the dirty data may appear in two general forms:

1. **Missed Readings and Unreliable readings** :- Both these problems are very much common in RFID application and they often happen in wireless communication. Wireless communication leads to dropped readings with faulty tag readers.

Few receptor based applications are interested in concept of Temporal Granules and Spatial Granules rather than individual readings or devices in terms of time and space. Temporal Granule is the smallest unit of time that an application operates. For example, in a retail store to count the number of items on each shelf, the temporal granule operates for 5 seconds. Spatial Granule items are based on some special spatial categories that are at lower level of spatial granule on which application operates.

2. **Data Duplicacy** :- Data duplicacy is considered as the most serious problem in sensor networks. Since RFID is based on sensor technology, data duplicacy is one of the most serious problem in RFID.

Data duplicacy or Data Redundancy can occur at two levels:

- **At Tag Readers level:-**
  
  Data Redundancy at Tag Readers level occurs when multiple readers send signals to a particular item at the same time. The optimal solution to this problem is when an item is scanned by a single reader then all the other readers should get turned off.

  The same problem may occur again when two readers’ reads a tag at a same fraction of time assuming that other tags are turned off. To solve this problem one reader can lock the tag while reading it causing other readers to wait. When that reader releases the tag other reader can read it.

  The other solution would be to maintain a counter in a reader which counts the number of times a tag has been scanned. When a tag is brought in vicinity with readers, all reader issues a tag count on all its covered tags. The reader which issues the highest tag count will lock the tag.

- **At Data level:-**

  Volume of data for RFID is usually larger, that is, few readers can perform over 100 readings per second, and RFID data as compared to other data streams is less useful. For example, in case of traffic monitoring and financial applications each record is useful for future analysis. On the other hand in case of RFID data we should be able to identify data that have been repeated many times. Suppose in a store an item is kept on a shelf for the whole day, then that item will continuously send data to RFID management system. The possible solution to this problem would be to keep snapshots of specific reader and compare the snapshots every time the item is scanned. Here we assume that movement of item from one point to another point on same shelf is not considered as an event.

5. **FUTURE OF RFID**

RFID technology is used for automatic identification and data collection.

A consumer product could be tracked right from manufacturing to the retail store and down to the consumer home with the use of RFID technology, though it has not yet been proven, but if successful it could be the indication of a revolution in the field of consumer goods supply chain.
Another application of RFID is in the identification of hospital staff and patient.

6. BOTTLENECKS IN THE USAGE OF RFID
1. Technological Problems with RFID
2. Security and moral values Problems

6.1. TECHNOLOGICAL PROBLEMS
6.1.1 Problems with Unique RFID Standard
Different manufacturers implement RFID in different ways; global standards are still being worked on. RFID devices should never be allowed to leave their network, it can cause problems for the companies.

Consumers also face problems with the RFID standards. For and instance ExxonMobil’s Speed Pass system in order to access it they’d have to pay. And if all companies have their individual Speed Pass system, then a consumer would need to carry many different devices with them.

6.1.2. RFID systems can be easily disturb
RFID System employ the electromagnetic spectrum ,so it could be disturb if we use at high frequency as well as the emitted frequencies may be problem for end users.

6.1.3. Collision of RFID Readers
As tag information is read by readers so some time if more than one reader send signal to same tag for the information then tag is unable to respond appropriately to two queries simultaneously . To overcome this, many systems use an anti-collision protocol [also called singulation protocol]. These protocols enable the tags to take turn in transmitting to a reader.

6.1.4. Collision between RFID Tags
The presence of many tags in a small area creates a hassle [i.e. tag collision]; so it may be difficult for the readers to read information about all tags. So there may be possibility of loss of data.

6.2. SECURITY, PRIVACY AND ETHICS PROBLEMS WITH RFID
6.2.1. Incorrect information read by Readers
As readers can read the information of tag through a distance also so if the consumer has purchased the goods and paid the cost still reader may read the information while the customer will be walking out of store which will lead to incorrect or extra information which may not be required by the Data Management system.

6.2.2. RFID tags cannot be removed easily
Another fault of the RFID system being the difficulty of removal of the RFID tags. Some of these tags may be palpably embedded inside a product where viewing them is again a hassle. Nowadays, RFID tags are printed onto the product hence making them non disposable.

6.2.3. RFID tags read without one being aware of it
Anyone who owns an RFID tag reader may easily read the tags embedded in your clothes and other products without you even being aware of it. So it is the security issue as some one with a reader can read the tag information of your goods your bank account numbers or so on.

6.2.4. High Radius of RFID leads to privacy problems
Usually, RFID readers are designed keeping in view the distance between the tag and reader to be minimum. But we can use the high gain antenna so the range of reading the tag information will be increased and the reader can read the tags at a distance also which may breach the privacy.

CONCLUSION
As explained above, RFID forms the backbone of goods security system now days, and had it not been for RFID, theft of goods would still be a major bottleneck for all the big stores and haute couture departments etc.

The problems that RFID poses i.e. security, privacy ethical problems etc all get obsolete once one looks at the positive side of RFID. Besides, solutions for the abovementioned problems are being looked and if solutions to the problems like RFID tag not coming off when one uses it or it being traced even in the privacy of one’s home are found, RFID is sure to become the apogee of commercial security systems.

REFERENCES


[10]. Derakhshan, R., Orlowska, M., Li, X.: RFID Data Management: Challenges and Opportunities. IEEE
International Conference on RFID 2007, Grapevine, Texas, USA (2007)


[15]. Han, J., Gonzalez, H., Li, X., Klabjan, D.: Warehousing and Mining Massive RFID Data Sets. 2006 Int. Conf. on Advance Data Mining and Its Applications (ADMA’06), China, August (2006)