Journal of Data Science 6(2008), 369-388

Radio Frequency Identification: A new Opportunity for Data Science

Vijay Wadhwa and Dennis K. J. Lin
The Pennsylvania State University

Abstract: Radio Frequency Identification (RFID) has taken center stage at retail and consumer products forum. RFID is not a new technology; it has been in use for many years. In this paper, we first review RFID technology and the components that form the backbone of the RFID system. Next, we demonstrate the usefulness of RFID in supply chain and present some data mining challenges in RFID. Finally a real-life case study is used to illustrate how organizations are using RFID data.

Key words: Data mining, RFID, supply chain management.

1. Introduction and Historical Perspective

Radio Frequency IDentification (RFID) serves as the basis for an automated data collection system. RFID has rapidly taken center stage at retail and consumer-products forums around the globe (Hardgrave, 2006). RFID is neither a fad nor a new technology; consumers have been using RFID tags at gas stations and highway tollbooths for years; what's new is its increasing affordability – and, hence, scalability – and the commitment to use the technology to drive cost savings and improved efficiency. RFID is a generic term for technologies that use radio waves to automatically identify people or objects (Auto-ID Center, 2002). RFID originated in the 1940's as a means of distinguishing friendly aircraft from enemy aircraft. Large powered RFID tags, or transponders, were placed on friendly aircraft. When interrogated by a radar signal, these transponders would give the appropriate response to identify the carrying aircraft as *friendly*. This IFF (Identify: Friend or Foe) system was the first obvious use of RFID (Stockman, 1948).

The RFID system, at the front-end, consists of a tag, which is embedded, in an object or item. This tag uses a microchip to store information and an antenna to transmit the stored information. An RFID reader is used to convert the radio waves from the RFID tag into a form that can be passed on to a computer. The computer system, which forms the back-end of the RFID system, interprets this data and passes the information on to other systems like the Inventory Management System (IMS), Enterprise Resource Planning system etc (Finkinzeller, 2000). The basic components for RFID technology are described in Figure 1:

1. RFID Tag: An RFID tag is made up of a microchip attached to an antenna. It is placed on the pallet or stock keeping unit (SKU). There are different types of tags for example, active and passive tags, read-write vs. read only tags. Each type of tag has a specific application.

2. Reader: Two most common type of RFID readers are fixed and handheld reader.

3. Antennae: RFID readers use an antenna to communicate to the RFID tag through the tag's antenna. Some readers have integral antenna while others can have various types and sizes of antenna fitted to them.

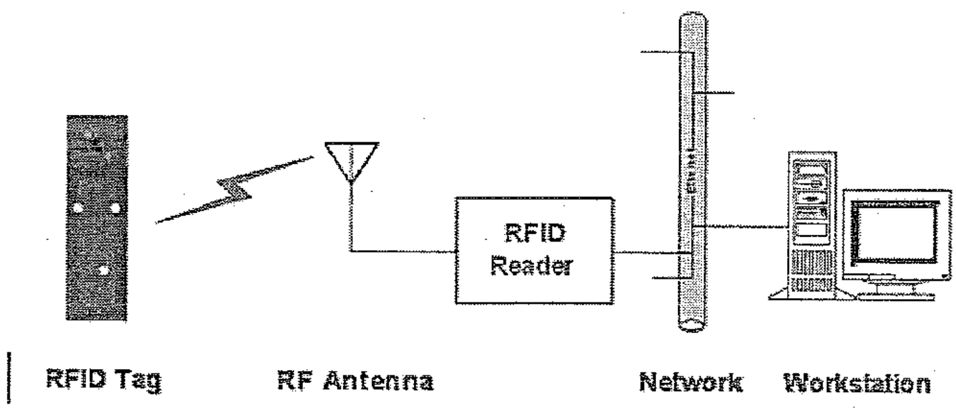

Figure 1: RFID functionality

Interest in RFID technology is particularly acute for retail and consumer goods industry because of Wal-Mart. In June 2003 Wal-Mart mandated its top 100 suppliers to be RFID compliant. Top 100 suppliers of Wal-Mart began shipping tagged products in 2005. The next 200 suppliers began shipping RFID enabled products in January 2006 and the next 300 suppliers in January 2007 (Delen, 2007) .This has ranging implications for retail and consumer goods supply chain, (Atkinson, 2004):

- Wal-Mart's top 100 suppliers are the world's largest non-automotive consumer goods manufacturers. They will become RFID compliant to support Wal-Mart.

- Wal-Mart's retailing competitors will follow the lead soon.

This paper is organized as follows: In section 2, we discuss EPC code which is used to uniquely identify every object in a supply chain; Section 3 deals with the impact of RFID on supply chain and compares RFID with Barcode. We also discuss a theoretical model for managing inventory in supply chains. Section 4 discusses the challenges in mining the RFID data. Section 5 illustrates a real life case study of analyzing RFID data at IBM. Section 6 provides some intelligent applications of RFID and finally we have conclusion and suggestions for future work in Section 7.

2. Electronic Product Code

EPC-global (EPCGlobal Inc) is an organization that is leading the drive to standardize and promote the Electronic Product Code (EPC). EPC combines RFID with a network system to capture and enable real-time identification and sharing of information throughout the supply chain. The EPC is a unique number that identifies a specific item in the supply chain. The EPC is stored on a RFID tag. Once the EPC is retrieved from the tag, it can be associated with dynamic data such as from where an item originated or the date of its production.

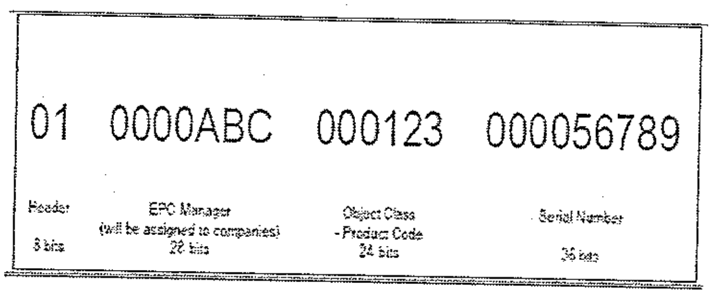

Figure 2: EPC structure

The EPC is a number made up of a header and three sets of data, as shown in Figure 2. The header identifies the EPC's version number – this allows for different lengths or types of EPC later on. The second part of the number is the EPC Manager – the manufacturer of the product the EPC is attached to - for example 'The Coca-Cola Company'. The third, called object class refers to the exact type of product -for example 'Diet Coke 330 ml can'. The fourth is the serial number, unique to the item – this tells exactly which 330 ml can of Diet Coke we are referring to. This makes it possible to identify each and every unique item in the supply chain. Currently, most RFID tags contain 96 bits of data in the form of serialized global trade identification numbers (SGTIN) for identifying cases or serialized shipping container codes (SSCC) for identifying pallets (Delen, 2007).

3. RFID and Supply Chain Management

The use of RFID in supply chain is a perfect example of how information technologies and business processes re-shape each other and produce solutions where the two can no longer be distinguished. Stand alone, RFID looks interesting but does not show immediately a return-of-investment. An analogy can be drawn from the Internet and World Wide Web. As a point-to-point communications tool, Internet bears no big difference from the traditional EDI from a business point of view. As a broadcasting method, the Web looks like a fancier version of newspaper and printed catalogs. When companies such as Amazon created an entirely new business that could not have existed without the Web, the true value of the technology started to emerge. For example, now vendors can use the Web Service technology to coordinate their marketing and supply-a practice no one could have dreamed of when we first looked at the Web just a few years back. The true value of the technology is still demonstrating itself in different forms, revamped or invented, in this case, XML-based EDI and Google. Whoever gets the integration of a new technology and their business processes right, will make that integration their core competitive advantage long before followers try to mimic and catch up; and each case would be as different for each business as they are brilliant in their creativity and execution.

3.1 Use of RFID and barcode in supply chains

For organizations with a global supply chain and vast distribution and retail network, benefits of RFID are numerous. The increases in data capture capability, data integrity and supply chain visibility will considerably reduce costs and increase the supply chain efficiency (Lin, 2007). Although both barcode and RFID can help drive down the supply chain costs, both are different technologies that can work well in conjunction with each other. Bar code is an optical technology and RFID is a radio technology.

There are many differences between RFID and barcode; for instance, RFID tags can be read through any surface except metal whereas barcodes have to be placed on the surface. RFID tags can be programmed on the fly and can provide real time data scan, whereas barcodes cannot be reprogrammed. RFID tags have a much higher data storage capacity which makes it suitable candidate for identifying every object in a supply chain, different barcodes can have different data storage capacity but they are not suitable for item level identification. On the flip side barcodes have a much lower operating cost compared to RFID. For example each RFID tag costs about 25 cents whereas a barcode costs about 1 cent. RFID implementation involves several other expenses like readers; RFID enables warehouses etc (rfidjournal.com). Although RFID can provide several distinct advantages, implementation of RFID system for any application needs careful understanding of the current business processes and the desired role RFID would play to enhance them. The distinct advantages of RFID comes from the saving in labor and increased throughput or efficiency. Two of the biggest advantages of RFID are supply chain visibility and theft reduction (Rappold, 2003). Although barcodes have been long used in supply chain RFID is gaining on them. A few primary differences between RFID and Barcodes are shown in Table 1 (Raza *et al.*, 1999, Delen 2007).

Table 1: Comparison of RFIID with barcode

Subject	Barcode	Passive RFID	Active RFID
Power Supply	Has no power source	Radio waves	Battery powered
Modification of data	Data cant be modified	Can be modified	Can be modified
Data Security	Minimal security	Security can be changed	Highly secure
Memory capacity	Limited	Up to 64 KB	Up to 8 MB
Life span	Short, easily damaged	Indefinite	Depends on battery
Read range	Line of sight	Not constrained by line of sight	Not constrained by line of sight
Interference	Optical barriers between the barcode and reader	Environments that affect radio waves	Limited interference
Cost	Low	Medium (about 25 cents)	High

3.2 Managing inventory using RFID

The fundamental benefit of RFID is in the finer visibility that it can provide (Lee, 2004). Finer visibility means we can observe more individual items and more states. Without RFID, we have coarse information, based upon which the solution either does not solve the problem, or has to incorporate more tolerance, which leads to higher variance in decisions, and thus shows bull-whip effect. With RFID, we can have more detailed information, if we can know which details matters (through Statistical Process Control), we will be able to react more precisely, and cut down the jitter throughout the supply chain.

In inventory management, there are two stochastic processes: the inspection process, which examines the current inventory level, and the replenishment (ordering) process, which replenishes or places orders. Both require some time to complete (lead time). Both the inspection and replenishment processes can be discrete or continuous, which gives us four combinations, yet in the traditional inventory control theory, there are only two control theories. In continuous review, both inspection and ordering are continuous. In periodic review, both are discrete. RFID calls for a new type of model that is continuous in inspection, but periodic in replenishment.

Let us consider a base stock policy of $(0, L)$ for a particular product, i.e., if the inventory level reaches zero, then replenish it back to L. Here we assume instantaneous replenishment i.e., lead time is zero. Suppose that the inventory inspection happens at time $nT, n = 1, 2, 3, \ldots$, for some inspection cycle time T. It is clear that each time the inventory is brought back to L, the process probabilistically renews itself. When the inventory level drops to zero before an inspection, stock-out appears and we may have lost sales. We can calculate the expected number of lost sales between replenishment.

$$\text{Average number of lost sales} = \sum_{i=1}^{\infty} iP[\text{ number of lost sale } = i].$$

$$\begin{aligned} &P[\text{number of lost sales} = i] \\ &= \sum_{n=1}^{\infty} P[\text{ lost sales happen in } t, (n-1)T < t \leq nT, \text{ and number of lost sales } = i] \\ &= \sum_{n=1}^{\infty} P[(N((n-1)T) < L \cap (N(nT) = L + i)], \end{aligned}$$

where $N(t)$ is the cumulative demand (since the last replenishment) at time t. Then,

$$\begin{aligned} &P[N((n-1)T < T) < L, N(nT) = L + i] \\ &= \sum_{j=0}^{L-1} P[N((n-1) = j, N(nT) = L + i] \\ &= \sum_{j=0}^{L-1} P[N(nT) = L + i | N((n-1)T] P[N((n-1)T = j] \\ &= \sum_{j=0}^{L-1} P[N(T) = L + i - j] P[N((n-1)T) = j]. \end{aligned}$$

Suppose the demand $N(t)$ is a Poisson process with rate m, substitution yields

$$P[\text{ number of lost sales } = i] = \sum_{n=1}^{\infty} \sum_{j=1}^{L-1} e^{-2mT} \frac{(mT)^{L+i-1}(m(n-1)T)^J}{j!(L+i-j)!}.$$

So,

$$\begin{aligned}\text{average number of lost sales} &= \sum_{i=1}^{\infty} iP[\text{ number of lost sale } = i] \\ &= \sum_{i=1}^{\infty}\sum_{n=1}^{\infty}\sum_{j=1}^{L-1} e^{-2mT}\frac{i(n-1)^j(mT)^{L+i}}{j!(L+i-j)!}.\end{aligned}$$

The average lost sales depends on T, which is the inspection cycle time. Inspection is an expensive business process. By adopting RFID, we can significantly reduce T, and reduce the lost sales.

3.3 Impact of supply chain strategy on IT

RFID applications are further complicated when examined under different integration strategies of the supply chain. For example, a tightly integrated supply chain may find centralized network architecture brings the best performance. On the other end of the spectrum, a supply chain with members who sometimes have to compete with each other may find the only solution is a peer-to-peer model. By *IT architecture*, we mean a set of cohesive and comprehensive design decisions that have the greatest impact on the actually IT systems when they are built and used. A good IT architecture requires not only the mastering of underlying technologies but also a profound understanding of business strategies and operations management. A very tangible research objective is: given certain degree of integration of a supply chain, find the best IT architecture that accomplishes, for various parties in the supply chain, the secure and fast discovery and retrieval of appropriate information associated with the data in RFID tags.

4. RFID Data Mining

As a product moves through the supply chain from the supplier to manufactures to the final customer, it may pass through various RFID readers whih will capture the tag data. In a typical supply chain RFID readers will be placed at several important locations; receiving dock, shipping area, conveyor belt, storage facility, shop floor and sales floor to name a few. In its most elementary form the RFID data would consist of EPC number, Entity type (Supplier, retailer, etc) where the tag was scanned, date and time of scan and the reader location (Delen, 2007). Data collected in this format can be analyzed to improve supply chain operations.

Figure 3 illustrates the flow of material in a typical supply chain. Tracking the product at various points in the supply chain can provide exact information on flow-time of the product. Analyzing the flow-time can provide critical insights into the forecasting system and also the lead-time of different products (Delen, 2007). For example from Figure 3 it can be seen that α, β and γ respectively measure the time spent by a product in a warehouse, time spent in transit (from warehouse to retail store) and the time spent in a retail store. The measurement of these times can also provide insight into supply chain efficiency.

Another unique advantage of analyzing RFID data is to understand the precise movement of products from one supply chain entity to another. For example, the optimal product movement in a retail store should be the arrival of the product at receiving dock and then to sales floor. Through the technique of *process mining* (Agrawal, 2006), the actual movement of the product can be compared to theoretically *optimal* movement. Process mining helps in analyzing business processes based on event logs.

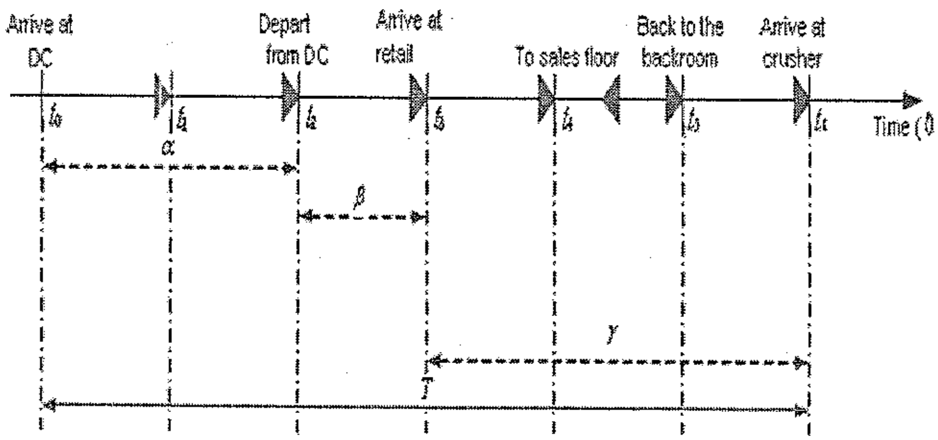

Figure 3: Movement of product in a supply chain (Delen, 2007)

As companies move from piloting to deploying RFID technology, they have to be prepared to deal with a variety of data management issues. RFID projects will generate huge amount of data. There has to be techniques designed to differentiate between "good data" and "bad data" or more appropriately "relevant data" and "redundant data" (Chawate, 2002). Not only the amount of data generated is huge, but also the time in which in this data is to be analyzed is critical. Data mining can help alleviate these problems. Data mining also referred to as knowledge discovery is the process of analyzing data from different perspectives and summarizing it into useful information – information that can be used to increase revenue, cuts costs, or both. Data mining allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified.

Challenges of RFID data mining

In mining the RFID data we need methods of collecting, cleaning and updating data, and may be updating the data. With this in mind the data-miner has to decide which inferences should be made eagerly and which inferences can be made lazily. A summary of challenges faced in data mining are listed below:

- The data may not be held at one location, through its travel across the supply chain, the information about the product will be held in repositories of different entities; therefore the first issue is timely collection of data.
- The data mining cannot be the traditional form of data-mining, RFID data mining in many cases will have to be real-time; real time data analysis is very critical, unless and until we can do real time data mining we cannot realize significant RFID advantages.
- The way the data will be obtained through RFID is also a possible problem. Let us assume there are four different supply chain entities; if all the four use different middleware or ERP system, the question is how we obtain data in a format that can be common across all the entities is a potential problem.

Gonzaleze *et al.* (2006) have proposed a method to construct compressed probabilistic workflows which capture trends in RFID data.

5. IBM Experience

IBM is in the process of implementing RFID at two locations; as a part of the alliance between IBM and Supply Chain Consortium at Penn state, we analyzed the RFID data provided by IBM.

5.1 Business process

IBM uses various chemicals in its manufacturing processes. The company needs to keep track of the movement of these chemicals within the vicinity of its manufacturing site from the moment the chemicals enter the site until the moment they are used and their container disposed of.

There are two primary reasons for accurately tracing data of these chemicals. First, some of the chemicals are hazardous and the company executive management has mandated a record of trace be kept. Second, the quality of these chemicals has a direct impact on the manufacturing processes and final products. Process engineers and quality engineers require a detailed record of what chemicals are used at which point of the manufacturing processes at what exact time.

The chemicals move in three main stages at a manufacturing site, which is a campus of 1-mile radius. Each campus has a central warehouse. At the first stage, the chemicals arrive at a receiving dock of the warehouse. At the second stage, the chemicals are transported by small vehicles from the warehouse to a storage room in a manufacturing building. There are many buildings around the warehouse. Each building may have a few storage rooms. At the third stage, the chemicals are moved from a storage room to a pumping station (room) in the same building. A pumping station has several pumps through which the chemicals enter the manufacturing processes.

Chemicals in liquid form are stored and transported in a drum and chemicals in gas form in cylinders. A drum is a 55 gallon plastic container. Cylinders are made of metal. Each drum and cylinder has a barcode id that uniquely identifies that container and its content. Each pump has a unique a barcode label. Storage rooms have their unique labels too.

In the current practice, workers use hand-held barcode readers to scan the barcode labels on the containers and pumps, and to enter other relevant information by hand. The mobile barcode readers use Wi-Fi technology to relay the real-time raw data to a premise server, which applies some business logic to the data and provide the final trace information of the chemical containers. At the final stage, before a container is hooked up to a pump, the barcodes on both the container and pump are scanned, as well as the worker's badge. Only after the premise server verifies the valid match between the chemical and pump, is the worker given the Ok to proceed with the hook-up.

A pilot project was put into place to use RFID technology for raw data gathering in lieu of barcode. About three percent of the containers are tagged with an RFID ID number, which is the same as its bar code. Fixed RFID readers are installed in various locations. The pumps don't have RFID tags due to the interference caused by the metal pumps. One of the goals that can be achieved through this analysis is that IBM would be able to set standards for time and motion studies, both for the employees and operations; this would greatly improve the process efficiency.

5.2 Data format

We have three different data files, first, the raw data captured by RFID readers. Second, processed data based on raw data captured by both RFID and barcode readers. The processed data represents movement of the chemicals; raw RFID data forms a subset of processed data. The third data set is metadata that further define the location terms used in the other two files. A snapshot of the data is attached in Table 2-4.

The data provided by IBM is in three different files. The *rawdata* contains raw data captured by RFID readers. The *movedata* contains processed data based on raw data captured by both RFID and barcode readers. The processed data represents movement of the chemicals. The valid data in rawdata.txt forms a subset of the *movedata*. The third file, *locationdata,* is metadata that further defines the location terms used in the other two files.

The *rawdata* (Table 2) has the following format.

- UID1: A unique id generated by the premise server for a row, which is a raw RFID read.
- TAGID: A unique RFID tag id on a container. This is same as the CONTAINERID in movedata.
- LOCATION ID: An id for a location, which is normally a storage room.
- READER ID: An id for a reader. When each room has just one reader, the reader id is the same as the location id.
- DTS: A day time stamp. It is a Linux UTC timestamp.

UID1 is a unique id generated for each row, UID1 and the TAGID forms the primary key in the *rawdata* table. IBM started by trying to fit the container id (e.g. BURL200612345) into a 96 bit tag; then at a later stage of implementation moved to 128 bit tag so there are a lot of leading zeros. For example,

- The container id 00 00 00 00 00 00 00 04 25 55 24 C2 00 60 24 32 Contains the following data:

  Site: BURL (4255524C) ascii conversion.
  Year: 2006 treated as hex digits even though the year is decimal.
  Container ID: 02432 treated as hex digits even though it is decimal.

- A similar logic follows for the East Fishkill site

  Site: EAST (45415354) ascii conversion
  Year: 2006 treated as hex digits even though they are decimal
  Container ID: 02432 treated as hex digits even though they are decimal

The *movedata* (Table 3) has the following format.

- TOBUILDING: The id of the building to which the container is moving.
- TOLOCATION: The id of the room to which the container is moving.

- DATE, TIME: The date and time stamp generated by the readers and not the servers.
- CONTAINERID: A unique Barcode or RFID id on a container, the container id is the same as in the rawdata file
- UNIQID: A unique id generated by the server for this row, which is a movement observed by an RFID reader or barcode reader.

TO BUILDING is the id for the building, every building has a unique id assigned to it. The RFID fixed reader moves require directional logic to function properly i.e., for the column TOLOCATION, if the last transaction was "in-trans" then the next one is "into the building".

The *locationdata* (Table 4) has the following format.

- LOCATION ID: Location id for a location on campus. Location ID is a foreign key in *rawdata* and *movedata* files.
- DESCRIPTION: Description of the location.
- CONTAINER LOCATION ID: The prefix used in container IDs.
- CONTACT_NAME: The name of the employee at the time of scanning.
- LOCATIONIN: Inside of a location.
- LOCATIONOUT: Outside of a location.

Table 2: IBM RFID data 1 (Rawdata)

UID1	TAGID	LOCATION_ID	READER_ID	DTS
A422520079121593	B	C04	BTV.R004	1142252097776
A1422529851381599	B	C17	BTV.R017	1142252988109
A422529930411601	B	C16	BTV.R016	1142252995451
A43814093125534	B	C10	BTV.R010	1143814094235

A=ibmrfidsrvr-savant11, B=00000004255524C200602432, C=BTV.R0

Table 3: IBM RFID data-2 (Movedata)

TOBUILDING	TOLOCATION	DATE	TIME	CONTAINER_ID	UNIQID
BURLSITE	BURL-IN-TRANS	3/13/06	07:14.0	D	A422520137031594
BURL973	M10-CHEM-DOCK	3/13/06	07:29.0	D	A422529923941600
BURL973	M10-CHEM-DOCK	3/13/06	07:29.0	D	A422529990491602
BURL973	M10-CHEM-DOCK	3/31/06	09:08.0	D	A438140991805345
BURL974	BTV-MT-CYLS	3/31/06	14:21.0	D	A438331864385439

A= ibmrfidsrvr-savant11, D=BURL200602432

Table 4: IBM RFID data-3 (Locationdata)

(A)	(B)	(C)	(D)	(E)	(F)
BTV.R001	BTV Reader 1 E1 Dock	BURL	A	No EPT	BURL974-E1-CHEM-DOCK
BTV.R002	BTV Reader 2 Chem Mix	BURL	A	No EPT	BURL971-G4-CHEM-MIX
BTV.R003	BTV Reader 3 Decon	BURL	A	No EPT	BURL971-D4-DECON
BTV.R004	BTV Reader 4 A11 Cyl Dock	BURL	A	No EPT	BURL974-A11-CYL-DOCK

(A)=LOCATION_ID, (B)= DESCRIPTION,
(C)= CONTAINER_LOCATION_ID, (D)= CONTACT_NAME,
(E)= LOCATIONIN, (F)= LOCATIONOUT.

5.3 Data analysis and data mining

The biggest hurdle in analyzing IBM data was the inconsistency in data gathering, the data had many missing fields, messy columns, incorrect information, which is of course expected since it's a pilot project. The first step in analyzing the data is to carry out a comprehensive data massaging; in the next step we converted Unix datetime stamp to windows date for further analysis. We did some preliminary analysis on the data set, some of which cannot be revealed due to confidentiality agreement with IBM. Some highlights of the analysis are as follows:

1. One reader is the dominant reader i.e. one reader scans most number of parts. It is found that Reader 1 is the dominant reader, almost all the tags are read by reader 1

2. On an average each part was scanned abut 18 times, i.e., number of scans between the entry of the part into the system and exit is about 18.

3. We also calculated flow time of a part through the system. Flow-time of the part through the system is calculated by taking the difference between the first and the last RFID read. By flow time we mean the flow time through the RFID, the part could remain in the system beyond the last RFID read but will only be scanned by barcodes. We found that average flow time is about 44.1 sec. The flow time of a part can be correlated to product demand or the rate of product handling.

4. Last but not the least we tried to get an estimate of RFID read accuracy. A recent article on e-business news reported that METRO group in Germany have reported RFID read rates of around 99 percent (this involves

tags, hardware and middleware together). The objective in this part of the analysis is to calculate the RFID read accuracy for this pilot project. The methodology for calculating the read rates is as follows: RFID accuracy is obtained by comparing the rawdata against processed data. We found the RFID accuracy to be in the range of 75%.

5.4 Lesson learned

The biggest hindrance in carrying out a comprehensive data analysis was the fact that the data was gathered without any specific objective in mind; the primary objective as we understand was to make RFID work; therefore the data set is not structured in a manner to carry out analysis specific to RFID. There are many instances when the logic behind RFID fails. Our aim through this case study is to illustrate how RFID works in a manufacturing environment. This analysis is by no means data intensive; it is rather focused on the understanding the process behind making RFID work effectively. Tables 2 and 3, for example, depict the movement of container *BURL200602432*. It is shown that

1. The container left Reader 4 (location id: BTV.R004) on the cylinder dock so BURL-IN-TRANS is a correct RFID read.

2. It arrived at the Reader 17 (location id: BTV.R017) on the M10-CHEM-Dock so that transaction is also correct.

3. Then the Solvent Room Reader 16 (location id: BTV.R016) read the tag because someone opened the door to that room which adjoins the M10- CHEM-Dock. The correct location should have been M9-SOLVENT-ROOM if RFID was working properly.

4. In the end reader 10 (location id: BTV.R010) did its job and moved the container from L9 Gas Room to the M10-Dock. At this stage some employee put the tags into the sleeves on the cylinders and they rest against the metal so Reader 17 which should put container in transit and Reader 4 which should receive it back at the Cylinder dock never saw the tag. This shows the potential RFID problem when reading it on metallic surfaces.

5. The last move transaction comes from a handheld barcode scanner. Hence RFID works but extra effort has to be spent on making RFID work.

Another limitation encountered in the project was the limited understanding of the process behind the data set. For example, it is difficult to establish relation between readers and their locations without understanding the building where readers are placed.

For carrying out a more effective analysis data needs to be gathered with some objective in mind. Two most commonly cited uses of RFID are supply chain visibility and theft reduction. Besides these RFID has been successfully implemented in improving supply chain efficiency. The data has to be structured to target one of these objectives only then we will start to see tangible advantages of RFID data.

6. Intelligent Applications of RFID

There is a general widespread belief that RFID can be handled in the same way as barcode, although there are a few similarities between the two technologies, there are a great number of differences to. Both RFID and Barcode are identification technologies that hold data that is accessed by some type of readers. In reality both RFID and barcode could complement each other well and be used in many applications side by side. Barcode is particularly limited in its application, as it needs line of sight and the data processing rate is very slow. Also barcodes are susceptible to damage and abrasion which can lead to *no-read*. Barcode technology is a read-only technology and so it cannot provide a real-time locating system that can monitor the inventory in the pipeline, whereas with RFID a line of sight is not required and the tags can be read through wood, plastic and any other material except metal.

There are different RFID technologies available today which can be used in a wide variety of applications; depending upon the application one or more technology can be applicable in a particular situation. For example, a retail store could use passive tags with handheld readers, while a warehouse could use active tags to store information about location of expensive and fast moving items, the real-time location of inventory would need a more expensive system with powerful antennae and tags. Through extensive research in the area of RFID application we found the use of RFID is on an ever increase; from supply chain management to hospitals to animals and forests. Below is a summary of some of the applications of RFID (Texas Instruments 2002); by no means exhaustive this summary is designed to give the readers a flavor of the potentials of RFID (Angeles 2005, Doyle 2004, RFID Journal, Finkinzeller 2000).

6.1 Applications in warehousing and distribution

Picking and Sortation System within a Warehouse

Sanacorp a German pharmaceutical wholesaler has a service policy where the customer order is ready within one hour of receiving it; to expedite the fulfillment process and to increase customer satisfaction, Sanacorp started using RFID tags on the totes which travel on a conveyor. The antennae are located under the conveyor belt at picking locations; when an order is initiated, the unique ID of the tote is associated with that specific order in the database. As the tote goes through its fulfillment route, the ID code indicates 'stop' or 'go'. The system can also be used to tracks orders, and detects bottlenecks. A central database coordinates all processes. The general benefits to Sanacorp include reduction in maintenance and reduction in the error rate.

Tracking Material Movement In and Out of a Warehouse

Chevrolet Creative Services uses a *red light, green light* system based on RFID to control and track the 3,500 crates coming and going from their Wixom, Michigan storage warehouse. The crates contain materials needed for specific tradeshows. RFID tags are mounted on each crate carrying a unique ID. The bay doors are equipped with RFID readers and readout antennas are located in the floor. When a crate passes over the antenna on its way to being loaded onto a truck, the tag ID is compared to a manifest held in a host database. In case of a match, a green light signals *go* to ship the crate. A *no match* activates a red light. The benefits include reduction in errors and improvement in speed, efficiency and productivity and at the same time better asset utilization and elimination of emergency shipping.

6.2 Applications in manufacturing

Improving Chip Manufacturers Yield

A number of semiconductor companies, like Motorola, SGS Thomson etc. use RFID in their clean rooms to gain control, improve quality and operator efficiency; by using RFID tags on the wafers (a part), the accuracy in processing of the wafers is greatly improved. Both wafer carriers and employees have RFID tags; at every only the correct combination of the tags would let the wafers pass forward. If there is not a match, then a warning signal emits and the equipment cannot be started until the correct lot is in place. The system also monitors the use of equipment to detect bottlenecks or inefficiencies. The advantages of such a system is higher capacity utilization, elimination of losses due to wrong operation sequence on a wafer lot, lower inventory costs, increased quality and process optimization and removal of bottlenecks.

6.3 Applications in airport security

RFID has found incredible use in security for baggage checking at the airports. *San Francisco airport* is using RFID tags to track and manage bags. The airport makes sure that the bags have gone through all the right security steps. A similar application is being installed at *London Heathrow Airport*.

6.4 Applications in retail

RFID tags keeping the shelves stocked (synchronized supply chains)

Wal-Mart has begun experimenting with individually tagged products to track their movements from manufacturer to warehouse to store aisle. This system would help in better demand planning and produce following advantages, reduction or elimination out-of-date stocks, decrease in the order lead time, higher inventory turns and decreased safety stock.

To summarize Table 5 further gives instances of RFID applications in various industry sectors. Alien Technologies, ConAgra Foods, Checkpoint Systems, CHEP, Defense Commissary Agency, Gillette, Intel, J.C. Penney, Johnson & Johnson, Kraft Foods, Kroger, Matrics, Philips Semiconductors, Procter & Gamble, Shaw's, Sensormatic Electronics, Symbol Technologies, Tesco, ThingMagic, Unilever, and Wal-Mart. The implementation of RFID system for any application needs careful u nderstanding of the current business processes and the desired role RFID would play to enhance them. This makes it necessary to not only understand the functionality of the technology but also choose the right technology which will build the desired capability to the current system while generating positive ROI.

Table 5: Summary of RFID applications

Industry Sector	Characteristics	Use of RFID
Retail (Wal-Mart, Mark & Spencer, Target, Tesco)	Misplaced products, inventory management and difficulties in on-shelf availability	Increased product visibility, Responsive production and Improved forecasts
Manufacturing (Ford, Honda, Erricson)	Managing the production philosophy (JIT, Kanban), getting the product out at right time	Improved accuracy in production, improved life cycle management, improved real-time inventory management
Warehouses (International Paper)	High labor costs, low productivity, Cross-docking	Increased speed of tracking parts, increased labor productivity, increased throughput
Health Care (Lancaster Hospital, CVS pharmacy)	Tracking patients, paper work reduces productivity, drug counterfeiting costs million of dollars	RFID tags on patients can track their movement and history, reduces counterfeiting by increased traceability

7. Conclusion

The paper gives an overview of RFID technology and its impact on supply chain The key to success of RFID doesn't lie in the technology; the key to success of RFID lies in the value that RFID data can generate (Delen, 2007). We identify several RFID data mining challenges and illustrate through an IBM case study the usefulness and potential problems in analyzing RFID data. Finally we surveyed some successful applications of RFID and put them in different categories to illustrate the widespread nature of RFID. Our work on RFID is by no means complete; as a matter of fact this is just the beginning.

7.1 Future work

Our research work is still at infancy. Currently we are investigating several issues regarding RFID data.

- A comparative analysis of various different network architectures in their suitability under various supply chain integration strategies for supporting the information retrieval based on Electronic Product Code (EPC) — the data contained in RFID tags.
- A set of policies for the visibility of RFID data based on the overall integration strategies of a supply chain.
- Various aggregation and abstraction data models for storing, querying, retrieving, and transmitting RFID data.

Acknowledgements

Jack Lee is a truly scholar and educator. The authors would like to acknowledge the impact of Jack's influence to our work. It has been a great pleasure to collaborate with Jack lee on organizing various conferences/programs, sharing his wisdom and principles in making final decisions. His untimely death was a dreadful loss and he will be missed in our society. We would also like to thank Dr. Jun Shu for his comments on the paper.

References

Agrawal, R, Gunopulos, D., Leymann, F. Mining process models from workflow logs. (1998). *Proceedings of the Sixth International Conference on Extending Database Technology*, 469-483.

Angeles, R, (2005). RFID technologies: Supply-Chain Application and Implementation Issues. *Information Systems Management* **22**, 51-65.

Atkinson, W, (2004). Tagged: The risks and rewards of RFID technology. *Risk Management* **51**, 12-18.

Auto-ID Center. (2002). Technology Guide, Auto-Id Center. www.autoidcenter.org.

Chawathe, S. S., Krishnamurthy, V., Ramachandran, S. and Sarma, S. (2004). Managing the RFID data. *Proceedings of the 30th VLPB Conference*. Toronto, Canada.

Delen, D., Hardgrave, B. and Sharda, R. (2006). RFID for better supply chain management through enhanced information visibility. Forthcoming in Production and Operations Management.

Doyle, S, (2004). Auto-ID technology in retail and its potential application in marketing. *Journal of Database Marketing & Customer Strategy Management* **11**, 274-279.

EPCglobal Inc. www.epcglobalinc.org. Excerpts from the website.

Finkinzeller, K., (2000). *RFID Handbook, Radio Frequency Identification Fundamentals and Application*. John Wiley.

Gonzalez, H., Han, J. and Li, X. (2006). Mining Compressed Commodity Workflows from Massive RFID Data Sets. *Proceedings of the 2006 Conference on Information and Knowledge Management (CIKM'06)*. Arlington, Virginia, November 2006, 162-171.

Hardgrave, B. C. and Miller, R. (2006). The Myths and Realities of RFID. *International Journal of Global Logistics and Supply Chain Management* **1**, 1-16.

Lee, M., Cheng, F. and Leung, Y. T. (2004). Exploring the impact of RFID on supply chain dynamics. *Proceedings of the 2004 Winter Simulation Conference*.

Lin, Dennis K. J. and Wadhwa, Vijay (2008). Efficiency gets a new identity: Barcode and RFID in the supply chain. *Quality Progress* **41**, 20-24.

Rappold, J. (2003). The risks of RFID. *Industrial Engineer* **35**, 37-38.

Raza, N., Bradshaw, V. and Hague, M. (1999). Applications of RFID technology. *IEE Colloquium on RFID Technology* **1**, 1-5.

RFID Journal, http://www.rfidjournal.com, Excerpts from the website http://www.rfidjournal.com/article/articleview/207, accessed 7/06/08.

Stockman, H. (1948). Communication by means of reflectal power. *Proceedings of Institute of Radio Engineers*, 1196-1204.

Texas Instruments, (2002). Radio Frequency Identification Solutions. Excerpts from the website, http://www.ti.com/rfid/.

Received December 29, 2007; accepted May 4, 2008.