Optimizing Operational Efficiency and Enhancing Data Reliability using Effective and Adaptive Cleaning Approach for RFID in Healthcare

Anny Leema.A
Assistant Professor, Department of Computer Applications,
B.S.Abdur Rahman University, Vandalur, India
Research Pscholar , Karpagam University,Covai.

Dr.Hemalatha.M
Head, Dept of Software Systems,
Karpagam University, Coimbatore, India

ABSTRACT:

RFID technology with the special ability of multiple object identification without the constraint of line-of-sight provides a perfect way for real time data collection and object identification. The healthcare industry involves critical activities where small mistakes could cause huge loss of life and incur massive financial losses. The data captured by RFID readers are usually of low quality and may contain many anomalies. Data quality has become increasingly important to many organizations. In order to provide reliable data to RFID application it is necessary to clean the collected data. SMURF is a declarative and adaptive smoothing cleaning technique for unreliable RFID data. However it does not work well when tag moves rapidly in and out of reader's communication range. The errors need to be cleansed in an effective manner before they are subjected to warehousing. Factors such as inter tag distance , tag-antenna distance, number of tags in the read range of antenna, reader communication range, velocity of tag movement affect the data cleaning result. Our proposed algorithm considers these factors and also the missing tag information, tags that are mistakenly read as present dynamically in determination of the size of slide window. Thus with the aid of the planned data cleaning technique we can bring down the health care costs, optimize business processes, streamline patient identification processes and improve patient safety.

KEYWORDS:

RFID technology, significance of data quality, RFID middleware systems, cleaning methods, Data cleaning approaches

1. INTRODUCTION

Radio Frequency Identification (RFID) is a means to identify and track objects using radio frequency transmission. An RFID system consists of readers and tags. Readers use radio signals to communicate with the tags. Tags may be active (battery powered) or passive (powered by the reader's signals). A tag consists of a microchip applied to a miniaturized antenna, which carries a "unique tag ID" but can be programmed with other information. The reader captures the signals, decoded it and sends it to host computers. In both the reader and tag, the antenna can be sized and shaped in different ways. Because of the small size of the tag, RFID tag can be designed to fit almost any situation. Placement of the tag and reader is now no longer critical since there is no contact or line of sight is required; the RFID system allows great freedom of movement. RFID includes hardware, middleware and software components.

This technology is widely used in diverse application such as supply chain automation, asset tracking, medical/Health Care applications, people tracking, Manufacturing, Retail, Warehouses, and Livestock Timing. Of these, Medical/Health care applications are of more importance because minute errors in it can cost heavy financial and personal losses. For hospitals and healthcare systems, increasing the operational efficiency is the primary target. It is a tough task to keep up the effectiveness and monitor each and every patient [2]. However, utilization of RFID (Radio Frequency Identification) technology in addition to reducing the health care costs facilitates automating and streamlining patient identification processes in hospitals and use of mobile devices like PDA, smart phones, design of health care management systems etc.,

2. RFID IN HEALTHCARE-ADVANTAGES

Healthcare management system has several RFID applications. Drug faking applications give a unique code and identity to medicine by placing RFID tags [3]. Fake medicines can be easily distinguished in case of illegal distribution. By using RFID tags, the expired medicines can also be kept under control. The information about the status (cleaned/sterilized/unsterilized) of medical devices can be obtained by supplying those devices with RFID tags and placing RFID readers at the entrance of storing chambers. Thus, assists to avoid devastating errors that may occur on utilizing contaminated instruments. The RFID tags can help saving time by providing the exact location of the medical devices inside the rooms/chambers. The seal and packaging of medicines may be uncertain. Read/write RFID tags can be extremely useful in resolving these problems. By placing the information related to the seal time, packaging date etc [1] on the RFID read/write tags they can be examined whenever required. It is also used to improve the security of a hospital by issuing tags for both employee and patient so that when a restricted area is entered an alarm would be triggered.

By monitoring and recording the blood pressures and heart beat rates, the RFID tags can also convey information about the health status of patients [6]. This can assist the nurses/caretakers to have a better control on the patient's health and also enable the faster recovery of patients. RFID technology can be formed as an essential part of healthcare.

The risk factors of human error are reduced by RFID in the healthcare industry [5].

3. WAREHOUSING RFID DATA

Data warehousing is the only viable solution for providing strategic information. In fact, they are very large repositories that integrate data coming from operational databases of several enterprise sectors for decisional analysis. There are three areas of software research into management of RFID data (as illustrated in figure1). The first area deals with the secure fetching and management of online tag associated information. The second area is related to removal of errors that are present in RFID because of error inaccuracies. The third area is concerned with providing OLAP operations over huge RFID data sets by constructing suitable multi-dimensional warehouse [4]. The major challenge of data warehousing is data cleaning [1] and our work focus towards it. Data cleaning is necessary for improving the quality of data so that it becomes "fit for use" by users.

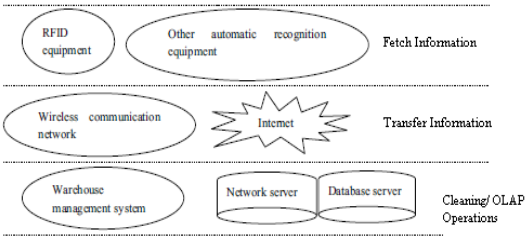


Fig.1 Three steps in management of RFID data

4. SIGNIFICANCE OF DATA QUALITY

Data quality has become increasingly important to many firms as they build data warehouses and focus more on customer relationship management. For health care organizations, data is central to both effective health care and to financial survival. Data quality was concerned with accuracy, precision and timeliness. Recently the Institute of Medicine shocked the public with a report that 98,000 people die every year due to medical errors [3]. Some of the errors are the result of missing or bad information about drugs, orders and treatments. Poor data quality has adverse effects at the operational, tactical and strategic levels of an organization. At the operational level, poor data reduces customer satisfaction, increases operational costs and reduces employee job satisfaction. Quality of data in the data warehouse is dependent on the quality of data in the various source systems and on the quality of the extraction, cleansing, transformation and transfer processes that make up the source- to- target transformation. Therefore data quality problems are identified and steps can be taken to monitor and improve the quality.

5. RFID MIDDLEWARE SYSTEMS

Middleware becomes a very hot topic. The observed read rate in real-world RFID deployments is the range 60-70% and 30% of the tag readings are routinely dropped. In order to correct reading errors, and allow data streams to meet the high level information requirements, we deploy the RFID middleware system between reader and applications. The concept behind this middleware is using a non-overlapping static window or an overlapping sliding window RFID approach which interpolates dropped readings for every tag within the time window on the collected data stream. The standard data-cleaning mechanism for today's RFID middleware systems is a temporal "Smoothing filter". The goal is to reduce or eliminate dropped readings by giving each tag more opportunities to be read within the smoothing window and then providing data to a higher level end application.

SELECT distinct tag_id FROM RFID_stream [RANGE ' 3 sec'] GROUP BY tag_id

In order for RFID technology to become feasible, RFID middleware must be able to produce reliable streams describing the physical world.

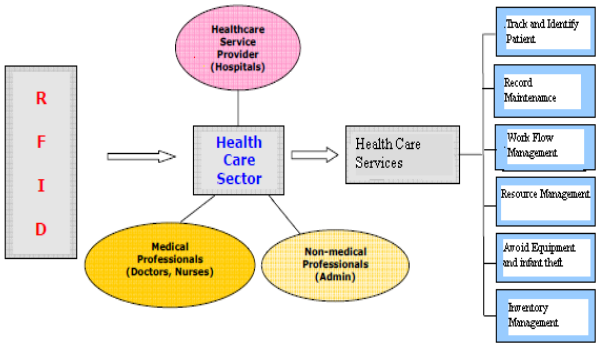
6. Hospital/Healthcare takes off with RFID

RFID is an amazing technology holds an unlimited number of benefits within healthcare industry. In this section first we describe how simulation can be done; next we describe how to build an elegant hospital environment and finally the provided healthcare services.

6.1. Simulation in Healthcare

Simulation has long been used as a decision support tool in various sectors. It is especially suited to the analysis of healthcare organizations due to its ability to handle high complexity and variability which is usually inherent in this sector. Experimentation of different workflows, staffing decisions and what-if analysis are all promising applications of simulation in healthcare. However, a typical simulation study requires deliberate data collection effort over a considerably long period of time. Fig [2] shows our proposed health care services.

Fig. 2 Healthcare Services



In order to provide the above said health care services the followings are required:

Issue of medical card attached RFID tag.

The collection and store of much kind of data.

The extraction of individual ID through RFID reader for abstracting individual information from RFID tag to enter within a limited area.

The acquisition and verification of security information.

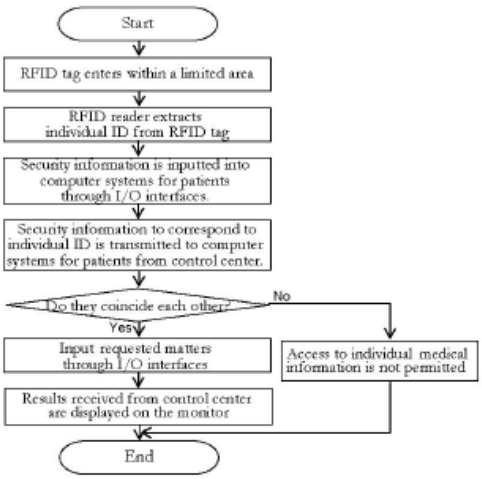
The receipt of requests through input/output interfaces.

The search of requested data according to extracted individual ID

The provision of searched data.

The following figure depicts the guidance of Medical Service.

The flowchart [1] depicts the mode of services for out patients.



Flowchart [1] Services done for out patients

7. RFID DATA CLEANING

7.1 Data Cleaning

The major challenge of data warehousing is data cleaning. It is necessary for improving the quality of data so that it becomes "fit for use" by the users.

7.2 Issues in Data Cleaning

Lack of Completeness

RFID readers capture only 60-70% of all tags that are in the vicinity.

Smoothing of data is done to rectify the loss of intermediate messages.

Temporal Nature of data or tag dynamics

RFID tags are in motion and that is what makes them more difficult to handle.

But motion of a tag causes dropping of messages.

RFID data streams are very fast and are huge in number

Hence filtering is important before sending them to database

7.3 Types of Errors

The errors occur in the process of data capture are false negatives, false positives, and duplicates. False negative is that tags are in the vicinity of a reader but not detected by it. False positive refers to that a tag is not present but captured. Besides RFID tags to be read, additional unexpected reading are generated. Duplicate readings refers to tags are in the scope of a reader for a long time and are read by the reader multiple times. It also occurs due to the tags in the overlapped areas read by multiple readers

8. DATA CLEANING APPROACHES

The approaches in RFID data cleaning are studied by many authors. They are fix-size slide window , EPS (Extensible sensor stream processing) , and SMURF (Statistical smoothing for unreliable RFID data). Fix-size slide window approach uses the time window to determine whether the data is right or not. In EPS (Extensible Sensor Stream Processing) the static size of the window is the limitation of the approach because large window induces false positives and small window cannot fill false negatives. ESP organizes receptor stream cleaning into a cascade of five programmable stages: Point-Smooth-Merge- Arbitrate-Virtualizes. ESP captures the context of temporal and spatial application layers by introducing the concepts of temporal and spatial granularity. SMURF is introduced to reduce the reading dropping rate on the statistical basis, and incorporating cleaning logic in applications produce more reliable RFID data stream. It dynamically adjusts the size of the window to pre-treat the RFID data. However SMURF does not work well when tag moves rapidly in and out of reader's communication range. Through an adaptive smoothing filter for cleaning raw RFID data streams, this greatly increases the complexity of the application. To ensure that all tags are in the reader's detection range are read, the smoothing window must be large enough to correct for reader unreliability. Small window sizes will easily omit some tags, causing a false negative. On the other hand, adopting large windows will limit the detection of the tag's movements (or reader's movement) resulting in a false positive (Jeffery et al., 2008).

SMURF has weakness in dealing with the size of slide window. Another weakness is determining the parameter delta. SMURF assumes that tag does not move frequently. When it frequently moves in and out of a reader's communication range, SMURF does not work well and it leads to higher error rate. The table no.1 depicts the comparative study of the existing cleaning approach and our proposed technique.

S.No	Title	Description
1	SMURF-RFID data cleaning based on adaptive window	SMURF is one of the recognized data cleaning approaches and it does not have good performance when tag moves rapidly in and out of readers' communication range.
2	RFID data cleaning based on Adaptive Window	This algorithm is better than SMURF for large tag moving speed. Error rate is lower than SMURF but not fully removed. False positive errors are also not removed using this algorithm.
3	bSpace: A Data Cleaning Approach for RFID Data Streams Based on Virtual Spatial Granularity	A new "smoothing filtering" approach named bSpace is proposed, which is based on the concept of virtual spatial granularity. bSpace uses a Bayesian estimation algorithm to fill up false negatives, and uses the rules which we define to solve false positives. Complexity is high.

4	Optimizing Operational Efficiency and Enhancing Data Reliability using effective cleaning approach for RFID in Healthcare	SMURF (Statistical Smoothing for Unreliable RFID data) does not work well when tag moves rapidly. To solve the problem, we propose an improved algorithm based on adaptive window. Factors, such as reader communication range, reading frequency, velocity of tag movement, affect the data cleaning result. Our new algorithm considers these factors dynamically in determination of the size of slide window. A new method is also proposed to fill data. Simulation shows our approach deals with RFID data more efficiently and accurately.

9. PROPOSED METHODOLOGY

Obviously, to ensure a tag can be captured, window size must be large enough to fill missing readings. On the other hand if the window size is too large, the data is filled incorrectly and results in false positives. In real world many factors that affect the reliability of data. They are Inter tag distance, the orientation of the tags with respect to the antenna, tag - antenna distance, number of tags in the read range of antenna and the speed of the tagged objects etc. The proposed technique will consider missing tag information, tags that are mistakenly read as present and also the above said factors. The cleaning is planned to perform in an adaptive manner. As it is made adaptive, the cleaning operation will automatically adjust itself to focus more on more erroneous data and less on less erroneous data. By performing this adaptive kind of operation, the computational complexity is expected to be less. Hence, an effective cleaning will be accomplished for the RFID data. This will pave the way for an effective means of data warehousing system that will keep the RFID data safe for future mining..

10. CONCLUSION

RFID plays an essential role in all the sub-domains of the applications in health care applications. Among them, RFID technology dominates in tracking the patients under treatment. But, there may be errors and redundancies in the obtained RFID data from all the readers. The effectiveness in cleaning the RFID data in healthcare sectors remains a concern, even though a number of literary works are available. The probability of occurrence of the error due to the wrong consideration about the presence of the tags is high. This happens because of the electronic devices that are operating in the same RF range. The RFID readers may read an unavailable tag because of the interference generated by the malicious electronic device. This gives rise to erroneous reading by the corresponding RFID reader. To a maximum, the dirty data that are read because of these errors may even leads to patients' death. The errors need to be cleansed in an effective manner before they are subjected to warehousing. SMURF is one of the recognized data cleaning approaches. However it does not have good performance when tag moves rapidly in and out of reader's communication range. In this paper, we have proposed an improved algorithm which considers the factors such as reader communication range, tag antenna distance, number of tags in the read range of antenna and the speed of the tagged objects etc. Thus the errors are cleansed in an effective manner before they are subjected to warehousing and it will pave the way for an effective means of data warehousing system that will keep the RFID data safe for future mining.

11. REFERENCES

[1] Sudarshan S. Chawathe, Venkat Krishnamurthy, Sridhar Ramachandran, and Sanjay Sarma, "Managing RFID Data", In Proceedings of the 30th VLDB Conference, pp.1189-1195, 2004.

[2]. A. Lahtela, "A Short Overview of the RFID Technology in Healthcare", Systems and Networks Communications, 2009, ICSNC '09, In Proceedings of Fourth International Conference on, pp.165-169, 2009

[3] O. Shoewu and O. Badejo, "Radio Frequency Identification Technology: Development, Application, and Security Issues", The Pacific Journal of Science and Technology, Vol.7, No.2, November 2006.

[4]. Panos Vassiliadis,Christoph Quix,Yannis Vassiliou and Matthias Jarke, "Data warehouse process management", In proceedings of the 12th international conference on advanced information systems engineering, Vol. 26,Issue 3, pp.205 - 236, 2001.

[5] J. Fisher and T. Monahan, "Tracking the social dimensions of RFID systems in hospitals", In Proceedings of International Journal of Medical Informatics, Vol. 77, Issue 3, pp. 176 -183, 2007.

[6]Bettina Fazzinga, Sergio Flesca, Elio Masciari and Filippo Furfaro, "Efficient and effective RFID data warehousing", In proceedings of the 2009 International Database Engineering & Applications Symposium,pp. 251-258 , 2009

AUTHORS BIOGRAPHIES

A.Anny Leema completed MCA., MPhil., and pursuing PhD in Karpagam University. She is working as an Assistant professor in the department of MCA in B.S.Abdur Rahman University. She has ten years of experience in teaching. She has presented twelve papers in National conferences and two in International conference. She has published three papers in International Journals.Her area of interest is Data Mining, Cybercrime and E-learning. She is currently doing her research in the warehousing RFID under the guidance of Dr. Hemalatha., Assistant Professor and Head , dept of software systems in Karpagam University.

Dr. Hemaltha completed MCA., MPhil., PhD in Computer Science and currently working as a Asst Professor and Head , dept of software systems in Karpagam University. She has ten years of experience in teaching. She has published several papers in International journals and presented seventy papers in National conferences and five international conference .Her area of research is Data mining, Software Engineering, bioinformatics, Neural Network. Currently she is working to improve the performance of Spatial Datasets in Databases. She also act as reviewer in several journals .