http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (21): 6044-6049, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.6044.6049

An Extensible Model for Uncertain RFID Data in Supply Chains

[1,2]Dong Xie, [3,4]Song Tang, [2]Jie Xiao, [3]Huijun Wang and [1]Jianbo Wang
[1]Department of Computer Science and Technology, Hunan University of Humanities, Science and Technology, Loudi, China
[2]Department of Information Science and Engineering, Hunan First Normal University, Changsha, China
[3]Department of Network Media, Hunan Mass Media Vocational Technical College, Changsha, China
[4]Institute of Network Media and Public Sentiment,
Hunan Mass Media Vocational Technical College, Changsha, China

Abstract: Radio Frequency Identification (RFID) technology is widely used to trace objects. However, RFID application systems can not effectively and efficiently use massive uncertain data. This study considers some properties of objects captured by sensor and GPS and proposes a comprehensive extensible model for uncertain data according to key features of RFID data which is suitable to different application scenarios. The model can effectively and efficiently store different RFID data according to key features of RFID data and supports a variety of queries for tracking and tracing RFID objects.

Key words: Supply chain, RFID, relational database, uncertain data

INTRODUCTION

RFID technology (Nath *et al*., 2006) is a flexible and relatively low-cost solution for tagging and wireless identification in a large number of businesses. Every tagged RFID object has a unique identifier which can be automatically identified and collected to generate massive data by RFID readers. In the recent years, the technology has become increasingly popular in manufacturing industries such as Supply Chain Management (SCM) (Ilic *et al*., 2009). In supply chains, RFID applications cross multi-enterprises such as manufacture, wholesaler, retailers and customers and obtain tagged RFID products including their specifications and positions and so on. RFID applications are available for users to track and trace historical trajectories and concrete positions of RFID objects over the Internet according to their related position information denoted as data lineages. This will significantly affect our life styles.

Despite different type of RFID applications, their general features are same. Information of tagged RFID objects are read in certain time and concrete positions by certain RFID readers. Generally, tracing applications focus on concrete positions at different time, no information of RFID readers. Although most of RFID data are precise, RFID applications still face a significant challenging problem for managing RFID data. Since RFID devices are the intrinsic sensitivity to environmental factors, RFID readers might miss objects' certain positions and frequently read and create massive redundant data such as unchanged object positions. These readings significantly complicate the tasks of determining object position. We summarize types of RFID data as the followings:

- **Insufficient data (InsD):** Since RFID is an identification technology for identifying individual objects; this might not capture containment relationships of inter-objects
- **Missing data (MisD):** RFID readings are noisy with observed read rates below 100% in actual deployments. Missed readings result in lack of objects' certain positions. RFID objects without RFID data will lose in the RFID application system. This means that missing data need be inferred by related algorithms according to some conditions
- **Inconsistent data (IncD):** Similarly, since radio frequencies exist the intrinsic sensitivity to environmental factors, different RFID readers may obtain different values for same RFID objects. Inconsistencies inevitably exit in RFID data and these inconsistent data violate integrity constraints such as Functional Dependencies (FDs) and reference dependencies. It is difficult to identify a correct value for a RFID object with different values. Data cleaming technologies can remove noisy data to store certain data. However, the technologies are semiautomatic and they need human intervention

Corresponding Author: Song Tang, Hunan Mass Media Vocational Technical College, China

- **Redundant data (RedD):** As wired RFID readers read frequently in powered environments, they can create massive data. However, the stored data may include significant amounts of redundant information such as unchanged object positions. Therefore, the information need be filtered
- **Ghost data (GhoD):** Since radio frequencies in reading areas might cause that RFID readers obtain inexistent RFID objects

Generally, insufficient and missing data are main rates in above RFID data. Since RFID data are very massive, it is difficult to fill these data into RFID application systems. These data need be automatically inferred by related algorithms. In addition, since ghost data only exist in a single RFID application, it is necessary to differ from missing data. Redundant data collected by RFID readers from RFID objects are easy to be filtered useless data by identifying their unique identifiers.

These above data lead to two main problems: (1) RFID applications generally focus on concrete positions of RFID objects at different time rather than information of RFID readers; the users' queries cannot be directly operated on uncertain RFID data without preprocessing; (2) The traditional entity-relationship data model of relational databases might not effectively and efficiently process uncertain data for expressing RFID objects' data linage; a common storing RFID data model employs a triplet (EPC, reader-id, timestamp) which is difficult to store and process the above data. It is necessary to model uncertain RFID data under a unified model and to preprocess the uncertain data in different databases of different logistics enterprises.

In this study, we summarize key features of RFID applications and classify entities of RFID applications and distinguish several data such as static data, reading data and generating data. We propose a comprehensive extensible data model which is suitable to different application scenarios. The model considers some properties of objects captured by sensor and GPS. The model can effectively and efficiently store and query RFID data according to key features of RFID data.

RELATED WORK

In recent years, a few researches about the data model have been done on how to process the uncertain RFID data (e.g., RFID data cleaning technique) (Jeffery *et al*., 2006). However, few researches have been done on how to effectively model uncertain RFID data and efficiently processing various uncertain data.

It is difficult to model RFID data by using the traditional ER-model due to RFID data are temporal. Dynamic Relationship ER Model (DRER) (Wang and Liu, 2005) discusses two types of history data in the scenario of fixed readers: event-based and state-based. Applications should be further classified into a set of basic scenarios and develops constructs for modeling each scenario according to fundamental characteristics of RFID applications (Wang *et al*., 2010).

Moreover, it is inefficient to support path queries which may perform multi-self-joining of the table involving many related positions. Gonzalez *et al*. (2006) focus on group of objects and uses compression to preserve object transition relationships for reducing the join cost of processing path selection queries. They further develop a gateway-based movement graph model as a compact representation of RFID data sets (Gonzalez *et al*., 2010). To single objects, a movement path of an EPC-tagged object is coded as the position of readers and the order of positions denoted as a series of unique prime number pairs (Lee and Chung, 2011). However, the method only supports the product of the first 8 and 15 prime numbers respectively for 32 bits and 64 bits. The method doesn't code more long paths which might cause the overflow of the prime numbers. Moreover, the method is not applicable to express cyclic paths because an object may have the same position with different time intervals.

The above methods do not adequately considered on uncertainties of RFID data. Traditional queries should be redefined as specially queries over inconsistent data by Consistent Query Answer (CQA) (Arenas *et al*., 1999), but redefined queries might eliminate inconsistent data and don't track and trace objects with inconsistencies. Probabilistic Consistent Query Answer (PCQA), Lian *et al.* considers all possible repairs and possible worlds (Lian *et al*., 2010) and assumes that data are independent, but different RFID objects may be at a same position. Chen *et al.* (2010) proposes each internet of things (IoT) object using 7-tuples of the form <EPC, instance_id, attribute, value, time, probability, reliability> which provides a neat representation of IoT data in the form of probabilistic streaming data, but the work doesn't consider key features of RFID applications such as the position relationship.

PRELIMINARY

RFID-based supply chains: All readings are automatically read by RFID readers in the whole process of the RFID-based supply chain. Firstly, each object is tagged

Electronic Product Code (EPC) in the production line of the manufacturer and its product specifications are written into tags; then EPC-tagged objects are packed into cases in manufacturer warehouses. These cases are packed onto pallets at the manufacturer warehouses, where the EPC tags of both these cases and these containing pallets are read by RFID readers; next, pallets are loaded into trucks which depart manufacturers to dealers or retail stores. Pallets are unloaded from the trucks and cases are unpacked from pallets in zones of retail stores; finally, EPC-tagged objects are purchased by consumers.

Specially, sensors may be attached on EPC tags of objects or positions. In most periods of the supply chain, sensors may measure environmental parameters of RFID objects such as temperature and humidity. These are helpful to protect objects. For example, foods will go bad when their temperatures are more than a certain value. In addition, GPS sensors might be attached on EPC-tagged trucks for measuring their positions in the transportational process.

Key features of uncertain RFID data: Several key concepts are considered as the followings: temporal and position. All readings are related with timestamps that are important to form historical trajectories which determine data lineages for tracking and tracing RFID objects. Position information of RFID objects determine historical trajectories according to their time sequences.

Temporal: Tracking and tracing RFID objects are mostly temporal oriented. When RFID readings happen, same RFID objects are associated with different timestamps in different positions; different RFID objects are associated with same timestamps in a same position.

Position: One key feature is position. RFID readings form a series of different relative position data as data lineage for same RFID objects which are a key to track and trace RFID objects.

Entities in RFID applications: Many fundamental entities are directly used to RFID applications (Wang and Liu, 2005; Wang *et al*., 2010) as the followings:

- **Object:** EPC-tagged objects with unique code may be items, carries, cases, pallets
- **Reader:** EPC-tagged readers with unique code may be fixed or moveable
- **Sensor:** Sensors of RFID tags measure some parameters (e.g., temperature or position) of RFID objects and write measurements to their master RFID tags
- **Position:** Positions of EPC-tagged objects may be physical (e.g., point measured by GPS) or symbolic such as a warehouse, a distribution zone, retail market. A position may contain another

Moreover, we may summarize other entities by analysing the fundamental entities. These entities support activities of fundamental entities. Fox example, owners of positions and readers, properties and businesses of objects and types of sensors, etc.

Entity relationships: Entity relationships of RFID applications are similar to the traditional ER model which generates relational table according to correspondence between entities. However, inconsistent data that violate integrity constraints need be stored in RFID databases. This raises that relational table cannot completely employ the traditional method to create data model.

Since positions of RFID objects and reader are dynamically changed. Here we distinguish three types of data: static data, reading data and generating data.

Static data are stable and be rarely changed. The data are certain and consistent. Fox example, position and reader information are often unchanged but don't include missing data.

Reading data are directly obtained while RFID readers read RFID objects or sensors sense parameters. The data may be insufficient data, inconsistent data, redundant data and ghost data, but don't include missing data.

Generating data are tackled into some relational tables by related algorithms according to reading data. Since users focus on positions of objects at a certain time, so these data are very helpful for tracking and tracing. These data include missing data.

DATA MODEL

Data model: Our model will extend (Xie *et al*., 2012) for storing uncertain RFID data and tracing EPC-tagged objects (Agrawal and Biswas, 2012). The model will focus on a comprehensive expression for RFID data. There are several mapping types from the data model to relational tables according to common ER mapping rules. Entities are mapped directly as entity tables. A relationship is mapped as a table consisting of keys from among entities with the multi-multi relationship. A relationship from both entities without the multi-multi relationship is not mapped as a new table. Key attributes of master tables should be added into child tables as key attributes together key attributes of child tables (Fig. 1).

Static data include: READER, OBJECT, SENSOR and POSITION for fundamental entities and BUSINESS,

Table 1: Storing Uncertain Data for Different Stages in Supply Chains

Table		Deployment	Production	Warehouse	Transfer	Retail store
Tables for Static data	OBJECT		X	X		
	POSITION, READER, SENSOR, SENSORTYPE, OWNER, ITEMTYPE, BUSINESS	X				
Tables for Reading data	READING		X	X	X	X
	MEASURE_XYZ				X	
	MEASURE		X	X	X	X
Tables for Generating data	REA_BUS	X	X	X	X	X
	OBJ_POS		X	X	X	X

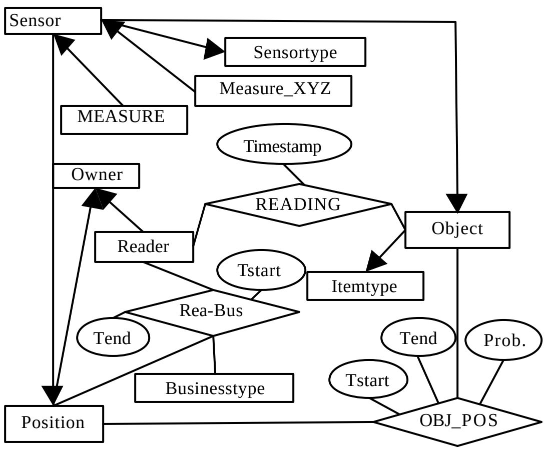

Fig. 1: Data model

ITEMTYPE, OWNER, SENSORTYPE for multi-value attributes of fundamental entities.
Reading data include: READING generated from OBJECT and READER, MEASURE_XYZ and MEASURE generated from SENSOR.

Generating data include: OBJ_POS generated from OBJECT and POSITION, REA_BUS generated from READER and BUSINESS. Plus an interval [tstart, tend] to express the period for a generating data or plus a timestamp to express the occurring time for reading data. In additional, in order to express uncertainties of RFID readers, positions of objects and sensors, we plus a prob attribute into REA_BUS, OBJ_POS and SENSOR respectively (Table 1).

Tables for static data: The READER (reader_id, name, owner_id) table stores EPC, name and owner ID of readers. A reader belongs to a certain owner. The OBJECT (object_id, item_id) table stores EPC and item type of objects. READER may join the OWNER table and OBJECT may join to ITEMTYPE.

The SENSOR (sensor_id, attachment, name, type, measurement_unit) table stores sensor ID, attachment with sensor, name, type and measurement unit of sensors. A sensor may be attached an object and they may be moved to other places together. Also, a sensor may be fixed at a certain position for checking parameters.

The POSITION (postion_id, name, parent_id, x, y, z, owner) table stores ID, name, parent ID of position, coordinates and owner of positions. The zone of the attribute parent_id includes the zone of the attribute postion_id. We can obtain the concrete position according to the concrete coordinates. The table need be added special attributes as continuous coordinates which might be identified by a position sensor (e.g., GPS). Similarly, a position belongs to an owner.

The table OWNER (id, name) stores owner information of readers and positions such as wholesalers or suppliers. The table BUSINESS (id, name) stores business types. Different readers may tackle different businesses such as packing and unloading. The table SENSORTYPE (id, name) stores type of sensors such as temperature or humidity. The table ITEMTYPE (id, name) stores type of item such as food or drug.

Tables for reading data: The table READING (reader_id, object_id, timestamp) stores raw reading data including reader's EPC, tag's EPC and the reading timestamp. MEASUREMENT (sensor_id, timestamp, values) stores measurement information of sensors at a certain time.

MEASURE_XYZ (sensor_id, timestamp, x, y, z) stores concrete coordinates at certain time according to the POSITION table's coordinate which is approximate to the coordinate of the MEASURE_XYZ table. The coordinate of POSITION should be correct if the coordinate of MEASURE_XYZ is not consistent with POSITION when the sensor is fixed,.

Tables for generating data: Generating data are the key to track and trace positions of objects for users.

The table REA_BUS (reader_id, tstart, tend, business, position) stores reader EPC, start time, end time and business of readers during certain time. There are special readers for every concrete transaction.

The table OBJ_POS (object_id, position_id, tstart, tend, prob) stores object EPC, position ID and period of readers for expressing the position histories. The attribute prob expresses EPC uncertainty at the position_id. Since uncertain data are missed, these data need be inferred their positions by related algorithms. READING stores data read by readers and the inferred data will be stored in OBJ_POS.

Querying uncertain RFID data: The model supports RFID queries with temporal constraints for tracking and tracing. We highlight that users used to want to query a concrete position and a time period for an object. We propose queries based on our model as the follows.

We will tracks the change history of an object's states and detecting missing objects. Q1 queries position histories of the object 'OBJ' from the table OBJ_POS. The attribute prob of 'OBJ' may be null. This means the value is not inferred and but it is inferred by related algorithms.

Q1. SELECT object_id, position_id, tstart, tend,prob
FROM OBJ_POS WHERE object_id ='OBJ'

Q2 queries the position while object 'OBJ' is lost according to the latest time (MAX (tstart)) of 'OBJ'. Similarly, the attribute prob may be null, so the tuple is certain. This means that the lost 'OBJ' doesn't be read by readers, so it need be inferred by algorithms.

Q2. SELECT object_id, position_id, tstart, tend,prob
FROM OBJ_POS
WHERE object_id ='OBJ'AND
tstart =(SELECT MAX(OBJ_POS. tstart)
FROM OBJ_POS
WHERE OBJ_POS.object_id ='OBJ')

Q3 queries moving time of the object 'OBJ' from position 'p1' to position 'p2'.

Q3. SELECT (a.tstart - o1.tend), prob
FROM OBJ_POS a, OBJ_POS b
WHERE a.object_id = 'OBJ'AND
b.object_id ='OBJ' AND
a.position_id ='p1' AND
b.position_id ='p2'

Q4 return objects in the last two hour.

Q4. SELECT object_id, prob FROM OBJ_POS
WHERE position_id = 'L04' AND tend = 'U'
AND tstart <= SYSDATE-(2/24)

Q5 queries the number of objects at the position 'L03' from 11/01/2013 to 31/01/2013.

Q5. SELECT count(object_id) FROM OBJ_POS
WHERE lid = 'L03' AND tstart >='11/01/2013'
AND tend <'31/01/2013'

CONCLUSION

This study summarizes several features of uncertain RFID data and proposes a comprehensive extensible model to manage uncertain RFID data which can effectively store different types of uncertain RFID data produced at different stages of supply chains. The model can support a few types of queries for tracking and tracing EPC-tagged objects in RFID applications. Our next work aims to compress uncertain RFID data.

ACKNOWLEDGMENTS

This study is supported by the Research Foundation of Education Committee of Hunan Province, China (09A046, 11B104, 12C0999, 13A046), the Hunan Provincial Natural Science Foundation of China (12JJ3057, 12JJ2040), the Hunan Science and Technology Plan of China (2010TT2055, 2011FJ3033, 2012GK3073, 2013NK3090), the Hunan Provincial Philosophy and Social Sciences Foundation of China (12YBA267), the Aid program for Science and Technology Innovative Research Team in Higher Educational Institute of Hunan Province and the Construct Program of the Key Discipline in Hunan Province, China.

REFERENCES

Agrawal, T. and P.K. Biswas, 2012. An optimized query tree algorithm in RFID inventory tracking-a case study evidence. Int. J. Comput. Sci., 9: 85-93.

Arenas, M., L. Bertossi and J. Chomicki, 1999. Consistent query answers in inconsistent databases. Proceedings of the 18th ACM Sigmod-Sigact-Sigart Symposium on Principles of Database Systems, May 31-June 3, 1999, Philadelphia, PA, USA., pp: 68-79.

Chen, L., M. Tseng and X. Lian, 2010. Development of foundation models for internet of things. Front. Comput. Sci. China, 4: 376-385.

Gonzalez, H., J.W. Han, H. Cheng, X.L. Li, D. Klabjan and T.Y. Wu, 2010. Modeling massive RFID data sets: A gateway-based movement graph approach. IEEE Trans. Knowl. Data Eng., 22: 90-104.

Gonzalez, H., J.W. Han, X.L. Li and D. Klabjan, 2006. Warehousing and analyzing massive RFID data sets. Proceedings of the 22nd International Conference on Data Engineering, April 3-7, 2006, Atlanta, GA, USA., pp: 83-83.

Ilic, A., T. Andersen and F. Michahelles, 2009. Increasing supply-Chain visibility with rule-based RFID data analysis. IEEE Int. Comput., 13: 31-38.

Jeffery, S.R., M. Garofalakis and M.J. Franklin, 2006. Adaptive cleaning for RFID data streams. Proceedings of the 32nd International Conference on Very Large Data Bases, September 12-15, 2006, Seoul, Korea, pp: 163-174.

Lee, C.H. and C.W. Chung, 2011. RFID data processing in supply chain management using a path encoding scheme. IEEE Trans. Knowl. Data Eng., 23: 742-758.

Lian, X., L. Chen and S.X. Song, 2010. Consistent query answers in inconsistent probabilistic databases. Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data, June 6-11, 2010, Indianapolis, IN, USA., pp: 303-314.

Nath, B., F. Reynolds and R. Want, 2006. RFID technology and applications. IEEE Pervasive Computing, 5: 22-24.

Wang, F. and P. Liu, 2005. Temporal management of RFID data. Proceedings of the 31st International Conference on Very Large Data Bases, October 4-6, 2005, Trento, Italy, pp: 1128-1139.

Wang, F.S., S. Liu and P. Liu, 2010. A temporal RFID data model for querying physical objects. Pervasive Mobile Comput., 6: 382-397.

Xie, D., Q.Z. Sheng and J.G. Ma, 2012. A temporal-based model of uncertain RFID data. Proceedings of the 7th International Conference on Computer Science and Education, July 14-17, 2012, Melbourne, VIC., pp: 847-850.