

A Framework of Process Mining for RFID Event Analysis

Kyuhyup Oh, Jae-Yoon Jung

**Department of Industrial and Management Systems Engineering
Kyung Hee University, Yongin-si, Gyeonggi-do, 446-701, Republic of Korea**

Abstract

As information systems and telecommunication devices are spread, many organizations accumulate a lot of events which are generated in performing business activities. The analysis of real-time data and events can play a critical role in implementing real-time enterprises and business intelligence. Recently, supply chain and manufacturing sectors have adopted ubiquitous environment that generate RFID and sensor events. However events analysis to discover important information is desirable. In this paper, we propose a framework of business process mining for RFID events analysis. We explain how raw RFID events can make business process data accessible for process mining. The results of this research can be widely used for RFID events analysis in various areas such as logistics and manufacturing systems.

Keywords

RFID, process mining, business process analysis

1. Introduction

As ubiquitous computing devices such as RFID and USN are spreading, they can provide useful information of analyzing business process. Recently, research on business process analysis includes the physical data which are generated from the devices in order to comprehend the logistics of products or the progress of human task in real-time. RFID (Radio Frequency Identification), which recently replaces bar codes, is used for the purpose of identification of objects by using radio waves. Ubiquitous technology is of interest in the manufacture, retail and procurement sectors. It plays the role of connecting enterprise information systems with the physical environment. With real-time collection and accumulation of RFID events, it can be used for the tracking of objects or the procedure of tasks in business process.

Many information systems such as enterprise resource planning (ERP), supply chain management (SCM), and business-to-business e-commerce execute business process by definition. They accumulate events during its execution in a system log. To analyze business process execution logs, research on business process mining appeared [1]. This technique consists in the extraction of information from processes logs that can be later used to innovate the process, to improve its design., to make performance worker operation analysis. The techniques used come from statistics, artificial intelligence, and social networks analysis [2-4].

In this paper, we present the framework of process mining for analysis of RFID events. This framework considers the collection of RFID events and its accumulation. It also allows the filtering of data by means of Complex Event Processing (CEP) technology. Finally, it can transform RFID events into the XML format of the process mining platform, called ProM.

Process mining using RFID event started with process model extraction [5]. Then events are widely used for SCM sector [6, 7]. Gerke et al. [6] adopted RFID process mining for the supply chain analysis based on EPCglobal RFID standards by using ProM.

In this paper we provide a twofold contribution. First, we propose the process mining framework using real-time RFID events. This framework takes RFID events and generates log files for process mining. Second, the proposed process mining system supports collection and filtering of raw RFID events by using CEP.

The paper is composed as follows. The framework of RFID process mining is presented in Section 2. To illustrate the proposed framework, an implemented system is introduced in Section 3. An example of RFID process mining is explained with some process mining techniques with the implemented system in Section 4. Finally, Section 5 concludes this paper.

2. RFID Process Mining

In this section, we describe the procedure of process mining using physical RFID data. The aim of RFID process mining is to understand and improve processes. The objects used for process mining can be RFID tagged items and workers. Figure 1 shows the procedure of RFID process mining, which consists of three steps: i) collecting and processing of RFID events, ii) transformation of RFID events into mining formats, and iii) process mining of RFID events.

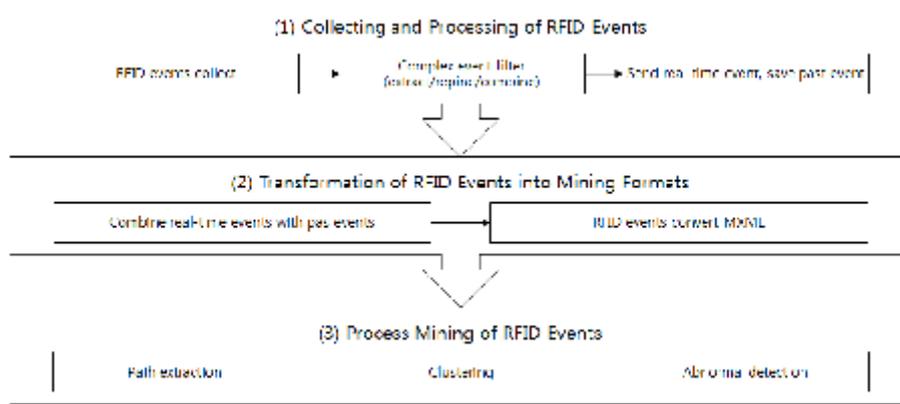


Figure 1: Framework of RFID process mining

2.1 Collecting and Processing of RFID Events

The first stage is to collect and process events which are generated from RFID readers. This stage user can set filter rule by CEP language. The rule can filter single or complex type of reader name, reader antenna name, tag id and so on. The reader sends RFID event readings to a message queue in the application server. Message queues are using for temporary saving RFID events. We use two message queues. The first queue is used for saving non-filtering event. The second queue use for saving filtered events. Filtering is done with CEP. The CEP engine can filter duplicated events and filtering by user-defined rules. Duplicated RFID events consist of repeated tags generated in the same reader. Filtered events are served when the converter requests events. Finally, filtered events are moved to database after a given period. The reasons to move events into a database, are that the message queue has a storage limit and the database is resource easy to query for past events.

2.2 Transformation of RFID Events into Mining Formats

The next stage is to combine real-time events in the message queue with past events in the database. The combined RFID events are converted into the Mining XML (MXML) format.

The analysis result is changed by past events. So we separated past events and real-time events. Past events are selected by user rules. The user can select events from database by type of date, time and reader's type. Selected events make the process mining more accurate. Then selected events and real-time events are combined for transformation.

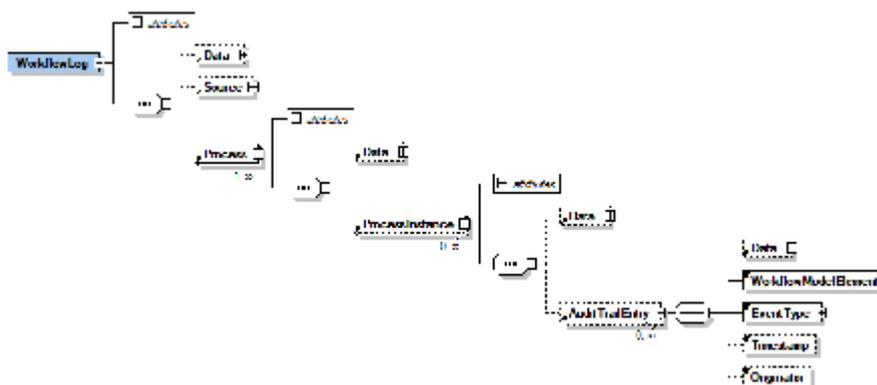


Figure 2: MXML schema

In order to make events accessible to mining algorithms, we do a transformation of events into the Mining XML (MXML) format. MXML is a generic XML based format to store process log files. The format can be used as an input file of the ProM platform. The ProM framework (<http://www.processmining.org>) is a process mining platform, which has a lot of plug-ins such as process model extraction, process analysis, process performance analysis, and social network analysis. The open source of the ProM platform is provided and an additional plug-in can be developed on the platform according to the users' purposes. Figure 2 shows the MXML schema. MXML root node of each log is the WorkflowLog, which can contain several Processes and optional information about the Source as well as further Data. Each ProcessInstance can have an arbitrary number of AuditTrailEntries. The AuditTrailEntries represent the events of the process. They are assumed to be in chronological order. The WorkflowModelElement describes the process activity. The EventType delineates the state change of the WorkflowModelElement. Every level of the hierarchy has the optional element Data, which can be used to store additional information

An RFID event has the following data: TagID, Reader Name, DiscoveryTime, Antenna, and ReadCount. TagID is a unique RFID ID, which are used to classify objects such as products, customers, and smart carts. ReaderName is the name of the RFID reader which generates RFID events. DiscoveryTime is the point in time when RFID events are generated. Antenna is the name of the RFID antenna which detects RFID tags. ReadCount is a number of counts that the reader recognized the same RFID tag.

According to the MXML structure, we have to assign each event at least to a WorkflowModelElement and a ProcessInstance. There are several data fields of an RFID event that can be directly mapped to the MXML fields. The mapping is illustrated in Table. 1.

Table 1: Mapping of event fields

RFID Event Attribute	Attribute Description	MXML Field
ReaderName	RFID Reader Name	Originator
TagID	RFID Tag ID	ProcessInstance ID
DiscoveryTime	RFID generated time	TimeStamp
Antenna	Antenna name	WorkflowModelElement
ReadCount	RFIDEvent generate Start/End	Event Type

2.3 Process Mining of RFID Events

The final stage is to analyze transformation events by process mining algorithms. The ProM platform support many plug-ins which implement the algorithms of process mining. The algorithms cover path extraction and analysis (tag attached object path analysis and extraction), similarity group classification (analyzed path classify with similar path), abnormal execution detection (ongoing process detect abnormal execution by past process), and so on.

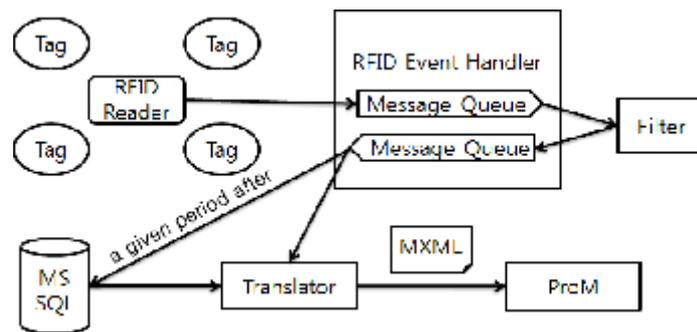


Figure 3: Overall procedure of RFID process mining

3. System Implementation

In this section, we describe implementation of the process mining framework presented in Section 2. Figure 3 shows the procedure of RFID process mining, which goes as follow: first, each tag is read by a RFID reader. Second, each reader sends RFID events to RFID Event Handler. Third, the RFID Events handler temporarily saves events in the first message queue. The handler then filters the duplicated events according to user-defined rules. Filtered events are saved in the second message queue. Events are retained in the second message queue for short periods (e.g. 1 week). After that period, filtered events are moved to into the database and deleted from the message queue. Fourth, the

translator combines real-time filtered events with past filtered events according to the selected user rules. Then the translator transforms events into the MXML. Finally, the process mining is done in ProM.

For the implementation, the Alien 9800 development kit made by Alien Technology is used for passive RFID event processing. This machine multi-read EPC Gen2 RFID tags. A reader can send RFID event to RFID event handler in XML.

The RFID event handler receives and filters RFID events. The events handler is implemented using GlassFish [8] and Esper engine [9]. GlassFish is used for saving events. GlassFish can exchange events and interact with another system using message queues. The Esper engine is for filtering events. Esper engine can filter duplicated events and filter them by using user-defined rules. Received RFID events are saved in the RFID event handler in the first message queue. Later, saved events are filtered by rules defined in Event Processing Language (EPL). These EPL-written rules select event streams from the message queue. Table 2 shows an example of an EPL statement used to filter duplicated events. Later, filtered events are saved in the handler at a second message queue. When the translator requests events, the handler sends events from second message queue. The saved messages are moved to database after a given period (1 week). We use MS-SQL as the database engine. MS-SQL procedures are used for inserting and selecting data in XML.

Table 2: EPL State for Event Filtering

```
INSERT INTO FilteredRFIDEventStream
SELECT ReaderName, TagID, TimeStamp, SUM(EventType)
FROM RFIDEventStream
GROUP BY TagID and ReaderName
HAVING EventType < 2
```

```
<?xml version='1.0' encoding='UTF-8'>
<ReaderName>Test Reader</ReaderName>
<ReaderType>Alien RFID Tag Reader, Model : ALR 9000</ReaderType>
<Process>192.168.0.7</Process>
<Timestamp>2009/12/12 12:12:12</Timestamp>
<Originator>Test Reader</Originator>
- <Alien RFID Tag>
  <TagID>0102 0304 0506 0709</TagID>
  <DiscoveryTime>2009/12/12 12:12:12</DiscoveryTime>
  <LastSeenTime>2009/12/12 12:12:12</LastSeenTime>
  <ReadCount>1</ReadCount>
</Alien RFID Tag>
- <Alien RFID Tag>
  <TagID>0103 0506 0708 0709</TagID>
  <DiscoveryTime>2009/12/12 12:12:18</DiscoveryTime>
  <LastSeenTime>2009/12/12 12:12:18</LastSeenTime>
  <ReadCount>0</ReadCount>
</Alien RFID Tag>
```

Figure 4: RFID event XML

```
<?xml version='1.0' encoding='UTF-8'>
<ProductInstance id='0102 0304 0506 0709'>
  <AuditTrailEntry>
    <Timestamp>2009/12/12 12:12:12</Timestamp>
    <Originator>Test Reader</Originator>
    <WorkflowModelElement>0</WorkflowModelElement>
    <EventType>Appear</EventType>
  </AuditTrailEntry>
</ProductInstance>
- <ProductInstance id='0103 0506 0708 0709'>
  <AuditTrailEntry>
    <Timestamp>2009/12/12 12:12:18</Timestamp>
    <Originator>Test Reader</Originator>
    <WorkflowModelElement>0</WorkflowModelElement>
    <EventType>Appear</EventType>
  </AuditTrailEntry>
</ProductInstance>
```

Figure 5: Transformed ProM Type MXML

The translator is used for combining and transforming events. First, the translator selects events by user-defined rules. User can set rules using SQL query or event types (e.g. date, time, RFID types). Second, the translator loads real-time events. Then translator combines loaded events with selected events. Finally, translator transforms events into MXML based on event mapping Table 1. Example, figure 4 is garnered RFID event XML by reader. This event is transformed MXML (see Figure 5).

Process mining can be achieved by using transformed MXML on ProM platform. ProM supports a lot of process mining tools. The results of example RFID process mining are explained in Section 4.

4. Example

We propose a minimarket store as an application scenario for effective real-time RFID analysis, schematically shown in Figure 6. We suppose each store has installed RFID readers. We also assume that each cart is a smart cart (i.e it includes RFID tags). Each customer starts at the entrance, and ends at the exit. The analysis consisted in studying the customer paths by means of RFID events. The analysis steps go as follows. First, the reader sends the cart's tag events to the handler. Second, the handler filters the events and sends the events to the translator. Third, the

translator combines selected past RFID events in stored in the database with real-time RFID events. Fourth, the translator transforms events into MXML. Finally, we do process mining using ProM by created MXML. The ProM has a lot of plug-ins. We used fuzzy mining plug-in for the analysis, and the result was illustrated in Figure 7. This result shows the customers cart path. This path depicts the sequence of movements that a customer followed. Every cart almost started at entrance counter, ended at exit counter. In a first path (1→ 3→ 9) customers are mainly bought foods. In a second path (0→ 2→ 4→ 7→ 8→ 9) customers bought food, cosmetics and appliances. And in a third path (0→ 5→ 6→ 9) customers bought clothes.

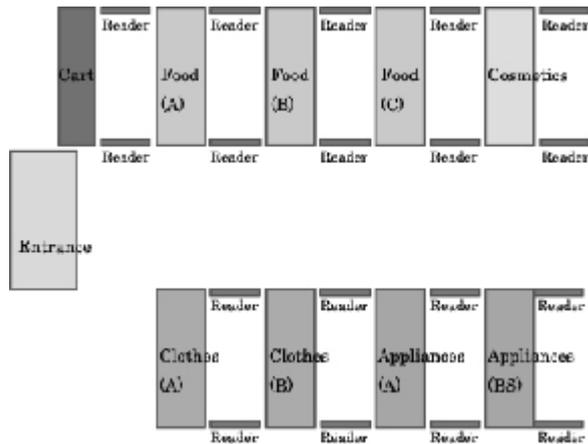


Figure 6: Example scenario of RFID Installed mart

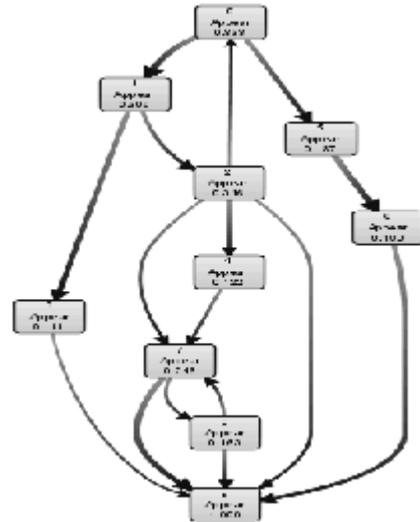


Figure 7: Fuzzy mining result of mart example

5. Conclusion

The proposed RFID process mining framework and its implementation, presented in this paper, provide a reference architecture for business process mining using real-time RFID events. The real-time RFID event processing allows companies to detect and react promptly to situations in a changing environment. We consider that RFID event process mining can be a helpful technique for real-time logistics tracking and logistics systems improvement.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2010-0015334).

References

- [1] van der Aalst, W. M. P., Weijters, A. J. M. M., and Maruster, L., 2004, "Workflow Mining: Discovering Process Models from Event Logs," *IEEE Transactions on Knowledge and Data Engineering*, 16(9), 1128-1142.
- [2] Alves de Medeiros, A. K., Weijters, A. J. M. M., and van der Aalst, W.M. P., 2007, "Genetic Process Mining : An Experimental Evaluation," *Data Mining and Knowledge Discovery*, 14(2), 245-304.
- [3] Jansen-Vullers, M. H., van der Aalst, W. M. P., and Rosemann, M., 2006, "Mining Configurable Enterprise Information Systems," *Data and Knowledge Engineering*, 56(3), 195-244.
- [4] van der Aalst, W. M. P., Reijers, H. A., and Song, M., 2005, "Discovering Social Networks from Event Logs," *Computer Supported Cooperative Work*, 14(6), 549-593.
- [5] Gonzalez, H., Han, J., and Li, X., 2006, "Mining Compress Commodity Workflows from Massive RFID Data Sets," *Proc. of Conference on Information and Knowledge Management*, 162-171.
- [6] Gerke, K., Claus, A., and Mending, J., 2009, "Process Mining of RFID-based Supply Chains," *Proc. of IEEE Conference on Commerce and Enterprise Computing*, 285-292.
- [7] Henry, H. B., and Lin, D. K. J., 2009, "RFID-Enabled Discovery of Supply Networks," *IEEE Transactions on Engineering Management*, 56(1), 129-141.
- [8] Glass Fish Homepage. <http://glassfish.dev.java.net>
- [9] Esper Engine Homepage. <http://esper.codehaus.org>