A Review on Complex Event Processing in RFID System

Siti Salwani Yaacob#, Hairulnizam Mahdin#, RD Rohmat Saedudin*, Mohd Hanafi Ahmad Hijazi$, Mohd Sanusi Azmi^, Nor Azlina Rais†

#Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Batu Pahat, 86400, Malaysia
E-mail: hairunb@uthm.edu.my

*Department of Information System, Telkom University, Bandung, Indonesia

$Faculty of Computing and Informatics, Universiti Malaysia Sabah, Malaysia

^Fakulti Teknologi Maklumat dan Komunikasi, Universiti Teknikal Malaysia Melaka, Malaysia

†Jabatan Pengajian Umum, Kolej Poly-Tech Mara, Batu Pahat, Johor, Malaysia

Abstract—The motivation of this paper is to review some of the leading research issues encountered by current Complex Event Processing (CEP) techniques. General CEP system undoubtedly laid a certain degree of uncertainties due to some unforeseen reasons such as inaccurate measurements through network failures or unpredicted interference from the system failures. A variety of traditional to the most current techniques was disclosed with some features, which are rarely considered in typical CEP problems. Besides, this paper also discusses a broad review of modern and future solutions, including techniques beyond mainstreams in complex event processing. Based on the critical analysis of prior techniques and solutions, therefore, we further the research and propose a solution for a new type of event processing engine especially in RFID for future reference.

Keywords— RFID; complex event processing; RFID event

I. INTRODUCTION

The term ‘RFID’ is not new anymore. Radio Frequency Identification (RFID) technology is a new common but fundamental technology behind every edge of business in the world. It is a wireless communication technology that used radio frequency waves for detecting, tracking, and searching objects. The implementation of RFID technology has been widely implemented as well as integrated into many other technology and applications such as logistics and supply chain visibility [1], [2], medical and healthcare [3], [4], and real-time location system [5].

Through billions of connections of tags to sensors via the internet to multiple platforms, the RFID technology based application will generate extraordinary volume data streams [6]. In the presence of a vast amount of sensor and RFID data, it is difficult, and even not possible to determine accurate data and match according to the needs of an enterprise system. The traditional data processing method is now unreasonable for RFID data processing as the features of RFID data are unstable, dynamic, massive and spatial-temporal correlation [7]. These require data to be filtered and correlated in complicated pattern detection process and convert to events that tell meaningful, actionable information to end applications. In recent years, the use of Complex Event Processing (CEP) has been gaining attention in the research of wireless sensor networks, particularly RFID applications. The primary challenge associated with sensor data is the nature of the application that contains data uncertainty, which is indirectly demanding real-time filtering and data analysis [8], [9]. The conventional CEP inevitably carry a certain degree of uncertainties, such as imprecision and incompleteness caused by the sensor environment and mediators [10]. These uncertainties are the results of insufficient and imprecise that leads to various challenges in uncertainty due to lack of knowledge and system errors.

In CEP, it mainly consists of event filtering and event detection [11]. Event filtering aims at errors for event detection. Event detection is about detecting and searching particular pattern and its relationship in the event stream. The CEP system then issues a new type of event, based on the pattern detected and its relationship so that meaningful event can be extracted. The system has to process the event data more or less as it arrives so that the appropriate action can be taken quickly. Take stock broker application as an
example, when the average price of stock’s market value falls below a limit value over a specified period, then a company can sell the stock in order to avoid losses. While in the banking system, if there are suspicious transactions happened at different geographical locations and at the same time, then the authority can release the report as a potential fraud. Based on the examples above, the CEP deals with processing multiple event streams with a goal to identify a particular pattern of events and its relationship in the event streams [12].

II. MATERIAL AND METHOD

A. The CEP mechanism in RFID

In complex event processing, the occurrence of RFID data from the RFID data streams is known as RFID events. Event stream and data stream have the same characteristics, but an event stream hides a specific meaning. Essentially, an event stream consists of events and every event includes event ID, event attributes, and the occurrence of time. Compare to data stream the stream does not comprise any specific meaning. The RFID event attributes are defined as follows [14]:

\[ \text{Event (TagId, ReaderId, Timestamp)} \]  \hspace{1cm} (1)

TagId represents the tag identifier, while ReaderId is the RFID reader identifier and Timestamp represents the timestamp when the reader read the tagged object.

In complex event processing, there are two types of events; primitive event and complex event [10]. In RFID applications, the primitive events are events generated during the interaction between readers and tagged objects. The primitive event will undergo multiple processes such as pattern detection, or aggregation before it is derived to complex events. The compound event is an inferred event based on its relationship between event pattern and event rules that occur over a specified period.

Besides the improvement of the precision of RFID readers, there are still void readings such as missed readings and false negative readings [18]. False negative readings or noise can be defined as the unexpected extra readings that are generated by the readers. This is due to the reason of interference, or system failure of some components. In addition, RFID data may hold meaningful redundant information due to duplicate readings in the event stream [19]. Duplicate readings are widespread in every RFID application. The most common reason is that the multiple readers read the same tag at high frequency.

In this study, we classify the inaccuracy of RFID data into two categories: unreliability and uncertainty. Here, unreliability refers to the invalid readings such as false positives and false negatives readings. While uncertainty refers to the failure of RFID reads from an application and system design perspective. Two types of data filtering

Complex events were created based on sequence SEQ (ABC).

The CEP becomes more mature, the challenge of real-time analysis becomes more complicated since the sensors and smart devices continue to collect more data from a different domain. Being able to react quickly in a mission-critical situation that can save companies from losses is one of the IoT goals. This is the reason why complex event processing becomes the industry’s leading solution.

Fig. 2 portrays a process of RFID events from event stream to event detection. The event streams consist of RFID events from different locations and times. Based on the table above, the RFID event holds essential contextual characteristics such as temporality, spatiality, and implicit semantics. RFID events are always temporal constrained [15]; the temporal distance between two events and the interval of a primitive event are demanding for event detection. Besides, RFID event is always in spatial localization [16]. The tags object is detected based on the localization of the tag for example in store at the supermarket that is represented by distance and angle. The events always need to be filtered as it naturally consists of uncertain due to environment and system errors. After filter, the events will be aggregate according to, for example, rule or sequence based on time. RFID has unexpressed meanings [17], which have to be semantically altered and aggregated to find its hidden meanings.

As in Fig. 1, a sequence of tagged items move through reader 1 and recognized by the reader as an event type A. The tagged items then randomly packed into the package and read by reader 2 as an event type B. These packages were moved into a warehouse and observed by the reader 3 and recognized as an event type C. From the architecture as shown in Fig. 1, a sequence was recognized as ABC.
should be considered for RFID data before they can be further processed: low-level data filtering and semantic data.

B. The CEP Model and Architecture

Complex event processing has three different levels [20]: (i) Filtering, (ii) Matching and (iii) Event Analysis. Based on Fig. 3, as events entering into the CEP engine, some pre-processing is essentially needed. It may include translating the source-specific event format to a format that is understandable by the CEP engine. Once pre-processing completed, events are enumerating over prerequisite queries and ready for pattern match process.

![Fig. 3 Phases in complex event processing](image)

The process of pattern matching is the most important component in the CEP engine. During pattern matching process, this component detects the predefined pattern and its relationships to be inferred or identified as a complex event. Event constructors, also known as event operators are used to express the relationships between events and correlate events to define complex events. In this phase, it associates several methods such as grouping and correlating, as well as filtering and aggregation to identify a specific pattern of events within or across streams. Based on Fig. 2, the matched pattern is then expressed as a complex event and sends to post-processing stream.

Meanwhile, in the derivation phase, this component detects more complex events using the information detected from matching subsets in event history. Once the filtering and grouping process completed, not all well-ordered and uncertain event from the post-processing stream can generate meaningful events. It is necessary to filter those potential data to go through complex event processing. In the meaningful event streams, it shows the relationships between data such as causality, membership timing and event-driven process. The definition of a meaningful event is varying depending on the different situation, requirement, and scenarios.

Once the detection of complex events completed, the semantics event then interpreted and can be merely integrated based on domain and applications. At this point, such action can be taken, for example, updating business dashboard, updating database, sending an alert by short message services or email, or even one can start a new piece of application work.

C. Related Works

The term CEP research domain is getting the hype that attracts the attention of researchers. The research literature shows that most of the researches focus on the system design and implementation of CEP based engine that can handle complex event queries over event streams in real time.

Some of the CEP systems such as SASE [21], CAYUGA [22], [23], CASCADIA [24], and RCEDA [25] consider that the event stream is errorless, or it has been filtered before processing starts. Even though the techniques equipped with significant and distinguished processing functions to the complex event, these researches do not consider uncertainty in the event stream and only effective to process and improve the RFID event streams. On the other hand, focusing on the uncertainty event stream, systems such as SASE, Cayuga, and RCEDA unable to detect potential complex events from probability event stream and utilize event stream with probability features, hence decrease the efficiency of event processing capability. While Cascadia systems, on the contrary focus on the event detection and event processing with uncertainty. However, for the reason that the research primarily focused on the probabilistic event model based on the natural RFID event, this research does not deliberate uncertainty in the compound event stream and review in detail and optimized the probability event in the compound event stream.

In recent times, a few techniques in complex have been introduced to handle complex event queries in event streams such as in real time. The Complex Event Processing Mechanism for Android Platform [26] has been introduced to controls and detect the RFID complex event in real-time and execute an event by implementing web page script. The CEP based on Android platform comprising two modules: the CEP module in the terminal and the CEP module in the server. Besides server, the CEP engine in the terminal is programmed in the Android devices. The authors assured the fact that using mobile devices solves data congestion at the server and increase the processing speed of native jobs in the terminal.

Besides CEP for the Android platform, the authors have proposed LiSEP [27], an integration of CEP engine based in a distributed system comprising a hierarchy of nodes organized in different levels. The research describes a technique based on CEP and RFID technologies to sustain clinical risk management by detecting potential critical conditions on the patient. LiSEP allows efficient load balancing of the event stream for the reason that the number of events reaching the top level is lesser compared to the number of events pass into the system. In this case, the response is vital in order to detect potential critical conditions.

Other than LiSEP, Hierarchical-Complex Event Processing (HCEP) [28] in wireless sensor networks was proposed by query planning and assign sub-expressions to the low-level nodes. By using query planning, the expression
will be separate into sub-expressions, and these sub-expressions will be allocated to each sensor and gateway. By processing sub-expressions, sensors and gateways send meaningful event, and in the meantime, their computing and storage capacity is fully utilized. The authors also describe even though the sensors need to do complicated event detection process, the amount of transmission data can reduce and on the contrary reduce the bandwidth sensors and gateways. By using proposed CEP engine also can minimize the response time as the primitive event can be processed as soon as it arrives at the engine.

Researchers have proposed Sequence Clustering-based Automated Rule Generation (SCARG) [29] for adaptive complex event processing. The proposed approach can automatically trigger new rules based on the decision-making history of domain experts by using sequence clustering and probabilistic graphical modeling. On the other hand, the proposed method also supports the automated rule updates. This makes the self-adaptive CEP system possible by conjointing the rule generation method with the current dynamic CEP systems.

A sophisticated event processing engine, LIDA-E [30] exploits rules based on the knowledge base to be implemented into an event processing algorithm for analysis. The event models an operator use to detect event pattern in order to support the rules for event selection and aggregation. The proposed operators based on rules are utilizing in order to develop proposed complex event processing system and architecture. The proposed system also adopts filtering agents to ensure the event transmission mechanism performs effectively. In this research, the authors believe that by implementing the proposed method and architecture gives better performance in both transmitting and analyzing especially on a large number of event streams.

Next, a complex event processing based on INFA-HTS (Improved Nondeterministic Finite Automaton-Hash Table Structure) [31] was proposed. This technique is an extension of the traditional nondeterministic finite automata (NFA) model to a new improved nondeterministic finite automata model (INFA) to capture the related RFID primitive events from the out-of-order event stream. To improve the storing and processing efficiency for the sizeable intermediate matching results that exist in detection processing, they proposed HTS (Hash Table Structure) to store and process the large intermediate result. To provide a fast output detection result, they used a hash table search technique to output the desired detection results instead of depth-first search technique.

Strong Match Principle and Weak Match Principle (SMP–WMP) aims to detect compound events in a random situation [32]. They also implement correlation coefficient technique (weight value) to determine the relationship between a random variable and observed data values. A correlation coefficient is a coefficient that emphasizes a quantitative measure of a few types of correlation and dependency, meaning statistical relationships between two or more random variables or observed data values. Each primitive event will assign weight value when they process uncertain complex events. When a primitive event’s value, relative to others event’s weight values, is smaller, it shows this basic event less effect on the complex event.

D. Comparative Analysis

Based on Table 2.0 below, these techniques were built centralized which results in the operation of sensors were not fully utilized and wireless network overload. In [26]-[28] and [30], the sensor often need to transmit unnecessary data and uncertainties that dissipates energy and broad bandwidth of the sensor. While in [26], [28] and [29], servers were encountered with pressure because they are responsible for all the processing logic. These can lead to wireless network overloads that caused data congestion at the server. While in [29], [31] and [32], these techniques mainly used sizeable sliding window in order to detect event redundancy. These can increase memory usage in the meantime affect the efficiency of event detection and pattern matching process in RFID event stream. Therefore, it is very challenging to meet the requirement of the applications, especially in the real-time processing system.

### Table I

<table>
<thead>
<tr>
<th>Techniques</th>
<th>Uncertainty Handling</th>
<th>Approach/Comments</th>
</tr>
</thead>
</table>
| [26] CEP for Android Platform | No 2014 | • Consist of two CEP module; Terminal and Server.  
• Data interact using WebView Control  
• Does not mention any event processing in the module. |
| [27] LiSEP | No 2014 | • Event model comprises four different levels that are hierarchically organized.  
• This model does not mention the event-filtering process. |
| [28] HCEP | No 2014 | • Automatic machine based hierarchical CEP  
• Assume sensors and gateway will filter redundant data. |
| [29] SCARG | Yes 2017 | • In order to filter uncertainties, the author compares the orders and similarities between primitive events that emerge along sequences. |
| [30] LIDA-E | No 2016 | • The filter only detects composite events from different information systems. |
| [31] INFA-HTS | Yes 2016 | • Filter data by using hash-mapping function whether it already exists in the hash array by searching for the event type. |
| [32] SMP - WMP | Yes 2018 | • Implement a coefficient function to identify the true event |
| Proposed Approach | Yes | • Using Kalman Filter and weight value to identify the relationship between events. |

III. RESULTS AND DISCUSSION

A reliable solution is demanded especially to overcome uncertainties in CEP technology. This is not just a critical problem, but challenging task as it involves some practical issues such as incomplete tags readings, environmental disturbances, and potential data forgery. Based on this
analysis, we would like to propose to use the Kalman Filter and correlation coefficient function (weight value) to overcome the uncertainties issues in the event stream. A complex event could unveil an opportunity or potential risk to particular businesses primarily based on RFID technology. The immediate reaction is critical to avoid such business from losses. Consequently, it is fundamental that the response reaction period is lesser than the processing time. As the aim of CEP applications is to execute low-latency and high-velocity processing, the CEP based technology need to enhance and to assure that the technology meets the performance goals.

In addition to the current techniques of CEP, during the event filtering phase, another powerful way to optimize the performance of event filtering is to use predictive analytics to determine the uncertainties. Predictive analytics is a form of advance analytic that uses historical data, statistical algorithms and machine learning techniques to determine the future outcomes. In a simple definition, it implements or uses a wide range of techniques to study data and discovers the knowledge of such implicit data. Predictive analytic techniques have been proven mutually beneficial to complex event processing. For example, it can be used to detect complex events and inferring rules in complex event processing. One of the predictive techniques includes Kalman filter and correlation function. Kalman filter is a set of arithmetic equation that brings an adequate computational explanation of the least squares method [33]. While correlation coefficient function is powerful as it supports estimation of past, current, and even future states even when the design system is still unknown. Kalman filters involve two repetitive stages: prediction update and measurement update. Fig. 4 shows the iterative steps in the Kalman Filter. The prediction stage is bound for projecting forward the current state and error estimates to obtain a prior estimate for the next time point. While the measurement stage is responsible for the feedback, incorporating a new measurement into the prior estimate in order to obtain a previous improved estimate.

![Fig. 4 Stages in Kalman Filter](image)

While pattern matching, also known as pattern detection, enables the analysis of event collections and the relationship between events. In other words, a conditional combination of event pattern that matches and it satisfies a particular pattern definition. In this technique, it involves a combination of several methods. Grouping and correlating, as well as filtering and aggregate, are some commons methods used in order to identify a specific pattern to events within or across streams. However, these methods tend to use large sliding windows and will consume large memory usage. In order to solve this issue, we propose to use hashable method in order to minimize the use of memory and time processing. Hashable uses arrays as a storage medium, and hash technique is responsible for generating an index where an element is to be inserted or located from. Based on these methods, we believe that this will increase the optimization of the CEP engine.

IV. CONCLUSIONS

The advancement of RFID technology has become a new common technology with the capability of identifying tagged objects without line of sight. This technology has enhanced much existing application as well as the research domain to a different level of intelligent. From this point forward, the RFID technology has become vital at every edge of the world. Nonetheless, despite the undoubted potential of RFID application, there are considerable numbers of fundamental research that has been enhanced and improved to cater with issues especially based on the RFID technology. This paper elaborates some of the current advances in RFID based technology that implements event processing engine.

In recent times, the implementation of CEP in the business process has become a new standard for the reason that the technology rapidly expands to many areas of research and domain. Based on this review paper, previous as well as current methods and techniques based on CEP engine have been deeply discussed. This paper concluded our paper summarizes previous and current technique, also pointing out exciting directions for future research. In conjunction with the current research efforts, some of the exciting research direction has been pointing out and in the meantime encourage future research based on CEP to focus more on the CEP architecture and infrastructure design.

ACKNOWLEDGMENT

Ministry Of Education Malaysia sponsors this work under Fundamental Research Grant (FRGS) vote number 1611.

REFERENCES


