2014 IEEE Fourth International Conference on Big Data and Cloud Computing

Risk Prediction System based on Risk Prediction Model with Complex Event Processing

Risk Prediction in Real Time on complex event processing engine

Yoon-Ki Kim
Department of Electrical Engineering
Korea University
Seoul, Republic of Korea
vardin@korea.ac.kr

Chang-Sung Jeong
Department of Electrical Engineering
Korea University
Seoul, Republic of Korea
csjeong@korea.ac.kr

***Abstract*— System failure mainly arises from various reasons. It would most likely become an immediate cause of accident. It is difficult to predict system failure in real-time, since it is hard to detect which subsystem caused the whole system failure. To track the system failure in real-time, we have to recognize immediately the failure of each subsystem. In this paper, we propose new a risk prediction model and risk prediction system architecture to trace the failure of system in real-time by using fault tree analysis (FTA) on complex event processing (CEP) engine. We shall show our new risk prediction model efficiently enables CEP system to predict the risk before any undesired event occurs, and to prevent disaster, accident and disease in advance.**

***Keywords*— Risk Prediction model, Complex Event Processing(CEP), Real-time Processing, Fault Tree Analysis(FTA).**

I. Introduction

Today, the more technology develops, the more a system become complicated. Hence, analyzing the cause of accident from system fault is difficult. System fault is derived from one or more subsystem fault. To predict a risk of system, we need to understand the failure logic that gives rise to the subsystem fault. Fault tree analysis (FTA) [1] is a top down, deductive failure analysis in which an undesired state of a system is analyzed using Boolean logic. This method is to understand how systems can fail. It cannot be implemented in real time. We propose the new risk prediction model using FTA can detect status of each subsystem with complex event processing (CEP) [2] in real time. In this paper, we present a risk prediction model offering a way easy to detect system fault based on FTA in four parts. Risk prediction model efficiently enables CEP system to predict the risk before undesired event occurs. Using this, we can prevent disaster, accident and disease in advance.

The outline of our paper is as follows: In section 2, we describe related works for introducing FTA to design a risk prediction model and CEP to process complex event to predict a risk in real time. In section 3, we present the risk prediction model enables system to predict a risk before accident occurs. In section 4, we explain the methodology for designing a risk prediction model based on FTA and converting risk prediction model into query in CEP language. In section 5, we explain the architecture of rick prediction system monitoring all subsystem status in real time.

II. Related Works

A. Fault Ttree Analysis(FTA)

FTA [3] is a method for identifying ways which hazards can lead to accident. Since no system is perfect, dealing with a subsystem fault is necessary, and any working system eventually will have a fault in some place. Fault Tree is a graphical representation of the interrelationships between equipment failures. A logic diagram shows how initiating events, at the bottom of the tree, through a sequence of intermediate events, can lead to a top event. An undesired effect is taken as the root ('top event') of logic tree. The logic to get to the right top events can be diverse. One type of analysis that can help with this is called the functional hazard analysis. There should be only one top event. Each situation that could cause that effect is added to the tree as a series of logic expressions. When fault trees are labeled with actual numbers about failure probabilities, FTA can calculate failure probabilities from fault trees. When a specific event is found to have more than one effect event, i.e. it has impact on several subsystems, it is called a common cause or common mode. Graphically speaking, it means this event will appear at several locations in the tree. Common causes introduce dependency relations between events. The probability computations of a tree which contains some common causes are much more complicated than regular trees where all events are considered as independent.

B. Complex Event Processing(CEP)

Complex Event Processing (CEP) [4] is a method of tracking and analyzing streams of information (data) about things that happen (event) in real time, and deriving a conclusion from them. CEP is event processing that combines data from multiple sources [5] to infer events or patterns that suggest more complicated circumstances. The goal of complex event processing is to identify meaningful events [6] (such as opportunities or threats) and respond to them as quickly as possible. And CEP engine is a processing system to process the

978-1-4799-6719-3/14 $31.00 © 2014 IEEE
DOI 10.1109/BDCloud.2014.43

IEEE computer society

event stream data in real time. As CEP engines, event correlation engines analyze a mass of events [7]. It uses event processing statement derive and aggregate information from one or more streams of events, to join or merge event streams, and to feed results from one event stream to subsequent statements. And CEP uses named windows which is a global data window that can take part in many statement queries, and that can be selected-from, inserted- into and deleted-from by multiple statements. Named windows are similar to a table in a relational database system.

III. Risk Prediction Model

To predict the risk before undesired event occurs, we must define the risk logically. In this section, we propose a risk prediction model based on FTA.

A. Risk Prediction Model

Generally, before the undesired event occurs, sub-events occur in advance [8]. From this understanding, if we define the complex event as a risk, we can detect the undesired event. Since it can be loaded on CEP enable data to process in real time, it can predict any risk or detect fault occurrence real time. We suggest a new risk prediction model in real-time on CEP engine as shown in Fig 1.

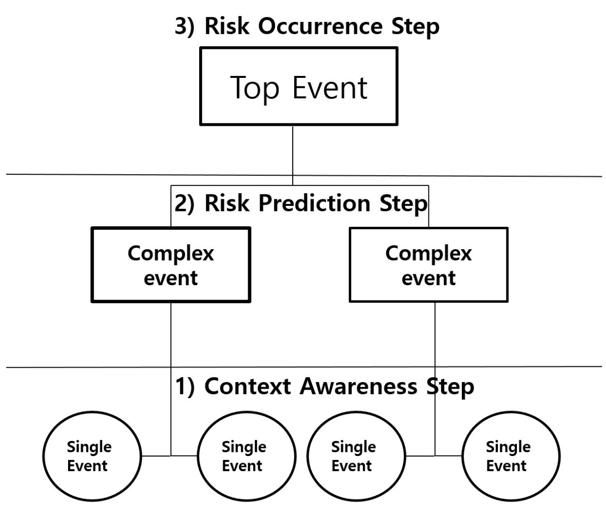

Fig. 1. Risk Prediction Model on CEP

B. Step for predicting risk

Top event is the ultimate target which should be prevented. Event contains an occurrence involving equipment faults or human errors, or external to the system. The following three steps are used to predict a risk.

1) Context Awareness Step

In the context awareness step, only single events occurs, and detected by the system. It may be an abnormal occurrence in one part of whole system. We do not know if it is a fatal cause of undesired top event or a temporary one for improving the system performance. However, in this step, we recognize all the abnormal events.

2) Risk Prediction Step

In the risk prediction step, several complex events occurs, and risk is predicted by the system. It take place before top event occurs. A complex event consists of several single events or other complex events from sub level. User predefines the undesired occurrence as a complex event which may cause a meaningful error. In risk prediction step, each complex event has a valid time attribute for filtering meaningless events. In this step, system gives an alarm.

3) Occurrence of Risk Step

In the risk occurrence step, top event occurs, and it must be prevented by the system. In this step, there is only one complex event called top event. It consists of one or more complex events from the lower step.

IV. Methodology

In this section, we describe the methodology for predicting or detecting faults based on risk prediction model. It consists of four steps which shall be explained bellows.

1) Express a system failure logic as risk prediction model based on FTA

Firstly, to understand system failure logic, define the undesired events which may occur. The primary events are typically used as follows.

- Single event - failure or error in a system component or element. It must be atomic.
- Complex event - consists of more than two single events.

Gate symbols describe the relationship between input and output events. The symbols are derived from Boolean logic symbols. The gates work as follows:

- OR gate - the output occurs if any input occurs
- AND gate - the output occurs only if all inputs occur

This deductive method which is used in a quantitative way, although it requires as an initial step a qualitative study of the system under consideration, just as any method of system analysis. The model is based on the combinations of failure of more basic system components, safety systems, and human reliability.

Fig. 2. shows a risk prediction model using Boolean logic . The top event derived from other complex events is the ultimate goal which we look for.

2) Convert risk predicton model into XML

To process the data on CEP engine, risk prediction model should be converted into CEP language. Since risk prediction model is expressed as a diagram, it cannot be used directly on CEP engine. For the using of a risk prediction model, we adopt EPL which is described in next section. EPL cannot express the hierarchical data. For this reason, risk prediction model should be converted into XML preferentially, which can express data as a tree type.

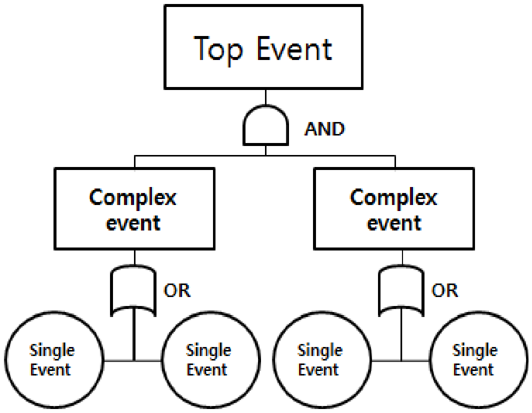

Fig. 2. A Risk Prediction Model Consists of Single and Complex Events

It provides the ability to navigate around the tree, selecting nodes by a variety of criteria [9]. XML contains both event and its attribute. XML data type is freely convertible to CEP language. The following is the rule of converting into XML.

- Top event is a kind of complex event assigned to uppermost hierarchy of XML. Each complex event or single event is assigned to under the top event in a hierarchical XML
- Complex event contains an attribute of relationship between each node as a Boolean logic, and valid time for finding meaningful event.
- Single event contains an attribute of event about stream source information.

We suggest the rules which enable risk prediction model to convert into XML. TABLE 1 indicates the rules within the tree representation for converting risk prediction model into XML.

TABLE I. CONVERTING RULES

Syntax	Risk Prediction Model Expression	Notes
Child	Single Event	Leaf node
Parent	Complex event	Top Event, Intermediate Event
Attribute	AND, OR, TIME, CONDITIONAL	Information for relationship between other nodes

Hierarchical data of node and relationship between each node in Fig 3. is expressed in XML as shown in Fig 2.

Complex event type has a Boolean logic attribute which expresses relationship between child nodes. On the other hand, single event type has an attribute about event stream source information and attribute of system status.

```
<?xml version="1.0 encoding="UTF-8">
<Top_event boolean="AND" time="30sec">
  <Complex_event_A boolean="OR" time="15 sec">
        <Single_event_A   event   stream="Sensor_A"
         Heart Rate="82>=">
        </Single_event_A>
        <Single_event_B event stream=" Sensor _A"
        Blood Vessel Tension ="2.0<=">
        </Single_event_B >
    </Complex_event_A>
    <Complex_event_B boolean="OR" time="20 sec">
        <Single_event_C event stream="Sensor_B"
         Stress Power="-120.0<=">
        </Single_event_C>
        <Single_event_D event stream=" Sensor _B"
        Differential Pulse Wave Index="-300.0<=">
        </Single_event_D>
    </Complex_event_B>
</Top_event>
```

Fig. 3. A Example of XML from FT

3) Extract the query in CEP language from XML

We suggest a method of extracting CEP language from XML. CEP language is similar to SQL(structured query language) [10]. We adopt EPL(Event processing Language) [11] as CEP language. EPL is a declarative language that combines information from one or more streams of event. EPL is similar to SQL in its use of the *select* clause and the *where* clause. However EPL statements instead of tables use event streams and a concept called views. Similar to tables in an SQL statement, views define the data available for querying and filtering. Views can represent windows over a stream of events. Views can also sort events, derive statistics from event properties, group events or handle unique event property values. The overall structure of EPL is:

```
SELECT      <event pattern>
WHERE       <qualification>
FROM        <Sensor_ID>.win:time()
```

First of all, EPL of single event queries are generated from XML made in the previous step. The following example expresses single_event_A in Fig 3. by using EPL.

SELECT	*Heart Rate.Sensor_A*
WHERE	*Heart Rate>=20*
FROM	*Sensor_A*

Complex event queries are generated from several single events in child nodes. Its declaration is generated based on single event declaration by using Boolean logic. The following example which expresses the complex_event_A in Fig 3. by using EPL.

SELECT	*Heart Rate.Sensor_A,*
	Blood Vessel Tension.Sensor_A
WHERE	*Heart Rate>=20 OR*
	Blood Vessel Tension <=2.0
FROM	*Sensor_A.win:time(15 sec)*

Therefore, we can define the event query of each node by expressing in EPL.

4) Execute the event queries on CEP engine

The complex event queries are loaded into CEP engine by query loader, and then executed on CEP engine.

V. Risk Prediction System Architecture

Our system architecture is designed to detect the faults not only in the subsystems but also in the whole system in real-time and to predict the overall system failures which could cause an accident. Figure 4. shows the risk prediction model in CEP engine. Our system consists of three components: Risk prediction model(RPM) Creator, Query Converter, and CEP Engine. The function of each component is as follows:

- RPM creator consists of Risk prediction model Creator and XML Creator XML is created based on RPM. Firstly user define the RPM. RPM has both node information and attribute. And RPM is converted into XML in this component.
- Query Converter consists of XML Analyzer and query Creator. XML Analyzer parses XML from Risk Prediction Creator. Query Creator creates event query by using EPL based on XML.
- CEP consists of Query Loader, Query Executor and Alarm Generator. It receives the event query from Query Converter. Event queries are loaded onto CEP engine initially. Query is executed by query Executor. When complex event occur, the Alarm Generator gives an alarm. CEP receives stream data continuously.

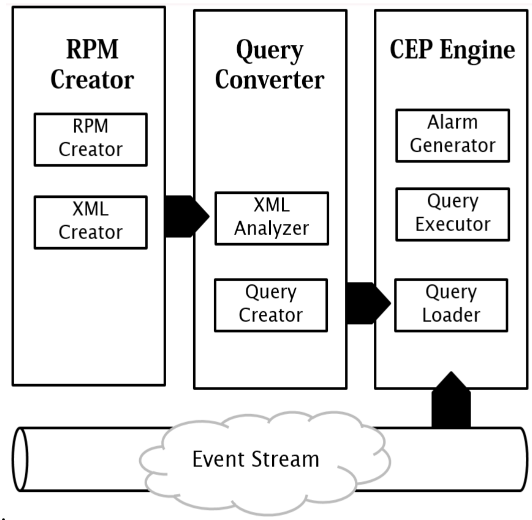

Fig. 4. Risk Prediction System on CEP engine

VI. Conclusions

In this paper, we presented a new risk prediction model to trace the failure of system in real-time with fault tree analysis (FTA) on complex event processing (CEP) engine. Also, we proposed the risk prediction system architecture for supporting risk prediction. Using this system, we can express the overall system logic easily, and track the system fault in real- time by using complex event processing engine. Since most of the accidents are caused by several subsystem faults, our system can efficiently prevent disaster, accident and disease in advance by detecting subsystem faults in real-time.

Acknowedgment

This research was supported by the MSIP (Ministry of Science, ICT & Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (NIPA-2014-H0301-14-1001) supervised by the NIPA (National IT Industry Promotion Agency), and Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF) funded by the MSIP (NRF-2013043678).

References

[1] Center for Chemical Process Safety (April 2008). “Guidelines for Hazard Evaluation Procedures (3rd edition ed.)”, Wiley.

[2] Luckham, David. “The power of events”, Vol. 204. Reading: Addison-Wesley, 2002,pp.1-2.

[3] “Fault tee analysis” http://en.wikipedia.org/wiki/Fault_tree_analysis.

[4] Luckham, David C. (2012). “Event Processing for Business: Organizing the Real-Time Enterprise” Hoboken, New Jersey: John Wiley & Sons, Inc.,. p. 3.

[5] Buchmann, Alejandro, and Boris Koldehofe. "Complex event processing." it-Information Technology Methoden und innovative

Anwendungen der Informatik und Informationstechnik 51.5 (2009),pp. 241-242.

[6] Bates, John, “John Bates of Progress explains how complex event processing works and how it can simplify the use of algorithms for finding and capturing trading opportunities”, Fix Global Trading, retrieved May 14, 2012

[7] Martin-Flatin, J.P., Jakobson, G. and Lewis, L. "Event Correlation in Integrated Management: Lessons Learned and Outlook", Journal of Network and Systems Management, Vol. 17, No. 4, December 2007.

[8] Lacey, Peter (2011). "An Application of Fault Tree Analysis to the Identification and Management of Risks in Government Funded Human Service Delivery" (pdf). Proceedings of the 2nd International Conference on Public Policy and Social Sciences. Retrieved 2013-07-09

[9] Bergeron, Randy (October 31, 2000). "XPath—Retrieving Nodes from an XML Document". SQL Server Magazine.

[10] Harkins, Susan Sales, and Martin WP Reid. "Structured Query Language."SQL: Access to SQL Server. Apress, 2002. 1-5.

[11] Event Processing Language. Available at http://esper.codehaus.org/index.html