Poster: Towards an Inexact Semantic Complex Event Processing Framework

Qunzhi Zhou
Department of Computer Science
University of Southern California
qunzhizh@usc.edu

Yogesh Simmhan
Ming Hsieh Department of Electrical Engineering
University of Southern California
simmhan@usc.edu

Viktor Prasanna
Ming Hsieh Department of Electrical Engineering
University of Southern California
prasanna@usc.edu

ABSTRACT

Complex event processing (CEP) deals with detecting real-time situations, represented as event patterns, from among an event cloud. The state-of-the-art CEP systems process events as plain data tuples and are limited to detect precisely defined patterns. Emerging application areas like optimization in smart power grids require CEP to incorporate semantic knowledge of the domain for easier pattern specification, and detect inexact patterns in the presence of uncertainties. In this paper, we present motivating use cases, discuss limitations of existing CEP systems and describe our work towards an Inexact Semantic Complex Event Processing (InSCEP) framework.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

General Terms

Design, Languages

Keywords

Complex event processing, Semantic Web, demand response

1. INTRODUCTION

Complex event processing deals with detecting real-time situations, represented as event patterns, from among an event cloud. In recent years, research into CEP has received much attention in the research community motivated by applications in domains like financial services [2] and RFID data management [4]. Many research prototypes, and several commercial systems such as ruleCore [1] and Esper [2] have been developed.

Demand response optimization (DR) in Smart Grid is an emerging application area for CEP [5]. Smart Grid is the modernization of power grid by integrating digital and information technologies, with the deployment of millions of sensors and smart meters to monitor energy use activities. DR, a cornerstone application of Smart Grid, deals with curtailing power load when a peak load is encountered. Continuous data relevant to DR emanating from various sources can be abstracted as events. These may be from smart appliances (*ThermostatChange* event), smart meters (*MeterUpdate* event), weather phenomena (*HeatWave* event) or consumer activity (*ClassSchedule* event). CEP can correlate these heterogeneous events to detect patterns that predict peak load occurrences or identify load curtailment opportunities for DR in a timely manner.

Limitations of existing CEP systems limit their uses in diverse information space like Smart Grid. Existing systems process events as relational data tuples. As such, event patterns can only be defined as a combination of attributes presented in event data. In addition, most CEP systems only support precise pattern matching, without any leeway to relax pattern constraints. However, uncertainty is an intrinsic feature of real world cyber-physical applications, where potentially incomplete and even incorrect information exist, yet need to be matched within certain bounds.

An effective CEP solution for DR optimization needs to extend traditional CEP systems in two aspects. First, it must be extensible to meet the organic growth of the Smart Grid information diversity with the provision to easily model and identify new events and event patterns by both domain experts and non-domain users. Second, it should capture uncertainties of events, and relax deterministic event patterns for inexact pattern detection.

2. USE CASES

We present example DR event patterns for load prediction, curtailment and monitoring, and use them to illustrate key features that our proposed Inexact Semantic Complex Event Processing (InSCEP) framework should provide. Consider in a campus micro grid, the DR application processes information coming from sensors and equipments that report their measurements or operations. We have the following patterns,

i. *Load Prediction:* A teaching building consumes 90% of its peak load, more than 5 classrooms have high probabilities of increasing from base load according to meter readings, class schedules and weather conditions.

ii. *Load Curtailment:* The thermostat in one office room is tuned 5 degrees lower than the average set point of thermostats in the same type of rooms which were tuned in the last 30 minutes.

[1] http://www.rulecore.com
[2] http://esper.codehaus.org

Copyright is held by the author/owner(s).
DEBS'11, July 11–15, 2011, New York, New York, USA.
ACM 978-1-4503-0423-8/11/07.

iii. *Load Monitoring:* Conservative curtailment patterns were applied, followed by a sequence of meter readings that indicate power load remains steady or increases.

Traditional CEP systems define patterns by specifying precise constraints of event data. However, the above examples illustrate the need to incorporate semantics and flexibility in pattern specification. The background knowledge of events from multiple domains (e.g. electrical systems, appliances, room scheduling, etc) need to be captured. In addition, flexibility has to exist to allow a limited number of errors or mismatches to still detect a relevant pattern. The need for specify such inexact patterns lies in two reasons, (1) component events can be probabilistic due to imprecise or incomplete observations, and (2) event pattern itself is uncertain and may have infinite acceptable equivalences. For instance, in the third example the sequence of meter readings need not strictly remain constant or monotonically increase. A small fraction of outsider readings should be tolerated.

3. INSCEP: SYSTEM DESCRIPTION

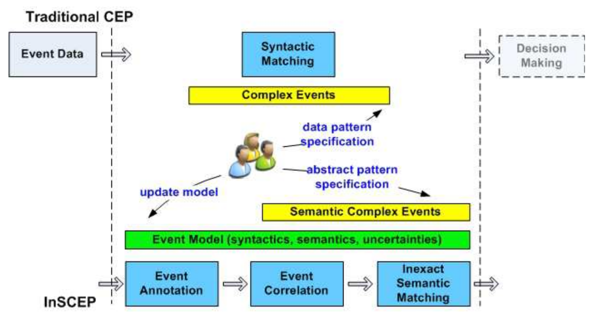

Figure 1: CEP (top) and InSCEP (bottom)

The proposed InSCEP framework for demand response optimization extends traditional CEP with Semantic Web technologies and uncertain pattern detection. As shown in Figure 1, our approach allows users to specify inexact event patterns based on a domain-specific event model at high level. As the system processes events, it *annotates source data* using the semantic event model, correlates events to *derive unobserved events* and then *performs inexact and semantic matching* to detect pattern occurrences.

Semantic and Uncertain Event Model

We develop an event model that captures semantics and uncertainties. Events are modeled using Semantic Web ontologies and represented using the Web Ontology Language (OWL). The event ontology is modular, for easy extension with new domain ontologies. The top level, core event ontology captures concepts and relationships between events. The notion of an event is classified into domain specific classes, like *ThermostatChange* event and *MeterUpdate* event. The second layer of the model are the subjective domain ontologies, which capture the Smart Grid domain entities, including grid networks, meters and appliances. The lowest level of the model captures entities in other relevant domains such as weather and transportation. Event correlations are modeled using rules represented using Semantic Web Rule Language (SWRL). We capture uncertainties of events by annotating probabilities to events and assigning confidences to correlation rules.

A formal event algebra is defined for specifying complex events that incorporate semantic and inexact query features. Some of the algebraic operators such as *Selection*, *Aggregation*, *Projection* and *Renaming* originate from relational algebra and are supported in existing CEP systems [1, 3]. The added expressive power of our algebra lies in the *Semantic* and *Inexact Selection* operators. *Semantic Selection* evaluates pattern constraints based on the semantic equivalence of attribute meanings captured by the event ontology instead of syntactic identical attribute values. *Inexact Selection* selects events and allows a limited number of mismatches to detect relevant patterns. A similarity function is associated with *Inexact Selection* to evaluate relevances between matching patterns and target patterns.

Inexact and Semantic Pattern Detection

Events in InSCEP are represented as RDF triples, accompanied with timestamps and probabilities. Algebraic expressions of complex events are mapped onto an extended version SPARQL query language.

We propose to extend existing automata and tree based matching algorithms to detect different types of inexact patterns. In particular, we generalize the NFA model proposed in [1] to match inexact sequential patterns: (1) state transitions are controlled by using both semantic predicates and similarity evaluations. A transition edge can be traversed *iff* the predicate evaluation is true and the accumulative mismatch is less than a predefined threshold; (2) it allows to skip intermediate states and transit to indirectly connected matching states.

4. DISCUSSION AND FUTURE WORK

CEP is a promising solution for demand response optimization in Smart Grid. Ease of use, expressive pattern definition and scalable pattern matching are key issues to address. We propose to enhance CEP by incorporating semantic and inexact query features to suite DR application needs. Currently we are implementing semantic and inexact pattern matching algorithm into an existing open source CEP system. Future work includes extending the simplified event uncertainty model, and combining CEP with pattern mining to assist automatic pattern discovery. The second problem is interesting because identifying meaningful event patterns manually in an diverse information space like Smart Grid is difficult. In addition, event patterns that correspond to certain interests such as predicting peak demand of a building may constantly change over time, for example, due to changes of owners and weather conditions.

5. REFERENCES

[1] A. Demers et al. Cayuga: A general purpose event monitoring system. In *CIDR*, 2007.

[2] A. Adi, D. Botzer, G. Nechushtai, and G. Sharon. Complex event processing for financial services. In *IEEE Services Computing Workshops*, 2006.

[3] Y. Diao, N. Immerman, and D. Gyllstrom. SASE+: An agile language for Kleene closure over event streams. Technical report, UMass, 2007.

[4] D. O'Keeffe and J. Bacon. Reliable complex event detection for pervasive computing. In *DEBS*, 2010.

[5] Y. Simmhan et al. An informatics approach to demand response optimization in Smart Grids. Technical report, USC, 2011.