Volume 3, Issue 10, October 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Mining Contract Documents to Explore Business Exceptions

G. Mahesh, K. R. K. Satheesh
(M. Tech), Department of Computer Science Engineering,
Madanapalle Institute of Technology & Science,
Kadiri Road, Angallu (Vill),
Madanapalle-517325, Andhra Pradesh, India

Abstract---Here we perform the investigation of the situation of business support selection and categorization according to an extended concept of trust and reputation. It's more popular now that the creation of value is the important tenet of service research. Particularly, the importance of understanding human behavior as a basis for support research is well-recognized, but isn't fundamentally reflected in the technical ways. In specific the quality of experience that a consumer derives from a site engagement depends on how the expectations are refined and met by the supplier during the process. Indeed, this really is well-known in marketing theory as a base of customer care whiles the service quality GAPS model. Here we propose a post mining process for mining company contract documents.

***Keywords—— Apriori, dataset, UDF's,* SQL-OR.**

I. Introduction

The quality of a site diamond would depend on the needs of the participants involved with the business. In specific the quality of experience that a consumer derives from a service engagement depends on how the consumer's expectations are filtered and how well they met the provider during the service engagement process. We postulate that integrating objectives plays a vital part in business contract service selection. We investigate the thought of objectives from a computational perspective. We declare that

(1) Expectations, despite being subjective, certainly are a well-formed, reliably computable notion and
(2) We can calculate targets and use them as a basis for increasing the effectiveness of service selection.

In the base of language processing, XibinGao et al.,[1] suggested an exploration method called Contract Miner that uses designs being a hint to locate support conditions at the phrase level. A learning focused technique and a fundamental structure recognizer too can extract sentences or other linguistic context like a text screen consisting of-the few quickly preceding and following sentences. The-difference between an old-fashioned design recognizer program and Contract Miner method is based on their reasons, abilities, and scopes. But, as ideas for finding exceptions Contract Miner is not considering the patterns themselves, but uses patterns. Next, the pattern recognizer normally works on the surface, and doesn't demand chunking of significant text units, consequently demanding significant individual energy in knowing the text. Meaningful support exceptions are automatically extracted by contract Miner at the term level.

1. Related Work:

Meneguzzi et al.'s attempt[2] could be the main European Union's framework of CONTRACT undertaking, an entire approach to approach, motive about, and enact electronic contracts. Commitment Miner matches the previous work-in two values. 2, Contract Miner introduces the normal company conditions that appear in a website and in this manner provides a foundation for confirming whether its enactments could adapt the conditions and that a particular contract is adequately strong. Groups and Wilson [3] distinguish between the enterprise and functional levels A learning-based strategy plus a fundamental structure recognizer too may extract sentences or alternative linguistic context such as a text screen containing the few immediately preceding and following sentences. The-difference between a mainstream design recognizer system and Contract Miner strategy lies inside their motives, abilities, and scopes. Next, the pattern recognizer generally operates on the area, and doesn't demand chunking of important text units, thus requiring substantial human power in understanding the text. Substantial service conditions are automatically produced by contract Miner in the phrase level. But, present methods on contracts and exceptions do not interface effectively with the 'heritage' of text based contracts, which can be just how all major organization is still being performed today. Study within the automated extraction of conditions from contracts has been unusual, as well as non-existent regardless of exhortations of investigators for instance Karlapalem [10] and Radha Krishna. Of the contract. At the functional level, and service-level agreements, a contract might be indicated as plans, permits. At-the organization level, a contract is picked by contract attorneys and done by the participating businesses. Problems are analyzed within the specific context of quality, representation, and identification. Molina Jimenez et al., [7] add style for exclusion quality. Poon [8] and Grosof indicate company contracts in RuleML, for that reason empowering brokers to appraise, mechanically produce, examine, and execute contracts and also to manage conditions. Klein et 's., [9] explain a technique for determining exceptions and finding appropriate responses for all these conditions. More, Klein et 's., [9] suggest taxonomy of conditions.

2. Service Event Extraction From Business Contracts

2.1 Service Events Extraction

The setting that people considered to remove service events is made in [1]. The pursuit of the same is follows. The preprocessing of the given business contract data is the initial step of the procedure. Within this step the formatting elements of the given information is removed and components plain text. More in step 2, the sentences from that plain text is likely to be extracted. In the resulting sentences of step2, the sentences that representing the events will undoubtedly be submitted and extracted to construct noun phrases and recognizes their relation with events, which is step3.

Within this approach, sentences which contains signal words such as 'may', 'shall', 'should' are produced as events. Further NLP scenarios are used to find the noun phrases and to chunk them. For detail by detail exploration of the methods, relate [1].

2.2 Service Event Importance Evaluation and Classification (SESEC) Method

Course Descriptor: The site specialist classifies the features involved with business rules will soon be classified in to different categories. The method of group as follow

1.The website specialist also starts to obtain the relationship between classes. The relationship may be between any two classes, such as

2.Based on characteristic properties, qualities is likely to be categorized into a class.

.Ex: if all of the attribute 'a' properties matched to school 'c' then

3.Initially courses will soon be derived on the basis of the properties; thus each class includes group of properties.

The proposed article exploration approach named "Exception Significance Evaluation and Classification" explained in detail by detail here. Table 1 presents the notations found in ESE Approach.

a. Relation between class and sub-class of other class

b.Relation between two immediate courses

c.Relation between two sub-classes

A xml based credit school descriptor will be prepared. Fig1 shows an illustration descriptor. Notations information equations can within table 1 that fallows.

```
<class-descriptor>
        <properties>
                <property id=1>
                        <name>prop1</name>
                </property>
                <property id=2>
                        <name>prop2</name>
                </property>
                <property id=3>
                        <name>prop3</name>
                </property>
        </properties>
        <attributes>
                <attribute name="item-1" id="1" properties="{list of property ids}" />
                <attribute name="item-2" id="2" properties="{list of property ids}" />
                <attribute name="item-3" id="3" properties="{list of property ids}" />
                .
                .
                <attribute name="item-n" id="n" properties="{list of property ids}" />
        </attributes>
        <classes>
                <class name="class-1" id="1" properties="{list of property ids}" />
                <class name="class-2" id="2" properties="{list of property ids}" />
                <class name="class-3" id="3" properties="{list of property ids}" />
                .
                .
                <class name="class-n" id="n" properties="{list of property ids}" />
        </classes>
        <child-classes>
                <!--parent value must be unique -->
                <child-class parent="class-id" child="{list of class ids}"/>
                .
                .
                <child-class parent="class-id" child="{list of class ids}"/>
        </child-classes>
```



```
        <relations>
                <!--lhs value must be unique -->
                <!--
                classes that related to a child-class also related to it's parent class
                -->
                <relation lhs="class-id" rhs="{list of class ids}" />
                .
                .
                <relation lhs="class-id" rhs="{list of class ids}" />
        </relations>
</class-descriptor>
```

Fig 1: Class-descriptor

Table1: Notations used in Exception Significance Evaluation Approach

1	r_{lhs}	source attributes of the Exception r
2	r_{rhs}	Target attributes of the exception r
3	RS	Exception set
4	cpc_c	Class properties count of class c
5	apc_a	Attribute property count of attribute a
6	psd	Property support degree
7	tp_c	Total properties of class c
8	ps_a $ps_a = \frac{apc_a}{cpc_c}$	Property support of attribute a of class c
9	$psd_a = \frac{ps_a}{tp_c}$	where $a \in c$
10	rs_c	Relation support of class c is max threshold value 1.
11	ARS	Attribute Relation support
12	$ARSD_i$	Attribute Relation support Degree of attribute set i.
13	If attributes a_i and a_j belongs to same class c	$ARS(a_i, a_j) \cong 1$, $where\ \{a_i, a_j\} \in c$
14	If attribute a_i belongs to class c_i and attribute a_j belongs to class c_j, c_i and c_j relation is true then	$ARS(a_i, a_j) = \frac{psd_{a_i} + psd_{a_j}}{rs_{c_i} + rs_{c_j}}$
15	If c_i and c_j relation is false	$ARS(a_i, a_j) = 0$,
16	$ARS(a_i, a_j) \cong ARS(a_j, a_i)$	Applicable in all cases such as both belongs to same class, both belongs to different classes that are not having relation and both belongs to different classes that are having relation
17	No of attribute pairs in an attribute set	pc

18	$pc = 0;$ $\sum_{i=1}^{n-1} pc + i \quad (or) \quad pc = \frac{n-1}{2} \times n$	pc : is total number of pairs in an attribute set involved in an exception. n : is total number of attributes in given exception
19	$pc_{r_{lhs}}$	Pair-count of attribute set, which is lhs of exception r .
20	$pc_{r_{rhs}}$	Pair-count of attribute set, which is rhs of exception r .
21	$pc_{r_{lhs} \cup r_{rhs}}$	Pair-count of attribute set that generated from $r_{lhs} \cup r_{rhs}$
22	$PS_i = \{p_1, p_2, ..., p_m\}$	Pair set that generated from attribute set i
23	$ARS(p_i)$	Attribute relation support of pair p_i .
24	$ARSD_i = \frac{\sum_{k=1}^{\lvert ps_i \rvert} ARS(p_k)}{\lvert ps_i \rvert}$	Attribute relation support degree of attribute set i And $p_k \in ps_i$, here p_k is k[th] pair of pair-set ps of attribute set i
25	RC_r	Relation confidence of exception r .
26	$ARSD_{r_{lhs}} = \frac{\sum_{k=1}^{\lvert PS_{r_{lhs}} \rvert} p_k}{\lvert PS_{r_{lhs}} \rvert}$	Attribute relation support degree of attribute set, which is lhs of exception r And $p_k \in ps_{r_{lhs}}$, here p_k is k[th] pair of pair-set ps of attribute set lhs of exception r
27	$ARSD_{r_{rhs}} = \frac{\sum_{k=1}^{\lvert PS_{r_{rhs}} \rvert} p_k}{\lvert PS_{r_{rhs}} \rvert}$	Attribute relation support degree of attribute set, which is rhs of exception r And $p_k \in ps_{r_{rhs}}$, here p_k is k[th] pair of pair-set ps of attribute set rhs of exception r
28	$ARSD_{r_{lhs} \cup r_{rhs}} = \frac{\sum_{k=1}^{\lvert PS_{r_{lhs} \cup r_{rhs}} \rvert} p_k}{\lvert PS_{r_{lhs} \cup r_{rhs}} \rvert}$	Attribute relation support degree of attribute set that generated from $lhs \bigcup rhs$ of exception r And $p_k \in PS_{r_{lhs} \cup r_{rhs}}$, here p_k is k[th] pair of pair-set ps of attribute set that generated from $lhs \bigcup rhs$ of exception r
29	$rc_r = \frac{ARSD_{r_{lhs} \cup r_{rhs}}}{ARSD_{r_{lhs}}}$	Relation confidence of r is the coefficient emerged as result when 30attribute support degree of all attribute involved in exception r is divided by attribute support degree of exception r 's lhs

ESEC algorithm:

Input: Rule set RS , Class Descriptor CD and relation confidence threshold rct

Output: Significant Rule set RS' which is subset of RS $RS' \subseteq RS$

Begin:

While RS is not empty

Begin:

Read a rule r from RS

Find property support ps_a and property support degree psd_a of each attribute a of r_{lhs} [Table 1 row: 8, 9]

Find property support ps_a and property support degree psd_a of each attribute a of r_{rhs} [Table 1 row: 8, 9]

Find unique two attribute pair set $PS_{r_{lhs}}$ from r_{lhs} [Table 1 row: 22]

Find Attribute relation support ARS_p for each pair p , where $p \in PS_{r_{lhs}}$ [Table 1; row: 13, 15, 15 and 16]

Find Attribute relation support degree $ARSD_{r_{lhs}}$ of r_{lhs} [Table 1; row: 24, 26].

Find unique two attribute pair set $PS_{r_{lhs}}$ from r_{lhs} [Table 1 row: 22]

Find Attribute relation support ARS_p for each pair p, where $p \in PS_{r_{rhs}}$ [Table 1; row: 13, 15, 15 and 16]

Find Attribute relation support degree $ARSD_{r_{rhs}}$ of r_{lhs} [Table 1; row: 24, 27].

Find unique two attribute pair set $PS_{r_{lhs} \cup r_{rhs}}$ from $r_{lhs} \bigcup r_{rhs}$ [Table 1 row: 22]

Find Attribute relation support ARS_p for each pair p, where $p \in PS_{r_{lhs} \cup r_{rhs}}$ [Table 1; row: 13, 15, 15 and 16]

Find Attribute relation support degree $ARSD_{r_{lhs} \cup r_{rhs}}$ of $r_{lhs} \cup r_{rhs}$ [Table 1; row: 24, 28].

Find unique two attribute pair set $PS_{r_{lhs} \cup r_{rhs}}$ from $r_{lhs} \bigcup r_{rhs}$ [Table 1 row: 22]

Find Attribute relation support ARS_p for each pair p, where $p \in PS_{r_{rhs} \cup r_{lhs}}$ [Table 1; row: 13, 15, 15 and 16]

Find Attribute relation support degree $ARSD_{r_{rhs} \cup r_{lhs}}$ of $r_{lhs} \cup r_{rhs}$ [Table 1; row: 24, 28].

Find Relation confidence rc_r of rule r

If $rc_r \geq rct$ then add rule r to resultant Exception set RS'

End

End

Fig 2: Exception Significance Evaluation algorithm

Property Support: No of attribute properties are matched to number of class properties to which that attribute belongs to [Table 1 row: 8].

Property Support degree: indicates the ratio of properties matched to class level properties [table 1 row: 9].

Ex: $psd_a = \frac{ps_a}{cpc_c}$; here a is an attribute of class c $[a \in c]$

Attribute Relation support: Indicates the strength of the relation between two attributes of an itemset that are considered as pair for equation [see table 1 row: 11, 13].

Pair Count: Total number of two attributes sets; here these attribute sets must be unique [see table 1, row 17, 18]

Attribute Relation support degree: is an itemset level measurement representing average relation strength of the attributes those belongs to an itemset [see table 1 row: 24]

Relation confidence: is a rule level measurement concludes the relation strength between left hand side itemset and right hand side itemset of a given rule[see table 1 row: 29]

1. Performance Analysis:

That section mainly focuses on providing evidence on alleging that the proposed post mining strategy SESEC is skilled enough to momentously classify service events of the business contracts. On the basis of the surveillance performed on the refining service exceptions and as domain specific segregating, proves that SESEC execution is much more noteworthy and important in regard to label the service events in business contracts.

JAVA 1.6_ 20th build was employed for success of the ARA together with PEPP under analysis. A workstation equipped with core2duo processor, 2GB RAM and Windows XP installation was made use of for analysis of the algorithms. The simultaneous reproduction was started to ultimately achieve the bond idea in JAVA.

Dataset Characteristics:

We used same experiments platform described in [1]. To evaluate the effectiveness of SESEC in extracting exception phrases, we manually annotate the following five (arbitrarily selected) manufacturing contracts from the Onecle repository. We consider a corpus of 2,647 contracts from Onecle for some evaluations. As Table 1 shows, our pattern sentences are prevalent in contracts across seven major domains of interest to services.

Table 1: Description of the corpus used in experiments

Contract Type	Contracts	Matches	Average
Licensing	1,364	3,838	2.8
Consulting	501	509	1.0
Outsourcing	9	21	2.3
Supply	207	733	3.5
Manufacturing	206	577	2.8
Purchase	142	591	4.1
Stock Options	218	1,153	5.3
Overall	*2,647*	*7,422*	*2.8*

We used service event significance evaluation and classificationaccuracy (the percentage of conceptually validevent categorization by the proposal) as the main performance measure. In addition to accuracy, we employed precision, recall, and F-measureused in the information retrieval fieldas performance measures; these additional measures are defined using Eq1, Eq2 and Eq3. . We measured the SESEC accuracyas follows:

$$\text{S(sesec)} = \frac{\text{Events correctly Evaluated as significant}}{\text{Events Actually Significant}}$$

$$pr = \frac{t_+}{t_+ + f_+} \quad \text{..... Eq1}$$

Here in Eq1 the pr indicates the precision, t_+ indicates the true positives and f_+ indicates the false positive

$$rc = \frac{t_+}{t_+ + f_-} \quad \text{..... Eq2}$$

Here in the Eq2, the ' rc ' indicates the recall, ' f_- ' indicates the false negative.

$$F = \frac{2 * pr * rc}{pr + rc} \quad \text{..... Eq3}$$

Here in the Eq3, ' F ' indicates the F-measure.

Table 1: Precision, recall and F-measure values for event significance evaluation and categorization done by human perception and by proposed SESEC approach.

	Precision	Recall	F-Mesure
Manual event significance analysis and categorization	0.9976	0.9983	0.998
Service Event Significance Evaluation and Categorization(SESEC)	0.9974	0.9979	0.99785

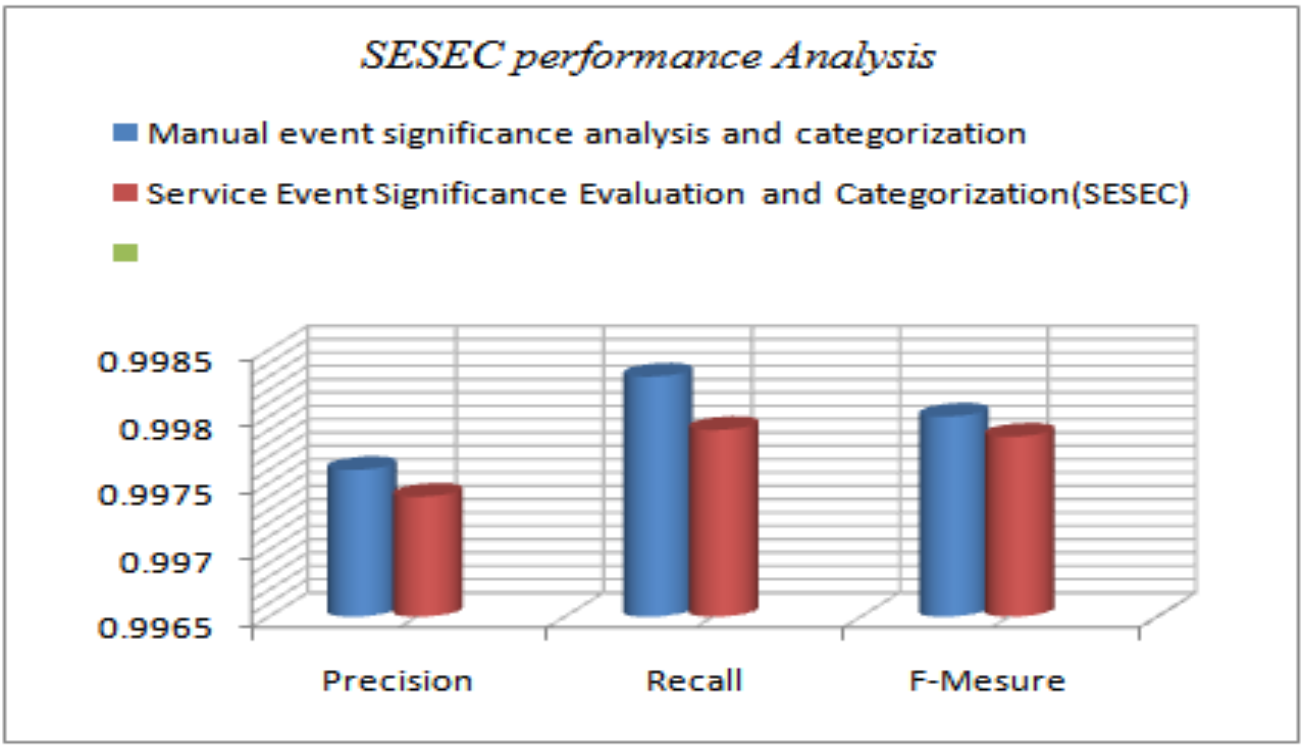

Fig1: SESEC performance

The compatibility of the proposed SESEC approach with manual process by domain experts is optimistic and considerable (see fig 1). Hence, SESEC is reliable text mining process to evaluate and categorize the events of the business contracts.

II. Conclusion

Here in this paper, we proposed a novel post mining process to evaluate the significance and categorize the events in business contracts. The process of extracting events from business contracts is taken from the recent work explored in [1], which is originally a text mining process that relies on NLP. Further we devised a post mining process called Service Event Evaluation and that evaluates the significance of the extracted events and categorizes the same. The experimental results explored the performance of the proposed SESEC as reliable and stable. Further this work can be extended to devise a machine learning model that scores the complexity of the business agreement by depending the event categories and significance

References

[1] XibinGao, Munindar P. Singh, PankajMehra, "Mining Business Contracts for Service Exceptions," IEEE Transactions on Services Computing, vol. 5, no. 3, pp. 333-344, Third Quarter 2012, doi:10.1109/TSC.2011.1

[2] F. Meneguzzi, S. Miles, C. Holt, M. Luck, N. Oren, N. Faci, and M. Kollingbaum, "Electronic Contracting in Aircraft Aftercare: A Case Study,"Proc. Seventh Int'l Conf. Autonomous Agents and MultiagentSystems,pp. 63-70, 2008.

[3] A. Arenas and M. Wilson, "Contracts as Trust Substitutes in Collaborative Business,"Computer,vol. 41, no. 7, pp. 80-83, July 2008.

[4] H. Ludwig, A. Keller, A. Dan, R. King, and R. Franck, "Web Service Level Agreement (WSLA) Language Specification," http:// www.research.ibm.com/wsla/WSLASpecV1-20030128.pdf, 2003.

[5] V. Tosic, B. Pagurek, K. Patel, B. Esfandiari, and W. Ma, "Management Applications of the Web Service Offerings Lan-guage (WSOL),"Information Systems, vol. 30, no. 7, pp. 564-586, 2005.

[6] G. Gangadharan, V. D'Andrea, R. Iannella, and M. Weiss, "ODRL Service Licensing Profile (ODRL-S)," Proc. Fifth Int'l Workshop Technical, Economic, and Legal Aspects of Business Models for Virtual Goods,2007.

[7] C. Molina-Jimenez, S. Shrivastava, and M. Strano, "Exception Handling in Electronic Contracting,"Proc. IEEE Conf. Commerce and Enterprise Computing (CEC),pp. 65-73, 2009.

[8] B. Grosof and T. Poon, "SweetDeal: Representing Agent Contracts with Exceptions Using XML Rules, Ontologies, and Process Descriptions,"Proc. 12th Int'l Conf. World Wide Web,pp. 340-349, 2003.

[9] M. Klein, C. Dellarocas, and J. Rodriguez-Aguilar, "A Knowledge-Based Methodology for Designing Robust Electronic Markets," Proc. First Int'l Conf. Autonomous Agents and Multiagent Systems (AAMAS),2002.

[10] P. Radha Krishna and K. Karlapalem, "Electronic Contracts,"IEEE Internet Computing,vol. 12, no. 4, pp. 60-68, July/Aug. 2008.

