Expert Systems With Applications 90 (2017) 332–343

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Enhancing semantic consistency in anti-fraud rule-based expert systems☆

María del Mar Roldán-García, José García-Nieto*, José F. Aldana-Montes

Dept. de Lenguajes y Ciencias de la Computación, University of Málaga, ETSI Informática, Campus de Teatinos, Málaga 29071, Spain

ARTICLE INFO

Article history:
Received 21 June 2017
Revised 18 August 2017
Accepted 19 August 2017
Available online 21 August 2017

Keywords:
Semantic model
Ontology reasoning
Rule-based expert system
Fraud detection expert systems

ABSTRACT

In this study, an ontology-driven approach is proposed for semantic conflict detection and classification in rule-based expert systems. It focuses on the critical case of anti-fraud rule repositories for the inspection of Card Not Present (CNP) transactions in e-commerce environments. The main motivation is to examine and curate anti-fraud rule datasets to avoid semantic conflicts that could lead the underpinning expert system to incorrectly perform, e. g., by accepting fraudulent transactions and/or by discarding harmless ones. The proposed approach is based on Web Ontology Language (OWL) and Semantic Web Rule Language (SWRL) technologies to develop an anti-fraud rule ontology and reasoning tasks, respectively. The three main contributions of this work are: first, the creation of a conceptual knowledge model for describing anti-fraud rules and their relationships; second, the development of semantic rules as conflict-resolution methods for anti-fraud expert systems; third, experimental facts are gathered to evaluate and validate the proposed model. A real-world use case in the e-commerce (e-Tourism) industry is used to explain the ontological knowledge design and its use. The experiments show that ontological approaches can effectively discover and classify conflicts in rule-based expert systems in the field of anti-fraud applications. The proposal is also applicable to other domains where knowledge rule bases are involved.

1. Introduction

Rule-based Expert Systems (RBESs) are the simplest form of artificial intelligence, which uses rules as the representation for encoding knowledge from a fairly narrow area into an automated system (Durkin, 1998). RBESs mimic the reasoning procedure of a human expert when solving a knowledge-intensive problem. A rule-based system consists of a set of `IF-THEN` rules, a set of *facts* and an interpreter controlling the application of the rules, given the facts. Rule-based systems are very simple models and can be adapted and applied to a wide set of different problems, whenever the domain of knowledge can be expressed in the form of `IF-THEN` rules.

In the case of fraud prevention and detection in e-commerce transactions, RBESs are used to identify customers' suspicious activities by automatically generating risk scoring reports of their transactions (Ketkar, Shankar, & Banwet, 2014). They analyze behaviors such as repetitive and quick access attempts, domestic/foreign transactions, and abnormal transactions compared with the customer‘s past behavior. A final decision is then delivered by the system, commonly: *Accept*, *Reject*, or *Revise*. A small subset of rules that might contribute to a negative risk assessment could be as follows (Ward, 2010): A single IP address has been used with multiple payment cards in the last few days; the shopper‘s billing address is more than “x” kilometers from the shipping address; the e-mail address has been flagged in a negative database (black list) of known fraud activity by other merchants participating in the same fraud detection strategy; the BIN (Bank Identification Number) on the payment card indicates the transaction comes from a high-risk country.

Using a combination of these and many other factors could benefit e-merchants, who are presently demanding autonomous expert systems, to quickly update their rule-bases and flag suspicious transactions (Wong, 2013). In the current market, there exist a series of tools that use rule-based knowledge engines for

---

☆ This work is partially funded by FP7 EU project SME-Ecompass under Grant No: 315637. It is also partially funded by Grants TIN2014-58304 (MINECO) and Regional projects P11-TIC-7529/P12-TIC-1519. Authors specially thanks to etravel.com and in particular to Orestis Papadopoulos to support this work with a set of private rules for testing and validation. José García-Nieto is recipient of a Post-Doctoral fellowship of “Captación de Talento para la Investigación” at Universidad de Málaga.

* Corresponding author.

E-mail addresses: mmar@lcc.uma.es (M.d.M. Roldán-García), jnieto@lcc.uma.es (J. García-Nieto), jfam@lcc.uma.es (J.F. Aldana-Montes).

http://dx.doi.org/10.1016/j.eswa.2017.08.036
0957-4174/

risk scoring of e-commerce fraud: Simility,[1] Subuno,[2] Riskfield,[3] and Trustev.[4] These tools are widely used not only for tracking and scoring transactions, but also for reporting statistics of the e-commerce site. However, these tools often concentrate on high level and generic sets of rules, without the possibility of considering new *ad hoc* rules specific to each e-commerce site. When these are provided, they are available only for commercial (non-free) versions, which are rarely accessible to SMEs or individual e-merchants.

In this context, the SME-Ecompass European initiative[5] aims to provide e-commerce SMEs with accessible tools for specialized fraud prevention and detection. These software facilities are built on a rule-based expert system for the risk scoring of Card Not Present transactions (CNP). The knowledge rule base can be easily updated by the e-merchant by inserting new rules specific to his/her own e-commerce site. Nevertheless, an increasing number of anti-fraud rules (and their combinations) often provoke the emergence of conflicting rules with semantic inconsistencies. In addition, anti-fraud expert systems face a major challenge as they operate in hostile conditions, as their anticipated inference capabilities are degraded with a continuously changing environment. As a consequence, these issues can lead the underpinning expert system to perform inefficiently (Grosan & Abraham, 2011), e.g., by accepting fraudulent transactions, while discarding harmless ones. Therefore, a key task in anti-fraud applications is to inspect and curate knowledge rule bases to avoid semantic inconsistencies, before delivering a final diagnosis.

With this motivation, an ontology-driven approach is proposed for semantic inconsistent detection and classification in rule-based expert systems. The proposed approach is based on Web Ontology Language (OWL) and Semantic Web Rule Language (SWRL) technologies to develop an anti-fraud rule ontology and to perform reasoning tasks, respectively. The three main contributions of this work are:

(i) creating a conceptual knowledge model, in terms of an OWL ontology, to describe anti-fraud rules and their relationships;
(ii) developing semantic SWRL rules as conflict/inconsistency detection methods for anti-fraud expert systems;
(iii) gathering experimental facts to evaluate and validate the proposal.

A real-world use case in the e-commerce (e-Tourism) industry is used to explain the ontological knowledge design and its uses. The experiments show that the proposed semantic approach can effectively discover and classify inconsistencies and conflicts in rule-based expert systems, in the field of anti-fraud applications.

The remainder of this article is organized as follows. The next section presents background concepts and related works in the literature. In Section 3, key concepts in a real-world anti-fraud expert system are explained. Section 4 describes the proposed semantic approach, giving details of the OWL Ontology and the reasoning model. The validation procedure is reported in Section 5. The main conclusions and future work are given in Section 6.

2. Background and literature overview

This section describes the main background concepts and reviews related works in the specialized literature.

Table 1
Basic OWL-DL semantic syntax used to formally define the proposed ontology.

Descriptions	Abstract syntax	DL syntax
Operators	*intersection*$(C_1, C_2, \cdots, C_n)$	$C_1 \sqcap C_2 \sqcap \cdots C_n$
	union$(C_1, C_2, \cdots, C_n)$	$C_1 \sqcup C_2 \sqcup \cdots C_n$
Restrictions	for at least 1 value V from C	$\exists V.C$
	for all values V from C	$\forall V.C$
	R is Symmetric	$R \equiv R^{-}$
Class Axioms	A *partial*$(C_1, C_2, \cdots, C_n)$	$A \sqsubseteq C_1 \sqcap C_2 \sqcap \cdots C_n$
	A *complete*$(C_1, C_2, \cdots, C_n)$	$A \equiv C_1 \sqcap C_2 \sqcap \cdots C_n$

2.1. Background concepts

- *Ontology.* Ontologies provide a formal representation of the real world by defining concepts and relationships between them (Gruber, 1993). In the context of the computer and information sciences, an ontology defines a set of representational primitives with which to model a domain of knowledge. These primitives are typically concepts (or classes), attributes (or properties), class members (class instances) and relationships (property instances). The definitions of the primitives include information about their meaning and constraints on their logically consistent application.

- *RDF.* Graphical language used to represent information about resources on the web (Staab & Studer, 2009). It is a basic ontology language. Resources are described in terms of properties and property values using RDF statements. Statements are represented as triples, consisting of a subject, predicate and object. The RDF Schema (RDFS) (Staab & Studer, 2009) "semantically extends" RDF to enable us to talk about classes of resources, and the properties that will be used with them.

- *SPARQL.* It is a query language for ontology models and databases, capable of extracting and manipulating information stored in RDF format. Essentially, SPARQL is a graph-matching query language that can be used to extract knowledge from a model like the one proposed in this article. Given a data source D, a query is a pattern, which is matched against D. The combinations of values resulting from this matching constitute the result of the query (Pérez, Arenas, & Gutierrez, 2009).

- *OWL.* In 2004, the W3C ontology working group (Dean & Schreiber, 2004) proposed OWL as a semantic markup language for publishing and sharing ontologies. From a formal point of view, OWL is equivalent to a very expressive description logic where an ontology corresponds to a Tbox (Gruber, 1993). This equivalence allows the language to exploit description logic researcher results. OWL extends RDF and RDFS. When compared to RDF models, OWL adds more vocabulary for describing properties and classes, among others: relationships between classes (e.g. disjointedness), cardinality (e.g. "exactly one"), equality, richer typing of properties, characteristics of properties (e.g. symmetry), and enumerated classes (McGuinness & Harmelen, 2004).

- *OWL-DL.* It is a syntactic variant of the SHOIN (D) description logic (Haase & Stojanovic, 2005) with a different terminology to OWL, which is based on RDF(S). Therefore, it supports data values, data types and data type properties. OWL-DL restricts OWL in two distinct ways (Horrocks & Patel-Schneider, 2003): first, some syntactic constructs, e.g., recursive descriptions in them are not permitted; second, classes, individuals and properties (respectively concepts, individuals and roles in description logics) must all be disjoint. In this approach, we use the OWL-DL syntax to formalize the proposed ontology for our semantic model. A description of the basic OWL-DL semantic syntax is shown in Table 1, where an informal logic syntax is represented (left-hand column) with regards to the corresponding OWL-DL equivalent (right).

---

[1] In URL https://simility.com/.

[2] In URL http://www.subuno.com/.

[3] In URL http://www.riskified.com/.

[4] In URL http://www.trustev.com/.

[5] SME-Ecompass FP7 European initiative http://www.sme-ecompass.eu/.