Information Sciences 328 (2016) 455–484

Contents lists available at ScienceDirect

Information Sciences

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

A feature-based trust sequence classification algorithm

Hamdi Yahyaoui*, Aisha Al-Mutairi

Computer Science Department, Kuwait University, Kuwait

ARTICLE INFO

Article history:
Received 28 December 2014
Revised 8 July 2015
Accepted 9 August 2015
Available online 9 September 2015

Keywords:
Service
Trust
Pattern
Sequence classification

ABSTRACT

Trust is a paramount factor in the development of service-based communities, where services continuously collaborate to successfully perform their tasks. Trust assessment helps users and services identify which partners to interact with. We tackle in this paper trust from an objective data mining perspective. We propose a novel feature-based approach to assess the trust behavior of a service. A trust behavior is represented as a sequence of trust observations during a certain time frame. By analyzing the possible trust behaviors of services, trust patterns are defined to describe trust sequences based on three criteria: its overall behavior, the starting behavior and ending behavior. Our approach spans over a rule based Prefix-Suffix Algorithm (PSA) for the classification of trust sequences. PSA computes new attributes to capture the chronological and structural nature of trust. Following a divide and conquer strategy, the trust sequence is divided into two parts each classified independently. PSA leverages some predefined merging rules to derive the class of the whole trust sequence from the classification results of these parts. We show the efficiency and accuracy of our approach by analytical and experimental evaluation.

1. Introduction

In human communities, a person usually has little trust in strangers and thus tends to avoid relying on them. Similarly in any service-based community with an increasing number of services, many are unidentified to each others and are considered as strangers [5]. Lack of trust hinders interactions between services. This is one of the reasons from which trust has drawn significant attention in the research community including e-commerce [88], social networks [50], P2P networks [43], etc. In the area of service computing, Chang et al. [15] present technologies for the assessment of trust and reputation in service-oriented environments. Malik and Bouguettaya [54] propose the assessment of a service's reputation using a Hidden Markov Model (HMM). Yahyaoui and Zhioua [86] suggest behavior analysis for trust evaluation of a service using a pattern-based technique with HMMs.

Zhang et al. present a classification scheme for trust in [90], which discriminates between objective and subjective trust. They describe objective trust as the trustworthiness of an entity when it is possible to measure the quality of the entity objectively. The subjective trust of an entity is measured by recommendations from other entities, based on previous experience and interactions with this entity. Our approach measures the trustworthiness of a service from an objective perspective.

Behavior analysis has been increasingly attracting attention from researchers in psychology, computer science, linguistics, neuroscience, and other scientific fields. Behavior analysis made great discoveries in e-business, social security surveillance, workplace safety, education and many applications. It is the answer to why and how behavior changes over time and studies the factors that have most influence on an entity's behavior. In computer science, behavior analysis is being adopted for problem

* Corresponding author. Tel.: +95624985325.
E-mail addresses: hamdi@cs.ku.edu.kw (H. Yahyaoui), aisha@cs.ku.edu.kw (A. Al-Mutairi).

http://dx.doi.org/10.1016/j.ins.2015.08.008
0020-0255/

solving and development. To help in system design, development, evaluation and improvement, machine learning techniques were employed in extracting data from console logs and error reports to describe system behavior and analyze it [7,26,84]. User clicks behavior in sponsored search is analyzed in [40] to help advertisers raise the quality of their ads. In trust related research, a framework for automatic analysis of malware behavior using machine learning is introduced in [67], and the abnormal internet behaviors are analyzed to detect mobile malware in [17]. Behavior modeling and behavior pattern analysis are addressed for web services in the bootstrapping approach presented in [86].

In this context, behavior can be defined as the recognizable pattern in a sequence of observations of activities or events [57]. Since service interactions play an important role in the assessment of its trustworthiness, they are collected over time as a sequence of observations describing its behavior. Behavior analysis for trust evaluation of services was first proposed by Yahyaoui and Zhioua [86]. In their work, a service is evaluated during a certain time frame in which a sequence of trust observations was built. The authors specified five trust patterns to identify the possible trust behaviors, then used a Hidden Markov Model (HMM) pattern-based approach to assign to each sequence one of predefined trust patterns.

As opposed to their model-based approach, we follow a feature-based approach for trust sequence classification. Each service is represented by a sequence of collected trust observations over a specified time interval describing its trust behavior. Data mining and machine learning algorithms offer a number of effective approaches to design sequence classifiers when a training set of labeled sequences is available [7]. To label the sequences based on behavior analysis, a feature-based classification technique is proposed and a rule-based algorithm is developed. Classification rules are very expressive and easy to understand giving clear conclusions, which justifies its selection as a classification approach in the trust behavior analysis. To build the classification rules, a number of features are defined and constructed from the sequences with strong relation to a number of predefined trust patterns. We identify the possible service behaviors from our perspective by specifying trust patterns extended from those defined in [86]. Our contributions can be summarized as follows:

1. We define eleven trust patterns describing the possible trust behaviors of trust sequences.
2. We generate eight attributes to discriminate trust sequences from one another.
3. We construct classification rules called Prefix-Suffix rules to assign to each trust sequence the trust pattern that best describes its behavior.
4. We propose a feature-based algorithm called the Prefix-Suffix algorithm to classify trust sequences using two sets of rules: the Prefix-Suffix rules and the general merge-rules. The algorithm splits a sequence into two parts classifying each part locally using Prefix-Suffix rules then merges the two classes using the general merge rules.

We consider that splitting the classification of a trust sequence into the classification of its subsequences is one of the appealing features of our approach. The devised Prefix-Suffix algorithm, which implements this splitting, enjoys a high classification accuracy rate. Another significant feature is the descriptive nature of the trust patterns defined and the expressiveness of the classification rules, which helps in understanding the underlying trust behaviors of the classified services.

The rest of this paper is organized as follows: We review the related work in Section 2. Section 3 discusses the feature-based sequence classification approach used to label the trust sequences. The approach consists of defining the trust patterns, generating the attributes, constructing the classification rules and applying the Prefix-Suffix Algorithm. In Section 4, an experimental study is presented to evaluate the results of labeling the trust sequences using the Prefix-Suffix Algorithm and the effect of feature selection on our approach. We conclude our work in Section 5 and present the possible future work.

2. Related work

Our literature review includes two parts: a brief review of trust assessment approaches and sequence classification methods.

2.1. Trust assessment

The issue of trust has been exhaustively discussed in literature from different perspectives for the purpose of trust-based decision making [30,48]. However, there is no clear consensus on how to define trust because it was studied in various domains for different contexts. Aljazzaf et al. [5] present a trust definition in the online world and according to their review, the most cited definition of trust is the one proposed by Mayer et al. [56] which states that "*Trust is the willingness of a party to be vulnerable to the action of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective to the ability to monitor or control that other party*".

Dragoni [25] presents a comprehensive review of trust-based web service selection approaches and categorizes these approaches according to their rationale. He discusses the limitations of each category. Wang and Vassileva [80] present a systematic review of different trust and reputation systems suggesting a typology for their classification from three perspectives: centralized vs. decentralized, persons/agents vs. resources and global vs. personalized. Josang et al. [36] provide an in-depth survey of trust and reputation systems for internet transactions. Li and Wang [48] present a thorough survey on trust evaluation in web applications and web services. They suggest different taxonomies to categorize trust evaluation approaches from different viewpoints: the trust evaluation technique, the trust structure and the trust bases. Wahab et al. [79] perform a comprehensive survey for trust and reputation models in the context of web services. They classify these models based on the web service architecture they target and the trust evaluation technique.