Expert Systems with Applications 41 (2014) 2269–2285

Contents lists available at ScienceDirect

Expert Systems with Applications

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

From blurry numbers to clear preferences: A mechanism to extract reputation in social networks ☆

Ramón Hermoso [a,*], Roberto Centeno [b], Maria Fasli [a]

[a] School of Computer Science and Electronic Engineering, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK
[b] Dpto. de Lenguajes y Sistemas Informáticos, UNED, c/Juan del Rosal 16, 28040 Madrid, Spain

ARTICLE INFO

Keywords:
Reputation
Pairwise elicitation
Tournaments
Social networks

ABSTRACT

Complex social networks are typically used in order to represent and structure social relationships that do not follow a predictable pattern of behaviour. Due to their openness and dynamics, these networks make participants continuously deal with uncertainty before any type of interaction. Reputation appears as a key concept helping users to mitigate such uncertainty. Most of the reputation mechanisms proposed in the literature are based on numerical opinions (ratings), and consequently, they are exposed to potential problems such as the subjectivity in the opinions and their consequent inaccurate aggregation. With these problems in mind, this paper presents a reputation mechanism based on the concepts of pairwise elicitation processes and knock-out tournaments. The main objective of this mechanism is to build reputation rankings from qualitative opinions, thereby removing the subjectivity problems associated with the aggregation of quantitative opinions. The proposed approach is evaluated with different data sets from the MovieLens and Flixster web sites.

1. Introduction

Complex Social Networks (CSNs) are typically used in order to represent and structure social relationships (Vega-Redondo, 2007). There are many types of CSNs depending on the type of social relationship they describe, such as professional links, friendships, leisure, and so on. The term *complex* refers to the impossibility of predicting the network behaviour over time due to the dynamism and complexity of the underlying social structure, as well as the large number of participants. As it is the case with any other complex system, here the concept of trust appears as the cornerstone in the process of selecting entities (partners or resources) in order to mitigate uncertainty.

Trust is typically assessed as a combination of local experiences and opinions gathered from others (reputation) (Fullam & Barber, 2007). In domains where repeated interactions among the same counterparts are rare, and hence, individuals cannot rely just on their own experiences to formulate an opinion or trust valuations about others (e.g., CSNs), reputation becomes more and more important. In particular, reputation has been applied in social networks ranging from a computational analysis point of view (fostering cooperation (Fu, Hauert, Nowak, & Wang, 2008) or exploiting the position of individuals in the social network to enhance reputation extraction (Pujol, Sangüesa, & Delgado, 2002)) to the study of sociological aspects (Raub & Weesie, 1990).

Traditional reputation mechanisms used in CSNs, such as: eBay,[1] Amazon,[2] TripAdvisor,[3] etc., usually deal with quantitative opinion exchange based on numerical ratings and textual feedback from the users. Putting aside textual feedback, this is out of the scope of this paper, such systems allow users to provide their ratings – in certain fixed intervals – which are subsequently aggregated resulting in an average value of reputation that any other user can then look up. The implicit assumption in such systems is that users understand and share an underlying trust model. However, the interpretation of the scale and its use might be subjective and dependent upon an individual's internal model, predisposition and preferences. Hence, the same rating given to an entity by two different users may have different meaning. For instance, two users may have rated the same movie "Ben-Hur" on a scale from 1–5 with a 4, but the same value may mean two completely different things: user u_i may be an optimist/enthusiastic user and may rate movies in general very high by using only the highest end of the scale, while u_j may be a pessimist/conservative user who normally uses the lower end of the scale and for him/her the rating of 4 may comprise an unusual high rating indicating strong likeness. Therefore, it may

☆ Work partially supported by the Spanish Ministry of Science and Innovation through the projects OVAMAH (Grant TIN2009-13839-C03-02; co-funded by Plan E) and "AT" (Grant CSD2007-0022; CONSOLIDER-INGENIO 2010) and by the Spanish Ministry of Economy and Competitiveness through the project iHAS (Grant TIN2012-36586-C03-02); and by the Region of Madrid through the eMadrid project (Grant S2009-TIC1650).

* Corresponding author. Tel.: +44 7580795230.
E-mail addresses: rhermoso@essex.ac.uk (R. Hermoso), rcenteno@lsi.uned.es (R. Centeno), mfasli@essex.ac.uk (M. Fasli).

[1] http://www.ebay.com.
[2] http://www.amazon.com.
[3] http://www.tripadvisor.com.

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.09.025

be said that most of the proposed reputation mechanisms might suffer from the problem of misinterpretation of the different subjective opinions received and their consequent inaccurate aggregation (Dalvi, Kumar, & Pang, 2013).

Along the same line, other studies, such as David and Pinch (2005), have shown how traditional reputation mechanisms might suffer from bias problems. For example, consider two users with different internal preferences over the same three hotels. Depending on the ratings given by them, the aggregation of their opinions into a reputation ranking may vary substantially, resulting in the most preferable hotel being a different one, even given ratings that represent their preferences relationships. That is, these systems are merely passive entities that expect user opinions in order to build accurate reputation. Therefore, they do not have proactive mechanisms allowing them to extract users' preferences.

All of these issues bring about the problems of subjectivity aggregation since different individuals might report different numerical opinions regarding the same outcomes, and easy manipulation of reputation extracted from users' opinions. To alleviate these problems, the work presented in this paper makes the following contributions. Firstly, we propose a reputation mechanism based on qualitative comparison instead of traditional numerical ratings. With this in mind we intend to show how we mitigate the effect of subjectivity in opinions compared to traditional systems by focusing on extracting and holding a preference ordering instead of averaging numerical opinions. We articulate the dynamics of the mechanism with a novel iterative approach that uses *knock-out tournaments* to allow *preference comparison*, as well as an accurate algorithm to aggregate the results from the previous comparisons into qualitative rankings. We exploit the advantages of the suggested mechanism and utilise it in order to extract reputation from CSNs working in a *proactive* manner, since it is the system the one which requests preferences from the users instead of expecting ratings from them. Finally, we present an extensive set of experiments using two well known real world data sets: MovieLens-100k[4] and Flixster.[5] We demonstrate that the mechanism approximates a known ground truth with reasonable resource consumption – in terms of requested opinions – under the same assumptions used in traditional mechanisms. We also illustrate how our mechanism outperforms traditional ones when there exists some sort of subjectivity in the opinion ratings as provided by users.

The rest of the paper is organised as follows: first, in Section 2 we present a set of preliminary concepts that our paper makes use of. Then we describe in detail the mechanism used to extract reputation in Section 3. In order to test the mechanism, we present a set of experiments in Section 4. Related work is put forward in Section 5. Finally, we summarise the paper and sketch avenues for future work in Section 6.

2. Preliminaries

In this section we introduce some basic concepts that we use throughout the paper. In particular, we define the concepts of pairwise elicitation processes and knock-out tournaments.

2.1. From opinion ratings to pairwise queries

As mentioned above, most of the reputation mechanisms developed so far allow users to share their opinions about entities (other users, items, etc.) in terms of quantitative representations of trust. For instance, in trust intervals of $[-1, 1]$ an opinion of 0.9 means a high trust value, while a value of -0.95 denotes high distrust. We claim that this might lead to some major shortcomings when aggregating opinions. Even when users do not cheat when communicating their opinions, subjectivity among users can lead to misinterpretations in reputation. With this in mind, and in order to alleviate these problems, we propose to use pairwise comparisons to extract the users' preferences. The rationale behind this approach is that, as it has been shown by some works like Balakrishnan and Chopra (2010), it is easier for users to state opinions when the queries compare objects in a pairwise fashion, than evaluating two different entities, separately. Consider for example the case of TripAdvisor. A user may find it easier to answer a question such as: "do you prefer hotel A or hotel B?" as opposed to "on a scale from 1 to 5, how do you rate hotel A and hotel B?". Therefore, pairwise queries seem to be a good tool for addressing the problem of misunderstanding the level of trust of a user in an entity, as well as, its consequent inaccurate aggregation. In fact, it may allow us to coax out users' preferences and aggregate them, even when they do not share the same trust representation.

2.2. Knock-out tournaments

The concept of a tournament is very often used in a variety of situations in human societies. They are used in many different situations such as social or commercial settings, e.g., hiring processes, sports, political elections, and so on. Tournaments consist of rounds during which several matches take place, and the results of those matches determine the individuals that get through to the next round, and so on, until there is a winner. Tournaments have also been used extensively for research and scientific competitions as well as entertainment. Examples include the Trading Agent Competition,[6] the Mario AI Benchmark (Karakovskiy & Togelius, 2012), and Robocup.[7] In this paper, we focus on a specific type of tournament called *knock-out tournament*. This model is commonly used in some sports competitions and consists of a tree-like structure of rounds of matches in which the winner of a match gets through the next round to play against the winner of a "sibling" match in the competition tree. From a voting theory perspective, a knock-out tournament could be seen as a sequential elimination voting protocol with pairwise elimination (Brams & Fishburn, 2002).

We will use the notion of tournament to gather the opinions of users in the network. We adhere to the definition of *knock-out tournament* by Vu, Altman, and Shoham (2009), by which $KT_P = (T, S)$ is a knock-out tournament where T is a binary tree with P the leaf nodes – or players. Note we consider only complete trees, thus there exist d rounds involving 2^d players. Fig. 1 shows an example of a knock-out tournament with 4 players. Dashed lines represent the paths of individuals getting through to the next rounds.

An important advantage of this type of selection procedure is that while a tournament is solving, depending on the domain, some implicit relationships may be extracted. For example, let us suppose that Fig. 1 represents a tournament for selecting the most valuable objects for a set of users. In the final round we can observe that the object A wins D, so it means that D is considered a more valuable object than A. In a previous round, the object A won B, so we implicitly could extract that the object D would win B and, thus, we could state that the former is considered also more valuable than the latter, even though objects B and D have not been matched up directly.

3. A reputation mechanism based on preference elicitation

In this section, we formalise the problem introduced previously, as well as the solution proposed based on a reputation mechanism

[4] http://www.grouplens.org/node/73.
[5] http://www.cs.sfu.ca/ sja25/personal/datasets/.
[6] http://tradingagents.eecs.umich.edu.
[7] http://www.robocup.org/.