Using Centrality Measures to Predict Helpfulness-based Reputation in Trust Networks

Using Centrality Measures to Predict Helpfulness-based Reputation in Trust Networks

PASQUALE DE MEO, University of Messina
KATARZYNA MUSIAL-GABRYS, Bournemouth University
DOMENICO ROSACI, University of Reggio Calabria
GIUSEPPE M.L. SARNÈ, University of Reggio Calabria
LORA AROYO, VU Amsterdam

In collaborative Web-based platforms, user reputation scores are generally computed according to two orthogonal perspectives: *(a) helpfulness-based reputation (HBR)* scores and *(b) centrality-based reputation (CBR)* scores. In HBR approaches, the most reputable users are those ones who post the most helpful reviews according to the opinion of the members of their community. In CBR approaches, a "who-trusts-whom" network – known as *trust network* – is available and the most reputable users occupy the most central position in the trust network, according to some definition of centrality. The identification of users featuring large HBR scores is one of the most important research issue in the field of Social Networks and it is a critical success factor of many Web-based platforms like e-marketplaces, product review Web sites and question-and-answering systems. Unfortunately, user reviews/ratings are often sparse and this makes the calculation of HBR scores inaccurate. In contrast, CBR scores are relatively easy to calculate provided that the topology of the trust network is known. In this paper we investigate if CBR scores are effective to predict HBR ones and, to perform our study, we used real-life datasets extracted from CIAO and Epinions (two product review Web sites) and Wikipedia and applied five popular centrality measures – Degree Centrality, Closeness Centrality, Betweenness Centrality, PageRank and Eigenvector Centrality – to calculate CBR scores. Our analysis provides a positive answer to our research question: CBR scores allow for predicting HBR ones and Eigenvector Centrality was found to be the most important predictor. Our findings prove that we can leverage trust relationships to spot those users producing the most helpful reviews for the whole community.

CCS Concepts: •**Information systems** → **Reputation systems; Social networking sites;** •**Human-centered computing** → **Social networks; Empirical studies in collaborative and social computing;** *Social network analysis;*

ACM Reference Format:
Pasquale De Meo, Katarzyna Musial-Gabrys, Domenico Rosaci, Giuseppe M.L. Sarné, Lora Aroyo, 2016. Using Centrality Measures to Predict Helpfulness-based Reputation in Trust Networks *ACM Trans. Internet Technol.* 9, 4, Article 39 (March 2010), 20 pages.
DOI: 0000001.0000001

1. INTRODUCTION

The success and the growth of many Web-based platforms crucially depend on *trust relationships* among their members: due to the lack of trust, users may not feel comfortable in sharing their ideas or interacting with other users and, on the long run, they are likely to drop out the platform [Bonchi et al. 2011; Hogg and Adamic 2004].

Author's addresses: P. De Meo, Department of Ancient and Modern Civilizations, University of Messina, Italy. K. Musial-Gabrys, Data Science Institute, Faculty of Science and Technology, Bournemouth University, UK. D. Rosaci, Department of Information Engineering, Infrastructure and Sustainable Energy, University of Reggio Calabria, Italy. G.M.L. Sarnè, Department of Civil Engineering, Energy, Environment and Materials, University of Reggio Calabria, Italy. L. Aroyo, Web & Media group, Department of Computer Science, VU University Amsterdam, The Netherlands.

DOI: 0000001.0000001

In a broad meaning, trust quantifies to what extent an individual is willing to rely on something (or somebody) with a feeling of relative security, despite bad consequences can derive from her/his choice [Jøsang et al. 2007]. Trust is strictly related to *reputation*, which is defined as a collective measure of trustworthiness associated with an individual as seen by the members of her/his community [Jøsang et al. 2007; De Meo et al. 2015].

On the basis of user-related data, we can consider two type of reputation scores, namely *helpfulness-based reputation scores* and *centrality-based reputation scores*.

Helpfulness-based reputation (HBR in short) scores rely on the fact that some platforms (e.g., online rating systems) allow their members to rate/comment items available in the platform. Other members can, in their turn, provide a feedback specifying how that comment/score is perceived as helpful [Zhang and Varadarajan 2006; Kim et al. 2006]. In this way, the reputation of a user is calculated by aggregating the feedbacks she/he got: for instance, if feedbacks were expressed as numerical values we could take the mean (or any operator like the median or the mode [Garcin et al. 2009]) of these values to calculate user reputation.

Centrality-based reputation scores (CBR in short) are based on the fact that some platforms allow their members to explicitly declare what users they trust. Because of trust relationships are asymmetric, the web of trust relationships can be viewed as a directed graph – called *trust network* – in which each user corresponds to a vertex and edges encode trust relationships [Hogg and Adamic 2004; Zhang and Mao 2014].

Trust propagates in the trust network and the reputation score of a user depends on the amount of trust flowing through the vertex associated with her/him in the trust network. In this way, individuals occupying the most important positions in the trust network are to be regarded as the most reputable users.

Graph theory provides powerful tools – called *centrality indices* – to determine the importance of a vertex [Newman 2010; Wasserman and Faust 1994]: in the following, we will understand the CBR score of an individual in the trust network as the centrality of the vertex associated with that individual.

Despite a number of researchers concentrated on designing and validating algorithms to calculate CBR and HBR scores, very little is known about the relationship between HBR and CBR scores. This paper aims at filling this gap and the following question guides our research: *Do CBR scores enable to predict HBR ones?*

To answer our question, we used four publicly available datasets: a dataset extracted from CIAO (a product review Web site) [Tang et al. 2013] and two datasets (called Epinions I and Epinions II) of different size extracted from Epinions [Tang et al. 2013; Richardson et al. 2003], another product review Web sites. The last dataset (called WIKI) is associated with administrator elections in Wikipedia [Leskovec et al. 2010]. For all datasets, we have access to the existing trust networks; in case of CIAO and Epinions I, we have also user ratings and helpfulness scores.

We used five, well-known, indices to compute centrality: the Degree Centrality [Wasserman and Faust 1994], Closeness Centrality [Bavelas 1948], Betwenness Centrality [Friedman 2001], Eigenvector Centrality [Bonacich and Lloyd 2001] and PageRank [Brin and Page 1998].

The main findings of our research can be summarized as follows:

(1) The organization of trust networks is far from random: we observed that trust links ensure path lengths were small; in addition, a large weakly connected component – often containing more than 90% of vertices – surfaced in all trust networks.
(2) Trust networks contain a relatively large number of triangles, which is an indicator of trust transitivity. In addition, CIAO, Epinions I and Epinions II display a large number of reciprocated trust links. A high level of reciprocity promote positive and

genuine interactions among platform members and motivate them to not leave the platform [Chen et al. 2009; Butler et al. 2007].

(3) We used *Gradient Boosting Regression* [Friedman 2001] to study the dependence of HBR scores from CBR ones. We used three different operators to aggregate helpfulness scores, namely the average, the median and the mode [Garcin et al. 2009]. We applied a ten-fold cross validation procedure to compute how much predicted HBR scores differ from actual ones and we used the Mean Square Error (MSE) to quantify this difference. We found that MSE was equal to 0.22 in case of CIAO, and 0.13 in case of Epinions I: therefore, CBR scores allow for predicting HBR ones and this aids in spotting users capable of producing helpful reviews even if they rated just few items.

(4) Eigenvector Centrality was always found to be the most influential variable to predict HBR scores. Therefore, we recommend the usage of Eigenvector Centrality to design software applications spotting highly reputable users.

Our study has deep practical implications:the ability of detecting users with large HBR scores is a fundamental ingredient to ensure the long-term success of e-marketplaces [Bonchi et al. 2011; Melnik and Alm 2002], to detect fraudulent behaviours [Neville et al. 2005] and to handle Customer Relationship Management (CRM) activities [Bonchi et al. 2011]. The paper is structured as follows: in Section 2 we provide some background material while in Section 3 we discuss related literature. Section 4 is devoted to illustrate the datasets used in this study and Section 5 investigates the structural properties of trust networks extracted from these datasets. Section 6 illustrates the relationship between HBR and CBR scores. In Section 7 we draw our conclusions and illustrate future works.

2. BACKGROUND AND BASIC CONCEPTS

In this section we introduce some basic concepts that will be extensively used throughout the paper. In particular, in Section 2.1 we focus on users interactions in Web-based platforms and how to model these interactions through graphs. Subsequently, in Section 2.2, we provide some basic concepts about graphs and centrality measures in graphs. We introduce centrality-based (CBR) and helpfulness-based reputation (HBR) scores in Section 2.3.

2.1. Modelling User Feedbacks in Web-based platforms

Most of the interactions in Web-based platforms involve many type of (often heterogeneous) entities like users, buyers, sellers, public/private institutions that usually have no prior relationships with each other.

Relationships among entities are built on the basis of user feedbacks, which can be classified into *implicit* (if we build them by tracking user behaviours and interactions) and *explicit* (if they are directly created by the users) [Musial and Kazienko 2013; De Meo et al. 2011].

Examples of implicit interactions are the frequency of private communications or the number of times a user comments/likes the posts or photos of another user [Bakshy et al. 2012].

The task of collecting implicit interactions is relatively easy and fully automatic [Mislove et al. 2007; Zhu 2010]. In the context of trust, however, implicit signals may introduce confounds because it is hard to discern a mere benevolent act from an explicit trust declaration. Explicit trust declarations may depend on the specific features of the platform: for instance, in online technological forums like StackOverflow, a user can award reputation points to the users who provide the most comprehensive answers [Hendrikx et al. 2015; Yang et al. 2014]. In product review Web sites (e.g., CIAO or

Epinions), a user can declare if she/he trusts another user in recognition of her/his activity as product reviewer [Tang et al. 2013].

Both implicit and explicit feedbacks can be modelled as graphs. In the next section we provide some basic terminology about graphs and we will introduce the concept of centrality in a graph.

2.2. Centrality Indices in Graphs

Implicit and explicit user feedbacks can be modelled as a graph $G = \langle V, E \rangle$, often called *social graph*. Here, V is the set of vertices and each vertex $v \in V$ represents a member of a Web-based platform. The set $E \subseteq V \times V$ is the set of edges; an edge $\langle u, v \rangle \in E$ specifies a kind of relationship like collaboration, friendship, trust and so on between individuals associated with vertices u and v respectively. In the following we will assume that edges encode trust relationships. Because trust relationships are often non-mutual, the graph G will be directed, i.e., if $\langle u, v \rangle \in E$ we have, in general, that $\langle v, u \rangle \notin E$. The centrality of an individual in G is a score assessing the importance of that individual. A large number of centrality indices are now available [Wasserman and Faust 1994]. In our study we adopted the following centrality indices:

— *Degree Centrality* - $DC(v)$ [Wasserman and Faust 1994]:

$$DC(v) = |\{w \in V : \langle v, w \rangle \in E\}| \tag{1}$$

— *Closeness Centrality* - $CC(v)$ [Bavelas 1948]:

$$CC(v) = \frac{1}{\sum_{u \in V} \mathtt{SP}(u, v)} \tag{2}$$

Here $\mathtt{SP}(u, v)$ is the length of the shortest path connecting vertices u and v.

— *Betweenness Centrality* - $BC(v)$[Freeman 1977]:

$$BC(v) = \sum_{v \neq u \neq w \in V} \frac{\sigma_{uw}(v)}{\sigma_{uw}} \tag{3}$$

Here, given any three distinct vertices v, u and w in V, $\sigma_{uw}(v)$ is the number of shortest paths from u to w and $\sigma_{uw}(v)$ be the number of the shortest paths from u to w passing through v.

— *Eigenvector Centrality* - $EC(v)$ [Wasserman and Faust 1994; Bonacich and Lloyd 2001]. $EC(v)$ is the v-th component of the vector $\vec{\mathbf{x}}$ defined as the solution of the following equation:

$$\mathbf{A}\vec{\mathbf{x}} = \lambda\vec{\mathbf{x}} \tag{4}$$

Here $\mathbf{A}$ is the adjacency matrix of G, i.e., $\mathbf{A}_{uv} = 1$ if and only if there is an edge from u to v, 0 otherwise. If we impose that $EC(v) \geq 0 \quad \forall v \in V$, the Perron-Frobenius theorem ensures that Equation 4 admits a unique solution $\vec{\mathbf{x}}^{\star}$ which is the leading eigenvector of $\mathbf{A}$ [Berman and Plemmons 1994].

— *PageRank* - $PR(v)$[Brin and Page 1998; Bonacich and Lloyd 2001]:

$$PR(v) = \frac{1-d}{|V|} + d \sum_{w \in N^{-}(v)} \frac{PR_w}{|N^{-}(w)|} \tag{5}$$

Here d a constant (generally set equal to 0.85) called *damping factor*.

2.3. Centrality-Based and Helpfulness-Based reputation scores

Centrality indices introduced in Section 2.2 can be applied to evaluate the reputation of a user. Specifically, if we consider a trust network and a centrality index c, the CBR score $\mathcal{C}_c(u)$ of u is the centrality score –calculated by means of c – of the vertex corresponding to u in the trust network. For instance, if we choose the PageRank introduced in Equation 5, we get a CBR score $\mathcal{C}_{PR}(u)$ for each member u of the trust network.

A further notion of reputation depends on the helpfulness of the ratings posted by u and the reputation of a user depends on how her/his ratings are perceived as useful from the members of her/his community. To this end, suppose that u can rate an item i and and let r_{ui} be such a rating. Suppose that any user $v \neq u$ can review r_{ui} by providing a *review score* $h_{vi}(u)$; let $\mathcal{I}(u) = \{h_{vi}(u) : v \neq u, v \in V\}$ be the set of review scores got by u. The *helpfulness-based reputation score* $\mathcal{H}(u)$ of u is defined as as the aggregate of the reviews that u got from her/his peers:

$$\mathcal{H}(u) = \mathrm{f}\left(\mathcal{I}(u)\right) \tag{6}$$

Here f is a suitable aggregating function. Various options for f are possible but the most common choices consist of taking the average, the median or the mode of $\mathcal{I}(u)$ [Garcin et al. 2009].

3. RELATED LITERATURE

In this section we first illustrate how centrality indices have been applied to compute reputation scores (see Section 3.1). We then discuss approaches to predicting review quality (see Section 3.2). Finally, in Section 3.3, we introduce two fundamental problems in the field of Social Networks, namely the link prediction and attribute inference problem and explain how they are related to the task of predicting HBR scores.

3.1. Computing Reputation Scores by means of Centrality Indices

Early approaches to computing reputation were based on the referrals/ratings an individual got from other members of her/his community [Jøsang et al. 2007] but reputation scores can be easily manipulated if groups of users/friends collude by maliciously providing positive feedbacks on a third party [Hogg and Adamic 2004].

Several authors suggested to integrate the referrals a user received with her/his social relationships to get a more robust evaluation of reputation scores [Pujol et al. 2002; Kamvar et al. 2003; Hogg and Adamic 2004; Hendrikx et al. 2015].

An interesting approach is *Eigentrust* [Kamvar et al. 2003], which targets at measuring peer reputation in P2P networks. Eigentrust assumes that trust scores are transitive and the trust score $ts(i, j)$ between a pair of peers p_i and p_j is equal to the number of times p_i downloaded an authentic file from p_j minus the number of times p_i downloaded a fake file from p_j. Trust scores are normalized and recorded in a trust matrix $\mathbf{T}$. The reputation of the peer p_i is the i-th component of the leading eigenvector $\vec{e}$ of $\mathbf{T}$.

[Chirita et al. 2004] also studied how to compute reputation in P2P networks: each peer p_i selects a set $\mathcal{S}_i$ of trusted peers whose trust ratings and past download histories are the most relevant to access the resources p_i needs. A personalized (and fully distributed) implementation of the PageRank algorithm is applied to compute the reputation of peers not belonging to $\mathcal{S}_i$.

Other authors assume that trust flows from few trustworthy users [Levien 2009; Golbeck and Hendler 2006; Sirivianos et al. 2014]. One of the early approaches to mention is the so-called *Advogato metrics* [Levien 2009], which solves a maximum flow problem to compute reputation scores. Such a procedure, however, may not scale well on large graphs and to this purpose [Sirivianos et al. 2014] provided an efficient yet accurate heuristic to approximate the optimal flow value. [Golbeck and Hendler 2006] apply a modified version of the Breadth First Search algorithm to infer multiple values of reputation for each user; these values are then aggregated by applying a voting algorithm to produce a unique value of reputation for each user.

Compared to approaches cited above, our study introduces several novelties. Unlike [Kamvar et al. 2003] and [Chirita et al. 2004], we consider communities of human users rather than P2P networks and, as a consequence, the creation of trust links between members of the same community relies on sociological and psychological factors. Furthermore, we advance the state-of-the-art by proving that we can supply the shortage of user contributed ratings/reviews to calculate HBR scores by means of CBR scores.

3.2. Predicting Review Quality

A large number of researchers studied the interplay of user ratings and trust scores in Web-based platforms [Ma et al. 2009; Golbeck 2006]. Most of the existing approaches are concerned with the *item rating problem*, i.e., they attempts at predicting the rating r_{ui} a user u is likely to confer to an unseen item i. Other community members can score r_{ui} or they can provide a short text explaining why they consider r_{ui} useful (or not). In case of text comments, however, the quality of provided review greatly varies: in fact, we can run into very detailed opinions but we can also encounter reviews which are the simple repetition of product specifications or spam content.

Due to such a variability, many researchers were concerned with the problem of automatically estimating the quality of reviews [Zhang and Varadarajan 2006; Kim et al. 2006; Liu et al. 2008]. Indeed, [Zhang and Varadarajan 2006] considered the number of superlative/comparative adjectives/adverbs appearing in a review and applied Support Vector Regression to predict review helpfulness. [Kim et al. 2006] also applied Support Vector Regression but they considered features like review length, unigrams and product ratings. [Liu et al. 2008] described a model to compute review quality in the movie domain. Here, three distinctive features influenced review quality, namely: *(a)* review expertise, *(b)* the writing style and *(c)* review timeliness. [Liu et al. 2008] applied Radial Basis Functions to model review expertise as well as writing style.

The approaches cited above focus only on textual features of reviews; other approaches suggest to include social information to predict review quality[O'Mahony and Smyth 2009; Lu et al. 2010]. [O'Mahony and Smyth 2009] calculated review quality in TripAdvisor by building a user-hotel bipartite graph and investigating its structural features. [Lu et al. 2010] suggested to explore the social network of a reviewer to get some insights about her/his expertise. Such an approach encodes the assumption that a user u_1 will admit a user u_2 in her/his social network if and only if the quality of reviews posted by u_2 is at least as high as those posted by u_1. High quality reviewers are those users expected to post high quality reviews.

Our approach substantially differs from the ones presented in this section: we rely on *explicit trust links* to compute user reputation; in contrast, the approach of [Lu et al. 2010] assumes that social ties are an *implicit* indication of trust and use such an information to infer reviewer expertise. As a further difference, all the approaches discussed in this section are concerned with the prediction of the quality of a review. In contrast, we wish to calculate the reputation of a user understood as her/his ability of posting helpful reviews. Finally, we aim at studying if such a definition of reputation is related to the centrality of an individual in a trust network.

3.3. Link Prediction and Attribute Inference

Our research is also related to two well known problems in Social Network Analysis literature, namely the *link prediction problem* and the *attribute inference* problem[Kumar et al. 2004; Fond and Neville 2010; Backstrom and Leskovec 2011; Yin et al. 2010; Gong et al. 2014].

Given a social network described, at the time instant t, by a social graph $G(t) = \langle V(t), E(t) \rangle$, the link prediction problem consists of guessing the structure of the social graph $G(t') = \langle V(t'), E(t') \rangle$ at the time instant $t' > t$.
The attribute inference problem can be defined as follows: given a vertex $v \in V(t)$, determine the value of some of its attributes (like gender, age or job).
Empirical studies show that users with similar attributes are likely to be connected [Kumar et al. 2004]; analogously, the social influence principle states that users who are connected through a social link are likely to share common attributes [Fond and Neville 2010]. Therefore, some authors suggested to make use of vertex attributes to predict link formation or, vice versa, to exploit graph topology to infer vertex attributes [Backstrom and Leskovec 2011].
[Yin et al. 2010] added extra-vertices to the social graph G to represent attributes, thus obtaining an augmented graph called *Social Attribute Network - SAN*; [Yin et al. 2010] applied Random Walks with Restart on the SAN to jointly predict social links and infer attributes. The SAN model was subsequently extended by [Gong et al. 2014] to include *negative attributes* (i.e., a vertex does not have that attribute) and *mutually exclusive attributes*. Link prediction and attribute inference problems resemble, to some extent, the research problem considered in our work: in fact, like the approaches cited above, we attempt at investigating the relationship between the topological structure of a trust network and the ability of each of its members to post helpful reviews.
Some important differences between our work and those cited in this section deserve to be highlighted: firstly, we focus on trust relationships which semantically differ from friendship relationships considered in the papers cited above. As a further difference, we observe that attributes associated with vertices reflect special personal traits or episodes of the life of an individual and, therefore, they exist independently of the presence of an individual in an online platform; by contrast, the HBR scores are strictly related to the behaviour of an individual in a specific platform and, then, the HBR score of an individual could vary if that individual switches from a platform to another one.

4. DESCRIPTION OF DATASETS

To perform our analysis, we used the following datasets:

— *CIAO*. CIAO is an online-shopping portal which provides its members a forum to review products. Reviews should help consumers to make decisions and they can be rated by other CIAO users on the basis of their *helpfulness* on a scale ranging from exceptional to off-topic.
— *Epinions*. Epinions was a consumer review website launched in 1999 and subsequently bought by eBay. It had a large catalogue of goods/services that consumers could review. All submitted reviews were published and they can be rated by other users. In our tests we considered two Epinions snapshots of different size, called *Epinions I* and *Epinions II* (see below for more details).
— *WIKI*. This dataset refers to the result of elections for Wikipedia administrators. A Wikipedia user can issue a public request for being promoted to the role of administrator; a public discussion or an election identify the users who are entitled to become administrators.

In all the considered platforms, the identification of the most reputable users has a clear implication on the development of the platform itself.

In case of CIAO and Epinions the reputation of a user depends on her/his ability of posting useful product reviews and, to this extent, reputable users contribute to the commercial success of the platform.

In case of Wikipedia, the most reputable users are Wikipedia administrators granted for blocking other users to prevent abuses or digital vandalism as well as to check the quality and completeness of Wikipedia entries. Administrators play, therefore, a fundamental role to the growth and quality of Wikipedia articles.

CIAO and Epinions I datasets have been collected in May 2011 [Tang et al. 2013][1] while Epinions II dataset has been collected in 2003 [Richardson et al. 2003][2]. CIAO and Epinions I have the same structure: in each dataset a tuple consists of a user ID, a product ID, a category ID (specifying the commercial category of that product), the rating assigned by the user to that product and, finally, the helpfulness of that review (defined as the average of all scores a review received). For each dataset, a list of trust relationships is available: each element of the list is a pair of user IDs such that the first user of the pair declared to trust the second one.

As for Epinions II, we managed only the *who-trusts-whom* network. As for WIKI, we used the data collected by [Leskovec et al. 2010]. WIKI was extracted from a Wikipedia dump (from January 2008) containing the results of $2\,794$ elections with total of $103\,663$ votes[3]. The active and passive electorate consists of $8\,298$ users. The WIKI dataset gives us a who-vote-whom network; only the trust network is available because users are not involved in any rating activity.

5. STRUCTURAL PROPERTIES OF CIAO, EPINIONS AND WIKI

Here we study the structural properties of trust networks associated with CIAO, Epinions I, Epinions II and WIKI datasets. Each trust network is represented as a directed graph $G = \langle V, E\rangle$ such that each vertex identifies a user and each edge models a trust relationship between the vertices it connects. In the following, we both describe the main properties of our trust networks and report the main practical implications they entail.

FINDING 1: *The organization of trust networks is far from random: almost all pairs of users are connected through paths which, on average, consist of few edges and there are few isolated users.*

From Table I we observe that Epinions I trust network contains $22\,167$ vertices (i.e., about 3 times the number of vertices in CIAO and 2.67 times the number of vertices in WIKI) and $355\,754$ edges (i.e., 3.18 times the number of edges in CIAO and 3.43 times the number of edges in WIKI). Epinions II is the biggest trust network with $75\,888$ vertices and $508\,837$ edges. From these data we can compute the *density* δ of each trust network, defined as the ratio of the actual number of trust links to the number of possible trust links, i.e.,$\delta = \frac{|E|}{|V|\times(|V|-1)}$. Density range is in the interval $[0, 1]$ and the smaller δ, the sparser G. CIAO trust network is the densest network and its density is about two orders of magnitude higher than that of Epinions II. However, δ is generally small for all networks under inquiry.

Two further structural features to consider are the *average path length* APL and the size of the *largest weakly connected component* WCC. In detail, the APL is defined as

$$\texttt{APL} = \frac{1}{|V| \times (|V| - 1)} \sum_{u \in V} \sum_{v \in V} \texttt{SP}(u, v)$$

[1] CIAO and Epinions I datasets are available at http://www.public.asu.edu/~jtang20/datasetcode/truststudy.htm

[2] Epinions II dataset is available at http://konect.uni-koblenz.de/networks/epinions

[3] WIKI dataset is available at https://snap.stanford.edu/data/wiki-Vote.html

Table I. Some statistics about the CIAO, Epinions I and II, and WIKI trust networks.

Parameter	Symbol	CIAO	Epinions I	Epinions II	WIKI
Number of vertices	$\|V\|$	7 376	22 167	75 888	8 298
Number of edges	$\|E\|$	111 781	355 754	508 837	103 689
Density	δ	0.002	0.0007	0.00008	0.0015
Average Path Length	APL	3.31	3.26	4.75	3.25
Average Clustering Coefficient	ACC	0.118	0.113	0.66	0.126
Size of Weakly Connected Component	WCC	7 305	17 990	75 877	7 066

Here, $\mathtt{SP}(u, v)$ is the distance between two vertices u and v and it is defined as the length of the shortest path connecting u and v. From Table I we notice that APL achieves its largest value (4.75) in case of Epinions II. For all other networks, the APL is around 3.

The *weakly connected component* of G is the largest connected subgraph G' of G. In all trust networks, a weakly connected component emerges and it consists of 99.03%, 81.11%, 99.98% and 85.15% of the whole user population in CIAO, Epinions I, Epinions II and WIKI respectively.

FINDING 2: *Trust networks contain a relatively large number of triangles, which is an indicator of trust transitivity.*

We studied if trust displays a *transitive* behaviour: given three users u, v and w such that u trusts v and v trusts w, we ask if u will trust w too. If this happens, we say that the vertices associated with u, v and w in the trust network will form a *triangle*. We then computed as *average clustering coefficient* ACC [Wasserman and Faust 1994], i.e., the ratio of the number of triangles in G to the total number of connected triplets of G (here, a connected triplet is a group of three vertices such that there exist at least two edges among the vertices).

Large values of ACC imply a high degree of transitivity and, from Table I, we observe that ACC peaks around 0.126: such a value is about 14 times bigger than the ACC measured on the Web graph and comparable to that of Facebook and YouTube [Mislove et al. 2007]. The measured ACC is unusually high if compared, for instance, with randomly generated graphs [Mislove et al. 2007] and suggests that network vertices tend to aggregate into dense clusters. Such an aggregation depends on mechanism of trust formation: the probability that a user u will trust a user v increases if there is a third user w who is trusted by u and, in her/his turn, trusts v.

FINDING 3: *The distribution of in-degree and out-degree in trust networks is uneven: few users intercept most of the trust links and few users yield most of the trust links. CBR scores, with the exception of CC, are also unevenly distributed.*

We computed the probability $P(k)$ that a vertex has at least k incoming/outgoing edges; we used the methodology described in [Clauset et al. 2009] and the toolbox illustrated in [Alstott et al. 2014] to check whether $P(k)$ is shaped as a power law, i.e., $P(k) \simeq k^{-\alpha}$, as in many real-world networks.

We found that a power law emerged for in-degree distribution in WIKI ($\alpha = 3.6$, p-value = 0.08) and for the out-degree in Epinions I ($\alpha = 3.44$, p-value = 0.003) and Epinions II ($\alpha = 2.8$, p-value = 0.005). In all other cases, a power law distribution was not appropriate to describe in-degree/out-degree distribution.

We also studied the distribution of CBR scores and we found that all these distributions (with the exception with that associated with CC) quickly fall apart because the vast majority of vertices displayed low CBR scores, independently of the network to analyse.

FINDING 4: *CIAO, Epinions I and Epinions II display a large number of reciprocated trust links.*

Our previous findings motivated a further question: do the vertices displaying a high in-degree also have a large out-degree? In graph theory, we define the *reciprocity* ρ as the proportion of mutual edges in a directed graph; large values of reciprocity have been empirically observed in Social Web platforms like Twitter because users frequently follow back their followers [Weng et al. 2010]. Reciprocity plays a role to boost the development of positive interactions among members of human societies [McAlexander et al. 2002] as well as to provide them with utilitarian benefits [Chen et al. 2009; Butler et al. 2007]. We found $\rho = 0.47, 0.38$ and 0.4 in CIAO, Epinions I and Epinions II; as for WIKI we observed that $\rho = 0.06$.

Such a result depends on the semantics of edges in each network. In particular, in WIKI, an edge encodes a vote a user casted to another one to promote her/him to the role of administrator but, in general we do not expect that this relationship will be reciprocated. In contrast, in CIAO and Epinions, each user accesses the reviews posted by other users and, if she/he finds these reviews useful, she/he can create trust connections to the users who issued the reviews. This enables a mutual reinforcement mechanism in which users trust each other; on the long run, a trust network will display relatively large values of reciprocity. Our finding suggest a practical recommendation for the owners of a Social Web platform: a Social Web platform should incorporate features to promote pairwise user interactions and establish mutual trust relationships. These relationship would motivate users to stay and be active in the platform.

FINDING 5: *Reviews posted by users in reaction to available ratings are sparse and, therefore, it is worth studying if CBR scores are useful to predict HBR ones.*

We analysed the distribution of reviews in case of CIAO and Epinions I datasets. For each user we applied Kernel Density Estimation [Bishop 2006] to calculate the probability $P(x)$ that a user gets exactly x reviews.

From Figures 1(a)-1(b) we have that the helpfulness scores are quite sparse in both CIAO and Epinions; therefore, a rudimentary aggregation of these scores would be uninformative, independently of the aggregation operator we decide to apply.

In the light of this result, we could get a great benefit if we could predict HBR scores as function of CBR ones. We will deal with this question in Section 6.

6. PREDICTING HELPFULNESS-BASED REPUTATION SCORES FROM CENTRALITY-BASED ONES

In this section we target at studying whether CBR scores are related to HBR ones. As we have at our disposal a pool X of centrality indices $\mathcal{C}_c(u)$ with $c \in X$, it is convenient to model the HBR score $\mathcal{H}(u)$ as function of $\mathcal{C}_1(u), \ldots, \mathcal{C}_{|X|}(u)$:

$$\mathcal{H}(u) = \mathrm{f}(\mathcal{C}_1(u), \ldots, \mathcal{C}_{|X|}(u)) \quad (7)$$

The simplest technique to learn f is *regression* [Bishop 2006]. We first investigated the correlation among CBR scores (see Section 6.1) and we found a high level of *collinearity*, i.e., predictors were related by an almost linear relationship. Collinearity does not affect the predictive power of a regression method but, because some predic-

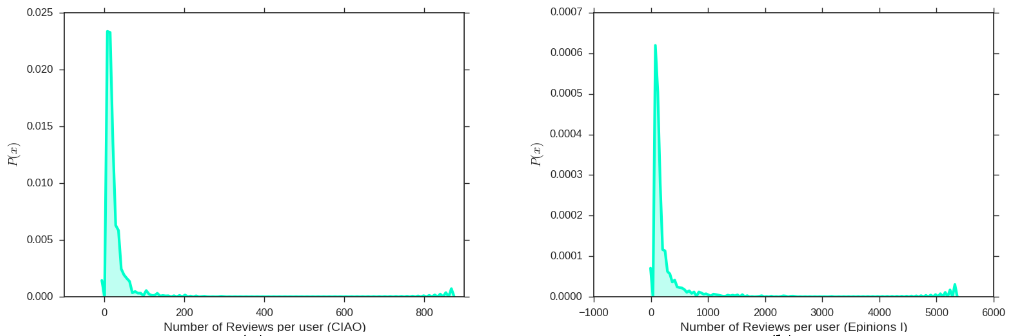


Fig. 1. **Left panel**: Probability $P(x)$ that a user has posted x reviews in CIAO. **Right panel**: Probability $P(x)$ that a user has posted x reviews in Epinions I.

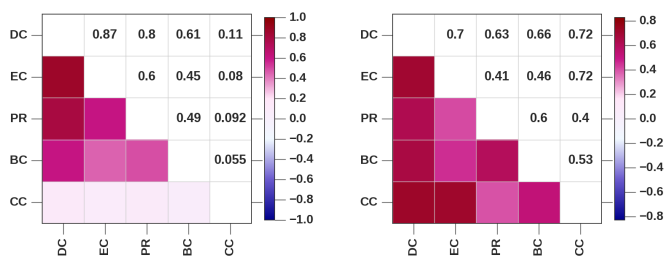


Fig. 2. **Left panel**: PCCs between CBR scores. **Middle Panel panel**: KCCs between CBR scores. **Right panel**: SCCs between CBR scores. We calculated correlation coefficients on the CIAO dataset. (Best viewed in color).

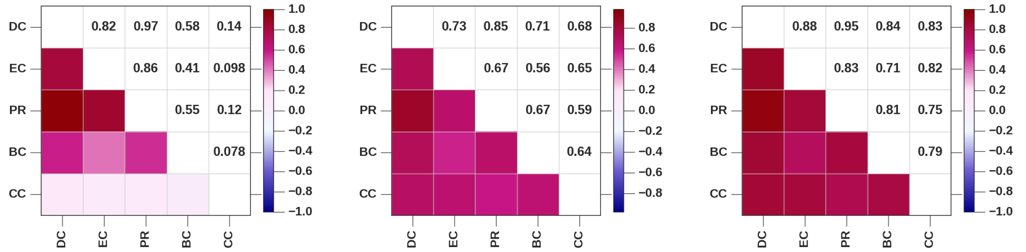
tors are redundant, it prevents us from clearly understanding the individual contribution of each CBR score in predicting $\mathcal{H}(u)$. Collinearity is not the only problem we need to deal with: to reduce the impact of *outliers*, we should opt for *robust regression methods* which include a regularization term to minimize the effect of outliers on $\mathcal{H}(u)$. In the next sections we discuss how we cope with collinearity and outliers.

6.1. Correlation between CBR scores

We applied the Pearson's Correlation Coefficient (PCC), the Kendall's tau Correlation Coefficients (KCC) and the Spearman's Correlation Coefficient (SCC) to calculate the correlation between CBR scores on CIAO, Epinions I, Epinions II and WIKI datasets.

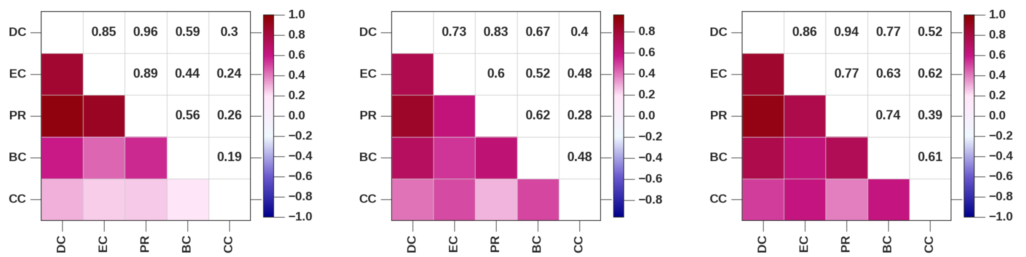
To check the statistical significance of our experiments we performed a permutation test [Good 2006] which showed that the obtained results were always statistically significant ($p < 0.001$). The results of our experiment are graphically reported in Figures 2-5.

The largest values of correlation coefficients were observed when we applied the Spearman's Correlation Coefficient; Spearman's Correlation Coefficient is, in fact, able to capture monotonic relationships between two continuous random variables and, then, generally high values of Spearman's Correlation Coefficient indicate an alignment among the rankings associated with different CBR scores. The Kendall's tau,



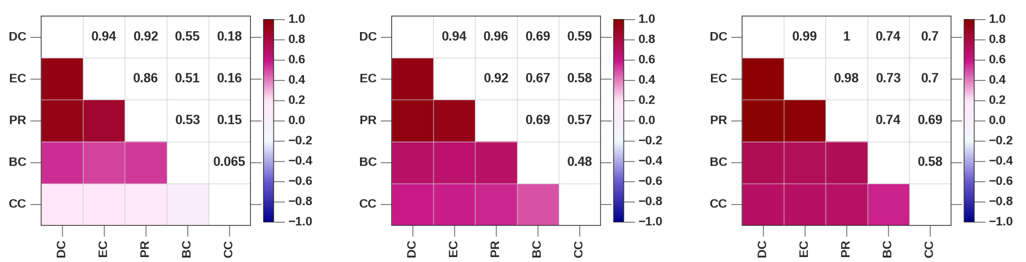
(a) (b) (c)

Fig. 3. **Left panel**: PCCs between CBR scores. **Middle Panel panel**: KCCs between CBR scores. **Right panel**: SCCs between CBR scores. We calculated correlation coefficients on the Epinions I dataset. (Best viewed in color).



(a) (b) (c)

Fig. 4. **Left panel**: PCCs between CBR scores. **Middle Panel panel**: KCCs between CBR scores. **Right panel**: SCCs between CBR scores. We calculated correlation coefficients on the Epinions II dataset. (Best viewed in color).



(a) (b) (c)

Fig. 5. **Left panel**: PCCs between CBR scores. **Middle Panel panel**: KCCs between CBR scores. **Right panel**: SCCs between CBR scores. We calculated correlation coefficients on the WIKI dataset. (Best viewed in color).

as the Spearman's Correlation Coefficient, takes the rankings produced by two different methods to calculate correlation but it is based on the number of concordant and discordant pairs in the two rankings. We notice that correlation coefficients between DC and EC (resp., DC and PR) are always larger than 0.79 (resp., 0.63), independently of the dataset we consider and the method applied to measure correlation. In many cases, the correlation between DC and EC (resp., DC and PR) was bigger than 0.9. An important exception displays if we calculate Pearson Correlation Coefficient between PR and CC scores as well as between BC and CC scores; after inspecting the distribution of CC scores we concluded that low values of PCCs depend on the fact that the Pearson Correlation Coefficient is more sensitive to noise in data than other methods. The outcome of our experiment confirms the findings of [Ronqui and Travieso 2015], despite they studied correlation between centrality scores in undirected networks. Our experiment reveals that CBR scores are consistent but they are also strongly related and, therefore, we need to cope with collinearity.

6.2. Gradient Boosting Regression

To deal with collinearity and outliers, we count on *robust regression methods*; these methods have been extensively studied in the latest years and, after some experiments, we found that *Gradient Boosting Regression - GBR* [Friedman 2001] was particularly suited to answer our research questions.

We opted for GBR for the following reasons: *(i)* GBR handles heterogeneous data in which features are measured according to different scales and this is particularly relevant in our application scenario because CBR scores are measured according to different scales; *(ii)* GBR does not separate the model fitting step with variable selection step, thus making the model it generates easy to interpret; *(iii)* GBR efficiently handles collinearity and, in particular, it allows to quantify the importance of each feature in predicting the target variable. In our application domain this means that we are able to find out what CBR score contributes the most to correctly predict $\mathcal{H}(u)$.

GBR is part of a more general family of techniques to perform classification and regressions tasks known as *boosting* [Elith et al. 2008]. Boosting creates a collection of trees to predict the values of HBR scores, the goal of boosting is to minimize a function – called *deviance* – which measures the distance between the true HBR score of a user and the predicted one. Various definitions of deviance are possible and, among them, we cite the *least-square loss function* (which uses the L_2 norm) and the *Huber Loss Function* [Friedman 2001] (which relies on the L_1 norm). After performing some initial experiments, we opted for the Huber Loss Function because it was less prone to overfitting.

GBR follows an iterative approach: in the first step, it creates only one regression tree that maximally reduces the deviance and it is grown by selecting one of the features – say x – and splitting it; at the second step, a tree is fitted with the goal of further reducing deviation, and that second tree could differ from the first one as for features selected for splitting and for split points. Existing trees are left unchanged as the model is enlarged. The final model is a linear combination of M trees, being M (*boosting iteration*) a fixed parameter controlling the complexity of the model.

A byproduct of the GBR algorithm is the ability of estimating the relative influence – called *relative importance* – of each feature in optimizing the deviance. Intuitively, the relative importance $i(x)$ of a feature x is computed as follows [Elith et al. 2008; Friedman 2001]: for each tree generated by the GBR algorithm, consider the number of times $t(x)$ in which x has been selected for splitting and we weight $t(x)$ by the squared reduction of deviance $\Delta(D)$ it generates. The relative importance of x is then defined as the average of the terms $t(x) \times \Delta(D)$ across all trees. The relative importance has been normalized so that the relative importance of the most important predictor equals 100 (with high numbers of $i(x)$ indicating stronger influence on the response).

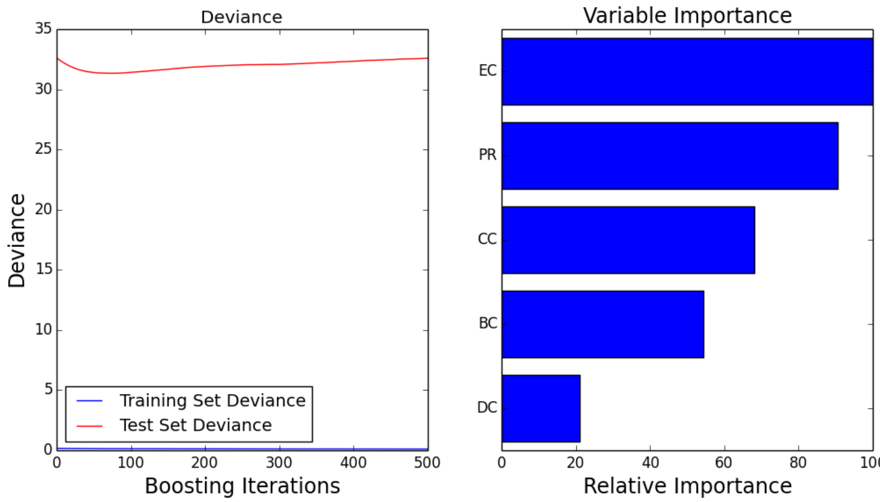


Fig. 6. **Left Panel**: Test and training set error vs. Boosting Iteration. **Right Panel**: Relative Feature Importance (CIAO Dataset). We used the avg operator to aggregate review scores.

6.3. Experimental Findings

To run our experiments we assumed that M varied up to 500. We used three different operators to aggregate helpfulness scores, namely the `avg`, `median` and `mode`.

We applied a ten-fold cross-validation procedure to measure the deviance associated with the test set error. In Figures 6 and 7, (left side), we reported the deviance in the training/test set as function of M when the `avg` operator is applied.

In the right side of Figures 6 and 7 we reported the relative importance of CBR scores. We repeated the experiment above by applying the `median` (see Figures 8 and 9) and the `mode` operator (see Figures 10 and 11), respectively. From our experiments we learned the following lessons:

(1) We measured the Mean Square Error (MSE) on the test set and we found that `avg` achieved the lowest MSE on both CIAO (0.22) and Epinions (0.13). As such, we conclude that CBR scores are actually able to predict HBR ones.
(2) On CIAO, the test set error associated with `avg` is roughly similar to that of `median` and `mode` aggregation operators. In all cases, the optimal number of trees $M^{\star}$ equals 100: if $M > M^{\star}$, in fact, the test error slightly increases due to overfitting; in addition, large values of M imply an increase in the computational costs of GBR.
(3) On Epinions I, the test set error of `avg` is twice lower than that of `median` which, in its turn, is 33% lower than the error in which the `mode` operator incurs. In addition, the test error when `median` and `mode` operator are applied is stable if M varies; if the `mode` operator is considered, the test error achieves its lowest value as $M = 80$ (and it remains stable if $M > 80$).
(4) In both CIAO and Epinions I, EC stands out as the most influential variable to predict HBR scores independently of the operator we used to aggregate helpfulness scores.

These results can be interpreted as follows: CIAO/Epinions members who find the reviews helpful are likely to create trust links to the member who posted them. As

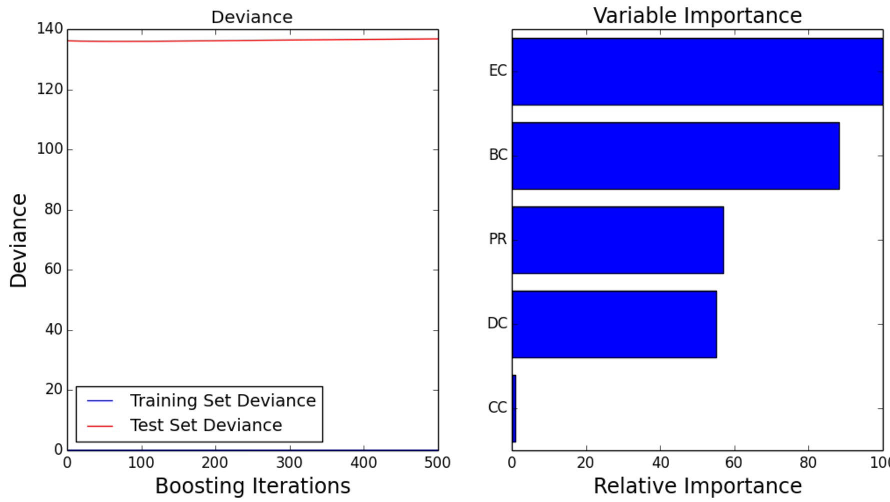


Fig. 7. **Left Panel**: Test and training set error vs. Boosting Iteration. **Right Panel**: Relative Feature Importance (Epinions I Dataset). We used the `avg` operator to aggregate review scores.

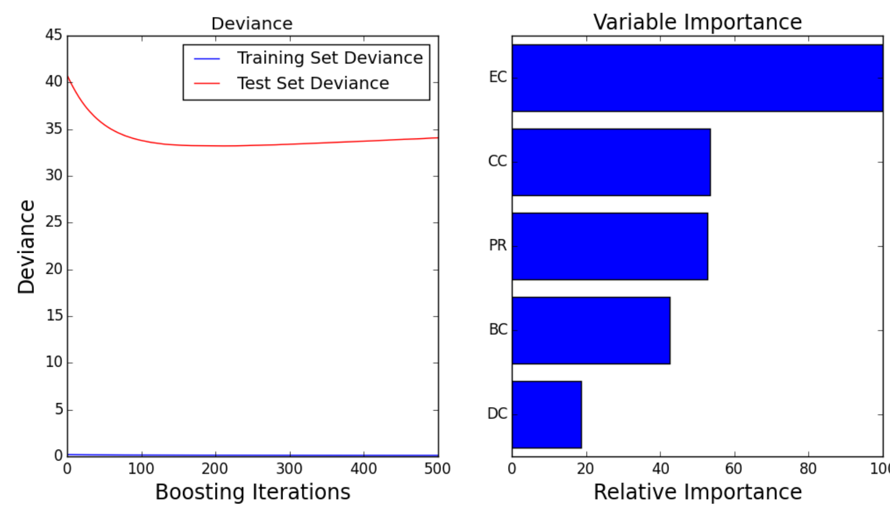


Fig. 8. **Left Panel**: Test and training set error vs. Boosting Iteration. **Right Panel**: Relative Feature Importance (CIAO Dataset). We used the `median` operator to aggregate review scores.

such, a user v with a high HBR score should thus occupy central position in the trust network.

7. CONCLUSIONS AND FUTURE WORKS

7.1. Summary of Results and Discussion

In this paper we considered two methods to compute user reputation scores in Web-based collaborative platforms. The first method, called centrality-based reputation (CBR) scores, relies on the knowledge of the web of trust connections (also known as *trust network*) between platform members and quantifies the reputation of a user as

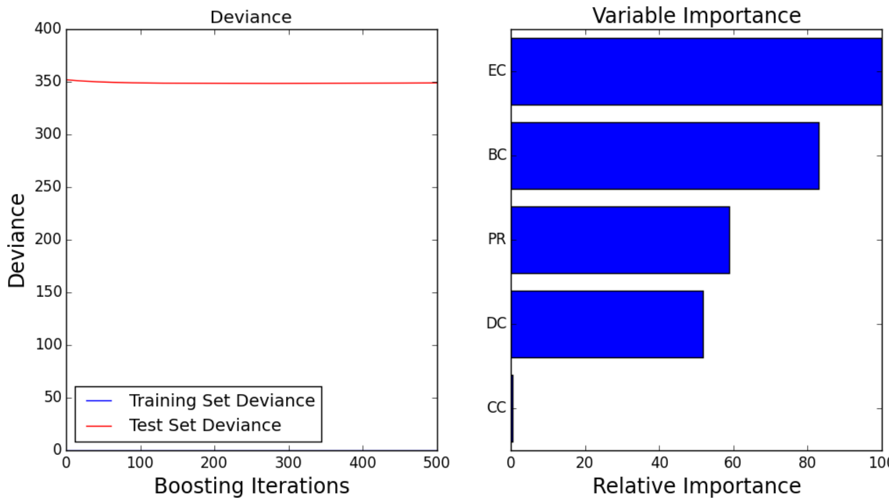


Fig. 9. **Left Panel**: Test and training set error vs. Boosting Iteration. **Right Panel**: Relative Feature Importance (Epinions I Dataset). We used the `median` operator to aggregate review scores.

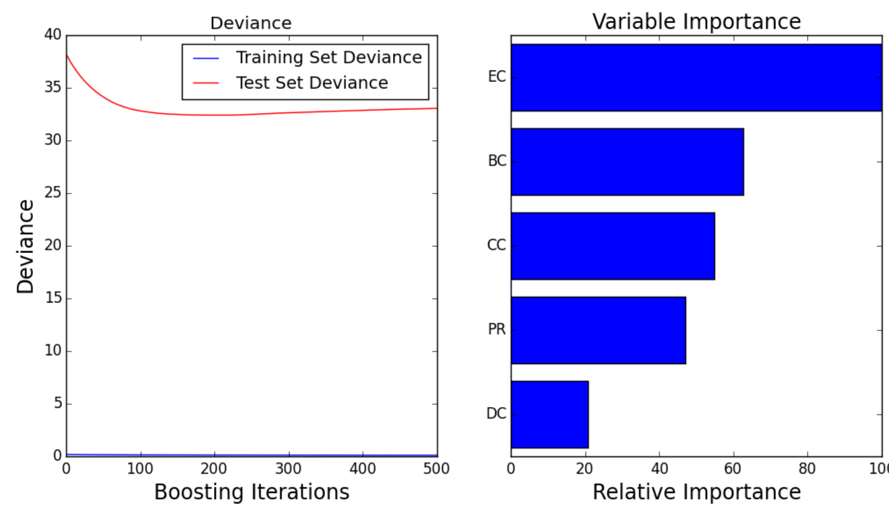


Fig. 10. **Left Panel**: Test and training set error vs. Boosting Iteration. **Right Panel**: Relative Feature Importance (CIAO Dataset). We used the `mode` operator to aggregate review scores.

her/his centrality in the trust network. The second method, called helpfulness-based reputation (HBR) scores, aggregates the feedbacks users provide in reaction to the comments/ratings posted by other platform members. The relationship between HBR and CBR scores in case of real-life trust networks has been only marginally investigated till now.

To answer our research question, we considered four real-life trust networks extracted from CIAO, Epinions and Wikipedia. All trust networks under investigation displayed a high level of reciprocity and transitivity and we also observed that trust links were created in such a way as to keep path length small; in addition, a large

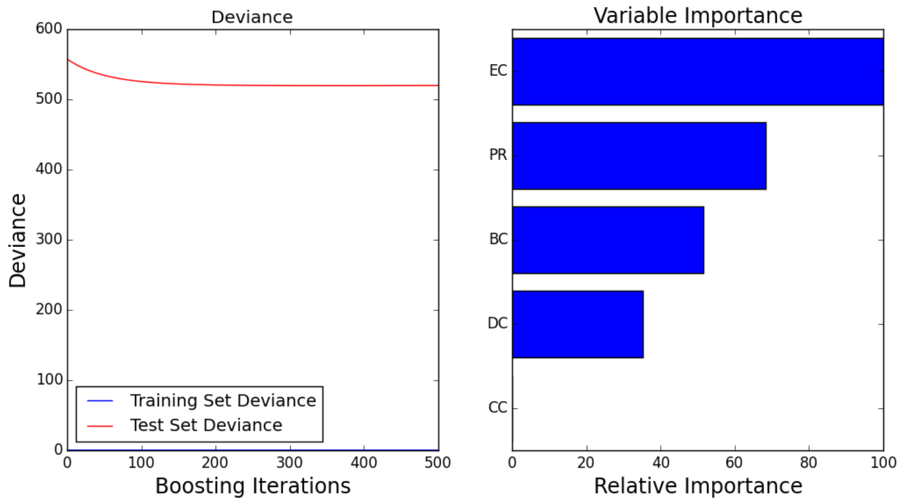


Fig. 11. **Left Panel**: Test and training set error vs. Boosting Iteration. **Right Panel**: Relative Feature Importance (Epinions I Dataset). We used the `mode` operator to aggregate review scores.

weakly connected component – often containing more than 90% of vertices – surfaced in all trust networks. We then computed CBR scores by applying five popular centrality measures (Degree Centrality, Closeness Centrality, Betweenness Centrality, Pagerank and Eigenvector Centrality). We found a high degree of correlation between CBR scores and this means that the user rankings generated by CBR scores were consistent.

As for CIAO and Epinions I dataset, we could manage user ratings and the helpfulness score associated with each rating. We used the *average*, the *median* and the *mode* operators to aggregate helpfulness scores and obtain the HBR score of each user. We applied Gradient Boosting Regression to shed light on the relationship between CBR and HBR scores. Our experiments provided evidence that CBR scores are good predictors of HBR ones and, independently of the trust network under inquiry and the operator used to aggregate reviews, the Eigenvector Centrality showed the best predictive power.

7.2. Limitations

The most important limitation of our work is about potential bias in the prediction task: a user u may be induced to trust a user v because v has posted many helpful reviews, and then, the creation of trust relationships would be a consequence of high HBR scores. If this would happen, we could conjecture that the formation of trust links is regulated by the same processes controlling the helpfulness of a user. Such a scenario could be problematic because CBR and HBR scores would be perfectly aligned and, then, we could not leverage the knowledge of network structure to spot reputable users. In practice, we observe that the notion of trust heavily depends on the application context and, then, in some domains the relationship between HBR scores and the creation of trust links may be stronger than in others. In addition, we observe that the formation of trust links is not only a consequence of the gratitude of u towards v but it may be influenced by other factors like the sense of belonging to the same community, pre-existing social relationships and the homophily principle (i.e., two individuals are more likely to trust each other if they share values and beliefs) [de Bunt et al. 2005]. Understanding how HBR scores contribute to the formation and growth of a trust network is a challenging but rewarding research stream that we plan to investigate in our future work. In line with these considerations, we observe that Eigenvector Centrality contributed the most to predicting HBR scores but we do not know to what extent it depends on the specific application domain on which we performed our tests (i.e., product review Web sites). We plan to consider further application domains and to study if Eigenvector Centrality is still the best predictor.

A further limitation of this work is that CBR scores reflect the importance of vertices in the trust network: vertices displaying the largest centrality also display the largest CBR scores. Further factors such as *structural holes* in social networks [Burt 1992] should be considered to more accurately predict HBR scores. A structural hole refers to a configuration in which there is a user v which is connected to at least two other users u and w, who are not connected each other. Users endowed with many structural holes occupy an advantageous position in a social network because they access more diverse information; to not lose their advantageous position, they should not put any effort in transforming weak interpersonal relationships into stronger ones but they should try to create new social links to access portions of the trust network out of their visibility. As a consequence, users with many structural holes should be less active in initiating new trust relationships. The contribution of structural holes in the prediction of HBR scores has not been investigated in this paper but we plan to consider them in future research.

A third limitation is about the computational effort required to calculate CBR scores. The computation of some centrality indices like BC is infeasible in large trust networks. In line with our previous works, we plan to consider randomized algorithms to efficiently but accurately compute centrality indices [De Meo et al. 2014].

Acknowledgment

Our deep appreciation and gratitude goes to all the reviewers who provided their insightful and constructive comments that help us to tremendously improve the quality of this paper. This work has been partially supported by (i) Program “Programma Operativo Nazionale Ricerca e Competitivitá” 2007-2013, project BA2Kno (Business Analytics to Know) PON03PE 00001 1, in “Laboratorio in Rete di Service Innovation”; (ii) Program “Programma Operativo Nazionale Ricerca e Competitivitá” 2007-2013, Distretto Tecnologico CyberSecurity funded by the Italian Ministry of Education, University and Research.

REFERENCES

ALSTOTT, J., BULLMORE, E., AND PLENZ, D. 2014. powerlaw: a Python package for analysis of heavy-tailed distributions. *PloS one 9,* 1, e85777.

BACKSTROM, L. AND LESKOVEC, J. 2011. Supervised random walks: predicting and recommending links in social networks. In *Proc. of the International Conference on Web search and Data Mining (WSDM 2011)*. ACM, Hong Kong, China, 635–644.

BAKSHY, E., ROSENN, I., MARLOW, C., AND ADAMIC, L. 2012. The role of social networks in information diffusion. In *Proc. of the International Conference on World Wide Web (WWW 2012)*. ACM, Lyon, France, 519–528.

BAVELAS, A. 1948. A mathematical model for group structures. *Human organization 7,* 3, 16–30.

BERMAN, A. AND PLEMMONS, R. 1994. *Nonnegative matrices in the mathematical sciences*. SIAM.

BISHOP, C. 2006. *Pattern recognition and machine learning*. Springer.

BONACICH, P. AND LLOYD, P. 2001. Eigenvector-like measures of centrality for asymetric relations. *Social Networks 23,* 3, 191–201.

BONCHI, F., CASTILLO, C., GIONIS, A., AND JAIMES, A. 2011. Social network analysis and mining for business applications. *ACM Transactions on Intelligent Systems and Technology 2,* 3, 22.

BRIN, S. AND PAGE, L. 1998. The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks 30,* 1-7, 107–117.

BURT, R. 1992. *Structural holes: The social structure of competition*. Harvard University Press.

BUTLER, B., SPROULL, L., KIESLER, S., AND KRAUT, R. 2007. Community effort in online groups: Who does the work and why. In *Leadership at a distance: Research in technologically supported work*, S. Weisband and L. Atwater, Eds. Lawrence Erlbaum, 171–194.

CHEN, J., ZHANG, C., AND XU, Y. 2009. The role of mutual trust in building members' loyalty to a C2C platform provider. *International Journal of Electronic Commerce 14,* 1, 147–171.

CHIRITA, P., NEJDL, W., SCHLOSSER, M., AND SCURTU, O. 2004. Personalized Reputation Management in P2P Networks. In *Proc. of the International Workshop on Trust, Security, and Reputation on the Semantic Web*. Springer, Hiroshima, Japan.

CLAUSET, A., SHALIZI, C., AND NEWMAN, M. 2009. Power-law distributions in empirical data. *SIAM review 51,* 4, 661–703.

DE BUNT, G. V., WITTEK, R., AND DE KLEPPER, M. 2005. The evolution of intra-organizational trust networks the case of a German paper factory: an empirical test of six trust mechanisms. *International Sociology 20,* 3, 339–369.

DE MEO, P., FERRARA, E., FIUMARA, G., AND PROVETTI, A. 2014. Mixing local and global information for community detection in large networks. *Journal of Computer and System Sciences 80,* 1, 72–87.

DE MEO, P., FERRARA, E., ROSACI, D., AND M.L. SARNÈ, G. 2015. Trust and compactness in social network groups. *IEEE T. Cybernetics 45,* 2, 205–216.

DE MEO, P., NOCERA, A., TERRACINA, G., AND URSINO, D. 2011. Recommendation of similar users, resources and social networks in a social internetworking scenario. *Information Sciences 181,* 7, 1285–1305.

ELITH, J., LEATHWICK, J., AND HASTIE, T. 2008. A working guide to boosted regression trees. *Journal of Animal Ecology 77,* 4, 802–813.

FOND, T. L. AND NEVILLE, J. 2010. Randomization tests for distinguishing social influence and homophily effects. In *Proc. of the International Conference on World Wide Web (WWW 2010)*. ACM, Raleigh, North Carolina, USA, 601–610.

FREEMAN, L. 1977. A set of measures of centrality based on betweenness. *Sociometry 40*, 35–41.

FRIEDMAN, J. 2001. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189–1232.

GARCIN, F., FALTINGS, B., AND JURCA, R. 2009. Aggregating reputation feedback. In *Proc. of the First International Conference on Reputation: Theory and Technology*. Springer, Arezzo, Italy, 62–74.

GOLBECK, J. 2006. Generating predictive movie recommendations from trust in social networks. In *Proc. of the International Conference on Trust Management (iTrust 2006)*. Pisa, Italy, 93–104.

GOLBECK, J. AND HENDLER, J. 2006. Inferring binary trust relationships in web-based social networks. *ACM Transactions on Internet Technology 6,* 4, 497–529.

GONG, N., TALWALKAR, A., MACKEY, L., HUANG, L., SHIN, E., STEFANOV, E., SHI, E., AND SONG, D. 2014. Joint Link Prediction and Attribute Inference Using a Social-Attribute Network. *ACM Transactions on Intelligent Systems Technology 5,* 2, 27.

GOOD, P. 2006. *Permutation, parametric, and bootstrap tests of hypotheses*. Springer Science & Business Media.

HENDRIKX, F., BUBENDORFER, K., AND CHARD, R. 2015. Reputation systems: A survey and taxonomy. *Journal of Parallel and Distributed Computing 75*, 184–197.

HOGG, T. AND ADAMIC, L. 2004. Enhancing reputation mechanisms via online social networks. In *Proc. of the ACM Conference on Electronic Commerce (EC 2004)*. ACM, New York, USA, 236–237.

JØSANG, A., ISMAIL, R., AND BOYD, C. 2007. A survey of trust and reputation systems for online service provision. *Decision Support Systems 43,* 2, 618–644.

KAMVAR, S., SCHLOSSER, M., AND GARCIA-MOLINA, H. 2003. The EigenTrust algorithm for reputation management in P2P networks. In *Proc. of the International Conference on World Wide Web (WWW 2003)*. ACM, Budapest, Hungary, 640–651.

KIM, S., PANTEL, P., CHKLOVSKI, T., AND PENNACCHIOTTI, M. 2006. Automatically assessing review helpfulness. In *Proc. of the International Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 423–430.

KUMAR, R., NOVAK, J., RAGHAVAN, P., AND TOMKINS, A. 2004. Structure and evolution of blogspace. *Communications of the ACM 47,* 12, 35–39.

LESKOVEC, J., HUTTENLOCHER, D., AND KLEINBERG, J. 2010. Predicting positive and negative links in online social networks. In *Proc. of the International Conference on World Wide Web (WWW 2010)*. Raleigh, North Carolina, USA, 641–650.

LEVIEN, R. 2009. Attack-resistant trust metrics. In *Computing with Social Trust*. Springer, 121–132.

LIU, Y., HUANG, X., AN, A., AND YU, X. 2008. Modeling and predicting the helpfulness of online reviews. In *Proc. of the International Conference on Data Mining (ICDM'08)*. IEEE, 443–452.

LU, Y., TSAPARAS, P., NTOULAS, A., AND POLANYI, L. 2010. Exploiting social context for review quality prediction. In *Proc. of the International Conference on World Wide Web (WWW 2010)*. ACM, ACM Press, Raleigh, North Carolina, USA, 691–700.

MA, H., KING, I., AND LYU, M. 2009. Learning to recommend with social trust ensemble. In *Proc. of the International Conference on Research and development in Information Retrieval (SIGIR 2009)*. ACM, ACM Press, Boston, MA, USA, 203–210.

MCALEXANDER, J., SCHOUTEN, J., AND KOENIG, H. 2002. Building brand community. *Journal of Marketing 66,* 1, 38–54.

MELNIK, M. AND ALM, J. 2002. Does a seller's ecommerce reputation matter? Evidence from eBay auctions. *Journal of Industrial Economics 50,* 3, 337–349.

MISLOVE, A., MARCON, M., GUMMADI, K., DRUSCHEL, P., AND BHATTACHARJEE, B. 2007. Measurement and analysis of online social networks. In *Proc. of the ACM SIGCOMM conference on Internet Measurement*. ACM, ACM Press, San Diego, USA, 29–42.

MUSIAL, K. AND KAZIENKO, P. 2013. Social networks on the Internet. *World Wide Web 16,* 1, 31–72.

NEVILLE, J., ŞIMŞEK, O., JENSEN, D., KOMOROSKE, J., PALMER, K., AND GOLDBERG, H. 2005. Using relational knowledge discovery to prevent securities fraud. In *Proc. of the ACM SIGKDD international conference on Knowledge discovery in data mining (SIGKDD 2005)*. ACM, ACM Press, Chicago, Illinois, USA, 449–458.

NEWMAN, M. 2010. *Networks: an introduction*. Oxford University Press.

O'MAHONY, M. AND SMYTH, B. 2009. Learning to recommend helpful hotel reviews. In *Proc. of the International Conference on Recommender Systems (ACM RecSys 2009)*. ACM, New York City, NY, USA, 305–308.

PUJOL, J., R. SANGÜESA, R., AND DELGADO, J. 2002. Extracting reputation in multi agent systems by means of social network topology. In *Proc. of the International joint conference on Autonomous agents and multiagent systems (AAMAS 2002)*. ACM, Bologna, Italy, 467–474.

RICHARDSON, M., AGRAWAL, R., AND DOMINGOS, P. 2003. Trust management for the Semantic Web. In *Proc. of the International Conference on Semantic Web (ISWC 2003)*. Springer, White Plains, NY, USA, 351–368.

RONQUI, J. AND TRAVIESO, G. 2015. Analyzing complex networks through correlations in centrality measurements. *Journal of Statistical Mechanics: Theory and Experiment 2015,* 5, P05030.

SIRIVIANOS, M., KIM, K., GAN, J., AND YANG, X. 2014. Leveraging social feedback to verify online identity claims. *ACM Transactions on the Web 8,* 2, 9.

TANG, J., HU, X., GAO, H., AND LIU, H. 2013. Exploiting local and global social context for recommendation. In *Proc. of the International Joint Conference on Artificial Intelligence (IJCAI 2013)*. Beijing, China, 2712–2718.

WASSERMAN, S. AND FAUST, K. 1994. *Social network analysis: Methods and applications*. New York: Cambridge University Press.

WENG, J., LIM, E., J. JIANG, J., AND HE, Q. 2010. Twitterrank: finding topic-sensitive influential twitterers. In *Proc. of the ACM international conference on Web Search and Data Mining (WSDM 2010)*. ACM, New York, USA, 261–270.

YANG, J., HAUFF, C., BOZZON, A., AND HOUBEN, G. 2014. Asking the right question in collaborative Q&A systems. In *Proc. of the ACM Conference on Hypertext and Social Media (HT 2014)*. ACM, Santiago, Chile, 179–189.

YIN, Z., GUPTA, M., WENINGER, T., AND HAN, J. 2010. LINKREC: a unified framework for link recommendation with user attributes and graph structure. In *Proc. of the International Conference on World Wide Web (WWW 2010)*. ACM, Raleigh, North Carolina, USA, 1211–1212.

ZHANG, R. AND MAO, Y. 2014. Trust prediction via belief propagation. *ACM Transactions on Information Systems 32,* 3.

ZHANG, Z. AND VARADARAJAN, B. 2006. Utility scoring of product reviews. In *Proc. of the ACM Conference on Information and Knowledge Management (CIKM 2006)*. ACM Press, Arlington, Virginia, USA, 51–57.

ZHU, Y. 2010. Measurement and analysis of an online content voting network: a case study of Digg. In *Proc. of the International Conference on World Wide Web (WWW 2010)*. ACM, Raleigh, North Carolina, USA, 1039–1048.