Contents

Chapter 1

Trust in Online Communities

Athirai A. Irissappane

Nanyang Technological University, Singapore

Jie Zhang

Nanyang Technological University, Singapore

1.1 Introduction

Internet World Stats[1] show that the number of online users worldwide has reached 2.75 billion as of March 2013, accounting for almost 38.8 percent of the global population. The increase has shown a huge impact on the growth of online communities such as e-commerce, social networks, content sharing sites, etc., especially in recent years. Trust has become a crucial factor for users who interact online, due to the limited web interface that does not allow to judge the trustworthiness of the interacting partner as in a typical face-to-face interaction. This is because online interactions are more impersonal, automated, provide fewer direct sensory cues, have less immediate gratification, entail more legal uncertainties, and present more opportunities for fraud and abuse [13]. This is even more the case with e-commerce and business transactions which deal with monetary value.

Reputation systems (trust models) promote online trust by identifying true reputation scores of entities (products/services/users) based on others' opinions. Such scores help to decide trustworthy interaction partners and engage in successful online business transactions. For example, in e-commerce, reputation systems collect, distribute, and aggregate feedback about the past

[1] http://www.internetworldstats.com/emarketing.htm

behavior of buyers and sellers in the system. Buyers who previously bought products from a seller share their experience, normally in the form of a numerical rating reflecting the level of satisfaction for the transactions with the seller. These ratings are aggregated to represent the seller's reputation. The reputation value is then used by other buyers to make decisions on which sellers to do business with. Reputation systems are particularly useful for users with no or very little experience in the interaction environment. These systems help people decide whom to trust, encourage trustworthy behavior, and deter participation by those who are unskilled or dishonest. Prominent reputation systems include those in commerce (*e.g.* eBay, Epinions), search (*e.g.* PageRank), blogs (*e.g.* Blogger), peer-to-peer networks (*e.g.* EigenTrust), etc. The following sections will outline the importance of trust and how trust evaluation is performed using reputation systems in various online communities.

1.2 Trust in E-Commerce Environments

Trust is of prime importance in e-commerce environments, because of the huge impact they create on online transactions. Research shows that 92 percent of people do research before making a purchase[2], 85 percent of users read online customer reviews and ratings before making purchases and 75 percent say that positive customer reviews make them trust a business more[3]. Trust in the form of ratings and reviews, promote or demote a product or service. Fig. 1.1 shows the rating system in *TripAdvisor.com*, signifying the extent of trust users have on a particular airline.

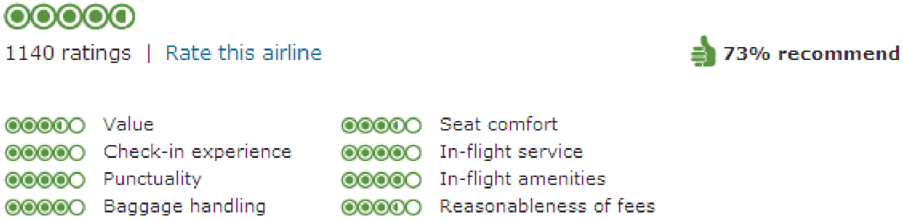

FIGURE 1.1: Rating system in *TripAdvisor.com*

In such open e-commerce environments, it is not easy to establish trust between interacting partners (buyers and sellers) because self interested sellers can act maliciously by not delivering products with the same quality as promised and buyers can provide misleading opinions [6], to promote some sellers (ballot-stuffing) or bad-mouth others. A reputation system collects

[2] http://reputationx.com/internet-reputation-management/
[3] http://www.marketingcharts.com/

feedback about participants' past behavior and assigns trust scores to each interacting partner. Doing so, it helps to select trustworthy sellers and buyers for successful transactions.

Several trust evaluation schemes for e-commerce systems have been proposed in literature. In the reputation system of eBay (*ebay.com*), both the buyers and sellers can provide feedback (rating $\in \{1, 0, -1\}$) on each other. A positive rating raises a buyer or seller's reputation score by 1 point, while a negative rating lowers the reputation score by 1 point. The feedback could also be in the form of text comments. In general, the observed ratings on eBay are positive. Resnick et al. [28] also found that there is a high correlation between buyer and seller ratings, suggesting that there is a degree of reciprocation of positive ratings and retaliation of negative ratings between buyers and sellers on eBay. This is problematic if obtaining honest and fair ratings is a goal, and a possible remedy could be to not let sellers rate buyers. However, as the market matured, sellers who have accumulated a lot of positive feedbacks were given higher scores, thus accounting for the reliability of the reputation system. Yahoo! Auction (*auctions.yahoo.com*), Amazon (*amazon.com*) and other auction sites extend eBay's reputation system by using different rating scales or aggregation schemes [27]. Specifically, Amazon allows sellers to be evaluated by buyers on a rating scale of $1 - 5$ stars, as well as to add a text comment. $4 - 5$ stars correspond to positive feedback, 3 stars represent neutral, and $1 - 2$ stars correspond to negative feedback. The overall rating is then calculated according to the average star rating for that particular seller. The Sporas system [41] calculates the trust score based on the ratings of transactions in a recent time period. In this method, the ratings of later transactions are given higher weights as they are more important in trust evaluation. The Histos system proposed in [41] is a more personalized reputation system compared to Sporas. Unlike Sporas, the reputation of a seller in Histos depends on who makes the query, and how that person rated other sellers in the online community. Some other fuzzy logic based reputation models also exist in literature, e.g., Song et al. [30] perform trust evaluation using fuzzy logic and their approach divides sellers into multiple classes of trust ranks (e.g., a 5-star seller or 4-star seller).

Reputation systems in e-commerce play a vital role in distinguishing honest behavior of buyers and sellers from malicious ones. However, reputation systems have widely become victims to the unfair rating problem, where advisors (i.e. buyers providing feedbacks) provide misleading opinions about sellers, to alter their trust scores. One such reported case is when Advertising Standards Agency launched an investigation into the popular review website *TripAdvisor.com* over malicious reviews due to which many small business from hospitality industry suffered. These business often received defamatory reviews (which sometimes could be a part of a coordinated attack) and almost lost their business[4].

[4] http://www.dailymail.co.uk/

Many trust schemes for multi-agent e-marketplaces have been proposed to deal with the unfair rating problem. The Beta Reputation System (BRS) [15] calculates seller reputation using a probabilistic model based on the beta probability density function, which can be used to represent probability distributions of binary events. The beta distributions are a family of statistical distribution functions that are characterized by two parameters α and β. The beta probability density function is defined as,

$$beta(p|\alpha,\beta) = \frac{\gamma(\alpha+\beta)}{\gamma(\alpha)\gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} \tag{1.1}$$

where, γ is the gamma function, $p \in [0,1]$ is a probability variable, and $\alpha, \beta > 0$. To calculate the reputation of a seller, ratings received by the seller are combined by simply aggregating the number of positive ratings (m), signifying that the seller is of high quality and the number of negative ratings (n), signifying that the seller is of low quality. The reputation of seller s, $R(s)$ is then calculated as the expected value of the beta probability distribution,

$$R(s) = E(p) = \frac{\alpha}{\alpha+\beta} \text{ ; where } \alpha = m+1 \text{ ; } \beta = n+1 \tag{1.2}$$

To handle unfair ratings provided by advisors, Whitby et al. [37] extend BRS to filter out those ratings that are not in the majority amongst other ones by using the iterated filtering approach. Feedback of each advisor to a seller s (consisting of both positive and negative ratings) is represented by a beta distribution. If the cumulated reputation of the seller s (obtained using the ratings of all advisors in the market) falls outside the q and $1-q$ quantile of the beta distribution formed by the advisor's ratings to the seller, then the advisor will be considered dishonest and filtered. However, the iterated filtering approach is only effective when a significant majority of ratings are fair and filters out the ratings that are not in the majority amongst others. From Fig. 1.2, we can see that when the calculated reputation of a seller falls outside the quantile ($q = 0.01$) region (0.01 and 0.99) of the beta distribution formed by a given advisor's ratings to the seller (with $m = 8$ and $n = 2$, respectively), the ratings from the advisor will be considered as unfairly high or unfairly low ratings and filtered.

Teacy et al. [31] propose the TRAVOS model to discount unfair ratings by modeling the trustworthiness of advisors based on their personal experience with the advisors' ratings. This approach is also based on the beta probability density function. It copes with unfair ratings by accomplishing two tasks: 1) it estimates the accuracy of the current feedback (ratings of 1 or 0) provided by the advisor about the seller, by evaluating the buyer's personal experience with the advisor's previous advice. More specifically, it divides the interval of [0, 1] into *bin* number of equal bins. It then finds out all the previous advice provided by the advisor that is similar to the advice being currently given by the advisor. The two pieces of advice are similar if they are within the same bin. The accuracy of the current advice will be the expected value of the beta

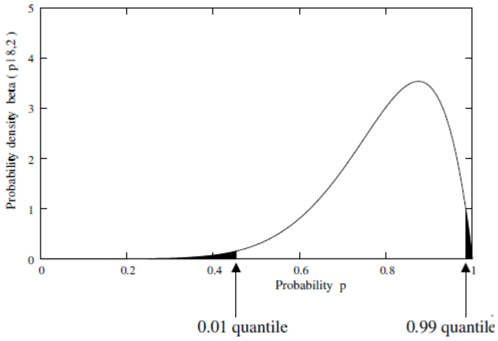

FIGURE 1.2: 1% and 99% quantiles of $beta(p|8,2)$

probability density function representing the amount of the successful and unsuccessful interactions between the buyer and the seller, when the buyer follows the previous advice; 2) the approach then adjusts the advisors ratings according to the obtained accuracy.

The Personalized approach proposed by Zhang and Cohen [43] combines buyers' personal experience and the public knowledge held by the system, to model the trustworthiness of the advisors. Private reputation of an advisor a, $R_{pri}(a)$ is calculated (using the beta probability density function) by comparing the advisor's ratings with the buyer's personal ratings regarding the commonly rated sellers. If the ratings are similar, a higher reputation value is achieved. Public reputation $R_{pub}(a)$ is estimated by comparing the advisor's ratings with other advisors' ratings regarding all sellers. The overall reputation of an advisor is then given by,

$$R(a) = wR_{pri}(a) + (1 - w)R_{pub}(a) \tag{1.3}$$

where, w is the reliability of the private reputation, calculated based on the minimum number of rating pairs needed to be confident about the private reputation value $R_{pri}(a)$ and the maximal acceptable level of error. A similar approach is adopted to calculate sellers' reputation i.e. by obtaining a weighted average of the private and public reputation values for sellers.

Several other approaches have also been proposed to deal with unfair ratings. Dellarocas [6] proposed a clustering-based algorithm to separate the advisor's ratings into two clusters (the cluster including lower ratings and the cluster including higher ratings). The ratings in the higher cluster are considered as unfairly high ratings and are discarded. However, this approach cannot effectively handle unfairly low ratings. The iCLUB approach [20] adopts a clustering technique (DBSCAN) to filter out dishonest advisors based on local and global information. Specifically, for a target seller, if advisors' ratings are not in the cluster containing the evaluating buyer's ratings, the advisors are considered to be dishonest. When the buyer has no direct experience (local information) with the target seller, the same process is applied on the nontarget sellers to identify the dishonest advisors. Yu and Singh [39] propose a distributed trust model to deal with real ratings. The Dempster-Shafer theory of evidence [19] is used as the underlying computational framework. A real rating is divided into three disjoint parts by predefined threshold settings and they are allocated into belief, disbelief and uncertainty, respectively. The weighted majority algorithm (WMA) is adopted to adjust the trustworthiness of advisors. If an advisor's opinion to the commonly rated sellers is not same as a buyer's experience, the buyer will decrease its trust value towards the advisor. The BLADE approach of Regan et al. [26] applies Bayesian learning to reinterpret advisors' ratings instead of filtering the unfair ones. The BLADE model allows the buyer to learn other advisors' evaluation functions on different features of the services delivered by sellers, by analyzing their ratings. This makes it possible to adjust the advisors' opinion, thereby coping with subjectivity and deception.

In recent times, trust evaluation is frequently based on many criteria, for example in *TripAdvisor.com*, an airline is rated based on 8 criteria (Fig. 1.1): value, check-in experience, punctuality, baggage handling, seat comfort, in-flight service, in-flight amenities and reasonableness of fees. Such detailed trust evaluation benefits users who may have different preferences for the various evaluation criteria. However, it increases the complexity of the trust evaluation engine to compute the multi-criteria trust score and identify the sellers and buyers exhibiting malicious behavior. Some schemes for multi-criteria trust modeling have also been proposed. Griffiths [9] introduced a multi-dimensional trust model tailored to a specific domain with a specific set of criteria. Each criterion is scored as a real number, and heuristics are proposed to update the score based on the buyer's direct experience. The weighted product model [1], which is a standard multi-criteria decision making technique, is used to combine the different criteria values while calculating the overall reputation of the seller. Each criteria score is raised to the power equivalent to its relative weight according to the evaluating buyer's preferences, while calculating the reputation of seller s, $R(s)$, given by,

$$R(s) = \prod_{i=1}^{n} [R_{c_i}(s)]^{w_{c_i}} \tag{1.4}$$

where, c is the evaluation criteria, $R_c(s)$ is the reputation score for the seller s on criteria c, and w_c is the weight (denoting the buyer's preference) on c. Reece et al. [25] model the seller's reputation by estimating the expected utility of a contract (based on various criteria) which is obtained by determining: 1) the probability that each contract dimension will be successfully fulfilled and 2) the correlations between these estimates. The Dirichlet distribution is used to calculate the probabilities and correlation. If $n_{c_1}, n_{c_2}, n_{c_3}$, etc., represent the number of outcomes for which each of the individual criteria c_1, c_2, c_3, etc., were successfully fulfilled, then in terms of standard Dirichlet parameters,

$\alpha_i = n_{c_i} + 1, \alpha_0 = \sum_i \alpha_{c_i} + 2$. The probability of the contract dimension c_i being successfully fulfilled, $p(o_{c_i} = 1)$ and variance V_{c_i} is given by,

$$p(o_{c_i} = 1) = \frac{\alpha_i}{\alpha_0};\ V_{c_i} = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(1 + \alpha_0)} \tag{1.5}$$

Apart from calculating the estimate of the probability that any contract dimension will be successfully fulfilled, the uncertainty and correlations in these probabilities is also calculated using the covariance matrix. It is found that using the Dirichlet formalism to calculate the multi-criteria trust score is more accurate than using multiple independent Beta distributions for each criteria (ignoring the correlations), because ignoring the correlation between the success probabilities of each criteria will lead to a miscalculation in estimating the uncertainty in the probability of each contract dimension being fulfilled. However, the Reece model evaluates only the reputation of sellers and simply assumes that advisor honesty can be modeled by extending trust models like TRAVOS [31]. Thus, we can see that the multi-criteria trust schemes, presented above do not address the problem of filtering dishonest advisors in multi-criteria environments to cope with the unfair rating problem. Irissappane et al. [14] propose a biclustering based approach to detect the malicious behavior of advisors providing misleading opinions specific to multi-criteria environments. Here, each buyer is assigned a set of biclusters, obtained by clustering advisors who are honest to a subset of criteria. Such a mechanism effectively identifies dishonest advisors, who provide honest ratings to some criteria, while acting malicious on others.

1.3 Trust in Search Engines

Given a query, a search engine identifies the relevant pages on the web and presents the users with the links to such pages, typically in batches of $10 - 20$ links. Once the users see relevant links, they may click on one or more links in order to visit the corresponding pages. For many commercial web sites, an increase in search engine referrals translates to an increase in sales, revenue, and, one hopes, profits.

The early search engines such as Altavista simply presented every web page that matched the key words entered by the user, which often resulted in too many and irrelevant pages being listed in the search results. This is because some web-masters may promote web sites in a spam-like fashion by filling web pages with large amounts of commonly used search key words as invisible text or as meta-data in order for the page to have a high probability of being picked up by a search engine, no matter what the user searched for. To address this problem, current web search engines use link-based reputation systems (e.g. PageRank) to measure the importance of web pages and rank them in the order of their reputation scores.

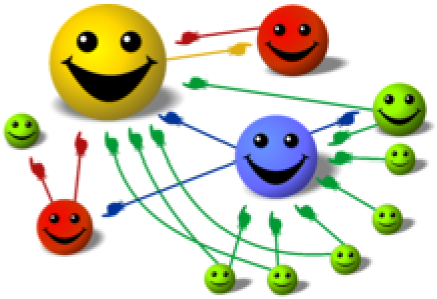

FIGURE 1.3: PageRank: The size of each face is proportional to the total size of the other faces which are pointing to it

PageRank proposed by Page et al. [24] is a widely used scheme by Google to rank the best search results based on a page's reputation. The reputation of a page (also called as PageRank) is based on the number and reputation of other pages which are pointing at it. Fig. 1.3 illustrates how PageRank works. Here, the size of each face represents the reputation of a page[5] and is proportional to the total reputation of the other pages pointing to it. In fact, this can be described as a trust scheme, because the collection of hyperlinks to a given page can be seen as public information that can be combined to derive a reputation score. PageRank applies the principle of trust transitivity to the extreme because rank values can flow through looped or arbitrarily long hyperlink chains. If P is a set of hyperlinked pages containing pages u and v, $N(u)$ is the set of web pages pointing to u, $\widetilde{N}(v)$ is the set of web pages that v points to and E is some vector over P corresponding to the source of the rank, then the PageRank of u is given by [24, 16],

$$R(u) = cE(u) + c \sum_{v \in N(u)} \frac{R(v)}{|\widetilde{N}(v)|} \tag{1.6}$$

where, c is chosen such that $\sum_{u \in P} R(u) = 1$. The term $cE(u)$ gives the rank based on the initial rank and the term $c \sum_{v \in N(u)} \frac{R(v)}{|\widetilde{N}(v)|}$ gives the rank as a function of the hyperlinks pointing to u.

PageRank has reduced the problem of spam-like pages to a certain extent because a high reputation is also needed in addition to matching key words, in order for a page to be presented to the user, while displaying the search results. However, as we know, the common problem in reputation systems is manipulation; strategic users may arrange links attempting to boost their own reputation scores. On the web, this phenomenon is called link spam, and is usually targeted at PageRank. Though Google's PageRank can deal with this

[5] http://en.wikipedia.org/wiki/PageRank

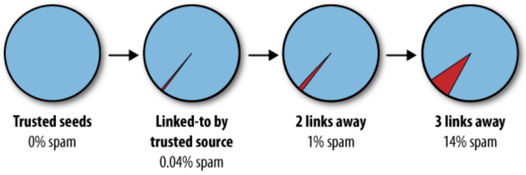

FIGURE 1.4: TrustRank

issue to some extent, still users can manage to obtain in-links to boost their own PageRank, and can also achieve this goal by carefully placing out-links. There are also certain works which specifically address the issue of link spam. TrustRank [11] addresses the problem of link spam by exploiting the intuition that good pages i.e. those of high quality are very unlikely to point to spam pages or pages of low quality. TrustRank propagates trust from the seed set of good pages, recursively to the outgoing links. The trust value is reduced as one moves further and further away from the good seed pages as shown in Fig. 1.4[6]. The BadRank algorithm, SpamRank algorithm, Anti-Trust Rank algorithm also deal with this issue [18].

1.4 Trust in P2P Information Sharing Networks

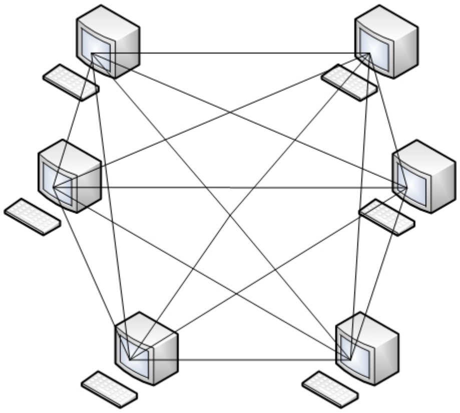

FIGURE 1.5: Peer to Peer (P2P) networks

In Peer-to-Peer (P2P) networks, peers communicate directly with each other to exchange information and share files [23] in a decentralized manner (Fig. 1.5). All peers are both consumers and providers of resources and can access each other directly without intermediary peers [35]. In an open P2P

[6] http://programming4.us/website/1596.aspx

system, peers often have to interact with unknown peers and need to manage the risks involved in these interactions, as some peers might be buggy or malicious and cannot provide services with the quality that they advertise. For example, to download a file, a requesting peer should choose from a given list of peers that can provide the requested file and download it. The requesting peer then has to check the downloaded file for any malicious content and if it actually corresponds to the requested file (i.e. the requested content). If the file is corrupted, it needs to download the file again. In traditional P2P systems, little information is given to the peers to help in the selection process. Since there is no centralized node to serve as an authority to monitor and punish the peers that behave badly, malicious peers have an incentive to provide poor quality services for their benefit because they can get away. Therefore, P2P systems are highly vulnerable to various types of attacks (denial-of-service attacks, etc.). To protect themselves from malicious intentions, requesting peers should be able to identify trustworthy peers for communication, which is quite challenging in such highly dynamic networks.

The issue of trust has been actively studied in Peer-to-Peer (P2P) information sharing networks (e.g. [5, 17, 35]). Trust in P2P systems allows peers to cooperate, and obtain in the long term an increased utility for the participating peers. Here, the requesting peer needs to enquire the trust data of a serving peer (target peer) from other peers, who may have transacted with the serving peer [17, 22, 38]. The computation of the trust level of the serving peer from the collected trust ratings is then performed by the requesting peer rather than a central management server, because of the decentralized architecture of the P2P system. However, the major challenge in building such a trust mechanism is to effectively cope with various malicious behavior of peers such as providing fake feedback about other peers. Another challenge is the method of implementation of the trust system in P2P networks. Most existing trust schemes for P2P systems require a central server for storing and distributing the reputation information. Building a decentralized P2P trust management system that is efficient and scalable is quite cumbersome.

EigenTrust [17] is a renowned reputation management algorithm for P2P networks. It adopts a binary rating system, and aims to collect the local trust values of all peers to calculate the global trust value of a given peer. The local trust peer i has on peer j is given by,

$$sat_{ij}^{loc} = sat(i,j) - unsat(i,j) \tag{1.7}$$

where, $sat(i,j)$ and $unsat(i,j)$ represent the number of satisfactory and unsatisfactory transactions i previously had with j. The local trust value is normalized using Eqn. 1.8 and the global trust is then obtained by aggregating the normalized local trust values from all peers, weighted by their trustworthiness (Eqn. 1.9).

$$R_{ij}^{loc} = \frac{max(sat_{ij}^{loc}, 0)}{\sum_j max(sat_{ij}^{loc}, 0)} \tag{1.8}$$

$$R_{ik}^{global} = \sum_j R_{ij}^{loc} R_{jk}^{loc} \tag{1.9}$$

The core of the protocol is the normalization process (Eqn. 1.8), where the trust ratings held by a peer are normalized to have their sum equal to 1. Although it has some interesting properties, this normalization may result in the loss of important trust information. For example, the normalized trust values may not distinguish between a new peer and a peer with whom peer i had a bad experience. Also, it assumes that there are some peers in the market who are already known to be trustworthy.

Xiong et al. [38], propose a more efficient solution called PeerTrust to effectively evaluate the trustworthiness of peers and identify various malicious behaviors. They introduce three basic trust parameters (i.e. the feedback that a peer receives from other peers, total number of transactions that a peer performs and credibility of the feedback sources) and two adaptive factors (i.e. transaction context factor to differentiate between transactions and community context factor to address community specific issues) in computing the trustworthiness of peers. Then, they define some general trust metrics and formulas to aggregate these parameters into a final trust value, given by,

$$T(u) = \alpha \times \sum_{i=1}^{I(u)} S(u,i) \times Cr(p(u,i)) \times TF(u,i) + \beta \times CF(u) \tag{1.10}$$

where, $T(u)$ is the trust value of peer u, $I(u)$ is the total number of transactions performed by peer u with all other peers, $S(u,i)$ is the normalized satisfaction u receives in the i^{th} transaction, $p(u,i)$ is the participating peer in the i^{th} transaction, $Cr(p(u,i))$ is the credibility of peer $p(u,i)$, $TF(u,i)$ is the adaptive transaction context factor and $CF(u)$ is the adaptive community context factor. α and β denote the normalized weight factors for the collective evaluation and community context factor. The implementation architecture [38] of PeerTrust is shown in Fig. 1.6. There is no central database and the trust data needed to compute the trust measure for peers is stored across the network in a distributed manner. Each peer has a trust manager which 1) provides feedback to the appropriate peers using the data locator; 2) evaluates the trustworthiness of a peer by collecting data from other peers.

Damiani et al. [5] propose an approach for evaluating the reputation of peers through a distributed polling algorithm and the XRep protocol before initiating any download action. This approach adopts a binary rating system and is based on the Gnutella[7] query broadcasting method. The following steps are used in the process: 1) resource searching: initiator p sends a query message for searching resources, peers matching that request respond with a query hit; 2) vote polling: p polls its peers about the reputation of a resource r and the set T of serving peers that offer it, peers wishing to respond send back a

[7] http://www.gnutella.com/.

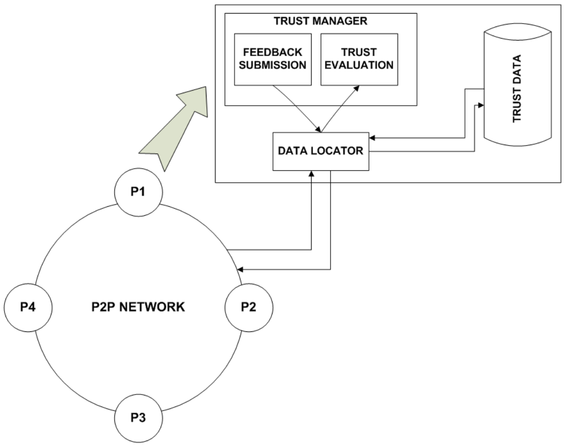

FIGURE 1.6: PeerTrust system architecture

poll reply; 3) vote evaluation: p selects a set of reliable voters and contacts them directly regarding their opinion about r and T; 4) best servant check: p contacts the best serving peer s to check the fact that it exports resource r; 5) resource download: finally p selects s, downloads a resource r, checks its integrity and updates its opinion based on the downloaded resource. Other P2P reputation systems include that of Marti et al. [22], who propose a voting system that collects responses from other peers regarding a target serving peer. The final reputation value is calculated by aggregating the values returned by the responding peers and the requesting peer's experience with the target peer. Zhou et al. [44] explore a power-law distribution in peer feedbacks, and develop a reputation system with a dynamic selection of a small number of power nodes that are the most reputable in the system.

1.5 Trust in Service-Oriented Environments

In service-oriented computing (SOC), service clients interact with service providers for services or transactions. From the point view of service clients, the trust status of a service provider is a critical issue to consider, particularly when the service provider is unknown to them. Typically, the trust evaluation is based on the feedback provided by service clients, about the quality of the service providers. In SOC environments, it is more feasible for the central trust management server(s) to compute the trust values and respond them as services to the requesting clients.

In the literature, the issue of trust has also received much attention in the field of service-oriented computing. Vu et al. [32] present a model to evaluate service trust by comparing the advertised service quality and that actually delivered. If the advertised service quality is as good as the delivered service quality, the service is reputable. The model consists of two phases: 1) service discovery: a list of web services with similar functionalities as required by the user is obtained from a matchmaking framework; 2) service ranking: it ranks the obtained services based on their predicted QoS values, taking into consideration the explicit quality requirements of users in the queries. For this, user reports on the QoS of all services over time are collected. The predicted QoS values are also based on the quality promised by the service providers, while still considering trust and reputation issues. Wang et al. [34] propose a fuzzy reputation scheme, for trust evaluation in SOC environments. The trust value of a service provider at time t_{k+1} is given by,

$$R_{k+1}(s) = \begin{cases} min(1, R_k(s) + \theta \times \Delta), \text{if } \Delta \geq 0 \\ max(0, R_k(s) + \theta \times \Delta), \text{if } \Delta < 0. \end{cases} \tag{1.11}$$

where, $\Delta = R_{k+1}(s) - R_k(s)$ is the difference between the actual rating given to the service provider and the previously predicted trustworthiness. $0 \leq \theta \leq 1$ is the impact factor determining the impact of the recent change Δ on the trust calculation. $\theta = \lambda \times f'(R_k(s))$, where $\lambda > 0$ and $f'(R_k(s))$ is the derivative of the curve function $f(R_k(s))$. The curve function $f(R_k(s))$ depicts the trust evaluation function for the service provider, obtained over a period of time. To more accurately reflect the trust status, than a mere numerical value $R_{k+1}(s)$, 5 fuzzy sets ’very low’, ’low’, ’moderate’, ’high’ and ’very high’, along with their membership functions are set up to categorize $R_{k+1}(s)$ into trust ranks.

Malik et al. [21] propose a decentralized technique to facilitate trust-oriented selection and composition of web services. The trust value of a service provider s (based on various evaluation criteria), $R(s)$ is given by,

$$R(s) = \frac{\sum_i [\frac{\sum_c (\Phi_c(s,i) \times \Psi_c)}{\sum_c \Psi_c} \times \lambda \times Cr(i)]}{\sum_i Cr(i)} \tag{1.12}$$

where, i represents the other service clients in the market, $\Phi_c(s,i)$ is the rating given by i to s for the evaluation criteria c, Ψ_c is the preference of the client evaluating s, for criteria c. λ denotes the reputation fader, to give more weights to recent ratings and $Cr(i)$ is the credibility of the client i.

Wang et al. [36] describe a super-agent based framework for web service selection, where service clients with more capabilities act as super-agents. These super-agents maintain reputation information of the service providers and share such information with other service clients, which have less capabilities than the super-agents. Also, super-agents maintain communities and build community-based reputation for a service provider based on the opinions from all community members (service clients in a community) that have similar interests and judgement criteria as the super-agents or the other community members. A reward mechanism is also introduced to create incentives for super-agents to contribute their resources (to maintain reputation and form communities) and provide truthful reputation information.

While most of the works on trust evaluation in SOC have focused on accurately predicting trust scores, Conner et al. [4] present a trust model that allows each service client (with different trust requirements) to use different scoring functions over the same feedback data for customized evaluations. Rather than assuming a single global trust metric like many existing reputation systems, they allow each service client to use its own trust metrics to meet its local trust requirements. They also propose a novel scheme to cache the calculated trust values based on recent client activity.

1.6 Trust in Social Networks

The proliferation of web-based social networks has led to new innovations in social networking, particularly by allowing users to describe their relationships beyond a basic connection. A social network is a set of people, connected by a set of social relationships such as friendship, co-working or information exchange [7]. People share information, express opinions, exchange ideas, make friends, and therefore form social networks. Some of the famous social networking sites include $Facebook.com, Twitter.com, Linkedin.com$, etc. The relationships in web-based social networks are more complex than social network models traditionally studied in the social sciences because users can make a variety of assertions about their relationships with others. For example, users may state how well they know the person to whom they are connected or how much they trust that person. These expanded relationships mean that analysis of the networks can take the new information into account to discover more about the nature of the relationships between people. Also, lots of companies have launched their social media marketing programs on social networking sites, which usually center on efforts to create content that attracts attention and encourages readers to share it with their social networks. For example, large companies such as Adidas have established their communities on social network sites (e.g, $Facebook.com$). Through these communities, users are encouraged to browse and discuss the product information, which can promote the brand reputation of corresponding companies. However, the continuously growing size of users and amount of information with widely varying quality in social networks have also raised the important concern of trust among users, about whom to trust and which information to trust.

Social networks are mainly represented as connected graphs with (directed/undirected) edges representing human-established trust relations (e.g., friend relations) as shown in Fig. 1.7. The computational problem of trust is to determine how much one member in the social network should trust another member to whom they are not directly connected. To solve this issue, trust propagation, during which the trust of a target member can be estimated from the trust of other connected members in the social network, is widely used.

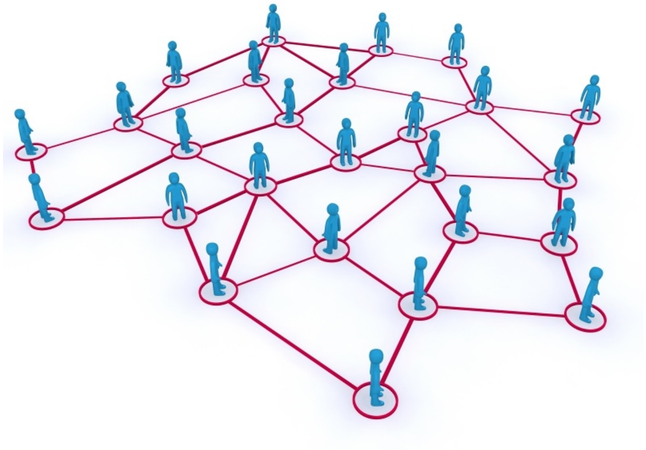

FIGURE 1.7: A social network: Edges represent the relationships between the individuals

There are usually many social trust paths between two members, who are unknown to one another (in large-scale social networks, there could be thousands of social trust paths between members). In addition, some social information, such as social relationships between members and the recommendation roles of members, can have a significant influence on the trust evaluation. Thus, evaluating the trustworthiness of a target member based on all the available social trust paths becomes quite challenging and time consuming. However, one can also search the optimal path yielding the most reliable trust propagation result from multiple paths. We call this the optimal social trust path selection problem, which is still a challenging research problem in this field.

In the literature, most of the trust evaluation schemes mainly exploit the social network structure and the social interactions between members [29] to select (optimal) social trust paths and accurately determine the trust scores. Buskens [2] observed that high density in a social network (i.e. high interconnectedness between members) can yield high level of trustworthiness and that members with high out-degree will have higher levels of trust. Caverlee et al. [3] propose a social trust model that exploits both social relationships and feedbacks for trust evaluation. Members in the social network provide feedback ratings after they have interacted with other members. The trust manager combines these feedback ratings to compute the social trust of the members. The member's feedback is also weighted by their link quality (high link quality indicates more links with members having high trust ratings). Golbeck et al. [8] present trust propagation algorithms based on binary ratings. They mainly consider three main concepts for trust evaluation: transitivity, asymmetry and personalization. To illustrate their trust inference scheme, we will consider a social network as shown in Fig. 1.8, in which source A has to infer the trust value of F. The source A will first poll each of the neighbors to

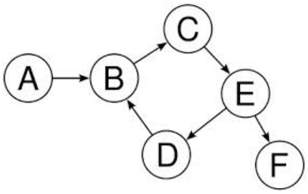

FIGURE 1.8: A social network: Node A needs to infer the trust value of F

which it has given a positive rating. Each trusted neighbor (B in this case) will return its rating for the target F. The source will then average these ratings to obtain the inferred reputation rating of F. Each of the source's neighbors will use this same process to come up with their reputation ratings for F. If there is a direct edge connecting them to the target, the value of that edge is used; otherwise, the value is inferred. In [12], Hang et al. propose an algebraic approach for propagating trust in social networks, including a concatenation operator for the trust aggregation of sequential invocation, an aggregation operator for the trust aggregation of parallel invocation, and a selection operator for trust-oriented multiple path selection.

As described above, most existing works for trust inference in social networks use the concept of trust propagation. However, experience with real-world trust systems such as those in Epinions and eBay suggest that distrust is at least as important as trust. To deal with this issue, Guha et al. [10], develop a framework, which uses both trust and distrust propagation, for trust inference in social networks.

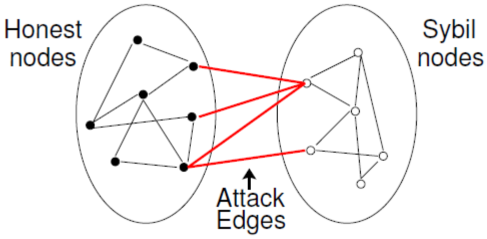

FIGURE 1.9: A social network showing honest and sybil nodes

Walter et al. [33], identify that network density, similarity of preference between members, and sparseness of knowledge about the trustworthiness of recommendations are crucial factors for trust-oriented recommendations in social networks. However, such trust-oriented recommendations can be attacked in various ways, such as sybil attacks, where the attacker creates unlimited number of false identities to provide feedback and modify the trust score. Yu et al. [40] present SybilGuard, a protocol for limiting the corruptive influences of sybil attacks, which depends on the established trust relationships between members in a social network. Their model is based on the observation that the edges connecting the honest members in the social network and the sybil region (called attack edges), are independent of the number of sybil identities created and is limited by the number of trust relation pairs between the sybil nodes and the honest members. If the malicious members create too many sybil identities, the graph becomes strange i.e. a small set of edges (attack edges) disconnects a large number of sybil nodes as shown in Fig. 1.9.

Zhang et al. [42] propose a scheme to combat sybil attacks in social networks by leveraging on both trust and distrust information. A sybil seed selection algorithm is presented to produce reliable sybil seeds, in combination with current social network-based sybil defence schemes. Moreover, a graph pruning strategy is introduced to reduce the attack ability near honest seeds, by exploiting local structure similarity between neighboring nodes. Finally, a ranking mechanism based on a variant of the PageRank algorithm is presented to combine trust and distrust together, in order to determine the trustworthiness of nodes in the social network and nodes with less trustworthiness score are more likely to be sybils.

1.7 Discussion

Trust evaluation in online communities has become crucial, mainly because of the risk associated while interacting online. But, it is often difficult to assess the trustworthiness of interaction partners, because computer mediated communication restricts a wide range of existing cues which allow people to easily assess trustworthiness in a physical interaction. The main aim of any reputation system is to distinguish between high and low quality products/services/users by collecting evidence from other members in the online community. However, the nature of the various online communities (e-commerce, search, P2P networks, social networks, etc.) also impose challenges, which the reputation system needs to deal with.

Online communities widely differ in their structure and content. Thus, reputation systems for the various online communities also differ in their trust evaluation methodology, to be suitable to the application environment. For example, in e-commerce systems buyers and sellers engage in business transactions and the main aim of reputation systems is to select trustworthy sellers by obtaining feedback from trustworthy advisors. Reputation systems for e-commerce environments can operate in a centralized or a decentralized manner, both of which have their advantages and disadvantages. In P2P networks, the main aim of reputation systems is to identify a trustworthy peer. Also, the reputation systems are mainly decentralized in order to suit the P2P network topology. In social networks, the members are represented by nodes of a graph, with edges representing the connectedness and the reputation systems should find a reliable path between a source node and target node, which are not directly connected. Apart from these differences, the threat models for reputation systems in the various online communities also seem to vary. While in e-commerce systems, unfair rating attacks and malicious seller behavior are of prime importance, in search engines, the reputation systems mainly need to deal with link spams. In P2P systems, again unfair rating behavior is a major concern. In social networks, sybil attacks pose a severe threat and the reputation systems proposed need to address this issue. Thus, the different nature of each online community demands different trust modelling schemes to address their specific needs and challenges.

Though there is a volume of literature on the theory and applications of reputation systems in online communities, research still needs to focus on the potential fields of improvement, addressing the specific challenges imposed by the online communities and the vulnerabilities of the reputation systems.

Bibliography

[1] Percy Williams Bridgman. *Dimensional analysis*. Yale University Press, 1922.

[2] Vincent Buskens. The social structure of trust. *Social networks*, 20(3):265–289, 1998.

[3] James Caverlee, Ling Liu, and Steve Webb. Socialtrust: Tamper-resilient trust establishment in online communities. In *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital libraries*, pages 104–114. ACM, 2008.

[4] William Conner, Arun Iyengar, Thomas Mikalsen, Isabelle Rouvellou, and Klara Nahrstedt. A trust management framework for service-oriented environments. In *Proceedings of the 18th International Conference on World Wide Web*, pages 891–900. ACM, 2009.

[5] Ernesto Damiani, De Capitani di Vimercati, Stefano Paraboschi, Pierangela Samarati, and Fabio Violante. A reputation-based approach for choosing reliable resources in peer-to-peer networks. In *Proceedings of the 9th ACM Conference on Computer and Communications Security*, pages 207–216. ACM, 2002.

[6] C. Dellarocas. Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. In *Proceedings of the 2nd ACM Conference on Electronic Commerce*, pages 150–157, 2000.

[7] Laura Garton, Caroline Haythornthwaite, and Barry Wellman. Studying online social networks. *Journal of Computer-Mediated Communication*, 3(1), 1997.

[8] Jennifer Golbeck and James Hendler. Inferring binary trust relationships in web-based social networks. *ACM Transactions on Internet Technology*, 6(4):497–529, 2006.

[9] N. Griffiths. Task delegation using experience-based multi-dimensional trust. In *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multiagent Systems*, 2005.

[10] Ramanthan Guha, Ravi Kumar, Prabhakar Raghavan, and Andrew Tomkins. Propagation of trust and distrust. In *Proceedings of the 13th*

International Conference on World Wide Web, pages 403–412. ACM, 2004.

[11] Zoltán Gyöngyi, Hector Garcia-Molina, and Jan Pedersen. Combating web spam with trustrank. In *Proceedings of the Thirtieth International Conference on Very Large Data Bases*, volume 30, pages 576–587. VLDB Endowment, 2004.

[12] Chung-Wei Hang, Yonghong Wang, and Munindar P Singh. Operators for propagating trust and their evaluation in social networks. In *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems*, volume 2, pages 1025–1032, 2009.

[13] Milena M Head and Khaled Hassanein. Trust in e-commerce: evaluating the impact of third-party seals. *Quarterly Journal of Electronic Commerce*, 3:307–326, 2002.

[14] Athirai A. Irissappane, Siwei Jiang, and Jie Zhang. A biclustering-based approach to filter dishonest advisors in multi-criteria e-marketplaces. In *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems*, 2014.

[15] Audun Jøsang and Roslan Ismail. The beta reputation system. In *Proceedings of the 15th Bled Electronic Commerce Conference*, pages 41–55, 2002.

[16] Audun Jøsang, Roslan Ismail, and Colin Boyd. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43(2):618–644, 2007.

[17] Sepandar D Kamvar, Mario T Schlosser, and Hector Garcia-Molina. The eigentrust algorithm for reputation management in p2p networks. In *Proceedings of the 12th International Conference on World Wide Web*, pages 640–651. ACM, 2003.

[18] Vijay Krishnan and Rashmi Raj. Web spam detection with anti-trust rank. In *AIRWeb*, volume 6, pages 37–40, 2006.

[19] Henry E Kyburg Jr. Bayesian and non-bayesian evidential updating. *Artificial Intelligence*, 31(3):271–293, 1987.

[20] S. Liu, J. Zhang, C. Miao, Y.L. Theng, and A.C. Kot. iCLUB: An integrated clustering-based approach to improve the robustness of reputation systems. In *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*, 2011.

[21] Zaki Malik and Athman Bouguettaya. Rateweb: Reputation assessment for trust establishment among web services. *The VLDB Journal-The International Journal on Very Large Data Bases*, 18(4):885–911, 2009.

[22] Sergio Marti and Hector Garcia-Molina. Limited reputation sharing in p2p systems. In *Proceedings of the 5th ACM Conference on Electronic Commerce*, pages 91–101. ACM, 2004.

[23] Loubna Mekouar, Youssef Iraqi, and Raouf Boutaba. Reputation-based trust management in peer-to-peer systems: taxonomy and anatomy. In *Handbook of Peer-to-Peer Networking*, pages 689–732. Springer, 2010.

[24] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. *Technical report, Stanford Digital Library Technologies Project*, 1998.

[25] S. Reece, A. Rogers, S. Roberts, and N.R. Jennings. Rumours and reputation: Evaluating multi-dimensional trust within a decentralised reputation system. In *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems*, 2007.

[26] K. Regan, P. Poupart, and R. Cohen. Bayesian reputation modeling in e-marketplaces sensitive to subjectivity, deception and change. In *Proceedings of the National Conference on Artificial Intelligence*, 2006.

[27] Paul Resnick, Ko Kuwabara, Richard Zeckhauser, and Eric Friedman. Reputation systems. *Communications of the ACM*, 43(12):45–48, 2000.

[28] Paul Resnick, Richard Zeckhauser, John Swanson, and Kate Lockwood. The value of reputation on ebay: A controlled experiment. *Experimental Economics*, 9(2):79–101, 2006.

[29] Wanita Sherchan, Surya Nepal, and Cecile Paris. A survey of trust in social networks. *ACM Computing Surveys (CSUR)*, 45(4):47, 2013.

[30] Shanshan Song, Kai Hwang, Runfang Zhou, and Yu-Kwong Kwok. Trusted p2p transactions with fuzzy reputation aggregation. *Internet Computing, IEEE*, 9(6):24–34, 2005.

[31] W.T.L. Teacy, J. Patel, N.R. Jennings, and M. Luck. TRAVOS: Trust and reputation in the context of inaccurate information sources. *Autonomous Agents and Multi-Agent Systems*, 12(2):183–198, 2006.

[32] Le-Hung Vu, Manfred Hauswirth, and Karl Aberer. Qos-based service selection and ranking with trust and reputation management. In *On the Move to Meaningful Internet Systems 2005: CoopIS, DOA, and ODBASE*, pages 466–483. Springer, 2005.

[33] Frank Edward Walter, Stefano Battiston, and Frank Schweitzer. A model of a trust-based recommendation system on a social network. *Autonomous Agents and Multi-Agent Systems*, 16(1):57–74, 2008.

[34] Yan Wang, Kwei-Jay Lin, Duncan S Wong, and Vijay Varadharajan. Trust management towards service-oriented applications. *Service Oriented Computing and Applications*, 3(2):129–146, 2009.

[35] Yao Wang and Julita Vassileva. Trust and reputation model in peer-to-peer networks. In *Proceedings of the Third International Conference on Peer-to-Peer Computing*, pages 150–157. IEEE, 2003.

[36] Yao Wang, Jie Zhang, and Julita Vassileva. A super-agent based framework for reputation management and community formation in decentralized systems. *Computational Intelligence*, 2014.

[37] A. Whitby, A. Jøsang, and J. Indulska. Filtering out unfair ratings in bayesian reputation systems. In *Proceedings of the International Joint Conference Autonomous Agents and Multiagent Systems Workshop on Trust in Agent Societies*, 2004.

[38] Li Xiong and Ling Liu. Peertrust: Supporting reputation-based trust for peer-to-peer electronic communities. *IEEE Transactions on Knowledge and Data Engineering*, 16(7):843–857, 2004.

[39] B. Yu and M.P. Singh. Detecting deception in reputation management. In *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems*, 2003.

[40] Haifeng Yu, Michael Kaminsky, Phillip B Gibbons, and Abraham Flaxman. Sybilguard: defending against sybil attacks via social networks. *ACM SIGCOMM Computer Communication Review*, 36(4):267–278, 2006.

[41] Giorgos Zacharia and Pattie Maes. Trust management through reputation mechanisms. *Applied Artificial Intelligence*, 14(9):881–907, 2000.

[42] Huanhuan Zhang, Chang Xu, and Jie Zhang. Exploiting trust and distrust information to combat sybil attack in online social networks. In *Proceedings of the 8th IFIP WG 11.11 International Conference on Trust Management*, 2014.

[43] J. Zhang and R. Cohen. Evaluating the trustworthiness of advice about seller agents in e-marketplaces: A personalized approach. *Electronic Commerce Research and Applications*, 7(3):330–340, 2008.

[44] Runfang Zhou and Kai Hwang. Powertrust: A robust and scalable reputation system for trusted peer-to-peer computing. *IEEE Transactions on Parallel and Distributed Systems*, 18(4):460–473, 2007.