CORE Metadata, citation and similar papers at core.ac.uk
Provided by Kent Academic Repository

Operational Models for Reputation Servers

D.W. Chadwick

Computing Laboratory, University of Kent, Canterbury, UK, CT2 7NF
d.w.chadwick@kent.ac.uk

Abstract. This paper devises a classification system for reputation systems based on two axes, namely: who performs the evaluation of a subject's reputation, and how the information is collected by the reputation system. This leads to 4 possible operational models for reputation systems, termed the Voting Model, the Opinion Poll Model, the MP Model and the Research Model, each of which is then analyzed. Finally, the paper postulates the inherent trustworthiness of each operational model, and concludes with a hypothesis of how these systems might evolve in the future.

1 Introduction

There is considerable interest in online reputation systems [1]. Many commercial web sites are employing different sorts of online reputation systems today. E-Bay [2] is probably the most well known example. QXL.com has a similar system. Slashdot.org also has its own rating system for ranking postings, based on the reputation (actually called *karma* by Slashdot) of the submitter. Prior to the evolution of online reputation systems the financial and banking world has had its own off-line reputation systems for very many years, in the shape of credit rating bureaus. For example, Dun and Bradstreet, now a global company, originally started in New York on 20th July1841.

Users use reputation systems to determine the trustworthiness of the people, organizations or services they want to do business with. They determine their trustworthiness in part from the information obtained from these reputation systems, and in part from their own knowledge, feelings, intuitions etc. For unknown entities, reputation systems will contribute most, whilst for well known entities they will contribute much less. But this begs the question "how trustworthy is the reputation system itself?" In order to determine the trustworthiness of a potential transaction partner one first needs to be able to trust the reputation system that provides you with information about the potential partner. One clearly can have little trust in the reputations dispensed by a fraudulent, biased or badly managed reputation system. But if one were able to gain similar reputation scores for a potential partner from two or more unrelated reputation systems, then one could have more confidence or trust in the reputation of that potential partner and in each of the dispensing reputation systems. Further, if one knew the methods, or operational models, by which each of the reputation systems operated, then this might help a user to determine how reliable or trustworthy are the reputations dispensed by it.

The aim of this paper therefore is to classify the different types of reputation system that could exist, based on their operational models, and to postulate the

P. Herrmann et al. (Eds.): iTrust 2005, LNCS 3477, pp. 108 – 115, 2005.
© Springer-Verlag Berlin Heidelberg 2005

inherent trustworthiness of these different types of reputation system, based solely on their operational models. The hypothesis is that some operational models are inherently more trustworthy than others, and therefore users can reasonably expect that the reputation information provided by them will be more reliable and trustworthy than that provided by other types of reputation system.

The rest of this paper is structured as follows. Section 2 proposes a classification scheme for the types of reputation system based on their operational models, and determines that there are 4 different types of reputation system. Sections 3 to 6 then analyze the properties of each of these models in more detail and postulates how they might be implemented. Section 7 discusses the inherent trustworthiness of each of these operational models and provides a ranking of them. Section 8 concludes and hypothesizes how these systems might evolve in the future.

2 Operational Models

This paper proposes two main axes for categorizing reputation servers. The primary axis distinguishes between who performs the evaluation of a subject's reputation based on the available information. The choice is between the actors who participate in transactions with the subject, or the reputation system itself. In the former case, each actor uses its own algorithm for computing the reputation of a subject. The reputation system then simply collates these values. In the latter case the reputation system takes raw data from the actors and uses its own algorithms to compute the reputation of the subjects.

The second axis distinguishes how the information is collected by the reputation service prior to collation and publishing the reputations. The reputation service can either gather and collate the data itself (the data pull mode) or the actors can spontaneously send data to the reputation service (the data push mode).

When these two axes are combined together we get the 2x2 matrix shown in Figure 1 below. Each of the four combinations has been given a name for ease of reference and this name is a metaphor to depict the primary characteristics of the operational model.

	Data Push	Data Pull
Actor evaluation	Voting model	Opinion Poll model
Reputation Server evaluation	MP[1] model	Research model

Fig. 1. The four operational models

Each of the four operational models is now discussed in more detail below.

[1] MP stands for Member of Parliament, an elected representative to the UK legislative body called the House of Commons.

3 The Opinion Poll Model

In this operational model, the reputation server actively collects reputation data from the actors. Each actor performs its own evaluation about the reputation of a subject, using its own algorithm, and based on its experiences of performing transactions with the subject. In real life humans do this sort of evaluation all the time about the shops they frequent, the people they meet, the political parties they vote for etc. The role of the reputation system is to find an appropriate sample of the actors and to gather and summarize the data, preferably using a simple publicly available algorithm. In the physical world opinion poll companies regularly collect this sort of information and publish the results. The most difficult operational aspect of this model for electronic reputation systems to implement is to discover who and where the actors are and how to contact them. In the physical world, when the actors are members of the public, opinion poll companies use electoral roles or telephone directories to determine who the actors are and where they live, or they simply stop a random sample of people in the street as they are passing by. If an opinion poll company is contracted by an organization to evaluate its reputation from the perspective of its customers, then the organization might provide the reputation service with its customer lists from which to obtain a sample of actors.

Engineering the latter in the virtual world is not difficult, because a directory of actors is available, but engineering the former is much more difficult since it involves finding out the opinions of the public. The difficulty lies in the fact that the equivalent of electoral rolls or public directory services do not exist on the Internet. The nearest thing we currently have to the electoral role is the Domain Name System (DNS) [3]. This lists all the publicly accessible services on the Internet, along with their IP addresses. The DNS is core to the functioning of the Internet. However, the DNS does not as yet hold reputation information or addresses of where it might be found by opinion poll servers. Two possibilities exist for this. One would be to define a new type of DNS resource record (RR), say the reputation (RP) record, that holds the name and reputation of the Internet service being reputed. The metric of the reputation e.g. a Likert scale, would need to be standardized. Actors would then write this DNS record into their DNS entry for each service they were assigning a reputation to. Opinion poll servers could then scan the DNS to sample these records and derive collective reputations for entities. The second method would be to define a new protocol for the gathering of reputation information from actors, say the Reputation Gathering Protocol (RGP), and then to register this protocol with the Internet Assigned Numbers Authority[2] (IANA) and get a well known port allocated to it. Once this is achieved, actors can simply register their RGP servers in the DNS, using the existing WKS RR [3], and this will allow opinion poll servers to contact them.

However both schemes suffer from a number of disadvantages. Firstly the granularity is wrong, since the DNS can only publish (reputation) information about Internet services, and not about individuals or organizations (unless they have their own Internet sites). Secondly the DNS is already heavily over-utilized and performs rather poorly. The IETF tries to keep tight control over it, and therefore is highly

[2] See www.iana.org

unlikely to sanction the definition of this new RR type or protocol, especially if it would cause significant performance penalties on existing DNS users. Finally, and most significantly, the DNS does not provide a scanning or search capability. DNS clients have to already know the DNS name of the entity they want to look up, before contacting the DNS to get its IP address. In the general case opinion poll servers won't know the DNS names of the actors they want to poll. Thus we need a search and discovery service that will allow opinion poll servers to search for all actors, or a subset of actors that meet pre-defined search criteria such as: size of business, no of transactions undertaken, currency of the data etc. before contacting the DNS. This implies that we need to either define schema for existing directories such as UDDI [10] or LDAP [11] (or both), or uses Web search engines such as Google, to enable this searching to take place. Web search engines already trawl the Internet for information, and so build their own internal directories of web pages available on the Internet. How complete these directories are depends upon the trawling methods used. The DNS on the other hand is guaranteed to be complete, since it holds the names of all publicly available services.

As can be seen from the above discussion, there are still a significant number of problems to be solved before opinion poll reputations systems become widely available.

4 The Voting Model

In this model, the actors evaluate the reputation of subjects, using their own algorithms and information, and then forward their decisions to a central Voting server. The role of the voting server is simply to collect messages that arrive, collate and summarize them (again using a simple publicly available algorithm), and then publish the results when asked. E-bay is one example of this type of reputation server in use today. Various shopping mall web sites also allow customers to register their votes about how well the stores in the mall are performing their various aspects of service provision, for example, timeliness of goods delivery, and quality of after sales service etc. This operational model is much easier to implement than the opinion poll model, since implicit in the voting model is a voter registration list, meaning that the system already has a full list of all actors that are allowed to vote. Thus a voting server does not need to have access to an external actor discovery or directory service, unlike the opinion poll model. Voting servers either keep their own lists of authorized actors, as in e-Bay, or have some way of authenticating a voter or a vote, as in e-voting systems [9]. They allow only these actors to lodge their votes with them. Very often these lists will be commercially sensitive, as in customer lists, and, if they contain personal data, will be protected by data protection legislation. Thus electronic voting type reputation systems are highly unlikely to make their lists public, or available to opinion poll servers. Therefore if a reputation subject is known to two or more voting type reputation systems, their reputation in each is likely to be calculated by different sets of actors.

5 The Research Model

In this model, the reputation server actively searches for information about subjects, and then evaluates it and publishes the results. The operations of the reputation server are complex and difficult to engineer. Not only does the reputation server have the problem of finding the actors, as in the Opinion Poll model, but also it has to determine what raw information to solicit from them and how to process and evaluate this in order to compute the reputations of the subjects. Such processes and algorithms are likely to be proprietary and commercially valuable.

Several examples of this model exist in the physical world, for example both Standard and Poor[3] and Dun and Bradstreet[4] provide credit ratings, and they are now global companies. Clearly this operational model can lead to a successful business model – if the reputation results are valuable they can be sold at a profit. If a client is considering whether to enter into a business venture with a subject or not, then knowing the subject's reputation can be worth a large amount of money to a client. But precisely because the server's operations are complex, the algorithms used to process the raw information are proprietary, and the results commercially valuable, then it is highly likely that the algorithms and processes used to calculate the reputations will be commercially sensitive and not open to public scrutiny. The implications of this on the trustworthiness of these types of reputation system are discussed below.

6 The MP Model

In this context, MP stands for Member of Parliament, a person elected to the UK House of Commons to represent a constituency. MPs should represent their constituency, but often they do not. When it comes to voting on issues in the House of Commons, they either usually follow the party line, or if a free vote is allowed, on such issues as capital punishment or hunting with dogs, they follow their own conscience. So even if constituents have sent them lots of letters imploring them to vote one way, they may quite freely decide to vote the opposite way.

A reputation server following the MP model, will be sent (pushed) raw data about subjects by the actors. Some of this may be data about transactions an actor has undertaken with a subject, others might be subject reputations evaluated by the actors themselves. The MP server will typically not have an actor list, and therefore not be able to tell which actors are genuine and which are not. Regardless of this, the MP server determines which data to use, which data to discard, and which other private information to use as well. Then using its own, usually unpublished, proprietary algorithms, it computes the reputation of subjects and publishes the results. The reputation results will primarily be based on the subjectivity of the MP server while the point of view of the actors submitting information may be ignored. Clearly this operational model is the most suspect in terms of reliability and trustworthiness.

[3] See http://www.standardandpoors.com/

[4] See http://www.dnb.com/us/

7 Trustworthiness of Reputation Servers

If one is relying on the reputations provided by a reputation server, one needs to ask how reliable or trustworthy are the reputations that the reputation server is providing. We now appraise the inherent trustworthiness or reliability of each of the four operational models.

The Opinion Poll model is inherently the most trustworthy and reliable, since the individual reputation scores have been calculated by very many actors. The opinion poll server merely needs to collate and sum the scores into an overall reputation for each subject, using a simple publicly available algorithm. Therefore it is very difficult for the reputation server to skew the results, unless it is configured to discard particular inputs, or bias the way actor selection is performed. Furthermore, it is difficult for an individual actor to try to skew the reputation of a subject, since they have no control over whether they are polled by the reputation server or not, and even if they are, they should be in a minority of one (a correctly operated opinion poll server should never poll the same actor twice, and each actor should only be registered once with the system). This indicates that the actor lists should have strong integrity protection against unauthorized modifications, and should have strong registration methods to prevent an actor registering multiple times. Of course, if many actors conspire together to inflate or deflate a subject's reputation, this is very difficult to protect against. Each actor's reputation scores are, in principle, available to be collected by any opinion poll server via the publicly available actor lists, therefore the resulting computed reputations are more easily validated and the results more easily repeatable by any opinion poll server (or any actor or subject for that matter). Therefore the published reputation results of any single opinion poll reputation server are not contingent on the reputations of the reputation server since their outputs can be independently validated. Any reputation server that was noted for publishing different results from other reputation servers would soon be ostracized and its results ignored.

The Voting model should provide the next most trustworthy set of results. The individual reputation scores have similarly been determined by many actors as in the Opinion Poll model, and therefore it should be difficult to skew the results, though it is not impossible. Because the list of actors is not public, it is not possible to independently validate the composite reputation results, nor is it possible for another reputation service to repeat the results. A dishonest reputation server could skew the results by discarding votes it did not like or by changing them. It was widely reported in the US that an electronic voting machine in Fairfax, Virginia "lost" one vote per hundred for a particular candidate [7]. Another recent paper showed how the results of the 2000 US presidential election could have been tipped either way by simply changing one vote per electronic voting machine from democrat to republican or vice versa [4]. Ballot stuffing, by either the reputation server itself, or by a group of colluding actors, could insert false votes to increase or decrease a subject's reputation. We are all familiar with this type of activity in real life, for example, the 2004 Ukrainian presidential election was accused of huge fraud [5] and had to be rerun. Dishonest actors can register false reputations. The consequences can be severe. Research has found that the appearance of a less than 100% positive reputation on e-Bay can seriously effect the price a seller is able to fetch for an item [8]. Actors that know how the summation algorithm of a reputation system works are able to fix the system. For example it is possible to get a high reputation on e-Bay by buying lots of worthless items that cost only a few cents. Once an actor has obtained a falsely high reputation it is possible to commit fraud on unsuspecting customers. For example, a Welsh schoolboy obtained £45,000 by selling non-existent electrical items through e-Bay [6]. So whilst electronic reputation systems that use the voting model, such as the one in e-Bay, should be able to provide trustworthy reputations of its subject, in reality, due to ballot stuffing or other nefarious actions by both the actors and the reputation system itself, this is not always the case, and the operational model has some inherent weaknesses in it.

The Research model provides the next lower level of trustworthy results. Because the reputation results are difficult to arrive at, both from a data collection and computational viewpoint, it is very difficult for actors to reproduce the results without an impossibly large investment of capital and time. Therefore an actor is left with no choice but to either trust or distrust the reputation service that is providing the results. Consequently it will take an appreciable amount of time for the reputation of a research model reputation service to be established, since trust in it will need to evolve over time and as its client base increases. Research model reputation systems might be expected to devote a considerable amount of their resources to ensuring the trustworthiness of the underlying data that they use, since ultimately their reputation will depend upon this. However, once their reputations have been established, they will become a great, if not the greatest, of the assets that the reputation systems possess. For example, it has taken Dun and Bradstreet over 160 years to build their reputation to the level it is now. We might therefore expect it will be many years before these type of electronic reputation servers will become a common feature of the Internet, unless companies with existing high reputations, such as D&B, move into this electronic world.

Finally, the MP model is inherently the least trustworthy and reliable of them all. This is because MP model reputations systems, like research model systems, will decide which data to keep, which data to discard, and how to evaluate it. But unlike research model systems, that actively solicit data from actors, so as to ensure that enough raw data is collected to compute meaningful results, MP model systems simply passively collect whatever information is provided to them, and might even be pre-programmed with particular biases that will skew whatever reputations they compute. The more trustworthy MP type reputation servers will be open to public scrutiny and will publish their algorithms and summaries of the raw data that they have used in their calculations (within the limits of the data protection act). But in general there is no requirement to do this, and therefore the trustworthiness of MP model reputation servers will at best be variable. It is for these reasons that we do not believe that this model is a viable one for reputation systems.

8 Conclusion

We have presented four different operational models for reputation servers, and evaluated the inherent trustworthiness of each model. We have shown that opinion poll model reputation servers, that use publicly accessible reputation information, and that summarize the results in an open and transparent way using publicly available algorithms are inherently the most trustworthy, whilst those reputation systems that do not disclose either the source of their data or the algorithms they use to compute reputations are inherently the least trustworthy.

Practical experience today with electronic reputation systems is limited to just the voting model type, since there are good business and operational reasons why this type of system has evolved. They are co-located with the e-business (usually of type B2C) that is hosting them, and they serve to enhance trust in the e-business itself. The actor lists (the consumers of the e-business) are readily available to the reputation system, making information collection easy, whilst the published reputations of the actors serve to increase the business of the hosting site itself. However, in a B2B world, where virtual organizations are continually being formed and dissolved from different sets of actors, and where there is no mandatory central hosting site such as e-Bay, then we believe that the opinion poll model for reputation systems will be the most effective and trustworthy to deploy. But before this can become a reality, there needs to be a publicly available actor list/directory service from which opinion poll servers can extract the sample of actors to poll for their opinions. In the longer term, research model reputation systems may become established, and this could be hastened by organizations that already enjoy a high reputation in the physical world moving into the electronic one.

References

[1] Paul Resnick , Ko Kuwabara , Richard Zeckhauser , Eric Friedman. “Reputation systems”, Communications of the ACM, v.43 n.12, p.45-48, Dec. 2000

[2] Resnick, P., and Zeckhauser, R. (2002) “Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System”. In The Economics of the Internet and E-Commerce. Michael R. Baye, editor. Volume 11 of Advances in Applied Microeconomics, JAI Press.

[3] P.V. Mockapetris. “Domain names - implementation and specification” RFC 1035. Nov 1987

[4] Anthony Di Franco et al. “Small Vote Manipulations Can Swing Elections”. Communications of the ACM, vol 47, no 10, pp 43-45. Oct 2004

[5] The Times Online. “Opposition overcomes 'total fraud' to claim victory in Ukraine elections”. http://www.timesonline.co.uk/article/0,,3-1369632,00.html (viewed 24 Nov 04)

[6] Guardian Unlimited “Sharks target bargain-hungry surfers” http://www.guardian.co.uk/uk_news/story/0,,1328767,00.html (viewed 24 Nov 04)

[7] Washington Post “Fairfax Judge Orders Logs Of Voting Machines Inspected” http://www.washingtonpost.com/ac2/wp-dyn?pagename=article&contentId=A6291-2003Nov5 (viewed 25 Nov 04)

[8] D. Houser, J. Wooders, “Reputation in Auctions: Theory, and Evidence from eBay”, University of Arizona. Paper under review.

[9] Q He, Z Su. “A new practical secure e-voting scheme”. IFIP SEC’98, Austrian Computer Society, 1998, pp. 196–205. Also available from http://www.cs.huji.ac.il/~ns/Papers/He-Su.ps.gz

[10] OASIS. “UDDI Version 3.0.2”. Oct 2004. Available from http://uddi.org/pubs/uddi_v3. htm

[11] Wahl, M., Howes, T., Kille, S. “Lightweight Directory Access Protocol (v3)”, RFC 2251, Dec. 1997