A presentation over:

A Survey of Trust and Reputation Systems for Online Service Provision

by: Audun JØsang, Roslan Ismail, and Colin Boyd

Presented by Kyle Willett

Abstract:

Trust and reputation systems represent a significant trend in decision support for Internet mediated service provision. The basic idea is to let parties rate each other, for example after the completion of a transaction, and use the aggregated ratings about a given party to derive a trust or reputation score, which can assist other parties in deciding whether or not to transact with that party in the future. A natural side effect is that it also provides an incentive for good behaviour, and therefore tends to have a positive effect on market quality. Reputation systems can be called collaborative sanctioning systems to reflect their collaborative nature, and are related to collaborative filtering systems. Reputation systems are already being used in successful commercial online applications. There is also a rapidly growing literature around trust and reputation systems, but unfortunately this activity is not very coherent. The purpose of this article is to give an overview of existing and proposed systems that can be used to derive measures of trust and reputation for Internet transactions, to analyze the current trends and developments in this area, and to propose a research agenda for trust and reputation systems.

Citation:

Josang, A., R. Ismal, and C. Boyd (2007). A Survey of Trust and Reputation Systems for Online Service Provision, Journal of Decision Support Systems, 43(2):618-644

Outline of Presentation

- Section 1: Introduction
- Section 2: Define trust and reputation
- Section 3: Trust and Reputation relationship as security mechanisms
- Section 4: Collaborative filtering and reputation
- Section 5: Trust Classes
- Section 6: Four categories of reputation and trust semantics
- Section 7: Centralized and distributed architectures for reputation

Outline of Presentation Continued

- Section 8: Reputation computation methods (Day 1 Goal)
- Section 9: Reputation systems used in commercial applications
- Section 10: Description of main problems in reputation systems
- Section 11: Ending discussion

Section 1: Introduction

- Online transactions differ from those of in person transactions because of the inherent asymmetry in the transaction, the seller has all the power so to say.
- The nature of online transactions obscure the traditional metrics used to establish if a brick and mortar store is trustworthy. Example: a brick and mortar store takes time to establish a web site takes very little time to set up.
  - These reasons make it hard to determine rather or not a particular online venue is trustworthy or not and is why this trust issue is receiving so much attention from an academic point of view.

- The authors of this paper wrote it in part because of the rapidly growing interest in this topic and because they felt that the prior overviews used inconsistent terminology.

Section 2: Define trust and reputation

- Two kinds of trust: reliability and decision trust
- Reliability Trust: Trust is the subjective probability by which an individual, A, expects that another individual, B, performs a given action on which its welfare depends.
- Decision Trust: Trust is the extent to which one party is willing to depend on something or somebody in a given situation with a feeling of relative security, even though negative consequences are possible.

- The authors mention that the prior mentioned definitions are not as simple as they seem.
  - For example: Trust in an individual is not necessarily enough to enter into a state of dependence with a person. In other words, that danger might seem to the agent an intolerable risk.
- The authors mentions that only a few papers deal with trust and that in economic circles there are some who reject trust as a computational model.
- Someone by the name of Williamson argues that the notion of trust should be avoided when modeling economic interactions, because it adds nothing new, and that well known notions such as reliability, utility and risk are adequate ad sufficient for that purpose.
- Williamson argues however that personal trust is still important for modeling, and that non-computation models for trust can be meaningful for studying certain relationships.

- Concerning reputation, the authors mention two aspects: trust because of good reputation and trust despite of bad reputation.
- These two statements shed light on the fact that trust is often made with outside information, knowledge about the relationship that is not general know, instincts and feelings, etc.
- Reputation can also be considered as a collective measure of trustworthiness based on referrals from the community.

- Research in Trust and Reputation Systems should have two foci:
  - Finding adequate online substitutes for traditional cues in the physical world and identifying new elements specific to the online applications which are suitable for measurements.
  - Taking advantage of IT and Internet to create efficient systems for collecting information and deriving measurements of trust and reputation in order to aid decision making and improve online markets.

- These simple principles invite rigorous research in order to answer some fundamental questions: What information elements are most suitable for deriving measures of trust and reputation in a given application? How can these information elements be captured and collected? What are the best principles for designing such systems from a theoretic and from a usability point of view? Can they be made resistant to attacks of manipulation by strategic agents? How should users include the information provided by such systems into their decision process? What role can these systems play in the business model of commercial companies? Do these systems truly improve the quality of online trade and interactions? These are important questions that need good answers in order to determine the potential for trust and reputation systems in online environments.

- According to a cited reference in the paper, Resnick, a reputation system must have the following:
  1. Entities must be long lived, so that with every interaction there is always an expectation of future interactions.
  2. Ratings about current interactions are captured and distributed.
  3. Ratings about past interactions must guide decisions about current interactions.

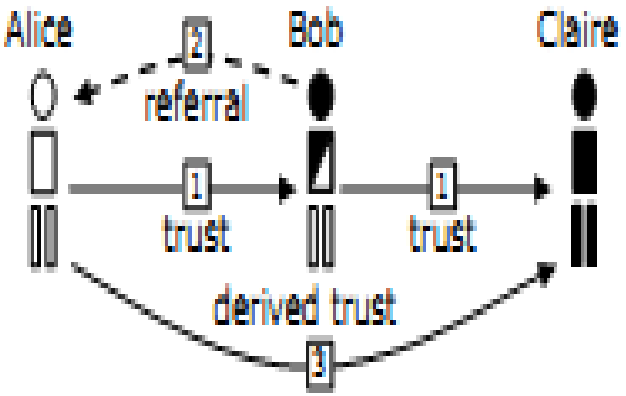

Fig. 1. Trust transitivity principle

Example of how trust is derived. (Fig. 1 from paper)

Section 3: Trust and Reputation relationship as security mechanisms

- In a general sense, the purpose of security mechanisms is to provide protection against malicious parties.
- In many situations we have to protect ourselves from those who offer resources so that the problem in fact is reversed. Information providers can for example act deceitfully by providing false or misleading information, and traditional security mechanisms are unable to protect against this type of threat. Trust and reputation systems on the other hand can provide protection against such threats.

- To summaries this section the author basically says that a computer system that appears to have robust security appears more trust worthy to the user. Listing known security vulnerabilities and using encryption techniques make the system appear to me more trust worthy.

Section 4: Collaborative filtering and reputation

- Collaborative filtering systems are a mechanism that shares traits with a reputation system but they are different at the same time.
- Collaborative filtering systems (henceforth CF) are a mechanism that attempts to take into consideration that different people have different tastes.

- If two separate people rate two items similarly then they are called neighbours in CF terminology.
- This new fact can be used to recommend to one something that the other liked, a technique called a recommender system.
- This takes the opposite assumption of reputation systems which assume that all people will judge the same performance or transaction consistently.

- The example provided by the article is that in CF systems users might rate a video or music file differently based on tastes but one containing a virus would be universally rated poorly.
- Another caveat about CF vs reputation systems is that CF systems assume an optimistic world view and reputation systems assume a pessimistic world view.
- In specifics, CF systems assume all participants are trustworthy and sincere, meaning that all participants report their genuine opinion.
- Conversely, reputation system assume that participants will try to misrepresent the quality of services in order to make more profit and will lie to achieve said goals.
- This duel opposing nature of these type systems can make it very advantageous to combine them as will be explored in the study of Amazon in section 9 which does this to some extent.

Section 5: Trust Classes

- Types of Trust classes:
  - Provision
  - Access
  - Delegation
  - Identity
  - Context
- Paper mentions them in order to get specific about trust semantics.
- Paper focuses on provision trust so it is emphasized.

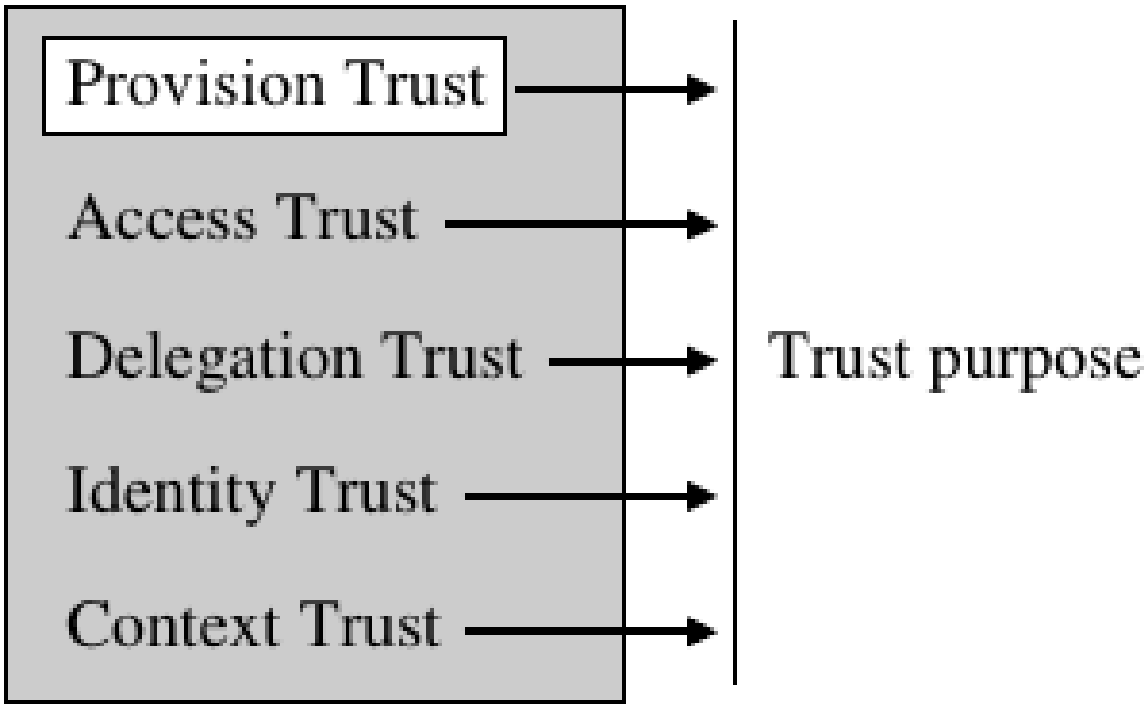

Fig. 2. Trust Classes (Grandison & Sloman 2000)

- Provision trust describes the relying party's trust in a service or resource provider. It is relevant when the relying party is a user seeking protection from malicious or unreliable service providers. I extrapolated from the paper that this is the type of trust that would be studied in business through subjects like contract law.
- Access trust describes trust in principals for the purpose of accessing resources owned by or under the responsibility of the relying party. This relates to the access control paradigm which is a central element in computer security.
- Delegation trust describes trust in an agent (the delegate) that acts and makes decision on behalf of the relying party.
- Identity trust describes the belief that an agent identity is as claimed. Identity trust systems have been discussed mostly in the information security community. An example mentioned in the paper is PGP encryption technology.
- Context trust describes the extent to which the relying party believes that the necessary systems and institutions are in place in order to support the transaction and provide a safety net in case something should go wrong. Factors for this type of trust can for example be critical infrastructures, insurance, legal system, law enforcement and stability of society in general.

- Trust purpose is a concept that can be used to express any trust class mentioned above. It defines the scope of a trust relationship. One example could be “to be a good car mechanic” which would go under provision trust in classification.
- An important relationship of the above mentioned classes is that provision trust can not exist without identity trust. In the absence of identity trust it is only possible to have a baseline provision trust in an agent or entity.

Section 6: Four categories of reputation and trust semantics

- Four types of trust semantics:
  1. Subjective and specific
  2. Subjective and general
  3. Objective and general
  4. Objective and specific

Section 6:

- specific means that it relates to a specific trust aspect such as the ability to deliver in time
- General means that it relates to an average of all aspects.
- A subjective measure means that an agents provides a rating based on subjective judgment
- objective means that the rating was determined by objectively measuring the trusted party on a formal assessment.

Section 6:

Table 1
Classification of trust and reputation measures.

	Specifi c, vector based	**General, synthesised**
Subjective	Survey questionnaires	eBay, voting
Objective	Product tests	Synthesised general score from product tests, D&B rating

Section 6:

- Subjective and specific measures are for example used in survey questionnaires where people are asked to express their opinion over a range of specific issues.
- Subjective and general measures example are eBay's reputation system
- Objective and specific measures are used in technical product tests where the performance or the quality of the product can be objectively measured.
  - Example: Washing machines can be tested according to energy consumption, noise, washing program features etc.
- Objective and general measures can be computed based on a vector of objective and specific measures.
  - In product tests, it is common to derive a general score which can be a weighted average of the score of each characteristic.

Section 6:

- General pitfalls with all subjective measures is that it is difficult to protect against unfair ratings and that slander accusations can occur with subjective trust in an entity.
- General advantages with objective measures is that correctness can be verified by other or automatically generated with an autonomous system.

Section 7: Centralized and distributed architectures for reputation

- Types of architecture:
  1. Centralized reputation systems
  2. Distributed reputation systems

Section 7:

- In centralized reputation systems
  - Information about the performance of a given participant is collected as ratings from other members in the community who have had direct experience with that participant. The central authority (reputation center) that collects all the ratings typically derives a reputation score for every participant, and makes all scores publicly available. Participants can then use each other's scores, for example, when deciding whether or not to transact with a particular party. The idea is that transactions with reputable participants are likely to result in more favorable outcomes than transactions with disreputable participants.

Section 7:

- A and B are two transaction partners that are considering another transaction in the present After each transaction the agents provide ratings about each other's performance and the reputation center collects the ratings and updates each agents score as a function of the received ratings. Updated scores are provided for each agent online for all agents to see and can be used to see rather or not to engage in a transaction with a particular agent.

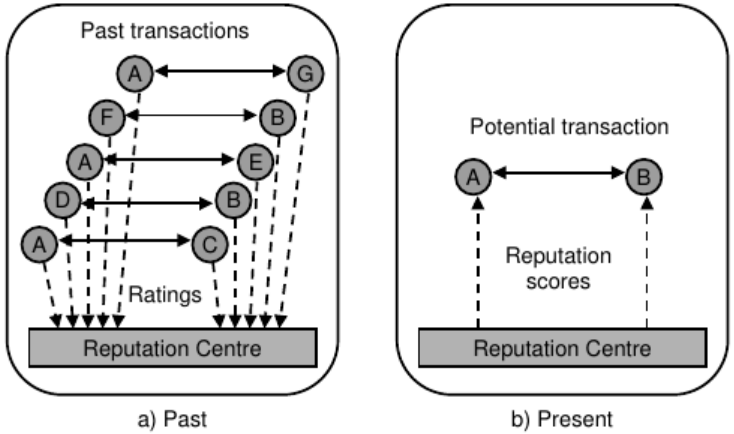

Fig. 3. General framework for a centralised reputation system

Section 7:

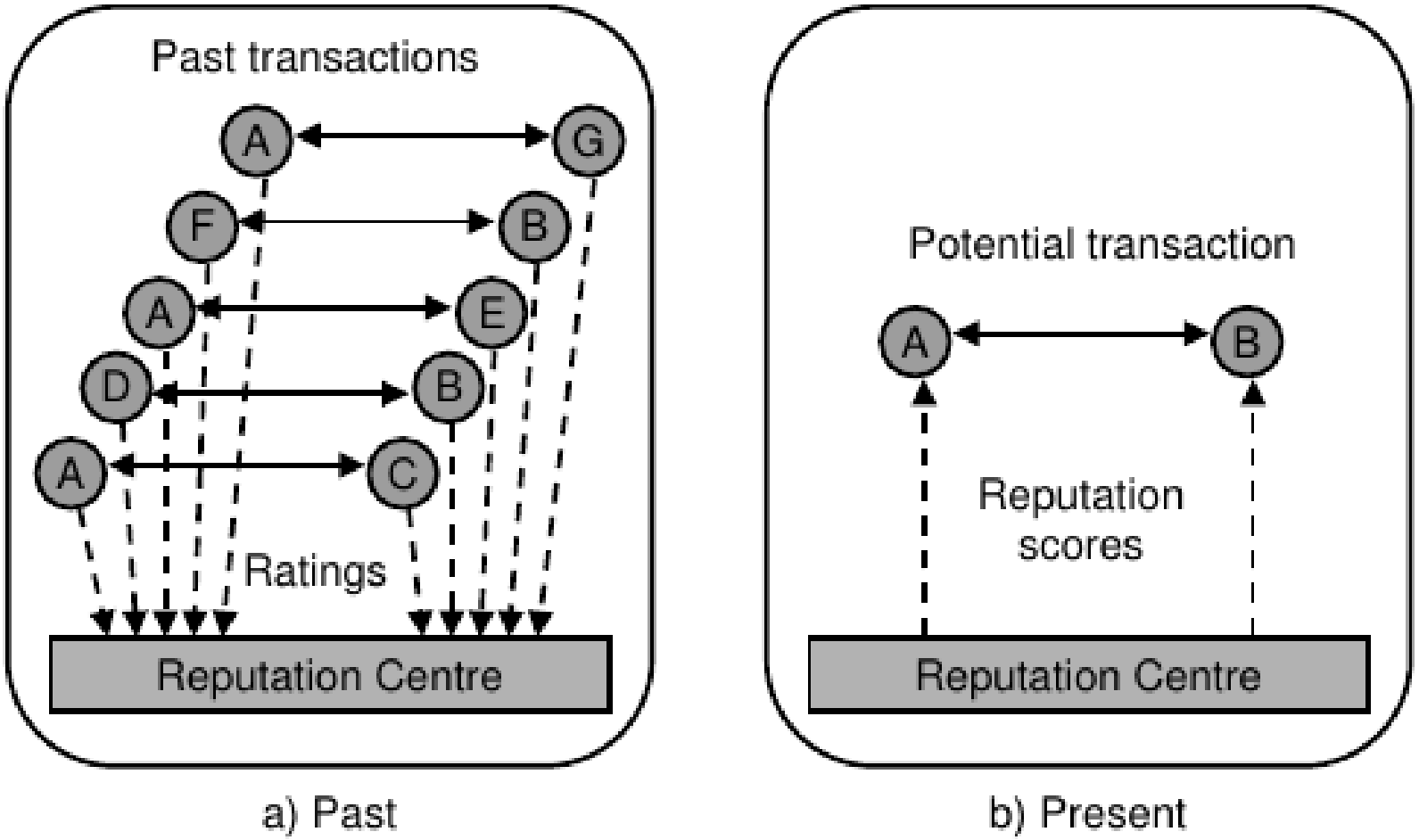

Fig. 3. General framework for a centralised reputation system

- The two big takeaways from centralized reputation systems are as follows:
  1. Centralized communication protocols allow participants to provide ratings about transaction partners to the central authority and to obtain reputation scores of potential transaction partners from the central authority.
  2. A reputation computation engine used by the central authority to derive reputation scores

- There are some environments where a centralized system doesn't work
  - Use a distributed system in this case
- Two approaches:
  - Distributed store of reputation
  - Each agent keeps track of its interactions with each other agent and provides this information upon request to other agents.

Section 7:

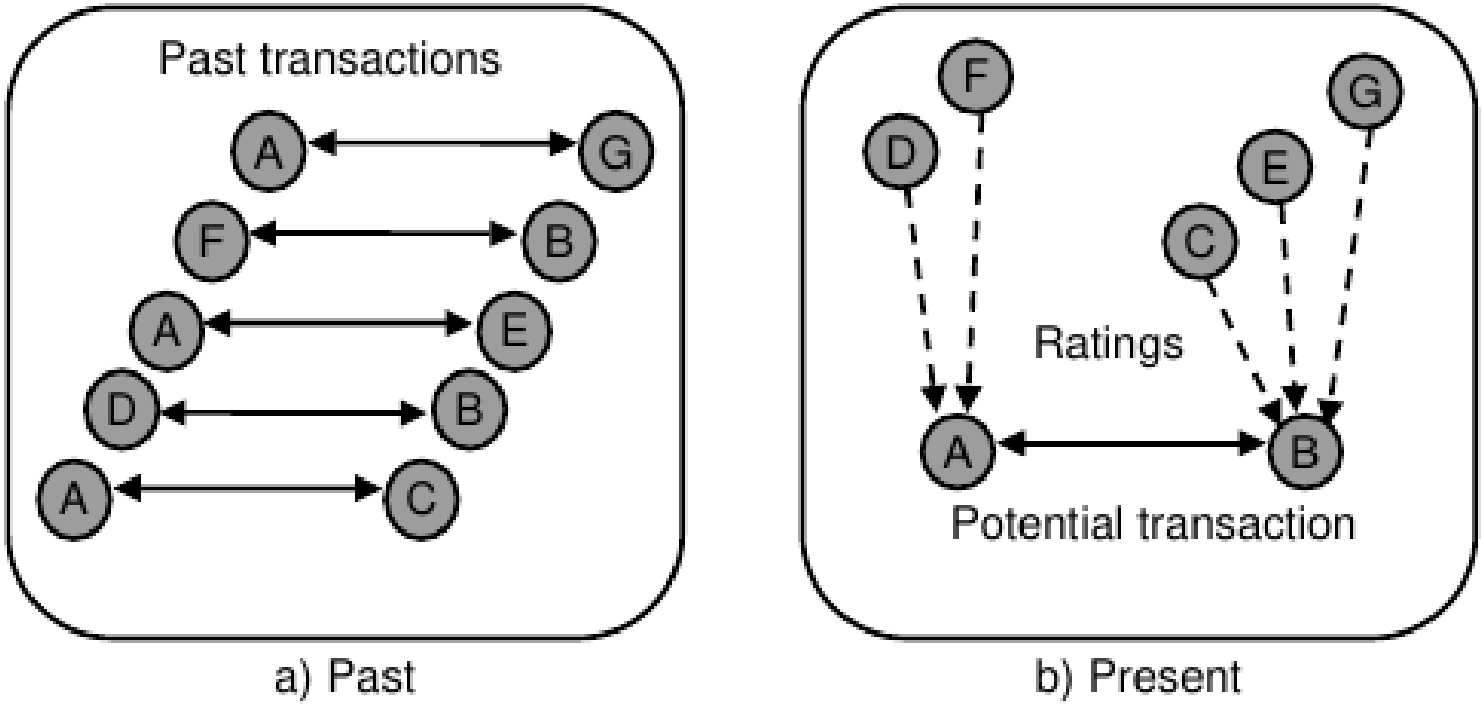

Fig. 4. General framework for a distributed reputation system

- The two big takeaways from distributed systems:
  1. A distributed communication protocol that allows participants to obtain ratings from other members in the community.
  2. A reputation computation method used by each individual agent to derive reputation scores of target parties based on received ratings, and possibly on other information.

- Peer to Peer networks: a great example
  - A brief description:
    - A peer to peer network is a network where every node is both a server and a client called a *servent* in this paper.
  - Two phases of operation:
    1. Search
    2. download
    - Search can be assisted by a centralized solution

Section 7:

- P2P networks have a range of security issues including the ability to be served malicious software and the fact that some entities poison the well so to say when it comes to software like music.
- A reputation system helps to mitigate this content poisoning that has been used by some, the paper notes the music industry as an example here.
- This is why the authors of this paper point out that many authors over the years have pointed out P2P networks as a prime candidate for reputation systems.

- Reputation systems for P2P networks provide two things:
  1. To determine which servents are most reliable at offering the best quality resources
  2. To determine which servents provide the most reliable information with regard to the previous point
- In a distributed environment like this, impossible to collect ranking from everyone
  - A subset of ratings is collected from each servent's "neighborhood"

Section 8: Reputation computation methods

- Reputation systems generally based on public information not private
- Some use private and public
- Private information is considered more reliable

- Types of Computation Methods:
    1. Summation and Average of Ratings
    2. Bayesian Systems
    3. Discrete Trust Models
    4. Belief Models
    5. Fuzzy Models
    6. Flow Models

- Summation and Average methods:
  1. Simplest method: sum the number of positive rankings and negative separately, total score is positive – negative.
     - Used by Ebay's reputation forum
     - Advantage: simple to understand
     - Disadvantage: primitive, provides poor picture
  2. Next advanced method: compute average of all rankings
     - Basis of system used by Amazon
  3. Most advanced method of this type: compute weighted average
     - Weights vary based on raters trustworthiness, age of ranting, and other factors.

- Bayesian Systems
  - Take as input binary ratings and compute something called a beta probability density function
    - The update reputation score is called the a posteriori score
    - Represented by ordered pair (α,β)
      - Alpha and Beta represent positive and negative reputation rakings respectively
  - Pros:
    - Theoretically sound basis for computation
  - Cons:
    - Too complex for average person to understand.

Section 8:

Prior slide case in point: papers description of how complicated the Beta Probability is to understand!

The beta-family of distributions is a continuous family of distribution functions indexed by the two parameters α and β. The beta PDF denoted by $\text{beta}(p \mid \alpha, \beta)$ can be expressed using the gamma function Γ as:

$$\text{beta}(p \mid \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} \quad \text{where } 0 \leq p \leq 1,\ \alpha, \beta > 0 \qquad (1)$$

with the restriction that the probability variable $p \neq 0$ if $\alpha < 1$, and $p \neq 1$ if $\beta < 1$. The probability expectation value of the beta distribution is given by:

$$\text{E}(p) = \alpha/(\alpha + \beta). \qquad (2)$$

- Discrete Trust Models
  - Trustworthiness of agent can be described by 4 states:
    - Very trustworthy, trustworthy, untrustworthy, and very untrustworthy
    - Uses look up tables and an upgrade downgrade action system
  - Advantage:
    - Easier for humans to interface with
  - Disadvantage:
    - Not sound computationally, uses heuristics instead (basically means practical, not optimal)

- Belief Models
  - Definition: Belief theory is a framework related to probability theory, but where the sum of probabilities over all possible outcomes not necessarily add up to 1, and the remaining probability is interpreted as uncertainty.
  - Deals with transitivity
  - Ratings are valid if they result from a transitive trust chain of “sufficient length” according to a defined limit
    - Side note: this method was proposed by the author of the paper.

Section 8:

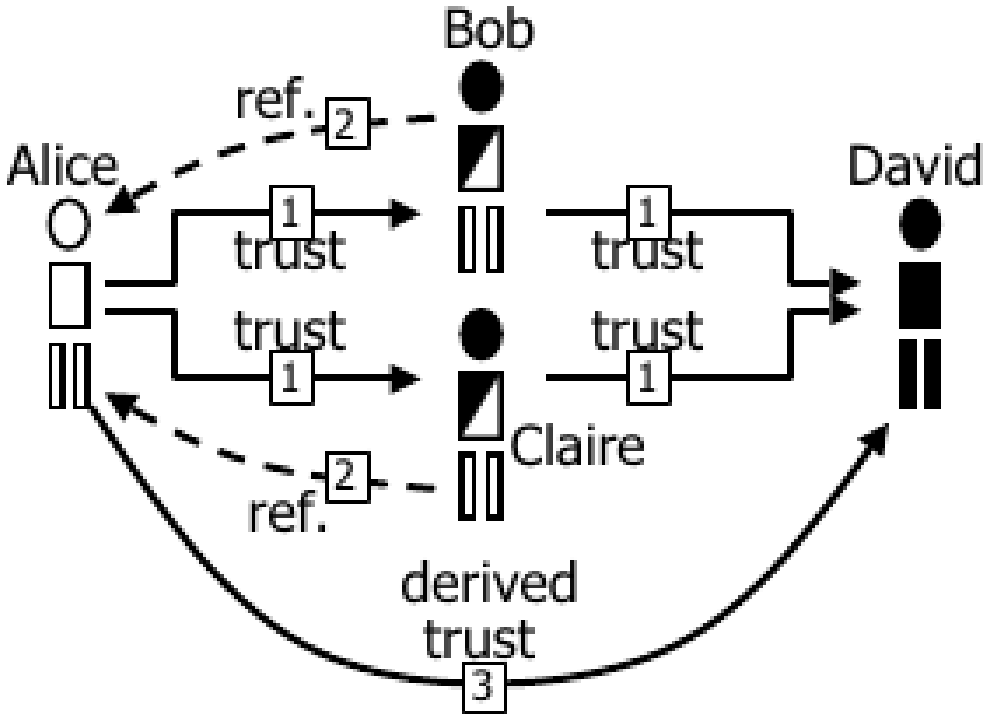

Fig. 6. Deriving trust from parallel transitive chains

- Fuzzy Models:
  - Talked about almost in passing in the paper and not gone into very much detail
    - Called a “linguistically fuzzy concepts”
  - Has a concept called individual reputation that is from private information
  - Concept called social reputation from public information
  - Concept called context dependent reputation that is derived from contextual information

- Flow Models:
  - Category of systems that use computation via transitive interaction through looped or arbitrary long chains
  - Uses constant weights for community reputation
    - Participants can only increase their reputation at the cost of others
    - Used by Google
  - Some exceptions to the constant requirement
    - One example is EigenTrust model, used in P2P networks
    - Repeated multiplication and aggregation of scores until community scores converge to stable values

Section 9: Reputation systems used in commercial applications

- Sections goal: describe the most well known systems in use in the real world
  1. Ebay
  2. Amazon
  3. Google
- Other systems are described in this section but I purposely made the command decision to omit for a few reasons.
- Will also mention relative time problem with paper in my opinion

Section 9:

- Ebay:
- Reputation system is called Feedback Forum
  - Options for feedback are 1, -1, 0
  - Option for text comments
  - Centralized system
- Calculation logic
  - Sum of positive, sum of negative, simple diffrence
  - 6 month, past month, 7 day statistics kept

- Ebay continued:
- Empirical study done by Resnick in 2002
  - Observed rankings are surprisingly overall positive
  - 51.7% of the time buyers provide feedback
  - 60.6% of the time sellers provide feedback
  - <1% is negative, <.5% is neutral, 99% is positive
  - Correlation between buyer/seller rankings
    - Indicative of reciprocity

- Ebay concluded:
  - Criticisms
    - Primitive and misleading system
    - 100 positive and 10 negatives is viewed the same as 90 positive and 0 negatives which the author finds wrong
  - Pros
    - User participation makes users feel good about the system and generally people tend to trust the system
    - ‘Ballot stuffing’ is very limited due to how the market works

- Amazon:
  - Uses reviews that consist of 1 to 5 star rating supplemented with text description
  - Users can vote on review as helpful or not helpful
  - Reviews can be sorted by user as newest first, most helpful, or highest rating
  - Not all information is publicly available concerning system
  - Rating system for reviews and how helpful they have been

- Criticisms:
  - Seemingly extreme criticism by the authors from a variety of angles on how Amazon's system is vulnerable
  - Votes on answers are only tied to a browser's cookie session so a reset of that lets users vote again
  - Indicators of extreme ballot stuffing has been discovered
  - 'Cat fights' occur between reviews to see who will be the top reviewer
  - Authors final note: not a robust scheme

- Google:
  - System is called PageRank, differs from early systems that used binary logic or pattern matching of keywords
  - Ranks page according to how many pages point to it.
    - Side note: authors compare google to a system used by AltaVista a company that has been defunct since 2013.

Section 9:

- Google: mathematical description of how PageRank algorithm works from paper.

Definition 4 *Let P be a set of hyperlinked web pages and let u and v denote web pages in P. Let $N^-(u)$ denote the set of web pages pointing to u and let $N^+(v)$ denote the set of web pages that v points to. Let E be some vector over P corresponding to a source of rank. Then, the PageRank of a web page u is:*

$$R(u) = cE(u) + c \sum_{v \in N^-(u)} \frac{R(v)}{|N^+(v)|} \tag{4}$$

where c is chosen such that $\sum_{u \in P} R(u) = 1$.

In [52] it is recommended that E be chosen such that $\sum_{u \in P} E(u) = 0.15$. The first term $cE(u)$ in Eq.(4) gives rank value based on initial rank. The second term $c \sum_{v \in N^-(u)} \frac{R(v)}{|N^+(v)|}$ gives rank value as a function of hyperlinks pointing at u.

- Google PageRank continued:
  - Google does not elaborate on the private side of things about how their system works but it is purposely designed to be expensive to influence the algorithm
  - PageRank dramatically reduced ‘ballot stuffing’ in the search engine world
    - Previously this was possible by using hidden text and meta data
  - PageRank is transitive in the extreme with some studies showing infinite length hyperlink chains being possible

Section 10: Description of main problems in reputation systems

- Problems and proposed solutions
  1. Low incentive for ratings
  2. Bias towards positive ratings
  3. Unfair ratings
  4. Change of identities
  5. Quality over time
  6. Discrimination
  7. Ballot Box Stuffing

- Low incentive for ratings
  - Most systems handle ratings after the transaction and there is little incentive for each party to rate the other
  - Niceness and fear of retaliation are contributing factors
  - No direct benefit
- Solutions?
  - Authors admit that there is no known solution to this problem and that the two proposed solutions are all financial based, i.e. pay people to rate each other.

Section 10:

- Bias towards positive ratings
  - Study by Resnick and Zeckhauser found that in general .6% and 1.6% from buyers and sellers respectively where negative
    - Seems illogical to think this represents real world
  - Causes:
    1. Fear or reprisals
    2. Law suits
    3. Exchange of courtesies
    4. reciprocity
- Solutions?
  - Anonymous rating mechanism and/or cryptographically secured rating mechanism
    - I personally see this as having a whole new set of problems by itself. Sociological and Psychological experiments show this is a bad idea in my opinion.

Section 10:

- Unfair ratings
  - Authors desire to prevent unfair negative as well as unfair positive reviews
    - Party relying on review can not control the sincerity
- Solutions?
  - Two classes
    1. Endogenous Discounting of Unfair Ratings
       - Give low weight to resumed unfair ratings, assumes that a property determinedable by statistics exists to find unfair ranting.
         - Method proposed that uses collaborative filtering of rater groups according to ratings of same object
    2. Exogenous Discounting of Unfair Ratings
       - System where external reputation of the rater is used to weight ratings. Assumes poorly rated raters give unfair ratings
       - Kind of needs private information
       - Several proposed solutions exist using different statistical models

- Change of identities
  - Reputation systems are based on axiom that identities are long lived
  - Authors propose hypothetical where a party has suffered significant loss of reputation
    - It is in agents best interest to cut ties with identity and start fresh but best interests of community are diametrically opposed to this
- Solutions?
  - System called ZMM scheme designed, used in KasBah agents MIT multi-agent system
    - System penalizes new users for changing identities
    - Since it penalizes all newcomers it is unfair because there are good and bad newcomers

- Quality over time
  - Reputation tends to move towards an equilibrium
  - Unsure if authors believe this to be good or bad
  - Authors want to discount the past
- Solutions?
  - Methods of discounting the past include:
    1. Forgetting factor
    2. Aging factor
    3. Fading factor
    4. Longevity factor
    5. Reinforcement learning

Section 10:

- Discrimination
  - Issues related to similar conditions in 10.3 section about unfair ratings

- Ballot Box Stuffing
  - More than legitimate number of ratings
  - Closely related to unfair ratings problem
  - Online version of this problem can have negative vote issue (authors don’t elaborate on point)
- Solutions?
  - Best solution is to only allow registered users to vote

Section 11: Ending discussion

- Metrics of a quality and sound reputation computation engine:
  1. Accuracy of long term performance
  2. Weighing towards current behavior
  3. Robustness against attacks
  4. Smoothness

- Hard problems:
    1. Robustness against attacks
    2. Unfair ratings
    3. Ballot stuffing
- Why they are hard?
  - Based on subjectivity
  - Not one solution fits all
    - Will always be minutia about the situation, pros and cons
  - Perhaps most frustrating: reason to provide rating in first place
    - Seemingly no rational reason to provide a rating
  - Doubts about reliability

- Despite these problems reputation systems are still relied upon, why?
  - System doesn't need to be robust- value is elsewhere
  - Participants just need to 'believe' it works
  - Can be considered beneficial if it provides two things:
    1. Stoning – swiftly reacts against bad behavior
    2. Label initiation dues – imposing of cost to get established

Section 11:

- Conclusion concluded
  - Authors speculate that value of reputation system might be in its aspect as a social network, it just attracts people to a particular website
  - Robustness may not even matter, might even be desirable from a business prospective!
  - Authors extrapolate that growing amount of research and literature on topic of reputation systems means it is important topic going forwards
- Authors finish conclusion by point out that this is a field still controlled by pioneers
  - Meaning systems are typically designed from scratch – not building upon prior work
  - No coherence from academic community
  - Authors hope to see more collaboration in the future and the field ‘settle down’ so to say as well as more commercial implementations