Trust, Reputation and Fairness in Online Auctions

Adam Wierzbicki

21/11/2006

Polish-Japanese Institute of Information Technology

Plan of the talk

- Background
  - Online auctions
  - Trust management
  - Reputation systems
- An example of a new reputation algorithm and its evaluation
- A second example of negative discrimination
- Changing existing definitions of trust management and trust

Online Auctions and the Internet economy

- E-commerce market growing at a rate of about 25%
- Online auctions generate 15% of all online sales
- eBay has 56 million active users
  - Annual transactions surpass $23 billion
  - About 1 million transactions daily
- Online auctions account for over 40% Internet fraud
  - This is the most popular type of Internet fraud
  - Annual loss about $14 million

The Trust Management Problem

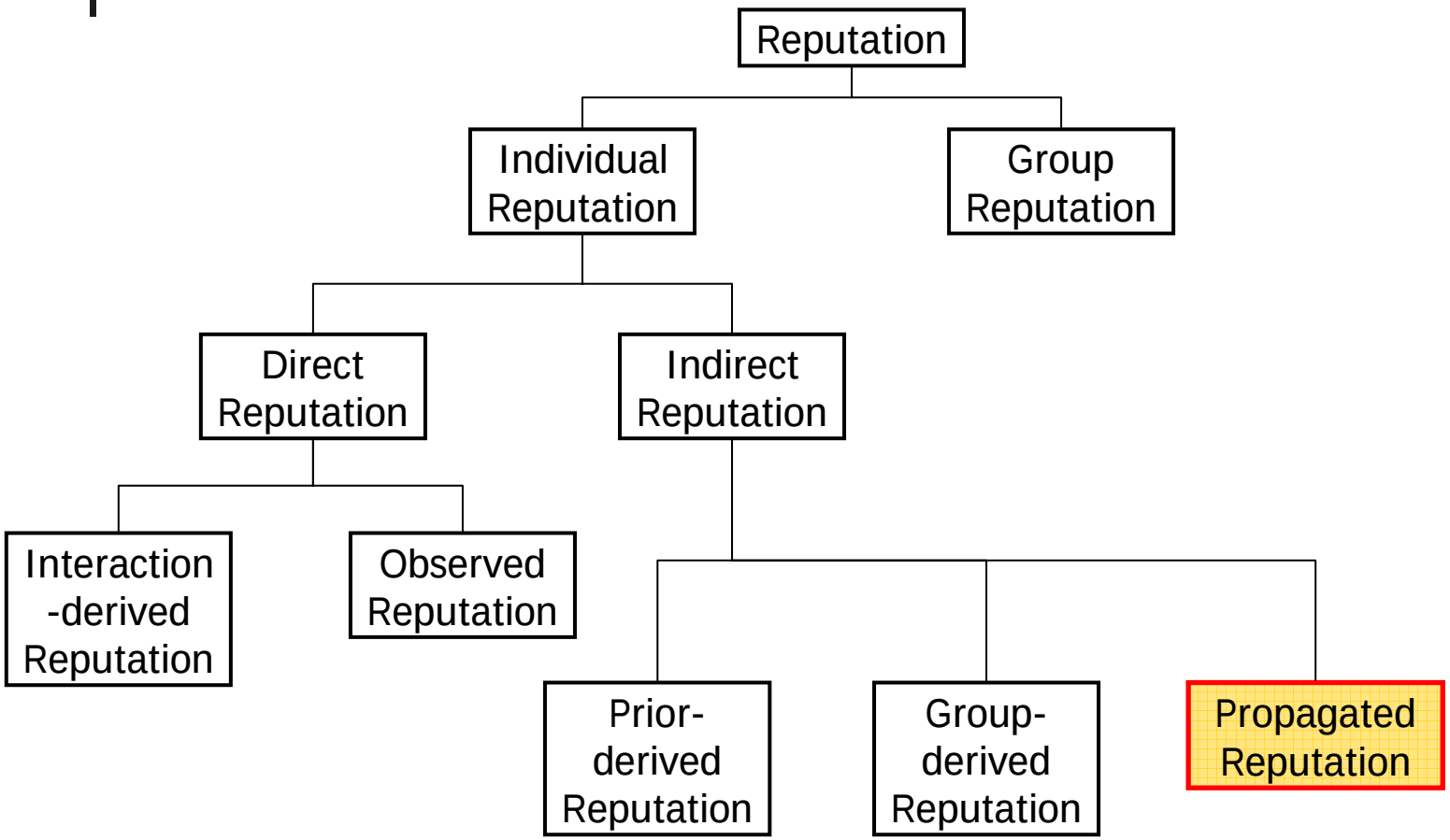
- How can an e-commerce trader know what to expect about the behavior of his partner in a transaction, when he has no record of previous encounters?

- More generally:
  How can an agent determine what to expect regarding the behavior of other agents under uncertainty?

A solution: reputation systems

- *Reputation*: perception that an agent creates through past actions about its intentions and norms
- A reputation system is a trust management system that uses reputation
- Examples of reputation systems in e-commerce:
  - eBay
  - Allegro
  - Amazon

Types of reputation systems



- Indirect, propagated reputation: majority principle

Reputation in online auctions

- Based on feedback
  - After a transaction, traders post feedbacks
  - Feedback can be positive, negative or neutral
- Risk assymetry of buyers and sellers
  - Sellers have little risk, since they require advance payment
  - Buyers have high risk, since they may not receive goods or receive goods of poor quality
  - Result: buyer reputation is less relevant to sellers than seller reputation is to buyers
- Experience assymetry of buyers and sellers
  - Sellers are usually longer in business and participate in more transactions than buyers
  - Thus, they are better known than buyers

Weaknesses of reputation systems

- Vulnerable to:
  - First-time cheating
  - Coalition attacks
  - Discrimination attacks
- Weakness of feedback systems
  - Many users do not post negative feedbacks
  - Out of inexperience or fear of retaliation
    - This is especially relevant for buyers
  - Frequently, all feedback is missing

Simple reputation algorithms

- The simplest algorithms are used in practice (eBay)
- Reputation is a proportion of positive feedbacks to all feedbacks about an agent
  - Drawbacks: is not context-dependent
  - Is vulnerable to all known reputation weaknesses
- Frequently, reputation is not computed. Users are provided with a list of all or most recent feedbacks about another user.
  - Reputation is evaluated implicitly by a human

Evaluation of reputation systems

- Paradigm of Prisoner's Dilemma
  - In this game, the sum of payoffs is highest when no one cheats
- Utilitarian concept: a reputation system is good if a sum of all agents' utilities is highest
  - This type of evaluation ignores the possibilities of cheating present in reputation systems
- **Our claim: a reputation (trust management) system should promote fair agent behavior!**

Plan of the talk

- Background
  - Online auctions
  - Trust management
  - Reputation systems
- An example of a new reputation algorithm and its evaluation
- A second example of negative discrimination
- Changing existing definitions of trust management and trust

A new reputation algorithm with implicit feedbacks

- The concept of implicit feedbacks
  - A user may send no feedback but mean a negative feedback
  - In Allegro, over 98% received feedbacks are positive
- Do implicit (negative) feedbacks exist?
  - About 30% of transactions have no feedback
  - Are these implicit feedbacks?

Existence of implicit feedback

- Majority strategy: if a user usually sends feedback, the omission of feedback could be an implicit feedback
  - There are very few users who continue to send few feedbacks. Thus, for experienced users, the lack of feedback is an implicit (negative) feedback
- Ochini (cosine) coefficient strategy: calculate a similarity of a user's feedback profile to the profile of a hypothetical user who never gives feedback
  - Results: less than 10% users are similar
  - Over 90% users are dissimilar. These users provide implicit (negative) feedbacks

A reputation algorithm that uses implicit feedback

- A simple extension of eBay's algorithm
- For every user, let:
  - n : number of transactions
  - m: number of feedbacks
  - m^+: number of positive feedbacks
  - $m^- = n - m$: number of missing feedbacks
  - $\rho = m^+ / (\alpha m^- + m)$: **user's reputation**
- If α is zero, then we have eBay's algorithm
- In practice, need to have an exponential smoothing of ρ

Evaluation of the reputation algorithm – the simulator

- Reputation system simulated faithfully
  - But, only positive or negative reports
- User behavior is realistic
  - Users take into account reputation of partners
  - Users decide whether to report, depending on type of report
  - Users can cheat in reports, as well as in transactions
  - Users can use transaction strategies that depend on known history
- Auction has been simplified
  - Random choice of partner depending on reputation, from a small set
  - Transaction simulated by a game: PD or zero-sum

Simulation parameters – 1

- Each simulated agent is characterized by:
  - His game strategy and its parameters
    - Tit-for-tat with reputation threshold, usually 0.5
    - Random cheating, usually 0.6
  - Probabilities of cheating in reports
- The reputation system can use various algorithms:
  - Simple ratio of positive reports to all reports
  - Algorithm with implicit feedbacks, for various values of α

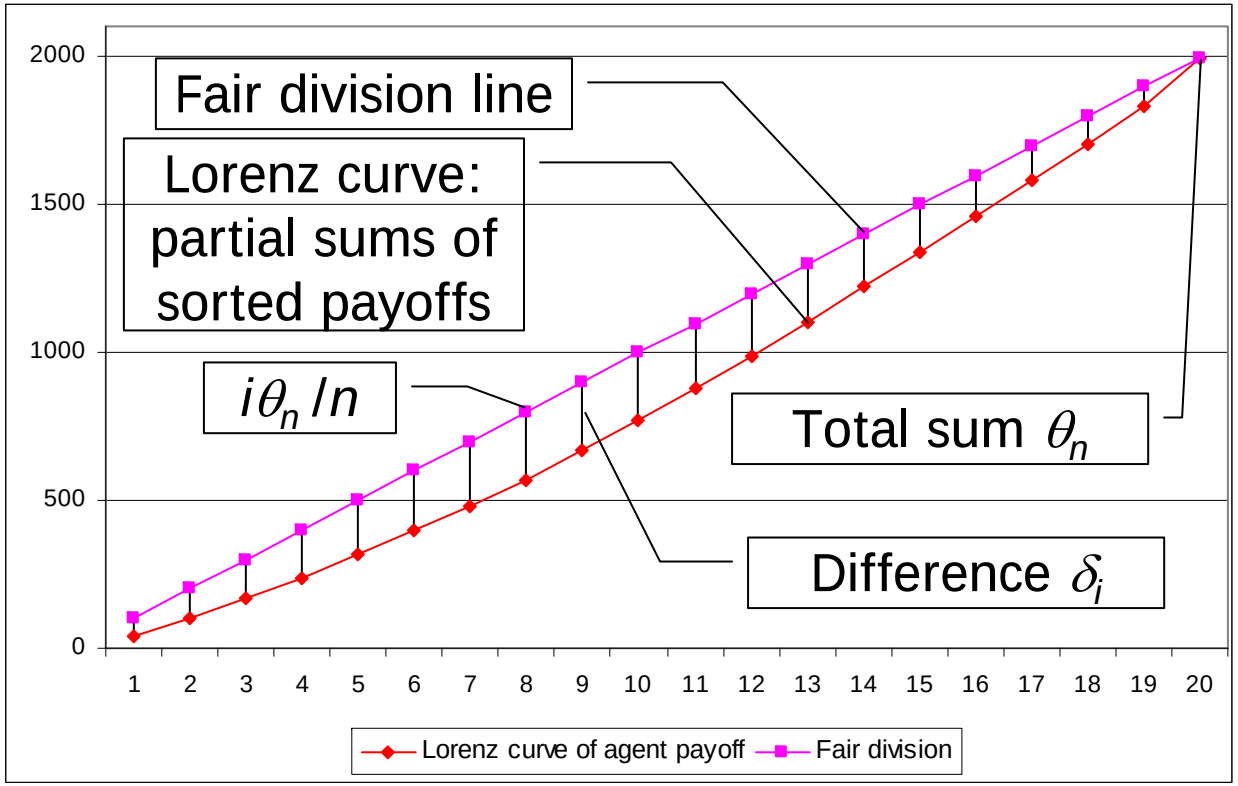
Simulation parameters – 2

- Agent set: 300 agents divided into two groups:
  - „Good agents”: 66% of all agents
  - „Bad agents”: the rest
  - Two reporting behaviors:
    - Perfect feedback
    - Poor feedback: good reports with probability 0.66, bad reports with probability 0.05
- 10 simulation runs
  - Present averages and 95% confidence intervals

Evaluation of the reputation algorithm – the criteria

- Use fairness to evaluate the algorithms!
- Criteria:
  - Average payoff of „good” agent
  - Average payoff of „bad” agent
  - Fairness of payoffs of „good” agents
    - Evaluated using the Gini coefficient

The Gini coefficient



- Gini = $\dfrac{\Sigma \delta_i}{2\theta_n}$

Comparison of reputation algorithms

	Good agents' avg. payoff	Bad agents' avg. payoff	Gini of good agents' payoffs	95% confidence intervals Gini	Payoff
Perfect reports	101,76	22,66	0,70	0,63-0,76	100-103,5
Poor reports, alpha=0	96,45	54,20	0,51	0,45-0,58	93,6-99,3
Poor reports, alpha=0.05	99,08	23,03	0,75	0,66-0,85	96,1-102
Poor reports, alpha=0.1	100,41	23,52	0,67	0,58-0,76	98,8-101,9
Poor reports, alpha=0.2	100,74	22,64	0,74	0,66-0,82	98-103,4
Poor reports, alpha=0.3	99,72	22,60	0,83	0,73-0,92	97,7-101,7

- Conclusion: Should take fairness into account!

Plan of the talk

- Background
  - Online auctions
  - Trust management
  - Reputation systems
- An example of a new reputation algorithm and its evaluation
- A second example of negative discrimination
- Changing existing definitions of trust management and trust

A second experiment – the power of negative discrimination

- Aim of experiment: demonstrate effect of abusing reputation on fairness
- Negative discrimination: cheating a minority of selected agents
- Real-world parallel: newcomers enter a market of hostile agents
  - Old agents always cooperate with each other
  - Old agents always cheat newcomers
  - Newcomers are a minority

Experiment scenarios

- 10 agents
- Agents 1, 2, 3 are newcomers
  - Play Reputation Tit-for-tat with everyone
- Rest of agents are old agents
  - Cheat agents 1, 2, 3, cooperate with each other – scenario 1 (**discrimination**)
  - Play Reputation Tit-for-tat – scenario 2 (**cooperation**)
- Perfect reporting in all scenarios
- 200 transactions, 10 iterations
  - Maximum outcome for all in PD: 1200

Results – payoffs

- Criterion: total payoff of old agents

	Discrimination	Cooperation
Average	**867**	**847**
Conf. Interval 95%	822-910	838-856

- Conclusion: it pays off to cheat
  - Usually, the payoff is much higher than the average

Results – reputation

- Reputation of old and new agents under discrimination scenario
  - In cooperation scenario, all agents have reputation of 100 (maximum)

	New agents	Old agents
Average	**30**	**75**
Conf. Interval 95%	24-34	72-78

- Conclusion: old agents have higher reputation
  - Because they cheat a minority

Conclusion from second experiment

- Fairness is a big issue in reputation systems
- Better reputation algorithms are needed
  - A solution for negative discrimination: controlled anonymity (Dellarocas)
- What to do about the iterated PD?
  - Here, the total payoff drops
  - But is that enough to demonstrate fairness problems?
  - Would it be better to use a zero-sum game?

Plan of the talk

- Background
  - Online auctions
  - Trust management
  - Reputation systems
- An example of a new reputation algorithm and its evaluation
- A second example of negative discrimination
- Changing existing definitions of trust management and trust

Conclusion: reconsider definition of trust

- Trust definition of Gambetta, 1988: Trust is the subjective probability by which an agent A expects another agent B to perform a particular action, on which A's welfare depends.
- New trust definition:
  Trust is the subjective probability by which an agent A expects another agent B to behave fairly.

Conclusion: reconsider definition of trust management

- Trust management problem:
  - How can an agent determine whether another agent will behave fairly under uncertainty?
- How to promote fairness in a community?
  - ...in a global, electronic community?
    - e-auctions, e-commerce, e-games, P2P apps
  - ...of selfish, deceitful peers?
  - ...in an environment without trusted central control?

Definitions of fairness

- How is fairness defined?
  - Very generally: obeying a set of predefined rules (contract, agreement)
  - Less generally: obeying a set of social norms
    - Trust as a normative notion: Elgesem, 2006
    - Our contribution:
      - practical use of fairness in evaluation of trust management algorithms/systems
      - Definition of trust in electronic systems, not just social (or AI) systems

Why is risk not enough?

- Previous definitions of trust presupposed a situation of uncertainty (risk)
- The new definition does also, but:
  - Consider a transaction with an outcome dependent on random factors
  - Consider a transaction where the partners may be dissatisfied even if both are fair (for example, poor bidding strategies)
- These are examples of situations that require consideration of fairness to evaluate trust

Related work

- D. Elgesem, Normative Structures in Trust Management, Trust Management (iTrust 2006), Springer, LNCS 3986, 2006
- C. Dellarocas, Immunizing Online Reputation Reporting Systems Against Unfair Ratings and Discriminatory Behavior, In Proc. of the 2nd ACM Conference on Electronic Commerce, Minneapolis, MN, USA, October 17-20, 2000
- M. Morzy and A. Wierzbicki, The Sound of Silence: Mining Implicit Feedbacks to Compute Reputation, to appear in Proc. 2nd international Workshop on Internet & Network Economics (WINE'06), Springer, LNCS, 2006