Relating Reputation and Money in On-line Markets

Ashwin Swaminathan[♣], Renan Cattelan[♢], Ydo Wexler[♡],
Cherian Varkey Mathew[♠], and Darko Kirovski[♡]

[♣] University of Maryland, College Park, MD, USA
[♢] Universidade de São Paulo, São Carlos, SP, Brazil
[♠] Indian Institute of Technology, Kanpur, India
[♡] Microsoft Research, Redmond, WA, USA

Contact: darkok@microsoft.com

Technical Report MSR-TR-2008-023
February 2008

Microsoft Research
One Microsoft Way Redmond, WA 98052, USA
http://research.microsoft.com

Relating Reputation and Money in On-line Markets

ABSTRACT

Reputation in on-line economic systems is typically quantified using counters that specify positive and negative feedback from past transactions and/or some form of transaction network analysis that aims to quantify the likelihood that a network user will commit a fraudulent transaction. These approaches can be deceiving to honest users from numerous perspectives. We take a radically different approach with a goal to guarantee to a buyer that a seller cannot disappear from the system with profit following a set of transactions that total a certain monetary limit. Even in the case of stolen identity, an adversary cannot produce illegal profit unless a buyer decides to pay over the suggested sales limit.

1. INTRODUCTION

In the recent decade C2C[1] markets have flourished on the Web via numerous economic opportunities with eBay as the iconic example [1]. Significant amount of fraud in such markets occurs as consumers build up their reputation either via several fabricated or small-cost transactions, to target a final fraudulent high-cost transaction after which they disappear from the market. Detailed statistics about the prevalence of such transactions are not available publicly. The market leader, eBay, claims that one in thousand listed products ends up being a bait for a fraudulent transaction [2]. At first, this figure seems encouraging, however, it relates to "listed" products only, not sold. Also there is no quantifiable data on the cost of fraud on popular on-line marketplaces such as eBay and Amazon.com. In 2006, consumers reported to the Internet Crime Complaint Center (IC^3) approximately US$90M of losses to auction-related fraud in the US only [3]. We speculate that the problem is vastly underreported and important, and pursue a novel, both logistically and technically, consumer reputation system for on-line markets.

If we assume that fraud is a sudden shift in behavior by a seemingly "honest" economic entity, then we can conclude that by definition, fraud prediction is an ill-defined problem. To address this issue, we define seller reputation as a monetary value that a buyer should feel comfortable paying, knowing that by committing fraud the seller still cannot make profit from its existence in the market. The objective is to quantify reputation using a deterministic economic value as opposed to a probabilistic predictor. Even in the case of stolen identity, an adversary cannot produce illegal profit unless a buyer decides to pay over the suggested sales limit. The sales limit of an individual seller is built using a record of her transaction fees, verifiable types of transaction costs (insurance, arbitration, shipping, etc.), and deposits. To further strengthen buyer's perspective, we enable each seller to establish a reimbursement fund used as a guarantee that defrauded buyers will get fully or partially reimbursed. Here, we present the novel deterministic reputation system, we outline a strategy for managing sales limits that maximizes selling power in the market, and propose a probabilistic strategy for risk assessment that aims at helping buyers estimate the risk of paying for a product or service over the sales limit. The latter effort follows more closely the related work in fraud-prediction.

As a simple motivational example, we remind the reader that in off-line markets reputation of small retailers is typically gained by investments in the location and decor of the retail store as well as with years of conducting trustworthy business. Robust reputation is valuable to merchants as it can enable them to select a desired spot on the volume vs. pricing curve for marketed products/services. Equivalently, in our system a seller must invest funds either in the form of a deposit or transaction fees to offset her maximum selling power at a given moment. Thus, we speculate that the system is acceptable from seller's perspective while it spurs confidence with prospective buyers.

1.1 Contributions

Existing reputation systems reviewed in the Appendix, model reputation using a probabilistic system with an objective of providing side information to help users predict malicious transactions. We depart from this traditional method for quantifying reputation, and aim at pricing reputation. From buyer's viewpoint, we offer three tools to strengthen buyer's confidence in (not) participating in an on-line transaction: *i*) a reimbursement fund, *ii*) a monetary limit on pricing that guarantees that even if the seller committed fraud on all pending transactions, she would still not walk out of the economic ecosystem with profit, and *iii*) a simple user interface to a tool that presents to the user the price vs. the probability that the seller commits a fraudulent transaction. Clearly, proposition *i*) is not novel as escrows and/or investment insurance have been part of trade markets for centuries. We build our proposal on top of such an insurance system with propositions *ii*) and *iii*). To the best of our knowledge this is the first reputation system tailored to on-line markets that exhibits such features.

[1] Consumer-to-consumer (C2C) e-commerce involves electronically facilitated transactions between consumers via a third party.

2. THE ECONOMIC MODEL

In this section, we formally introduce a market network and define a two-party transaction as an economic function in the system. Let $\mathbb{C} = \{c_1, \ldots, c_N\}$ be a cardinality-N set of nodes in a graph $\mathcal{G}$, where each node c_i models a distinct consumer. For now we describe a transaction as an exchange of economic value between a buyer and a seller; we define a simple transaction model later, in Subsection 3.2. In the case of a C2C market the buyer pays using a cash equivalent for a product or service offered by the seller. Any node in the graph can be a buyer or a seller in a transaction.

We formally define a **committed transaction** $t(c_i, c_j)$ between a buyer c_i and a seller c_j as a weighted directed edge $c_i \rightarrow c_j$ where the weight $w(c_i, c_j) \equiv w_{ij}$ is a real non-zero scalar such that:

- $w_{ij} > 0$: transaction was executed at the satisfaction of both the seller and the buyer with w_{ij} equal to the transaction costs.
- $w_{ij} < 0$: transaction was fraudulent with $-w_{ij}$ proportional to the cash equivalent paid by the buyer. The buyer suffered financial loss.

We denote as $\mathbb{T}$ and $\mathbb{W}$ the sets of all edges and their weights in the market graph $\mathcal{G}$.

We model **pending transactions** in the network as a set $\mathbb{P} = \{\mathbf{p}_1, \ldots, \mathbf{p}_N\}$ of arrays of values available for sale at corresponding nodes. A transaction is pending until its buyer and seller do not reach a closure on their satisfaction with the transaction; then the transaction becomes committed. An array $\mathbf{p}_i = \{p_1, \ldots, p_{L_i}\}$ is a list of L_i values that seller c_i is currently selling. We allow that products' prices form using an arbitrary negotiation mechanism. Each individual price, p_k, is formed as an asking price (if the seller does not have a buyer yet) or as a winning bid (in case there exists an arbitrary auctioning mechanism). In order for a buyer to learn about a specific product sold by any seller, we allow arbitrary marketing strategies in our model.

Finally, note that our model does not consider the reputation of winning bidders, i.e., nodes with the highest offer for a specific item sold by a seller. Thus, it does not link specific nodes with products itemized in $\mathbf{p}_i$. As opposed to related work where node connectivity is used to construct a reputation model (e.g., [12]), in our system the reputation of current bidders does not affect the reputation score of the seller, thus, this limitation of our model is appropriate. Systems that target detection of shill bidding typically rely on this type of data. Although this is not the goal of our paper, we still mention that this data can be tracked by modeling pending transactions the same way as committed transactions with the necessary relinking to model outbidding.

The considered economic network model includes the directed weighted graph $\mathcal{G}(\mathbb{C}, \mathbb{P}, \mathbb{T}, \mathbb{W})$, where pending transactions are still negotiated. Based upon this model, in this paper we construct the proposed reputation system.

2.1 Model Accuracy

At present time, most popular online markets such as eBay and Amazon.com have built large economic ecosystems that could be used to quantify certain parameters in the model. The first anticipation is that N tends to be rather large for these systems. For instance, eBay recorded around 82 million active users in 2006; this number has been increasing by around 15% every year [4].

Linking our model to an existing marketplace network is a difficult task from several perspectives. First, the number of transactions on marketplaces such as eBay or Amazon.com, is growing at a faster rate than a modest academic crawler could possibly browse. Second, fair random sampling of exceptionally large graphs is a problem of well-known difficulty [19]. Since we do not base our core primitives for building user reputation on network features such as average fan-in, fan-out, etc., we decided to constrain our marketplace snapshot in Section 5 to address transactions with negative feedback, *not* to determine our model parameters so to accurately mimic a typical marketplace network.

2.2 Buyer's Feedback

Typically both participants in a transaction provide feedback to each other. The feedback score, i.e., reputation, is recorded for public viewing and typically summarized in the form of positive and negative *points*. Although several reputation scores systems have been proposed [12, 13], they are not fool-proof – buyers can still be easily deceived by fraudulent sellers who have a very good reputation score. Trivial approaches to building up a positive transaction history include: fabricated transactions with friends or non-existent consumers (e.g., established using stolen identities) or, in the case which is the most difficult to prevent using probabilistic recommendation systems, relatively long-term honest sales behavior until a "major" fraudulent transaction fetches significant profits for the adversary.

Here it is important to stress that our model *does not* address the negative feedback that a buyer could receive. Such feedback is typically posted for failure to pay an item that the buyer won in an auction. Although the seller suffers financial loss due to delay of sale, this type of transaction outcome is still not considered fraudulent. For example, the Amazon.com Marketplace does not display buyer feedback on its system, thus lets sellers treat all buyers equally [5]. To that extent, we note that reputation of a buyer as a reliable payee could be handled efficiently using existing reputation systems and we chose not to address this issue.

2.3 On-Line Dispute Resolution Systems

Our reputation system complements existing on-line dispute resolution (ODR) systems such as SquareTrade [6, 7], in the extent that it aims at preventing/handling fraudulent transactions that SquareTrade cannot handle due to seller non-cooperation. Needless to say, ODR and insurance systems are orthogonal with respect to reputation systems in their effect on the marketplace as they address mostly non-fraudulent disputes. Thus, for brevity and simplicity in this paper we do not analyze ODR and insurance systems.

3. REPUTATION QUANTIFIERS

In addition to the IC^3 report, a recent survey by the National Fraud Information Center (NFIC) presented statistics that in 2006, an average loss for an Internet fraud reported to NFIC totalled US$1512 [8]. The top two types of fraud: auctions and general merchandise, accounted for 34% and 33% of all reported fraudulent activity with an average loss of US$1331 and US$1197 per case respectively. Clearly, losses experienced by consumers undermine the popularity of online markets such as eBay or the Amazon.com Marketplace.

Existing reputation systems are in place in such markets to predict fraudulent activity [1, 5] – however, they are not fool-proof. To address this issue, we propose a reputation system whose objective is not to probabilistically aid prediction of fraud which is common practice – but to assure buyers of deterministic pricing tactics that cannot profit the seller in case of a fraudulent transaction. We model seller's reputation using the following two monetary values: a **sales limit** and a **reimbursement fund**.

Definition 1. ***Sales limit*** *α_i for a specific user c_i is an upper bound on pricing $\mathbf{p}_i$ such that if c_i commits fraud on each item offered in $\mathbf{p}_i$ she can still not profit from her existence in the market as a consumer.*

By definition, α_i is set such that c_i could not make profit in the system if:

$$\sum_{\forall p_j \in \mathbf{p}_i} p_j \leq \alpha_i. \tag{1}$$

Definition 2. ***Reimbursement fund*** *β_i for a specific user c_i is a sum of money that can be used to offset losses to buyers who participate in pending transactions with c_i in case c_i commits fraud.*

In our system, each seller c_i chooses the value of β_i according to her required selling power. In general, consumers feel comfortable bidding to products from c_i knowing that any fraud would get fully reimbursed if pricing is such that:

$$\sum_{\forall p_j \in \mathbf{p}_i} p_j \leq \beta_i. \tag{2}$$

If pricing on $\mathbf{p}_i$ is over β_i and fraud is committed with losses greater than β_i, FIFO is one possible fair algorithm for using the reimbursement fund among the defrauded consumers. A reimbursement fund could be implemented via escrow accounts, transaction insurance, etc. Such funds are certainly not a novel mechanism to protect buyers; the introduction of a sales limit, $\alpha_i \geq \beta_i$, and subsequent techniques to construct and use it, is our contribution in this work.

3.1 Risk Taking

In our system, a buyer could get defrauded if she chooses to pay a price that sets $\pi_i = \sum_{\forall p_j \in \mathbf{p}_i} p_j$ over β_i. We consider two cases. First, the buyer is unlikely to encounter a fraudulent seller as long as she chooses to pay below seller's sales limit, $\alpha_i \geq \beta_i$. Although the seller could certainly defraud such a buyer, he would still not gain any profit. Repeating the iteration "open new user account, build reputation, then defraud" would still not profit an adversary because she would have to invest substantial funds to build the reputation of the fabricated user in each iteration to finally claim back these funds at the expense of innocent buyers. In the remainder of this section, we introduce an algorithm for constructing a sales limit as well as a technique for time-sharing sales limits, i.e., risk, among market participants in order to boost their selling power.

Second, if the buyer decides to pay over α_i, the risk of encountering a fraudulent transaction can be quantified depending upon the adversary's profit, $\pi_i - \alpha_i$. In Section 4, we introduce a simple, yet intuitive empirical model that aims to predict a fraudulent transaction based upon the incentive: $\pi_i - \alpha_i$.

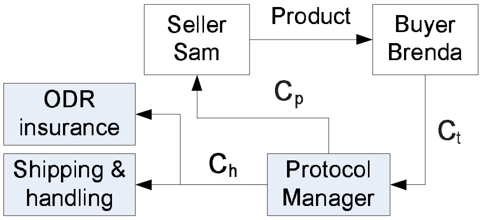

Figure 1: Transaction model: entities and involved costs.

3.2 The Transaction Model

The proposed reputation quantifier α_i is computed based upon consumer's prior transaction record. In order to establish an algorithm for its computation, we first adopt a simple transaction model. Fig. 1 illustrates a diagram of the basic system model. We first review the transaction costs. The cost of an individual transaction $\mathcal{C}_t = \mathcal{C}_p + \mathcal{C}_m + \mathcal{C}_h$, paid by the buyer, is composed of three entities:

- the **product price**, $\mathcal{C}_p$, which represents the total amount of money after all costs received by the seller,
- the **protocol manager fee**, $\mathcal{C}_m$, is paid to the mediator in the transaction, e.g., eBay or Amazon.com,
- the **miscellaneous fee**, $\mathcal{C}_h$, which includes other fees such as: arbitration insurance, shipping and handling, taxes, etc. The protocol manager (PM) may orchestrate some of these activities. All miscellaneous fees that can be verified by a trusted party (e.g., PM) are used to establish participants' sales limits.

Now we model an actual transaction between a seller, Sam, and a buyer, Brenda. Once the negotiation has completed, Brenda pays the amount due, $\mathcal{C}_t$, to Sam who now pays the transaction fee, $\mathcal{C}_m$, to the PM, both using an arbitrary payment system. The PM could offer a payment service for market participants[2] to simplify payments and reduce transaction fees. As opposed to the reimbursement fund which could be implemented as an escrow account, the PM does not serve as an escrow for the cash flow between the market participants. After receiving the transaction fee, the PM updates the accounts of all parties involved. Next, Sam is now required to deliver the merchandise to Brenda. Here is a list of considered outcomes upon merchandise delivery (or failure to):

{P} **positive feedback** – Brenda is satisfied with the outcome of the transaction; she compliments Sam.

{N} **negative feedback** – Brenda is dissatisfied with the received product; the participants in the transaction decide to resolve the situation as follows:

{N.1} **no refund** – Sam accepts negative feedback and Brenda does not initiate the refund process. This would be typical for a transaction with low $\mathcal{C}_p$.

{N.2} **refund/return** – Sam refunds Brenda for previously returned merchandise. If this process is closed at Brenda's satisfaction, the transaction record is deleted including Sam's negative feedback.

[2] This would be an equivalent to handling a transaction via eBay and PayPal.

{N.3} **dispute** – occurs in all other cases. This is the most interesting case, as it involves arbitration and resources for refunding the plaintiff.

A recent study suggests that 41% of seller-targeted disputes occur because sellers do not describe their products accurately which results in complaints by buyers once they receive the products [9]. In about half of such cases, the buyer chooses not to return the product; hence, we estimate {N.1} to account for about one fifth of all {N} cases. A survey of 225 {N.2} and {N.3} disputes on eBay in 1999 [10] points to $\approx 25\%$ of disputes that were resolved at mutual success, $\approx 25\%$ of them at impasse, and the remainder never entered the resolution stage although it was available for free as part of a study. Thus, to the best of our knowledge we conclude that detailed statistics about dispute estimates are unavailable; however, existing data points to a solid likelihood that {N.3} cases are relatively common, yet due to unavailability of inexpensive and efficient ODR systems they remain underreported.

3.3 Computing the Reputation Quantifiers

In this subsection, we evaluate how transaction outcomes affect buyer's and seller's reputation quantifiers. The global objective for the developed algorithms is to maximize the selling power in the system. Achieving this objective is important as it minimizes sellers' investments to reach a specific selling power, hence boosts the market economy, which on the side of the market organizers results in higher profit.

Case {P}. We are motivated by the fact that for a specific transaction $t(c_B, c_S)$ all overhead costs, $\mathcal{C}_o = \mathcal{C}_t - \mathcal{C}_p$, can be attributed to both the buyer and the seller. In this section, we assume that all miscellaneous costs can be verified by PM. This assumption may not always be true – for example, shipping costs, if not paid for via PM's payment system typically cannot be provably verified. The PM would subtract expenses that cannot be verified from $\mathcal{C}_o$ before applying them to seller's and buyer's reputation quantifiers.

We could take a stance that the buyer pays a fair market price for the product that includes $\mathcal{C}_o$ and that the seller is the one paying for transaction costs (i.e., in an off-line market this is certainly the case as the buyer pays the street price and the merchant offsets all costs to retain profit). In such a setting, after each committed transaction, seller's sales limit increment due to a transaction t equals:

$$\alpha_S(t(c_B, c_S)) = w_{BS} = \mathcal{C}_o. \tag{3}$$

In essence, $\mathcal{C}_o$ is the "loss" that the seller has with respect to fair market price. In another variant, the buyer and the seller could negotiate a shared application of the transaction cost during negotiation. This sets up a more general case for computing sales limits:

$$\alpha_S(t(c_B, c_S)) = w_{BS} = \varrho \mathcal{C}_o, \tag{4}$$

$$\alpha_B(t(c_S, c_B)) = w_{SB} = (1 - \varrho)\mathcal{C}_o, \tag{5}$$

where $0 \leq \varrho \leq 1$ is a parameter that scales the application of costs to buyer's and seller's sales limits. Note that in this case, an edge $t(c_S, c_B)$ directed $c_S \rightarrow c_B$ is added to $\mathbb{T}$ with an appropriate weight factor.

Case {N.1}. Only seller's reputation is affected by this case. Here, seller's sales limit is reduced by:

$$\alpha_S(t(c_B, c_S)) = w_{BS} = -\mathcal{C}_p, \tag{6}$$

if the PM can verify all miscellaneous costs. If this is not the case, all non-verified costs are also subtracted from seller's sales limit.

Case {N.3}. Disputes in on-line transactions are typically resolved using PM's or third party's ODR systems [7, 6]. Costs related to ODR are included in $\mathcal{C}_h$ as insurance against this outcome. Possible outcomes for the ODR process are:

(i) resolution in favor of one of the participants in the transaction; then this case is resolved as **{P}**, **{N.1}**, or **{N.2}** with respect to the sales limit.

(ii) impasse; a bargaining impasse occurs when the two sides negotiating an agreement are unable to reach an agreement and become deadlocked[3]. This situation is difficult to handle because possible solutions can hurt the party who is innocent. Certainly, entities who plan on participating in a transaction with either c_S or c_B should know that they have been involved in this dispute. As long as the dispute is in impasse, seller's sales limit is affected as defined in case **{N.1}**, (6) and buyer's record shows participation in a deadlocked dispute.

(iii) lack of co-operation in the ODR process by the seller, c_S; typically a consequence of fraud. Such an outcome of a transaction would reduce the sales limit of c_S as defined in case **{N.1}**, (6).

Fraud (case iii) is committed by sellers in vast majority of cases. One way for a buyer to cause a serious misconduct is to complain about a received product, agree to a return for refund, and then return a different, less valued product to the seller. Such cases are exceptionally infrequent and would result in a criminal investigation.[4] Our system does not protect sellers from such events.

The overall sales limit for a specific consumer, c_i, is then computed as follows:

$$\alpha_i = \sum_{\forall t(c_j, c_i) \in \mathbb{T}_i} w_{ji} + \beta_i, \tag{7}$$

where $\mathbb{T}_i$ is a subset of all edges in $\mathbb{T}$ with c_i as a destination.

Equation (7) includes the reimbursement fund, β_i, that c_i establishes to insure customers from potential fraud (see Def.2). Typically, a new seller would deposit a specific amount $\beta_i(0)$ into its account with the PM to start up its reputation, i.e., an initial sales limit of $\alpha_i(0) = \beta_i(0)$. Succeeding sales would establish its sales limit. Then, c_i can balance the value of its reimbursement fund (this fund can be lowered or increased on-demand) and thus adjust its sales limit, to achieve a desirable selling power. The reimbursement fund is utilized by the PM in, for example, FIFO manner when a transaction fails.

3.3.1 Bidding

When a buyer, c_B, aims to bid for an item sold by c_S, the system presents several quantifiers to c_B: α_S, β_S, and the current pricing of all items sold by c_S: $\pi_S = \sum_{\forall p_i \in \mathbf{P}_S} p_i$.

[3] c. Wikipedia.

[4] Tracing perpetrators in this case is easier than in fraudulent transactions committed by sellers due to the undeniable availability of buyer's physical address.

Based upon these quantifiers, c_B can decide upon the risk she is willing to take while bidding on an item sold by c_S that would increase the total price of his offering to π_{new}. For example, if $\pi_{new} > \alpha_S$, c_B can ask c_S to increase his α_S by increasing his reimbursement fund so that she can bid comfortably knowing that c_S cannot make profits in case he decides never to deliver the product. Similarly, c_S can eliminate any risk in her bid by asking c_B to set $\beta_S = \pi_{new}$.

3.4 Time-Sharing Sales Limits

On-line markets based upon reputation systems usually consist of a few users who are predominantly sellers and the remaining majority of users who are predominantly buyers. Thus, we offer a supplemental algorithm for computing sales limits with an objective to enable consumers establish higher (up to twice as large) sales limits at a *risk.* Higher sales limits in the economy translate to increased selling power, hence higher profits for everyone involved.

Here, for a specific executed $\{\mathbf{P}\}$-transaction $t(c_B, c_S)$, c_B and c_S create an agreement to distribute the costs of t, $\mathcal{C}_o(t)$, **on-demand** so that at any time:

$$\alpha_B(t) + \alpha_S(t) = \mathcal{C}_o(t), \tag{8}$$

where $\alpha_X(t)$ denotes a portion of the verifiable cost $\mathcal{C}_o(t)$ for transaction t, that is used to build up the sales limit $\alpha_X = \sum_{\forall t \in \mathbb{T}_X} \alpha_X(t)$ of c_X. User c_X participated in t either as a buyer or as a seller.

Under the agreement, if at a specific moment, only one of the participants in t, say c_B, is selling an item then $\alpha_B(t) = \mathcal{C}_o(t), \alpha_S(t) = 0$. Note that this flexibility comes at risk for c_S. If c_B commits a fraudulent transaction and her sales limit gets affected while she was using more than $\frac{1}{2}\mathcal{C}_o(t)$ to boost α_B, the reduction in her sales limit may proportionally, and possibly entirely, reduce the amount $\mathcal{C}_o(t)$ shared by the two parties and thus, affect $\alpha_S(t)$ as in (8). Consequently, when committing to t with time-shared costs, both participants agree to take on this risk. Since the reputation system offers preventive services against fraud, we anticipate that this risk is low and typically worth the increased selling power, in particular for new or infrequent system users.

3.4.1 Sales Limit Computation

We now formally present how c_B and c_S time-share the transaction cost $\mathcal{C}_o(t)$. When a prospective buyer c_D wants to establish the sales limit of c_S, the system displays:

$$\alpha_S = \beta_S + \sum_{\forall t \in \mathbb{T}_S} \alpha_S(t), \tag{9}$$

where the values $\alpha_S(t)$ are "grown" as much as possible within each $\{\mathbf{P}\}$-transaction t in $\mathbb{T}_S$ with time-sharing of sales limits. The costs of remaining transactions within $\mathbb{T}_S$ are accumulated as defined in Subsection 3.3. In the remainder of this subsection, for simplicity and brevity, we assume that all transactions in $\mathbb{T}_S$ are time-shared and $(\forall S)$ $\beta_S = 0$. We first define the following scalar:

$$\widehat{\alpha_S} = \sum_{\forall t \in \mathbb{T}_S} \alpha_S(t), \tag{10}$$

and establish the goal of minimizing the potential market-wide profit from fraud:

$$R = \sum_S \pi_S - \widehat{\alpha_S}. \tag{11}$$

Since fraud is typically not a wide-spread phenomenon, there exists demand to address it locally within the market network. Thus, we want to establish a set of rules that govern the fairness of the cost allocations, i.e., that should encourage participants to use time-sharing by guaranteeing that no participant can take a risk which is not proportional to her committed transactions.

Definition 3. ***The absolute fairness rule*** *asserts that for every seller, c_S, in the market:*

$$\sum_{t \in \mathbb{T}_S} \alpha_S(t) - \min \left\{ \pi_S, \frac{1}{2} \sum_{t \in \mathbb{T}_S} \mathcal{C}_o(t) \right\} \geq 0. \tag{12}$$

In other words, absolute fairness guarantees to each user that her sales limit will be built over time on-demand and will equal at least one half of the sum of costs for all her committed transactions. We now show how to compute the values $\alpha_S(t)$ with absolute fairness while minimizing R.

Consider a flow network $G(V, E)$ with a single source v_{source} and a single sink v_{sink}. Each market user, c_S, is assigned a node, v_S, which is connected to v_{source} with an edge of capacity π_S. For each transaction, $t(c_S, c_B)$, we construct a node $v_{S,B}$ and add two edges of infinite capacity: one from v_S to $v_{S,B}$ and another from v_B to $v_{S,B}$. Finally, we connect each transaction node to v_{sink} with an edge of capacity $\mathcal{C}_o(t(c_S, c_B))$. Figure 2 illustrates an exemplary flow network with four participants and five transactions.

A *legitimate sharing* of the cost of a transaction, $\mathcal{C}_o(t)$, between the seller and the buyer is such that the sum of shared sales limits obeys Equation (8).

Lemma 1. *Let f be a maximum flow in $G(V, E)$. Setting $\alpha_S(t(c_S, c_B))$ to be the flow in f from v_S through the edge $(v_{S,B}, v_{sink})$ is a legitimate set of values that minimizes R.*

Proof. First, the proposed values for $\alpha_S(t)$ are legal values as for every transaction, $t(c_S, c_B)$, the flow through the edge $(v_{S,B}, v_{sink})$ is at most $\mathcal{C}_o(t)$.

Now, every set of legitimate values to $\alpha_S(t)$ is also a legal flow in the network as $\alpha_S(t(c_S, c_B))$ can be pushed from the source through v_S and $v_{S,B}$ to the sink. As $\sum_{t \in \mathbb{T}_S} \alpha_S(t) \leq \pi_S$ and $\alpha_S(t) + \alpha_B(t) \leq \mathcal{C}_o(t)$ the flow in each edge is less than or equal to its capacity. Assuming by contradiction that there is a legitimate set of values $\alpha_S(t) = \zeta_S(t)$ such that $\sum_S \sum_{t \in \mathbb{T}_S} \zeta_S(t) > f(v_{source}, v_{sink})$, is in contrast to f being the maximum flow in the network. □

Maximal flow algorithms in networks with a single source and sink run in polynomial time in the network size. Complexity $\mathcal{O}(|V|^2 \log(|V|^2/|E|))$ is achieved by the push-relabel algorithm that uses dynamic trees [20]. In our application, G is constructed so that the number of edges is of order $|E| = \mathcal{O}(|V|)$. Therefore, the push-relabel and the Dinic's blocking flow algorithm with dynamic trees, [21], have an overall complexity of $\mathcal{O}(|V|^2 \log(|V|))$. A recent result by Goldberg and Rao reduces the complexity to $\mathcal{O}(|V|^{3/2} \log(|V|))$ [22]. According to Lemma 1, we know how to minimize R in polynomial time, and we are left with the problem of making the maximal flow obey the absolute fairness rule.

We address this problem near-optimally as follows. First, we replicate K times each transaction node in G and assign K times lower capacities to the edges going from the replicated nodes to the sink. In this new network, we run the classical max-flow algorithm by Edmond and Karp, [23], in a randomized manner. Once a specific participant c_S is selected as the first node in a shortest path through the residual network of G, we exclude its corresponding node, v_S, from the randomized round-robin until every node, v_X, that has participated in a transaction with c_S and such that (v_{source}, v_X) is not saturated, is visited at least once. The structure of the network helps us reduce the time needed to execute this algorithm as most shortest paths from the source to the sink are of length 3, with each of them saturating either an edge from source to a participant's node, or an edge from a transaction node to the sink.

Moreover, after solving the max-flow problem in the network once, only a marginal computational effort is needed to account for possible new events: adding an item for sale, successful new transaction, and a fraud which results in removing a participant and all of her previous committed transactions from the network.

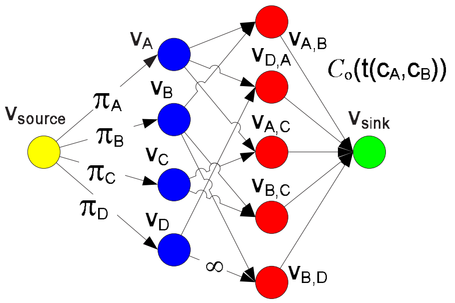

Figure 2: An example of a flow network with four participants and five transaction nodes.

- **Adding an item for sale by** c_S results in the increase of π_S. If $\alpha_S < \pi_S$ before this event, then no operation is needed. Otherwise, the algorithm then continues pushing the additional flow through the network via at least one additional iteration.

- **Successful new transaction** $t(c_S, c_B)$ results in reducing π_S and adding a transaction node and its K replicas. If $\alpha_S > \pi_S$ after this event, we push back $\alpha_S - \pi_S$ flow from the sink through the transaction nodes connected to v_S. The algorithm then continues to distribute the flow through the network using at least one additional iteration.

- **Fraud by participant** c_S results in removing the node v_S and all her transactions nodes $t(c_S, c_X)$ from the network. This can negatively affect any participant c_X and her portion of the flow in $t(c_S, c_X)$ should be pushed back.

In a typical on-line market, the continuous chain of transactions would result in an application of one of the previous three steps in an iterative manner. The absolute fairness is positively correlated with K, however at higher overhead to performance. Regardless of K, iteratively computing the maximum flow in the network after the above-mentioned events, negatively affects fairness by slowly accumulating fairness errors. To address this problem, we propose full computation of the maximum flow after a specific number of events in the network. In addition, if one wishes to compromise optimality for the sake of fast computation, the length of the shortest-path from source to sink can be limited to a constant, resulting in a linear time algorithm, which in a lightly constrained network could prove to be an efficient near-optimal option.

For large networks which could be potentially decentralized, it is important to consider algorithms with constant complexity or sub-linear in the size of the network. The key to building such algorithms is approaching absolute fairness in sub-optimal but localized manner. Here is an example of a localized cost redistribution algorithm:

$$\alpha'_S(t) = \begin{cases} \frac{1}{2}\mathcal{C}_o(t) & , \pi_S \geq \frac{1}{2}U_S \\ \frac{1}{2}\mathcal{C}_o(t) - \frac{1}{|\mathbb{T}_S|}\left(\frac{1}{2}U_S - \pi_S\right) & , \text{else} \end{cases} \qquad (13)$$

where $U_S = \sum_{t \in \mathbb{T}_S} \mathcal{C}_o(t)$. We locally utilize the available costs on per-transaction basis by setting $\alpha_S(t) = \mathcal{C}_o(t) - \alpha'_B(t)$ for each transaction $t(c_S, c_B)$ such that $\pi_S \geq \frac{1}{2}U_S$ or $\alpha_S(t) = \alpha'_S(t)$ otherwise. The algorithm could be generalized so that a certain neighborhood to c_S and c_B is exposed to further iterative redistribution. Localized algorithms allow for fast local updates of values when an event occurs. These updates involve only the value of a participant c_S and her neighbors with whom c_S has committed $\mathbf{\{P\}}$-transactions with time-sharing of sales limits.

Alternative fairness rules could be used while time-sharing sellers' sales limits, depending on the interests of the PM and the type of risk that it wishes to impose over market participants. One such example is the max-min fairness, in which the minimum $\widehat{\alpha_S}$ that a participant c_S achieves is maximized; secondly, the second lowest $\widehat{\alpha_{S'}}$ that a participant $c_{S'}$ achieves is maximized, etc. Max-min fairness allows setting $\widehat{\alpha_S}$ proportional to π_S, thus encouraging high level of transactions by some participants, while others, with lower π take on more risk. For brevity and simplicity, we do not propose any specific max-min fairness algorithms in this paper. However, we do note that minimizing R under this condition can be done using the already proposed flow network, G, with an optimization goal to maximize a multi-commodity flow, where every π_S is considered a distinct commodity and every edge $(v_{S,B}, v_{sink})$ can transfer only two commodities. The problem of maximal flow in multi-commodity networks is notoriously hard and the best known solution to the problem is a $(1-\epsilon)^{-3}$-approximation which takes $O(\epsilon^{-2}|E|^2 \log |E|)$ time [24].

3.4.2 Discussion

From the perspective of the adversary, time-sharing sales-limits could present an opportunity. By purchasing merchandise worth X monetary units, the adversary gains a maximum sales limit equal to the sum of all transaction fees, e.g., at eBay this amounts to $\frac{X}{20}$, for these purchases. As we speculate that it is unlikely for an adversary to spend twenty units of her own wealth prior to gaining one unit of fraudulent profit, we conclude that time-sharing is an effective mechanism to at most double the selling power of economic entities in the market. In addition, the proposed reputation system could enable pricing time-sharing of sales limits, and thus reach an equilibrium for the risk vs. profit from time-sharing sales limits in a market.

From the theoretical point of view, one could evaluate the sensitivity of the increase in sales limits depending on the market constraint: average current offering, $\overline{\pi}$, vs. average sum of fees from previous committed transactions, $\overline{\alpha}$. We assume that this analysis is of little practical importance for several reasons. First, in a well-established market, $\overline{\alpha} \gg \overline{\pi}$. Second, market networks are typically scale-free with a small group of "frequent sellers" generating large portion of the flow in the network. These nodes typically have an offering that is substantially smaller than their current sales limit. Thus, we expect that most of the demand for increase in sales limits from smaller sellers, will be sourced out from the "frequent sellers." Consequently, we anticipate that most sellers who demand an increase in their sales limits will highly likely succeed to double them with no adverse effects on the trade in the global market. To that extent, we do not present an experimental study on the efficacy of the proposed algorithms for time-sharing of sales limits.

3.5 Summary

In summary, the proposed algorithms for computing sales limits in a reputation network use "transaction losses" such as shipping and handling, insurance, protocol manager fees, etc., **not** transaction totals, to build up a value that quantifies user trustworthiness. Positive feedback, as it could be easily fabricated, does not cause that the value of the sold item affects seller's sales limit. By proposing a suite of algorithms that trade-off certain risk and optimized selling power with trustworthiness, we address the market demand for robust trades. Most importantly, our system is the first to offer **deterministic** guarantees to buyers in generally distrusted markets – as a consequence our system facilitates trade, offers new risk-taking opportunities, and should boost market pricing due to increased system security.

4. SELLER'S FRAUD MODEL

One disadvantage of deterministic reputation is the conformation towards the worst case. Since fraud is costly but still not frequent, we speculate that risk assessment technologies are still of value, in particular when bootstrapping the economic activity in the market. For instance, consider the scenario when a buyer, c_B, aims to bid for an item sold by c_S. In this case, the model presented in the previous section provides guarantees to c_B and as long as the price she offers, π, satisfies $\pi \leq \beta_S$, c_B cannot be defrauded. If $\pi > \beta_S$, c_B can ask c_S to increase his reimbursement fund. However, if c_S does not have the resources necessary to increase his reimbursement fund, c_B may not be willing to place a higher bid due to an increased likelihood of fraud. Such a system may end up in a bargaining impasse which is, on the average, a loss for all participants in the economic system. In order to facilitate bargaining through risk assessment, we introduce a novel price-dependent probabilistic reputation system. As a crucial component of this system, we introduce an additional reputation quantifier which we refer to as the **seller's fraud model**.

Definition 4. ***Seller's Fraud Model*** *is defined as a probability $\gamma_S(\mathbf{p}_S)$ that a seller c_S decides to defraud her current buyers based upon the pricing $\mathbf{p}_S$ of her product offering. The model is quantified using a function $f()$:*

$$\begin{aligned} \gamma_S(\mathbf{p}_S) &= \Pr[c_S \text{ commits fraud} | \mathbf{p}_S] \qquad (14) \\ &= f(\pi_S - \alpha_S) = f\left[\sum_{p_i \in \mathbf{p}_S} p_i - \alpha_S\right] \end{aligned}$$

over the profit that c_S would create if she would disappear from the market after charging for all listed products $\mathbf{p}_S$.

Before bidding for a product at a certain price, the buyer would be presented with a model that estimates the probability of a fraudulent transaction given the current offering of the seller and its pricing. The tool would offer normalized risk assessment based upon $f()$ trained on empirical market data. There exist numerous possibilities for creating efficient user interfaces to deliver the resulting probability, e.g., buyer would enter considered price into an HTML form field to observe the probability in question using a graphical display such as a pointer to a log γ-scale.

An example of an expected γ-model is illustrated in Figure 3. We observe that the probability is zero[5] for $\pi_S \leq \beta_S$, approximately zero for $\beta_S < \pi_S \leq \alpha_S$, and increasing in phase-transition style with the increase of seller's profits. The resulting probability converges towards Γ, the probability that one person commits fraud regardless of payout. This convergence is typical for any practical range of prices over $\pi_S - \alpha_S$.

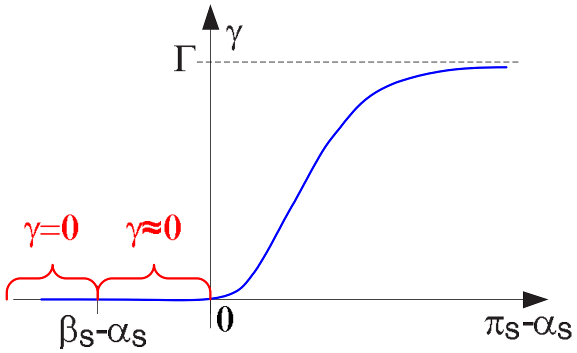

Figure 3: An example of the seller's fraud model.

4.1 Discussion

Interestingly, our γ-model does not consider the counts of positive and negative responses from previous customers nor any topology analysis of the transaction graph $\mathcal{G}$ – tools typical for traditional reputation systems. It does not need to. The fact that seller's existence in the economic ecosystem is reduced and accurately presented with only two parameters, α and β, renders other details about previous transactions, such as structure of the reputation tree, irrelevant. The fact is that a fraudulent seller has only one objective – to maximize profits during his existence in the on-line market. In most realistic scenarios this objective is amended to the desire to slip past detectors that would trigger criminal investigation. From that point of view, the only statistic which is crucial is the probability that, given a specific payout, the seller decides to fool her current buyers.

Obtaining function $f()$ empirically could be a difficult task. The problem lies in the fact that not all sellers are equal and some form of seller classification could be required. For example, sellers coming from countries with drastically different income levels should have different financial motives to commit a fraudulent transaction. Considering these facts, it is more realistic to expect that sellers are classified

[5] In the most strict sense this value is zero only approximately due to unexpected events that could prevent the seller from completing the transaction and that cannot be considered fraud.

to fit different behavioral models. Based upon seller's classification, the PM would select the appropriate γ-model and present it to prospective buyers. Classification algorithms have been researched well in several subfields of computer science, hence we point the interested reader to review some of the related work [11]. For brevity and simplicity, we do not focus on this aspect of our technology.

Finally, users who chose to use time-sharing of their transaction costs would use the γ-model in the same fashion as conservative sellers. Here, the risk is not only exhibited by the buyer but by the collaborating sellers as well. To issue a warning to a prospective buyer about the additional risk, the reputation system could show both the conservative and the time-shared α quantifier. Then, the buyer, fully informed, can assess the true risk and proceed with the pricing.

5. EMPIRICAL ANALYSIS

In this section, we present results from our empirical study. First, we describe the sampling method used to obtain a snapshot of real-world economic activity. Then, we present key statistics about our snapshot, followed by the main result: a seller's "fraud" model empirically obtained from real transactions. We built our datasets from public information available on existing on-line marketplaces. We remind the reader that payment channels available on existing on-line marketplaces are typically left up to user's free choice and are thus unsupervised. Based on the data we reviewed, in our empirical study we used only PM fees to build sellers' sales limits. Other transaction fees such as shipping and taxes, were not included in the construction of sales limits as modern on-line marketplaces typically do not receive receipts for such services.

5.1 Marketplace Sampling

We did not conduct our empirical study using a sampling technique that would model a marketplace network as accurately as possible. The primary objective of our work is to present an estimate of the seller's fraud model, $f()$ – and not to provide accurate modeling of marketplace networks. True random sampling of a large network is difficult as random walks, one of the most dominant techniques for this task, do not capture graph statistics accurately as they tend to visit well-connected nodes more often. In relatively sparse networks, even techniques that aim at uniform graph sampling are not sufficiently precise [19]. We acknowledge that it is difficult to create a sample of a large marketplace network that would precisely correspond to the true activity on its ecosystem and conclude that accurate statistics could be provided only by the marketplaces themselves.

Still, for the purpose of validating the ideas proposed in this paper, we decided to sample marketplaces with an objective of covering as many as possible sellers with no particular browsing objective (i.e., a greedy max-cardinality subset). Thus, we hoped that the statistics of a large subnet would provide an insight into quantifying $f()$. We followed a simple approach for extracting information from a given marketplace Web-site. We implemented a Web-crawler that automated the process, comprised of two stages.

In the first **data collection** stage, by starting with a arbitrarily chosen user, c_S, we would gather her key statistics: both positive and negative fan-in and fan-out and the current product offering, $\mathbf{p}_S$. Extracting π_S was sped up by recording only the first 200 products of seller's current product offering as reported by the marketplace. As the total number of items for sale was known, we extrapolated π_S based upon the average price of the first 200 products. All monetary values listed in Euro and Pound sterling, were converted to US dollars. For simplicity, we ignored transactions credited with other currencies.

In the second **traversal** stage, we performed a *breadth-first* search of the user's feedback pages, collecting information about every single transaction she performed, negative or not. For each transaction, we collected the anonymized users' IDs, the related transaction amounts with associated time-stamps, and the type of feedback reported (positive, negative or neutral). The recursive traversal was done in breadth-first manner with no exceptions.

5.2 Snapshot Statistics

During our sampling of on-line marketplaces, we collected data on 10,096,731 transactions worth over US\$270M. Our Web-crawler downloaded and parsed a total of 140,000 Web-pages, collecting transaction information for a large number of anonymized user accounts. We partitioned the type of users encompassed by our extracted subnets as *complete* and *incomplete*. For the first class, we focused on computing/estimating the information on their π_i and α_i. Thus, we extracted their full set of transactions. The cardinality of the set of complete users was 44,830. In order to record all their transactions we needed to include information about the incomplete users, i.e., users who have participated in at least one transaction with a complete user. The total number of incomplete users in our dataset was 5,274,759. Transactions between two incomplete users were not included in the subnet. Thus, the total size of the extracted subnet was 5,319,589 accounts. The statistics about the fan-in and fan-out of nodes in the subnet are presented in Figure 4. One can observe, as expected, that a smaller number of sellers is generating large number of transactions in the marketplaces.

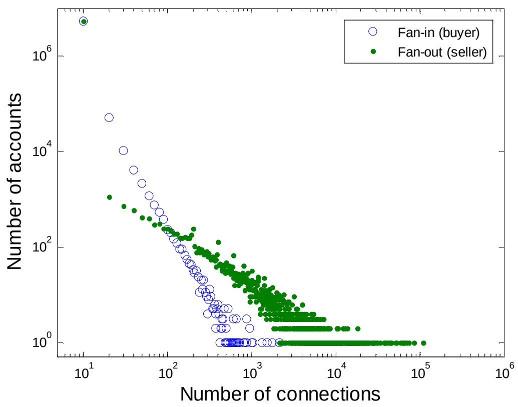

Figure 4: Histogram of the number of transactions nodes have achieved as buyers and sellers within the captured subnet.

Figures 5 and 6 describe the number of accounts in our sample that had a specific ratio and total number of transactions with negative feedback respectively. We observe, as expected, that most sellers do not initiate transactions that result in negative feedback. However, within the mass of users, we observe that there is a relatively large group of users who have recently generated substantial volume of neg-

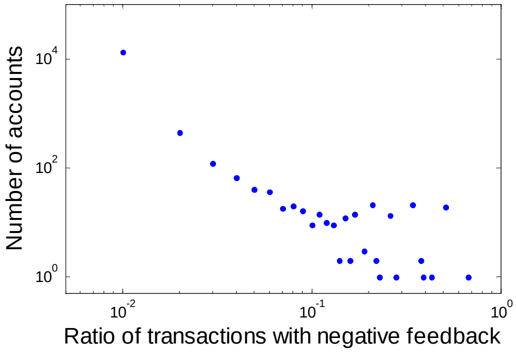

Figure 5: Histogram of the ratio of transactions that resulted in negative feedback within the captured subnet.

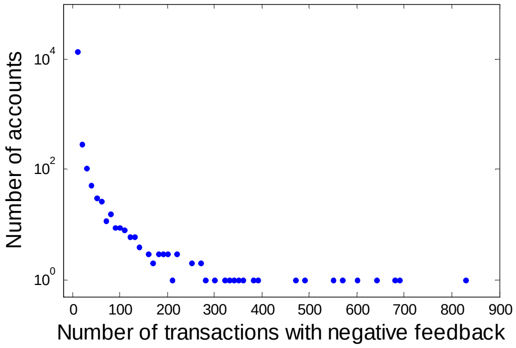

Figure 6: Histogram of the number of transactions that resulted in negative feedback within the captured subnet.

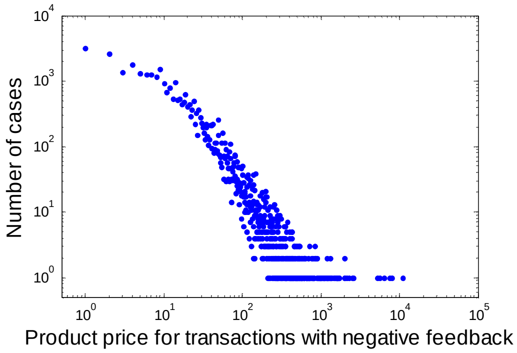

Figure 7: Histogram of prices for products exchanged in transactions that resulted in negative feedback within the captured subnet.

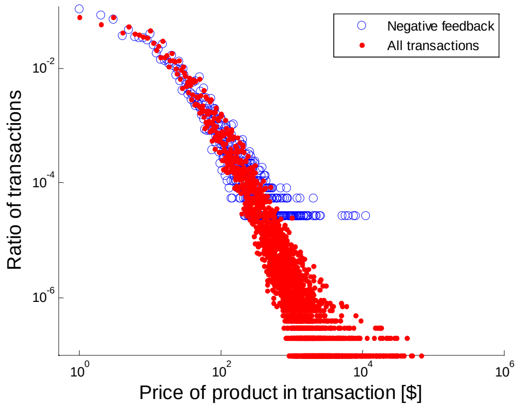

Figure 8: Ratio of transactions that result in negative and any feedback vs. the price of the product within the captured subnet.

ative feedback. Figure 7 illustrates the histogram of prices for products exchanged in transactions that resulted in negative feedback within the captured subnet. This plot is interesting as it points to a relatively low price of merchandise that is sold to the dissatisfaction of buyers – it also suggests that sellers aim at combining smaller profits from several transactions with negative feedback. In the subsequent subsection we conclude that our "fraud model" is particularly tailored to address this type of malicious activity.

Finally, we point to Figure 8 which plots the ratio of transactions that involve a product of certain price. We plot two sets of datapoints: for all transactions and for transactions with negative feedback only. One can observe from the curves that the pricing of merchandise that results in a transaction with negative feedback typically has similar pricing to positive-feedback transactions. This is reasonable as fraudulent transactions usually have "fair" or slightly lower pricing over the product bait in order to attract buyers.

5.3 Seller's "Fraud" Model

From the sampled subnet with approximately 10 million transactions we have constructed a statistical model for our seller's fraud model presented in Section 4. We must confirm that on-line marketplaces typically do not report actual fraudulent transactions on their Web-sites, rather report negative feedback. Therefore, to be more precise, we stress that our model represents accurately seller's *negative feedback* data from the sample. We speculate that it is strongly correlated to the actual fraud model – certainly, significant rise in the pdf of the obtained $f()$ model occurs at approximately the same value as the costs of fraudulent transactions reported to NFIC and IC^3.

Figure 9 presents a set of points that correspond to the log-fraction of transactions for which $\pi_i - \alpha_i$ corresponded to the abscissa, for which sellers received negative feedback. We have also provided a 6-th degree intrapolant for the collected data that outlines in a visually clear manner the probability of interest. Note that the curve is resemblant of the geometry anticipated in Figure 3. The variance of the intrapolated curve is greater at higher amounts (>US\$10,000) due to lack of data. Finally, the estimated model confirms the speculation that the probability of a fraudulent transaction rises strongly at $\pi_i - \alpha_i >$ US\$1000 and reaches more than 10 times higher values at $\pi_i - \alpha_i \approx$ US\$30K compared to transactions that executed when $\pi_i - \alpha_i <$ US\$1000. Informing consumers about this trend is the least that could be done, while the proposed remedies such as presenting a buyer with seller's $\{\alpha_i, \beta_i\}$ parameters would likely further improve robustness to fraudulent activity in on-line marketplaces. Since systems that aim at preventing fraud by pricing reputation commonly demand economic bootstrapping, in Section 3.4 we have proposed a technique for sharing reputation among sellers that boosts their selling power at certain quantifiable risks.

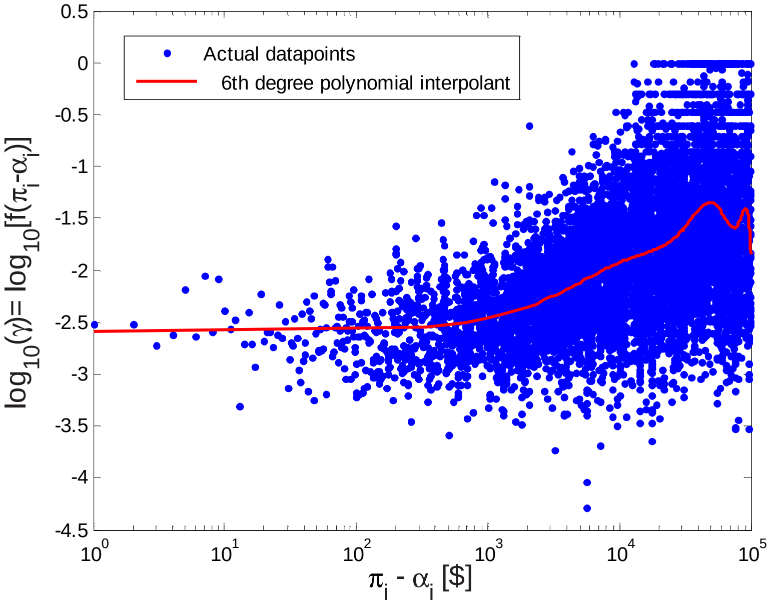

Figure 9: The seller's "fraud," i.e., negative feedback, model extrapolated from 10 million sampled transactions within the captured subnet.

6. SUMMARY

In this paper, we propose a methodology that aims at pricing reputation from seller's perspective. Buyers need not evaluate sellers' transaction trees and seek whether previous transactions were fabricated. We price reputation at three levels easily accessible to a common user:

- a limit on pricing that guarantees that the buyer will receive reimbursement in case seller commits fraud,
- a limit on pricing that states that the seller will not generate any profits from her existence in the marketplace if she commits fraud on all her current product offerings and disappears from the economic ecosystem,
- an user interface that can quantify for a buyer the risk of a fraudulent transaction when placing a higher price than the previous two limits.

In our system, even in the case of stolen identity, an adversary cannot produce illegal profit unless a buyer decides to pay over the suggested sales limits. We obtained relatively large subnets from actual on-line marketplaces in order to empirically quantify the key parameters in our scheme and demonstrate how it could be efficient if placed in an existing on-line marketplace.

7. REFERENCES

[1] EBay Inc. http://www.ebay.com.
[2] Chat with Rob Chesnut, Vice President of eBay's Trust & Safety Dept. http://pages.ebay.com/event/robc.
[3] Internet Crime Complaint Center. Internet Crime Report for 2006. http://www.ic3.gov/media/ /annualreport/2006_IC3Report.pdf.
[4] Ebay Inc. Annual report 2006. http://investor.ebay.com/annuals.cfm.
[5] The Amazon Marketplace. http://www.amazon.com.
[6] S. Abernethy. Building Large-Scale Online Dispute Resolution & Trustmark Systems. UNECE Forum on ODR, 2003.
[7] E. Katsh and L. Wing. Ten Years of Online Dispute Resolution (ODR): Looking at the Past and Constructing the Future. The University of Toledo Law Review, Vol.38, (no.1), pp.19–47, 2006.
[8] National Fraud Information Center. Top 10 Internet Scam Trends from NCL's Fraud Center, 2006. http://fraud.org/stats/2006/internet.pdf.
[9] I. MacInnes. Causes of Disputes in Online Auctions. Electronic Markets, Vol.15, (no.2), pp.146–157, 2005.
[10] E. Katsh, et al. E-commerce, E-disputes, and E-dispute Resolution: In the Shadow of "eBay Law." Ohio State J. of Dispute Resolution, Vol.15, (no.3), pp.705–734, 2000.
[11] L. Brieman, et al. Classification and Regression trees. Wadsworth & Brooks, 1984.
[12] S.D. Kamvar, et al. The EigenTrust Algorithm for Reputation Management in P2P Networks. WWW, 2003.
[13] L. Xiong, et al. PeerTrust: Supporting Reputation-Based Trust in P2P Communities. IEEE Trans. on Knowledge and Data Engineering, Vol.16, (no.7), 2004.
[14] G. Swamynathan. Reputation Management in Decentralized Networks. Technical report, UCSB, 2007.
[15] EBay dispute resolution. http://pages.ebay.com/ /services/buyandsell/disputeres.html.
[16] I. MacInnes. Understanding Disputes In Online Auctions. eCommerce Conference, 2004.
[17] C. Dellarocas. Immunizing Online Reputation Reporting Systems against Unfair Ratings and Discriminatory Behavior. ACM EC, 2000.
[18] F. Cornelli, et al. Choosing Reputable Servents in a P2P Network. WWW, 2002.
[19] M.R. Henzinger, et al. On near-uniform URL sampling. International World Wide Web Conference on Computer Networks, 2000.
[20] A.V. Goldberg and R.E. Tarjan. A new approach to the maximum flow problem. ACM STOC, 1986.
[21] E.A. Dinic. Algorithm for solution of a problem of maximum flow in networks with power estimation. Soviet Math. Doklady, 1970.
[22] A.V. Goldberg and S. Rao. Beyond the flow decomposition barrier. J. of the ACM, Vol.45, (no.5), pp.783–797, 1998.
[23] J. Edmonds and R.M. Karp. Theoretical improvements in algorithmic efficiency for network flow. J. of the ACM, Vol.19, (no.2), pp.248–264, 1972.
[24] M. Allalouf and Y. Shavitt. Centralized and Distributed Approximation Algorithms for Routing and Weighted Max-Min Fair Bandwidth Allocation. IEEE Workshop on High Performance Switching and Routing, 2005.

Appendix – Related Work

The open and anonymous nature of popular on-line markets such as eBay [1] and Amazon.com [5] makes them susceptible to numerous adversarial activities. Reputation has been widely accepted as means of establishing trust among participants to prevent malicious use.

eBay

EBay employs a simple feedback-based trust system. After each transaction, both the buyer and the seller can rate each other by assigning one of three possible ratings: satisfactory, neutral, and unsatisfactory. These ratings build up a feedback score (total number of positive transactions minus the total number of negative transactions) which serves as an indicator of users' transaction histories; and, could represent a valuable warning to new users who wish to interact with rated users. The higher the feedback score, the higher the reputation of the user. In addition, for each user, eBay reports the number of her transactions both as a buyer and as a seller, the percentage of positive feedback, and the actual feedback by the users with whom the user interacted. Using such information, users can decide whether to participate in a transaction with a specific user and also evaluate its risk. In [17], Dellarocas presents a survey of game theoretic and economics models for reputation management. Dellarocas also showed that reputation systems, which are only based on the sum of negative and positive ratings, are vulnerable to unfair rating attacks.

Reputation in Peer-to-Peer Systems

Several techniques have been proposed to quantify reputation in peer-to-peer systems. A survey of several existing methods can be found in [14]. P2PRep proposed by Cornelli et al. [18], focuses on providing a framework for reputation management without giving an explicit definition of a trust metric. EigenTrust [12] presents a method to minimize the impact of malicious peers on the performance of feedback-based reputation systems. EigenTrust associates each peer with a global trust value calculated using the left principal eigenvector of a matrix of normalized local trust values. Thus, EigenTrust takes into account the transaction history of the user to get his/her reputation score, and uses such information to identify malicious users. EigenTrust also introduces the notion of transitive trust: if a peer A trusts any peer X, it would also trust the peers trusted by X. PeerTrust [13] computes reputation scores as a function of five factors: the feedback obtained from other peers, total number of transactions, credibility of the feedback source, transaction context factor, and community context factor. The transaction context factor helps model such transaction properties as the size of the transaction, category, and time-stamp, and the community context factor helps provide incentives to obtain feedback.