Title	Revisiting public reputation calculation in a personalized trust model
Author(s)	Cohen, Robin; Wang, Peng Fei; Hu, Zehong
Citation	Cohen, R., Wang, P. F., & Hu, Z. (2017). Revisiting public reputation calculation in a personalized trust model. CEUR Workshop Proceedings, 2154, 13-24.
Date	2018
URL	http://hdl.handle.net/10220/47941
Rights	© 2018 The Author(s). This paper was published in CEUR Workshop Proceedings and is made available as an electronic reprint (preprint) with permission of The Author(s). The published version is available at: [http://ceur-ws.org/Vol-2154/]. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper is prohibited and is subject to penalties under law.

Revisiting Public Reputation Calculation in a Personalized Trust Model

Robin Cohen
University of Waterloo
Waterloo, Canada
rcohen@uwaterloo.ca

Peng Fei Wang
University of Waterloo
Waterloo, Canada
pfwang@uwaterloo.ca

Zehong Hu
Nanyang Technological
University, Singapore
huze0004@e.ntu.edu.sg

Abstract

In this paper, we present a strategy for agents to predict the trustworthiness of other agents based on reports from peers, when the public reputation reflected by majority opinion may be suspect. We ground our discussion in the context of Zhang's personalized trust model, where agents combine estimates of both private and public reputation, tempered by a representation of the trustworthiness of the peers providing ratings. We propose a change in how public reputation is calculated: instead of valuing consistency with the majority, we instead value consistency with the opinion adopted by a cluster of peers, chosen by a likelihood probability. We are able to show that our change reduces the number of negative transactions experienced by a new agent subscribing to a minority opinion, during a learning phase before private reputation dominates the calculations. In all, we offer a method for calibrating the benefit of discriminating more carefully among the peers being consulted for the trust calculations. We contrast briefly with other approaches advocating for a clustering of the set of peer advisors, and discuss as well related work for dealing with the challenge of misleading majority opinion when performing trust modeling. We also comment on the usefulness of our approach for practitioners designing intelligent systems to act as partners with human users.

1 Introduction

In this paper we propose a method for modifying trust modeling algorithms which depend on a calculation of public opinion. The problem we are concerned with is that majority opinion will always dominate minority opinions if we only value consistency with the majority. Grounding our discussion in the context of the Personalized Trust Model developed by Zhang and Cohen [13][12], we propose an alternate method for integrating advisor ratings, one that leverages a clustering of stated opinions. We outline a few distinct variations of this method, each time showing the results of simulations which help to calibrate the benefit introduced when replacing the current formula for calculating public reputation. We end with a discussion of the contributions of our work, contrasting our approach with those of others who have considered clustering within trust modeling for multiagent system environments, and highlighting the benefits for practitioners designing systems to act as trustworthy partners with human users. As such, we offer some important insights into how to promote effective human-AI collaboration, in a way that will put users more at ease in accepting these intelligent agent partners.

2 Background

Within the artificial intelligence subfield of multiagent systems, a group of researchers are focused on how best to model the trustworthiness of peers. This is of value in settings such as e-commerce marketplaces, where intelligent agents may assist buyers in selecting reliable sellers, based in part on reports provided about those sellers from other buying agents. Determining the most trustworthy partners in a multiagent system is also critical for assembling the most effective team of agents to carry out actions on behalf of a human user. In essence, an agent's effectiveness in modeling the trustworthiness of its peers is central to humans trusting that agent to provide advice and to carry out actions on their behalf.

Zhang and Cohen [13] [12] develop a personalized approach to trust modeling, where agents have the flexibility of weighting the value of both private and public knowledge of peers, when estimating trustworthiness, used in deciding which peers to accept as partners (for example in electronic commerce transactions, where buyers are selecting the most trustworthy sellers to do business with). For the remainder of this paper, we refer to the agents being considered as partners, with their trustworthiness modeled, as sellers; this labelling is consistent with that used by Zhang in his papers, as his work was originally framed for e-marketplace scenarios. The methods that we propose here can apply equally well in any other contexts where trust modeling is employed by agents, in order to select among their peers.

Zhang's model is intended to address the issue of unfair ratings and is grounded in the concept of beta reputation functions [5], learning probabilistically whether an agent will be trustworthy, estimating whether a fair rating will likely be provided by that agent. Trust scores are calculated as $\frac{\alpha}{\alpha+\beta}$ where $\alpha = N_{pos} + 1$ and $\beta = N_{all} - N_{pos} + 1$, N_{pos} is the number of positive or consistent ratings and N_{all} is the number of total ratings. So as the number of positive or consistent ratings increase, so will the trustworthiness of the agent.

In Zhang's model, one examines the rating for a seller provided by a peer acting as an advisor (0 as untrustworthy and 1 as trustworthy), judged within a limited time window, and measures whether the rating provided by this peer is consistent with the majority of ratings of that seller provided by others within that timeframe. A central unit has access to all ratings provided, for this measurement. Zhang comments "Determining consistency with the majority of ratings can be achieved in a variety of ways, for instance averaging all the ratings and seeing if that is close to the advisor's rating, which is the method used in our experiments..." More precisely public reputation used in Zhang's experiments were calculated as:

```
for each item in set of all items
    num_pos = number of ratings that were positive
    num_neg = number of ratings that were negative
    if num_pos > num_neg
        majority_opinion = 1
    else
        majority_opinion = 0
    for each advisor with ratings of item
        mark rating as consistent if it equals majority_opinion
```

The consequence of this approach, is that new users who have a minority opinion will always have an untrustworthy public reputation of other users, who share the same minority opinion. They must count on their private reputation of these advisors to obtain accurate item ratings. Unfortunately it can take an extraordinary high amount of negative transactions to build up a sufficient private reputation to make this change.

3 Proposed Solution

In this section, we outline a modified version of Zhang's model. We begin with a motivating example, to highlight some challenges and proceed to demonstrate what the modified version offers as improvements for the trust modeling. Our first discussion centres on a baseline case, where agents exhibit no variability in their preferences.

3.1 Motivating Example

In this world there are two sellers selling different products. Each seller advertises their products as being good. Good in this world can be either high quality, or cheap (two different opinions). There are currently 5 advisors in the market, advisors 1-3 prefer high quality while 4-5 prefer cheap (so the high quality opinion dominates cheap). Seller 1 sells cheap items, while seller 2 sells high quality items. So advisors 1-3 would think seller 1 is untrustworthy and that seller 2 is trustworthy, while advisors 4-5 would think the opposite. Let each advisor have 64 ratings of each seller (representing a series of distinct buying decisions). So advisors 1-3 will each have 64 ratings of 0 for Seller 1 and each have 64 ratings of 1 for Seller 2. Advisors 4-5 will each have 64 ratings of 1 for Seller 1 and each have 64 ratings of 0 for Seller 2.

The public reputation of an advisor is calculated as $\frac{N_p+1}{N_{all}+2}$ where N_p is the number of consistent ratings and N_{all} is the number of all ratings by that advisor. To calculate the number of consistent ratings, we must find the majority rating for each seller, which is the opinion with the most number of ratings. Seller 1 has 128 positive ratings, and 192 negative ratings so the majority rating for Seller 1 is untrustworthy. Seller 2 has 192 positive ratings, and 128 negative ratings so the majority rating for Seller 1 is trustworthy.

Thus to find the number of consistent ratings we just match the ratings for each seller to the majority rating; if they are the same, then it is a consistent rating. Advisors 1-3 have 128 consistent ratings, giving each of these advisors a public reputation of 0.99231. Advisors 4-5 have 0 consistent ratings giving them a public reputation of 0.00769

The advisor and seller trust score calculations are done using the Zhang model, briefly summarized here. The trustworthiness of an advisor is a weighted average of their private and public reputation: $Tr(A) = wR_{pri}(A) + (1-w)R_{pub}(A)$. Private reputation of A can be calculated as follows: $\alpha = N_f + 1, \beta = N_u + 1$ $R_{pri}(A) = E(Pr(A)) = \frac{\alpha}{\alpha+\beta}$, where $Pr(A)$ is the probability that A will provide fair ratings to B, and $E(Pr(A))$ is the expected value of the probability. These values are determined by examining agreement between A and B when rating common sellers in the past. An additional element restricts comparisons to limited time windows. The public reputation of A is estimated based on its ratings and other ratings for the sellers rated by A. In a similar fashion, the overall trustworthiness of a seller is: $Tr(S) = wR_{pri}(S) + (1-w)R_{pub}(S)$ where the private reputation of the seller is estimated once more through the beta probability density functions, as $R_{pri}(s) = \frac{\sum_{i=1}^{n} N^b_{pos,i}\lambda^{i-1}+1}{\sum_{i=1}^{n}(N^b_{pos.i}+N^b_{neg,i})\lambda^{i-1}+2}$. The public reputation of a seller not only considers overall positive and negative evaluations but also discounts each advisor rating by the overall trustworthiness of that advisor, as calculated above. Dempster-Shafer theory is introduced to set the discounting of advisors. Details are provided in [13].

We now introduce buyer b, who prefers cheaper products. So b would prefer Seller 1 over Seller 2. However b is going to use the trust score to determine which seller they will buy from. When buyer b purchases something from Seller 2, they will always leave a negative rating. This will continue until Seller 1 is rated to be higher than Seller 2. Below we show the results of this scenario.

From Figure 1, which shows the relationship between the seller trust scores and the number of transactions, it can be seen that it takes 54 negative transactions before buyer b finally rates Seller 1 as more trustworthy. From Figure 2, which shows the change in advisor trust scores, we see that advisors 1-3 are still rated to be more trustworthy than advisors 4-5, even after 54 negative transactions.

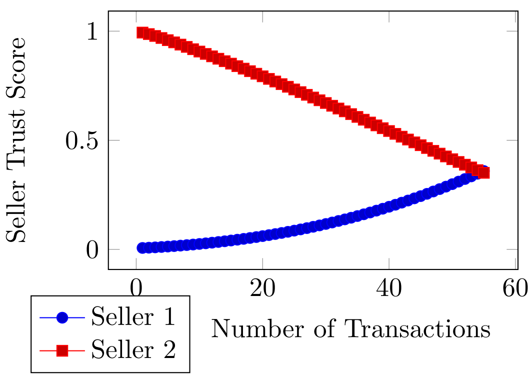

Figure 1: Seller Trust Score change of Zhang Model

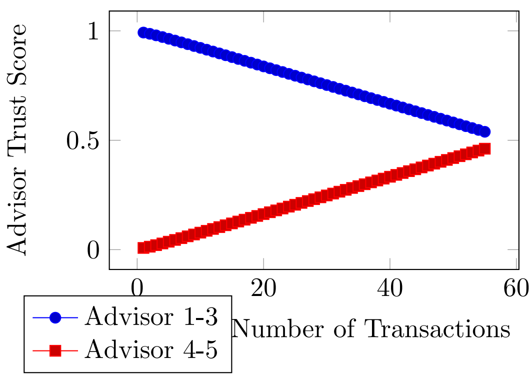

Figure 2: Advisor Trust Score change of Zhang Model

3.2 Redefining a consistent rating

The above scenario illustrates the issue of only recognizing the majority opinion as correct and not learning even when transactions show otherwise. To resolve this, we will redefine the concept of a consistent rating. A consistent rating will be defined as a rating that agrees with the majority rating in the chosen truthful cluster. The truthful cluster is an opinion cluster that has been chosen to be true for a particular advisor at time t. The opinion clusters are clusters of advisors that share the same opinion. They are found in this manner:

1. Create for each advisor A an opinion array defined as a N dimensional array where N is the number of sellers in the system, and the value of each dimension is the private reputation of $seller_n$ for advisor A, where the private reputation is defined as $\frac{N_{pos}+1}{N_{all}+2}$ [13]. An example of the opinion array created for advisors 1-5 is show in table 1.

2. Use the opinion array as the advisor's position in Euclidean space and cluster the advisors according to their position in space (in this example k-means clustering was used which tries to minimize the mean distance between the nodes and cluster centroids)

Table 1: Opinion Array of Advisors

	Seller 1	Seller 2
A_{1-3}	0.01515152	0.98484848
A_{4-5}	0.98484848	0.01515152

To illustrate how the private reputation is calculated for A_{1-3}:
For Seller 2, A_{1-3} have 64 positive ratings, and no negative ratings so private reputation $= \frac{64+1}{65+2} = 0.98484848$. Running k-means (k = 2) on the advisors will split advisors 1-3 into one cluster and 4-5 into another cluster. Cluster 1 will contain advisor 4-5 and Cluster 2 will contain advisor 1-3

The selection process of choosing a truthful cluster out of the opinion clusters is given as: **a)** Assign the initial probability of truth as the percent size of the cluster defined as $\frac{\text{number of advisors in the cluster}}{\text{total number of advisors}}$ **b)** Using the probabilities of truth of each cluster, select one of the clusters as 'truthful' **c)** Define consistent rating to be a rating that is the same as the majority rating of the members of the 'truthful' cluster **d)** Calculate reputation scores the same as previous **e)** If the transaction was marked as negative, reduce the 'truthful' clusters probability of truth by some factor (we used $e^{-0.3}$). So Cluster 1's initial probability of truth will be $2/5 = 0.4$ and 0.6 becomes 0.444490932409 when discounted by $e^{-0.3}$. Table 2 lists how the probability of truth can change for the clusters with each new transaction.

Table 2: Change in Probability of Truth

	Cluster 1	Cluster 2
T_0	0.4	0.6
T_1	0.555509067591	0.444490932409
T_2	0.670713018344	0.329286981656
...	...	...
T_5	0.945569228026	0.0544307719736

By breaking down the different opinions in a system, we get two different possible majority ratings seen in table 3. This new definition of what constitutes a consistent rating alters how public reputation is calculated, as now the minority opinion can be considered in the public reputation calculation. It allows opinions to be penalized if believing in this opinion results in a negative transaction, which will allow the public reputation to change in response to past transactions. If cluster 1 is chosen as the Truthful Cluster, then Seller 1 will have 128 positive ratings and Seller 2 will have 128 negative ratings. If cluster 2 is chosen as the Truthful Cluster, then Seller 1 will have 192 negative ratings and Seller 2 will have 192 positive ratings.

Table 3: Majority Rating for each Seller with Different Truthful Clusters

	# pos	# neg	Majority
S_1	128	0	Trustworthy
S_2	0	128	Untrustworthy

a) Cluster 1 Truthful

	# pos	# neg	Majority
S_1	0	192	Untrustworthy
S_2	192	0	Trustworthy

Cluster 2 Truthful

Applying this modified model on the previous scenario shows the improvements in redefining the concept of consistent rating. In Figure 3 and 4, we see the trust scores of the Sellers and Advisors swing back and forth as we switch between the 2 different clusters of truth, (cluster 1 favours Seller 1 and Advisors 4-5, while cluster 2 favours Seller 2 and Advisors 1-3). We see that as time goes on, the swings subside and more plateaus form where Seller 1 and Advisors 4-5 are declared to be more trustworthy. This happens because cluster 2's probability of truth is penalized due to the negative transactions, which is seen in Figure 5 where the Probability of Truth of Cluster 2 starts out high but decreases each time a negative transaction occurs. Figure 6, which shows which cluster is chosen as the cluster of truth, also shows how cluster 1 becomes the favoured cluster as more transactions occur. The end result is a reduction in the number of negative transactions. Figure 7 displays the number of negative transactions for 600 iterations of the scenario. The average number of negative transactions is 8.1, with a max of 11. This is significantly better than the 54 negative transactions of the original Zhang Model.

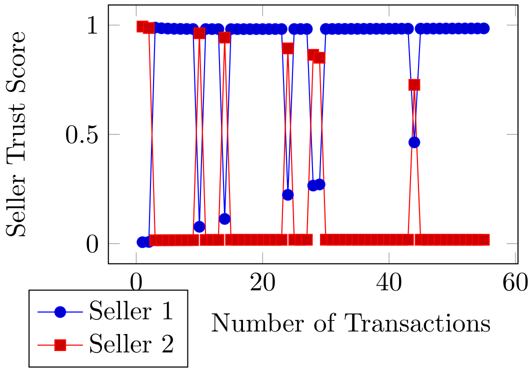

Figure 3: Seller Trust Score Change of Modified Zhang Model

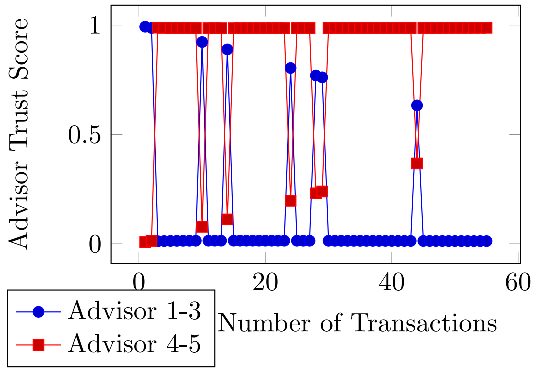

Figure 4: Advisor Trust Score Change of Modified Zhang Model

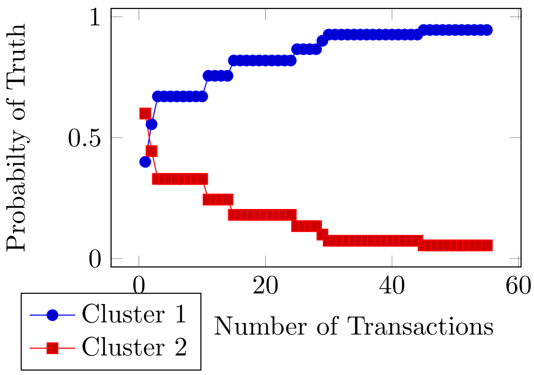

Figure 5: Probability of Truth Change Change of Modified Zhang Model

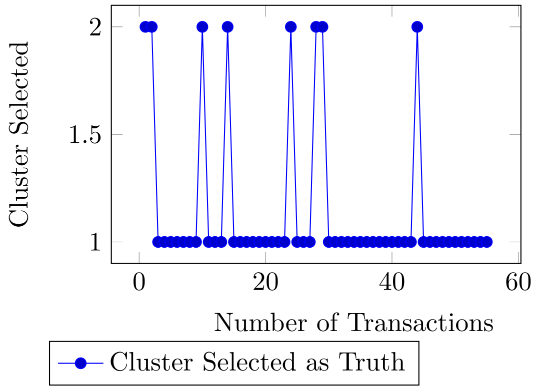

Figure 6: Cluster Selected of Modified Zhang Model

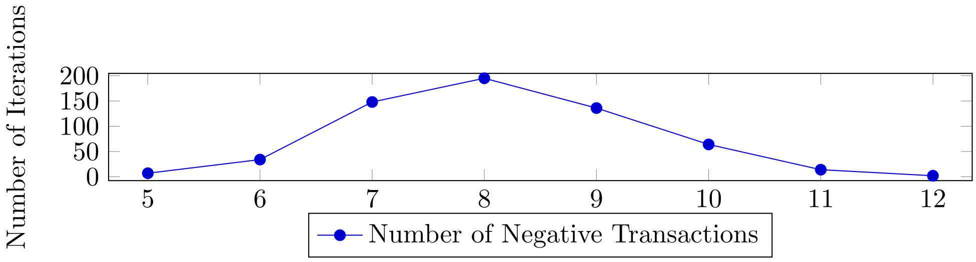

Figure 7: Distribution of Negative Transactions of Modified Zhang Cohen Model over 600 Iterations

4 Adjusting our Modified Model to Extended Scenarios

The obvious advantage that our approach offers is that buyers can now probabilistically accept minority opinions as truthful. With the previous example using this approach the buyer b will only make around 9 wrong purchases before they reach 54 ratings (where the private trust score takes over completely). This is a huge improvement over the 54 wrong purchases of the previous solution. So when the new buyer believes in a minority opinion this new approach will always perform better than the previous approach. A new weakness that our approach introduces is that buyers who believe in the majority opinion can now accept minority opinions as truthful. So previously a new buyer who holds a majority opinion will always get the right trust rating because the public reputation reflects their values. But in our current example setup, this new buyer will receive around 7 incorrect trust ratings before the 54 purchase limit. In the example setup we still see a total decrease in the number of incorrect trust ratings, but in other setups this might no longer be true. In order to examine the robustness of our approach to scenarios where agent behaviour is less consistent, we move on to a case where the same seller may sometimes satisfy and sometimes dissatisfy the buyer.

In this scenario, we will use the same values of the first 2-opinions scenario, but now the sellers have a 20% chance of doing the opposite. So Seller 1 has a 80% chance of satisfying Buyer b and 20% chance of not satisfying b, while Seller 2 has a 20% chance of satisfying Buyer b and 80% chance of not satisfying b. This is to simulate a more realistic scenario, as having the same opinion makes it more likely that Seller 2 will give Buyer b a positive transaction but it should not happen 100% of the time. Similarly while it is unlikely Seller 1 will satisfy Buyer b, it should still happen occasionally.

We first apply the Zhang model to this new scenario. From Figure 8 and 9, we see that under this scenario it takes longer for Buyer b to eventually rate Seller 1 as more trustworthy but the general shape of the graph is similar to the first scenario. Figure 10, shows the number of negative transactions for 600 iterations of the scenario. The average number of negative transactions is 80, with a max of 90 and a minimum of 66. This is an average increase of 26 negative transactions.

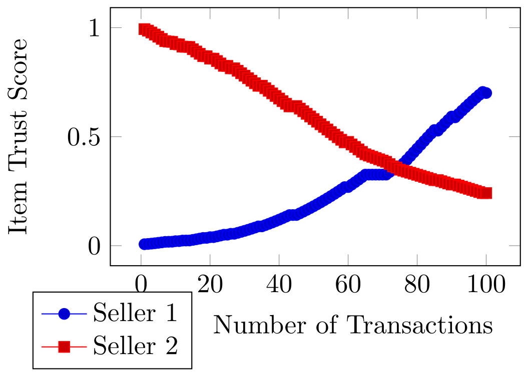

Figure 8: Seller Trust Score change of Zhang Model

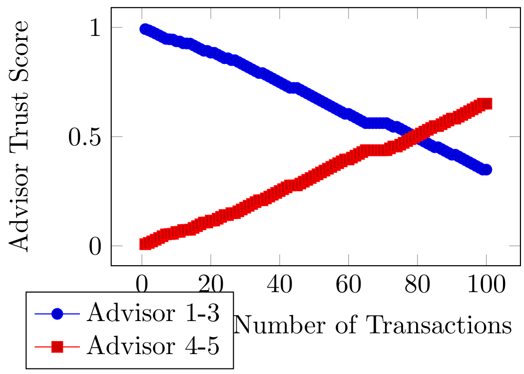

Figure 9: Advisor Trust Score change of Zhang Model

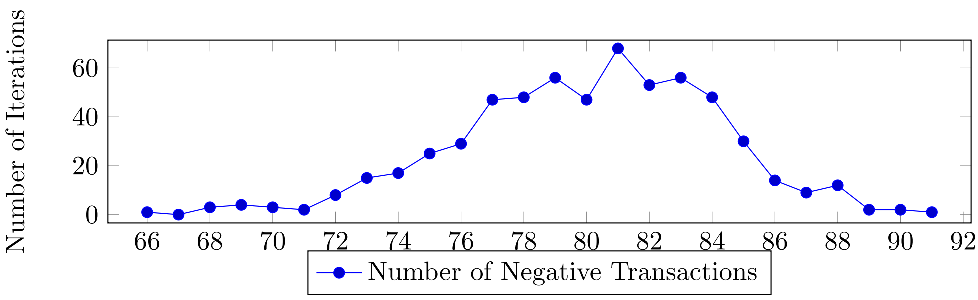

Figure 10: Distribution of Negative Transactions of New Model over 600 Iterations

When we applied our modified model to this new scenario, we noticed an issue. Advisor trust scores began to oscillate (Figure 11). The problem with only penalizing for negative transactions with no reward for positive transactions makes it challenging to properly separate the trustworthy opinion cluster from the untrustworthy one (Figure 12), once private reputation ends up dominating the trust scores (as both the wrong and the correct opinion cluster are no longer influencing the calculation).

To solve the problems encountered in the previous scenario 3 changes were made to the modified model.

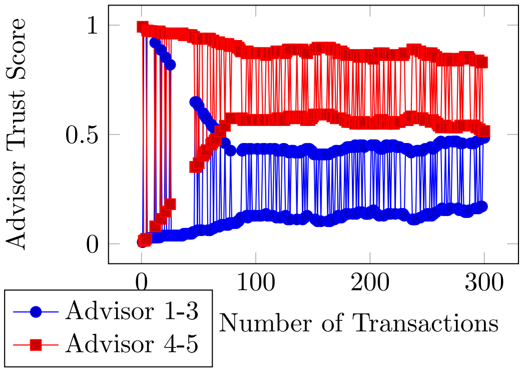

Figure 11: Advisor Trust Score Change of Modified Zhang Cohen Model

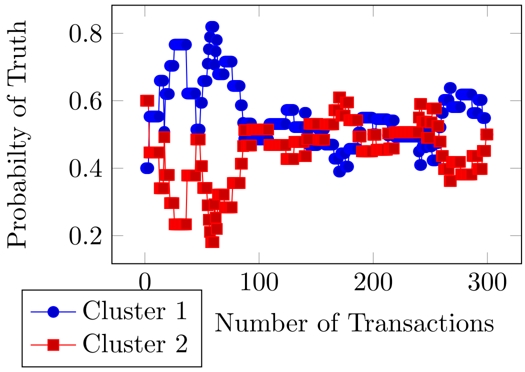

Figure 12: Probability of Truth Change of Modified Zhang Cohen Model

1. Instead of randomly choosing an opinion cluster using the probabilities of truth to be the truthful cluster, we will now pick the truthful cluster as the opinion cluster with the highest probability of truth (ie the most likely truthful cluster). This will reduce the oscillations that occur and make the trust model deterministic. Now cluster 2 will be picked 100% of the time as the truthful cluster rather than 60% of the time as before, because Cluster 1 will have a probability of truth of 0.4 while Cluster 2 will have a probability of truth of 0.6.

2. Reward clusters by increasing their probabilities of truth during positive transactions which will further help separate trustworthy and untrustworthy clusters.

3. Discount the change in the probabilities of truth by how much public reputation is used in determining the trust score of the seller. (ie when only private reputation is used to determine trust scores, probabilities of truth no longer change) So opinion clusters are not rewarded/punished when their information is not used in trust score calculations

Using this new model we will rerun the previous scenario. The figures below show that the modified model performs much better. Figures 13 and 14 show that the trust scores of the Sellers and Advisors no longer oscillate after an initial learning period, settling on trust scores that are representative of the underlying scenario. Notice that the modified model tends to rate the untrustworthy Seller to be lower than the original model; this is because under the modified model there are fewer private transactions with Seller 2 (what we wanted) which results in a lower trust score as the public trust score is lower than the private (no positive ratings vs 20 positive ratings). Figure 15 shows how the opinion clusters separate from each other and how after private reputation takes over, the clusters' probability of truth no longer changes. Figure 16 shows the cluster selection process as cluster 1 becomes the most likely truthful cluster over time. And Figure 17 displays the reduction of negative transactions.

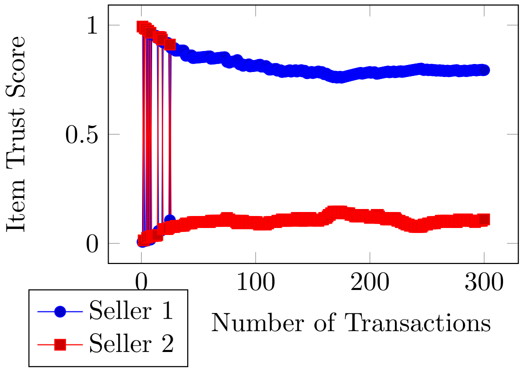

Figure 13: Seller Trust Score change of Modified Zhang Model

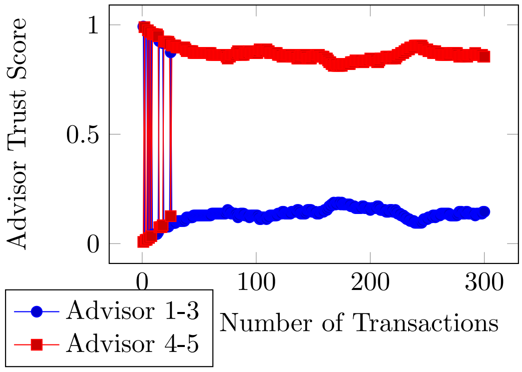

Figure 14: Advisor Trust Score Change of Modified Zhang Model

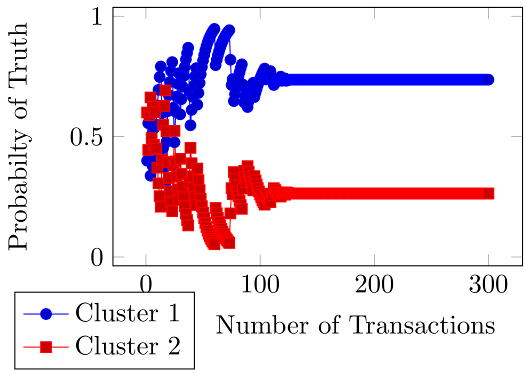

Figure 15: Probability of Truth Change of Modified Zhang Model

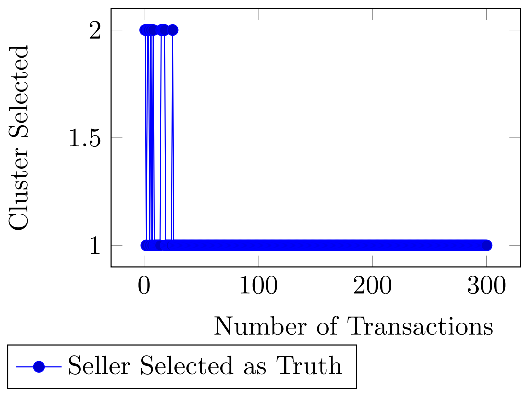

Figure 16: Seller Selected By Modified Zhang Model

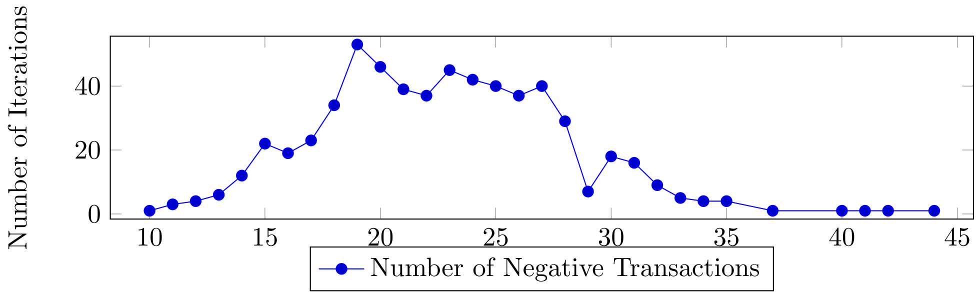

Figure 17: Distribution of Negative Transactions of Modified Model over 600 Iterations

5 Related Work on Cluster-Based Trust

Several other cluster based trust models also exist. The integrated clustering-based approach (iCLUB) by Liu is a trust model that is structured much like the personalized model, evaluating trust using either a private or public approach. iCLUB uses DBSCAN as the clustering algorithm of choice. DBSCAN is a density based clustering algorithm that automatically determines the number of clusters in the data. The ratings that each particular buyer B has of seller S is stored as a row vector with N elements, where N is the number of different rating levels (ie if one can rate from 1-5, the the row vector will have 5 elements). Each element will represent the number of ratings of that rating level that buyer B has submitted (ie if B submitted ratings: 5, 4, 4, 4, 2; then the vector will be [0,1,0,3,1]). iCLUB breaks down into a local approach and a global approach:

1. Local Approach:
   In the Local approach, buyer B is trying to evaluate seller S. Taking all the rating vectors of all advisors (and B) that have submitted a rating for S, cluster the rating vectors using DBSCAN. The advisors in the same cluster as the buyer B's rating vector will be marked as being fair while all other advisors in other clusters will be discarded as unfair.

2. Global Approach:
   In the Global approach, buyer B is trying to evaluate seller S, but has no transactions with S. First construct a set of trusted advisors, by performing the Local approach on sellers that B has interacted with, getting the trusted users for each seller, then taking the union of all the trusted advisors to form set W. Taking all advisors that have a transaction with S, cluster them using DBSCAN. The advisor cluster with the highest number of intersections with set W is taken as the trustworthy cluster, and the advisors in this cluster are trusted. If B has no ratings at all, then the largest cluster is taken to be the trustworthy cluster.

By integrating both approaches, we arrive at iCLUB. When buyer B is evaluating seller S, let ϵ be a threshold defined by the particular instance of the model. Use the local approach if the number of ratings made by B reviewing S, is greater than ϵ, otherwise use the global approach.

We see that iCLUB uses clustering with the same motivation as our modified approached. They both seek to exploit the fact that advisors that evaluate sellers in a similar fashion will see each other as more trustworthy as well and depend on the fact that clustering algorithms can reliability determine similar advisors given their rating vectors. The main difference between iCLUB and our modified approach is the type of buyer that the trust model is useful for. Each buyer goes through 3 stages: **1:** The first stage is when the buyer initially enters the network, which means they have no ratings with any and all sellers **2:** The second stage is when the buyer has had a meaningful number of transactions with other sellers but do not with a particular seller that the buyer is evaluating. **3:** The third and final stage is when the buyer has had a meaningful number of transactions with the particular seller the buyer is trying to evaluate. Looking at the three different stages of buyers, it becomes clear that iCLUB is mostly only valid for buyers in the second and third stage. The reason for this is when a completely new buyer enters the market, iCLUB will first simply direct them to the majority which is susceptible to majority opinion as discussed in this paper until they reach the specified ϵ. The strength of iCLUB only emerges once the buyer has had some meaningful transactions and thus has an initial network of trusted users. Meanwhile buyers in the first stage where iCLUB is lacking, are precisely the type of buyers we are targeting with our modified approach. Our clustering approach does not assume any linkages between the buyer and advisors and tries to find the most trustworthy cluster through a small number of transactions made by the buyer.

The Biculus [4] clustering model uses a biclustering method with the main goal of detecting which advisors are fair and unfair in a rating system that uses multiple criteria (ie like TripAdvisor where you can rate based on room, service, location, value, etc). Biculus uses 2 main identifiers to determine if an advisor is fair or unfair, the first identifier is distance from what they defined to be the central buyer, and correlations between criteria. By performing biclustering using a certain distance measure between the rating vector of seller S by buyer B and those of this seller by the advisors B, certain advisors are filtered out as untrustworthy. Advisors are also rejected as untrustworthy when a correlation difference between advisors and a buyer according to a set of selected seller features are greater than a threshold.

Using these 2 different methods of detection, Biculus has shown that it is capable of defeating a variety of different attacks by identifying untrustworthy advisors. Biculus also uses the idea that advisors that rate sellers similarly are more trustworthy, and expands it beyond iCLUB by splitting the clusters so it can consider each criteria separately and together. In both simulated experiments and real experiments Biculus has shown itself to be more robust in detecting untrustworthy advisors than iCLUB, but even so, Biculus does not consider buyers in the first stage and does not have a mechanism that addresses these buyers.

The two clustering models above both exhibit the use of local data to form clusters instead of global data. Local data in this context refer to ratings of only the seller in question while global data refers to the ratings of all sellers. iCLUB for instance finds trusted advisors locally, by clusters based on the ratings of a particular seller and selecting the cluster containing the buyers rating vector as trustworthy. Then to expand to a global context, iCLUB selects clusters of advisors found locally through a measure of intersectionality with the trusted local advisors selected from other sellers. Our modified approach on the other hand tries to identify global clusters by clustering advisors based on all of their ratings. The intuition here is that an advisor will rate different sellers based on the same criteria and so using all of an advisors ratings, we can form more accurate clusters. The problem is that a) an advisor could rate sellers based on different criteria b) it is very computationally exhaustive to maintain global clusters as the advisors interact through more transactions.

To correct these issues, we could consider shifting our approach to clusters based on local data instead. In this modification, when buyer B wants to evaluate seller S, B will obtain the ratings of all advisors who have rated S and construct their rating vector as levels of the ratings with each dimension being the number of ratings in the level. The buyer will then cluster the advisors based on their rating vectors and cluster the advisors using a clustering algorithm like DBSCAN. The buyer will then select clusters and evaluate clusters the same way as the previous modifications. The problem that arises from this is that each buyer has to deal with the problem of reconciling the old clusters that they have with possible new clusters if more ratings as time goes on creates new clusters or changes old clusters and that purchases between different sellers no longer give information about the respective advisors. We present our extended model below.

6 Dynamic Settings

The previous simulations provided a view of how new buyers will act in the face of a very static marketplace where the other advisors do not change their ratings nor participate in any new transactions. We seek to improve this simulation by allowing the advisors to participate in further transactions and even change their behaviour. In this experimental set up there are 60 initial buyers in the market who will act as the advisors and 2 sellers (seller 2 and seller 1). The buyers are divided into 2 groups, group 1 which has 40 buyers will prefer seller 2 over seller 1, while group 2 which has 20 buyers will prefer seller 1 over seller 2. The sellers will satisfy their preferred group 80% of the time, and the other group 20% of the time. Each buyer will be seeded with an initial 100 ratings choosing uniformly randomly between seller 1 and 2. After the initial setup, we introduce the buyer B, who has no rating history, and prefers seller 1. At this point we increment time in distinct timesteps where every agent in the system makes a transaction with the seller they determine is most trustworthy. The 60 advisors will use the personalized trust model to calculate trust scores because they have sufficient experience in the network, while the buyer B will use the modified personalized approach because he has history in the network.

Figure 18 which shows the seller trust scores plotted over the number of transactions for this new system. We see from figure 18 that buyer B quickly switches from Seller 2 to Seller 1, and while there are dips, the trust score of Seller 1 eventually approaches 0.8, while the reputation of Seller 2 approaches 0.4. Notice that Seller 2 now has a higher trust score than the static simulation this is likely because the public reputation scores are not as extreme as Seller 2 still fulfils some of the transactions to those who do not prefer Seller 2. Figure 19, which displays the probabilities of truth for the two clusters still look very similar to the same diagram for the static simulation. But due to more randomness we see that the separation of the truthfulness of the two clusters takes a much longer time to occur. Figure 20 shows that while the separation of the numeric truthfulness score between the clusters takes a longer time to occur, cluster 1 is still selected as the truthful cluster very quickly.

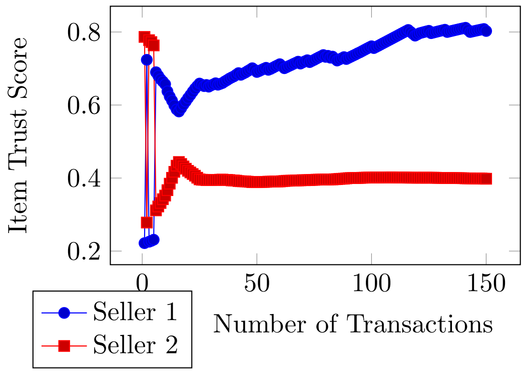

Figure 18: Seller Trust Score change of modified Zhang Model under improved simulations

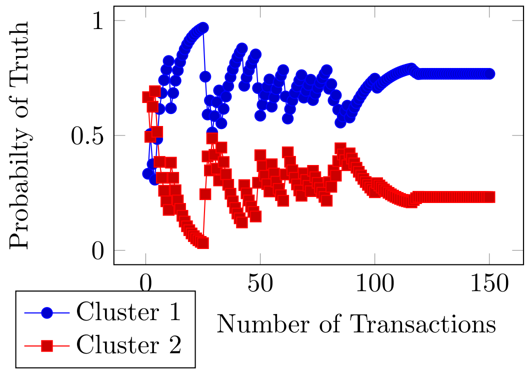

Figure 19: Probability of Truth Change of Modified Zhang Model under improved simulations

Here we see the cluster chosen as truthful, and we still see that quick change from cluster 2 to cluster 1.

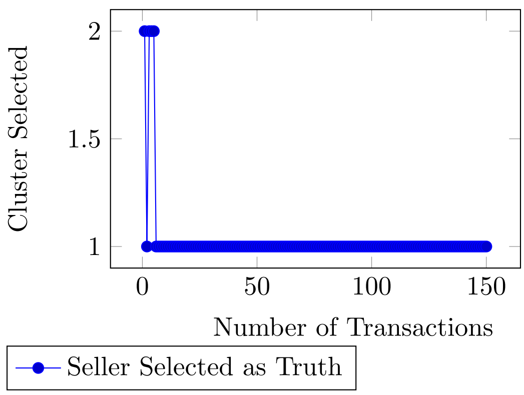

Figure 20: Seller Selected By Modified Zhang Model under improved simulations

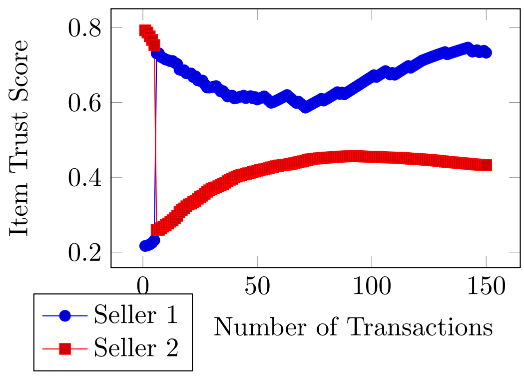

Figure 21: Seller Trust Score change of local modified Zhang Model under improved simulations

Next we will use the local clustering approach instead in this simulation. The local approach is implicated around the idea that clustering all users using all items is too computationally intensive so we will only cluster the relevant users of the item we are rating instead. To do this, we assign a cluster trust score to each user initialized to 1. When we are evaluating a particular item, we will cluster the users who have rated this item according to the private score the user assigns to the item. To calculate the truth score of the cluster, we simply add the cluster trust score of each user in each cluster, and the highest scoring cluster is the most likely trustworthy cluster. After the transaction succeeds, a successful transaction will result in an increase in the cluster trust score of each user in the truthful cluster, on a failure the trust scores of all the users will be decreased. Figure 21 shows the trust score change of the items over time using the local approach. Since all advisors rate all items in this simulation (this would not be true in the real world) we would expect this diagram to actually perform similarly to the global approach. It seems it takes longer to switch clusters, even though it still happens relatively fast, and an initial dip in trust score occurs for much longer. More investigation is needed to verify if the dip is due to randomness or if is an error in the simulation.

7 Discussion and Conclusion

There has been a wealth of literature to date from trust modeling researchers within the multiagent systems community. We chose to focus on the model of Zhang and Cohen in our study, because it is grounded in the prevalent paradigm within the community, that of beta reputation functions and probabilistic reasoning about expected trustworthiness of peers (as first proposed in the BRS system [5] and later expanded upon with greater focus on the role of advisors within the TRAVOS system [11]. The model of Zhang and Cohen emphasizes well the value of blending private and public reputation calculations effectively.

How best to capture the public reputation of agents is another topic of considerable debate. Some interesting work has proposed, for instance, coping with folklore that may be promoted by an ignorant majority, by using credibility values to recalibrate the alpha and beta parameters of the beta reputation function [8]. Zhang himself has worked with colleagues in order to produce more sophisticated variations of his original model, including work that leverages clusters in order to improve the public reputation calculation, as cited in this paper already [4][7].

A useful way of viewing the approach we have outlined in this paper is as a method for understanding, at a high level, the differences introduced into the experiences of agents within a multiagent system, when their trust calculations are dependent on a consistent majority. The simulations provided and the results produced from our realworld studies generate output of the extra burden assumed by agents while the trust calculations are not yet adjusted to the ideal values. Practitioners can employ these methods within their own systems and focusing on their preferred parameters, in a fairly straightforward manner (without having to encode more complex variations for the clustering solutions, for example, as outlined in [4][7]).

We acknowledge a concern with collusion in multiagent systems, whereby groups of agents make an effort to display consistent behaviour in an effort to eventually cheat or mislead other agents in the environment. Kerr [6] has in fact argued that clustering based methods may be leveraged in order to analyze agent behaviour, to detect coalitions of self-promoting agents. It would be interesting to examine as future work how various cluster-based trust modeling approaches, including ours, affect strategies for addressing collusion, as well.

For future work, immediate steps include exploring other options for the clustering solutions. For example, we could integrate a linear increase in a cluster's truth probability when a truthful transaction occurs. We would also lock a chosen cluster as "truthful" for all future transactions until an untruthful transaction occurs. We are also planning to project our model into experiments with larger number of agents and possibly operating with realworld data (for example from Epinions): these scenarios typically have issues with sparsity of ratings and with trust relations between peers being undeclared. The work of Etuk [1] may be valuable in providing insights into how to run tests in realworld contexts; the authors also advocate the use of clustering and may suggest adjustments to our own model as well. In fact, we view the application of trust modeling to settings of online social networks as an especially valuable direction for future work. Golbeck and Kuter [3] discuss the term social trust where majority opinion becomes especially critical. Misinformation in these networks is a particularly topical concern, with various models offering insights into how to cope [2][9]. Our exploration of majority opinion was in fact a precursor to a more general interest in adjusting trust models due to the presence of leading influencers (agents with special status in the majority opinion). It would also be interesting for future work to study related work on how best to elicit opinions from others in a way that measures trust with proper care. The approach of Simari [10] has some measures that are similar to ours, though employed for a different context (ontological query answering); the challenges explored by these authors may help to inform directions forward for our research.

In all, we offer here a step forward in the discussion of majority opinion as part of trust modeling solutions, one that suggests alternate heuristics for the calculations. The simulations provided here provide calibrations, allowing practitioners to prefer a longer or shorter learning curve with their agents, depending on the nature of the transactions and the importance of analyzing the base of peers in full detail. Our approach is especially valuable for initial phases where agents lack experience with other agents in the community. We believe that promoting the use of trust modeling within agent-based solutions for human users will have the effect of engendering more trust in these intelligent agent assistants, as well. In this respect, our work contributes to this ongoing effort.

References

[1] Anthony Etuk et al. "Strategies for Truth Discovery Under Resource Constraints". In: *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*. AAMAS '15. 2015.

[2] J. Golbeck. "Trust and Nuanced Profile Similarity in Online Social Networks". In: *ACM Trans. Web* 3.4 (2009), 12:1–12:33.

[3] J. Golbeck and U. Kuter. "The ripple effect: change in trust and its impact over a social network". In: *Computing with Social Trust*. Springer, 2009, pp. 169–181.

[4] A. Irissappane and J. Zhang. "Filtering unfair ratings from dishonest advisors in multi-criteria e-markets: a biclustering-based approach". In: *Autonomous Agents and Multi-Agent Systems* 31.1 (2017), pp. 36–65.

[5] A. Josang and R. Ismail. "The beta reputation system". In: *Proceedings of the 15th bled electronic commerce conference*. 2002.

[6] R. Kerr and R. Cohen. "Detecting and identifying coalitions". In: *Proc. AAMAS 2012*. 2012.

[7] S. Liu et al. "An integrated clustering-based approach to filtering unfair multi-nominal testimonies". In: *Comput. Intell.* 30.2 (2014), pp. 316–341.

[8] N. Sardana and R. Cohen. "Modeling Agent Trustworthiness with Credibility for Message Recommendation in Social Networks". In: *Proc. AAMAS 2014*. 2014.

[9] N. Sardana and R. Cohen. "Validating trust models against realworld data sets". In: *Proc. PST 2014*. 2014.

[10] Gerardo I. Simari et al. *Ontology-Based Data Access Leveraging Subjective Reports*. Springer Briefs in Computer Science. Springer, 2017. ISBN: 978-3-319-65228-3. DOI: 10.1007/978-3-319-65229-0. URL: https://doi.org/10.1007/978-3-319-65229-0.

[11] W. Teacy et al. "Travos: Trust and reputation in the context of inaccurate information sources". In: *Autonomous Agents and Multi-Agent Systems* 12.2 (2006), pp. 183–198.

[12] J. Zhang and R. Cohen. "Design of a mechanism for promoting honesty in e-marketplaces". In: *Proc. AAAI 2007*. 2007.

[13] J. Zhang and R. Cohen. "Evaluating the trustworthiness of advice about seller agents in e-marketplaces: A personalized approach". In: *Electronic Commerce Research and Applications* 7.3 (2008), pp. 330–340.