

# Improving the Use of Advisor Networks for Multi-Agent Trust Modelling

Joshua Gerner, Jie Zhang<sup>†</sup>, Robin Cohen

David R. Cheriton School of Computer Science, University of Waterloo, Canada

<sup>†</sup>School of Computer Engineering, Nanyang Technological University, Singapore

Email: jgerner@uwaterloo.ca, †zhangj@ntu.edu.sg, rcohen@uwaterloo.ca

**Abstract**—This paper provides an approach for improving the trust modelling of users when a social network of advisors is employed (for example when advisors are recommending the most trustworthy service providers). We present three important improvements to trust modelling, two directly relating to the size of the network (through either the use of a threshold or by setting a maximum network size) and a third (advisor referrals) which focuses on ensuring that the advisors have attained an appropriate level of expertise, for the advice that they provide. Experimental results confirm the value of our methods when choosing parameters in a principled manner, leading to improvements in the accuracy of trust modelling (shown in the context of electronic marketplaces). In all, this research provides insights into how to set the size and composition of social networks, in order to effectively integrate the advice of peers when modelling the trust of users.

**Index Terms**—Multi-Agent Systems, Trust Modelling, Social Networks, E-Commerce applications

## I. INTRODUCTION

In environments where users seek the advice of a social network of peers in order to make decisions, it is often important to be modelling the trustworthiness of those peers. This scenario of user modelling is typically framed as a multi-agent system, where each user is represented by an agent, and then the agents perform trust modelling, in order to propose effective actions for users to take. Current research in trust modelling using social networks of advisors (e.g., for electronic marketplaces) has traditionally proposed methods for limiting the number of peers that are consulted, to retain only the most trustworthy advisors. In general, however, these methods do not typically offer a principled methodology to set the values that serve to restrict the size of the social network that is consulted.

In this paper, we propose an approach for setting the size and composition of social networks of advisors. Our overall proposal is for social networks to be restricted in size initially and then adjusted in order to ensure that all peers have suitable expertise to offer needed advice. This includes a particular approach for integrating what we refer to as advisor referrals to replace peers with insufficient experience. This is intended to respect the advice of [1] that networks should not be so small as to exclude many experienced advisors but not too large as to be open to untrustworthy advice that reduces overall accuracy.

The methods that we propose to restrict the size of networks are either through a maximum number of peers or

through some threshold of trustworthiness. In contrast with other researchers, however, we investigate how best to set the values of these parameters empirically. In particular, we demonstrate how both with varying numbers of untrustworthy sellers and with varying size of e-marketplaces, our particular approach for setting the size of the social network offers trust modelling improvements that lead to the selection of more appropriate business partners. We are able to show as well how the integration of our approach for using advisor referrals to maintain appropriate experience levels of peers provides additional benefits to improve the modelling of agent trust. We conclude with a reflection on how our research contrasts with that used in collaborative filtering recommender systems and with other approaches for setting network size, and propose some directions for future research.

## II. IMPROVING TRUST MODEL ACCURACY

One significant trust model that serves as an appropriate context for our research is the Personalized Trust Model developed by Zhang [2]. In this section, we summarize this model and then use it as the initial backdrop for subsequent experimentation in order to demonstrate the value of our particular approach.

A buyer agent, denoted by  $b$ , first constructs a measure of private reputation of each advisor  $a$ , based on that advisor's ratings for sellers that  $b$  has previously dealt with, and representing an estimation of the probability that  $a$  will give fair ratings to  $b$ , using the beta family of probability density functions. In a similar fashion, the buyer then calculates the public reputation of an advisor  $a$ , or the probability that an advisor will provide "consistent" ratings. The overall trustworthiness of the advisor is then calculated as a weighted sum of the private and public reputations, based on the number of interactions between  $a$  and  $b$  (more interactions means that the private reputation will receive a greater weight). The overall trustworthiness of a seller is also calculated based on the buyer's private experience with the seller and also on the buyer's view of the seller's public reputation, where now the advice from each advisor is discounted, based on its trust value; the most trustworthy seller is considered as recommended by the advisors and thus selected by the buyer.

The original PTM did not state any restrictions on the size of the advisor network, which would imply that all advisors should be included, regardless of how trustworthy the buyer

had modelled them to be. Although the evaluation in [2] ultimately took into account only “trustworthy” advisors, i.e. those trusted by the buyer with a value of 0.5 or above, there remained some potential for improvement. Thus we are proposing three modifications to the original model that would optimize the size and composition of the advisor network to yield a more accurate trust model. Specifically, we suggest that the size of the advisor network be restricted using one of two methods: either (a) setting a maximum number of advisors (or *max\_nbors*, also known as *k*-nearest neighbours or *k*NN) to be included in each buyer’s advisor network, and taking those with the highest trustworthiness, or (b) setting a trustworthiness threshold, such that only those advisors that the buyer trusts at or above the given threshold would be included in the advisor network. Care must be taken in both cases, however, to ensure that the network is not so small as to exclude many experienced advisors, and not so large as to include many of the untrustworthy advisors [1].

The third method, advisor referrals, which assumes the prior use of one of the other two methods, would allow the system to make use of a wider range of experience throughout the population of advisors even if the chosen network is relatively small. In essence this mechanism works as follows: Having restricted the advisor network to some size *n*, and given a seller *s* that is under consideration, we seek to make use of *n* advisors that have each had some previously-determined number of past experiences with (i.e., purchases from) *s*. Any agent from the original advisor network that has met that level of experience will be used. If a particular advisor had not met this level, however, the system will examine the advisor’s own advisor network for a replacement advisor that has had a sufficient number of experiences with *s*. If no suitable advisor can be found, the advisor networks of the advisors just considered will be examined, and so on until either an acceptable replacement advisor is found, or a pre-determined maximum search level is exceeded (in which case the size of the advisor set used is reduced by one). If a replacement advisor is used, the trust assigned to it is the same value previously calculated by the buyer when it was initially determining which advisors to include in its advisor network.

We now adopt the simulation approach taken in [2] to verify the effectiveness of our proposals. Specifically we simulate an environment consisting of one buyer, 80 advisors, and 100 sellers, where the sellers are evenly divided into ten groups, each having a probability of dishonesty between zero and 0.9. The buyer and the advisors each randomly selects and rates one seller for each of the 80 days of the simulation, such that no seller is rated more than once by a single buyer or advisor. Finally, given these ratings, the buyer calculates the trustworthiness values corresponding to each of the sellers. These tests are performed for two values of the percentage of lying (dishonest) advisors (or LA), specifically 30% and 60%, and repeated a total of ten times for each possible combination.

These results are shown in Figure 1. In these figures, the *x*-axis represents the various values of the applicable network-

limiting parameter, either *max\_nbors* or thresholding (note that throughout this paper, the *x*-axes of these and other graphs are not necessarily to scale). The *y*-axis, meanwhile, indicates the mean absolute error (or MAE) for each simulation – that is, the average of the absolute differences between the trust value calculated for each seller and an “ideal” trust model (error = 0), in which the trust assigned to each seller corresponds exactly to the level expected given the indicated probability of dishonesty. For example, a seller that was assigned a 0.3 probability of dishonesty would ideally be trusted with value 0.7; if the simulation generated instead a trust value of 0.65 for that seller, there would be an error of 0.05 with respect to that seller.

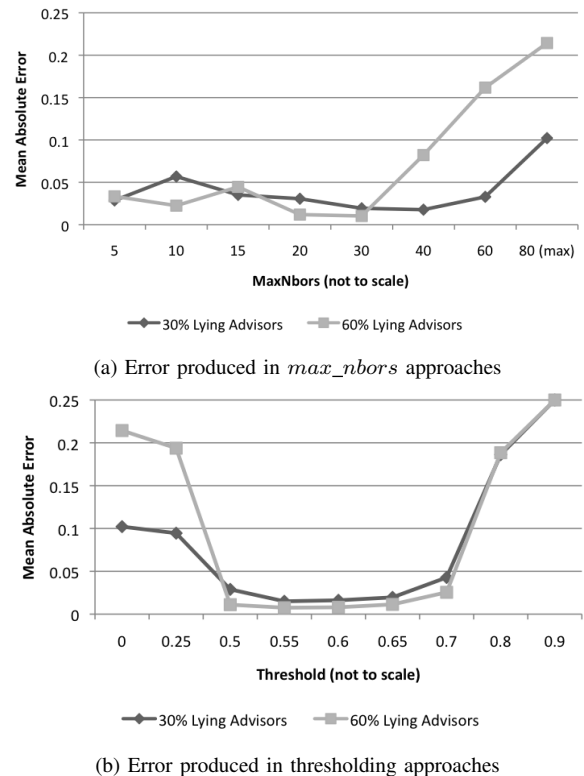


Fig. 1: Mean absolute error of various variants of PTM using *max\_nbors* or thresholding

As shown in these figures, the choice of parameter in either case is important; we generally found the best results with a “middle-of-the-road” parameter (e.g. a threshold of 0.55, or limiting the number of advisors to 40 out of a population of 80), as opposed to one that was either too relaxed or too restrictive.<sup>1</sup> Of these two options, thresholding performs better in two aspects. First, with an MAE of 0.015 in the 30% lying advisors scenario, and an MAE of 0.0075 with 60% lying advisors, using a threshold of 0.55 performed better than any of the *max\_nbors* parameters tested. Second, thresholding is not affected significantly by changes in the number of lying advisors, whereas in some instances performance dropped for

<sup>1</sup>This is in contrast to Zhang’s proposed arbitrary threshold of 0.5 [2].

maximum-number cases as the percentage of lying advisors was increased from 30% to 60%.

Although our tests were not exhaustive with regards to the possible choices of either the  $max\_nbors$  or threshold parameters, thresholding nevertheless appears to be the superior method at this stage. One explanation for these results is that setting a  $max\_nbors$  value makes the advisor network more likely to include untrustworthy advisors – even if the advisors are the  $k$  “best” advisors, not all of them will necessarily be trustworthy, particularly in the 60% lying advisor scenario. This in turn will reduce the accuracy of the seller trust model. By contrast, using thresholding will only include advisors that have reached or exceeded the applicable threshold, regardless of how many there are; it seems that this will lead in turn to more accurate results.

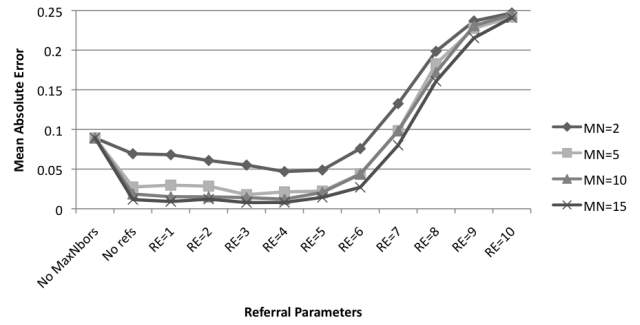
We note that in certain cases, including most of the threshold cases, the error for the 60% lying advisors case will be less than that seen for 30% lying advisors, given the same threshold or  $max\_nbors$  value. While one would expect it to be more difficult to avoid error in the case where 60% of the advisors are lying, our methods in fact turn out to be even more valuable, in these cases where the need to address possible inaccuracy is greater.

We now modify our simulation procedure slightly to test the use of advisor referrals. The parameters and test conditions are the same, except that we reduce the number of sellers to 40, and increase the number of simulation days to 120. We also adopt pure random selection for the sellers, such that buyers rate each seller a variable number of times (on average three), whereas previously they could rate each seller at most once. As noted above, in order to make use of referrals, we must set some value that represents the minimum number of experiences (i.e., purchases) that an advisor has had with the seller under consideration in order to be included in the trustworthiness calculations for that seller. In the simulations that follow, we refer to this value as the minimum “referral experience” (or  $RE$ ).

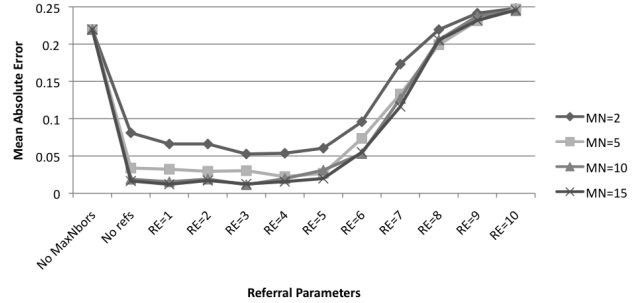
These results are shown in Figure 2, which displays similar data to that in Figure 1. In this case, Figures 2a and 2b show how the accuracy of the trust model is affected when a maximum number of advisors is set and the  $RE$  value is varied, with each plotted series representing a single  $max\_nbors$  setting. Likewise, Figures 2c and 2d show similar results when a threshold is applied and the  $RE$  value is varied; each series here represents a single threshold parameter.

Our results indicate that referrals offer fairly modest benefits when the number of advisors had already been optimized using one of the other two methods (for example,  $max\_nbors = 15$  or  $threshold = 0.55$ ). However, our evaluation indicates that referrals can serve to improve the accuracy of the trust model if the size of the advisor network is even more limited, i.e. a very low maximum number of advisors, or a very high trustworthiness threshold. In these cases, referrals will be successful in making use of advisors in the system relevant to the seller being examined.

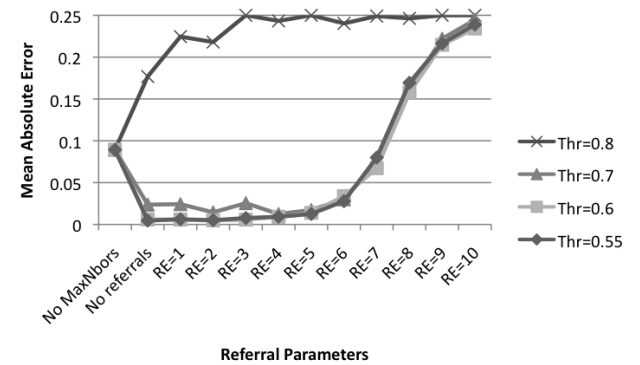
For example, while using  $max\_nbors = 2$  without referrals



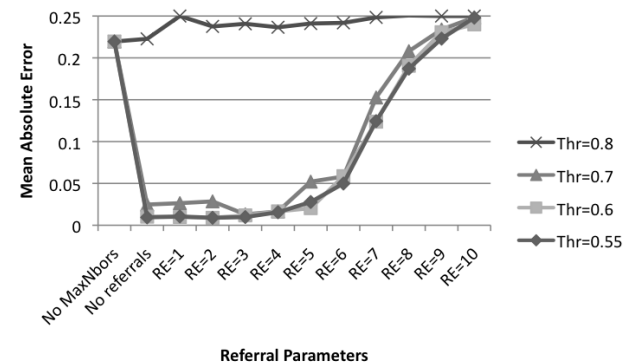
(a) Error comparison for various  $max\_nbors$  + referrals approaches with 30% LA



(b) Error comparison for various  $max\_nbors$  + referrals approaches with 60% LA



(c) Error comparison for various thresholding + referrals approaches with 30% LA



(d) Error comparison for various thresholding + referrals approaches with 60% LA

Fig. 2: Evaluation of effects of referrals on small advisor networks for PTM.

in the 60% lying advisors had an MAE of 0.081, allowing for referrals with an  $RE$  value of 4 led to a significantly smaller MAE of 0.054. Although this does not overcome the benefits of using a larger  $max\_nbors$  value – for example, setting  $max\_nbors = 5$  without using referrals resulted in an MAE of 0.034 – these results are still much closer in terms of accuracy.

### III. APPLICABILITY TO ALTERNATIVE MODEL

To demonstrate the generality of our approach, we also apply it to another probabilistic trust model, TRAVOS [3]. Under the TRAVOS approach, an agent models trust by using a beta pdf to determine the expected value of the probability that an interaction itself and another agent will be completed successfully – i.e., that the contract is fulfilled – given the set of outcomes of the past interactions between the agents. The agent then uses a separate metric to measure its confidence in the trust value just computed. If the confidence is not sufficiently high, the advice of third parties may be considered as well, although some of these opinions may be filtered out if the reputations of the applicable advisors are not themselves sufficiently high. However, there are two important distinctions between PTM and TRAVOS: PTM uses both private and public knowledge regarding all sellers, whereas TRAVOS uses only the private knowledge regarding some selected sellers. Moreover, the method used by TRAVOS to aggregate ratings provided by certain advisors is more complex, reducing the effect of ratings from less trustworthy advisors.

To demonstrate the effectiveness of our optimizations, we perform similar sets of experiments to those performed in Section II. These simulations for the  $max\_nbors$  and thresholding optimizations use an environment consisting of one buyer, 80 advisors, and 100 sellers with varying probabilities of dishonesty. During the simulation, the buyer and each advisor both randomly select and rate a total of 80 sellers. Finally, the buyer calculates the trustworthiness values corresponding to each of the sellers. These tests are performed for two values of the percentage of lying advisors, 30% and 60%. The results of these experiments are shown in Figure 3. Each figure shows two graphs, indicating how each model performs for both of the tested levels of lying advisors; as with the graphs shown previously, the data points map the applicable  $max\_nbors$  or threshold parameter on the  $x$ -axis to the mean absolute error (MAE) of the trust model under that scenario on the  $y$ -axis.

These figures indicate mixed results with regards to the effect of applying these modifications to TRAVOS. Consider that an unrestricted network (as represented by the far right of the  $max\_nbors$  figure, or the far left of the thresholding figure, within Figure 3) will yield a mean absolute error value between 0.15 and 0.25, and hence relatively low accuracy. In comparison, most of the models incorporating either  $max\_nbors$  or a threshold will have a smaller error value, and thus improved accuracy over an unrestricted network.

We notice that in Figure 3, the graphs representing the TRAVOS model have a zig-zag shape. Apart from possible inaccuracies in TRAVOS’s modelling of advisor trustworthiness, one theory for this performance, based on our examination

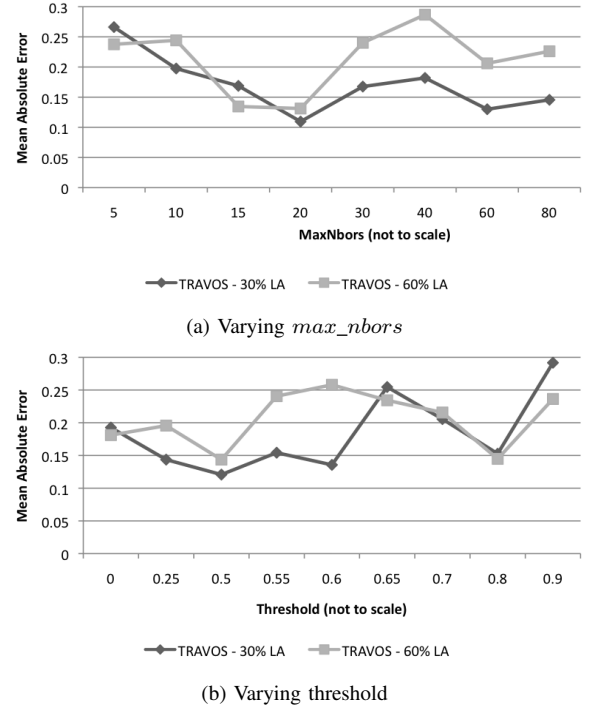
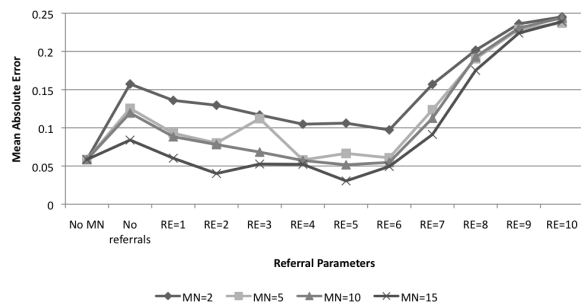


Fig. 3: Mean absolute error when applying our methods to TRAVOS at 30% and 60% LA

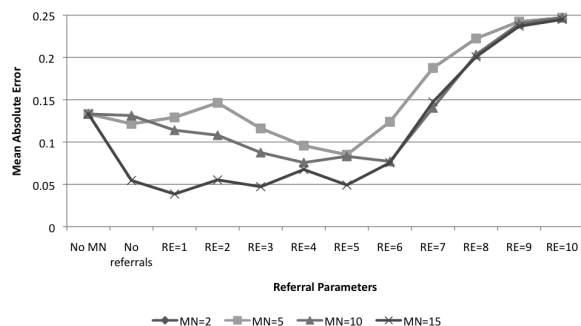
of the advisor trust values produced by our simulations, is that these values sometimes cluster into small ranges – for example, several trust values of approximately 0.42 may be generated, and several more with value of approximately 0.5, but none between 0.43 and 0.49. The implication is that a relatively minor change in the threshold value might serve to eliminate a number of advisors with the same level of trust, in turn affecting the overall trust model of sellers, possibly significantly. This reinforces the value of our proposed approach, to set an effective value for  $max\_nbors$  or thresholding, shown here through experimental methods. To this end, we are able to say that in general, TRAVOS works best in this scenario when  $max\_nbors$  is set to 20, or when a threshold of 0.5 is set.

We next look at examining the effect of advisor referrals using TRAVOS. Again, as with our work in the previous section, this is performed using a modified version of the above scenario; in this case the number of sellers is reduced to 40, and each buyer or advisor may submit 120 seller ratings, with no limit on the number of times each seller could be chosen.

The results for these tests are shown in Figures 4 and 5. As with the earlier figures, these are summary graphs which indicate the mean absolute error obtained for various combinations of minimum referral experience ( $RE$ ) and  $max\_nbors$  / threshold parameters; each series represents a different  $max\_nbors$  or threshold value, while the  $x$ -axis indicates the corresponding  $RE$  value. Like the results for PTM (see Figure 2), these graphs show that for low values of  $max\_nbors$ , where the advisor network size is very small,

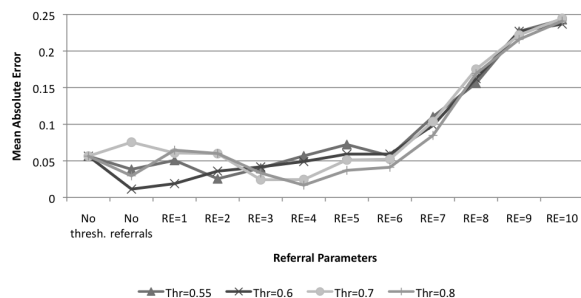


(a) Comparison for 30% LA

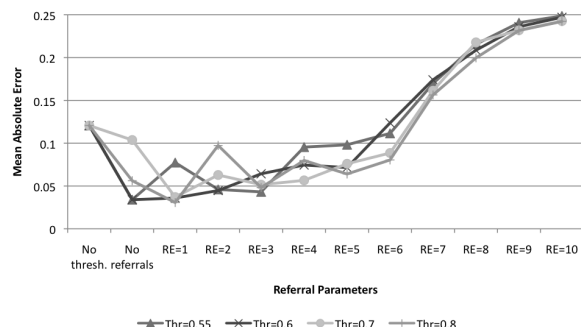


(b) Comparison for 60% LA

Fig. 4: Comparison of mean absolute error in TRAVOS using advisor referrals when varying  $max\_nbors$  and the minimum level of referral experience ( $RE$ )



(a) Comparison for 30% LA



(b) Comparison for 60% LA

Fig. 5: Comparison of mean absolute error in TRAVOS using advisor referrals when varying the trust threshold and the minimum  $RE$  level

using referrals will provide a reduction in error (that is, a trust model with improved accuracy). When using thresholding, similar reductions in error were observed by adding advisor referrals to networks using high threshold values (and hence having a small size). However, reductions in error were also occasionally seen for larger networks (those produced by using smaller thresholds); such improvements were rarely seen when applying referrals to large networks using PTM.

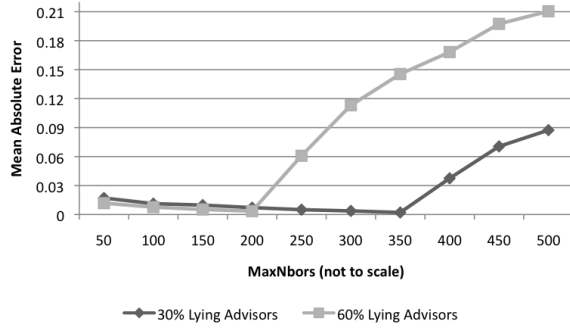
The high error that is seen in PTM when applying larger threshold values is generally due to the buyer having modelled very few, if any, of the advisors with such a high threshold, which leads in turn to insufficient information to model the trust of sellers (and assigning the default trust value of 0.5). In comparison, during our simulations, TRAVOS would assign high trust values, on the order of 0.8 or 0.9, to advisors more frequently, potentially because that model uses a more fine-grained model of advisor trust based on the advisor, the buyer, and the seller under consideration (whereas PTM calculates an overall value based only on the buyer and advisor). Accordingly, setting a high threshold would not affect the amount of information available to TRAVOS in the same way that it would PTM, leading to the more accurate results in this case.

#### IV. EFFECTS FOR LARGER POPULATIONS

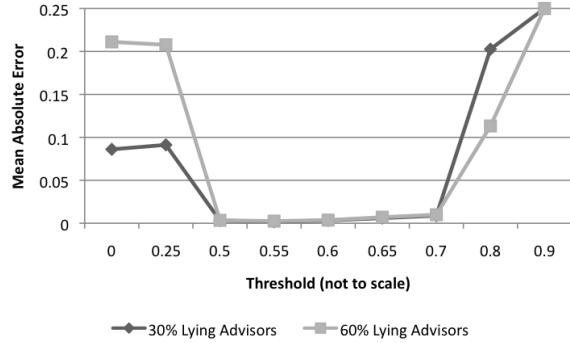
We also perform our simulations using PTM for a large advisor population size of 500. These tests otherwise maintain the same test conditions used for our earlier tests in Section II in terms of the number of sellers and the duration of the simulation. Again, however, we run the simulations with two values of the advisor lying percentage – 30% and 60% – and with several values of  $max\_nbors$  and thresholds. The results of these simulations, in terms of the mean absolute error of the trust model under each simulation as plotted against the  $max\_nbors$  or threshold value used, are indicated in Figure 6.

It seems clear in observing the results from Figure 6 that the results of applying different threshold values are reasonably consistent despite the change in advisor population. On the other hand, comparing the corresponding  $max\_nbors$  tests, as shown in this figure and in Figure 1, is somewhat trickier. It is clear that for both of the tested advisor populations, setting some  $max\_nbors$  value that is somewhat less than half of the advisor population size will result in a reduction in trust modelling error. However, while a value of  $max\_nbors$  of 30 is optimal for an advisor population of 80, the optimal value when the total population is 500 is much larger, at about 200 (which suggests that the value should not simply be set in absolute terms).

To find the solution to the  $max\_nbors$  issue, we remark that the two figures (6 and 1) have some visual similarity. This suggests it may be more appropriate to compare the two results in terms of  $max\_nbors$  as a proportion of the total advisor population. An effort to do this is provided as Figure 7. (Note that some of the data in this figure was interpolated from the simulations since comparable proportions were not



(a) Comparison when varying  $max\_nbrs$



(b) Comparison when varying threshold

Fig. 6: Mean absolute error applying our methods to PTM with advisor population of 500

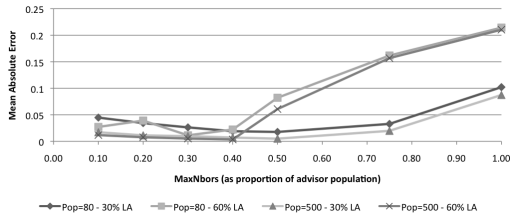
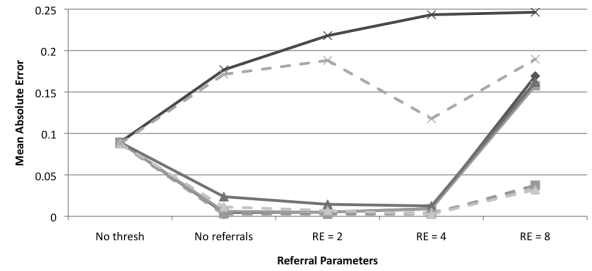


Fig. 7: Comparison of mean absolute error when varying  $max\_nbrs$  (as proportion of advisor population)

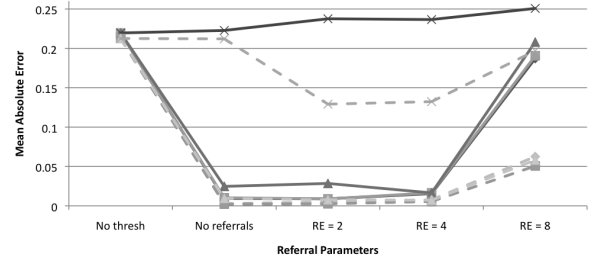
used in both sets of experiments.) This figure confirms that when setting  $max\_nbrs$  as a proportion of the total advisor population, the accuracy of the trust model is relatively consistent from one population size to the next.

We now turn to testing advisor referrals with PTM for the large advisor population case, again using the modified scenario used for the advisor referral simulations performed in the earlier sections, except with an advisor population of 500. Due to limited resources and the extended time required to perform simulations with the larger population, we elected to restrict the number of simulations performed for these cases to the minimum needed to indicate how the trust prediction accuracy changes as the  $RE$  increases.

We first consider how the larger population performs when thresholding and referrals are used in combination, as shown in Figure 8. The results are not identical for the two population



(a) Comparison when varying threshold (30% LA)



(b) Comparison when varying threshold (60% LA)

Fig. 8: Comparison of mean absolute error in PTM using advisor referrals, varying threshold and minimum RE, when advisor population is 500

sizes – there is a generally smaller trust-modelling error for the larger population when high  $RE$  values are used – but for the cases of greatest interest, specifically small advisor networks resulting from high thresholds, using modestly-chosen  $RE$  values, there are still improvements when adding referrals to these networks. Indeed, for a threshold of 0.8, the positive effects of adding referrals are much more pronounced in the larger population than in the 80-advisor scenario, particularly for  $RE = 4$ . We attribute this to the fact that more highly-trusted advisors will be available in the larger population, which would make a significant difference considering that perhaps only one or two advisors would survive the thresholding process using the smaller population.

Next we look at using  $max\_nbrs$  and referrals in combination, using  $max\_nbrs$  values of similar proportions relative to the population size, as shown in Figure 9. In this case, applying both techniques to a large network results in very similar, and in some cases (particularly for  $RE = 8$ ) much lower accuracy error compared to the smaller network. It seems safe to conclude that using referrals with a large advisor population will not only be effective in general, but that it will yield trust modelling accuracy at least as good as that obtained with a smaller population.

These results suggest to us that our proposed techniques –  $max\_nbrs$ , thresholding, and advisor referrals – together with our principled methods for setting their respective parameters, can be expected to help model trust more accurately

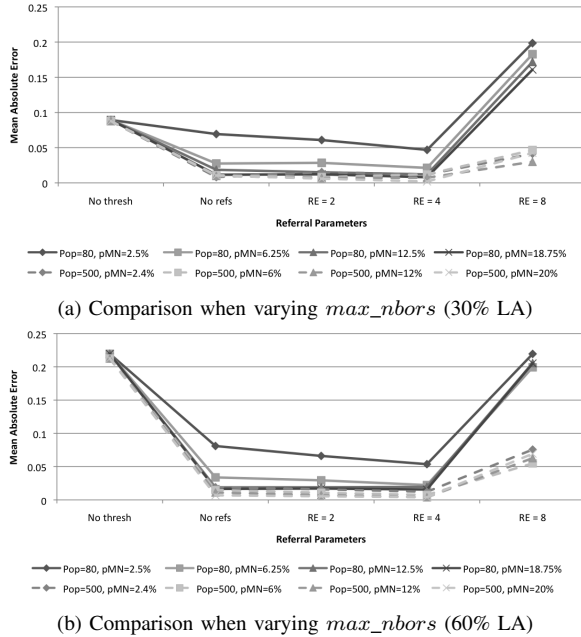


Fig. 9: Comparison of mean absolute error in PTM using advisor referrals, varying (proportional)  $max\_nbors$  and minimum RE, when advisor population is 500

in other trust approaches, at least those similar to the PTM and TRAVOS – for example, the Beta Reputation System [4]. Moreover, although the exact “optimal” parameters will likely differ from one system to the next, our results suggest that once the applicable threshold or  $max\_nbors$  value has been determined for one population, the same values (in the case of  $max\_nbors$ , the same proportion) can be used for other populations. In turn this could simplify these calculations greatly since it may only take a small population, perhaps 20 or smaller, to accurately determine the optimal threshold or (proportionate)  $max\_nbors$  values. Finally, our findings indicate that as the size of the advisor population increases, the benefits of using advisor referrals with regards to the accuracy of the trust model will also increase.

## V. DISCUSSION AND FUTURE WORK

In this paper, we have outlined three potential improvements to trust modelling – trustworthiness thresholding, maximum number of advisors, and advisor referrals – all of which aim to improve the accuracy of the recommendations for trustworthy agents derived from a buyer’s advisors. These three improvements can be used with different trust modelling methods, specifically the Personalized Trust Model and the TRAVOS model, as demonstrated in our study. We have also demonstrated that our proposed approach is sufficiently robust that it can be applied to offer improvements, even to large-sized populations of agents. The positive results outlined above suggest that other researchers should be able to adopt these methods when seeking to improve their own trust models. Towards this end, we have also clarified the experimental

framework which can be used to derive appropriate parameter values.

We have seen that either using trustworthiness thresholding or setting a maximum number of advisors will provide an improvement to the accuracy of the trust model. We have also seen that, in cases where the size of the advisor network is very small, using referrals may help to further improve the accuracy of this model.

In all cases, the parameters to be used should be modestly sized – allowing a reasonable number of advisors to be used, without including a large number of advisors that contribute little to the calculations of the trust model. Additionally, our results suggest how to set the actual value of the parameters. In particular, they indicate that the range of 0.5 to 0.6 is optimal for threshold parameters, while a  $max\_nbors$  parameter should be set as roughly 25% to 40% of the total size of the population.

### A. Related Work

Our usage of the  $max\_nbors$  and thresholding approaches in this paper was inspired in large part by the evaluation of design parameters for collaborative filtering (CF) recommender systems in [1]. That work evaluated  $max\_nbors$  and thresholding, two methods that had already been discussed in the CF literature [5][6], with regards to the correlation between two agents in a recommender system. Like our work here, that work found that  $max\_nbors$  was effective in improving the accuracy of the recommendations, so long as the value chosen was large enough to include a sufficient number of experienced agents, and small enough to exclude those agents that little to the evaluation. Unlike our findings for trust, however, the authors claim that for such a system, a  $max\_nbors$  value between 20 and 50 should suffice for most “real-world” scenarios, without regard to population size. Moreover, they do not find *any* benefit of using thresholding in the correlation case.

Yu and Singh have also explored work related to the usage of agents in trust and reputation systems, being the first to offer the suggestion of an advisor referral mechanism for reputation management [7]. In later work exploring the variance of the number of agents used in trust modelling [8], they showed that in varying the number of “witnesses” (agents) used in generating a trust model, the prediction accuracy improves slightly under otherwise identical conditions. However, due to the nature of the model, the numbers of agents to consider was limited to the range of 1 to 6 – much smaller than the  $max\_nbors$  values we considered; moreover, their model used a large number of simulation cycles.

### B. Future Work

In Section IV, we evaluated our three methods with a larger population. Experimental results show that trustworthiness thresholding is not affected by the population size, while the proper parameter for the  $max\_nbors$  method can be selected as a proportion of the population size. Meanwhile, in larger populations, allowing for advisor referrals will have

a greater benefit when used with smaller advisor networks, as compared to the fairly modest improvements seen with smaller populations.

However, we also note practical limits on the advisor population sizes for which these methods can be used. In addition to the tests previously described, we also attempted simulations with an advisor population of 1000, with the intention of further increases to more accurately reflect a large-scale system. During these attempts it became apparent that such a large population leads to additional challenges with regards to the execution time and memory resources needed to generate and calculate data for all of the advisors. We suggest that these issues may be overcome, in part, by applying our limiting techniques to an existing subset of advisors, perhaps one that is randomly determined, as opposed to the entire population. This provides an interesting avenue for future research.

A related topic which would also be promising to explore is improving the performance of the trust models when using these proposed methods. Given that these methods will in many cases substantially reduce the size of the advisor network used to produce the trust model of sellers, some performance optimization of these methods could help to improve the overall performance of the trust model. In particular, if a performance or memory gain could be developed to favour very small advisor networks while using referrals, this improvement might make up for the slight difference in accuracy compared to using larger networks.

With regards to our referral mechanism, we noted earlier that there is a limit on the number of levels of advisors through which this algorithm will search when looking for an acceptable replacement advisor. Presently this is set arbitrarily, based on a prediction of how many levels will be needed to search all nodes. In future work, we might examine whether varying this limit might produce improved results for referrals.

Several other researchers have proposed methods of incorporating CF techniques into trust modelling or vice versa, in many cases simply substituting trust as the primary metric in place of similarity [9]. One novel method in the literature is  $k$ -nearest recommenders ( $k$ NR), which dynamically selects the best  $k$  neighbours that are able to provide information about a particular desired item [10]. Further research might examine additional connections between trust modelling and recommender systems, such as applying  $k$ NR to existing trust models such as PTM and TRAVOS.

Work has also been done on the measure of information gain obtained as more agents are introduced into a trust or reputation system, which may be an additional factor to consider when determining how large the size of the advisor network should be [11].

Finally, it would be useful to consider how the domain under consideration might affect the choice of both the trust model and the specific methods or parameters used to optimize it. While we have focused on electronic marketplaces in this paper, other models are used in different domains – as in modelling the trust between agents collaborating on a health-

related challenge [12] – and the usefulness of our proposed methods may vary from one domain to the next.

## REFERENCES

- [1] J. Herlocker, J. A. Konstan, and J. Riedl, "An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms," *Information Retrieval*, vol. 5, no. 4, pp. 1386–4564, October 2002.
- [2] J. Zhang and R. Cohen, "Trusting advice from other buyers in e-marketplaces: The problem of unfair ratings," in *Proceedings of the Eighth International Conference on Electronic Commerce (ICEC)*, 2006.
- [3] W. T. L. Teacy, J. Patel, N. R. Jennings, and M. Luck, "TRAVOS: Trust and reputation in the context of inaccurate information sources," *Autonomous Agents and Multi-Agent Systems*, vol. 12, no. 2, pp. 183–198, March 2006.
- [4] A. Jøsang and R. Ismail, "The beta reputation system," in *15th Bled Electronic Commerce Conference*, 2002.
- [5] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: An open architecture for collaborative filtering of netnews," in *Proceedings of the 1994 Conference on Computer Supported Collaborative Work*, 1994.
- [6] U. Shardanand and P. Maes, "Social information filtering: Algorithms for automating "word of mouth";" in *Proceedings of ACM CHI'95 Conference on Human Factors in Computing Systems*, 1995, pp. 210–217.
- [7] B. Yu and M. P. Singh, "A social mechanism of reputation management in electronic communities," in *Proceedings of Fourth International Workshop on Cooperative Information Agents*, 2000, pp. 154–165.
- [8] —, "Detecting deception in reputation management," in *AAMAS '03*, 2003.
- [9] P. Massa and P. Avesani, "Trust-aware collaborative filtering for recommender systems," in *Proc. of Federated Int. Conference On The Move to Meaningful Internet: CoopIS, DOA, ODBASE*, 2004.
- [10] N. Lathia, S. Hailes, and L. Capra, "Trust-based collaborative filtering," in *Joint iTrust and PST Conferences on Privacy, Trust Management and Security*, 2008.
- [11] C. Sierra and J. Debenham, "Information-based agency," in *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI)*, 2007.
- [12] A. Koster, J. Sabater-Mir, and M. Schorlemmer, "Engineering trust alignment: a first approach," in *Proceedings of the Thirteenth International Workshop on Trust in Agent Societies*, 2010.