Entropy-Based Recommendation Trust Model for Machine to Machine Communications

Saneeha Ahmed(✉) and Kemal Tepe

University of Windsor, Windsor, ON, Canada
{ahmed13m,ktepe}@uwindsor.ca

Abstract. In a vast data collection and processing applications of machine to machine communications, identifying malicious information and nodes is important, if the collected information is to be utilized in any decision making algorithm. In this process, nodes can learn behaviors of their peers in the form of recommendation from other nodes. These recommendations can be altered due to various motives such as bad-mouthing honest nodes or ballot stuffing malicious nodes. A receiving node can identify an incorrect recommendation by computing similarity between its own opinion and received recommendations. However, if the ratio of false recommendations is low, the similarity score will be insufficient to detect malicious misbehavior. Therefore in this paper, an entropy-based recommendation trust model is proposed. In this model, a receiving node computes the conditional entropy using consistency and similarity of received recommendations with respect to its own opinions. The computed entropy indicates the trustworthiness of the sender. The proposed model clearly distinguishes malicious nodes from honest nodes by iteratively updating trust values with each message. The performance of the model is validated by a high true positive rate and a false positive rate of zero.

Keywords: Recommendation trust · Similarity · Entropy · Consistency · Connected vehicles

1 Introduction

Machine to Machine (M2M) communications is offered as a solution to collect vast amount of information from sensor and control actuators. In this setting, information is collected distributively, and decisions can be made at nodes by using the collected information. This information sharing enables nodes to benefit from their neighbors' experiences and to learn important information faster, such as identifying emergency events in connected vehicles, determining the quality of products in e-commerce, and making friends in social networks. In verifying or identifying false and malicious activity using the disseminated information from node, recommendation about other nodes can play an important role. However, recommendation schemes can be manipulated to badmouth honest nodes

Y. Zhou and T. Kunz (Eds.): ADHOCNETS 2016, LNICST 184, pp. 297–305, 2017.
DOI: 10.1007/978-3-319-51204-4_24

and ballot stuff malicious nodes in order to have a stronger influence in the network. By ballot stuffing and badmouthing, a malicious node can manipulate a novice receiver to exclude any important information coming from honest nodes or accept wrong information from malicious nodes. That is why trustworthiness of recommendations must be estimated before accepting them.

Recommendation Trust (RT) is often considered as feedback credibility [1] and its is defined as a measure of trustworthiness of a node's recommendation about another node. RT is estimated on the basis of similarity between an evaluator's own opinion and received recommendations. In case a node sends a false recommendation only for a few nodes, i.e. selective misbehavior, or switches between malicious and honest behaviors in an on-off manner, it will attain a very high similarity score. The evaluator may observe the pattern of recommendations over a period of time and determine how consistent is the sender in its recommendations. Consistency of information can allow a new evaluator to determine whether to trust a recommender when it provides information about new nodes.

In this paper, an entropy-based RT model is proposed which utilizes the consistency of information as well as its similarity with evaluator's own opinion. The evaluating node calculates the average entropy based on following two factors: (1) Jaccard similarity score [2] of the recommendations, (2) ratio of consistent information over the total number of recommendations. Similarity is defined as the fraction of recommendations of a node that are same as the opinion of the evaluator. Consistency is defined as the fraction of recommendations from a sender that does not change in consecutive messages. A lower entropy indicates higher trustworthiness of the source since the information sent by this source has less uncertainty. In order to predict the time dependent behavior of the source, the recent trust is derived from the observed entropy and previous trust value. With this proposed trust model, evaluator can identify the malicious nodes even if there is a selective misbehavior or an on-off attack.

Remainder of this paper is organized as follows: Sect. 2 provides a literature review, Sect. 3 provides details of the trust model Sect. 4 provides simulation results and Sect. 5 provides a conclusion and future works.

2 Literature Review

Recommendation systems have gained significant attention in M2M communications [3,4]. In these systems, nodes provide feedback about behaviors of other nodes in order to make communication secure and reliable. However, nodes can manipulate their recommendations in order to have a stronger influence in the network by supporting malicious nodes or by badmouthing honest nodes [5–7]. If the credibility of recommendations is not considered, an attacker can easily defame a target node by creating multiple fake identities to generate false recommendations [8]. In order to use recommendations in establishing trust for a seller in an e-commerce site or a user in a social network site, a personalized similarity metric is proposed in [9,10]. In these studies, similarity is a measure

of difference of satisfactions between two users over a set of common items. In [1], a user's recommendation is assigned as a weight equal to the similarity and the weighted recommendations provide the trust in the recommended user. Schemes in [1,9,10] are resilient against on-off attacks and effectively prevent unsuccessful transactions. The similarity measure has been adapted for a mobile ad hoc network in [11] which uses a fuzzy collaborative filter to restrain malicious recommendations from effecting the trust computation. The proposed collaborative filter uses the similarity between the nearest most similar neighbors to compute trust in a node. While the proposed trust mechanism improves the network throughput and packet drop ratio, it is not studied how reliable is the trust estimation. Moreover, the system is only tested under the honest majority scenario.

The infrequent badmouthing attacks are successful in defaming honest nodes or credible items. That is why a scheme to prevent those attacks is proposed in [12], where a drastic change in product rating with respect to time indicates such attacks. In [12], Dempster-Shaffer Theory is used in a system to identify malicious users and recover reputations of products or users. The system is tested under honest majority assumption and provided good receiver operation characteristics. However, if malicious users form a majority, the system may fail. In [12], recommendation values are real numbers which are naturally different from each other. This difference leads to oscillations while computing trust. These oscillations are undesirable in M2M communications and can be masked out by using binary recommendations similar to the one used in [13] for an e-commerce system. However, the scheme proposed in [13] only filters most deviant recommendations and ignores their sources. Hence, malicious nodes can survive and may corrupt a larger portion of the network to have a stronger influence on the network. That is why, in this paper RT is established in order to determine whether a node's recommendations can be trusted and included in obtaining the true nature of the neighbors. The proposed entropy-based RT model is explained in the following section.

3 Entropy-Based Recommendation Trust Model

In order to establish RT and use it to eliminate malicious recommenders, it is assumed that in a given network some of the nodes have opinions about each other and they update their opinions based on their experiences. These nodes communicate their opinions in the form of recommendation to their neighbors in order to assist them to make their own judgments about other nodes. In this scenario, malicious recommenders modify values of their opinions, where they report a binary "1" for a malicious node and a binary "0" for an honest node in order to misguide the evaluating node. Let us consider that an evaluator node E, having a set of opinions $\mathbf{G}$, receives recommendations $\mathbf{R^1}, \mathbf{R^2}, ..., \mathbf{R^d}$ with cardinality A, from sources $s_1, s_2, ..., s_d$, as shown in Fig. 1. In our work, the cardinality A was fixed for all cases. The opinions as well as recommendations comprise of binary values such that a binary "0" indicates that a negative recommendation and a binary "1" indicates a positive recommendation. The node E

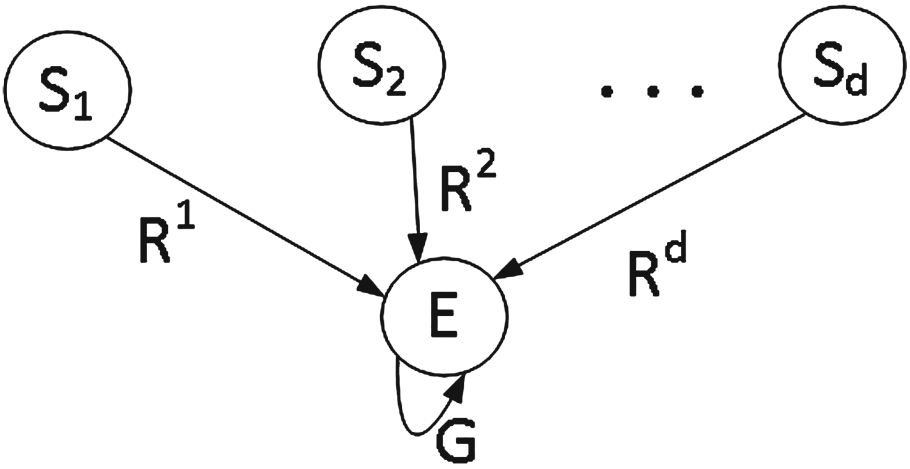

Fig. 1. Opinion of E and recommendations sent by neighbors

computes a Jaccard Similarity score [2] between its own opinion and the received recommendation from any node s_k as follows.

$$Sim(E, s_k) = \frac{|\mathbf{G} \cap \mathbf{R^k}|}{|\mathbf{G} \cup \mathbf{R^k}|}. \tag{1}$$

The similarity score computed in (1) is basically $P[G \cap R^k]$ where the universal set is assumed to be $\{G \cup R^k\}$. Hence, assuming that probability $p_{s_k} = Sim(E, s_k)$, then the entropy resulting from the similarity score is given by,

$$H_{s_k} = -p_{s_k} * log_A(p_{s_k}). \tag{2}$$

If the recommendation from the source s_k is similar to the evaluator E, the entropy will be low. However, if the similarity score is around 0.5, the entropy will be the maximum since the source is unreliable. In order to avoid detection, a malicious node may maintain a high similarity by giving incorrect feedback for a few nodes. Moreover such a node may only send incorrect feedbacks when they are most beneficial for the attacker, i.e. in a probabilistic manner. Hence the consistency of recommendations from this sender must be observed in subsequent messages in order to identify its true nature. Every time the node s_k sends a recommendation, E observes $\mathbf{R^k}$ for any inconsistencies in $\mathbf{R^k} = \{r_1, r_2, ..., r_A\}$. For simplicity, let us assume that symbols $r_1, r_2, ..., r_A$ always assume a binary value. Now let a given symbol r_i^k assume a value a_i in m_i out of N messages from sender s_k. Then the probability $p_{r_i}^k$ of r_i taking the value a_i in the next message is given by

$$p_{r_i}^k = P[r_i^k = a_i] = \frac{m_i}{N}. \tag{3}$$

The entropy of input symbols of this source is given by

$$H_{R_k} = \sum_{i=1}^{A} -p_{r_i} * log_A(p_{r_i}). \tag{4}$$

The trust of s_k for the message N is the average entropy of its similarity as well as entropy of input symbols and is given by

$$T_k(N) = \theta \cdot (1 - (H_{s_k} + H_{R_k})) + (1 - \theta) \cdot T_k(N - 1), \tag{5}$$

where θ is a weight parameter that recognizes the importance of recent entropy and previous trust. For the first set of recommendations from the sender s_k, the trust is only measured by the similarity, that is, $T_k(0) = H_{s_k}$. The value of trust can be normalized however in this work, if $T_k(N)$ is undefined or below zero, it is assigned a value of zero. If $T_k(N)$ is greater than one, it is assigned a value of one.

The performance of the system is measure in terms of true and false positive rates which are defined as follows:

True Positive Rate (TPR). TPR determines how correctly are the malicious nodes identified and it is given by

$$TPR = \frac{P_{M|M}}{P_{M|M} + P_{H|M}},$$

where $P_{M|M}$ is the probability of detecting a malicious node as malicious and $P_{H|M}$ is the probability of detecting a malicious node as honest.

False Positive Rate (FPR). FPR determines the error produced by misclassifying an honest node as malicious and it is given by

$$FPR = \frac{P_{M|H}}{P_{H|H} + P_{M|H}},$$

where $P_{H|H}$ is the probability of correctly identifying the honest nodes as honest and $P_{M|H}$ is the probability of misclassifying the honest nodes as malicious. Two parameters, TPR and FPR, will be used to study the performance of the proposed scheme in later sections.

4 Simulation Results

In order to test the trust model, a network of connected vehicles is assumed where cars travel together for some distance. In this scenario, one vehicle may encounter the same neighbors over and over and forms opinions about them. These opinions may change with time, however, for honest vehicles the changes will be infrequent. Nodes will report their opinions in the form of recommendations, in scheduled broadcast messages. Some of these senders act maliciously and modify

their recommendations about q of their neighbors. Malicious nodes send messages containing modified recommendations with a probability p. Honest nodes may misjudge some of the other nodes and amount of error thus introduced is assigned an error probability of 0.04. The error introduced by the honest nodes is treated as noise and does not affect the performance of the overall system.

It is assumed that the malicious nodes do not collude to provide incorrect recommendations about the same target nodes. However, it can be shown that even if the nodes collude, the consistency of the information and the similarity between evaluator and recommender can be used to identify malicious nodes and result will not be severely affected. Another assumption is that all malicious nodes misbehave with the same probability. This assumption would affect the true positive rate, since the threshold value for trust to classify nodes as malicious, will be affected.

Performance of the model will be tested in three aspects that include trust evolution with number of messages, true positive rates and false positive rate. First it will be studied how the trust of honest and malicious nodes evolve with the number of messages. For these experiments, malicious nodes constitute 50% of the nodes. These nodes send 100 messages in which they send their recommendations about 100 other nodes therefore the cardinality A is 100. The trust values of honest and malicious nodes are updated with each message, by averaging the previous trust, and combined entropy of similarity and consistency. For the simulations, equal weight is applied to the previous trust, and the combination of similarity and consistency, which means that $\theta = 0.5$. In the first message all nodes send correct information in order to attain a high initial trust value. First the evolution of trust is studied in Fig. 2. In Fig. 2 the honest nodes, despite of introducing noise in the information, attains a maximum trust value of over 0.999. However the malicious nodes send false recommendations about 10 out of 100 neighbors, i.e. $q = 10$, with different probabilities. In this scenario if p is low, the malicious node attains a relatively higher average trust value of about 0.6, although the trust of these nodes decreases gradually with the number of

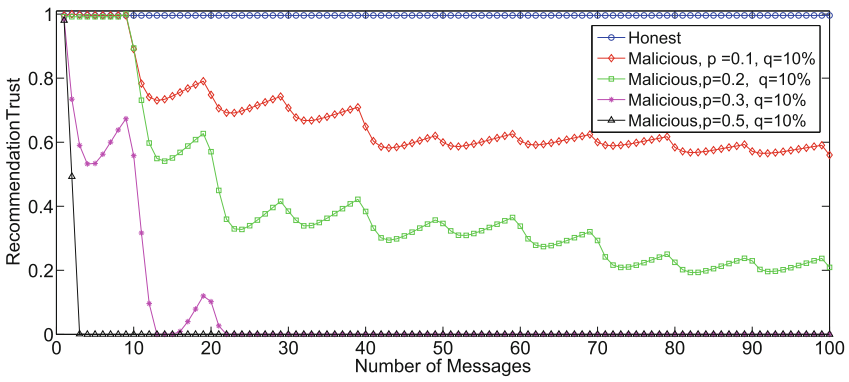

Fig. 2. Recommendation trust evolution

messages. It can be observed that at a low p, trust values of malicious nodes may increase with more correct messages, nevertheless the initial trust of the nodes can not be recovered completely. However if p is larger, the trust value is reduced at a much faster rate. The trust values of malicious nodes decline sharply with the first few messages and does not increase easily with correct messages. Thus the model identifies selective misbehavior and assigns low trust values to the malicious nodes even if the probability of incorrect information, p, is low.

Let us further examine the impact of number of false recommendations on trust evolution. For this experiment, a relatively larger value of q is considered in Fig. 3. In Fig. 3, the value of $q = 30\%$, which means that out of 100 nodes, recommendations for 30 nodes have been modified. With this setting, even if the malicious nodes send a very small number of messages with modified recommendations, hence p is small, the proposed scheme reduces their trust values drastically. Hence it is learned that if a large number of recommendations is modified then detecting malicious nodes becomes easier.

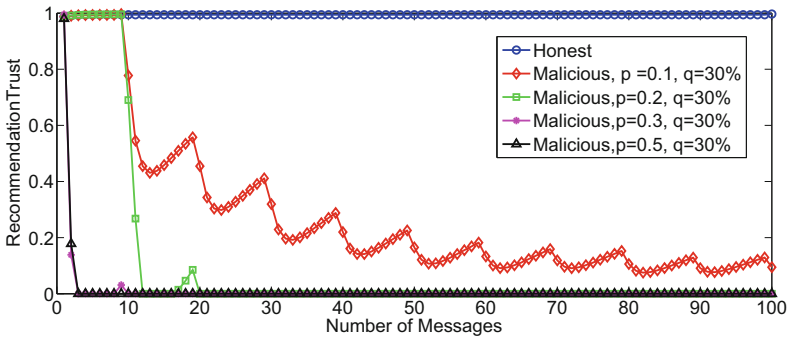

Fig. 3. Impact of increasing number of altered recommendations

In order to study the true positive rate, 50% of the nodes are configured as malicious and send false recommendations with different probabilities. The malicious nodes show selective misbehavior, such that they only ballot stuff or bad mouth q of its neighbors with a probability p and give honest recommendations about the others.

Figure 4 demonstrates that when p and q both are low, then the TPR is low, as the recommendations remain consistent most of the time and recommendations for only few neighbors are modified by the malicious nodes. These malicious nodes maintain a high similarity most of the time and show extremely infrequent misbehaviors. That is why, it is difficult to identify malicious nodes when p and q are low. However if the malicious activity becomes more obvious, either by increasing p or q or both, the true positive rate increases significantly. On the other hand, the trust of honest nodes increases with each message and therefore they are never misclassified. Hence the proposed trust model does not produce any false positives and the false positive rate is zero.

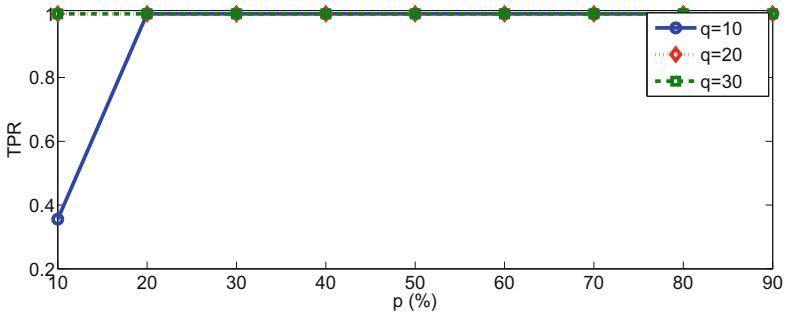

Fig. 4. TPR for a recommendation vector of length 100 bits where q bits are maliciously modified with a probability p

5 Conclusion

The essence of machine to machine communication is to identify reliable information even when some of the participants are malicious. In this study the malicious behavior has been modeled as false feedback about peers. These malicious peers are identified using the proposed entropy based trust model and their recommendations are excluded by the system. On the other hand the trust of honest nodes is increased with each interaction and their recommendations are used to predict the behavior of other nodes. How the system will make use of such recommendations is studied in our other research. The proposed trust model is tested only for binary recommendations. The scheme can be extended for non-binary recommendations which are used in most e-commerce systems.

References

1. Das, A., Islam, M.M.: SecuredTrust: a dynamic trust computation model for secured communication in multiagent systems. IEEE Trans. Dependable Secure Comput. **9**(2), 261–274 (2012)
2. Niwattanakul, S., Singthongchai, J., Naenudorn, E., Wanapu, S.: Using of Jaccard coefficient for keywords similarity. In: Proceedings of the International MultiConference of Engineers and Computer Scientists, vol. 1, p. 6 (2013)
3. Luo, J., Liu, X., Fan, M.: A trust model based on fuzzy recommendation for mobile ad-hoc networks. Comput. Netw. **53**(14), 2396–2407 (2009)
4. Kim, S., Kwon, J.: Recommendation technique using social network in internet of things environment. J. Korea Soc. Digital Ind. Inf. Manag. **1**(1), 47–57 (2015)
5. Ullah, Z., Islam, M.H., Khan, A.A., Sarwar, S.: Filtering dishonest trust recommendations in trust management systems in mobile ad hoc networks. Int. J. Commun. Netw. Inf. Secur. **8**(1), 18 (2016)
6. Khedim, F., Labraoui, N., Lehsaini, M.: Dishonest recommendation attacks in wireless sensor networks: a survey. In: 12th International Symposium on Programming and Systems (ISPS 2015), pp. 1–10. IEEE (2015)

7. Mousa, H., Mokhtar, S.B., Hasan, O., Younes, O., Hadhoud, M., Brunie, L.: Trust management and reputation systems in mobile participatory sensing applications: a survey. Comput. Netw. **90**, 49–73 (2015)
8. Yang, Y., Feng, Q., Sun, Y.L., Dai, Y.: Reptrap: a novel attack on feedback-based reputation systems. In: Proceedings of the 4th International Conference on Security and Privacy in Communication Networks, p. 8. ACM (2008)
9. Xiong, L., Liu, L.: Peertrust: supporting reputation-based trust for peer-to-peer electronic communities. IEEE Trans. Knowl. Data Eng. **16**(7), 843–857 (2004)
10. Srivatsa, M., Xiong, L., Liu, L.: Trustguard: countering vulnerabilities in reputation management for decentralized overlay networks. In: Proceedings of the 14th International Conference on World Wide Web, pp. 422–431. ACM (2005)
11. Luo, J., Liu, X., Zhang, Y., Ye, D., Xu, Z.: Fuzzy trust recommendation based on collaborative filtering for mobile ad-hoc networks. In: LCN, pp. 305–311. Citeseer (2008)
12. Liu, Y., Sun, Y., Liu, S., Kot, A.C.: Securing online reputation systems through trust modeling and temporal analysis. IEEE Trans. Inf. Forensics Secur. **8**(6), 936–948 (2013)
13. Whitby, A., Jøsang, A., Indulska, J.: Filtering out unfair ratings in bayesian reputation systems. In: Proceedings of 7th International Workshop on Trust in Agent Societies, vol. 6, pp. 106–117 (2004)