

A Strategy to Detect Colluding Groups by Reputation Measures

Attilio Marcianò^a, Domenico Rosaci^b and Giuseppe M. L. Sarnè^c

^aDepartment DIIES, University Mediterranea of Reggio Calabria, via Graziella snc, loc. Feo di Vito - 98123 Reggio Calabria

^bDepartment DIIES, University Mediterranea of Reggio Calabria, via Graziella snc, loc. Feo di Vito - 98123 Reggio Calabria

^cDepartment of Psychology, University of Milan Bicocca, Piazza dell'Ateneo Nuovo, 1, 20126 Milan, Italy

Abstract

Collusion is the malicious activity mostly frequent in agent-based recommender systems in which two or more agents agree with each other to mutually exchange high positive feedback in order to gain undue advantages by altering the correct computation of reputation measures in their agent communities. Therefore, identification of colluding agents is an important issue and several strategies have been developed to this purpose. Among them, the EigenTrust algorithm is well known, although it is limited by the necessity of knowing a priori which agents are considered as trustworthy and the impossibility of recognizing several groups of colluding agents acting simultaneously and autonomously. The focus of this paper is dealing with the above issues and, to this end, we will present a strategy to support EigenTrust both providing the necessary inputs about pre-trusted agents and recognizing groups of malicious agents. In particular, we combined EigenTrust with a clustering process in order to suitably grouping the agents according to their reputation scores. We carried out a preliminary tests which have shown promising results about the effectiveness of the proposed strategy.

Keywords

Agent Groups, Clustering, Collusion, EigenTrust, Reputation Measures

1. Introduction

The social skills of software agents make them able to support complex social relationships inside their communities [1, 2]. In particular, carrying out trust-based activities in social communities [3] it is assumed to be an effective solution to improve the quality of social interactions [4] limiting malicious behaviors. Therefore, the existence of mutual high levels of trustworthiness among the members of agent communities can be assumed as a preemptive conditions for carrying out satisfactory agents' activities to realize therein [5].

The possibility to measure the trustworthiness of an actor has been widely described in the

WOA 2023: 24th Workshop "From Objects to Agents", November 6–8, Rome, Italy


✉ attilio.marciano@unirc.it (A. Marcianò); domenico.rosaci@unirc.it (D. Rosaci); giuseppe.sarne@unimib.it (G. M. L. Sarnè)

🌐 https://www.unirc.it/scheda_persona.php?id=1130 (A. Marcianò);

https://www.unirc.it/scheda_persona.php?id=696 (D. Rosaci); <https://www.unimib.it/giuseppe-maria-luigi-sarne> (G. M. L. Sarnè)

🆔 0000-0002-7229-5464 (A. Marcianò); 0000-0002-9256-9995 (D. Rosaci); 0000-0003-3753-6020 (G. M. L. Sarnè)

© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

literature [6, 7, 8, 9] and it is usually based on feedback released by counterparts in order to manifest their satisfaction degree about the occurred interaction. In this respect, a large number of agent societies have been designed to consider measures of trust and reputation based on feedback. This helps multi-agent systems (MAS) to become more resilient against misleading behaviors [10, 11] aimed at deceiving others [12].

Conveniently, trust can be defined as: "the quantified belief by the trustor with respect to the competence, honesty, security and dependability of the trustee within a specified context" [13], where trustor and trustee are who gives and who receives trust, respectively. Reliability, honesty and security are the usual dimensions on which trust lives, respectively referred to the capability of satisfying counterparts' expectations, limiting misleading behaviors and, finally, avoiding undesired activities like unauthorized accessing to reserved data. In the following, we will be specifically focused on the collusion. It is a common malicious behavior occurring when two or more malicious agents agree to provide high positive feedback to each other in order to alter the perception of their reputation in honest agents so that undue advantages can be gained in damage of the others [14]. Moreover, with the term trust we will identify a measure of reliability and honesty, while the term collusion will be referred to the concerted willingness of the agents to work together to distort trust measures, while in the following we will not consider security issues because they are orthogonal to the focus of this paper.

In particular, the trust is a subjective measure directly computed by the trustor agent about the trustee agent. If a trustor agent does not have a suitable knowledge about the trustee then the trustor must require the opinion of other agents. The measure of the trustworthiness that the whole community place on a given agent is identified with the term *reputation* (which is an indirect measure) and, in presence of malicious agents implementing selfish behaviors, including deceptions and frauds, the availability of reputation measures is essential to make a good choice of agents with which to collaborate. Unfortunately, the computation of reputation measures in the presence of several different groups of colluding agents, each one operating independently, can lead to a misrepresentation of agent reputation within the community, making the task of identifying malicious agents very difficult.

In the literature several strategies to detect malicious actors have been described (see Section 2), among which an effective proposal is represented by the well-known EigenTrust algorithm [15], also adopted by the Google search engine to rank Web pages [16]. However, the effectiveness of the EigenTrust algorithm mainly depends on knowing which agents are a priori considered trustworthy [17, 18] and this is not a trivial task in the presence of colluding agents. Moreover, unfortunately, this system is not robust to an orchestrated attack [19] and, therefore, does not work well and even more so in the presence of multiple groups of colluding agents when they act simultaneously within the same agent community. To tackle these limits of EigenTrust, in this paper we propose of:

- providing in an automatic way all the information about the trustworthiness of agents that are required by the EigenTrust algorithm;
- combining a clustering process with EigenTrust to group agents based on their reputation scores in order to identify groups of colluding agents.

Preliminary experiments carried out on a test case have provided promising results showing that our proposal is potentially able to identify several groups of colluded agents. Moreover, the

obtained experimental results highlighted as the effectiveness of our method depends on the dimension of the agent community and the percentage of colluded agents present therein.

The rest of the paper is organized as follows. In Section 2 we present some related work, while in Section 3 we introduce the proposed strategy to identify colluded agents belonging to several independent groups, In Section 4, we provide an example of application of our method and describe the simulations we carried out to validate it. Finally, in Section 5 some conclusions are drawn and our ongoing researches will be discussed.

2. Related Work

Given the multifaceted nature of reputation, it is an hard task that of providing its definition. In the following, we will refer to the definition of reputation derived by [20, 21], where reputation of an actor can be assumed to be the expectation about his/her future behaviors estimated on the basis of feedback about past his/her behaviors with regard to a specific context at a specific time made by other actors. In general, reputation influences future choices when they are based on expectations about future behaviors [22] and, in particular, it is central when a direct knowledge of the potential counterpart there not exist or it is not adequate, like large real or virtual communities. Moreover, the reputation knowledge reduces both information asymmetry between parties and the risks of deceptions [23].

To be effective, a reputation system (RS) needs to involve long-living entities, driving future decisions and living on feedback released by counterparts [24]. In turn, feedback tightly depends on their honesty in reflecting the real opinion of the trustor about the trustee and on its accuracy measured as difference with respect to the real trustee's trustworthiness [20].

Other aspects of a reputation system, particularly relevant in presence of very large communities, relate to information management in a centralized or distributed architecture [25, 26] and/or in a local or global approach [27, 28]. The choice among these aspects is driven by community size, cost, scalability, computational and storage overhead, complexity, and other properties. For a more complete overview on those and other properties of the reputation systems, the interested reader may refer to [29, 30, 31, 32, 33].

Typically, opinions about a trustee (usually in numerical form) are collected and aggregated, often through a weighted average, to update the trustee's reputation score within the community. In this light, in the popular eBay RS [34], sellers' reputation scores are updated using feedback issued by buyers, who must also interpret the reputation scores based on their personal attitudes toward risk. Despite its various updates over time, the resilience of this RS is critical being exposed to various threats [35, 36, 37].

Designed over an overlay network, two performing RSs are *i*) PeerTrust [38], a distributed RS capable to identify reliable and malicious peers by aggregating direct feedback, indirect feedback (taking into account the source reliability), number of peers' transactions and context and *ii*) Hypertrust [39], designed to work in large and competitive federations of utility computing infrastructure to discovery resources on the basis of reliability and reputation data.

The EigenTrust [40] RS assumes the transitivity of trust opinions and calculates the overall reputation of peers by aggregating and normalizing the trust representation of peers appropriately weighted by their trustworthiness. Trust values, arranged in a matrix, and converge

asymptotically to the matrix eigenvalues. However, several strategies have been proposed to maliciously manipulate EigenTrust [41] scores and, likewise, several strategies have been proposed to improve the resilience of this [42, 43]. Against collusive activities this RS requires the presence of mentors whose opinions are assumed to be trustworthy, but to identify collusive actors there is no unambiguous criterion, e.g., an univocal threshold on reputation scores to divide honest peers from malicious ones.

We highlight how, to the best of our knowledge, [38], [39] and [40] are among the best-performing RSs, although profoundly different both in terms of the application scenario, which influences their design, and the metrics adopted to calculate the actors' reputation scores, which are in any case based on both direct and indirect contributions.

Finally, we would like to point out the existence of some RSs, such as FIRE [44], specifically designed for benign scenarios in the absence of malicious actors, which generally do not fit well in the real world.

More generally, RSs can be subject to a variety of attacks, of which collusion ones are common [45] and occur when two or more malicious actors secretly agree to engage in illicit activities aimed at changing the perception of their trustworthiness in the rest of the community.

In more detail, the more colluders there are, the less effective are the strategies aimed at identifying them. Such strategies are mainly based on [46]: (i) seeking and evaluating opinions that disagree with most other users; (ii) attending to sudden changes in reputation measures over time; (iii) assessing the accuracy and honesty of users giving feedback. In addition, to avoid specific colluders' countermeasures, RSs often combine multiple strategies and mechanisms together to increase their resilience toward collusive activities [47]. However, other strategies to detect collusion attacks have been proposed in the literature. For example, modeling the detection of colluders using game theory such as in [48] or with the use of blockchain technology such as in [49] where the blockchain also supports a reputation-based voting scheme in which candidate history and recommended opinion are considered or in [50] where messages and path redundancy are adopted based on the reliability and performance of nodes.

Finally, not rarely more groups of colluders can act in a community at the same time, but only few RSs consider such an issue as in [51] where the search is based on the search of tight peers' relationships or in [52] where the hypergraph theory is exploited to find whose vertices closely connected through hyperedges to find cluster of nodes tightly connected.

3. The Proposed Strategy

In a nutshell, EigenTrust operates in a social network scenario and its main idea is to calculate the reputation r_i of each member i as the sum of all the trust values t_{ij} that each other user j assigns to i , weighted with j 's reputation. In practice, r_i can be assumed to be the barycenter of all the t_{ij} . In short, EigenTrust exploits some information about known agents (mentors) considered as trustworthy by assigning them a very high reputation even though, at the same time, the reputation of all other agents is penalized by the algorithm. Although in this way EigenTrust can effectively detect malicious agents (among those agents not pre-trusted) the task of detecting mentors is not trivial.

To deal with this limitation of EigenTrust, our strategy is first to identify the best candidates

to be considered as colluding agents and, at the same time, indicate all the other agents (who are assumed not to be colluding) as mentors. Then, a new algorithm for organizing agents into clusters is introduced; agents belonging to the clusters with the worst average reputation scores will be considered as colluded if they assign high trust values to each other belonging to the same cluster, while from the other agents in the community they receive low levels of trust.

We verified that our two-step strategy is very effective in presence of a single cluster of colluding agents, but in presence of multiple clusters of colluding agents, each operating independently from the other clusters, our strategy is able to identify only the one with the worst average reputation as a cluster of colluders. To overcome this drawback, a suitable threshold was introduced into the algorithm in order to consider as colluded all the clusters with an average reputation lower than this threshold. Next, the algorithm works in an iterative way by removing from the community the clusters (i.e., agents) thus identified, then the reputations of the remaining agents will be recalculated in order to identify additional clusters of colluded agents. The iterative process will end when all remaining clusters have an average reputation greater than the threshold used by our algorithm.

It is important to point out two current limitations of our approach. The first limitation is that the threshold under which a group should be considered colluded is decided arbitrarily by the social network administrator; while in some aspects this is reasonable, we are operating for automatically extrapolating information from the social community to make this process less arbitrary. The second limitation is related to the clustering algorithm employed, which in advance requires to know the number of clusters to be formed. Again, we are operating to overcome this limitation both by verifying the performance of other clustering algorithms that do not require such information in advance, and by trying to extrapolate from the social network the information necessary to reliably estimate the number of clusters to be used. Both of the above issues are the subject of our ongoing research.

4. Experiment

4.1. The Recursive Reputation Model

Let us consider a social network of n members, each one identified in an univocal manner by an integer i , $1 \leq i \leq n$. In such a context, the *trust* perceived by the member j about the member i is represented by the real number t_{ij} , $0 \leq t_{ij} \leq 1$, $\forall i, j = 1, \dots, n$, while the *reputation* r_i that i has in the social network is computed as the sum of all the trust values t_{ij} , $j = 1, \dots, n$, referred to a trustor j , suitably weighted by the reputation r_j of j . More formally:

$$r_i = \frac{\sum_{j=1}^n t_{ij} r_j}{\sum_{j=1}^n r_j}, \quad i = 1, \dots, n. \quad (1)$$

Let us denote with $T = [t_{ij}]$ the *trust matrix*. It can be assumed to be the transpose of the weighted adjacency matrix $A = [a_{ij}]$ of a directed graph \mathcal{G} . In \mathcal{G} each node corresponds to a user and a non negative value a_{ij} is assigned to the edge (i, j) to represent the trust value placed by the member i on the member j . Therefore, the Equation (1) can be reformulated as:

$$Tr = r, \quad \|r\|_1 = 1, \quad (2)$$

where $r = (r_1, \dots, r_n)^T$ is the *reputation vector* and the 1-norm is defined as the sum of the absolute values of the vector components. However, to guarantee the reputation uniqueness, the vector r in (2) must be normalized. Moreover, we impose equal to 1 the sum of the trust values t_{ij} assigned by each member j to the other social network members because, in this way, the matrix T will result column-stochastic.

A solution to the eigensystem problem given by Equation (2) can be equivalently reformulated in the computation of the stationary distribution for a Markov chain [53], which is represented by the matrix T , named transition matrix. In this case, if we assume that T has only positive elements, by applying the Perron Frobenius Theorem, then $\lambda = 1 = \rho(T)$ is a simple eigenvalue of T (the other ones are less than 1 in modulus), and there exists a unique vector $r \in \mathbb{R}$, $\|r\|_1 = 1$, such that $Tr = \rho(T)r = r$ that is a unique positive reputation vector.

A modified version of (2) is the well-known *PageRank model*:

$$(\alpha T + (1 - \alpha)vu^T)r = r \quad (3)$$

where $0 \leq \alpha \leq 1$, u is the vector of all ones, while v is a non-negative vector with unitary 1-norm, that is $u^T v = 1$ and the vector v is known as *teleportation vector*. In the original PageRank algorithm $v = \frac{1}{n}u$ so that when $\alpha \neq 0, 1$, the existence and uniqueness of the solution to (3) is confirmed.

In the next subsection, we propose a strategy that allows to detect malicious users based on the matrix T .

4.2. Detection of malicious users

In this Section we present the proposed strategy to provide the EigenTrust algorithm with the inputs on pre-trusted agents and the information to identify groups of malicious agents, directly from the trust matrix $T = [t_{ij}]$ as defined in Equations (1) and (2).

More in detail, let i, j a pair of agents classified as *malicious* because we assume they collude when (i) t_{ij} and t_{ji} are high and they are similar from each other and (ii) differently, the sum of the remaining trust values belonging to the rows i and j is low. In other words, in the matrix T the values of t_{ij} and t_{ji} are high while these users receive low values from the majority of the other members.

To group the social network users, a *Spectral* clustering [54] has been adopted. This algorithm exploits the spectral transformation of a data similarity matrix to separate data in an Euclidean space and arrange them in a more complex way. A main advantage of this algorithm is to not make strong assumptions on the form of clusters so that it is suitable to solve a wide range of problems. Moreover, it is computationally efficient also on large data sets because when the similarity is chosen (a not trivial step) only the solving of a linear problem is required, without risking local minima or algorithm restarts.

From a practical viewpoint, for a set of data points (i.e., agents), assigned the similarity $s_{ij} \geq 0$ between all the agents pairs i, j , we cluster similar data points into more groups. We represent the data as a graph $G = (V, E)$, where each vertex represents an agent i and each edge represents the similarity value s_{ij} between vertices i and j . In particular, two agents are connected if the similarity is positive and, in this case, the weight of the edge will be s_{ij} .

In other words, two agents will be connected if they will have a high similarity degree. Now, the clustering problem can be restated based on the similarity graph in order to find the partition of G such that edges occurring between different groups will be denoted from very low weights (this means that they are dissimilar). Furthermore, the edges within a group will be denoted from high weights, to mean that the agents belonging to the same cluster are similar among them.

Based on the construction of the similarity graphs we model the local neighborhood relationships between the data points. In such a manner the link between users belonging to the same category (malicious or honest) will be evident, with $s_{ij} = \frac{\frac{t_{ij}+t_{ji}}{2}}{0.1 + |t_{ij} - t_{ji}|}$, where $S = \frac{\frac{T+T^t}{2}}{0.1 + |T - T^t|}$. Note that by construction, S is a symmetric matrix and its elements s_{ij} are characterized by high values if the users i and j are similar, low values if the users i and j are not similar.

More deeply, in such a way, we exploit the key principle of the Spectral Clustering techniques of considering the data as vertices of a graph and weighting the connections by the similarity between two vertices. On this basis, we assume the data of the training set as an approximation of a topological space which can be studied through the spectral properties of a matrix called the Laplacian, in order to perform an appropriate partitioning.

Accordingly to [54], the applied Spectral Clustering algorithm is the following.

Input: Similarity matrix $S \in \mathbb{R}^{n \times n}$, number k of clusters to construct.

1. Compute not normalized Laplacian $L = D - W$ where D is the degree matrix and W is the weighted adjacency matrix equal to S .
2. Compute the normalized Laplacian $L_{sym} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}}$.
3. Compute the first k eigenvectors u_1, \dots, u_k of L_{sym} .
4. Let $U \in \mathbb{R}^{n \times k}$ be the matrix containing the vectors u_1, \dots, u_k as columns.
5. Form the matrix $F \in \mathbb{R}^{n \times k}$ from U by normalizing the rows to norm 1, that is set $f_{ij} = \frac{u_{ij}}{(\sum_k u_{ik}^2)^{\frac{1}{2}}}$.
6. For $i = 1, \dots, n$ let $y_i \in \mathbb{R}^k$ be the vector corresponding to the i -th row of F .
7. Cluster the points $(y_i)_{i=1, \dots, n}$ with k -means algorithm into A_1, \dots, A_k .

Output: Cluster C_1, \dots, C_k with $C_i = \{j | y_j \in A_i\}$.

The choice of the number k of clusters provided in input to the algorithm will depend on both the number of users and the nature of social network. Moreover, the average reputation associated with each cluster is computed, according to Equation (1), as:

$$\tilde{r}_i = \frac{\sum_{j \in C_i} r_j}{|C_i|}, \quad \forall i = 1, \dots, k.$$

Observe as the malicious users' reputation is generally very low, although it can vary significantly based on context and the actions they carried out. In this respect, to discriminate honest from colluding users, based on some preliminary tests, the threshold δ has been set to $\delta = 0.46/n$ by taking into account the number n of network users.

Exploiting δ we have identified, the clusters C_1, \dots, C_m with $\tilde{r}_i \leq \delta$, with $m \leq k$ as clusters of colluding agents. Then the agents belonging to the clusters from C_1 to C_m identified as malicious colluding agents are removed from the network and the reputations of the remaining agents is recalculated as well as the average reputations of the remaining clusters. The process is reiterated until all the remaining clusters will have a reputation greater than δ .

To test the correctness of our method we carried out some preliminary tests simulating a little social network of 500 users. We did not limit ourselves to considering only honest or malicious users, because in reality users often have multiple characteristics, often not well defined, representing this a real challenge to test our method. For example, we considered malicious colluding users who receive medium-high trust values or those who receive high trust values from only a few. At the same time, we considered honest users who receive low trust values from the majority of users. In addition, we tried generating the test social network using different statistical distributions but found no significant differences, probably due to the small size of the social network. This aspect will be clarified in future research when medium to large networks will be tested.

More specifically, with Matlab we generated a dataset and the results obtained varied according to the percentage of malicious colluding agents in the network. As shown in the Table 1 and as we can see better in the Figures 1 and 2, the detected malicious clusters, therefore the malicious agents, compared to what is identified by the strategy, increase as the percentage of effective ones increases. While, the percentage of false positives, i.e. of honest agents evaluated as malicious by the proposed strategy, decreases as the number of effective malicious agents in the community increases.

Finally, with the information obtained regarding the malicious users present in the community, we were able to calculate the reputation of the users by applying EigenTrust. The previously identified malicious ones can be considered as non-pretrusted.

Let us consider the case in which the actual malicious ones are 25%. The Figure (3) represents the initial reputation obtained by solving Equation (2) and the reputation obtained by applying EigenTrust. We desire to highlight how EigenTrust significantly lowers the reputation of malicious users, while the reputation of honest users, as we know, recognized as false positive, is slightly distorted, not marking a clear difference between very good and less good honest users.

However, we consider this as the result of preliminary tests that were only meant to verify the potential of the proposed strategy. In this regard, we can consider the results obtained by applying the proposed strategy as promising.

Number of users	% of effective malicious	% of identified malicious agents	false positive rate
500	10%	72%	28%
	15%	72%	28%
	20%	80%	20%
	25%	86%	14%

Table 1
Experimental test for a social network of 500 users

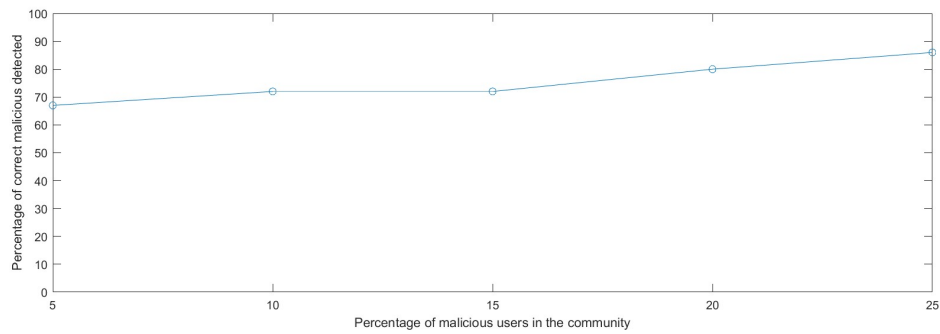


Figure 1: Percentage of identified malicious agents with clustering technique.

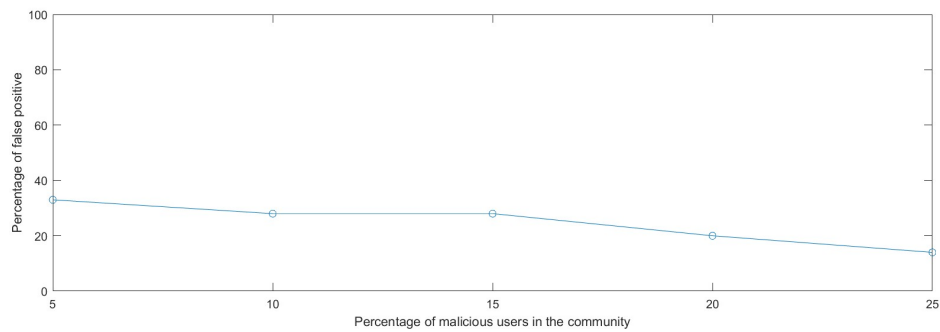


Figure 2: Percentage of honest users viewed as malicious

5. Conclusions

In this paper, we have focused on the problem of identifying in a social network, more groups of colluded agents.

To this aim, among the systems proposed in the literature, we have considered the well-known algorithm EigenTrust, that is recognized as one of the most effective solution to measure the reputation in a set of social agents. However, the EigenTrust algorithm is limited by the

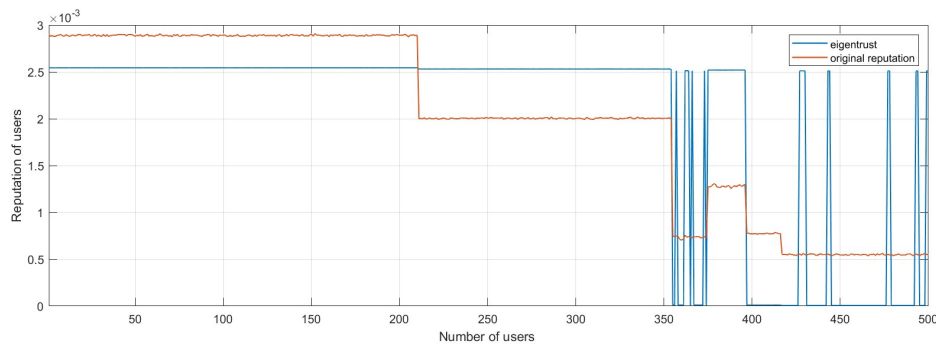


Figure 3: Users reputation in comparison

necessity of knowing a priori which agents are considered as trustworthy and the impossibility of recognizing several groups of colluding agents acting simultaneously and autonomously.

To address the problems above, in this paper we have proposed a different strategy, able to suggest, in an automatic fashion, information about the trustworthiness of agents, as required by EigenTrust, and combining a clustering stage with EigenTrust to group agents exploiting their reputation scores to detect groups of colluded agents. Then a recursive strategy has been implemented to identify each group of colluded agents and, iteration by iteration, cleaning the social network from the presence of such malicious actors.

Preliminary tests, performed by simulating a little social network, have highlighted that our method is effective in identifying several groups of colluded agents. Results have shown as such effectiveness is connected to the dimension of the social community and the percentage of colluded agents present therein.

Forthcoming research will be focused on testing our method on very large community also to test its scalability, accuracy and complexity, refine the choice of δ threshold and study a strategy for automatically assigning the number of clusters k to be examined, also by testing other clustering algorithms.

Acknowledgment

This work has been partially developed with the financial support of the: *i*) Project CAL.HUB.RIA funded by the Italian Ministry of Health, Project CUP: F63C22000530001. Local Project CUP: C33C22000540001; *ii*) Italian Research Center on High Performance Computing, Big Data and Quantum Computing (ICSC) funded by EU – NextGenerationEU (PNRR-HPC, CUP:C83C22000560007); *iii*) Multilayered Urban Sustainability Action (MUSA) funded by EU – NextGenerationEU (PNRR-MUSA, CUP:H43C2200550001).

References

- [1] C. Castelfranchi, F. D. Rosis, R. Falcone, S. Pizzutilo, Personality traits and social attitudes in multiagent cooperation, *Applied Artificial Intelligence* 12 (1998) 649–675.
- [2] M. Rheu, J. Y. Shin, W. Peng, J. Huh-Yoo, Systematic review: Trust-building factors and implications for conversational agent design, *International Journal of Human–Computer Interaction* 37 (2021) 81–96.
- [3] A. Dorri, S. Kanhere, R. Jurdak, *Multi-agent systems: A survey.*, 2018.
- [4] P. E. Petruzzi, J. Pitt, D. Busquets, Electronic social capital for self-organising multi-agent systems, *ACM Transactions on Autonomous and Adaptive Systems (TAAS)* 12 (2017).
- [5] D. Rosaci, G. M. L. Sarné, S. Garruzzo, Integrating trust measures in multiagent systems, *International Journal of Intelligent Systems* 27 (2012) 1–15.
- [6] A. Ahmed, K. Abu Bakar, M. I. Channa, K. Haseeb, A. W. Khan, A survey on trust based detection and isolation of malicious nodes in ad-hoc and sensor networks, *Frontiers of Computer Science* 9 (2015) 280–296.
- [7] F. Messina, G. Pappalardo, D. Rosaci, C. Santoro, G. M.L. Sarné, A trust-based approach for a competitive cloud/grid computing scenario, in: *Intelligent Distributed Computing VI: Proceedings of the 6th International Symposium on Intelligent Distributed Computing-IDC 2012*, Calabria, Italy, September 2012, Springer, 2013, pp. 129–138.
- [8] S. M. Sajjad, S. H. Bouk, M. Yousaf, Neighbor node trust based intrusion detection system for wsn, *Procedia Computer Science* 63 (2015) 183–188.
- [9] F. Messina, G. Pappalardo, D. Rosaci, G. M.L. Sarné, A trust-based, multi-agent architecture supporting inter-cloud vm migration in iaas federations, in: *Internet and Distributed Computing Systems: 7th International Conference, IDCS 2014*, Calabria, Italy, September 22-24, 2014. *Proceedings 7*, Springer, 2014, pp. 74–83.
- [10] W. Fang, W. Zhang, W. Chen, T. Pan, Y. Ni, Y. Yang, Trust-based attack and defense in wireless sensor networks: a survey, *Wireless Communications and Mobile Computing* 2020 (2020) 1–20.
- [11] R. Govindaraj, P. Govindaraj, S. Chowdhury, D. Kim, D.-T. Tran, A. N. Le, A review on various applications of reputation based trust management., *International Journal of Interactive Mobile Technologies* 15 (2021).
- [12] A. J. Bidgoly, B. T. Ladani, Benchmarking reputation systems: A quantitative verification approach, *Computers in Human Behavior* 57 (2016) 274–291.
- [13] T. Grandison, M. Sloman, Trust management tools for internet applications, in: *Trust Management: First International Conference, iTrust 2003 Heraklion, Crete, Greece, May 28–30, 2003 Proceedings 1*, Springer, 2003, pp. 91–107.
- [14] Y. Rizk, M. Awad, E. W. Tunstel, Decision making in multiagent systems: A survey, *IEEE Transactions on Cognitive and Developmental Systems* 10 (2018) 514–529.
- [15] S. D. Kamvar, M. T. Schlosser, H. Garcia-Molina, The eigentrust algorithm for reputation management in p2p networks, in: *Proceedings of the 12th international conference on World Wide Web*, 2003, pp. 640–651.
- [16] P. Berkhin, A survey on pagerank computing, *Internet mathematics* 2 (2005) 73–120.
- [17] M. Cotronei, S. Giuffrè, A. Marciàno, D. Rosaci, G. M. L. Sarné, Detecting collusive agents by trust measures in social iot environments: A novel reputation model, in: *Security, Trust*

- and Privacy Models, and Architectures in IoT Environments, Springer, 2022, pp. 43–61.
- [18] M. Cotronei, S. Giuffrè, A. Marcianò, D. Rosaci, G. M. L. Sarnè, Identifying colluding actors in social communities by reputation measures, in: International Conference on Applied Intelligence and Informatics, Springer, 2022, pp. 347–359.
 - [19] L.-C. Canon, E. Jeannot, J. Weissman, A dynamic approach for characterizing collusion in desktop grids, in: 2010 IEEE International Symposium on Parallel & Distributed Processing (IPDPS), IEEE, 2010, pp. 1–12.
 - [20] F. Azzedin, Taxonomy of reputation assessment in peer-to-peer systems and analysis of their data retrieval, *The Knowledge Engineering Review* 29 (2014) 463–483.
 - [21] D. H. Mcknight, M. Carter, J. B. Thatcher, P. F. Clay, Trust in a specific technology: An investigation of its components and measures, *ACM Transactions on management information systems (TMIS)* 2 (2011) 1–25.
 - [22] J. Wu, D. Balliet, P. A. Van Lange, Reputation, gossip, and human cooperation, *Social and Personality Psychology Compass* 10 (2016) 350–364.
 - [23] G. Bolton, B. Greiner, A. Ockenfels, Engineering trust: reciprocity in the production of reputation information, *Management science* 59 (2013) 265–285.
 - [24] P. Resnick, K. Kuwabara, R. Zeckhauser, E. Friedman, Reputation systems, *Communications of the ACM* 43 (2000) 45–48.
 - [25] A. Josang, R. Ismail, C. Boyd, A Survey of Trust and Reputation Systems for Online Service Provision, *Decision Support System* 43 (2005) 618–644.
 - [26] F. Hendrikx, K. Bubendorfer, R. Chard, Reputation systems: A survey and taxonomy, *Journal of Parallel and Distributed Computing* 75 (2015) 184–197.
 - [27] G. Fortino, L. Fotia, F. Messina, D. Rosaci, G. M. L. Sarnè, Trust and reputation in the internet of things: state-of-the-art and research challenges, *IEEE Access* 8 (2020) 60117–60125.
 - [28] P. De Meo, L. Fotia, F. Messina, D. Rosaci, G. M. L. Sarnè, Providing recommendations in social networks by integrating local and global reputation, *Information Systems* 78 (2018) 58–67.
 - [29] S. A. Ghasempouri, B. T. Ladani, Modeling trust and reputation systems in hostile environments, *Future Generation Computer Systems* 99 (2019) 571–592.
 - [30] K. Hoffman, D. Zage, C. Nita-Rotaru, A survey of attack and defense techniques for reputation systems, *ACM Computing Surveys (CSUR)* 42 (2009) 1–31.
 - [31] A. Jøsang, J. Golbeck, Challenges for robust trust and reputation systems, in: *Proceedings of the 5th International Workshop on Security and Trust Management (SMT 2009)*, Saint Malo, France, volume 5, Citeseer, 2009.
 - [32] P. Resnick, R. Zeckhauser, F. E., K. Kuwabara, Reputation systems, *Communication of ACM* 43 (2000) 45–48.
 - [33] S. Vavilis, M. Petković, N. Zannone, A reference model for reputation systems, *Decision Support Systems* 61 (2014) 147–154.
 - [34] <http://www.ebay.com> , 2022, ????
 - [35] L. Cabral, A. Hortacsu, The dynamics of seller reputation: Evidence from ebay, *The Journal of Industrial Economics* 58 (2010) 54–78.
 - [36] S. C. Hayne, H. Wang, L. Wang, Modeling reputation as a time-series: Evaluating the risk of purchase decisions on ebay, *Decision Sciences* 46 (2015) 1077–1107.
 - [37] P. Resnick, R. Zeckhauser, Trust among strangers in internet transactions: Empirical

- analysis of ebay's reputation system, in: *The Economics of the Internet and E-commerce*, Emerald Group Publishing Limited, 2002.
- [38] L. Xiong, L. Liu, Peertrust: Supporting reputation-based trust for peer-to-peer electronic communities, *IEEE transactions on Knowledge and Data Engineering* 16 (2004) 843–857.
 - [39] F. Messina, G. Pappalardo, D. Rosaci, C. Santoro, G. M. L. Sarné, A trust-aware, self-organizing system for large-scale federations of utility computing infrastructures, *Future Generation Computer Systems* (2015).
 - [40] S. Kamvar, M. Schlosser, H. Garcia-Molina, The eigentrust algorithm for reputation management in P2P networks, in: *Proc. of World Wide Web, 12th International Conference on*, ACM, 2003, pp. 640–651.
 - [41] R. Jansen, T. Kaminski, F. Korsakov, A. Saint Croix, D. Selifonov, A priori trust vulnerabilities in eigentrust, *Technical Report University of Minnesota* (2008).
 - [42] X. Fan, L. Liu, M. Li, Z. Su, Eigentrust⁺⁺: Attack resilient trust management, in: *8th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, IEEE, 2012, pp. 416–425.
 - [43] H. A. Kurdi, Honestpeer: An enhanced eigentrust algorithm for reputation management in p2p systems, *Journal of King Saud University-Computer and Information Sciences* 27 (2015) 315–322.
 - [44] T. Huynh, N. Jennings, N. Shadbolt, An integrated trust and reputation model for open multi-agent systems, *Autonomous Agents and Multi-Agent Systems* 13 (2006) 119–154.
 - [45] G. Swamynathan, K. C. Almeroth, B. Y. Zhao, The design of a reliable reputation system, *Electronic Commerce Research* 10 (2010) 239–270.
 - [46] Y. Liu, Y. Yang, Y. L. Sun, Detection of collusion behaviors in online reputation systems, in: *2008 42nd Asilomar Conference on Signals, Systems and Computers*, IEEE, 2008, pp. 1368–1372.
 - [47] Y. Sun, Y. Liu, Security of online reputation systems: The evolution of attacks and defenses, *IEEE Signal Processing Magazine* 29 (2012) 87–97.
 - [48] N. Zhang, W. Yu, X. Fu, S. K. Das, Maintaining defender's reputation in anomaly detection against insider attacks, *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 40 (2009) 597–611.
 - [49] J. Kang, Z. Xiong, D. Niyato, D. Ye, D. I. Kim, J. Zhao, Toward secure blockchain-enabled internet of vehicles: Optimizing consensus management using reputation and contract theory, *IEEE Transactions on Vehicular Technology* 68 (2019) 2906–2920.
 - [50] M. Yu, M. Zhou, W. Su, A secure routing protocol against byzantine attacks for manets in adversarial environments, *IEEE transactions on vehicular technology* 58 (2008) 449–460.
 - [51] C. Xu, J. Zhang, Combating product review spam campaigns via multiple heterogeneous pairwise features, in: *Proceedings of the 2015 SIAM International Conference on Data Mining*, SIAM, 2015, pp. 172–180.
 - [52] J. Hu, D. Fang, X. Wei, J. Xie, Colluder detection based on hypergraph decomposition, in: *2013 Ninth International Conference on Computational Intelligence and Security*, IEEE, 2013, pp. 630–634.
 - [53] M. Cotronei, S. Giuffrè, A. Marcianò, D. Rosaci, G. M. L. Sarné, Improving the effectiveness of eigentrust in computing the reputation of social agents in presence of collusion, *International Journal of Neural Systems* (2023) 2350063–2350063. doi:10.1142/

S0129065723500636.

- [54] A. Ng, M. Jordan, Y. Weiss, On spectral clustering: Analysis and an algorithm, *Advances in neural information processing systems* 14 (2001).