Detecting Hidden Communities in Online Auction Networks

Kai Zhu, Yong Guan and Lei Ying
Department of Electrical and Computer Engineering, Iowa State University
Email: {kzhu, guan, leiying}@iastate.edu

Abstract— **Online auction networks often use reputation-based systems to help users assess each other's honesty and integrity. Fraudsters, however, can collude with accomplices to accumulate bogus positive feedback to manipulate the reputation systems. In this paper, we model an online auction network with fraudsters as a random network with hidden communities (fraudsters and associated accomplices), and propose a maximum likelihood framework to detect the fraudsters. We develop an iterative message passing algorithm to heuristically solve the maximum likelihood detection problem. This algorithm identifies fraudsters and accomplices in a distributed fashion and is a scalable solution. The algorithm converges in a finite number of iterations and has very high detection rates according to our simulations.**

I. Introduction

Online auction networks, such as eBay and taobao.com, have become popular trading platforms, with a large variety of products available with competitive prices. Today, these networks have hundreds of billions dollars in trading volume, and hundreds of millions dollars in revenue.

While online auction networks have many advantages over traditional retail stores, many people are still reluctant to sell/buy products on these networks with the concern that sellers/buyers on these networks may not be reliable. To help users assess each other's honesty and integrity, online auction networks often use reputation-based systems. For example, eBay allows the seller and buyer to leave feedback to each other for each transaction and the feedback may be viewed by other users. A seller/buyer with more positive comments can be regarded as a more reliable user.

Fraudsters, however, can collude with accomplices to accumulate bogus positive feedback to manipulate the reputation systems, which makes it harder to evaluate a user's reliability according to the reputation (feedback). It has been observed in [1] that the fraudsters and accomplices are likely to form a dense bipartite core as the fraudsters receive most of the feedback from the accomplices, and are interested in receiving a large number of feedback comments as quick as she/he can. This structure property allows us to detect fraudsters based on the network structure instead of relying only on the reputations. In this paper, we model an online auction network with fraudsters as a random network with hidden communities (fraudsters and associated accomplices). We propose a maximum likelihood framework for detecting fraudsters. The main contributions of this paper include:

- We formulate a maximum likelihood detection problem for fraud detection based on a random graph model. The random graph model assumes the underlying graph is a bipartite graph (sellers/buyers) with an embedded complete bipartite core (fraudsters and accomplices). The model assumes the observed transaction graph (i.e., there is a link between a buyer and a seller if there is a transaction between them) is generated by the underlying graph. The maximum likelihood detection problem is to detect the embedded bipartite core based on the observed transaction graph.
- To solve the maximum likelihood detection problem, we propose an iterative message passing algorithm. The message passing algorithm exploits the bipartite nature of the network, where when the set of fraudsters/accomplices is fixed, a seller/buyer can be optimally classified as a fraudster/accomplice based on local information. The algorithm converges in a finite number of iterations.
- We conduct extensive simulations on synthesize networks, and compare the performance of the message passing algorithm with two existing clustering algorithms — Spectral Clustering [2] and Stochastic Flow Clustering [3]. From the simulations, we observe our algorithm has high detection rates and low false positive/negative rates; and significantly outperforms the two existing clustering algorithms.

A. Related Work

In [1], [4], the authors studied fraud detection in online auction networks, and observed the existence of bipartite cores between fraudsters and accomplices. The authors developed a fraud detection algorithm by combining user-level features and the standard sum-product brief-propagation algorithm [5], [6]. In this paper, we propose a random bipartite graph model for online auction networks and formulate the fraud detection problem as a maximum likelihood detection problem. We further develop an iterative message passing algorithm based on some structure properties of the maximum likelihood estimator. The random graph model in this paper is similar to the model used for community detection [7]–[9]. However, our detection algorithm relies on the bipartite structure of the auction networks, so is different from those in [7]–[9].

II. Model

Nodes in an online auction network can be classified into two categories: sellers, who supply goods; and buyers, who purchase goods. So an auction network can be represented as

978-1-4673-3140-1/12/$31.00 ©2012 IEEE

a bipartite graph where a link between a seller and a buyer indicates that there is at least one transaction between them. We let s_i denote the i^{th} seller and b_j denote the j^{th} buyer. Figure 1 is a simple example of the bipartite representation of an auction network. The figure shows that seller s_1 has sold items to buyer b_4, and seller s_2 has traded with both buyers b_2 and b_3.

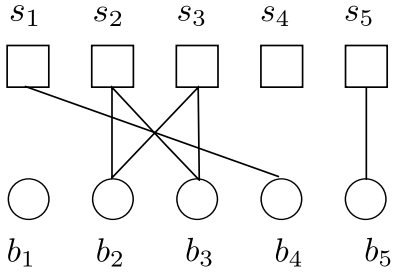

Fig. 1. A bipartite graph representation of online auction networks, where the rectangles represent sellers and circles represent buyers.

As we mentioned in the introduction, fraudsters can collude with accomplices to accumulate bogus positive feedback to manipulate reputation systems. So reputation (or feedback) along is not sufficient for assessing users' trustworthiness. We need to further understand the "value" of the feedback comments, which may be evaluated according to the structure of the network. In [1], it has been observed that fraudsters and accomplices tend to form dense bipartite cores since the fraudsters are interested in gaining as many positive comments as possible in a short period of time and the accomplices are heavily reused.

To detect potential fraudsters in auction networks, we need to identify these bipartite cores. However, fraudsters and accomplices can hide the bipartite cores by either purposely trading with honest users to behave like normal sellers/buyers, or purposely avoiding trading with a subset of fraudsters/accomplices to reduce the density of the bipartite cores.

To analytically study this problem, we assume the following model. We assume the network consists of n_s sellers and n_b buyers. Each seller s_i is associated with an indicator function x_i such that $x_i = 1$ if s_i is a fraudster and $x_i = 0$ otherwise. Similarly, buyer b_j is associated with an indicator function y_j such that $y_j = 1$ if b_j is an accomplices and $y_j = 0$ otherwise. For simplicity, we assume only one bipartite core in the network. We further assume that the single bipartite core consists of n_f fraudsters and n_a accomplices. In other words, $\sum_{i=1}^{n_s} x_i = n_f$ and $\sum_{i=1}^{n_b} y_i = n_a$. The vectors $\mathbf{x}$ and $\mathbf{y}$ contain the complete information of fraudsters and accomplices. So the fraud detection is to estimate $\mathbf{x}$ and $\mathbf{y}$.

Denote by $\mathcal{G} = \{\mathcal{S}, \mathcal{B}, \mathcal{L}\}$ the transaction network we observe, where $\mathcal{S}$ is the set of sellers, $\mathcal{B}$ is the set of buyers and $\mathcal{L}$ is the links between sellers and buyers. The adjacent matrix $\mathbf{W}$ associated with $\mathcal{G}$ is

$$W_{ij} = \begin{cases} 1, & \text{seller } s_i \text{ and buyer } b_j \text{ are connected;} \\ 0, & \text{otherwise.} \end{cases}$$

We assume $\mathbf{W}$ is randomly generated based on $\mathbf{x}$ and $\mathbf{y}$ such that

$$\Pr(W_{ij} = 1) = \begin{cases} p_{fa}, & \text{if } x_i = 1 \text{ and } y_j = 1 \\ p_{fh}, & \text{if } x_i = 1 \text{ and } y_j = 0 \\ p_{ha}, & \text{if } x_i = 0 \text{ and } y_j = 1 \\ p_{hh}, & \text{if } x_i = 0 \text{ and } y_j = 0. \end{cases} \quad (1)$$

The probabilities p_{fa}, p_{fh}, p_{ha} and p_{hh} are assumed to be known.

Let $\mathbf{w}$ be a realization of random matrix. Detecting the hidden bipartite cores can be formulated as the following maximum likelihood detection problem:

$$\arg\max_{\mathbf{x},\mathbf{y}} \Pr(\mathbf{w}|\mathbf{x}, \mathbf{y}). \quad (2)$$

While in theory, given $\mathbf{w}$ and $\mathbf{p} = (p_{fa}, p_{fh}, p_{ha}, p_{hh})$, we can always find $\mathbf{x}$ and $\mathbf{y}$ that maximize the likelihood. The number of possible configurations of $\mathbf{x}$ and $\mathbf{y}$ are $2^{n_s+n_b}$. So an exclusive search is computationally expensive when the network size is large. Therefore, the focus of this paper is to develop efficient and scalable detection algorithms.

The links in the auction network can be classified into four classes. We define the following notations:

- m_{fa} : number of links between fraudsters and accomplices,
- m_{fh} : number of links between fraudsters and honest buyers,
- m_{ha} : number of links between honest sellers and accomplices, and
- m_{hh} : number of links between honest sellers and accomplices.

With these notations, we have:

$$\begin{aligned} &\Pr(\mathbf{W} = \mathbf{w}|\mathbf{x}, \mathbf{y}) \\ =&\prod_{i=1,\cdots,n_s, j=1,\cdots,n_b} \Pr(W_{ij} = w_{ij}|\mathbf{x}, \mathbf{y}) \\ =&p_{fa}^{m_{fa}}(1-p_{fa})^{n_f n_a - m_{fa}} p_{fh}^{m_{fh}}(1-p_{fh})^{n_f(n_b-n_a)-m_{fh}} \\ &\times p_{ha}^{m_{ha}}(1-p_{ha})^{(n_s-n_f)n_a - m_{ha}} \\ &\times p_{hh}^{m_{hh}}(1-p_{hh})^{(n_s-n_f)(n_b-n_a)-m_{hh}}. \end{aligned}$$

III. An Iterative Message Passing Algorithm

In this section, we present an iterative message passing algorithm to solve (2). The algorithm is based on the following two propositions. Proposition 1 states that if the set of fraudsters are given, a buyer can be optimally classified according to the node degree, the number of links to fraudsters and the total number of fraudsters. Similar result holds when the set of accomplices are given. We next present the two propositions. The detailed proof is omitted due to the lack of space and can be found in [10].

Proposition 1: Assume $\mathbf{x}$ is given (i.e., n_f is known) and define

$$\mathbf{y}(\mathbf{x}) = \arg\max_{\mathbf{y}} \Pr(\mathbf{w}|\mathbf{x}, \mathbf{y}).$$

Let d_j denote the degree of node j and let d_j^f denote the number of links between buyer j and the fraudsters. Then, $y_j(\mathbf{x}) = 1$ (i.e., buyer j is classified as an accomplice) if

$$\kappa_1 d_j^f - \kappa_2(d_j - d_j^f) > \kappa_3 n_f - \kappa_4 n_s,$$

and $y_j(\mathbf{x}) = 0$ otherwise, where

$$\begin{aligned}
\kappa_1 &= \log \frac{p_{fa}(1-p_{fh})}{(1-p_{fa})p_{fh}} \\
\kappa_2 &= \log \frac{(1-p_{ha})p_{hh}}{p_{ha}(1-p_{hh})} \\
\kappa_3 &= \log \frac{(1-p_{ha})(1-p_{fh})}{(1-p_{fa})(1-p_{hh})} \\
\kappa_4 &= \log \frac{1-p_{ha}}{1-p_{hh}}.
\end{aligned}$$

□

Proposition 2: Assume $\mathbf{y}$ is given (i.e., n_a is known) and define

$$\mathbf{x}(\mathbf{y}) = \arg\max_{\mathbf{x}} \Pr(\mathbf{w}|\mathbf{x}, \mathbf{y}).$$

Let d_i denote the degree of node i and let d_i^a denote the number of links between seller i and the accomplices. Then, $x_i(\mathbf{y}) = 1$ (i.e, seller i is classified as a fraudster) if

$$c_1 d_i^a - c_2(d_i - d_i^a) > c_3 n_a - c_4 n_b,$$

and $x_i(\mathbf{y}) = 0$ otherwise, where

$$\begin{aligned}
c_1 &= \log \frac{p_{fa}(1-p_{ha})}{(1-p_{fa})p_{ha}} \\
c_2 &= \log \frac{(1-p_{fh})p_{hh}}{p_{fh}(1-p_{hh})} \\
c_3 &= \log \frac{(1-p_{fh})(1-p_{ha})}{(1-p_{fa})(1-p_{hh})} \\
c_4 &= \log \frac{1-p_{fh}}{1-p_{hh}}.
\end{aligned}$$

□

According to Proposition 1, given $\mathbf{x}$, the decision to determine whether buyer j is an accomplice is based on local information d_j^f, the number of links from buyer j to the fraudsters, and $d_j - d_j^f$, the number of links from buyer j to honest sellers. A buyer with large d_j^f and small $d_j - d_j^f$ is likely to be an accomplice since the buyer has more connections to the fraudsters than to the honest sellers.

Similarly, given $\mathbf{y}$ (so n_a is known), the decision to declare whether a seller i is a fraudster or not is based on local information d_i^a and $d_i - d_i^a$.

Based on these two propositions, we propose an iterative message passing algorithm. The algorithm consists of two steps — A-step and F-step. In an A-step, we fix the identities of sellers and decide the identities of buyers according to Proposition 1. In an F-step, we fix the identities of the buyers and decides the identities of sellers according to Proposition 2. We repeat these two steps iteratively until no node changes the identity. Next, we describe the algorithm in details.

Algorithm 1 Iterative Message Passing Algorithm:

- **Notation**
  - ϕ_i : the set of neighbors of node i
  - $z(t)$: the value of z at the t^{th} iteration
  - $\psi_{ij}(t)$: the message sent from node i to node j at the t^{th} iteration.
- **Initialization:** Set $x_i(0) = y_j(0) = 0$ except for the identified fraudster (say seller s_l), $x_l(0) = 1$, $n_a(0) = 0$ and $n_f(0) = 0$. Seller s_l sends a message $\psi_l(0) = 1$ to all its neighbors.
- Each iteration consists of two steps: F-step and A-step. Consider iteration $t > 0$.

  A-step:
  - The network updates $n_f(t)$ such that
  $$n_f(t) = n_f(t-1) + \sum_{i=1}^{n_s} \psi_i(t).$$
  - For each buyer,
    - Each buyer j updates $d_j^f(t)$ such that
    $$d_j^f(t) = d_j^f(t-1) + \sum_{i \in \phi_j} \psi_i(t).$$
    - Each buyer j check the following inequality
    $$\kappa_1 d_j^f(t) - \kappa_2(d_j - d_j^f(t)) > \kappa_3 n_f(t) - \kappa_4 n_s,$$
    and sets $y_j(t) = 1$ if the inequality holds and $y_j(t) = 0$ otherwise.
    - If $y_j(t) \neq y_j(t-1)$, seller j sends message
    $$\psi_j(t) = y_j(t) - y_j(t-1)$$
    to all the neighbors.

  F-step:
  - The network updates $n_a(t)$ such that
  $$n_a(t) = n_a(t-1) + \sum_{j=1}^{n_b} \psi_j(t-1).$$
  - For each seller,
    - Each seller i updates $d_i^a(t)$ such that
    $$d_i^a(t) = d_i^a(t-1) + \sum_{l \in \phi_i} \psi_l(t-1).$$
    - Each seller i check the following inequality
    $$c_1 d_i^a(t) - c_2(d_i - d_i^a(t)) > c_3 n_a(t) - c_4 n_b,$$
    and sets $x_i(t) = 1$ if the inequality holds and $x_i(t) = 0$ otherwise.
    - If $x_i(t) \neq x_i(t-1)$, seller i sends message
    $$\psi_i(t) = x_i(t) - x_i(t-1)$$
    to all its neighbors.
- The algorithm repeats F-step and A-step until no message is passed.

A pictorial description of the algorithm is given in Figure 2

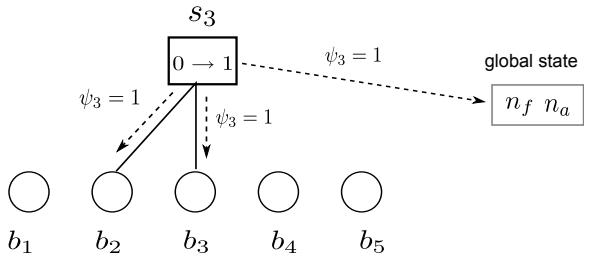

Fig. 2. When the state of seller s_3 changes from 0 to 1, i.e., the seller is declared to be a fraudster, a message $\psi_3 = 1$ is sent from seller 3 to its neighbors, buyer b_2 and b_3, and the global state (n_f, n_a).

The following theorem states that the iterative message passing algorithm terminates in a finite number of iterations.

Theorem 1: Let $d_{\max}$ denote the maximum node degree in the network. The iterative message passing algorithm requires $O(\max\{n_s, n_b\}(d_{\max}+1))$ message exchanges for each step, and converges in a finite number of steps. □

The detailed proof is omitted due to the lack of space and can be found in [10].

IV. Evaluation

In this section, we evaluate the message passing algorithm using simulations. We evaluate the detection rate, false positive rate and false negative rate of the algorithm. Denote by r_f, α_f and β_f the detection rate, false positive rate, and false negative rate of detecting fraudsters, respectively. The notations r_a, α_a, and β_a are similarly defined for detecting accomplices.

A. Synthesis Random Graph

We first validate the performance of the iterative message passing algorithm on random graphs.

The network consists of $5,000$ sellers and $5,000$ buyers. Among them, we randomly picked 10 sellers as fraudsters and 10 buyers as accomplices. The link formation probability is

$$p_{hh} = 0.01,\ p_{ha} = 0.01,\ p_{fh} = 0.001 \text{ and } p_{fa} = 0.9.$$

We set $p_{hh} = p_{ha}$ by assuming that the accomplices purposely behave like honest buyers to avoid being detected, which makes the detection difficult and is a worst case assumption. Furthermore, we assume that one fraudster has been identified.

First, we varied p_{fa} from 0.1 to 0.9. For each p_{fa}, the results are the average of the results on 100 randomly generated graphs. The detection of fraudsters is perfect ($r_f = 100\%$) for all p_{fa}s. Figure 3 shows the results of detecting accomplices. We can see that the detection is not perfect when p_{fa} is small, but still close to 80%. We note that even when the detection rate of accomplices is not 100% when p_{fa} is small, the detection rate of fraudsters remains to be perfect. This can be explained by the fact that the false positive rate of accomplices (α_a) is near zero even when p_{fa} is small. So the claimed accomplices are the actual accomplices. A subset of accomplices is then sufficient for the algorithm to detect the fraudsters.

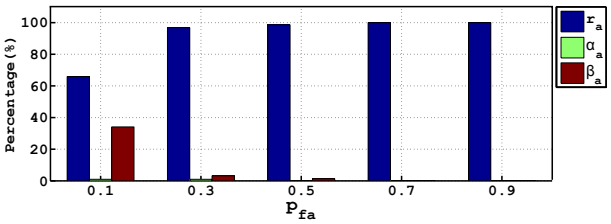

Fig. 3. The Results of the Message Passing Algorithm with p_{fa} Varied from 0.1 to 0.9

In Figure 4, we varied p_{fh} from 0.001 to 0.01. We observed that when $p_{fh} > 0.5p_{hh}$, the detection rate of fraudsters becomes very low (close to 10%). The detection of accomplices, however, is very accurate for all p_{fh}s. The low detection rate of fraudsters can be explained by the fact when the fraudsters receive a significant portion of comments from honest buyers, it is hard to tell fraudsters from honest sellers. On the other hand, the false positive rate α_f is low which means the detected fraudsters are actual fraudsters and with these fraudsters, we can still detect accomplices accurately.

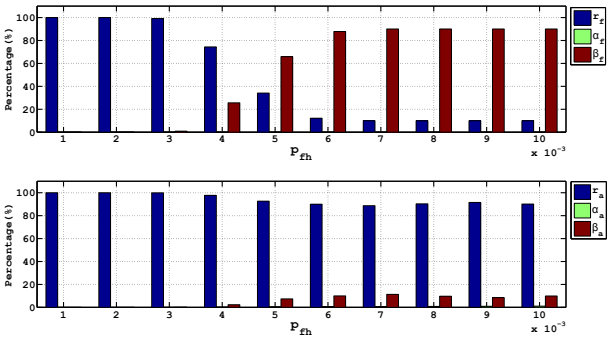

Fig. 4. The Results of the Message Passing Algorithm with p_{fh} Varied from 0.001 to 0.01

B. Comparing with other clustering algorithms

An alternative approach to detect the hidden community is to use the existing community detection algorithms to detect community structures of auction networks. In this set of simulations, we compared the iterative message passing algorithm (called MP) with two existing community detection algorithms: Spectral Clustering (SP) [2] and Stochastic Flows Clustering (SF) [3]. First, we conducted experiments on random graphs with the following link formation probabilities

$$p_{hh} = 0.1,\ p_{ha} = 0.1,\ p_{fh} = 0.01 \text{ and } p_{fa} = 0.9.$$

The networks contain 100 sellers and 100 buyers. Among them, we randomly picked 10 sellers as fraudsters and 10 buyers as accomplices.

We varied p_{fa} from 0.1 to 0.9 and averaged the results over 100 randomly generated graphs. The results are shown in Figure 5. The Stochastic Flow Clustering algorithm yields high

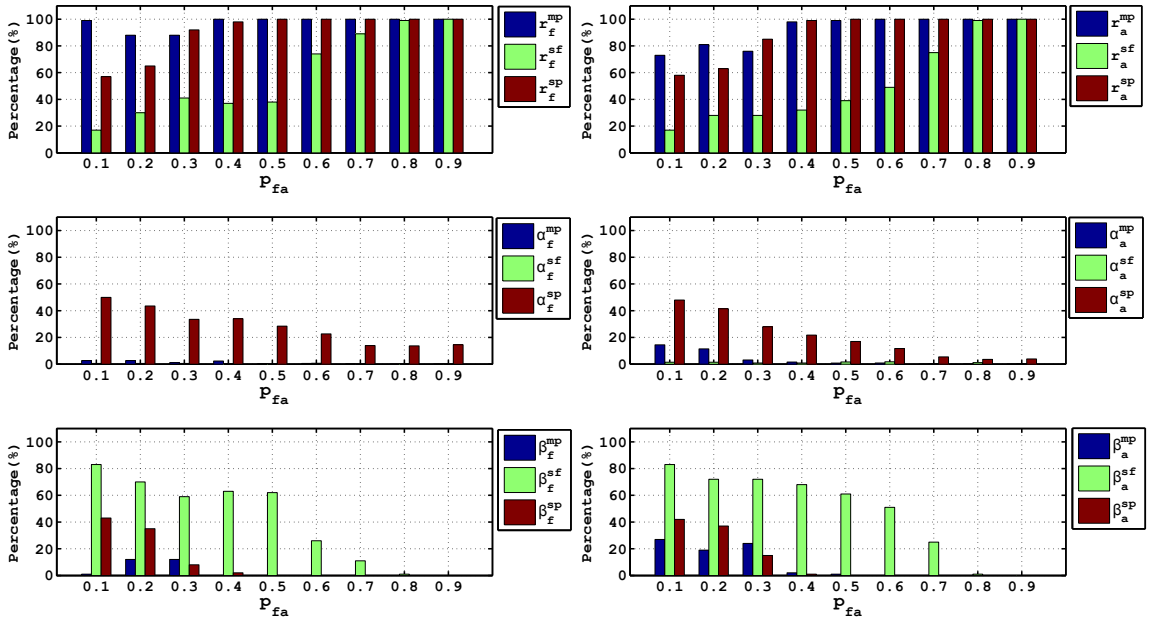

Fig. 5. Comparison of SP, SF and MP with p_{fa} Varied from 0.1 to 0.9

detection rates only when $p_{fa} \geq 0.8$, and the false negative rates are very high when p_{fa} is small. While the Spectral Clustering has a very high detection rate, the false positive rate is high as well (30% when $p_{fa} = 0.5$ and 15% when $p_{fa} = 0.9$.) In a summary, the iterative message passing algorithm significantly outperforms the other two algorithms in all three performance metrics.

C. Unknown link formation probabilities

In the simulations above, we assume the link formation probabilities are known. While the probability $p_{hh} \approx |\mathcal{L}|/n_s n_b$ when the number of fraudsters and accomplices are small, the other three link formation probabilities depend on the sets of fraudsters and accomplices and may be difficult to estimate. In this simulation, we generalize our algorithm so that it can work for the case where the link formation probabilities are unknown.

First, we randomly guess the link formation probabilities, and apply our algorithm to decide $\mathbf{x}$ and $\mathbf{y}$. Then, based on the $\mathbf{x}$ and $\mathbf{y}$ produced by the algorithm, we can estimate the link formation probabilities as follows:

$$p_{fa}(\tau+1) = \frac{m_{fa}(\tau)}{n_f(\tau)n_a(\tau)}$$
$$p_{fh}(\tau+1) = \frac{m_{fh}(\tau)}{n_f(\tau)(n_b - n_a(\tau))}$$
$$p_{ha}(\tau+1) = \frac{m_{ha}(\tau)}{(n_s - n_f(\tau))n_a(\tau)}$$
$$p_{hh}(\tau+1) = \frac{m_{hh}(\tau)}{(n_s - n_f(\tau))(n_b - n_a(\tau))}.$$

Then we can apply our algorithm again to decide $\mathbf{x}$ and $\mathbf{y}$, and then re-compute the link formation probabilities. We repeat this procedure until no change in both the community structure and link formation probabilities. The algorithm is called the modified iterative message passing algorithm.

In this set of simulations, the underlying random graphs are generated using the same configuration in Section IV-A. The initial value of p_{fa} is uniformly distributed over $(0.5, 1)$. Since the numbers of fraudsters and accomplice are small compare to honest users, we set the initial value p_{hh} to $|\mathcal{L}|/n_s n_b$. Based on p_{hh}, we further set $p_{fh} = 0.2p_{hh}$ and $p_{ha} = p_{hh}$ as initial values. Since there is no guarantee that the modified iterative message passing algorithm will converge, we say the detection fails if the algorithm does not converge after ten different initial configurations are used.

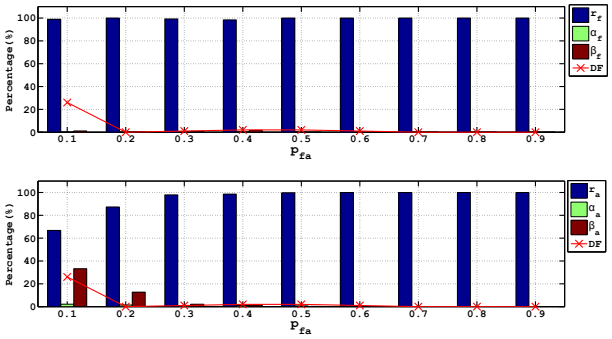

Fig. 6. The Results of the Modified Message Passing Algorithm with p_{fa} Varied from 0.1 to 0.9

First, we varied p_{fa} from 0.1 to 0.9. The results are presented in Figure 6, where DF is the frequency of detection failure, and the detection rate, false positive/negative rates are the averages over all cases where the algorithm converged. When p_{fa} is small, more than 20% graphs ended up with a detection failure. As p_{fa} increases, the algorithm converged in most cases and the performance is similar to that of Figure

3. Therefore, the algorithm will converge when fraudster and accomplice have enough connections. Furthermore, we varied p_{fh} from 0.001 to 0.01. The results are shown in Figure 7. The algorithm always converged when $p_{fh} < 0.005$ and has a high detection rate, similar to that of Figure 4.

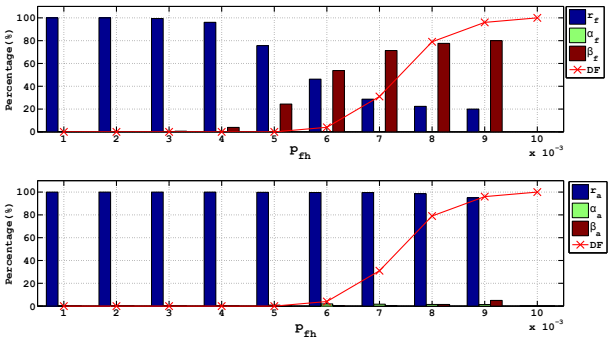

Fig. 7. The Results of the Modified Message Passing Algorithm with p_{fh} Varied from 0.001 to 0.01

In a summary, the simulation results shows that in most cases, the modified algorithm performs similarly with the original algorithm without knowing the link formation probabilities.

V. Conclusion

In this paper, we studied the fraud detection problem in online auction networks. We proposed a random network model and formulated the problem as a maximum likelihood detection problem. We developed a simple and scalable message passing algorithm to solve the maximum likelihood problem. From the simulations, we observed the message passing algorithm can accurately detect the fraudsters and accomplices in synthesis graphs and performs much better than clustering-based detection algorithms.

VI. Acknowledgements

Research supported in part by NSF Grants CNS-0644238 and CNS-0831470.

References

[1] D. H. Chau, S. Pandit, and C. Faloutsos, "Detecting fraudulent personalities in networks of online auctioneers," in *Proceedings of the 10th European conference on Principle and Practice of Knowledge Discovery in Databases*, 2006, pp. 103–114.

[2] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Proceedings of Advances in Neural Information Proceeding Systems*, 2001, pp. 849–856.

[3] V. Satuluri and S. Parthasarathy, "Scalable graph clustering using stochastic flows: applications to community discovery," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2009, pp. 737–746.

[4] S. Pandit, D. H. Chau, S. Wang, and C. Faloutsos, "Netprobe: a fast and scalable system for fraud detection in online auction networks," in *Proceedings of the 16th international conference on World Wide Web*, 2007, pp. 201–210.

[5] S. Aji and R. McEliece, "The generalized distributive law," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 325 –343, 2000.

[6] M. J. Wainwright and M. I. Jordan, "Graphical models, exponential families, and variational inference," *Found. Trends Mach. Learn.*, vol. 1, pp. 1–305, 2008.

[7] J. Copic, M. O. Jackson, and A. Kirman, "Identifying community structures from network data via maximum likelihood methods," *The BE Journal of Theoretical Economics*, vol. 9, no. 1, p. Article 30, 2009.

[8] M. E. J. Newman and Leicht, "Mixture models and exploratory analysis in networks." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, pp. 9564–9, 2007.

[9] T. A. B. Snijders and K. Nowicki, "Estimation and prediction for stochastic blockmodels for graphs with latent block structure," *Journal of Classification*, vol. 14, no. 1, pp. 75–100, 1997.

[10] K. Zhu, Y. Guan, and L. Ying, "Detecting hidden communities in online auction networks," 2012, Technical Report, Iowa State University.