Survey report of FRAUDAR: Bounding Graph Fraud in the Face of Camouflage

CS 249 Winter 2017
Team BTF

Lizhi Zeng(304593058)
Dui Lin(504759948)
Pei Jiang(604685278)
Zhiming Zhuang(204593426)
Yang Guo(104588741)

I. Introduction

As the rising popularity of social networks like Twitter, large online service providers such as Amazon, TripAdvisor, or Twitter have sooner-or-later inevitably to cope with fraudulent activities on their platforms to remain trustworthy and reliable in the eye of the general public. On these platforms, fraudsters intentionally misrepresent facts and circumstances about themselves, their products, or their services to attain unjustified benefits that are created based on erroneous beliefs of honest users. To evade detection, fraudsters constantly adapt their strategies and strive to blend into the environment as best as possible **[1]**. One cunning way to disguise fraudulent behavior is to use "camouflage" in which smart fraudsters appear to look "normal" by also having connections to honest objects. The detection of such fraudulent user accounts on large platforms is a challenging task, even with the aid of analytical models.

Previous state-of-the-art algorithms exploit the density signal and has shown good accuracy. For example, SpokEn has investigated community structure in large social Mobile Call graphs, and reported the presence of a surprising pattern, which in the term of EigenSpokes, when singular vectors of the graphs are plotted against each other **[2]**. They showed sufficient conditions for origin of EigenSpokes (well-knit communities in sparse graphs) as well as their persistence across several different Mobile Call graphs from various times and regions. And another algorithm named NetProbe, has proposed a novel way to model users and transactions on an auction site as a Markov Random Field **[3]**. They have also shown how to tune the well-known belief propagation algorithm so as to identify suspicious patterns such as bipartite cores. However, Smart fraudsters will also try to 'look normal', by adding links to popular items/idols (like famous singers/actors, or well-liked products) - this behavior is called "camouflage" in the recent literature, and unfortunately, previous algorithms can not handle the camouflage.

The focus of this paper is to spot fraudsters in the presence of camouflage or hijacked accounts. The authors propose FRAUDAR, an algorithm that is stated to have the following properties: camouflage-resistant, provides upper bounds on the effectiveness of fraudsters, and is effective in real-world data. Experimental results under various attacks show that FRAUDAR outperforms the top competitor in accuracy of detecting both camouflaged and non-camouflaged fraud. Additionally, in real-world experiments with a Twitter follower-followee graph of 1.47 billion edges, FRAUDAR successfully detected a subgraph of more than 4000 detected accounts, of which a majority had tweets showing that they used follower-buying services.

This survey first introduces some state-of-the-art algorithms and enumerates the pros and cons compared to Frauder. Then we conclude and explain the highlights of Frauder, namely how this novel algorithm can successfully detect the fraudsters under camouflage, and the

provable limit on undetectable fraud the algorithm provides. Finally, we summary the potential drawbacks of Frauder and propose future outlooks and our suggestions.

II. Current State-of-the-Art

Fraud detection has caused serious concern in recent years. Many existing methods attempt to detect fraud through review texts. However, these approaches are typically not robust: spammers can carefully select their review texts to avoid detection. Or even without knowledge of the detection system, they may mimic normal user reviews as closely as possible so as to make themselves look normal. And these kind of camouflage behaviors make it more difficult to detect fraudsters.

Basically, we can view this problem into a graph-based detection model. Graph-based methods detect groups of spammers or fake followers, often by identifying unexpectedly dense regions of the graph of users and products. Such approaches are potentially harder to evade, as creating fake reviews unavoidably generates edges in the graph.

Global methods:

Building on singular value decomposition (SVD), latent factor models, and belief propagation (BP), these model the entire graph to find fraud. SPOKEN **[2]** considered the "eigen-spokes" pattern produced by pairs of eigenvectors of graphs, and was later generalized for fraud detection. FBOX **[4]** builds on SVD but focuses on detecting attacks missed by spectral techniques. Several methods have used HITS-like ideas to detect fraud in graphs **[5]**. BP has been used for fraud classification on eBay **[6]**, and fraud detection. All of these methods have been successful in finding fraud but they offer no guarantees of robustness. **[7]** performs adversarial analysis for spectral algorithms, showing that attacks of small enough scale will necessarily evade detection methods which rely on the top k SVD components.

Dense subgraph mining:

Finding dense subgraphs has been an important focus of graph theory communities and has been studied from a wide array of perspectives **[8]**. Most closely related to ours is Charikar's work on finding subgraphs with large average degree **[9]**, which shows that subgraph average degree can be optimized with approximation guarantees. Variants have been proposed to efficiently find large, dense subgraphs **[10]**, with approximation guarantees. To our knowledge, however, this is the first work which adapts this theoretical perspective to the challenges of fraud detection and camouflage resistance, and achieves meaningful bounds for our application. Moreover, our work differs from these in its setting of bipartite graphs, and in the use of edge reweighting to further increase accuracy.

Local clustering methods:

Clustering is one of the classic problems in both machine learning and data mining, with a wide range of methods still being developed. A different direction for fraud detection focuses on local subgraphs, by analyzing the properties of ego nets to detect fraud **[11]**. COPYCATCH **[12]** and GETTHESCOOP **[13]** use local search heuristics to find relevant dense bipartite subgraphs. However, without guarantees on the search algorithm, the algorithms may not be robust to intelligent adversaries.

By comparing different methods, we can find that COPYCATCH, CatchSync, BP-based methods, SPOKEN, FBOX, GETTHECOOP, all these approaches can detest dense blocks but they don't have theoretical guarantees. Moreover, besides COPYCATCH, all other methods are not camouflage- resistant. This is to say, they could not detect the fake followers to add edge to honest users to make them less suspicious. Our proposed method FRAUDAR is the only one that matches all specifications.

III. Model and Algorithm

In this part, we discuss the prediction model with algorithm FRAUDAR and highlight the key contributions of this paper from four perspectives: (a) The introduction of suspicious metrics and their advantages; (b) The FRAUDAR algorithm based on greedy selection and priority tree; (c) Theoretical bound to guarantee the FRAUDAR can produce at least half optimal value; (d) Selection of particular suspicious metrics to make sure camouflage-resistance.

1. Metrics

As demonstrated in Figure 1, the paper introduces a novel family of metrics with intuitive axioms that provides several advantages as a suspiciousness metrics. Ultimately, this class of metrics allows proving formally that the density metric is camouflage-resistance and offers theoretical guarantees for the detection of fraud.

$$g(\mathrm{S}) = \frac{f(\mathrm{S})}{|\mathrm{S}|} \tag{1}$$

$$f(S) = fv(\mathrm{S}) + f\varepsilon(\mathrm{S}) = \sum_{i \in S} a_i + \sum_{i,j \in S \wedge (\mathrm{i,j}) \in \varepsilon} c_{ij} \tag{2}$$

There are many advantages to metrics of this form. Firstly, metrics of this form can be optimized in a way that is (a) scalable; (b) offers theoretical guarantees, and (c) is robust to camouflage, as we demonstrate in the rest of this paper. All three of these properties will hold due to the particular chosen form in (1) and (2). Secondly, metrics of this form obey a

number of basic properties or called axioms, that a reasonable suspiciousness metric should meet, as shown below:

AXIOM 1: (NODE SUSPICIOUSNESS). *A subset consisting of higher suspiciousness nodes is more suspicious than one consisting of lower suspiciousness nodes, if the other conditions are fixed.*

AXIOM 2: (EDGE SUSPICIOUSNESS). *Adding edges within a subset increases the suspiciousness of the subset if the other conditions are fixed.*

AXIOM 3: (SIZE). *Assuming node and edge weights are all equal, larger subsets are more suspicious than smaller subsets with the same edge density.*

AXIOM 4: (CONCENTRATION). *A subset with smaller size is more suspicious than one with the same total suspiciousness but larger size.*

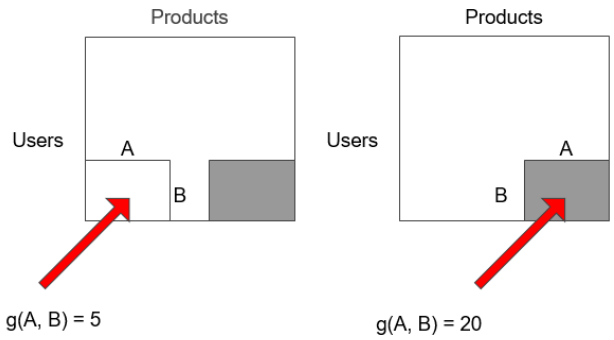

Figure 1: Suspicious Metrics given by A and B, the dark parts represent the abnormal dense subgraph.

2. Algorithm

The pseudo-code in Figure 2 has shown how the FRAUDAR algorithm works. This paper start with the entire set of nodes $U \cup W$, then repeatedly remove the node which results in the highest value of g evaluated on the remaining set of nodes. Formally, denote by X the current set they are optimizaing over; Initially they set $X = U \cup W$. Let $\Delta_i = f(\mathrm{X}\backslash\{\mathrm{i}\}) - f(\mathrm{X})$ be the change in f when we remove i from X . They then repeat this process; they recompute the values of Δ_j, then choose the next node to delete, and so on. This leads to a shrinking series of sets X over time, denote $X_0, \ldots\ldots, \mathrm{X}_{m+n}$ of size $\mathrm{m+n}, \ldots\ldots, 0$. At the end, they return the one of these that maximizes the density metric g.

Require: Bipartite $G = (\mathcal{U} \cup \mathcal{W}, \mathcal{E})$; density metric g of the form in (1)
procedure FRAUDAR (G, g)
  Construct priority tree T from $\mathcal{U} \cup \mathcal{W}$ ▷ see Section 4.2
  $\mathcal{X}_0 \leftarrow \mathcal{U} \cup \mathcal{W}$ ▷ suspicious set is initially the entire set of nodes $\mathcal{U} \cup \mathcal{W}$
  for $t = 1, \ldots, (m + n)$ **do**
    $i^* \leftarrow \arg\max_{i \in \mathcal{X}_i} g(\mathcal{X}_i \setminus \{i\})$ ▷ exonerate least suspicious node
    Update priorities in T for all neighbors of i^*
    $\mathcal{X}_t \leftarrow \mathcal{X}_{t-1} \setminus \{i^*\}$
  end for
  return $\arg\max_{\mathcal{X}_i \in \{\mathcal{X}_0, \ldots, \mathcal{X}_{m+n}\}} g(\mathcal{X}_i)$ ▷ return most suspicious set $\mathcal{X}_i$
end procedure

Figure 2: FRAUDAR Algorithm Pseudo-code.

The key fact that allows the algorithm to be efficient is the forms for f and g in (1) and (2). When i is removed, the only values of Δ_j which need to be updated are those where j is a neighbor of i. This is because for all other j, the expressions (1) and (2) ensure that Δ_j does not change. Hence, the updates are fast: for each $(i, j) \in \varepsilon$, over the lifetime of the algorithm they will perform at most one such update over this edge, for a total of $O(|\varepsilon|)$ updates. Here, using a priority tree can improve the performance significantly. Each update can be performed in $O(log|V|)$ time, totalling $O(|\varepsilon|log|V|)$ time.

3. Theoretical Bound

Up to now, we have shown that g can be optimized in near-linear time. In this section, we will show that when f and g are of the form (1) and (2), FRAUDAR is guaranteed to return a solution of at least half of the optimum value by the following theorem:

THEOREM: Let A and B be the set of users and objects returned by FRAUDAR. Then:

$$g(\mathrm{A} \cup \mathrm{B}) \geq \mathrm{g}_{OPT} \tag{3}$$

where g_{OPT} is the maximum value of g, i.e.

$$\mathrm{g}_{OPT} = \max_{A',B'} g(\mathrm{A}' \cup \mathrm{B}') \tag{4}$$

4. Edge Weights and Camouflage Resistance

The metrics they select are resistant to camouflage which means they don't allow fraudulent users to make themselves less suspicious by adding camouflage edges, i.e. edges toward honest objects.

The key idea of FRAUDAR is that instead of treating every edge equally, they assign a lower weight C_{ij} when the target object j has high degree.

$$C_{ij} = h(\mathrm{d}_j) = \frac{1}{\log(\mathrm{d}_j + \mathrm{c})} \tag{5}$$

This is because objects of very high degree are not necessarily suspicious (since highly popular objects commonly exist). Thus, this weighting, more specifically column-weighting, allow putting greater emphasis on objects within unexpectedly dense subgraphs, rather than just high degree objects.

IV. Outlooks

Although FRAUDAR algorithm can outperform many competitors and is proven to have effectiveness on the real-world data, we still believe there are some issues need to be addressed. First, the algorithm's behavior on bipartite graphs without any camouflages is not clear. It would be helpful to conduct experiments on such graphs and evaluate the algorithm's false positive rate. Second, the metrics used for scoring the suspiciousness of nodes and edges are ad-hoc. It is possible that different metrics result in better or worse results. Thus, as a future work, the authors can evaluate other information retrieval metrics and find a method for choosing the best metric for different bipartite graphs.

V. Team Member Contributions

- **Lizhi Zeng:** Related works in background research and survey writting;
- **Dui Lin:** Related works in Fraudar algorithm research and paper presentation;
- **Pei Jiang:** Related works in experiments research and survey writting;
- **Zhiming Zhuang:** Related works in edge weights and camouflage-resistance research and paper presentation;
- **Yang Guo:** Related works in experiments research and survey writting;

VI. References

[1] Baesens, B., Van Vlasselaer, V., & Verbeke, W. (2015). Fraud analytics using descriptive, predictive, and social network techniques: a guide to data science for fraud detection. John Wiley & Sons..

[2] Prakash, B. A., Sridharan, A., Seshadri, M., Machiraju, S., & Faloutsos, C. (2010, June). Eigenspokes: Surprising patterns and scalable community chipping in large graphs. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 435-448). Springer Berlin Heidelberg.

[3] Pandit, S., Chau, D. H., Wang, S., & Faloutsos, C. (2007, May). Netprobe: a fast and scalable system for fraud detection in online auction networks. In Proceedings of the 16th international conference on World Wide Web (pp. 201-210). ACM.

[4] Shah, N., Beutel, A., Gallagher, B., & Faloutsos, C. (2014, December). Spotting suspicious link behavior with fbox: An adversarial perspective. In Data Mining (ICDM), 2014 IEEE International Conference on (pp. 959-964). IEEE.

[5] Cao, Q., Sirivianos, M., Yang, X., & Pregueiro, T. (2012, April). Aiding the detection of fake accounts in large scale social online services. In Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation (pp. 15-15). USENIX Association.

[6] Ghosh, S., Viswanath, B., Kooti, F., Sharma, N. K., Korlam, G., Benevenuto, F., ... & Gummadi, K. P. (2012, April). Understanding and combating link farming in the twitter social network. In Proceedings of the 21st international conference on World Wide Web (pp. 61-70). ACM.

[7] Shah, N., Beutel, A., Gallagher, B., & Faloutsos, C. (2014, December). Spotting suspicious link behavior with fbox: An adversarial perspective. In Data Mining (ICDM), 2014 IEEE International Conference on (pp. 959-964). IEEE.

[8] Giatsidis, C., Thilikos, D. M., & Vazirgiannis, M. (2011, July). Evaluating cooperation in communities with the k-core structure. In Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on (pp. 87-93). IEEE.

[9] Charikar, M. (2000, September). Greedy approximation algorithms for finding dense components in a graph. In International Workshop on Approximation Algorithms for Combinatorial Optimization (pp. 84-95). Springer Berlin Heidelberg.

[10] Perozzi, B., Akoglu, L., Iglesias Sánchez, P., & Müller, E. (2014, August). Focused clustering and outlier detection in large attributed graphs. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 1346-1355). ACM.

[11] Beutel, A., Xu, W., Guruswami, V., Palow, C., & Faloutsos, C. (2013, May). Copycatch: stopping group attacks by spotting lockstep behavior in social networks. In Proceedings of the 22nd international conference on World Wide Web (pp. 119-130). ACM.

[12] Jiang, M., Cui, P., Beutel, A., Faloutsos, C., & Yang, S. (2014, May). Inferring strange behavior from connectivity pattern in social networks. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 126-138). Springer International Publishing.

[13] Jiang, M., Cui, P., Beutel, A., Faloutsos, C., & Yang, S. (2014, May). Inferring strange behavior from connectivity pattern in social networks. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 126-138). Springer International Publishing.