 CMC, vol.60, no.1, pp.317-333, 2019

# Network Embedding-Based Anomalous Density Searching for Multi-Group Collaborative Fraudsters Detection in Social Media

**Chengzhang Zhu[1, 2], Wentao Zhao[2, *], Qian Li[1], Pan Li[2] and Qiaobo Da[3]**

**Abstract:** Detecting collaborative fraudsters who manipulate opinions in social media is becoming extremely important in order to provide reliable information, in which, however, the diversity in different groups of collaborative fraudsters presents a significant challenge to existing collaborative fraudsters detection methods. These methods often detect collaborative fraudsters as the largest group of users who have the strongest relation with each other in the social media, consequently overlooking the other groups of fraudsters that are with strong user relation yet small group size. This paper introduces a novel network embedding-based framework NEST and its instance BEST to address this issue. NEST detects multiple groups of collaborative fraudsters by two steps. In the first step, to disclose user collaboration, it represents users according to their social relations. Then, in the second step, to identify the collaborative fraudsters, it detects the user groups with anomalous large group density in its representation space. BEST instantiates NEST by using a bipartite network embedding method to represent users and adopting a fast density group detection method based on the k-dimensional tree. Our experiments show BEST (i) performs significantly better in detecting fraudsters on four real-word social media data sets, and (ii) effectively detects multiple groups of collaborative fraudsters, compared to three state-of-the-art competitors.

**Keywords:** Fraudster detection, network embedding, social media.

## 1 Introduction

The reliability of social media content is becoming increasingly significant because social media heavily affects people every day. Unfortunately, a large proportion of social media content is proposed by fraudsters who collaborate to manipulate social opinions driven by huge profit and incentives of reputation [Mukherjee, Venkataraman, Liu et al. (2013); Xiang, Li, Hao et al. (2018)]. As a result, effectively detecting such collaborative fraudsters is critical and with great bossiness values [Akoglu, Chandy and Faloutsos (2013)].

Recent year has seen significant progress made in fraudsters detection. Current efforts

[1] Faculty of Engineering and Information Technologies, University of Technology Sydney, 15 Broadway, Ultimo NSW 2007, Australia.

[2] College of Computer, National University of Defense Technology, Changsha, China.

[3] School of Economics and Management, Beijing Institute of Graphic Communication, Beijing, China.

* Corresponding Author: Wentao Zhao. Email: wtzhao@nudt.edu.cn.

mainly focused on extracting fraudster indicators and/or features from users' behavior [Mukherjee, Liu and Glance (2012); Ye and Akoglu (2015); Hooi, Shin, Song et al. (2017)] or users' proposed content [Mukherjee, Venkataraman, Liu et al. (2013); Wang, Liu and Zhao (2017); You, Qian and Liu (2018)]. Because of the great distinguishing ability of anomalous behavior and content, these indicators and/or features have shown remarkable performance in detecting individual fraudsters [Rayana and Akoglu (2016)].

However, identifying fraudsters with collaborative manipulation is a challenging task. Specifically, the collaborative manipulation poses the two major challenges below: (i) The content of collaborative fraudsters may not be anomalous because the collaborative manipulation may dominate social opinions. (ii) The professional fraudsters will imitate the behavior of honest users to evade inspection [Hooi, Song, Beutel et al. (2016)]. These two challenges cause the failure of current behavior and content-based fraudsters detection methods in detecting collaborative fraudsters.

To detect collaborative fraudsters, the dense subgraph mining methods [Hooi, Song, Beutel et al. (2016); Hooi, Shin, Song et al. (2017); Wu, Hu, Morstatter et al. (2017); Liu, Hooi and Faloutsos (2017); Xiang, Shen, Qin et al. (2018); Xiang, Zhao, Li et al. (2018)] are the major solutions, which detect collaborative fraudsters according to the significant collaboration footprint. Specifically, the dense subgraph mining methods always detect collaborative fraudsters as the largest group of users who have the strongest relation with each other in the social media. However, in this way, they may overlook the other groups of fraudsters that are with strong user relation yet small group size. In reality, social media may contain multiple groups of collaborative fraudsters instead of only the largest group of collaborative fraudsters.

In this paper, we introduce a novel Network Embedding-based denSiTy subgraph mining (NEST for short) framework for multi-group collaborative fraudsters detection in social media. Specifically, NEST first represents users according to their social relations to disclose user collaboration. In this process, users who have similar activities will be embedded near to each other in the representation space. NEST then detects the user groups with anomalous large group density in its representation space to identify the collaborative fraudsters. Accordingly, any group of collaborative fraudsters with large joint activities can be effectively detected.

Essentially, this detection procedure simultaneously tackles three challenges brought by collaborative fraudsters: content domination, behavior camouflage, and multiple fraudsters groups, resulting in a robust and comprehensive collaborative fraudsters detecting result. In the first step, NEST solves the content domination and behavior camouflage problems by distilling user social relations which are reflected in users' joint activities. The rationale is that the cooperation of collaborative fraudsters to manipulate opinions cannot be avoided. In the second step, NEST discovers fraudsters groups by analyzing the outlier of group density in its representation space. The intuition is that the joint activities of collaborative fraudsters must be more frequent than honest users, but the number of fraudsters is much less than honest users.

We further implement NEST by proposing a Bipartitie networking Embedding-based fast denSiTy subgraph mining method based on the k-dimensional tree structure, termed BEST. Specifically, BEST first models the users and their activities as a bipartite network

as demonstrated in Fig. 1. In the bipartite network, the nodes on each side are users and activities, and a link refers to a user participates in an activity. Then, to comprehensively capture user collaborations, BEST represents users by embedding both the explicit and implicit relations in the bipartite network. Lastly, to fast detect the collaborative fraudsters, BEST builds a k-dimensional tree for the representation space and searches the anomalous density group based on the k-dimensional tree.

Accordingly, this paper makes two major contributions:

- We introduce a novel network embedding-based framework NEST for identifying collaborative fraudsters in social media. NEST represents users according to their social relations and detects fraudsters by analyzing the outlier of group density in the representation space. It results in a more reliable and comprehensive collaborative fraudsters detection, compared to existing dense subgraph mining-based solutions.
- We instantiate NEST to an effective and efficient multi-group collaborative fraudsters detection method, BEST, by introducing bipartite network embedding and k- dimensional tree-based anomalous density group searching. The bipartite network embedding captures both explicit and implicit user relations, and the k-dimensional tree-based method guarantees the efficiency of density groups searching.

Extensive empirical results show that (i) BEST performs significantly better in detecting fraudsters on four large real-world social media data sets; and (ii) BEST effectively detects multiple groups of collaborative fraudsters, compared to three state-of-the-art competitors.

## 2 Related work

### *2.1 Fraudster detection*

Current efforts on fraudster detection can be roughly classified into two categories: *individual characteristics-based methods and relational characteristics-based methods.*

The *individual characteristics-based methods* use the user proposed content and/or user's behavior to identify whether a user is a fraudster. The information used by these methods mainly include the statics and linguistic characteristics of a content [Li, Huang, Yang et al. (2011); Mukherjee, Kumar, Liu et al. (2013); Wang, Liu and Zhao (2017); You, Qian and Liu (2018)], and the historical actions of a user [Fei, Mukherjee, Liu et al. (2013); Mukherjee, Venkataraman, Liu et al. (2013)]. These individual characteristics are designed as features for fraudster detection [Jindal and Liu (2008); Lim, Nguyen, Jindal et al. (2010); Zhao, Resnick and Mei (2015); Li, Fei, Wang et al. (2017)]. However, as evidenced by Hooi et al. [Hooi, Song, Beutel et al. (2016)], the individual characteristics are not robust for collaborative fraudsters who jointly manipulate social opinions and fraudsters may imitate the behavior of honest users.

The *relational characteristics-based methods* capture user-activity, user-user, and activity-activity relations, typically via a graph [Pandit, Chau, Wang et al. (2007); Stringhini, Kruegel and Vigna (2010); Akoglu, Chandy and Faloutsos (2013); Junqué de Fortuny, Stankova, Moeyersoms et al. (2014); Akoglu, Tong and Koutra (2015); Shehnepoor, Salehi, Farahbakhsh et al. (2017)]. They hold an assumption that fake reviews are manipulated by groups of fraudsters. With this assumption, they assume a

group of fraudsters will have dense links to a group of manipulated activities (user-activity relation) [Akoglu, Chandy and Faloutsos (2013); Wang, Xie, Liu et al. (2011)], a group of fraudsters will co-occur in many activities (user-user relation) [Wu, Hu, Morstatter et al. (2017); Sun, Qu, Chakrabarti et al. (2005); Xu, Zhang, Chang et al. (2013)], and different manipulated activities will have overlapped linked fraudsters (activity-activity relation) [Hovy (2016)].

Although current methods show their strengths to disclose fraudsters, most of them fail to discover multiple groups of collaborative fraudsters in social network. In this paper, we propose a networking-embedding based framework NEST to fill the gaps of multi- group collaborative fraudsters detection. The proposed NEST achieves a more reliable and comprehensive detection by revealing users within density groups in its representation space, which delicately embeds the user's social relationships.

### *2.2. Network embedding*

Our proposed method is based on network embedding, which can be categorized into two types: matrix factorization (MF)-based and neural network-based methods.

MF-based methods involve linear [Cox and Cox (2000)] and nonlinear [Nedich and Ozdaglar (2008)] procedures in the embedding process. While the linear procedures adopt linear transformations, such as singular value decomposition (SVD) and multiple dimensional scaling (MDS), to generate low-dimensional embedding [Cox and Cox (2000)], the non- linear methods utilize nonlinear transformations, e.g. kernel PCA and manifold learning, to capture complicated data structures. However, both have high computational cost because of their eigen-decomposition operation on data matrix. Accordingly, these methods do not suit for large social network embedding.

Recently, neural network-based methods have shown the state-of-the-art performance. Followed by DeepWalk [Perozzi, Al-Rfou and Skiena (2014)] and Node2Vec [Grover and Leskovec (2016)], most of neural network-based methods reformulate a network embedding task as a word embedding task via performing truncated random walks in a network to convert the network to sentences. More recently, advanced work embeds both explicit and implicit relations in a network and shows its significance [Tang, Qu, Wang et al. (2015); Wang, Cui and Zhu (2016); Cao, Lu and Xu (2015); Xu, Wei, Cao et al. (2017)]. However, the above methods are not designed for social network embedding. They treat the nodes in a network homogeneously, and thus, cannot capture the difference between user and activity in social media. In addition, the truncated random walks used in these methods do not consider the user-activity joint distribution in social network.

In this paper, we instantiate NEST as an effective and efficient method, BEST, via a bipartite network embedding method. This Bipartitie network embedding method is tailored for social media. Accordingly, it captures user-activity relations better in its user representation space, which provides a solid foundation for collaborative fraudsters detection.

## 3 NEST for collaborative fraudster detection

NEST framework adopts a two-steps procedure to detect collaborative fraudsters in social media. The workflow of NEST framework is shown in Fig. 1. For a social media $S$ with

a set of users $U = \{u_1, u_2, \cdots, u_n\}$ and a set of activities $A = \{a_1, a_2, \cdots, a_m\}$, in the first step, NEST extracts a bipartite network $G$ from $S$ as $G = (U, A, E)$, where $U$ and $A$ are the nodes on the two sides of $G$, respectively, and $E \subseteq U \times V$ defines the inter-set edges. Here, each edge in $E$ carries a non-negative weight $w_{ij}$, reflecting the strength between a user $u_i$ and an activity $a_j$, and the $w_{ij}$ will be zero if the user $u_i$ does not join the activity $a_j$. Accordingly, the weights in the bipartite network can be represented by a $n \times m$ matrix $W = [w_{ij}]$. Then, NEST learns an embedding function $f(\cdot): U \to R^d$, which maps a user $u_i$ to a $d$ dimensional vector representation $\mathbf{u}_i$. The embedding function $f(\cdot)$ should capture and embed the social relations of users in the bipartite network into their representation space. In the second step, NEST finds the anomalous density groups in the user representation space and treats the users in the anomalous density groups as collaborative fraudsters.

Formally, NEST detects a set of collaborative fraudster groups $F = \{G'_1, \cdots, G'_k\}$ according to

$$dist(\mathbf{u}_i, \mathbf{u}_j) < \eta \ \ and \ \ |G'_l| \geq \varepsilon, \forall u_i, u_j \in G'_l, \forall l = 1, \cdots, k, \tag{1}$$

where $G'_l$ is a subgraph of $G$, $dist(\cdot,\cdot): R^d \times R^d \to R$ is a distance measure building on the user representation space, $|\cdot|$ refers a density measurement, and $\eta$ and $\varepsilon$ are two parameters which control the density range and the density anomalous degree, respectively. Essentially, NEST embeds the collaboration footprints of users into a vector space where users joined similar activities will be located together. Therefore, the density of a group of users in the vector space reflects the degree of collaboration between this group of users. The larger density a group of users has, the more collaboration between them. Because collaborative fraudsters may have much more cooperation [Hooi, Song, Beutel et al. (2016)], NEST can effectively detect collaborative fraudsters by searching the groups with anomalous density in the user representation space. Different from typical graph subgraph mining methods, which only disclose a single group of collaborative fraudsters, i.e., the users in the largest dense subgraph, NEST provides a more comprehensive detection result that contains multiple groups of collaborative fraudsters.

NEST has a good generalizability since it can be instantiated by specifying any network embedding method and any anomalous density groups searching method. We introduce an instance of NEST in next section and then verify its performance by empirical analyses.

## 4 A NEST instance: BEST

BEST instantiates NEST by a bipartite network embedding method catering for social net- work, and a *k*-dimensional tree-based anomalous density group searching method for efficient fraudsters detection.

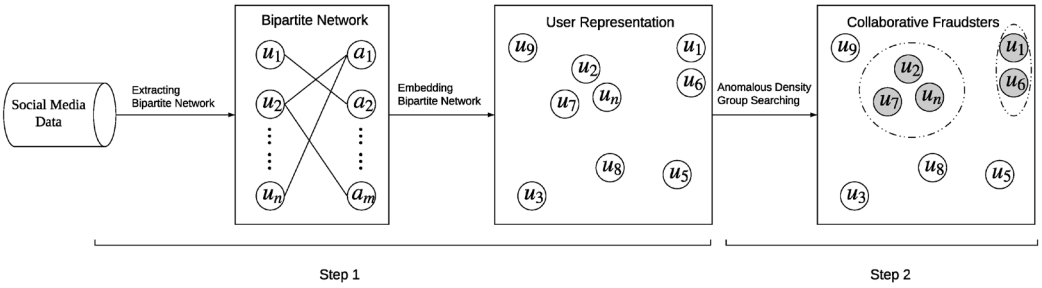



**Figure 1:** NEST Framework. In the first step, NEST extracts a bipartite network from social media data, and represents user into a vector space by embedding their social relation in the bipartite network. In the second step, NEST searches the anomalous density group of users in the representation space for collaborative fraudsters detection. The detected collaborative fraudsters are illustrated with a grey background, and their corresponding groups are highlighted by a dotted circle

### *4.1 Bipartite network embedding*

The network embedding reveals and embeds social relations of a user into the user's vector representation, which reflects the cooperation of users in social media. We introduce a bipartite network embedding method to jointly capture the explicit and implicate relations of users in social media.

#### *4.1.1 Explicit relations embedding*

The explicit relations refer to the direct links between users and activities, which reflect the activities a user jointed. If two users always joint similar activities, their similarity should be large in the representation space.

To preserve the explicit relations, we keep the preference of users in their representation space. Specifically, we measure the preference of a user in both social media and representation space, and make the preference of a user in representation space similar to that in social media. For the preference measurement in social media, we consider the probability of a user join in an activity. Given the bipartite network, this probability can be calculated as follows:

$$P(u_i, a_j) = \frac{w_{ij}}{\sum_{e_{ij} \in E} w_{ij}} \tag{2}$$

where $w_{ij}$ is the weight of edge $e_{ij}$. The measurement reflects the preference distribution of users. We follow the setting of word2vec to use the sigmoid function to measure the interaction of a user and an activity in their representation space in a probability space:

$$\hat{P}(u_i, a_j) = \frac{1}{1 + \exp(-\mathbf{u}_i^{\mathrm{T}} \mathbf{a}_j)}, \tag{3}$$

where $\mathbf{u}_i \in R^d$ and $\mathbf{a}_j \in R^d$ are the embedding vectors of $u_i$ and $a_j$, respectively. Then, we adopt KL-divergence to measure the difference between $P$ and $\hat{P}$, and optimize the

user and activity representation to minimize the KL-divergence as follows:

$$\min_{\mathbf{u},\mathbf{a}} \sum_{e_{i,j}\in E} P(u_i,a_j)\log(\frac{P(u_i,a_j)}{\hat{P}(u_i,a_j)}). \tag{4}$$

Considering $P(i,j)$ is a constant, minimizing the Eq. (4) equals to follows:

$$\min_{\mathbf{u},\mathbf{a}} -\sum_{e_{i,j}\in E} w_{i,j}\log\hat{P}(u_i,a_j). \tag{5}$$

*4.1.2 Implicit relations embedding*

The implicit relations refer to the relations between users and activities that are not directly connected. For two users, if there exist a path between them in the bipartite network, they may have an implicit relation, and the weight of the path reflects the strength of this implicit relation. However, counting the paths between two nodes in a bipartite network has a great high complexity, which is impracticable in social media.

Inspired by DeepWalk [Perozzi, Al-Rfou and Skiena (2014)], we also perform a truncated random walks on the network to generate nodes corpus as random walk paths, which contain higher order implicit relations between nodes. We move a step further to reconstruct the bipartite network $G$ as two networks where each network only contains users $G^{(u)}$ or activities $G^{(a)}$, and conduct random walks on these two transformed networks. It results in a stationary distribution of random walks on social media data [Gao, Chen, He et al. (2018)]. In $G^{(u)}$, $u_i$ and $u_j$ will have an edge $e_{ui,uj}$ if exists a $t_k$ that $e_{ui,ak}\in E$ and $e_{uj,ak}\in E$ where $E$ is the edge set of $G$. In $G^{(a)}$, $a_i$ and $a_j$ will have an edge $e_{ai,aj}$ if exists a $u_k$ that $e_{uk,ai}\in E$ and $e_{uk,aj}\in E$ where $E$ is the edge set of $G$.

The random walk paths generation procedure is illustrated in Algorithm 1, which generates a set of random walk paths $D^{(u)}$ of $U$, a set of random walk paths $D^{(a)}$ of $A$.

The implicit relations embedding aims to maximize the conditional probability of the context of a node. For user corpus $D^{(u)}$, it maximizes the conditional probability as follows:

$$\max_{\mathbf{u}} \prod_{\mathbf{u}_i\in S\wedge S\in D^{(u)}} \prod_{\mathbf{u}_c\in C_S(\mathbf{u}_i)} \frac{\exp(\mathbf{u}_i^{\top}\mathbf{u}_c)}{\sum_{k=1}^{n}\exp(\mathbf{u}_i^{\top}\mathbf{u}_k)}, \tag{6}$$

where $S$ refers to the sequence in the context, $C_S(\mathbf{u}_i)$ refers the context nodes of node $u_i$ in sequence $s$. Similarly, for activities corpus $D^{(a)}$ the implicit embedding maximizes the conditional probability as:

$$\max_{\mathbf{a}} \prod_{\mathbf{a}_j\in S\wedge S\in D^{(a)}} \prod_{\mathbf{a}_c\in C_S(\mathbf{a}_j)} \frac{\exp(\mathbf{a}_j^{\top}\mathbf{a}_c)}{\sum_{k=1}^{m}\exp(\mathbf{a}_i^{\top}\mathbf{a}_k)}, \tag{7}$$

**Algorithm 1:** Random walks paths generation for bipartite network

**Input:** Bipartite graph $G=(U,A,E)$, Weight matrix $W$ of $G$, maximum walk per vertex $l_{\max}$, minimum walk per vertex $l_{\min}$, walk stopping probability $p$.

**Output**: A set of random walk paths $D^{(u)}$ of $U$, and the $D^{(a)}$ of $A$.

**foreach** $u_i$ in $U$ **do**

1 **foreach** $u_j$ in $U$ **do**

2 **if** $\exists a, e_{u_i,a_k} \in E$ and $e_{u_i,a_k} \in E$ **then**

3 $E^{(a)} \leftarrow e_{a_i,a_j}$ ;

4 $w_{a_i,a_j} = \sum_{e_{u_k,a_i}, e_{u_k,a_j} \in E} w_{u_k,a_i} \cdot w_{u_k,a_j}$;

5 **end**

6 **end;**

7 $v_{right} \leftarrow$ BUILDKDTREE( $\mathbf{u}^{(2)}$ , depth+1) ;

**foreach** $a_i$ in $A$ **do**

9 **foreach** $a_j$ in $A$ **do**

10 **if** $\exists a, e_{u_k,a_i} \in E$ and $e_{u_k,a_j} \in E$ **then**

11 $E(a) \leftarrow e_{a_i,a_j}$ ;

12 $w_{a_j,a_j} = \sum_{e_{u_k,a_i}, e_{u_k,a_j} \in E} w_{u_k,a_i} \cdot w_{u_k,a_j}$;

13 **end**

14 **end**

15 $D^{(u)} = \{\}, D^{(a)} = \{\}$ ;

**foreach** $u_i$ in $U$ **do**

17 $l = \max\left(\frac{\sum_{e_{u_i,u_j} \in E^{(u)}} w_{u_i,u_j}}{\sum_{e_{u_k,u_l} \in E^{(u)}} w_{u_k,u_l}} \cdot l_{\max}, l_{\min}\right)$, $P = \{u_i\}$ ;

18 **for** $m$ =1 to $l$ **do**

19 Draw $e_{u_i,u_j} \sim \frac{w_{u_i,u_j}}{\sum_{e_{u_i,u_k} \in E^{(u)}} w_{u_i,u_k}}$

20 P $\leftarrow u_i, u_i = u_j$ ;

21 Draw $s \sim$ uniform(0, 1) ;

22 **if** $s \le p$ **then** break;

23 **end**

24 $D^{(u)} \leftarrow P$ ;

25 **end**

**foreach** $u_i$ in $U$ **do**

27 $l = \max\left(\frac{\sum_{e_{a_i,a_j} \in E^{(a)}} w_{a_i,a_j}}{\sum_{e_{a_k,a_l} \in E^{(a)}} w_{a_k,a_l}} \cdot l_{\max}, l_{\min}\right)$, $P = \{a_i\}$ ;

**for** $m$=1 to $l$ **do**
  Draw $e_{a_i,a_j} \sim \frac{w_{a_i,a_j}}{\sum_{e_{a_i,a_k} \in E^{(a)}} w_{a_i,a_k}}$
  $P \leftarrow a_i, a_i = a_j$ ;
  Draw $s \sim$ uniform(0, 1) ;
  **if** $s \le p$ **then** break;
**end**
$D^{(a)} \leftarrow P$ ;
**end**
**return** $D^{(u)}, D^{(a)}$.

BEST jointly considers the explicit and implicit relations embedding, forming a joint embedding objective function:

$$\min_{\mathbf{u},\mathbf{a}} -\alpha \sum_{e_{ij} \in E} w_{ij} \log \hat{P}(u_i, a_j) - \beta \prod_{u_i \in S \wedge S \in D^{(u)}} \prod_{u_c \in C_S(\mathbf{u}_i)} \frac{\exp(\mathbf{u}_i^{\mathrm{T}} \mathbf{u}_c)}{\sum_{k=1}^{n} \exp(\mathbf{u}_i^{\mathrm{T}} \mathbf{u}_k)} - \gamma \prod_{\mathbf{a}_j \in S \wedge S \in D^{(a)}} \prod_{\mathbf{a}_c \in C_S(\mathbf{a}_j)} \frac{\exp(\mathbf{a}_j^{\mathrm{T}} \mathbf{a}_c)}{\sum_{k=1}^{m} \exp(\mathbf{a}_j^{\mathrm{T}} \mathbf{a}_k)} \tag{8}$$

where $\alpha$ , $\beta$ and $\gamma$ are the hyper-parameters to trade-off the effects of the three components. This objective function can be effectively solved by stochastic optimization methods. By solving the objective function (8), BEST represents users into a vector space where user's social relations have been embedded.

### *4.2 K-dimensional tree-based anomalous density group searching*

To fast search the anomalous density group, BEST first builds a k-dimensional tree (kd-tree for short) for the user representation space, and then estimates the density around each user in that space. Finally, it adopts the criteria Eq. (1) in NEST to identify the anomalous density groups.

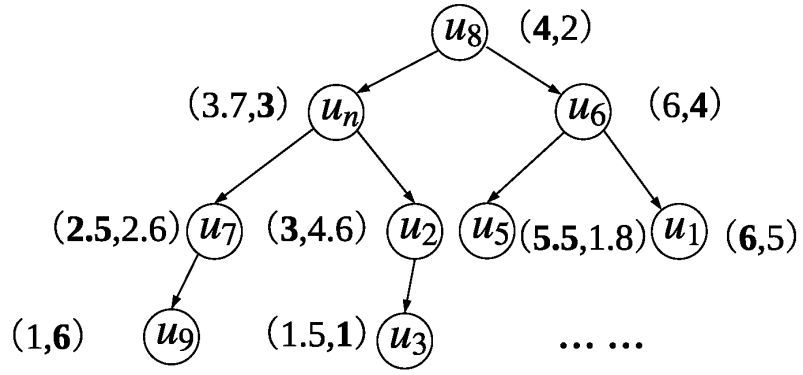



**Figure 2:** Example of kd-tree. The illustrated kd-tree is built on the user representation space shown in Fig. 1. Each level splits one dimension of the space into two parts

*4.2.1 Building kd-tree*

For user representation set $\mathbf{u} = \{\mathbf{u}_1, \mathbf{u}_2 \cdots, \mathbf{u}_n\}$, BEST builds a kd-tree, $v$, by Algorithm 2. As illustrated in Fig. 2, the kd-tree v is a binary tree storing the user representation with their structure information, which enables the fast searching of anomalous density groups.

**Algorithm 2:** Building *kd*-tree

**Procedure name:** BUILDKDTREE( $\mathbf{u}$ ,depth)
**Input:** A set of point u, and the current depth.
**Output:** The root of the *kd*-tree, $v$, storing $\mathbf{u}$

**if** $u$ *contains only one point* **then**
  **return** a leaf storing this point**.**
**else**
  $l \leftarrow$ depth%d+1;
  Split $\mathbf{u}$ into two subsets according to the median value $q$ in the $l$ th-dimension of the points in $\mathbf{u}$. Let $\mathbf{u}^{(1)}$ be the set of points which $l$ th-dimension value is smaller or equal to the $q$, and let $\mathbf{u}^{(2)}$ be the set of other points ;
  $v_{left} \leftarrow$ BUILDKDTREE( $\mathbf{u}^{(1)}$, depth+1) ;
  $v_{right} \leftarrow$ BUILDKDTREE( $\mathbf{u}^{(2)}$, depth+1) ;
  Create a node $v$ storing the $q$ in the $l$ th-dimension, make $v_{left}$ $v_{left}$ the left child of $v$, and make $v_{right}$ the right child of $v$ ;
  **return** v.
**end**

*4.2.2 Density estimation*

BEST estimates the density around each user in its representation space based on the kd-tree $v$ according to the Algorithm 3, where the function SEARCHKDTREE( $\mathbf{u}_i$ , $v$ , $\rho$ )

returns a set of users that around the user $\mathbf{u}_i$ within the range $\eta$ based on the kd-tree $v$. Essentially, BEST estimates the density around a user by the number of users close to the user within a certain distance in the representation space. If a user has a large density, the user should have a lot of collaborations with others. Accordingly, BEST uses the density as an important evidence to identify collaborative fraudsters.

**Algorithm 3:** Density estimation based on kd-tree

**Input :** A set of point $\mathbf{u}$, the kd-tree $v$ and $\eta$.

**Output:** A set of densities around each user $\rho$, a set of user sets $S$.

$\rho \leftarrow \{\}$
**foreach** $u_i$ in **u do**
    $S_i \leftarrow$ SEARCHKDTREE($u_i$, $v$, $\eta$);
    $\rho_i \leftarrow |S_i|$;
    $\rho \leftarrow \rho \cup \{\rho_i\}$;
    $S \leftarrow S \cup \{S_i\}$;
    **end**
**return** $\rho, S$.

*4.2.3 Collaborative fraudsters detection*

BEST detects collaborative fraudsters after estimating density around users in the user representation space. Specifically, it treats the density larger than a threshold $\varepsilon$, e.g. five times of the averaged density, as anomalous, and assigns the users in the density areas as fraudsters. The procedure is summarized in the Algorithm 4.

**Algorithm 4:** Collaborative fraudsters detection

**Input :** A set of densities around each user $\rho$, a set of user sets $S$, a threshold $\varepsilon$

**Output:** A set of fraudster users $F$ ..

$\rho \leftarrow \{\}$
**foreach** $S_i$ in $S$ **do**
    **if** $\rho_i > \varepsilon$ **then**;
        $F \leftarrow F \cup S_i$
**end**
**return** $F$

## 5 Experiments

### *5.1 Data sets*

The experiments are carried on two large scale real word social media data sets, including Yelp restaurant and Yelp hotel data sets used in Mukherjee et al. [Mukherjee, Venkataraman, Liu et al. (2013)]. All the activities in these data sets have been assigned authenticity labels given by commercial filters.

### *5.2 Evaluation metrics*

We evaluate their performance by three metrics - *precision*, *recall*, and *F-score*. While precision evaluates the fraction of true fraudsters among detected fraudsters, recall reflects the fraction of true fraudsters that have been detected over the total amount of true fraudsters. The precision and recall should be jointly considered since fraudsters detection is an imbalance problem [Luca and Zervas (2016)], i.e., fraudsters are much less than honest users. Thus, we use F-score, which balances the precision and recall, as an averaged indicator. Higher F-score indicates a better performance of a fraudsters detection method. We report these three metrics per ground-truth honest user and fraudster classes to illustrate the performance for different categories. We further average them to show overall performance.

We follow the literature [Wang, Liu and Zhao (2017)] to use the results of the Yelp commercial fraud filter to evaluate the performance. Because the Yelp commercial fraud filter only give the authenticity labels of activities, we transform the authenticity labels to the honest labels of users as the ground-truth. Considering the fraud activities distribution per each user assigned by the commercial filters, we assign the fraudster label to a user if more than 80% of the activities of the user have been labeled as fraud. The rationale is that we need to filter the false positive made by the commercial filters [Li, Chen, Liu et al. (2014)]. In other words, we assume that a user with a higher proportion of the assigned fraud activities will be more likely a real fraudster.

### *5.3 Parameters settings*

In the experiments, we set the parameters of BEST as follows. To balance the explicit and implicit social relations, we set the hyper-parameters $\alpha$ , $\beta$ , and $\gamma$ is the network embedding objective function Eq. (8) as 0.5, 0.25, and 0.25, respectively. We train the network embedding by Adam [Kingma and Ba (2014)] with embedding dimension 128 and batch size 32. For the density estimation, we set the distance range $\eta$ as 1. For the anomalous density detection, we set the threshold s as the five times of the averaged density. For the parameters in the compared methods, we take their recommended settings.

### *5.4 Evaluation of BEST effectiveness on fraudster detection*

#### *5.4.1 Experimental settings*

BEST is compared with two state-of-the-art competitors: Frauder [Hooi, Song, Beutel et al. (2016)] and HoloScope [Liu, Hooi and Faloutsos (2017)] in detecting collaborative fraudsters. These two competitors are both based on dense subgraph mining, but with different setting on the graph construction.

- *Fixed weighting dense subgraph mining-based method - FRAUDER* [Hooi, Song, Beutel et al. (2016)]. FRAUDER is a fraudsters detection method by dense subgraph mining. To detect camouflage and hijacked accounts, it adopts a fixed weighting strategy.
- *Dynamic weighting dense subgraph mining-based method-HoloScope* [Liu, Hooi and Faloutsos (2017)]. HoloScope uses information from graph topology and temporal

spikes to detect groups of fraudsters, and employs a dynamic weighting approach to allow a more accurately fraud detection.

*5.4.2 Findings-BEST significantly improving fraudsters detection performance, especially recall*

The precision, recall and F-score of BEST, Frauder, and HoloScope are reported in Tab. 1. Overall, BEST significantly outperforms the competitors. It improves 21.8% and 10.03% compared with the best-performing method in terms of F-score on two data sets.

**Table 1:** Collaborative fraudsters detection performance of different methods

| Data Info. | | | BEST | | | HoloScope | | | FRUADER | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Data | Category | #Review | Precision | Recall | F-Score | Precision | Recall | F-Score | Precision | Recall | F-Score |
| Hotel | Honest User | 420,785 | 0.68 | 0.96 | 0.80 | 0.64 | 0.6 | 0.62 | 0.64 | 0.98 | 0.77 |
| | Fraudster | 267,544 | 0.85 | 0.32 | 0.46 | 0.42 | 0.46 | 0.44 | 0.82 | 0.11 | 0.31 |
| | Overall | 888,329 | 0.75 | 0.71 | 0.67 | 0.55 | 0.55 | 0.55 | 0.71 | 0.65 | 0.55 |
| Restaurant | Honest User | 461,490 | 0.66 | 0.87 | 0.75 | 0.51 | 0.95 | 0.66 | 0.63 | 0.95 | 0.76 |
| | Fraudster | 326,981 | 0.74 | 0.35 | 0.48 | 0.74 | 0.12 | 0.21 | 0.74 | 0.21 | 0.33 |
| | Overall | 788,741 | 0.69 | 0.65 | 0.64 | 0.63 | 0.52 | 0.43 | 0.68 | 0.64 | 0.58 |

### *5.5 Evaluation of BEST-generated user representation quality*

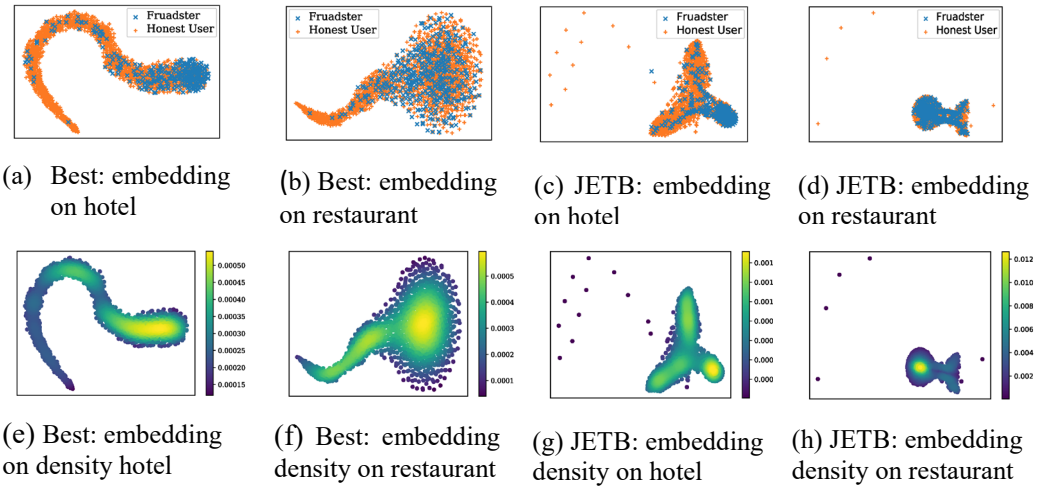
*5.5.1 Experimental settings*

We visualize the user representation in a two-dimensional space trough TSNE [Maaten and Hinton (2008)]. To evaluate the user representation quality, we plot the ground-truth labels of each user at their positions in the representation space. A high-quality user representation will enable a dense distribution for the collaborative fraudsters. The behavior representation generated by BEST is compared with that generated by JETB [Wang, Liu and Zhao (2017)], which is the state-of-the-art user representation method for fraudsters detection.

*5.5.2 Findings-BEST generated user representation embeds fraudsters into groups with anomalous high density*

The user representations generated by BEST and JETB are visualized in Fig. 3. In the

JETB generated representation space, the users with large density are not consistent to the ground-truth fraudster label. In contrast, the density of BEST generated representation is consistent with the ground-truth fraudsters distribution. This qualitative illustrates that BEST effectively captures the social relation of users in social media, which is essential for the collaborative fraudsters detection.




(a) Best: embedding on hotel (b) Best: embedding on restaurant (c) JETB: embedding on hotel (d) JETB: embedding on restaurant

(e) Best: embedding on density hotel (f) Best: embedding density on restaurant (g) JETB: embedding density on hotel (h) JETB: embedding density on restaurant

**Figure 3:** User representation with density of different methods on Yelp-hotel and Yelp-restaurant. The sub-figures (a), (b), (c), (d) contain the user representation information with the ground-truth labels, and the sub-figures (e), (f), (g), (h) show the density in the representation space

## 6 Conclusion

This paper introduces a network-embedding collaborative fraudsters detection framework NEST and its instance BEST. They perform an anomalous density searching procedure on a network embedding space which enables the detecting multiple groups of collaborative fraudsters. Two large real-world data sets demonstrate the performance of BEST is substantially better than the state-of-the-art competitors.

**Acknowledgements:** The work is supported by National Natural Science Foundation of China under Grant No. U1811462.

## References

**Akoglu, L.; Chandy, R.; Faloutsos, C.** (2013): Opinion fraud detection in online reviews by network effects. *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 13, pp. 2-11.

**Akoglu, L.; Tong, H.; Koutra, D.** (2015): Graph based anomaly detection and description: a survey. *Data Mining and Knowledge Discovery*, vol. 29, no. 3, pp. 626-688.

**Cao, S.; Lu, W.; Xu, Q.** (2015): GraRep: learning graph representations with global

structural information. *Proceedings of the ACM International on Conference on Information and Knowledge Management*, pp. 891-900.

**Cox, T. F.; Cox, M. A.** (2000): Multidimensional scaling. *Chapman and Hall/CRC*.

**Fei, G.; Mukherjee, A.; Liu, B.; Hsu, M.; Castellanos, M. et al.** (2013): Exploiting burstiness in reviews for review spammer detection. *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 13, pp. 175-184.

**Gao, M.; Chen, L.; He, X.; Zhou, A.** (2018): BiNE: bipartite network embedding. *Proceedings of the International ACM SIGIR Conference on Research & Development in Information Retrieval*, pp. 715-724.

**Grover, A.; Leskovec, J.** (2016): node2vec: scalable feature learning for networks. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 855-864.

**Hooi, B.; Shin, K.; Song, H. A.; Beutel, A.; Shah, N. et al.** (2017): Graph-based fraud detection in the face of camouflage. *ACM Transactions on Knowledge Discovery from Data*, vol. 11, no. 4, pp. 44:1-44:26.

**Hooi, B.; Song, H. A.; Beutel, A.; Shah, N.; Shin, K.; Faloutsos, C.** (2016): FRAUDAR: Bounding graph fraud in the face of camouflage. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 895-904.

**Hovy, D.** (2016): The enemy in your own camp: how well can we detect statistically-generated fake reviews-an adversarial study. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, vol. 2, pp. 351-356.

**Jindal, N.; Liu, B.** (2008): Opinion spam and analysis. *Proceedings of the ACM International WSDM Conference*, pp. 219-230.

**Junqué de Fortuny, E.; Stankova, M.; Moeyersoms, J.; Minnaert, B.; Provost, F. et al.** (2014): Corporate residence fraud detection. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1650-1659.

**Kingma, D. P.; Ba, J.** (2014): Adam: a method for stochastic optimization. *arXiv preprint arXiv:1412.6980.*

**Li, F.; Huang, M.; Yang, Y.; Zhu, X.** (2011): Learning to identify review spam. *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 2488-2493.

**Li, H.; Chen, Z.; Liu, B.; Wei, X.; Shao, J.** (2014): Spotting fake reviews via collective positive-unlabeled learning. *Proceedings of the IEEE International Conference on Data Mining*, pp. 899-904.

**Li, H.; Fei, G.; Wang, S.; Liu, B.; Shao, W. et al.** (2017): Bimodal distribution and co-bursting in review spam detection. *Proceedings of the International Conference on World Wide Web*, pp. 1063-1072.

**Lim, E. P.; Nguyen, V. A.; Jindal, N.; Liu, B.; Lauw, H. W.** (2010): Detecting product review spammers using rating behaviors. *Proceedings of the ACM International Conference on Information and Knowledge Management*, pp. 939-948.

**Liu, S.; Hooi, B.; Faloutsos, C.** (2017): Holoscope: topology-and-spike aware fraud detection. *Proceedings of the ACM International Conference on Information and Knowledge Management*, pp. 1539-1548.

**Luca, M.; Zervas, G.** (2016): Fake it till you make it: reputation, competition, and yelp review fraud. *Management Science*, vol. 62, no. 12, pp. 3412-3427.

**Maaten, L. v. d.; Hinton, G**. (2008): Visualizing data using t-SNE. *Journal of Machine Learning Research*, vol. 9, pp. 2579-2605.

**Mukherjee, A.; Kumar, A.; Liu, B.; Wang, J.; Hsu, M. et al.** (2013): Spotting opinion spammers using behavioral footprints. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 632-640.

**Mukherjee, A.; Liu, B.; Glance, N.** (2012): Spotting fake reviewer groups in consumer reviews. *Proceedings of the International Conference on World Wide Web*, pp. 191-200.

**Mukherjee, A.; Venkataraman, V.; Liu, B.; Glance, N. S.** (2013): What yelp fake review filter might be doing? *Proceedings of the International AAAI Conference on Web and Social Media*, pp. 409-418.

**Nedich, A.; Ozdaglar, A.** (2008): A geometric framework for nonconvex optimization duality using augmented lagrangian functions. *Journal of Global Optimization*, vol. 40, no. 4, pp. 545-573.

**Pandit, S.; Chau, D. H.; Wang, S.; Faloutsos, C.** (2007): Netprobe: a fast and scalable system for fraud detection in online auction networks. *Proceedings of the International Conference on World Wide Web*, pp. 201-210.

**Perozzi, B.; Al-Rfou, R.; Skiena, S.** (2014): Deepwalk: Online learning of social representations. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 701-710.

**Rayana, S.; Akoglu, L**. (2016): Collective opinion spam detection using active inference. *Proceedings of the IEEE International Conference on Data Mining*, pp. 630-638.

**Shehnepoor, S.; Salehi, M.; Farahbakhsh, R.; Crespi, N.** (2017): Netspam: a network-based spam detection framework for reviews in online social media. *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 7, pp. 1585-1595.

**Stringhini, G.; Kruegel, C.; Vigna, G.** (2010): Detecting spammers on social networks. *Proceedings of the Annual Computer Security Applications Conference*, pp. 1-9.

**Sun, J.; Qu, H.; Chakrabarti, D.; Faloutsos, C.** (2005): Neighborhood formation and anomaly detection in bipartite graphs. *Proceedings of the IEEE International Conference on Data Mining*, pp. 1-8.

**Tang, J.; Qu, M.; Wang, M.; Zhang, M.; Yan, J. et al.** (2015): Line: large-scale information network embedding. *Proceedings of the International Conference on World Wide Web*, pp. 1067-1077.

**Wang, D.; Cui, P.; Zhu, W.** (2016): Structural deep network embedding. *Proceedings of the 22nd ACM SIGKDD international Conference on Knowledge Discovery and Data Mining*, pp. 1225-1234.

**Wang, G.; Xie, S.; Liu, B.; Philip, S. Y.** (2011): Review graph based online store review spammer detection. *ICDM*, pp. 1242-1247.

**Wang, X.; Liu, K.; Zhao, J.** (2017): Handling cold-start problem in review spam detection by jointly embedding texts and behaviors. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, vol. 1, pp. 366-376.

**Wu, L.; Hu, X.; Morstatter, F.; Liu, H.** (2017): Adaptive spammer detection with sparse group modeling. *Proceedings of the International AAAI Conference on Web and Social Media*, pp. 319-326.

**Xiang, L.; Li, Y.; Hao, W.; Yang, P.; Shen, X.** (2018): Reversible natural language watermarking using synonym substitution and arithmetic coding. *Computers, Materials & Continua*, vol. 55, no. 3, pp. 541-559.

**Xiang, L.; Shen, X.; Qin, J.; Hao, W.** (2018): Discrete multi-graph hashing for large-scale visual search. *Neural Processing Letters*.

**Xiang, L.; Zhao, G.; Li, Q.; Hao, W.; Li, F.** (2018): TUMK-ELM: A fast unsupervised heterogeneous data learning approach. *IEEE Access*, vol. 6, pp. 35305-35315.

**Xu, C.; Zhang, J.; Chang, K.; Long, C.** (2013): Uncovering collusive spammers in Chinese review websites. *Proceedings of the ACM International on Conference on Information and Knowledge Management*, pp. 979-988.

**Xu, L.; Wei, X.; Cao, J.; Yu, P. S.** (2017): Embedding of embedding (EOE): Joint embedding for coupled heterogeneous networks. *Proceedings of the ACM International Conference on Web Search and Data Mining*, pp. 741-749.

**Ye, J.; Akoglu, L.** (2015): Discovering opinion spammer groups by network footprints. *Proceedings of the European Conference on Machine Learning*, pp. 267-282.

**You, Z.; Qian, T.; Liu, B.** (2018): An attribute enhanced domain adaptive model for cold- start spam review detection. *Proceedings of the International Conference on Computational Linguistics*, pp. 1884-1895.

**Zhao, Z.; Resnick, P.; Mei, Q.** (2015): Enquiring minds: early detection of rumors in social media from enquiry posts. *Proceedings of the International Conference on World Wide Web*, pp. 1395-1405.