

Anomaly, reciprocity, and community detection in networks

Hadiseh Safdari,* Martina Contisciani,† and Caterina De Bacco‡
Max Planck Institute for Intelligent Systems, Cyber Valley, Tuebingen 72076, Germany
(Dated: February 2, 2023)

Anomaly detection algorithms are a valuable tool in network science for identifying unusual patterns in a network. These algorithms have numerous practical applications, including detecting fraud, identifying network security threats, and uncovering significant interactions within a dataset. In this project, we propose a probabilistic generative approach that incorporates community membership and reciprocity as key factors driving regular behavior in a network, which can be used to identify potential anomalies that deviate from expected patterns. We model pairs of edges in a network with exact two-edge joint distributions. As a result, our approach captures the exact relationship between pairs of edges and provides a more comprehensive view of social networks. Additionally, our study highlights the role of reciprocity in network analysis and can inform the design of future models and algorithms. We also develop an efficient algorithmic implementation that takes advantage of the sparsity of the network.

I. INTRODUCTION

Anomaly detection algorithms are a crucial tool in the study of networks. These algorithms are designed to identify unusual or unexpected patterns in the data, which can provide valuable insights into the structure and function of a network [1, 2]. For instance, anomalous edges in a network may indicate the presence of a structural flaw or a potential problem, such as a vulnerability to attack. By detecting and analyzing these anomalies, we can gain a better understanding of the network and potentially identify ways to improve its performance or security [3]. In addition, anomaly detection algorithms can be used to monitor networks in real-time, allowing researchers to quickly identify and respond to potential issues as they arise.

Anomalies are often difficult to define precisely because they can vary depending on the context and the system being analyzed [4]. For example, in a network of online transactions, an anomaly could be a sudden spike in the number of transactions coming from a single user [5]. In this case, the regular behavior in the system would be the typical number of transactions coming from a single user, and any deviation from this pattern would be considered as anomaly. Hence, one of the main obstacles in detecting anomalies in networks is determining what is considered “normal” (or “regular”) behavior. To overcome this challenge, we must create a null model which is a realistic representation of the network data. This null model provides a standard against which we can compare the network data and identify anomalies.

Relevant approaches to address this problem include statistics-based methods, which fit a statistical model to the network data [6, 7]. Among these, generative models [8–10] make assumptions about the processes that drive network formation and evolution to generate synthetic

network data. By using these approaches, we can define null models that are tailored to the specific characteristics of the network under study. This is the approach we take here.

In this work, we focus on plain networks, which only contain information about the presence or absence of connections between individuals, and do not include any additional information. One approach to perform anomaly detection in these binary and single-layer networks is to use the structure of the graph to identify patterns and detect deviations from them [1]. These structural patterns can be divided into two categories: patterns based on the *overall structure* of the graph, and patterns based on the *community structure* of the graph. Methods in the first category rely on the global properties of the graph [11], such as the distribution of node degrees or the overall connectivity of the network. On the other hand, methods in the second category perform anomaly detection by focusing on the local properties of the graph, such as the membership of nodes in communities [12, 13]. Hence, with the second approach, we assume that the null model reflects a community structure that can be identified through latent variables, a process known as community detection task [14]. Thus, by considering the community structure, anomalous behavior can be determined in this context. For example, a friendship between two individuals from different groups, such as high school classmates and college classmates, could be considered anomalous. We recently developed a model (ACD) that performs anomaly detection by using community structure [15], where anomalous edges are those that deviate from regular patterns determined by community structure. As a result, this model outputs both node memberships and edge labels identifying them as legitimate or anomalous.

Accurately identifying anomalies is deeply connected with the chosen null model determining what regular patterns are. As a consequence, it is important to consider other possible mechanisms for tie formation, beyond community structure. For instance, reciprocity, another fundamental structural feature in networks [16–18], refers

* hadiseh.safdari@tuebingen.mpg.de

† martina.contisciani@tuebingen.mpg.de

‡ caterina.debacco@tuebingen.mpg.de

to the mutual exchange of resources or actions between individuals or groups. This can include actions such as returning a favor, sharing information or resources, or collaborating on a project. For example, in a social network, if two individuals consistently like and comment on each other's posts, this could be considered reciprocity. In a business network, if two companies frequently refer customers to each other, this could also be considered reciprocity. Mathematically, it is calculated as the ratio of the number of reciprocated edges to the total number of edges in the graph. Recent works [19, 20] have shown that including reciprocity effects in the modeling of community patterns results in more accurate and expressive generative models. This has the potential to improve the performance of an anomaly detection model for networks as well.

In this work, we develop a probabilistic generative model that we refer to as Community Reciprocity Anomaly Detection (CRAD) algorithm, that performs anomaly detection by proposing a null model based on both community structure and reciprocity. Intuitively, our model regards as regular ties those who follow the group membership and reciprocity effects, and as anomalous ties those whose formation process is not aligned with these two mechanisms. Notice that node memberships, reciprocity effect, and anomalous edges are all unknown processes. Our model is able to infer them from data by representing them as latent variables in a probabilistic model.

More specifically, we model the existence of ties between pairs of nodes using a bivariate Bernoulli distribution. This has the crucial statistical property that independence and uncorrelatedness of the component random variables are equivalent [21], which facilitates the derivation of a closed-form joint distribution of a pair of edges. Furthermore, both the marginal and conditional distributions are Bernoulli distributions, enabling closed-form analytical expressions. This facilitates downstream analysis and also improves model performance, as shown in [19].

II. THE MODEL

We are given an adjacency matrix, \mathbf{A} as our observed data, with entries indicating the presence or absence of an edge from node i to node j , represented by $A_{ij} = 1$ or $A_{ij} = 0$, respectively. Pairs of directed edges between

two nodes (i, j) are defined as $A_{(ij)} = (A_{ij}, A_{ji})$. We consider binary data, thus $A_{(ij)} \in \{0, 1\}^2 = \{0, 1\} \times \{0, 1\}$, and directed networks, i.e., in general $A_{ij} \neq A_{ji}$. We aim at classifying any such pair as either regular or anomalous, accounting for community structure and reciprocity effects. For this, we introduce a Bernoulli random variable that represents the binary label of being anomalous or not as a random variable:

$$\sigma_{(ij)} \sim \text{Bern}(\mu) , \quad (1)$$

where $\sigma_{(ij)} = 0, 1$ if the pair $A_{(ij)}$ is regular or anomalous, respectively. In this work we assume that edges between any pair of nodes must be either anomalous or regular. Mathematically, this means that the matrix σ with entries σ_{ij} is symmetric, i.e., $\sigma_{ij} = \sigma_{ji}$. These latent variables must be learned from data, as anomalies are not known in advance. They also determine the mechanism from which the pair of edges are drawn. The hyper-parameter $\mu \in [0, 1]$ controls the prior distribution of $\sigma_{(ij)}$.

With these main ingredients in mind, we can proceed to characterize the joint probability distribution of pairs of edges. Assuming to know the label $\sigma_{(ij)}$ for a given pair of edges, we denote the pair joint probability $p_{nm}^{(\ell)} = P^{(\ell)}(A_{ij} = n, A_{ji} = m)$, where $n, m \in \{0, 1\}$ and $\ell \in \{r, a\}$ denotes the label being regular or anomalous, respectively. We then consider the joint probability distribution of a pair of edges as a bivariate Bernoulli distribution:

$$P(A_{(ij)}, \sigma_{(ij)}) = P(A_{ij}, A_{ji}, \sigma_{(ij)}) = P(A_{ij}, A_{ji} | \sigma_{(ij)}) P(\sigma_{(ij)})$$

$$= P^{(a)}(A_{ij}, A_{ji} | \theta_a)^{\sigma_{(ij)}} P^{(r)}(A_{ij}, A_{ji} | \theta_r)^{1-\sigma_{(ij)}} P(\sigma_{(ij)} | \mu)$$

$$= \left[[p_{11}^{(a)}]^{A_{ij} A_{ji}} [p_{10}^{(a)}]^{A_{ij}(1-A_{ji})} [p_{01}^{(a)}]^{(1-A_{ij})A_{ji}} [p_{00}^{(a)}]^{(1-A_{ij})(1-A_{ji})} \right]^{\sigma_{(ij)}}$$

$$\times \left[[p_{11}^{(r)}]^{A_{ij} A_{ji}} [p_{10}^{(r)}]^{A_{ij}(1-A_{ji})} [p_{01}^{(r)}]^{(1-A_{ij})A_{ji}} [p_{00}^{(r)}]^{(1-A_{ij})(1-A_{ji})} \right]^{1-\sigma_{(ij)}}$$

$$\times \mu^{\sigma_{(ij)}} (1-\mu)^{1-\sigma_{(ij)}} , \quad (2)$$

where θ_r and θ_a denote parameters specific to the two distributions $P^{(r)}$ and $P^{(a)}$. The parameters $p_{nm}^{(\ell)}$ must satisfy $\sum_{n,m=0,1} p_{nm}^{(\ell)} = 1$ to have valid probability density functions.

Following the notation as in [19, 21], we can rewrite the full joint probability density function in Eq. (2), as the product,

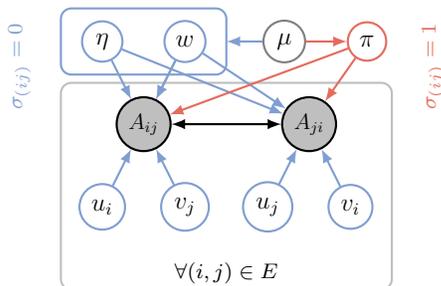
$$P(\mathbf{A}, \sigma) = \prod_{(i,j)} \left[\frac{\exp \left\{ A_{ij} f_{ij}^{(a)} + A_{ji} f_{ji}^{(a)} + A_{ij} A_{ji} J_{(ij)}^{(a)} \right\}}{Z_{(ij)}^{(a)}} \times \mu \right]^{\sigma_{(ij)}} \left[\frac{\exp \left\{ A_{ij} f_{ij}^{(r)} + A_{ji} f_{ji}^{(r)} + A_{ij} A_{ji} J_{(ij)}^{(r)} \right\}}{Z_{(ij)}^{(r)}} \times (1-\mu) \right]^{1-\sigma_{(ij)}} , \quad (3)$$

where $p_{00}^{(\ell)} = 1/Z_{(ij)}^{(\ell)}$, and $Z_{(ij)}^{(\ell)}$ is the normalization constant for the regular or anomalous edges, for $\ell \in \{r, a\}$; $f_{ij}^{(\ell)}$, $f_{ji}^{(\ell)}$, and $J_{(ij)}^{(\ell)}$ are the natural parameters of their density functions. The interaction term $J_{(ij)}^{(\ell)}$ appears in order to capture reciprocity. It allows to have a joint pair distribution $P(A_{ij}, A_{ji} | \sigma_{(ij)})$ that is not simply the product of two independent distributions $P(A_{ij} | \sigma_{(ij)}) \times P(A_{ji} | \sigma_{(ij)})$, as it is usually assumed in cases where reciprocity (or other properties involving more than one variable) is not taken into account explicitly. These parameters can be expressed in terms of the probability of occurrence of edges as follows:

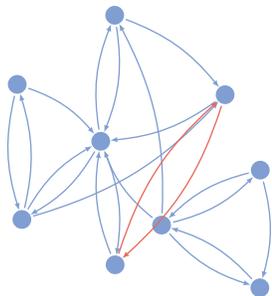
$$f_{ij}^{(\ell)} = \log \left(\frac{p_{10}^{(\ell)}}{p_{00}^{(\ell)}} \right), \quad f_{ji}^{(\ell)} = \log \left(\frac{p_{01}^{(\ell)}}{p_{00}^{(\ell)}} \right),$$

$$J_{(ij)}^{(\ell)} = \log \left(\frac{p_{11}^{(\ell)} p_{00}^{(\ell)}}{p_{10}^{(\ell)} p_{01}^{(\ell)}} \right), \quad \ell = \{r, a\}. \quad (4)$$

We aim at modeling reciprocity when two edges are regular, as this can be the result of a reasonable tie formation mechanism involving two nodes, e.g., exchanging



(a) Graphical model representation.



(b) Network example.

FIG. 1: Model visualization. (a) Graphical model: the entry of the adjacency matrix A_{ij} is determined by the community-related latent variables u, v, w and the reciprocity parameter η (blue); and by the anomaly-related parameters π (orange) and the hyper-prior μ (grey). (b) Example of a possible realization of the model: blue edges display interactions based on community and reciprocity and the orange ones are anomalous.

favors or cooperative behaviors. For anomalous edges instead, it is less clear what would reciprocity mean, hence we remain agnostic to it and assume that the edges $i \rightarrow j$ and $j \rightarrow i$ are independent when they are anomalous. In other words, the existence of the anomalous edge A_{ji} has no influence on its reciprocated edge A_{ij} , which is also anomalous. To reflect this mathematically, we set $J_{(ij)}^{(a)} = 0$. This follows the properties of multivariate Bernoulli distributions, where independence and uncorrelatedness are equivalent phenomena [21]. As the correlation between the pair of edges (A_{ij}, A_{ji}) is captured by $J_{(ij)}^{(\ell)}$, when $J_{(ij)}^{(\ell)} = 0$, the pair (A_{ij}, A_{ji}) is uncorrelated. In addition, we assume a symmetric structure of $f^{(a)} = f_{ij}^{(a)} = f_{ji}^{(a)}$ for all anomalous edges.

To summarize the steps of our proposed generative model: we first draw hidden labels for the edges, determining them being regular or anomalous; then, we draw pairs of edges (A_{ij}, A_{ji}) from a specific form of distribution depending on the edges' labels. Formally, the generative model is:

$$\sigma_{(ij)} \sim \text{Bern}(\mu) \quad (5)$$

$$A_{(ij)} \sim \begin{cases} \frac{\exp\{(A_{ij} + A_{ji})f^{(a)}\}}{Z_{(ij)}^{(a)}} & \text{if } \sigma_{(ij)} = 1 \\ \frac{\exp\{A_{ij}f_{ij}^{(r)} + A_{ji}f_{ji}^{(r)} + A_{ij}A_{ji}J_{(ij)}^{(r)}\}}{Z_{(ij)}^{(r)}} & \text{if } \sigma_{(ij)} = 0 \end{cases} \quad (6)$$

Up to this point, we focused on reciprocity and how to incorporate it into our model via the interaction term $J_{(ij)}^{(r)}$. Now, we turn our attention to community structure, another important mechanism that we believe regulates tie formation of regular edges. Conversely, we assume that communities have no influence on anomalous edges. To formalize this, we utilize similar model specifications as outlined in [19], and we incorporate community structure through latent variables embedded in the natural parameters of the joint pair distribution $P^{(r)}(A_{ij}, A_{ji})$. In detail, we assume the tie formation depends on communities and reciprocity for regular edges, and only on anomaly parameter for anomalous ties.

$$f_{ij}^{(r)} = \log \lambda_{ij}, \quad f_{ji}^{(r)} = \log \lambda_{ji}, \quad (7)$$

$$J_{(ij)}^{(r)} = \log \eta, \quad (8)$$

$$f^{(a)} = \log \pi, \quad (9)$$

where

$$\lambda_{ij} = \sum_{k,q=1}^K u_{ik} v_{jq} w_{kq}, \quad (10)$$

regulates how mixed-membership community structure determines tie formation in directed networks, as in [22]. We provide a schematic visualization of these contributions in Fig. 1. The normalization parameters are

obtained by enforcing the normalization constraint using the above definitions, so that $Z_{(ij)}^{(a)} = (\pi + 1)^2$ and $Z_{(ij)}^{(r)} = \lambda_{ij} + \lambda_{ji} + \eta\lambda_{ij}\lambda_{ji} + 1$.

The parameters λ and η play important roles in our model of community-reciprocity structure. λ captures the mixed-membership aspect, while η is the pair-interaction coefficient that regulates the formation of pairs of edges between nodes. The K -dimensional vectors u_i and v_i represent the out-going and in-coming communities of node i , respectively. The entries in these vectors, $u_{ik} \geq 0$ and $v_{jq} \geq 0$, represent the weights assigned to each community, where K is the number of communities. The value of K can be either specified as input or selected using model selection criteria, such as cross-validation [22]. The affinity matrix w_{kq} controls the structure of the communities, with higher values on the diagonal indicating more assortative communities. The formation of anomalous edges is derived by the latent parameter $\pi > 0$, as in the limit $\pi \rightarrow 0$ the probability of the existence of an anomalous edge converges to zero (see Appendix A for more details on derivations). All of these parameters, along with μ , are included in the latent parameter set $\Theta = \{\{u_i\}, \{v_i\}, \{w_{kq}\}, \eta, \pi, \mu\}$ that will be inferred from data. In addition to point estimates of these parameters, our model returns a posterior estimate for the edge label variable $\sigma_{(ij)}$ in the form of a Bernoulli posterior distribution of parameter $Q_{(ij)}$. This is also the estimated expected value of the edge label. We provide more details in Sec. III.

Our model assumes that community structure drives the process of formation of a regular edge, and that the regular edges between a pair of nodes depend on each other explicitly according to the value of η . If $J_{(ij)}^{(r)} = 0$ (when $\eta = 1$), the probability of the edges between nodes i and j is determined solely by their respective communities. On the other hand, a positive value of $J_{(ij)}^{(r)}$ (when $\eta > 1$) increases the probability of the existence of both $i \rightarrow j$ and $j \rightarrow i$, while a negative value (when $0 < \eta < 1$) decreases it.

By utilizing properties of the bivariate Bernoulli distribution [19, 21], we obtain a closed-form solution for the expected value of an edge (see for more details Appendix A):

$$\mathbb{E}[A_{ij}] = (1 - Q_{(ij)}) \frac{\lambda_{ij} + \eta\lambda_{ij}\lambda_{ji}}{Z_{(ij)}^{(r)}} + Q_{(ij)} \frac{\pi}{1 + \pi}. \quad (11)$$

This result is useful in link prediction experiments, in that we can score edges based on the values calculated from Eq. (11) and use these to compute prediction metrics such as the area under the receiver operating curve (AUC), we illustrate this in Sec. IV A.

Algorithm 1 CRAD : EM algorithm.

Input: network $\mathbf{A} = \{A_{ij}\}_{i,j=1}^N$,
number of communities K .

Output: memberships $u = [u_{ik}]$, $v = [v_{ik}]$; network affinity matrix $w = [w_{kq}]$; pair-interaction coefficient η ; anomaly parameter π ; prior on anomaly indicator μ .

Initialize $\Theta : (u, v, w, \eta, \pi, \mu)$ at random.

Repeat until $L(\Theta)$ convergence:

1. Calculate ρ and Q (E-step):

$$\rho_{ijkq} \sim \text{as in Eq. (B13)},$$

$$Q_{ij} \sim \text{as in Eq. (B22)}.$$

2. Update parameters Θ (M-step):

i) for each node i and community k update memberships:

$$u_{ik} = \frac{\sum_j Q_{(ij)} (1 - Q_{(ij)}) A_{ij} \rho_{ijkq}}{\sum_j \left[\frac{\sum_q (1 - Q_{(ij)}) (1 + \eta \lambda_{ij}) v_{jq} w_{kq}}{\lambda_{ij} + \lambda_{ji} + \eta \lambda_{ij} \lambda_{ji} + 1} \right]}$$

$$v_{ik} = \frac{\sum_j Q_{(ij)} (1 - Q_{(ij)}) A_{ij} \rho_{jikq}}{\sum_j \left[\frac{\sum_q (1 - Q_{(ij)}) (1 + \eta \lambda_{ij}) u_{jq} w_{kq}}{\lambda_{ij} + \lambda_{ji} + \eta \lambda_{ij} \lambda_{ji} + 1} \right]}$$

ii) for each pair (k, q) update affinity matrix:

$$w_{kq} = \frac{\sum_{i,j} (1 - Q_{(ij)}) A_{ij} \rho_{ijkq}}{\sum_{i,j} \left[\frac{(1 - Q_{(ij)}) (1 + \eta \lambda_{ij}) u_{ik} v_{jq}}{\lambda_{ij} + \lambda_{ji} + \eta \lambda_{ij} \lambda_{ji} + 1} \right]}$$

iii) update pair-interaction coefficient:

$$\eta = \frac{\sum_{(i,j)} (1 - Q_{(ij)}) A_{ij} A_{ji}}{\sum_{(i,j)} (1 - Q_{(ij)}) \left[\frac{\lambda_{ij} \lambda_{ji}}{\lambda_{ij} + \lambda_{ji} + \eta \lambda_{ij} \lambda_{ji} + 1} \right]}$$

iv) update anomaly parameter:

$$\pi = \frac{\sum_{(i,j)} Q_{(ij)} (A_{ij} + A_{ji})}{\sum_{(i,j)} Q_{(ij)} (2 - A_{ij} - A_{ji})}$$

v) update prior on anomaly indicator:

$$\mu = \frac{1}{N(N-1)/2} \sum_{(i,j)} Q_{(ij)}.$$

III. INFERENCE

Our ultimate goal is to determine Θ , the latent parameters of the model. To do this, we maximize the posterior probability $P(\Theta|\mathbf{A}) = \sum_{\sigma} P(\sigma, \Theta|\mathbf{A})$. Instead of directly maximizing this probability, it is more computationally efficient to maximize the log-posterior, as the maxima of the two functions are equivalent:

$$\begin{aligned}
L(\Theta) &= \log P(\Theta|\mathbf{A}) = \log \sum_{\sigma} P(\sigma, \Theta|\mathbf{A}) \\
&\geq \sum_{\sigma} q(\sigma) \log \frac{P(\sigma, \Theta|\mathbf{A})}{q(\sigma)}, \quad (12)
\end{aligned}$$

where we defined $q(\sigma)$, a variational distribution that must sum to 1. Our maximum likelihood approach involves the use of an expectation-maximization (EM) algorithm in which we alternately update different sets of parameters of our model. More specifically, we first update the variational distribution parameters (E-step), ρ and \mathbf{Q} , and then maximize $L(\Theta)$ with respect to Θ (M-step). This process is repeated until $L(\Theta)$ converges, signifying the completion of the optimization process. The full procedure is outlined in Algorithm 1 (see Appendix B for more details on the inference task). The computational complexity of the algorithm is $O(N^2)$, primarily due to the terms in the dense matrix $Q_{(ij)}$ that are not multiplied by the sparse adjacency matrix A_{ij} . As \mathbf{Q} is crucial for identifying anomalous edges, its presence may make the model infeasible for large systems. Investigating ways to reduce this complexity, for instance by making its representation sparse, is an interesting avenue for future work.

IV. RESULTS

A. Synthetic datasets

We validate our model on synthetic datasets, generated with the generative algorithm in appendix C. The studied datasets consist of $N = 500$ nodes, with an average degree of $\langle k \rangle = 60$. The number of communities is set to $K = 3$, and the pair-interaction coefficient, η , has a range of values. The anomaly density (ratio of anomalous edges to total number of edges) is varied within the interval $\rho_a \in [0, 1]$. We compare CRAD with JointCRep [19], which is what CRAD reduces to if we had not considered anomalies, i.e., when $\mu = 0$ and $\lim \pi \rightarrow 0$. This allows to focus on observing the impact of considering the existence of anomalous edges in a given dataset.

In order to determine the effectiveness of our proposed model, which is based on the concept of community structure, we first evaluate its ability to accurately identify the memberships of individuals within a community. To accomplish this, we measure the cosine similarity (CS) between the ground truth and inferred community memberships vectors. The CS has values in $[0, 1]$, with CS = 1 indicating the best performance. For this task, we also run a Bayesian Poisson matrix factorization (BPMF) algorithm [23]. BPMF is a scalable algorithm for factorizing sparse matrices and provides a useful comparison for our proposed algorithm. We run all algorithms on synthetic datasets generated by CRAD (see Appendix C for more details). The results, as illustrated in Fig. 6 (a)

and (b), show that when the proportion of anomalous edges in the dataset is relatively low, BPMF outperforms our proposed algorithm. However, when the number of anomalous edges is above 50% of the total number of edges, our algorithm is still able to detect community structure with a reasonable level of accuracy. Additionally, it can be observed that CRAD performs the same as JointCRep; with both models having higher performance for smaller values of the anomaly density, ρ_a . This behavior is expected, as for higher values of ρ_a , the community structure plays a weaker role in the formation of edges.

It is worth mentioning that the primary objective of the current research is to develop the capabilities of JointCRep through the incorporation of anomaly detection functionality, rather than focusing on further improving its community detection abilities or recovering reciprocity parameter. Therefore, our focus is on assessing and optimizing the model’s anomaly detection potential. For this, we measure the AUC on edges, i.e., on a binary matrix that stores what edges are true anomalies, and use as scores the inferred $Q_{(ij)}$. From our results, illustrated in panels (e) and (f) in Fig. 6, we find that CRAD demonstrates good performance in the detection of anomalous edges across a range of anomaly densities. Furthermore, the integration of reciprocity effects is enhancing performance, compared to a model (ACD) where there is no such effect [15].

In addition to evaluating anomaly detection, we are also interested in assessing the ability of CRAD to identify missing edges, also known as link prediction task. In these experiments we employ a 5-fold cross-validation approach, where the dataset is split in five sets of data. In each realization, four of these groups are utilized as a training set to infer the parameters Θ . The remaining group is used as a test set, where the score for each pair (A_{ij}, A_{ji}) in the matrix is evaluated to compute the AUC. By iteratively varying which group serves as the test set, we obtain a total of five trials per realization. The final AUC value is determined by averaging the results of these trials. The score of an edge is calculated using the closed-form expression for its marginal probability, as described in Eq. (11). As shown in panels (c) and (d) in Fig. 6, an increase in the reciprocity parameter results in an increase in the AUC for both CRAD and JointCRep, however, we observe a bigger improvement in terms of AUC of CRAD over the competitive algorithms. These results indicate that our model becomes more effective in link prediction tasks for higher values of reciprocity.

B. Real World datasets

In order to assess the practical utility of our model, we investigate its usage on a variety of real-world data covering applications as food-sharing between bats, social support interactions in a rural community, email exchanges, and online dating. Their sizes range from $N = 19$ to $N = 3562$, see Table I in Appendix E for a summary

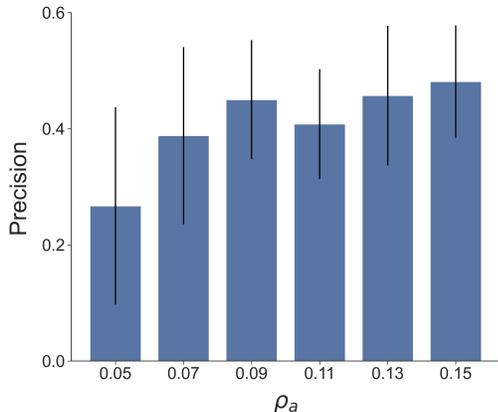


FIG. 2: **Precision in detecting the injected edges in the vampire-bat network.** The precision increases by the increase in the number of anomalous edges injected in the network, i.e., anomaly density in the dataset, ρ_a . The results is the average over 10 randomly injected sets of edges, bars are standard deviations. Here we use the initialization $\pi = 0.1$.

description.

Injecting anomalous edges To evaluate the accuracy and precision of the model in detecting anomalous edges, we first need to know the true label of edges, being anomalous or regular. However, one of the challenges in this regard is the lack of data containing explicit anomalies. To address this challenge, we conduct an experiment where we inject n random edges between nodes in a real dataset and label them as anomalous. We vary n to evaluate the impact of anomaly density $\rho_a = n/E$ on model performance. We then run our model on this manipulated dataset and infer the expected value $\mathbb{E}[\sigma_{(ij)}] = \hat{Q}_{(ij)} \in [0, 1]$ for the edge labels, which also indicates the likelihood that the edges between two nodes are anomalous. Based on this, we assign labels to the edges. In this specific experiment, we label the first n pairs (i, j) with the highest values of $\hat{Q}_{(ij)}$ as anomalous edges. We measure the precision as performance metric, this is the fraction of inferred anomalous edges which are correctly classified (in our case—since we fix the number of inferred anomalous to be equal to the number of injected anomalous edges—this also corresponds to recall, i.e., the fraction of true anomalous edges that are inferred as such).

1. Smaller datasets

a. Vampire bat network The vampire bat network is a complex and dynamic social structure in which individual vampire bats form connections and share food with one another [24]. The bats have a remarkable ability to detect the body heat of other bats, even in complete darkness, allowing them to locate potential food sources

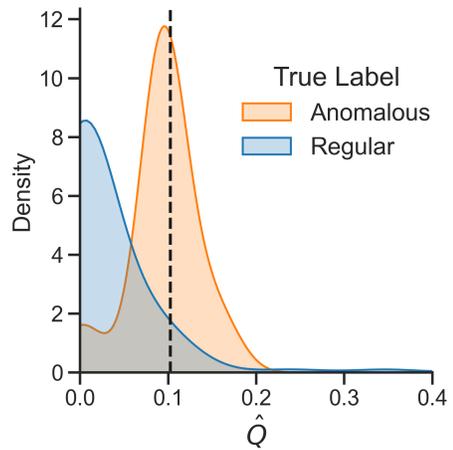


FIG. 3: **Anomaly Detection in the vampire-bat network.** We show the distribution of $\hat{Q}_{(ij)}$, i.e., the probability of a pair of edges (i, j) being anomalous, as estimated by CRAD. We distinguish the true regular and anomalous edges with different colours, blue and orange, respectively, to highlight their different inferred distributions. Here, $\rho_a = 0.09$ and $\pi = 0.1$. We measure a precision of 0.4. For this, we label as anomalous the fraction of ρ_a edges with highest $\hat{Q}_{(ij)}$. The vertical dashed line denotes the minimum $\hat{Q}_{(ij)}$ observed in this set of anomalous edges.

and potential recipients for food sharing. When a bat finds food, it will often regurgitate some of it and share it with other members of its network. This behavior, known as reciprocal altruism, is essential for the survival of the group, as it ensures that all members have access to food even when they are unable to find it themselves. The decision of who to feed is likely to be influenced by both the genetic relatedness of the individuals involved and their history of reciprocal sharing. Given this, we expect that reciprocity will play a significant role in determining which individuals form close social ties within this network. As such, when examining this dataset, it will be important to carefully consider this effect. In our analysis, we use the data obtained from [24] and remove isolated nodes. The network consists of $N = 19$ nodes, $E = 103$ edges and has high reciprocity of 0.64. In addition, we fix $K = 2$ as in [19].

As shown in Fig. 2, our model’s ability to detect anomalies improves when there is a higher concentration of anomalies in the dataset. The plot depicts the precision in detecting the anomalous edges, for a range of anomaly density, ρ_a . In a more specific case, Fig. 3 provides an example of how CRAD can be used for anomaly detection in the vampire-bat dataset. In this example, a set of edges with $\rho_a = 0.09$ were embedded in the system. In this figure, the entries of the estimated \hat{Q} matrix, which represent the probability of edges being anomalous, are categorized based on their true labels and assigned different colours to highlight their different inferred distributions of \hat{Q} . The plot clearly shows two

different distributions, one highly peaked at $\hat{Q}_{(ij)} = 0$ and the other peaked around $\hat{Q}_{(ij)} = 0.1$. The first corresponds to regular edges, which are thus correctly identified as such, while the latter are the injected anomalies, which are indeed assigned a higher probability of being anomalous. While there are few regular edges that have a high \hat{Q} , we observe that a significant density of anomalous edges is concentrated at $\hat{Q}_{(ij)} > 0.1$, indicating that the model is correctly assigning them as anomalous. Quantitatively, we measure precision and recall values of 0.4, obtained by labelling as anomalous the fraction of $\rho_a = 0.09$ edges with highest $\hat{Q}_{(ij)}$. Even though a small fraction of regular edges are classified as anomalous and vice versa, these numbers show that overall the algorithm is doing well at detecting the injected anomalies.

b. A Nicaraguan community The next dataset represents the social support network of indigenous Nicaraguan horticulturalists [25]. The original dataset is self-reported network data. Ties are reported by several individuals and these may be in disagreement with each other. Hence, we process it using VIMuRe algorithm [17], which estimates probabilistically an underlying network structure from self-reported network data, provided by multiple reporters, accounting for reciprocity. The summary description of the estimated network by VIMuRe can be seen in Table I. In addition, it estimates the reliability $\theta > 0$ of each individual reporter, with higher values denoting over-reporting. Reliabilities can be correlated to anomalies in that we expect that unreliable reporters may report non-existence ties which we interpret as anomaly.

To assess this, we run VIMuRe twice. The first time, we run its default version and use it to collect estimates of reporters' reliabilities. The second time, we run it in a modified version where we fix the reliability parameters to a neutral value, assuming that all reporters are reliable. We use this output, the estimated network in this modified version, as input for CRAD. In this way, we aim at observing proxies for anomalous edges: these are some of the edges that are involving unreliable reporters, as estimated in the first run of VIMuRe. Our model labels anomalies on edges, instead in this dataset we have information on nodes (their reliabilities). We can build a correspondence between these two types of information by assuming that edges connected to the most unreliable reporters would have highest value in the estimated \hat{Q} matrix. To quantify this match, we assign a value $\hat{Q}_i = \max_{j \in \partial i} \hat{Q}_{(ij)}$ to each reporter i , where ∂j is its neighbourhood, being the maximum probability that one of its connecting ties is anomalous.

We expect \hat{Q}_i to be high for nodes that have a high unreliability θ . We find indeed a positive correlation of 0.46 between θ_i and \hat{Q}_i , as shown in Fig. 4. In particular, we observe that the edge (76,3) between the two most unreliable nodes has the maximum observed value of $\hat{Q}_{(ij)} = 0.36$, which is consistent with the findings reported in [17]. Notice that we expect this correlation

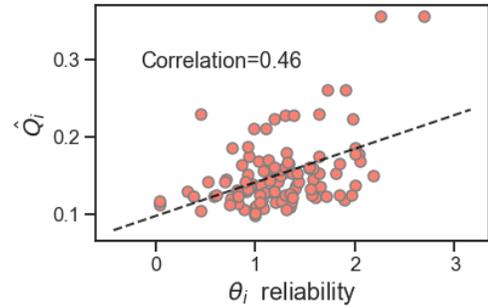


FIG. 4: **Anomaly Detection in a Nicaraguan social support network.** We show a scatterplot of \hat{Q}_i (the maximum probability that one of the connecting ties of node i is anomalous), as estimated by CRAD, against θ_i , reporters' reliabilities, as estimated by VIMuRe algorithm. The correlation is calculated as the Pearson coefficient, the dashed line is a linear fit to the data. Positive correlation signals that nodes that are more unreliable (high θ_i) tend to have an edge that is more likely to be labeled as anomalous among its connections.

to further increase if we were able to account explicitly for anomalies on nodes (instead of on edges). In this case, one could envision adapting our formalism to assign random variables σ_i to nodes, which may result in less tractable distributions and thus higher complexity, but may be more appropriate for applications in which nodes act consistently as anomalous. We leave this as an open question for future work.

2. Larger datasets

In this section, we test our algorithm on UC Irvine and POK messages, as examples of larger datasets. In each case, we randomly select and add 10% additional edges, labeled as anomalous. The CRAD algorithm consistently produces reliable results in detecting anomalies in both datasets.

a. UC Irvine messages The network of UC Irvine messages is composed of messages sent between users of an online community of students from the University of California, Irvine [26]. Each node in this communication network represents a user and each directed edge represents a message that was sent from one user to another. Our model consistently identified anomalies in this dataset with a high level of accuracy, as evidenced by a particularly high peak in the distribution of $\hat{Q}_{(ij)}$ corresponding to anomalous edges in Fig. 5 (a). This is also quantified with a precision value of 0.63 in the confusion matrix shown in Fig. 5 (b).

b. Network of online dating The POK dataset is a large data set containing the messages exchanged by users within the online dating POK community. The results depicted in Fig. 5 (c-d) demonstrate the strong

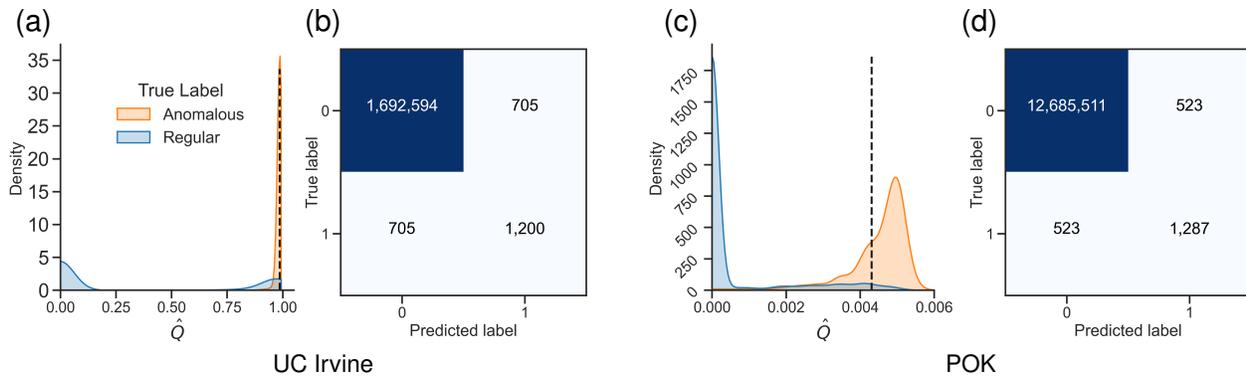


FIG. 5: **Anomaly Detection in the UC Irvine and POK networks.** We show the distribution of $\hat{Q}_{(ij)}$, i.e., the probability of a pair of edges (i, j) being anomalous ((a),(c)) and the confusion matrix ((b),(d)), as estimated by CRAD, for the UC Irvine (left) and POK (right) datasets. We distinguish the true regular and anomalous edges with different colours, blue and orange, respectively, to highlight their different inferred distributions. Here, $\rho_a = 0.1$ and $\pi = 0.3$. We measure a precision of 0.63 for UC Irvine and of 0.71 for POK network. The vertical dashed line denotes the minimum $\hat{Q}_{(ij)}$ observed in this set of anomalous edges.

performance in identifying and reconstructing anomalous edges. Figure 5 (c) illustrates how, also in this case, the distribution of \hat{Q} values for the anomalous edges is peaked around higher values. While this distribution is more distributed (i.e., has higher variance) than the analogous one observed for UC Irvine, here we observe the distribution corresponding to regular edges being more peaked around zero. This means that in this case the model is distinguishing more clearly the regular edges, with the consequence of obtaining a higher precision of 0.71, as shown in the confusion matrix of Fig. 5 (d). Taken together, these results support the efficacy of our classification methodology.

V. CONCLUSION

We introduce an expressive generative model to detect edge anomalies in networks that takes into account community membership and reciprocity as main mechanisms driving tie formation. By leveraging these two effects, it is able to detect what edges deviate from a regular behavior and estimate their probability of being anomalous. This inference is performed in a joint learning of edge anomalies and mixed-memberships of nodes in communities, thus allowing practitioners to flag potential irregular edges while providing an interpretable community structure.

In contrast to common models for anomaly detection that rely on metadata on edges or nodes, our model takes as input only the adjacency matrix and estimates anomaly labels on the edges. It is an unsupervised model, meaning it does not require any input label to train it. These features make it particularly relevant in cases where extra information is not available—which is the

case for many networked datasets—where the applicability of many machine learning methods for anomaly detection is significantly limited. As an example, traditional models for anomaly detection in financial transactions often rely on metadata such as transaction amount, location, and merchant information [5, 27, 28]. Instead, our model only requires the adjacency matrix of the transactions, which represents the connections between different account holders.

One key feature of our model is that it provides a joint probability for the existing pairs of edges between any pairs of nodes, allowing for the inclusion of reciprocity in the model, a relevant property in many directed networks. Furthermore, our model allows for mixed community membership, meaning that nodes can belong to more than one community. This is a more realistic representation of data structures compared to models that assume a single community membership for each node.

There are numbers of ways that our model could be further improved. As mentioned above, our model takes little information in input, only the network’s adjacency matrix. A natural next step would be to extend the current model to account for extra information as node attributes, using ideas from generative models with both communities and attributes [29, 30]; or to consider techniques from semi-supervised learning [31], in case of availability of labels on a subset of the edges. Furthermore, we can envision that, for rich and large datasets, deep learning architectures for anomaly detection [32–34] may be competitive methods. However, one could imagine extending standard architectures by combining them with the main ingredients of our model, in datasets where communities and reciprocity matter. The robust performance in detecting anomalies in real data with no extra information suggests that by combin-

ing these insights with complex deep architectures may make the latter more expressive and thus boost predictive power.

Another type of extra information that is present in many real datasets is time [35]. Edges can be timestamped and this could be used to improve estimates of anomalies. Hence, future work could be directed at generalizing our model to dynamical networks, for instance by combining insights from generative models for dynamic networks with communities [36–38].

It is important to note that the inferred labels for edges in our model should be treated as estimates rather than definitive conclusions. These labels should be used with caution in the study of a network, as further investigation may be necessary to fully understand the nature of anomalous edges. However, our model can provide valuable insights for practitioners to better understand and interpret the networks they are studying, especially when combined with their specialized knowledge and understanding of the data at hand.

ACKNOWLEDGMENTS

All the authors were supported by the Cyber Valley Research Fund. The authors thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting Martina Contisciani.

APPENDIX A: DETAILED DERIVATIONS

Anomalous edges: As in the formation of anomalous edges, the reciprocated edges are independent, we apply the condition $J_{(ij)}^{(a)} = 0$, therefore from Eq.(4), we find

$$\frac{p_{11}^{(a)} p_{00}^{(a)}}{p_{10}^{(a)} p_{01}^{(a)}} = 1 \quad \Rightarrow \quad p_{11}^{(a)} = \frac{p_{10}^{(a)} p_{01}^{(a)}}{p_{00}^{(a)}} . \quad (\text{A1})$$

Moreover, $f_{(ij)}^{(a)} = f_{(ji)}^{(a)} = f^{(a)} \implies p_{10}^{(a)} = p_{01}^{(a)} = p^{(a)}$ and

$$f^{(a)} = \log \pi = \log \frac{p^{(a)}}{p_{00}^{(a)}} \quad \Rightarrow \quad p^{(a)} = \pi p_{00}^{(a)} . \quad (\text{A2})$$

Using the normalization condition, $p_{00}^{(a)} + p_{10}^{(a)} + p_{01}^{(a)} + p_{11}^{(a)} = 1$, and the results of Eqs. (A1)–(A2), we find the explicit mapping between the latent variables and the instances of $P^{(a)}(A_{ij}, A_{ji} | \theta_a)$ in Eq. (2),

$$p_{00}^{(a)} = \frac{1}{Z_{ij}^{(a)}}, \quad p_{10}^{(a)} = p_{01}^{(a)} = \frac{\pi}{Z_{ij}^{(a)}}, \quad p_{11}^{(a)} = \frac{\pi^2}{Z_{ij}^{(a)}} , \quad (\text{A3})$$

where the normalization constant is:

$$Z_{ij}^{(a)} = (1 + \pi)^2 . \quad (\text{A4})$$

Regular edges: In order to find the explicit mapping between the latent variables and the instances of $P^{(r)}(A_{ij}, A_{ji} | \theta_r)$ in Eq. (2), we follow the same procedure as in [19],

$$p_{01}^{(r)} = \frac{\lambda_{ji}}{Z_{(ij)}^{(r)}} \quad (\text{A5})$$

$$p_{10}^{(r)} = \frac{\lambda_{ij}}{Z_{(ij)}^{(r)}} \quad (\text{A6})$$

$$p_{11}^{(r)} = \frac{\eta \lambda_{ij} \lambda_{ji}}{Z_{(ij)}^{(r)}} \quad (\text{A7})$$

$$p_{00}^{(r)} = \frac{1}{Z_{(ij)}^{(r)}} , \quad (\text{A8})$$

where the normalization constant is:

$$Z_{(ij)}^{(r)} = \lambda_{ij} + \lambda_{ji} + \eta \lambda_{ij} \lambda_{ji} + 1 . \quad (\text{A9})$$

Having these mappings, we can construct the marginal and conditional distributions of the ties. Thus, the marginal and conditional distributions of A_{ij} have the following densities, respectively:

$$P(A_{ij}) = \left[[p_{10}^{(r)}]^{A_{ij}} [p_{00}^{(r)}]^{(1-A_{ij})} + [p_{11}^{(r)}]^{A_{ij}} [p_{01}^{(r)}]^{(1-A_{ij})} \right] \times (1 - \mu) + \left[[p_{10}^{(a)}]^{A_{ij}} [p_{00}^{(a)}]^{(1-A_{ij})} + [p_{11}^{(a)}]^{A_{ij}} [p_{01}^{(a)}]^{(1-A_{ij})} \right] \times \mu , \quad (\text{A10})$$

$$\begin{aligned}
P(A_{ij}|A_{ji}) &= \frac{[p_{11}^{(r)}]^{A_{ij} A_{ji}} [p_{10}^{(r)}]^{A_{ij} (1-A_{ji})} [p_{01}^{(r)}]^{(1-A_{ij}) A_{ji}} [p_{00}^{(r)}]^{(1-A_{ij}) (1-A_{ji})}}{P(A_{ji})} \times (1 - \mu) \\
&+ \frac{[p_{11}^{(a)}]^{A_{ij} A_{ji}} [p_{10}^{(a)}]^{A_{ij} (1-A_{ji})} [p_{01}^{(a)}]^{(1-A_{ij}) A_{ji}} [p_{00}^{(a)}]^{(1-A_{ij}) (1-A_{ji})}}{P(A_{ji})} \times \mu
\end{aligned} \tag{A11}$$

APPENDIX B: INFERENCE

Our goal is, given two mechanisms responsible for edge formation, first to determine the values of the parameters $\Theta = \{\{u_{ik}\}, \{v_{ik}\}, \{w_{kq}\}, \eta, \pi, \mu\}$, which determine the relationship between the anomaly indicator $\sigma_{(ij)}$ and the data, and then, given those values, to estimate the indicator $\sigma_{(ij)}$ itself.

We have the posterior:

$$P(\sigma, \Theta | \mathbf{A}) = \frac{P(\mathbf{A} | \sigma, \Theta) P(\sigma | \mu) P(\Theta) P(\mu)}{P(\mathbf{A})}. \tag{B1}$$

Summing over all the possible indicators we have:

$$P(\Theta | \mathbf{A}) = \sum_{\sigma} P(\sigma, \Theta | \mathbf{A}), \tag{B2}$$

which is the quantity that we need to maximize to extract the optimal Θ . It is more convenient to maximize its logarithm, log-posterior, as the two maxima coincide. We use the Jensen's inequality:

$$L(\Theta) = \log P(\Theta | \mathbf{A}) = \log \sum_{\sigma} P(\sigma, \Theta | \mathbf{A}) \geq \sum_{\sigma} q(\sigma) \log \frac{P(\sigma, \Theta | \mathbf{A})}{q(\sigma)}, \tag{B3}$$

where $q(\sigma)$ is a variational distribution that must sum to 1. In fact, the exact equality happens when,

$$q(\sigma) = \frac{P(\sigma, \Theta | \mathbf{A})}{\sum_{\sigma} P(\sigma, \Theta | \mathbf{A})}, \tag{B4}$$

this definition is also equivalent to maximizing the right-hand-side of Eq. (B3) w.r.t. q .

Finally, we need to maximize the log-posterior with respect to Θ to get the latent variables. This can be done in an iterative way using Expectation-Maximization algorithm (EM), alternating between maximizing w.r.t. q using Eq. (B4) and then maximizing Eq. (B23) w.r.t. Θ . In this work, we only fix priors for the σ_{ij} (Bernoulli distributions with parameter μ). For this variable we can thus estimate full posterior distributions; instead for the other parameters our model outputs point estimates. This could be modified by suitably specifying priors also for the reciprocity or community-related parameters. In this case, one could easily obtain maximum a posteriori (MAP) estimates with calculations similar to those reported here.

Defining $Q_{(ij)} = \sum_{\sigma_{(ij)}} q(\sigma_{(ij)}) \sigma_{(ij)}$, the expected value of $\sigma_{(ij)}$ over the variational distribution, we obtain,

$$\begin{aligned}
L(\Theta) &= - \sum_{\sigma} [q(\sigma) \log q(\sigma)] + \sum_{(i,j)} \left\{ (1 - Q_{(ij)}) \left(A_{ij} f_{ij}^{(r)} + A_{ji} f_{ji}^{(r)} + A_{ij} A_{ji} J_{(ij)}^{(r)} - \log Z_{(ij)}^{(r)} \right) \right. \\
&\quad \left. + Q_{(ij)} \left((A_{ij} + A_{ji}) f^{(a)} - \log Z_{(ij)}^{(a)} \right) + Q_{(ij)} \log \mu + (1 - Q_{(ij)}) \log(1 - \mu) \right\}, \tag{B5}
\end{aligned}$$

and having Eqs. (7-10),

$$\begin{aligned}
L(\Theta) = & - \sum_{\sigma} [q(\sigma) \log q(\sigma)] + \\
& + \sum_{(i,j)} \left\{ (1 - Q_{(ij)}) \left(A_{ij} \log \sum_k u_{ik} v_{jq} w_{kq} + A_{ji} \log \sum_k u_{jk} v_{iq} w_{kq} + A_{ij} A_{ji} \log \eta \right. \right. \\
& \left. \left. - \log \left[\sum_{k,q} u_{ik} v_{jq} w_{kq} + \sum_{k,q} u_{jk} v_{iq} w_{kq} + \eta \sum_{k,q} u_{ik} v_{jq} w_{kq} \sum_{k,q} u_{jk} v_{iq} w_{kq} + 1 \right] \right) \right. \\
& \left. + Q_{(ij)} \left((A_{ij} + A_{ji}) \log \pi - 2 \log(\pi + 1) \right) + Q_{(ij)} \log \mu + (1 - Q_{(ij)}) \log(1 - \mu) \right\}. \tag{B6}
\end{aligned}$$

Derivative of log-posterior w.r.t η ,

$$\frac{\partial L(\Theta)}{\partial \eta} = \frac{1}{\eta} \sum_{(i,j)} (1 - Q_{(ij)}) A_{ij} A_{ji} - \sum_{(i,j)} (1 - Q_{(ij)}) \frac{\lambda_{ij} \lambda_{ji}}{\lambda_{ij} + \lambda_{ji} + \eta \lambda_{ij} \lambda_{ji} + 1} \stackrel{!}{=} 0 \tag{B8}$$

leads to a fixed-point equation,

$$\eta = f(\eta) = \frac{\sum_{(i,j)} (1 - Q_{(ij)}) A_{ij} A_{ji}}{\sum_{(i,j)} (1 - Q_{(ij)}) \left[\frac{\lambda_{ij} \lambda_{ji}}{\lambda_{ij} + \lambda_{ji} + \eta \lambda_{ij} \lambda_{ji} + 1} \right]}, \tag{B9}$$

which can be solved numerically with fixed-point methods. Alternatively, one can use root-finding methods to solve directly Eq. (B8) in η .

The equations for the remaining parameters need to be solved using Jensen's inequality, and using $\log x < x$ to obtain $-\log x > -x$

$$L(\Theta) \geq - \sum_{\sigma} [q(\sigma) \log q(\sigma)] + \tag{B10}$$

$$\begin{aligned}
& + \sum_{(i,j)} \left\{ (1 - Q_{(ij)}) \left(A_{ij} \sum_{k,q} \rho_{ijkq} \log \left(\frac{u_{ik} v_{jq} w_{kq}}{\rho_{ijkq}} \right) + A_{ji} \sum_{k,q} \rho_{jikq} \log \left(\frac{u_{jk} v_{iq} w_{kq}}{\rho_{jikq}} \right) \right. \right. \\
& \left. \left. + A_{ij} A_{ji} \log \eta - \left[\sum_{k,q} u_{ik} v_{jq} w_{kq} + \sum_{k,q} u_{jk} v_{iq} w_{kq} + \eta \sum_{k,q} u_{ik} v_{jq} w_{kq} \sum_{k,q} u_{jk} v_{iq} w_{kq} + 1 \right] \right) \right\} \\
& + Q_{(ij)} \left((A_{ij} + A_{ji}) \log \pi - 2 \log(\pi + 1) \right) + Q_{(ij)} \log \mu + (1 - Q_{(ij)}) \log(1 - \mu) \tag{B11}
\end{aligned}$$

$$\tag{B12}$$

and the equality holds when

$$\rho_{ijkq} = \frac{u_{ik} v_{jq} w_{kq}}{\sum_{k,q} u_{ik} v_{jq} w_{kq}}. \tag{B13}$$

We derive community parameters, for example we start by considering u_{ik}

$$\frac{\partial L(\Theta)}{\partial u_{ik}} = \sum_j \left[(1 - Q_{(ij)}) \left[A_{ij} \sum_q \rho_{ijkq} \frac{1}{u_{ik}} - \sum_q v_{jq} w_{kq} - \sum_q \eta v_{jq} w_{kq} \lambda_{ji} \right] \right] \stackrel{!}{=} 0 \tag{B14}$$

and we finally obtain

$$u_{ik} = \frac{\sum_{jq} (1 - Q_{(ij)}) A_{ij} \rho_{ijkq}}{\sum_j \left[\frac{\sum_q (1 - Q_{(ij)}) (1 + \eta \lambda_{ji}) v_{jq} w_{kq}}{\lambda_{ij} + \lambda_{ji} + \eta \lambda_{ij} \lambda_{ji} + 1} \right]}. \tag{B15}$$

We find similar expression for v_{ik} and w_{kq} :

$$v_{ik} = \frac{\sum_{jq} (1 - Q_{(ij)}) A_{ji} \rho_{jiqk}}{\sum_j \left[\frac{\sum_q (1 - Q_{(ij)}) (1 + \eta \lambda_{ij}) u_{jq} w_{qk}}{\lambda_{ij} + \lambda_{ji} + \eta \lambda_{ij} \lambda_{ji} + 1} \right]} \quad (\text{B16})$$

$$w_{kq} = \frac{\sum_{i,j} (1 - Q_{(ij)}) A_{ij} \rho_{ijkq}}{\sum_{i,j} \left[\frac{(1 - Q_{(ij)}) (1 + \eta \lambda_{ij}) u_{ik} v_{jq}}{\lambda_{ij} + \lambda_{ji} + \eta \lambda_{ij} \lambda_{ji} + 1} \right]} . \quad (\text{B17})$$

For the π it yields to the following:

$$\pi = \frac{\sum_{(i,j)} Q_{(ij)} (A_{ij} + A_{ji})}{\sum_{(i,j)} Q_{(ij)} (2 - A_{ij} - A_{ji})} . \quad (\text{B18})$$

Similarly for μ :

$$\frac{\partial L(\Theta)}{\partial \mu} = \sum_{(i,j)} \frac{1}{\mu} Q_{(ij)} - \frac{1}{1 - \mu} \sum_{(i,j)} (1 - Q_{(ij)}) \stackrel{!}{=} 0 \quad (\text{B19})$$

yielding

$$\mu = \frac{1}{N(N-1)/2} \sum_{(i,j)} Q_{(ij)} . \quad (\text{B20})$$

To evaluate $q(\sigma)$, we substitute the estimated parameters inside Eq. (B4):

$$\begin{aligned} q(\sigma) &= \frac{\prod_{(i,j)} \left[\frac{\exp\{(A_{ij} + A_{ji}) f^{(a)}\}}{Z_{(ij)}^{(a)}} \times \mu \right]^{\sigma_{(ij)}} \left[\frac{\exp\{A_{ij} f_{ij}^{(r)} + A_{ji} f_{ji}^{(r)} + A_{ij} A_{ji} J_{(ij)}^{(r)}\}}{Z_{(ij)}^{(r)}} \times (1 - \mu) \right]^{1 - \sigma_{(ij)}}}{\sum_{\sigma_{(ij)}} \prod_{(i,j)} \left[\frac{\exp\{(A_{ij} + A_{ji}) f^{(a)}\}}{Z_{(ij)}^{(a)}} \times \mu \right]^{\sigma_{(ij)}} \left[\frac{\exp\{A_{ij} f_{ij}^{(r)} + A_{ji} f_{ji}^{(r)} + A_{ij} A_{ji} J_{(ij)}^{(r)}\}}{Z_{(ij)}^{(r)}} \times (1 - \mu) \right]^{1 - \sigma_{(ij)}}} \\ &= \prod_{(i,j)} \frac{\left[\frac{\exp\{(A_{ij} + A_{ji}) f^{(a)}\}}{Z_{(ij)}^{(a)}} \times \mu \right]^{\sigma_{(ij)}} \left[\frac{\exp\{A_{ij} f_{ij}^{(r)} + A_{ji} f_{ji}^{(r)} + A_{ij} A_{ji} J_{(ij)}^{(r)}\}}{Z_{(ij)}^{(r)}} \times (1 - \mu) \right]^{1 - \sigma_{(ij)}}}{\sum_{\sigma_{(ij)}=0,1} \left[\frac{\exp\{(A_{ij} + A_{ji}) f^{(a)}\}}{Z_{(ij)}^{(a)}} \times \mu \right]^{\sigma_{(ij)}} \left[\frac{\exp\{A_{ij} f_{ij}^{(r)} + A_{ji} f_{ji}^{(r)} + A_{ij} A_{ji} J_{(ij)}^{(r)}\}}{Z_{(ij)}^{(r)}} \times (1 - \mu) \right]^{1 - \sigma_{(ij)}}} \\ &= \prod_{(i,j)} Q_{(ij)}^{\sigma_{(ij)}} (1 - Q_{(ij)})^{(1 - \sigma_{(ij)})} , \end{aligned} \quad (\text{B21})$$

where

$$\begin{aligned} Q_{(ij)} &= \frac{\exp[(A_{ij} + A_{ji}) f^{(a)} - \log Z_{(ij)}^{(a)}] \mu}{\exp[(A_{ij} + A_{ji}) f^{(a)} - \log Z_{(ij)}^{(a)}] \mu + \exp[f_{ij}^{(r)} A_{ij} + f_{ji}^{(r)} A_{ji} + J_{(ij)}^{(r)} A_{ij} A_{ji} - \log Z_{(ij)}^{(r)}] (1 - \mu)} \\ &= \frac{\exp[(A_{ij} + A_{ji}) \log \pi - 2 \log(\pi + 1)] \mu}{\exp[(A_{ij} + A_{ji}) \log \pi - 2 \log(\pi + 1)] \mu + \exp[A_{ij} \log \lambda_{ij} + A_{ji} \log \lambda_{ji} + \log \eta A_{ij} A_{ji} - \log Z_{(ij)}^{(r)}] (1 - \mu)} \\ &= \frac{\frac{\pi^{(A_{ij} + A_{ji})} \mu}{Z_{(ij)}^{(a)}}}{\frac{\pi^{(A_{ij} + A_{ji})} \mu}{Z_{(ij)}^{(a)}} + \frac{\lambda_{ij}^{A_{ij}} \lambda_{ji}^{A_{ji}} \eta^{A_{ij} A_{ji}} (1 - \mu)}{Z_{(ij)}^{(r)}}} . \end{aligned} \quad (\text{B22})$$

Notice that this is exactly the expected value w.r.t. the variational distribution as previously defined.

1. Convergence criteria

The EM algorithm consists of randomly initializing u, v, w, η, π, μ , then iterating Eqs. B13, B22, B15-B17, B9, B18, B20, until the convergence of the following log-posterior,

$$\begin{aligned}
L(\Theta) &= \log P(\Theta|A) \geq \sum_{\sigma} q(\sigma) \log \frac{P(\sigma, \Theta|A)}{q(\sigma)} \\
&= -\sum_{\sigma} q(\sigma) \log q(\sigma) + \sum_{\sigma} q(\sigma) \{ \log P(A|\sigma; \Theta) + \log P(\sigma|\mu) \} \\
&= -\sum_{\sigma} q(\sigma) \log q(\sigma) \\
&\quad + \sum_{\sigma(ij)} q(\sigma(ij)) \left\{ \sum_{(ij)} \left[(1 - \sigma(ij)) \left(A_{ij} f_{ij}^{(r)} + A_{ji} f_{ji}^{(r)} + A_{ij} A_{ji} J_{(ij)}^{(r)} - \log Z_{(ij)}^{(r)} \right) + \sigma(ij) \left((A_{ij} + A_{ji}) f^{(a)} - \log Z_{(ij)}^{(a)} \right) \right. \right. \\
&\quad \left. \left. + \sigma(ij) \log \mu + (1 - \sigma(ij)) \log(1 - \mu) \right] \right\} \\
&= -\sum_{(i,j)} \left[Q_{(ij)} \log Q_{(ij)} + (1 - Q_{(ij)}) \log(1 - Q_{(ij)}) \right] \\
&\quad + \sum_{(i,j)} \left\{ (1 - Q_{(ij)}) \left(f_{ij}^{(r)} A_{ij} + A_{ji} f_{ji}^{(r)} + A_{ij} A_{ji} J_{(ij)}^{(r)} - \log Z_{(ij)}^{(r)} \right) + \right. \\
&\quad \left. + Q_{(ij)} \left((A_{ij} + A_{ji}) f^{(a)} - \log Z_{(ij)}^{(a)} \right) + Q_{(ij)} \log \mu + (1 - Q_{(ij)}) \log(1 - \mu) \right\} + const \quad , \tag{B23}
\end{aligned}$$

where we neglect *const*, constant terms due to the uniform priors. To calculate $q(\sigma)$, we used Eq. (B21), i.e., a Bernoulli distribution.

APPENDIX C: GENERATIVE MODEL

Being generative, the model can be used to generate synthetic networks with anomalies. For this, one should sample the latent parameters $\Theta = (u, v, w, \eta, \pi, \mu)$, then sample σ given the parameters. Finally, given the σ and the latent parameters, the adjacency matrix \mathbf{A} could be constructed. For a given set of community parameters as the input parameters, [15, 22], the expected number of anomalous and non-anomalous edges are $N^2 \mu \frac{\pi}{(1+\pi)}$, and $\mathbb{E}[M] = (1 - \mu) \sum_{i,j} \frac{\lambda_{ij} + \eta \lambda_{ij} \lambda_{ji}}{Z_{(ij)}^{(r)}}$, respectively. Assuming a desired total number of edges E , we can thus multiply π, μ and M by suitable sparsity constants that tune: i) the ratio of anomalous edges to the total number of edges, $\rho_a = N^2 \mu \frac{\pi}{(1+\pi)} / (N^2 \mu \frac{\pi}{(1+\pi)} + (1 - \mu) \mathbb{E}[M]) \in [0, 1]$; ii) the success rate of anomalous edges π . Once these two are fixed, the remaining sparsity parameter for the matrix M , is estimated as:

$$E(1 - \rho_a) = (1 - \mu) \sum_{i,j} \frac{\zeta \lambda_{ij} + \eta \zeta \lambda_{ij} \zeta \lambda_{ji}}{\zeta \lambda_{ij} + \zeta \lambda_{ji} + \eta \zeta \lambda_{ij} \zeta \lambda_{ji} + 1} \tag{C1}$$

which can be solved with root-finding methods.

APPENDIX D: RESULTS ON SYNTHETIC NETWORKS

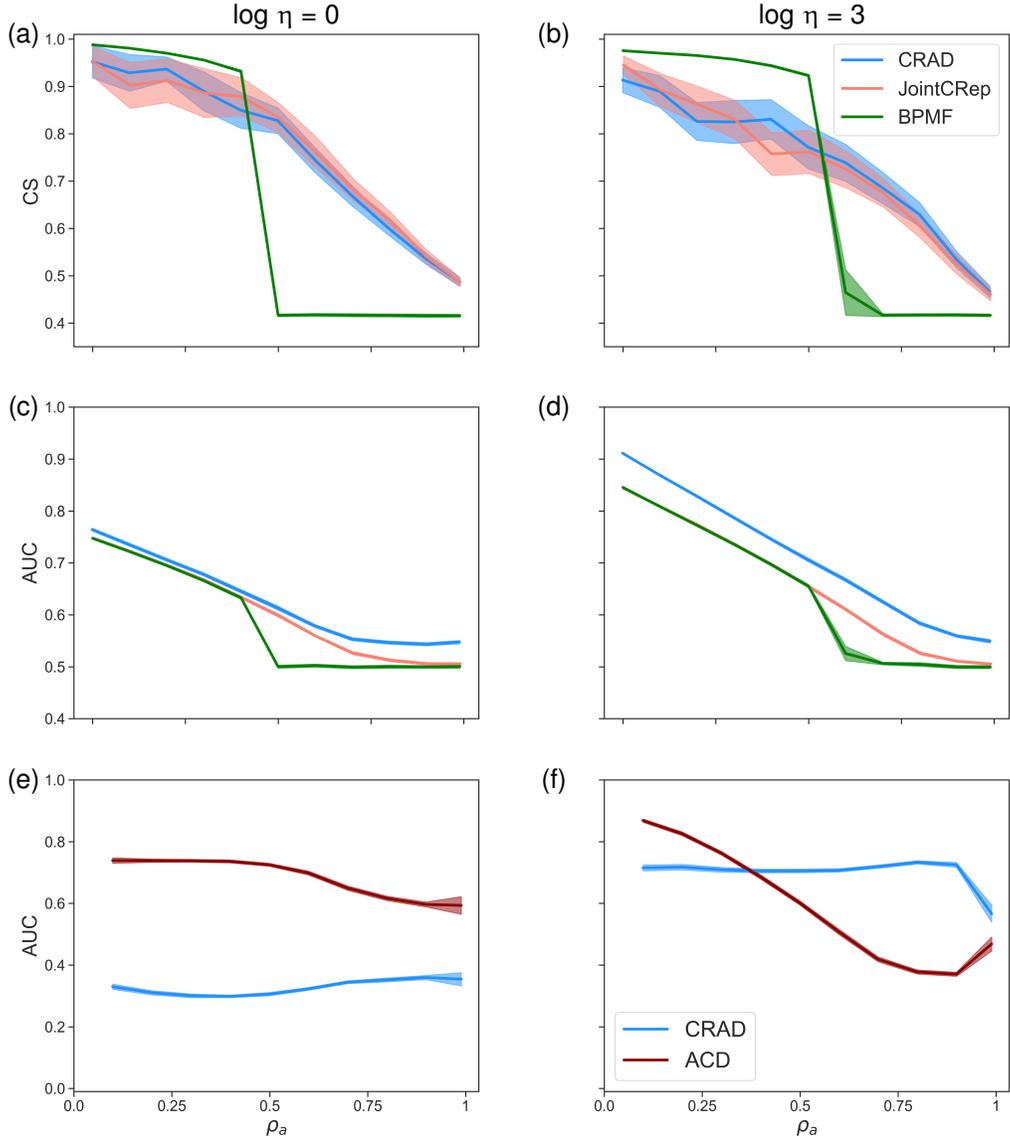


FIG. 6: Community detection, link prediction, and anomaly detection on synthetic network datasets. (a)-(b) We compare the performance of CRAD against JointCRep and BPMF algorithms in community detection, as measured by cosine similarity (CS); and (c)-(d) in link prediction tasks, as measured by AUC on held-out data. In addition, (e)-(f) we test the ability to detect anomalies against a model that does not include a reciprocity effect (ACD), as measured by the AUC on a binary dataset that contains what edges are regular and what are anomalous. The datasets have $N = 500$, average degree $\langle k \rangle = 60$, $K = 3$. The first column is for networks generated without reciprocity, $\log \eta = 0$, while the second column is for networks with positive reciprocity, $\log \eta = 3$. In the x -axis we vary ρ_a , the ratio of anomalous edges over the total number of edges. Lines and shades around them are averages and standard deviations over 10 network samples, respectively.

APPENDIX E: REAL DATA: DATASET DESCRIPTION

Table I provides a summary of the key characteristics of the studied datasets. The dataset of UC Irvine messages and Online dating (POK0) have undergone pre-processing that involved the removal of self-loops, retaining only nodes with both incoming and outgoing edges, and using only the giant connected components.

TABLE I: Real-world datasets description.

Network	Abbreviation	N	E	$\langle k \rangle$	reciprocity	Ref.
Vampire bat	vampire bat	19	103	10.8	0.64	[24]
A Nicaraguan community	Nicaraguan	108	1517	14.05	0.11	[25]
UC Irvine messages	uc-social	1302	19044	29.3	0.68	[26]
Online dating	POK	3562	18098	10.2	0.78	[39]

- [1] L. Akoglu, H. Tong, and D. Koutra, Graph based anomaly detection and description: a survey, *Data Mining and Knowledge Discovery* **29**, 626 (2015).
- [2] V. J. Hodge and J. Austin, A survey of outlier detection methodologies, *Artificial Intelligence Review* **22**, 85 (2004).
- [3] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, Network anomaly detection: Methods, systems and tools, *IEEE Communications Surveys and Tutorials* **16**, 303 (2014).
- [4] D. M. Hawkins, *Identification of outliers (monographs on statistics and applied probability)*, Vol. 11 (Springer, London:Chapman & Hall, 1980).
- [5] E. Aleskerov, B. Freisleben, and B. Rao, Cardwatch: a neural network based database mining system for credit card fraud detection, in *Proceedings of the IEEE/IAFE 1997 Computational Intelligence for Financial Engineering (CIFEr)* (1997) pp. 220–226.
- [6] V. Chandola, A. Banerjee, and V. Kumar, Anomaly detection: A survey, *ACM Comput. Surv.* **41**, 10.1145/1541880.1541882 (2009).
- [7] E. Eskin, Anomaly detection over noisy data using learned probability distributions, in *Proceedings of the Seventeenth International Conference on Machine Learning, ICML '00* (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2000) p. 255–262.
- [8] A. Goldenberg, A. X. Zheng, S. E. Fienberg, and E. M. Airolidi, A survey of statistical network models, (2010).
- [9] A. Bojchevski and S. Günnemann, Bayesian robust attributed graph clustering: Joint learning of partial anomalies and group structure, in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32 (2018).
- [10] D. Savage, X. Zhang, X. Yu, P. Chou, and Q. Wang, Anomaly detection in online social networks, *Social Networks* **39**, 62 (2014).
- [11] K. Henderson, T. Eliassi-Rad, C. Faloutsos, L. Akoglu, L. Li, K. Maruhashi, B. A. Prakash, and H. Tong, Metric forensics: A multi-level approach for mining volatile graphs, in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '10 (Association for Computing Machinery, New York, NY, USA, 2010) p. 163–172.
- [12] V. Nikulin and T.-H. Huang, Unsupervised dimensionality reduction via gradient-based matrix factorization with two adaptive learning rates, in *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, Proceedings of Machine Learning Research, Vol. 27, edited by I. Guyon, G. Dror, V. Lemaire, G. Taylor, and D. Silver (PMLR, Bellevue, Washington, USA, 2012) pp. 181–194.
- [13] B. Perozzi and L. Akoglu, Discovering communities and anomalies in attributed graphs: Interactive visual exploration and summarization, *ACM Trans. Knowl. Discov. Data* **12**, 10.1145/3139241 (2018).
- [14] S. Fortunato and D. Hric, Community detection in networks: A user guide, *Physics reports* **659**, 1 (2016).
- [15] H. Safdari and C. De Bacco, Anomaly detection and community detection in networks, *Journal of Big Data* **9**, 122 (2022).
- [16] S. Wasserman, K. Faust, *et al.*, *Social network analysis: Methods and applications*, Vol. 8 (Cambridge university press, 1994).
- [17] C. De Bacco, M. Contisciani, J. Cardoso-Silva, H. Safdari, D. Baptista, T. Sweet, J.-G. Young, J. Koster, C. T. Ross, R. McElreath, D. Redhead, and E. A. Power, Latent network models to account for noisy, multiply-reported social network data (2021).
- [18] L. D. Molm, The structure of reciprocity, *Social Psychology Quarterly* **73**, 119 (2010), <https://doi.org/10.1177/0190272510369079>.
- [19] M. Contisciani, H. Safdari, and C. De Bacco, Community detection and reciprocity in networks by jointly modelling pairs of edges, *Journal of Complex Networks* **10**, 10.1093/comnet/cnac034 (2022), cnac034, <https://academic.oup.com/comnet/article-pdf/10/4/cnac034/45291210/cnac034.pdf>.
- [20] H. Safdari, M. Contisciani, and C. De Bacco, Generative model for reciprocity and community detection in networks, *Phys. Rev. Research* **3**, 023209 (2021).
- [21] B. Dai, S. Ding, G. Wahba, *et al.*, Multivariate bernoulli distribution, *Bernoulli* **19**, 1465 (2013).

- [22] C. De Bacco, E. A. Power, D. B. Larremore, and C. Moore, Community detection, link prediction, and layer interdependence in multilayer networks, *Phys. Rev. E* **95**, 042317 (2017).
- [23] P. Gopalan, J. M. Hofman, and D. M. Blei, Scalable recommendation with hierarchical poisson factorization., in *UAI* (2015) pp. 326–335.
- [24] G. G. Carter and G. S. Wilkinson, Food sharing in vampire bats: reciprocal help predicts donations more than relatedness or harassment, *Proceedings of the Royal Society B: Biological Sciences* **280**, 20122573 (2013).
- [25] J. Koster, Family ties: the multilevel effects of households and kinship on the networks of individuals, *Royal Society Open Science* **5**, 172159 (2018), <https://royalsocietypublishing.org/doi/pdf/10.1098/rsos.172159>.
- [26] J. Kunegis, KONECT – The Koblenz Network Collection, in *Proc. Int. Conf. on World Wide Web Companion* (2013) pp. 1343–1350.
- [27] Y. Lucas, P.-E. Portier, L. Laporte, S. Calabretto, O. Caelen, L. He-Guelton, and M. Granitzer, Multiple perspectives hmm-based feature engineering for credit card fraud detection, *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing* 10.1145/3297280.3297586 (2019).
- [28] V. Van Vlasselaer, C. Bravo, O. Caelen, T. Eliassi-Rad, L. Akoglu, M. Snoeck, and B. Baesens, Apatate: A novel approach for automated credit card transaction fraud detection using network-based extensions, *Decision Support Systems* **75**, 38 (2015).
- [29] M. E. Newman and A. Clauset, Structure and inference in annotated networks, *Nature communications* **7**, 1 (2016).
- [30] M. Contisciani, E. A. Power, and C. De Bacco, Community detection with node attributes in multilayer networks, *Scientific reports* **10**, 1 (2020).
- [31] X. Zhu and A. B. Goldberg, Introduction to semi-supervised learning, *Synthesis lectures on artificial intelligence and machine learning* **3**, 1 (2009).
- [32] K. Liu, Y. Dou, Y. Zhao, X. Ding, X. Hu, R. Zhang, K. Ding, C. Chen, H. Peng, K. Shu, G. H. Chen, Z. Jia, and P. S. Yu, Pygod: A python library for graph outlier detection, *arXiv preprint arXiv:2204.12095* (2022).
- [33] K. Liu, Y. Dou, Y. Zhao, X. Ding, X. Hu, R. Zhang, K. Ding, C. Chen, H. Peng, K. Shu, L. Sun, J. Li, G. H. Chen, Z. Jia, and P. S. Yu, Bond: Benchmarking unsupervised outlier node detection on static attributed graphs, *arXiv preprint arXiv:2206.10071* (2022).
- [34] E. Müller, P. I. Sánchez, Y. Mülle, and K. Böhm, Ranking outlier nodes in subspaces of attributed graphs, in *2013 IEEE 29th International Conference on Data Engineering Workshops (ICDEW)* (2013) pp. 216–222.
- [35] N. Masuda and R. Lambiotte, *A guide to temporal networks* (World Scientific, 2016).
- [36] X. Zhang, C. Moore, and M. E. Newman, Random graph models for dynamic networks, *The European Physical Journal B* **90**, 1 (2017).
- [37] H. Safdari, M. Contisciani, and C. De Bacco, Reciprocity, community detection, and link prediction in dynamic networks, *Journal of Physics: Complexity* **3**, 015010 (2022).
- [38] T. P. Peixoto and M. Rosvall, Modelling sequences and temporal networks with dynamic community structures, *Nature communications* **8**, 1 (2017).
- [39] H. Makse, POK Dataset (<https://hmakse.cuny.cuny.edu/>).