entropy

MDPI

*Article*

# Reconstructing Sparse Multiplex Networks with Application to Covert Networks

**Jin-Zhu Yu [1], Mincheng Wu [2], Gisela Bichler [3], Felipe Aros-Vera [4] and Jianxi Gao [5,6,*]**

[1] Department of Civil Engineering, University of Texas at Arlington, Arlington, TX 76019, USA
[2] State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310058, China
[3] School of Criminology and Criminal Justice, California State University, San Bernardino, CA 92407, USA
[4] Department of Industrial and Systems Engineering, Ohio University, Athens, OH 45701, USA
[5] Department of Computer Science, Rensselaer Polytechnic Institute (RPI), Troy, NY 12180, USA
[6] Network Science and Technology Center, Rensselaer Polytechnic Institute (RPI), Troy, NY 12180, USA
* Correspondence: jianxi.gao@gmail.com

**Abstract:** Network structure provides critical information for understanding the dynamic behavior of complex systems. However, the complete structure of real-world networks is often unavailable, thus it is crucially important to develop approaches to infer a more complete structure of networks. In this paper, we integrate the configuration model for generating random networks into an Expectation–Maximization–Aggregation (EMA) framework to reconstruct the complete structure of multiplex networks. We validate the proposed EMA framework against the Expectation–Maximization (EM) framework and random model on several real-world multiplex networks, including both covert and overt ones. It is found that the EMA framework generally achieves the best predictive accuracy compared to the EM framework and the random model. As the number of layers increases, the performance improvement of EMA over EM decreases. The inferred multiplex networks can be leveraged to inform the decision-making on monitoring covert networks as well as allocating limited resources for collecting additional information to improve reconstruction accuracy. For law enforcement agencies, the inferred complete network structure can be used to develop more effective strategies for covert network interdiction.

**Keywords:** multiplex networks; partially observable networks; interlayer dependency; network completion; expectation–maximization



**Citation:** Yu, J.-Z.; Wu, M.; Bichler, G.; Aros-Vera, F.; Gao, J. Reconstructing Sparse Multiplex Networks with Application to Covert Networks. *Entropy* **2023**, *25*, 142. https://doi.org/10.3390/e25010142

Academic Editor: Stanisław Drożdż

Received: 24 November 2022
Revised: 7 January 2023
Accepted: 8 January 2023
Published: 10 January 2023



## 1. Introduction

Multiplex networks, in which each layer consists of the same set of nodes but different sets of links (Figure 1), are a powerful tool for describing and analyzing the connectivity from interactions of different types among the same set of entities in complex systems [1]. Examples include social networks [2], multi-modal transportation networks [3], multi-tier supply networks [4], etc. Due to the interplay between network structure and dynamics, the knowledge of complete multiplex network structure is essential to a deep understanding of the dynamic behavior of multiplex networks and predicting future interactions between nodes in the networks [5]. Nonetheless, in many practical applications, particularly covert networks, it is extremely difficult to obtain sufficient data for constructing the complete structure of all layers of a multiplex network due to limited observing resources and privacy concerns [6].

Reconstructing the complete topology (i.e., all nodes and links) from incomplete topologies (partially observed nodes and links) is an exceedingly challenging task primarily because (i) the number of missing nodes and links will be very large; (ii) sparse networks (e.g., illicit supply networks) usually have much lower link density than typical (overt) social networks, thereby it is often inappropriate to use many prediction methods developed for denser networks with a balanced ratio of positive and negative labels; and

(iii) observations are limited to task-specific activity, excluding historic relations that are instrumental to network resilience [7,8].

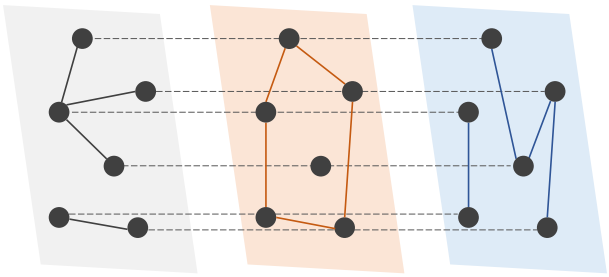

**Figure 1.** A schematic of a multiplex network with three layers and six nodes. In a multiplex network, the nodes are shared across layers, and links in each layer represent one type of relation between nodes. An interlayer link exists between any node and its counterparts in other layers.

On the application side, this study focuses on illicit networks, such as drug and human trafficking networks. The reason is that illicit networks compromise national security and prosperity as well as public health and safety [9]. Despite considerable advances in understanding and disrupting the operations of covert networks, a deep understanding of the structure and dynamics of covert networks remains lacking. Although the structure is crucial for understanding, monitoring, and controlling covert networks, this topological information is always incomplete [10,11]. There are multiple reasons: (i) Investigations typically focus on current operations with the aim of documenting the involvement of specific individuals in specific criminal activities. (ii) Potential criminal actors may be identified and are then often omitted from the network due to essential legal principles and civil rights protections. (iii) Criminal actors take measures to obfuscate their involvement, hiding, or disguising activities and relations because they are wary of attack by agents of the criminal justice system, as well as competitors or rivals. By modeling illicit networks as a multiplex and reconstructing the structure of multiplex covert networks, this study can be leveraged to help monitor, control, and interdict the illicit activities on covert networks.

The rest of this paper is structured as follows. Section 2 presents the literature review. Details on the EMA framework are presented in Section 3. Details on the real-world multiplex networks used to validate the proposed EMA framework are given in Section 4. The numerical experiments and the resultant results are reported in Section 5, followed by the discussion on the implication of this study on disrupting illicit networks in Section 6 and the concluding remarks in Section 7.

## 2. Literature Review

Network reconstruction methods can be divided into three categories: similarity-based approach, probabilistic (statistical) approach, and algorithmic (machine learning) approach.

Structural similarity-based methods assume that the nodes tend to connect to other nodes with a higher level of similarity computed by a certain distance function [12]. These approaches give similarity scores for missing or unobserved links between any pair of two nodes, then links with high similarity scores will be predicted to exist. For example, in [13], the eigenvectors of the layer adjacency matrix are used to measure the layer topological similarity and then the topological similarity (element of the layer similarity matrix) of the unconnected nodes is used to predict if a link exists between them. Berlusconi et al. [14] and Calderoni et al. [15] apply multiple similarity metrics, such as common neighbor and resource allocation, to identify missing links in a criminal network. They demonstrated that these similarity metrics can identify possible missing links in criminal networks with noise or incomplete information. However, the performance of similarity-based methods depends on the structure of specific networks [16]; therefore, prior knowledge about the networks under study is usually required to improve the predictive performance [13]. This approach is infeasible for the network reconstruction problem in this study because

the nodes and links to be inferred are completely missing without any connection with observed nodes. Therefore, all scores of structural similarity between unobserved nodes and observed nodes are zeros.

Statistical approaches for network structure inference assume a network has a known structure, and then a parametric or nonparametric model is built to fit the structure, such as the stochastic block model or its variants, as well as the Dirichlet network distribution model [17–20]. Then, the model parameters are estimated using statistical methods, such as maximum likelihood estimation. The model parameters are then used to compute the formation probability of each unobserved link, such as the stochastic block model and the hierarchical structure model. In particular, Kim and Leskovec [21] developed a scalable Kronecker Expectation–Maximization (KronEM) approach to reconstruct a single-layer network with a known number of missing nodes by alternating between inferring the unobserved part of the network (E-step) and estimating the optimal model parameters (M-step) until convergence. The network structure is assumed to be described by the Kronecker graph model, and this approach does not require attributes about nodes or edges. Wu et al. [22] use a Bayesian approach to reconstruct two-layer multiplex networks from the available aggregate topology of all layers (the aggregation of all the layers into a monoplex layer) and partial observations of nodes in each layer, and the inference is conducted by leveraging the Expectation–Maximization algorithm. However, statistical approaches would be computationally challenging for large networks, especially when a large number of samples are required for inference [21].

The previous two approaches are based on computing a score for each unobserved link by defining a similarity or a probability function. However, the inference of network structure can also benefit from other algorithmic (machine learning) approaches, including supervised learning and optimization techniques. In particular, Zhang and Chen [23] develop graph neural network (GNN) models to predict missing links. However, the method is mainly developed for single-layer networks. Gao et al. [24] developed a network embedding method called interactive learning across relations that exploits existing multiple types of relational data to predict inductive links (new or unobserved relation) between existing nodes instead of inter- or intra-layer links in a known layer.

On the application side, a growing number of researchers are resorting to multilayer network models to study illicit supply chains [11,25–28] since these models allow actors to be placed in different layers representing different characteristics of covert networks. Going beyond mapping, only the co-offending activity, multilayer models help to reveal more about the social processes that connect people, which can be used to infer missing information [29,30]. However, to the best of our knowledge, a few studies have attempted to infer their multiplayer structure (topology), even though the complete structure of multiplex illicit networks is typically missing, as discussed earlier.

In this study, we develop an EMA framework built on the EM algorithm. The EMA framework is advantageous for reconstructing covert networks because:

1. The sparsity of links in covert networks, i.e., the negatives (no links between nodes) significantly outnumber the positives, leads to challenges in inference and thus significantly affects the predictive performance of classical learning-based methods [31].
2. It can infer the links that connect unobserved nodes of the network. This is hardly possible for methods built on structural similarities, such as Adamic–Adar index, Jaccard index, and Resource Allocation index [32], because the links to be inferred usually connect unobserved nodes that are isolated from the observed part of the network.
3. It embeds observed criminal activity within the latent social structure required to sustain operations within a dynamic operational context [33–37]. This informs disruption efforts by improving estimates of network resiliency [38,39].

The EMA framework is an extension to the EM framework proposed in a recent work [22], but the EMA framework is more generic because it can be applied to multiplex networks with more than two layers and unknown aggregate topology, which is more common for real-world incomplete multiplex networks.

## 3. Methodology

### *3.1. Problem Description*

Mathematically, a multiplex network can be described by a set $\boldsymbol{Z} = \{\boldsymbol{Z}^\ell | \, \ell = 1, 2, \ldots, m\}$, where $m$ is the number of layers. $\boldsymbol{Z}^\ell$ consists of the complete set of nodes $\mathcal{N}$ and the complete set of edges $\mathcal{E}^\ell$ of relation type $\ell$. For incomplete multiplex networks, only some parts (partial topologies) of each layer are observed, i.e., a fraction $c$ of nodes and the associated links of the complete network. If we denote the partially observed topologies as $\boldsymbol{X} = \{\boldsymbol{X}^\ell | \, \ell = 1, 2, \ldots, m\}$, then we have $\boldsymbol{X}^\ell \subseteq \boldsymbol{Z}^\ell$. In this study, we use the same notation for topology and the adjacency matrix that encodes the connectivity of the topology. We consider undirected and unweighted multiplex networks; therefore, the observed and complete adjacency matrices for each layer of a multiplex network are symmetric and binary (i.e., $\boldsymbol{Z}^\ell_{ij}, \boldsymbol{X}^\ell_{ij} \in \{0, 1\}$, $\forall i, j \in \mathcal{N}$). Formally, the problem of this study can be given as follows.

**Problem.** *Consider an undirected and unweighted multiplex network wherein each layer $\ell$ have the same set of nodes $\mathcal{N}$ but a different set of links $\mathcal{E}^\ell$. The multiplex network reconstruction problem is to infer the complete topology of all layers $\boldsymbol{Z}$ given partial topologies of all layers $\boldsymbol{X}$.*

Note that the classical link prediction problem [32], in which only links are missing and typically not all links connected a node are missing, is a special case of the network reconstruction problem, in which both nodes and links can be missing, i.e., all links connected to a missing node are missing as well. As such, the network reconstruction problem is naturally more difficult than the link prediction problem.

### *3.2. Expectation–Maximization–Aggregation Framework*

This section provides details on the proposed EMA framework for multiplex network reconstruction, including the classical EM algorithm, the network generation model, the integration of aggregate topology into the EM framework, as well as the corresponding algorithms.

#### 3.2.1. Expectation–Maximization

The EM algorithm [40] is an iterative approach for estimation problems wherein the value of model parameters and latent variables depend on each other. For the network reconstruction problem, where the task is to find the most probable topology of the missing subnetwork, it is required to first estimate the parameters of the network model, which establish a connection between the observed subnetworks $\boldsymbol{X}$ and the complete multiplex network $\boldsymbol{Z}$. This framework can be formulated as follows.

We here consider a model parameterized by $\boldsymbol{\Theta}$. Given the observed subnetworks $\boldsymbol{X}$ of the complete multiplex network $\boldsymbol{Z}$, the task of classical maximum likelihood estimation is to estimate $\boldsymbol{\Theta}$. However, since the complete observations $\boldsymbol{Z}$ are not available, we need to marginalize it out. Let $q(\boldsymbol{Z})$ be the probability density distribution of $\boldsymbol{Z}$, i.e., $\sum_{\boldsymbol{Z}} q(\boldsymbol{Z}) = 1$ and $q(\boldsymbol{Z}) \geq 0$, then we obtain the likelihood of the observed topology given the model

$$\ln p(\boldsymbol{X} \mid \boldsymbol{\Theta}) = \ln \sum_{\boldsymbol{Z}} p(\boldsymbol{X}, \boldsymbol{Z} \mid \boldsymbol{\Theta}) \tag{1a}$$

$$= \ln \sum_{\boldsymbol{Z}} q(\boldsymbol{Z}) \cdot \frac{p(\boldsymbol{X}, \boldsymbol{Z} \mid \boldsymbol{\Theta})}{q(\boldsymbol{Z})} \tag{1b}$$

$$\geq \sum_{\boldsymbol{Z}} q(\boldsymbol{Z}) \ln \left( \frac{p(\boldsymbol{X}, \boldsymbol{Z} \mid \boldsymbol{\Theta})}{q(\boldsymbol{Z})} \right). \tag{1c}$$

The last inequality follows from Jensen's equality. Specifically, $\ln \mathbb{E}[x] \geq \mathbb{E}[\ln x]$ since $\ln x$ is a concave function. The equality is attained when $\frac{p(\boldsymbol{X},\boldsymbol{Z}|\boldsymbol{\Theta})}{q(\boldsymbol{Z})}$ is a constant, i.e., $q(\boldsymbol{Z}) \propto p(\boldsymbol{X},\boldsymbol{Z} \mid \boldsymbol{\Theta})$. Typically, $q(\boldsymbol{Z})$ is set to the posterior distribution of $\boldsymbol{Z}$, thus we get

$$q(\boldsymbol{Z}) = \frac{p(\boldsymbol{X},\boldsymbol{Z} \mid \boldsymbol{\Theta})}{\sum_{\boldsymbol{Z}} p(\boldsymbol{X},\boldsymbol{Z} \mid \boldsymbol{\Theta})} = p(\boldsymbol{Z} \mid \boldsymbol{X},\boldsymbol{\Theta}). \tag{2}$$

The lower bound given by Equation (1c), referred to as the evidence lower bound (ELBO), is used because it is difficult to directly optimize Equation (1a) due to the missing information [41]. The EM algorithm then proceeds by alternating between the following two steps until convergence [40]:

1. E-step: updating $q(\boldsymbol{Z})$ by setting it to the posterior of $\boldsymbol{Z}$ given the current $\boldsymbol{\Theta}$, which is used to construct the local ELBO on the log-likelihood.
2. M-step: updating $\boldsymbol{\Theta}$ by maximizing the ELBO, i.e., solving for $\boldsymbol{\Theta}$ in

$$\frac{\partial}{\partial \boldsymbol{\Theta}} \sum_{\boldsymbol{Z}} q(\boldsymbol{Z}) \ln p(\boldsymbol{X},\boldsymbol{Z} \mid \boldsymbol{\Theta}) = 0. \tag{3}$$

#### 3.2.2. Network Generation Model

In this study, the parameters in the model for generating the topology of each layer include the degree sequences of layers $\boldsymbol{d}$ and the adjacency of the aggregate topology from all layers of the multiplex network $\boldsymbol{A}$. The degree sequence of a layer is used to generate the topology of that layer using the configuration model [42,43]. According to this model, for a network with a given degree sequence, the probability of a link between two nodes with degree $\boldsymbol{d}_i$ and $\boldsymbol{d}_j$ is $\frac{\boldsymbol{d}_i \boldsymbol{d}_j}{2|\mathcal{E}|-1}$ where $|\mathcal{E}|$ is the number of edges. Therefore, after estimating the degree sequences of all layers, we can probabilistically infer the complete topology of each layer. The adjacency matrix of the aggregate topology is included in the model parameters because we add an aggregation step to the EM framework to obtain the aggregated topology of all layers. In this way, information on the interlayer dependence of the multiplex network is leveraged to improve the reconstruction accuracy. Without this step, the topologies of layers are independent of each other and are hence reconstructed separately.

#### 3.2.3. Full Algorithm for the EMA Framework

This section presents details about how to reconstruct multiplex networks using the EMA framework, including the E-step, M-step, and A-step.

In the E-step, the task is to obtain the posterior of $\boldsymbol{Z}$ given by Equation (2). According to the configuration model, the probability of a link between nodes $i$ and $j$ in layer $\ell$ before it is not observed yet can be given by

$$p_{ij}^{\ell} = \frac{\boldsymbol{d}_i^{\ell} \boldsymbol{d}_j^{\ell}}{2|\mathcal{E}^{\ell}| - 1} \tag{4}$$

Considering that $\frac{\boldsymbol{d}_i^{\ell} \boldsymbol{d}_j^{\ell}}{2|\mathcal{E}^{\ell}|-1}$ can be greater than 1 and certain entries are observed, Equation (4) is changed to

$$p_{ij}^{\ell} = \begin{cases} \boldsymbol{X}_{ij}^{\ell}, & \text{if } i \in \mathcal{V}_{\text{obs}} \text{ and } j \in \mathcal{V}_{\text{obs}} \\ \min(1, \frac{\boldsymbol{d}_i^{\ell} \boldsymbol{d}_j^{\ell}}{2|\mathcal{E}^{\ell}|-1}), & \text{otherwise} \end{cases} \tag{5}$$

Then, $q\left(\boldsymbol{Z}_{ij}^{\ell}\right)$ is expressed as

$$q(\boldsymbol{Z}_{ij}^{\ell}) = (p_{ij}^{\ell})^{\boldsymbol{Z}_{ij}^{\ell}} \left(1 - p_{ij}^{\ell}\right)^{1-\boldsymbol{Z}_{ij}^{\ell}}, \boldsymbol{Z}_{ij}^{\ell} \in \{0,1\}, \tag{6}$$

Next, in the M-step, given the current $q\left(\boldsymbol{Z}_{ij}^{\ell}\right)$, we solve for $\boldsymbol{d}_i^{\ell}$ for any $i \in \mathcal{N}$, $\ell \in \mathcal{L}$ as follows. In the ELBO, the term involving $\boldsymbol{d}_i^{\ell}$ is

$$\sum_{\boldsymbol{Z}_{ij}^{\ell}} q(\boldsymbol{Z}_{ij}^{\ell}) \ln \left( \frac{\left(\frac{\boldsymbol{d}_i^{\ell}\boldsymbol{d}_j^{\ell}}{2|\mathcal{E}^{\ell}|-1}\right)^{\boldsymbol{Z}_{ij}^{\ell}} \left(1-\frac{\boldsymbol{d}_i^{\ell}\boldsymbol{d}_j^{\ell}}{2|\mathcal{E}^{\ell}|-1}\right)^{1-\boldsymbol{Z}_{ij}^{\ell}}}{q(\boldsymbol{Z}_{ij}^{\ell})} \right) \tag{7}$$

To simplify the term involving $\boldsymbol{d}_i^{\ell}$ in Equation (7), we let $a_j = \frac{\boldsymbol{d}_j^{\ell}}{2|\mathcal{E}^{\ell}|-1}$. Then differentiate with respect to an arbitrary $\boldsymbol{d}_i^{\ell}$ and we obtain

$$\sum_{\boldsymbol{Z}_{ij}^{\ell}} \frac{\partial}{\partial \boldsymbol{d}_i^{\ell}} \left( q(\boldsymbol{Z}_{ij}^{\ell}) \cdot \ln \left[ (a_j \boldsymbol{d}_i^{\ell})^{\boldsymbol{Z}_{ij}^{\ell}} (1-a_j \boldsymbol{d}_i^{\ell})^{1-\boldsymbol{Z}_{ij}^{\ell}} \right] \right) \tag{8a}$$

$$= \sum_{\boldsymbol{Z}_{ij}^{\ell}} q(\boldsymbol{Z}_{ij}^{\ell}) \cdot \frac{\boldsymbol{Z}_{ij}^{\ell} - a_j \boldsymbol{d}_i^{\ell}}{\boldsymbol{d}_i^{\ell}(1-a_j \boldsymbol{d}_i^{\ell})} \tag{8b}$$

$$= \frac{\mathbb{E}\left[\boldsymbol{Z}_{ij}^{\ell}\right] - a_j \boldsymbol{d}_i^{\ell}}{\boldsymbol{d}_i^{\ell}(1-a_j \boldsymbol{d}_i^{\ell})} \tag{8c}$$

Let Equation (8c) = 0 and we get

$$\boldsymbol{d}_j^{\ell}\boldsymbol{d}_i^{\ell} = \left(2|\mathcal{E}^{\ell}|-1\right)\mathbb{E}\left[\boldsymbol{Z}_{ij}^{\ell}\right], \ \forall i \in \mathcal{N}. \tag{9}$$

To ensure that the solution to $\boldsymbol{d}_i^{\ell}$ does not involve $\boldsymbol{d}_j^{\ell}$, we sum both sides of the above equation over $j$ and obtain

$$\boldsymbol{d}_i^{\ell} = \frac{2|\mathcal{E}^{\ell}|-1}{2|\mathcal{E}^{\ell}| - \boldsymbol{d}_i^{\ell}} \sum_{j \in \mathcal{N}} \mathbb{E}\left[\boldsymbol{Z}_{ij}^{\ell}\right], \ \forall i \in \mathcal{N}. \tag{10}$$

Solving for $\boldsymbol{d}_i^{\ell}$, we have

$$\boldsymbol{d}_i^{\ell} = |\mathcal{E}^{\ell}| - \sqrt{|\mathcal{E}^{\ell}|^2 - 2|\mathcal{E}^{\ell}| \sum_{j \in \mathcal{N}} \mathbb{E}\left[\boldsymbol{Z}_{ij}^{\ell}\right] + \sum_{j \in \mathcal{N}} \mathbb{E}\left[\boldsymbol{Z}_{ij}^{\ell}\right]}. \tag{11}$$

Note that Equation (10) has two solutions to $\boldsymbol{d}_i^{\ell}$, but the other solution is discarded because it is greater than the number of links in layer $\ell$. For $\boldsymbol{d}_i^{\ell}$ to have a real-valued solution, it is easy to show that the following condition should be satisfied

$$\left(|\mathcal{E}^{\ell}| - \sum_{j \in \mathcal{N}} \mathbb{E}\left[\boldsymbol{Z}_{ij}^{\ell}\right]\right)^2 \geq \left(\sum_{j \in \mathcal{N}} \mathbb{E}\left[\boldsymbol{Z}_{ij}^{\ell}\right]\right)^2 - \sum_{j \in \mathcal{N}} \mathbb{E}\left[\boldsymbol{Z}_{ij}^{\ell}\right], \ \forall i \in \mathcal{N}, \tag{12}$$

Since $|\mathcal{E}^{\ell}| - \sum_{j \in \mathcal{N}} \mathbb{E}\left[\boldsymbol{Z}_{ij}^{\ell}\right] \geq 0$, we have

$$\frac{|\mathcal{E}^{\ell}|}{\sum_{j \in \mathcal{N}} \mathbb{E}\left[\boldsymbol{Z}_{ij}^{\ell}\right]} \geq \sqrt{\frac{\sum_{j \in \mathcal{N}} \mathbb{E}\left[\boldsymbol{Z}_{ij}^{\ell}\right] - 1}{\sum_{j \in \mathcal{N}} \mathbb{E}\left[\boldsymbol{Z}_{ij}^{\ell}\right]}} + 1, \ \forall i \in \mathcal{N}, \tag{13}$$

which indicates that a layer should not have giant hub nodes that are incident with over 50% of the edges of a layer. This condition is typically satisfied for sparse multiplex networks. For layers with a sufficiently large number of edges but no giant hub node, $|\mathcal{E}^{\ell}| \gg \boldsymbol{d}_i^{\ell}$, thus the solution to Equation (10) can be approximately given by

$$\boldsymbol{d}_i^\ell \approx \sum_{j\in\mathcal{N}} \mathbb{E}\left[\boldsymbol{Z}_{ij}^\ell\right] = \sum_{j\in\mathcal{N}} p_{ij}^\ell,\ \forall i \in \mathcal{N}. \tag{14}$$

In the A-step, the aggregate topology from topologies of all layers in the multiplex network is obtained using the OR aggregation mechanism [22]. Specifically, the respective entries in the aggregate adjacency matrix should be equal to those calculated from the observed part of the adjacency matrix for each layer, which is expressed as

$$\boldsymbol{A}_{ij} = \varphi(\boldsymbol{Z}_{ij}^1, \ldots, \boldsymbol{Z}_{ij}^{|\mathcal{L}|}) = 1 - \prod_{\ell\in\mathcal{L}} (1 - \boldsymbol{Z}_{ij}^\ell) \tag{15}$$

where $\varphi(\cdot)$ is the aggregation function. Apparently, $\boldsymbol{A}_{ij} = 1$ when there exists at least one $\ell$ such that $\boldsymbol{Z}_{ij}^\ell = 1$ and $\boldsymbol{A}_{ij} = 0$ only when $\boldsymbol{Z}_{ij}^\ell = 0$ for all $\ell$. The aggregate topology (where nodes are connected) is used to update the estimated complete topology of each layer as follows. According to the Bayes' Theorem, we have

$$P(\boldsymbol{Z}_{ij}^\ell = 1 \mid \boldsymbol{A}_{ij} = 1) = \frac{P(\boldsymbol{A}_{ij} = 1 \mid \boldsymbol{Z}_{ij}^\ell = 1) \cdot P(\boldsymbol{Z}_{ij}^\ell = 1)}{P(\boldsymbol{A}_{ij} = 1)} \tag{16a}$$

$$= \frac{1 \cdot p_{ij}^\ell}{1 - \prod\limits_{h\in\mathcal{L}} (1 - p_{ij}^h)} \tag{16b}$$

In the implementation, this means that the latest $p_{ij}^\ell$ for all $i, j$, which is equivalent to $P(\boldsymbol{Z}_{ij}^\ell = 1 \mid \boldsymbol{A}_{ij} = 1)$ at this point, is updated according to Equation (16b).

The complete EMA algorithm is summarized in Algorithm 1. Note that we can treat the aggregate topology as part of the parameters for generating the multiplex network, i.e., $\Theta = (\boldsymbol{A}, \boldsymbol{d})$. From this perspective, the EMA framework is reduced to an EM framework and the A-step is regarded as a component of the M-step for estimating the optimal model parameters. This is because given the current estimate of the unobserved part of the multiplex network, the aggregate topology obtained by Equation (15) is the only solution and hence the optimal solution to the aggregate topology.

**Algorithm 1:** EMA for multiplex network reconstruction

**Input:** Error tolerance $\epsilon_T$, maximum number of iterations $Iter_{\max}$, number of unique nodes in the multiplex network $|\mathcal{N}|$, and partially observed topology $\boldsymbol{X}$.
**Output:** Reconstructed complete topology $\boldsymbol{Z}$ for the multiplex network.

```
Initialize the degree sequences and predicted adjacency matrices: d_i^ℓ ~ U(1, |N|), ∀i ∈ {1,...,|N|} and Z_ij^ℓ ~ U(0,1), ∀i ∈ {1,...,|N|}.
for iter in 1 to Iter_max do
    E-step: Update q(Z) according to Equation (6).
    A-step: Update Z using the aggregate topology according to Equation (16).
    M-step: Update d according to Equation (14).
    Calculate the error ε
    if ε > ε_T then Continue
    else Break
    end if
end for
```

## 4. Datasets

This section presents the details about the real-world multiplex network we use in validating the EMA framework. Covert multiplex networks include the drug trafficking multiplex network and the Sicilian Mafia multiplex network. To demonstrate that the EMA framework is general, we validate it against the other two models on overt multiplex networks, including the London transportation multiplex network and the *C. elegans* neural

multiplex network. This set of networks provide an opportunity to test the EMA framework under varying fractions of observed components of multiplex networks and permit the comparison between social networks, biological, and physical networks.

### *4.1. Drug Trafficking Multiplex Network*

The dataset on drug trafficking networks in this study was obtained from the 2007 'E' Division Provincial Threat Assessment (PTA) report generated by Criminal Intelligence Service Canada and the Royal Canadian Mounted Police (RCMP) [39]. This report contains information about all individuals associated with the criminal operations of 129 different groups. This extensive data collection effort maps a community of drug trafficking operations, revealing how individuals interact within and between groups to sustain regional illicit drug supply and distribution. Individual-level information includes demographic characteristics, a description of three years of drug trafficking activity, their role in drug trafficking, and their current and historical relationships among each other, including co-offenders (individuals who commit crimes together), legitimate business partners (including friends), their kinship, and involvement in a formal criminal organization (including enemies) (Figure 2). This multiplex network integrates information obtained from official records management systems, surveillance, and non-criminal justice information sources. The layers for co-offenders and legitimate business partners are used to build a two-layer multiplex network (representing surveillance data capturing current activity, criminal or otherwise), the layers for co-offenders, legitimate business partners, and formal criminal organization are used to build a three-layer multiplex network (adding social embedding in a community of criminal actors), and the four-layer multiplex network adds familial relations (adding trusted relations that bolster network resiliency). It is obvious from the figure that both layers do not have giant hub nodes that are incident with over 50% of the edges of a layer. Some key network statistics of each of the four layers are presented in Table 1. We can see from the table that each layer has a low link density ($\rho$), low average degree ($\langle d \rangle$), and the average size of the connected component ($\langle |CC| \rangle$). Although the size of the greatest connected component is not small, particularly the one in the co-offender network, the coefficient of variance (CoV, i.e., ratio of standard deviation and mean) of the size of the connected component is small, indicating that members in the drug trafficking community are generally isolated.

**Table 1.** Network statistics of each layer in the drug trafficking multiplex network. FCO is the shorthand for formal criminal organization.

| **Layer** | $\lvert\mathcal{N}\rvert$ | $\lvert\mathcal{E}\rvert$ | $\rho(\times 10^{-3})$ | $\langle d \rangle$ | $\langle \lvert CC \rvert \rangle$ | $\lvert GCC \rvert$ | **CoV of** $\lvert CC \rvert$ |
|---|---|---|---|---|---|---|---|
| Co-offender | 1645 | 1808 | 1.32 | 2.19 | 19.00 | 1024 | 5.71 |
| Legitimate | 1022 | 1041 | 2.00 | 2.04 | 10.99 | 462 | 4.43 |
| FCO | 560 | 597 | 3.81 | 2.13 | 14.36 | 150 | 2.24 |
| Kinship | 399 | 308 | 3.88 | 1.54 | 3.22 | 25 | 0.99 |

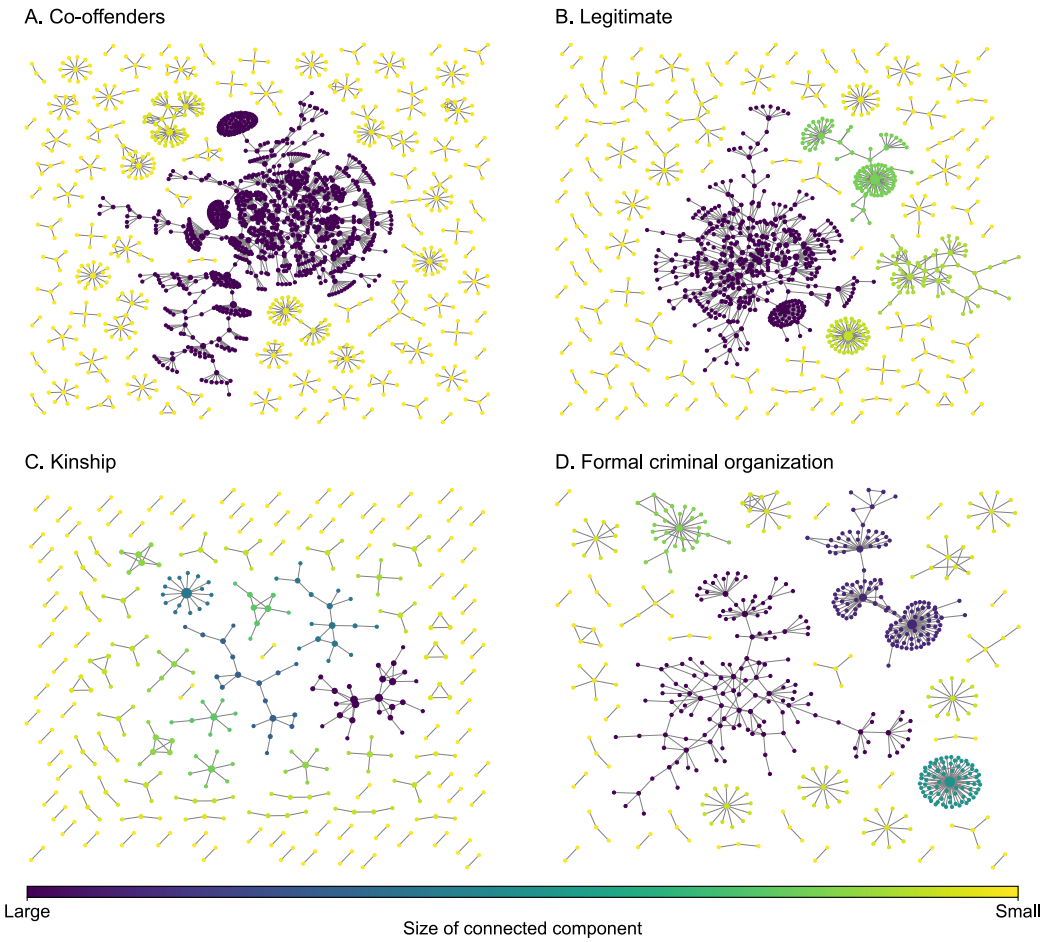


**Figure 2.** Layers of the drug trafficking multiplex network. The size of a node is proportional to its degree.

### *4.2. Sicilian Mafia Multiplex Network*

The Sicilian Mafia multiplex network (Figure 3) was derived from the pre-trial detention order issued by the local court in 2007 in response to a major investigation [44]. The Italian civil law system is substantively different than federalist case law systems (such as the U.S. system), and as such, the court documents filed in support of the detention order provide sufficient information to map individual-level involvement in criminal enterprise activities [15]. Network information captures five years of observations (2003 to 2007) for two Mafia clans, known as the Mistretta family and the Batanesi clan. These familial clans are embedded within a community of illicit entrepreneurs collectively referred to as the Sicilian Mafia. This data set includes two types of layers: (i) meeting data among suspected (101 nodes and 256 edges) and (ii) phone calls among individuals (100 nodes and 124 edges). These layers represent different modes of communication commonly observed during surveillance activity in support of major investigations. Some key statistics of each layer are presented in Table 2. It can be observed from the figure that both layers do not have giant hub nodes as well.

**Table 2.** Network statistics of each layer in the Sicilian Mafia multiplex network

| **Layer** | $\lvert\mathcal{N}\rvert$ | $\lvert\mathcal{E}\rvert$ | $\rho(\times 10^{-3})$ | $\langle d\rangle$ | $\langle\lvert CC\rvert\rangle$ | $\lvert GCC\rvert$ | **CoV of** $\lvert CC\rvert$ |
|---|---|---|---|---|---|---|---|
| Meeting | 101 | 256 | 55.77 | 5.07 | 19.02 | 86 | 1.76 |
| Call | 100 | 124 | 27.45 | 2.48 | 23.50 | 85 | 1.51 |

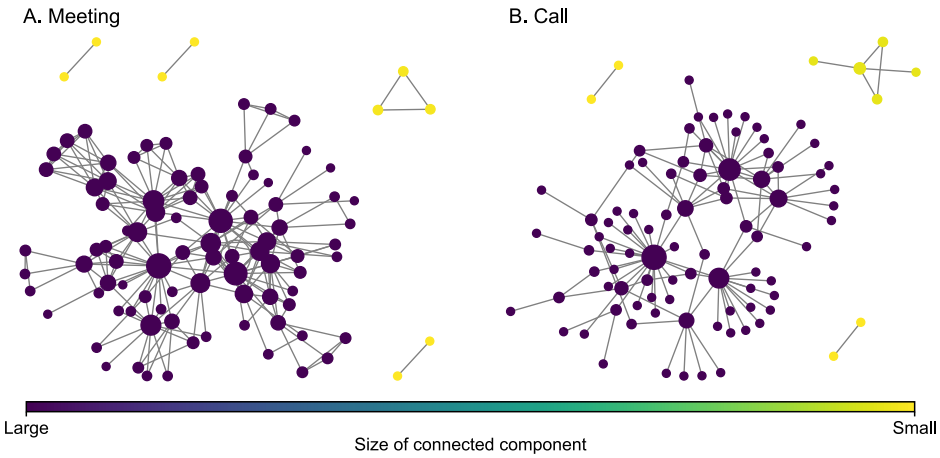


**Figure 3.** Layers of the Sicilian Mafia multiplex network. The size of a node is proportional to its degree.

### *4.3. London Transportation Multiplex Network*

The London transportation multiplex network [45] (Figure 4) consists of underground (260 nodes and 225 edges), overground (81 nodes and 62 edges), and lightrail (44 nodes and 34 edges). The overground and light rail layers are selected to construct a two-layer multiplex network. The third layer adds the underground edges. While the modalities (layers) of transit systems are designed to intersect forming a fully connected giant component, single layers will not always form a connected component. Moreover, this network is intentionally incomplete in that buses that may link different modalities (e.g., underground to light rail) were not included. Some key statistics of each layer are presented in Table 3. Each layer of this transportation multiplex network has many connected components and the size of the greatest connected component is not very large compared to the number of nodes in each layer.

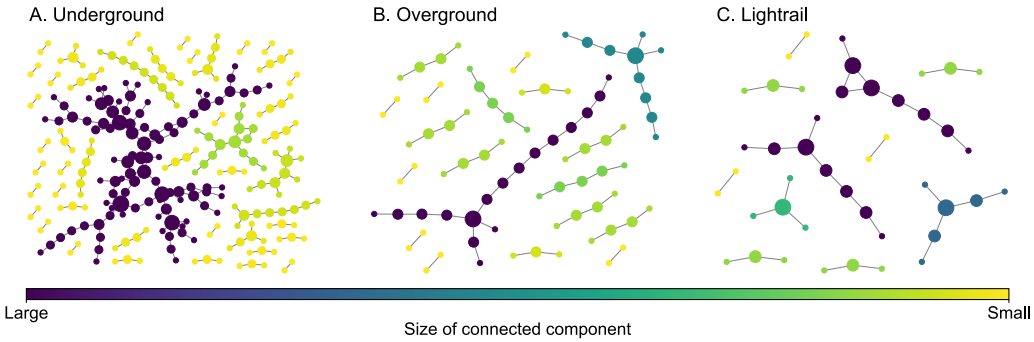


**Figure 4.** Layers of the London transportation multiplex network. The size of a node is proportional to its degree.

**Table 3.** Network statistics of each layer in the London transportation multiplex network.

| **Layer** | $\|\mathcal{N}\|$ | $\|\mathcal{E}\|$ | $\rho(\times 10^{-3})$ | $\langle d \rangle$ | $\langle \|CC\| \rangle$ | $\|GCC\|$ | **CoV of** $\|CC\|$ |
|---|---|---|---|---|---|---|---|
| Underground | 260 | 225 | 6.68 | 1.73 | 5.65 | 98 | 2.48 |
| Overground | 81 | 62 | 19.14 | 1.53 | 4.26 | 17 | 0.83 |
| Lightrail | 44 | 34 | 35.94 | 1.55 | 4.00 | 8 | 0.54 |

### *4.4. C. elegans Multiplex Network*

The *C. (Caenorhabditis) elegans* neural multiplex network [46] (Figure 5) consists of three layers for different types of synaptic junctions, including electric (242 nodes and 450 edges), chemical (260 nodes and 869 edges), and polyadic (277 nodes and 1666 edges). Layers for

chemical and polyadic junctions are selected to construct a two-layer multiplex network. Some key statistics of each layer are presented in Table 4. Compared to the covert multiplex networks introduced earlier, all layers of the *C. elegans* multiplex network are much denser because each layer has a higher density. In each layer, the size of the greatest connected component is large. In particular, all the nodes in the polyadic layer are connected because the size of the greatest connected component is equal to the number of nodes.

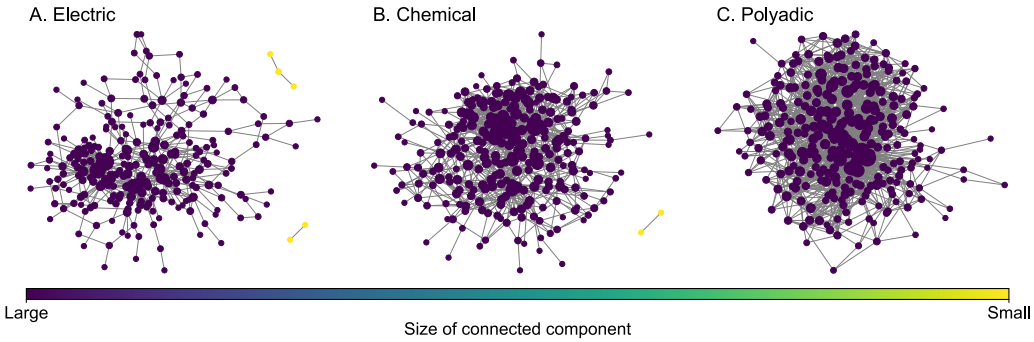


**Figure 5.** Layers of the *C. elegans* multiplex network. The size of a node is proportional to its degree.

**Table 4.** Network statistics of each layer in the *C. elegans* multiplex network.

| **Layer** | $\|\mathcal{N}\|$ | $\|\mathcal{E}\|$ | $\rho(\times 10^{-3})$ | $\langle d \rangle$ | $\langle \|CC\| \rangle$ | $\|GCC\|$ | **CoV of** $\|CC\|$ |
|---|---|---|---|---|---|---|---|
| Electrical | 242 | 450 | 15.43 | 3.72 | 80.67 | 237 | 1.37 |
| Chemical | 260 | 869 | 25.81 | 6.68 | 130.00 | 258 | 0.98 |
| Polyadic | 277 | 1666 | 43.58 | 12.03 | 277.00 | 277 | 0.00 |

## 5. Experiments

### *5.1. Experiment Setup*

For each multiplex network, we use 50 random simulations at each fraction of missing nodes to obtain the average predictive performance (we find that 50 repetitions can already achieve a stable mean value). When a node is not observed, the links connected to this node are also considered unobserved. For both EMA and EM, the error tolerance $\epsilon_T$ is set to $1 \times 10^{-5}$, i.e., the solution is considered converged if the difference in the mean absolute error (MAE) of the predicted adjacency matrices (unobserved entries only) at two consecutive iterations for all unobserved links is lower than the tolerance.

Because covert networks are sparse, the associated datasets are typically imbalanced toward negatives (no links between nodes), we employ the Matthews Correlation Coefficient (MCC) and G-mean to measure the prediction (classification) accuracy [47,48]. MCC can be computed as

$$\text{MCC} = \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP}+\text{FP}) \times (\text{TP}+\text{FN}) \times (\text{TN}+\text{FP}) \times (\text{TN}+\text{FN})}} \tag{17}$$

where TP represents true positives, TN represents true negatives, FP represents false positives, and FN represents false negatives. Note that the value of MCC can be negative. G-mean is given by

$$\text{G-mean} = \sqrt{\frac{\text{TP}}{\text{TP}+\text{FN}} \times \frac{\text{TN}}{\text{TN}+\text{FP}}} = \sqrt{\text{Recall} \times \text{Specificity}} \tag{18}$$

We also include the random model (RM), i.e., a binary uniformly random classifier, as a baseline model. Among all possible links, except for the links among observed nodes, the RM selects a number (i.e., the number of links left) of links in a uniformly random manner as the predicted links.

### 5.2. Results

We first check the performance of EMA under different fractions of missing components. The predictive performances measured by G-mean and MCC for different multiplex networks under 10% to 90% of missing components are presented in Figures 6–9 where $c$ represents the fraction of observed components. We can see that EMA achieves the best MCC regardless of the multiplex networks and the fraction of missing components. In terms of G-mean, EMA outperforms EM in all cases and random model when the observed fraction of components is large. Note that the trends of G-mean and MCC over different fractions of observed components obtained by EMA and EM are close to each other because the values of FP $\times$ FN are much smaller than TP $\times$ TN.

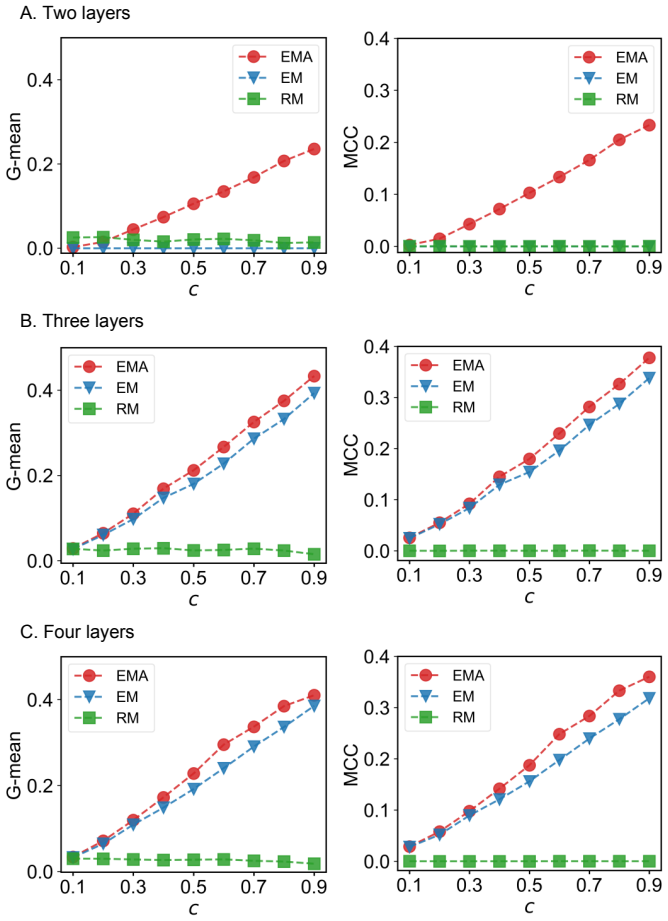


**Figure 6.** G-mean and MCC for drug trafficking networks.

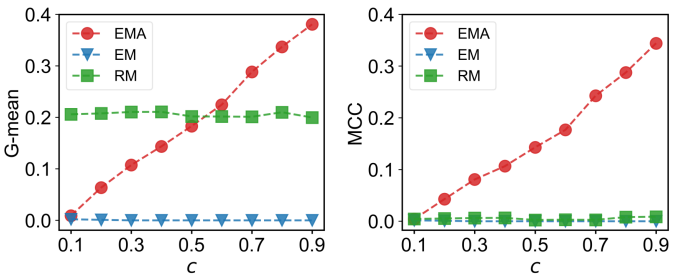


**Figure 7.** G-mean and MCC for Sicilian Mafia networks.

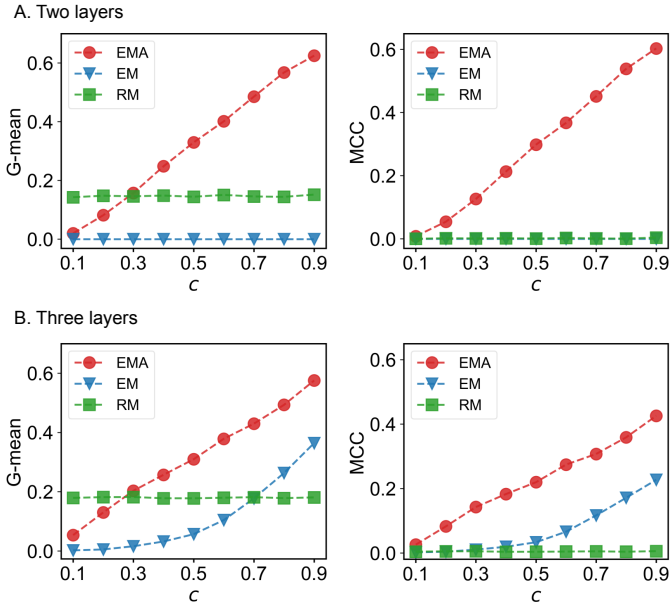


**Figure 8.** G-mean and MCC for *C. elegans* neural networks.

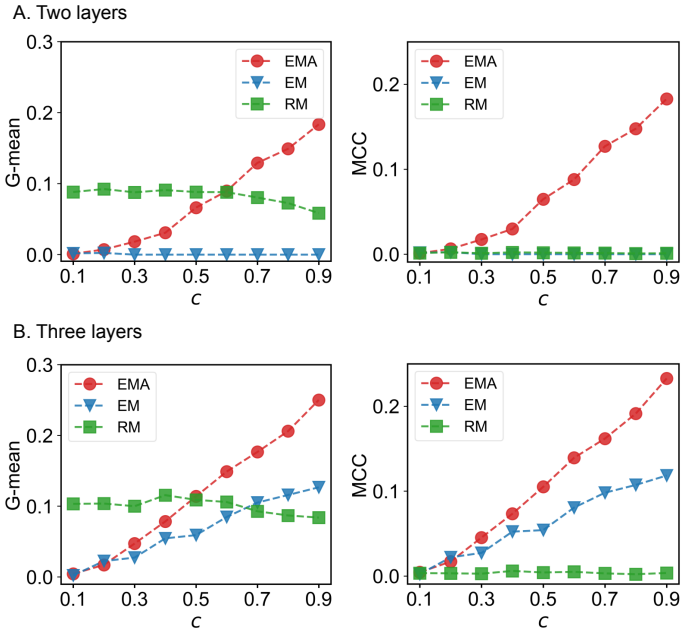


**Figure 9.** G-mean and MCC for London transportation networks.

Next, we examine the impact of the number of layers on the EMA and EM approaches. As the number of layers increases from two to three in drug trafficking (Figure 6), *C. elegans* (Figure 8), and London transport networks (Figure 9), the performance improvement of EMA when compared to EM in both G-mean and MCC decreases, indicating that the benefits of adding the aggregation step drops. This is because when there are more layers, there are more choices of values for individual entries in the layers, making a positive entry in the aggregate topology less informative about the positive value of each layer.

We also compare the convergence of EMA and EM algorithms on the drug trafficking multiplex networks (Figure 10) and London transportation multiplex networks (Figure 11). Regardless of the size of multiplex networks and the fraction of components observed, EM converges more quickly than EMA. In the worst-case tested, EMA can reach convergence within 40 iterations given a convergence error of $1 \times 10^{-5}$. However, if the convergence error can be set to a larger value, such as $1 \times 10^{-4}$, then EMA can achieve convergence much

faster (within 10 iterations across the networks considered). Therefore, both algorithms can be applied to moderately large-scale multiplex networks.

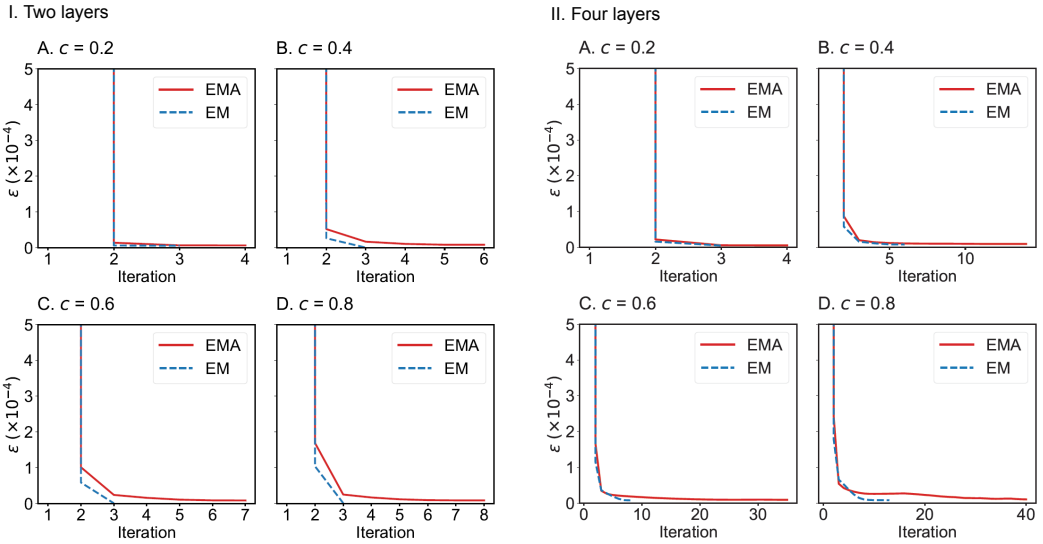


**Figure 10.** Convergence of EMA and EM on the two-layer (**I**) and four-layer (**II**) drug trafficking multiplex networks under different fractions of observed components: (**A**) 0.2; (**B**) 0.4; (**C**) 0.6; (**D**) 0.8.

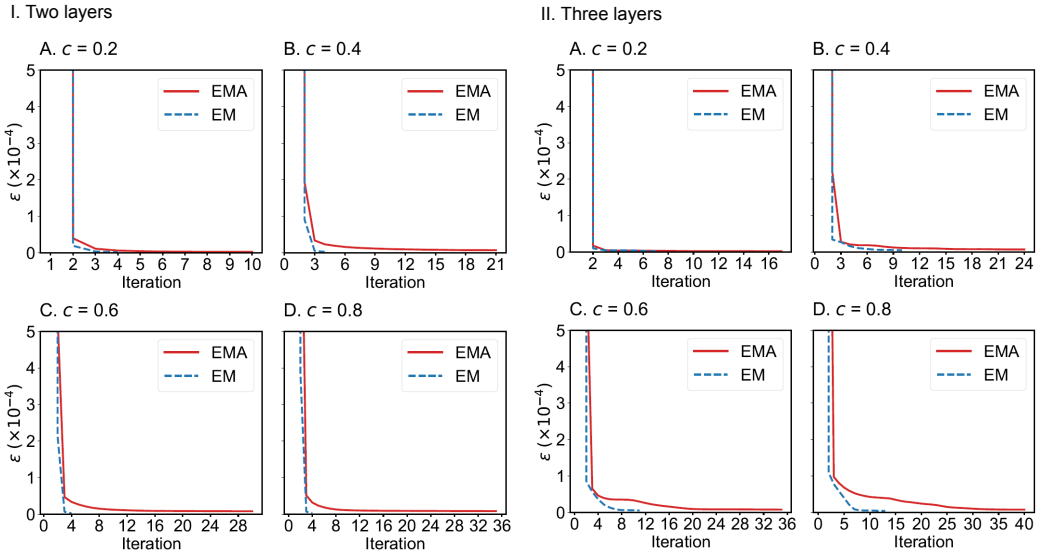


**Figure 11.** Convergence of EMA and EM on the two-layer (**I**) and three-layer (**II**) London transportation multiplex networks under different fractions of observed components: (**A**) 0.2; (**B**) 0.4; (**C**) 0.6; (**D**) 0.8.

## 6. Discussion

From a criminal justice perspective, mapping networks using available data on illicit activities is typically not sufficient to obtain the complete structure of the covert networks under study. As a result, networks of criminal operations are always incomplete, thereby compromising the potential practical contributions network science can make toward developing effective interdiction strategies. Two major implications for Disrupting Illicit Networks arise from this investigation for reconstructing a more complete structure of multiplex networks.

First, with regard to the broad context of disrupting covert networks, the experiments support arguments that criminal enterprise is intimately embedded within a complex social system wherein a pair of individuals can be connected by multiple relations or a

criminal actor could tap into different social networks to resolve a pressing issue [35,49,50]. For example, a drug smuggler uses legitimate business assets (helicopter and employed pilot) to transport illicit products (multiple relations) and, in a pinch, the smuggler could launder proceeds through a relative's real-estate business (temporary use of a different type of relation for criminal ends). The extent to which associates and kin are aware of these illegalities is not necessarily relevant; what is critical is that this interwoven social fabric must be fully understood before effective interdiction strategies can be developed and implemented. Results on the performance of the EMA framework for the drug trafficking network and the Sicilian mafia network show that networks generated from the surveillance of current criminal activity are insufficient. Interpolating latent connections by adding at least one more layer that captures historic criminogenic relations (i.e., formal organized crime membership), business or professional associations, or purely social and familial connections significantly improves network completeness. This finding coincides with applied network criminology [34,39] and crime opportunity theory [51,52].

Second, while inference can improve network completeness, such modeling frameworks do not constitute a silver bullet for all methodological issues. Continued efforts are needed to improve data collection protocols to enhance the capacity to generate more complete information about covert operations and the complex social systems within which they are embedded. To this aim, our experiments show that identifying additional layers composing a multiplex network makes the EMA algorithm more accurate. Supporting prior arguments to establish methodological conventions regarding the use of criminal justice system data in network applications [53,54], this finding confirms that the key to uncovering more of the covert multiplex network is to tap different layers of information sources [34,51]. However, additional tests are needed to understand the nature of information benefits. For example, how does the dependency between layers, particularly in social networks, impact information benefits? When does identifying additional layers of covert networks benefit more than identifying more components within known layers? Finally, given that the covert networks were mapped from observations made over many years, does the model offer more information benefits when used with shorter observations, such as six months?

## 7. Conclusions

In this study, we develop the EMA framework for reconstructing multiplex networks and validate this approach against the EM approach and the random model on several multiplex networks given different fractions of the observed part of the complete networks. We find that EMA achieves the best predictive performance in terms of G-mean and MCC over different levels of available observed parts of the true networks. The findings are consistent across experiments conducted on different types of multiplex networks (social and non-social). This indicates that integrating the estimation of the aggregate topology, which encodes the interlayer dependency of multiplex networks, can improve the reconstruction accuracy. By inferring a more complete structure of covert multiplex networks, the EMA framework can support interdiction efforts aimed at disrupting criminal operations, especially when the observations are only available for two or three types of interactions. As the fraction of observed actors and the type of interactions among actors increases, the added benefits of the EMA framework when compared to the EM framework decrease. In cases where quick decisions are required for interdicting multiplex with many layers, the EM framework is recommended due to a lower computational load.

Future work will consider the use of graph neural networks to reconstruct sparse multiplex networks. Although GNN has been widely applied to predict missing links [23,55,56], research efforts that apply GNN to reconstruct multiplex networks are limited. As such, it is valuable to compare the reconstruction accuracy of EMA and GNN on multiple datasets and identify where each approach performs the best. Furthermore, since the accuracy of the EMA framework reduces for multiplex networks with many layers, future research can explore how to improve the predictive accuracy using network generative models

with a prescribed level of interlayer dependency [57–60], and use the dynamics of the networks [61].

**Author Contributions:** Conceptualization, J.-Z.Y. and J.G.; methodology, J.-Z.Y., M.W. and J.G.; software, J.-Z.Y.; formal analysis, J.-Z.Y., M.W., G.B. and J.G.; data curation, J.-Z.Y., M.W. and G.B.; writing—original draft preparation, J.-Z.Y.; writing—review and editing, J.-Z.Y., M.W., G.B., F.A.-V. and J.G.; supervision, J.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the USA National Science Foundation under Grant No. 2047488, and the Rensselaer-IBM AI Research Collaboration. M.W. is partially supported by the National Natural Science Foundation of China under Grant No. 62203388.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The code and datasets used for the numerical experiments are available at https://github.com/jinzhuyu/multiplex_recon (accessed on 9 January 2023). The *C. elegans* neural networks dataset is originally obtained from https://comunelab.fbk.eu/data.php (accessed on 1 September 2022). The London transportation network dataset is originally obtained from [45]. The Sicilian Mafia networks dataset is originally obtained from https://zenodo.org/record/3938818#.Y2azsnbMKUl (accessed on 1 September 2022). The drug trafficking networks dataset is originally obtained from [39].

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## Notations

**Sets & Indices**

| | |
|---|---|
| $i$, $j$ | Indices of nodes |
| $\ell$ | Index of layers |
| $\mathcal{L}$ | Set of layers |
| $\mathcal{N}$, $\mathcal{E}$ | Sets of nodes and links |
| $\mathcal{V}_{obs}$ | Set of observed nodes |
| $\boldsymbol{X}$ | Set of the observed topologies (by layer) of a multiplex network |
| $\boldsymbol{Z}$ | Set of complete topologies (by layer) of a multiplex network |

**Variables & Parameters**

| | |
|---|---|
| $c$ | Fraction of observed nodes in a layer |
| $\boldsymbol{d}$ | Degree sequences |
| $\langle d \rangle$ | Average degree |
| $m$ | Number of layers |
| $U$, $V$ | Generic random variables |
| $\boldsymbol{X}^{\ell}$ | Observed adjacency matrix (subnetwork) for layer $\ell$ |
| $\boldsymbol{Z}^{\ell}$ | Adjacency matrix (complete topology) for layer $\ell$ |
| $\boldsymbol{A}$ | Adjacency matrix for the aggregate topology of a multiplex network |
| $\lvert CC \rvert$, $\lvert GCC \rvert$ | Size of a connected component and size of the greatest connected component |
| $\langle \lvert CC \rvert \rangle$ | Average size of connected components |
| $\epsilon$, $\epsilon_T$ | Convergence error and error tolerance |
| $p_{ij}$ | Probability of a link between node $i$ and node $j$. |
| $q(\boldsymbol{Z})$ | Posterior distribution for the complete multiplex network |
| $\rho$ | Link density |
| $\boldsymbol{\Theta}$ | Model parameters |

## References

1. Gao, J.; Bashan, A.; Shekhtman, L.; Havlin, S. *Introduction to Networks of Networks*; IOP Publishing Ltd.: Bristol, UK, 2022.
2. Zhu, X.; Ma, J.; Su, X.; Tian, H.; Wang, W.; Cai, S. Information spreading on weighted multiplex social network. *Complexity* **2019**, *2019*, 5920187. [CrossRef]
3. Solé-Ribalta, A.; Gómez, S.; Arenas, A. Congestion induced by the structure of multiplex networks. *Phys. Rev. Lett.* **2016**, *116*, 108701. [CrossRef]

4. Kim, M.K.; Narasimhan, R. Exploring the multiplex architecture of supply networks. *Int. J. Supply Chain Manag.* **2019**, *8*, 45–64.
5. Aleta, A.; Tuninetti, M.; Paolotti, D.; Moreno, Y.; Starnini, M. Link prediction in multiplex networks via triadic closure. *Phys. Rev. Res.* **2020**, *2*, 042029. [CrossRef]
6. Tran, C.; Shin, W.Y.; Spitz, A.; Gertz, M. DeepNC: Deep generative network completion. *arXiv* **2019**, arXiv:1907.07381.
7. Gao, J.; Barzel, B.; Barabási, A.L. Universal resilience patterns in complex networks. *Nature* **2016**, *530*, 307–312. [CrossRef]
8. Liu, X.; Li, D.; Ma, M.; Szymanski, B.K.; Stanley, H.E.; Gao, J. Network resilience. *Phys. Rep.* **2022**, *971*, 1–108.
9. Xu, J.; Chen, H. The topology of dark networks. *Commun. ACM* **2008**, *51*, 58–65. [CrossRef]
10. Hosseinkhani, J.; Chuprat, S.; Taherdoost, H. Discovering criminal networks by Web structure mining. In Proceedings of the 2012 7th International Conference on Computing and Convergence Technology (ICCCT), Seoul, Republic of Korea, 3–5 December 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 1074–1079.
11. Pourhabibi, T.; Ong, K.L.; Kam, B.H.; Boo, Y.L. DarkNetExplorer (DNE): Exploring dark multi-layer networks beyond the resolution limit. *Decis. Support Syst.* **2021**, *146*, 113537. [CrossRef]
12. Martínez, V.; Berzal, F.; Cubero, J.C. A survey of link prediction in complex networks. *ACM Comput. Surv. (CSUR)* **2016**, *49*, 1–33. [CrossRef]
13. Abdolhosseini-Qomi, A.M.; Jafari, S.H.; Taghizadeh, A.; Yazdani, N.; Asadpour, M.; Rahgozar, M. Link prediction in real-world multiplex networks via layer reconstruction method. *R. Soc. Open Sci.* **2020**, *7*, 191928. [CrossRef]
14. Berlusconi, G.; Calderoni, F.; Parolini, N.; Verani, M.; Piccardi, C. Link prediction in criminal networks: A tool for criminal intelligence analysis. *PLoS ONE* **2016**, *11*, e0154244.
15. Calderoni, F.; Catanese, S.; De Meo, P.; Ficara, A.; Fiumara, G. Robust link prediction in criminal networks: A case study of the Sicilian Mafia. *Expert Syst. Appl.* **2020**, *161*, 113666. [CrossRef]
16. Lü, L.; Zhou, T. Link prediction in complex networks: A survey. *Phys. A Stat. Mech. Its Appl.* **2011**, *390*, 1150–1170. [CrossRef]
17. Herlau, T.; Schmidt, M.N.; Mørup, M. Infinite-degree-corrected stochastic block model. *Phys. Rev. E* **2014**, *90*, 032819. [CrossRef]
18. Williamson, S.A. Nonparametric network models for link prediction. *J. Mach. Learn. Res.* **2016**, *17*, 7102–7121.
19. Stanley, N.; Bonacci, T.; Kwitt, R.; Niethammer, M.; Mucha, P.J. Stochastic block models with multiple continuous attributes. *Appl. Netw. Sci.* **2019**, *4*, 1–22. [CrossRef]
20. Wu, M.; Li, Z.; Shao, C.; He, S. Quantifying multiple social relationships based on a multiplex stochastic block model. *Front. Inf. Technol. Electron. Eng.* **2021**, *22*, 1458–1462. [CrossRef]
21. Kim, M.; Leskovec, J. The network completion problem: Inferring missing nodes and edges in networks. In Proceedings of the 2011 SIAM International Conference on Data Mining, Mesa, AZ, USA, 28–30 April 2011; SIAM: Philadelphia, PA, USA, 2011; pp. 47–58.
22. Wu, M.; Chen, J.; He, S.; Sun, Y.; Havlin, S.; Gao, J. Discrimination reveals reconstructability of multiplex networks from partial observations. *Commun. Phys.* **2022**, *5*, 163. [CrossRef] [PubMed]
23. Zhang, M.; Chen, Y. Link prediction based on graph neural networks. *Adv. Neural Inf. Process. Syst.* **2018**, *31*. Available online: https://proceedings.neurips.cc/paper/2018/hash/53f0d7c537d99b3824f0f99d62ea2428-Abstract.html (accessed on 7 January 2023).
24. Gao, M.; Jiao, P.; Lu, R.; Wu, H.; Wang, Y.; Zhao, Z. Inductive link prediction via interactive learning across relations in multiplex networks. *IEEE Trans. Comput. Soc. Syst.* **2022**, *Early Access*. [CrossRef]
25. Baycik, N.O.; Sharkey, T.C.; Rainwater, C.E. Interdicting layered physical and information flow networks. *IISE Trans.* **2018**, *50*, 316–331. [CrossRef]
26. Kosmas, D.; Sharkey, T.C.; Mitchell, J.E.; Maass, K.L.; Martin, L. Interdicting restructuring networks with applications in illicit trafficking. *arXiv* **2020**, arXiv:2011.07093.
27. Shen, Y.; Sharkey, T.C.; Szymanski, B.K.; Wallace, W.A. Interdicting interdependent contraband smuggling, money and money laundering networks. *Socio-Econ. Plan. Sci.* **2021**, *78*, 101068. [CrossRef]
28. Ficara, A.; Fiumara, G.; De Meo, P.; Catanese, S. Multilayer network analysis: The identification of key actors in a Sicilian Mafia operation. In *International Conference on Future Access Enablers of Ubiquitous and Intelligent Infrastructures*; Springer: Cham, Switzerland, 2021; pp. 120–134.
29. Traganitis, P.A.; Shen, Y.; Giannakis, G.B. Topology inference of multilayer networks. In Proceedings of the 2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Atlanta, GA, USA, 1–4 May 2017; pp. 898–903.
30. Chen, C.; Tong, H.; Xie, L.; Ying, L.; He, Q. Cross-dependency inference in multi-layered networks: A collaborative filtering perspective. *ACM Trans. Knowl. Discov. Data (TKDD)* **2017**, *11*, 1–26. [CrossRef] [PubMed]
31. Wang, Y.; Aggarwal, C.; Derr, T. Distance-wise prototypical graph neural network in node imbalance classification. *arXiv* **2021**, arXiv:2110.12035.
32. Liben-Nowell, D.; Kleinberg, J. The link prediction problem for social networks. In Proceedings of the Twelfth International Conference on Information and Knowledge Management, Orleans, LA, USA, 3–8 November 2003; pp. 556–559.
33. Agreste, S.; Catanese, S.; De Meo, P.; Ferrara, E.; Fiumara, G. Network structure and resilience of Mafia syndicates. *Inf. Sci.* **2016**, *351*, 30–47. [CrossRef]
34. Smith, C.M.; Papachristos, A.V. Trust thy crooked neighbor: Multiplexity in Chicago organized crime networks. *Am. Sociol. Rev.* **2016**, *81*, 644–667. [CrossRef]

35. Malm, A.; Bichler, G. Using friends for money: The positional importance of money-launderers in organized crime. *Trends Organ. Crime* **2013**, *16*, 365–381. [CrossRef]
36. Kleemans, E.R.; Van de Bunt, H.G. The social embeddedness of organized crime. *Transnatl. Organ. Crime* **1999**, *5*, 19–36.
37. Duijn, P.A.; Kashirin, V.; Sloot, P.M. The relative ineffectiveness of criminal network disruption. *Sci. Rep.* **2014**, *4*, 4238. [CrossRef] [PubMed]
38. Krebs, V. Mapping networks of terrorist cells. *Connections* **2002**, *24*, 43–52.
39. Malm, A.; Bichler, G.; Van De Walle, S. Comparing the ties that bind criminal networks: Is blood thicker than water? *Secur. J.* **2010**, *23*, 52–74. [CrossRef]
40. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B* **1977**, *39*, 1–22.
41. Blei, D.M.; Kucukelbir, A.; McAuliffe, J.D. Variational inference: A review for statisticians. *J. Am. Stat. Assoc.* **2017**, *112*, 859–877. [CrossRef]
42. Bender, E.A.; Canfield, E.R. The asymptotic number of labeled graphs with given degree sequences. *J. Comb. Theory Ser. A* **1978**, *24*, 296–307. [CrossRef]
43. Newman, M. *Networks*; Oxford University Press: Oxford, UK, 2018.
44. Cavallaro, L.; Ficara, A.; De Meo, P.; Fiumara, G.; Catanese, S.; Bagdasar, O.; Song, W.; Liotta, A. Disrupting resilient criminal networks through data analysis: The case of Sicilian Mafia. *PLoS ONE* **2020**, *15*, e0236476. [CrossRef]
45. De Domenico, M.; Solé-Ribalta, A.; Gómez, S.; Arenas, A. Navigability of interconnected networks under random failures. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 8351–8356. [CrossRef]
46. Chen, B.L.; Hall, D.H.; Chklovskii, D.B. Wiring optimization can relate neuronal structure and function. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4723–4728. [CrossRef]
47. Branco, P.; Torgo, L.; Ribeiro, R.P. A survey of predictive modeling on imbalanced domains. *ACM Comput. Surv. (CSUR)* **2016**, *49*, 1–50. [CrossRef]
48. Boughorbel, S.; Jarray, F.; El-Anbari, M. Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. *PLoS ONE* **2017**, *12*, e0177678. [CrossRef] [PubMed]
49. Morselli, C.; Giguere, C. Legitimate strengths in criminal networks. *Crime Law Soc. Chang.* **2006**, *45*, 185–200. [CrossRef]
50. Diviák, T.; Dijkstra, J.K.; Snijders, T.A. Poisonous connections: A case study on a Czech counterfeit alcohol distribution network. *Glob. Crime* **2020**, *21*, 51–73. [CrossRef]
51. Bichler, G. *Understanding Criminal Networks: A Research Guide*; University of California Press: Oakland, CA, USA, 2019.
52. Kleemans, E.R.; De Poot, C.J. Criminal careers in organized crime and social opportunity structure. *Eur. J. Criminol.* **2008**, *5*, 69–98. [CrossRef]
53. Campana, P.; Varese, F. Listening to the wire: Criteria and techniques for the quantitative analysis of phone intercepts. *Trends Organ. Crime* **2012**, *15*, 13–30. [CrossRef]
54. Diviák, T. Key aspects of covert networks data collection: Problems, challenges, and opportunities. *Soc. Netw.* **2022**, *69*, 160–169. [CrossRef]
55. Trouillon, T.; Welbl, J.; Riedel, S.; Gaussier, É.; Bouchard, G. Complex embeddings for simple link prediction. In Proceedings of the International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016; pp. 2071–2080.
56. Wang, Z.; Ren, Z.; He, C.; Zhang, P.; Hu, Y. Robust embedding with multi-level structures for link prediction. In Proceedings of the International Joint Conference on Artificial Intelligence, Macao, China, 10–16 August 2019; pp. 5240–5246.
57. Nicosia, V.; Bianconi, G.; Latora, V.; Barthelemy, M. Growing multiplex networks. *Phys. Rev. Lett.* **2013**, *111*, 058701. [CrossRef]
58. Yu, J.Z.; Baroud, H. Modeling uncertain and dynamic interdependencies of infrastructure systems using stochastic block models. *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part B Mech. Eng.* **2020**, *6*, 020906. [CrossRef]
59. Wang, Y.; Yu, J.Z.; Baroud, H. Generating synthetic systems of interdependent critical infrastructure networks. *IEEE Syst. J.* **2021**, *16*, 3191–3202. [CrossRef]
60. Fügenschuh, M.; Gera, R.; Méndez-Bermúdez, J.A.; Tagarelli, A. Structural and spectral properties of generative models for synthetic multilayer air transportation networks. *PLoS ONE* **2021**, *16*, e0258666. [CrossRef]
61. Jiang, C.; Gao, J.; Magdon-Ismail, M. Inferring degrees from incomplete networks and nonlinear dynamics. In Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20), Yokohama, Japan, 7–15 January 2020; pp. 3307–3313.