An ensemble model to optimize modularity in dynamic bipartite networks

Neelu Chaudhary1 · Hardeo Kumar Thakur1 · Rinky Dwivedi2

Received: 12 August 2021 / Revised: 20 December 2021 / Accepted: 3 January 2022
© The Author(s) under exclusive licence to The Society for Reliability Engineering, Quality and Operations Management (SREQOM), India and The Division of Operation and Maintenance, Lulea University of Technology, Sweden 2022

Abstract Distinct non-random quantitative interactions at diverse timestamps formulate real-world dynamic complex networks. The most frequently used class of methods for discovering communities in dynamic networks is modularity optimization that evaluates the quality of the partition of network nodes into distinct communities. The bipartite networks have bipartite modularity and bipartite modularity optimization respectively. Newman’s modularity is a consistently used algorithm to evaluate modules of unipartite networks yet it is ineffective for assessing the division of bipartite networks with two types of vertices. Many community detection methods suggest bipartite modularity to accommodate this issue. They usually employ information about the existence or lack of interactions between nodes. In quantitative networks, weighted modularity is a potential approach for measuring the quality of community partitions (Lu et al. IEEE, 179–184, 2013). This study offers an ensemble model for detecting one-mode communities and optimizing modularity in dynamic bipartite weighted networks. By using collaborative weighted projection, bipartite networks get projected into two weighted one-mode networks. The results of experiments both on real-world dynamic network data and synthetic data demonstrate that the modularity of the method is significantly greater than that of current techniques and the communities discovered contain vertices of comparable kinds exhibiting the suggested algorithm’s performance is ample.

Keywords Dynamic networks · Bipartite modularity · One-mode projection · Communities · Clustering coefficient

1 Introduction

The specific and inevitable category of networks that separates nodes into two distinct types and permits the edges to get connected through the different types of nodes are called as bipartite networks. For many practical systems these before said networks acts as the intrinsic models such as—specialized fields (Purwitasari et al. 2020), fraud detection (Yan et al. 2018), consumer purchasers (Yoon et al. 2020) etc. In bipartite networks, the community structure reveals not just the extent of nodes relation in one part to the other part of the network moreover clusters the network in groups of similar nodes.

The instance the agglomerative method and their discordant coefficients graph (Girvan and Newman 2002) partitioning techniques and ordered clustering techniques have traditionally been used. Firstly, Newman and Girvan defined modularity, that has become a distinguishing criterion for evaluating partition quality. The partition of edges devised within the community comparing the part of edges established between communities is measure by modularity. To check the impermanence of connections created within two points, the authors developed a model.
that preserves the degree of network nodes. As a high rate of modularity shows a better score, this indicator has become a highly used topic in recent times to maximize through optimization techniques (Newman 2004). The optimal solution to the problem, however, is unwieldy, but numerous approximate algorithms get developed (Servis and Clark 2021; Bu et al. 2013; Blondel et al. 2008). The most effective and high-performance algorithm was the greedy approach initiated by Blondel called the Louvain algorithm (Blondel et al. 2008). There are two steps to the Louvain algorithm. The first step involves optimizing modularity locally in each node’s vicinity. Following this step, the nodes of the same communities get added into super nodes, building a new ground grain network. The process is iteratively carried out in progressively concentrated networks until node movement in the graph increases the score.

1.1 State of ART

The study of a vast number of salient topics is based on network representation. Community formation within a network that reflects the aggregation of entities with hugely distributed edges inside groups but limited connections between them, has been identified in several observed networks (Schaub and Peel 2020; Pesántez-Cabrera and Kalyanaraman 2017). To effectively recognize the community structure, several community identification methods get developed for complex networks (Cherifi et al. 2019; Singh et al. 2020; Zhang et al. 2009). Although community mining on static networks gets explored, it has not been substantial research for dynamic networks (Duan et al. 2009). A successful partition generates optimum values for several criteria and, even under different resolutions, may reveal the hierarchical order of community structure. A collaborative network (Yan et al. 2018) comprises a range of entities (e.g., organizations, persons), that in their operational environment, culture, social capital, and targets, are broadly autonomous, geographically distributed, and heterogeneous but work together to attain compatible goals. The structure, comportment, and changing dynamics of networks of independent entities that cooperate to accomplish similar or amicable objectives are the focus of the field of collaborative networks (Valkokari et al. 2020a). Many of the real-world examples get visualized as collaborative networks like professional virtual committees (Oaktari et al. 2018), business networks (Valkokari et al. 2020b), and scientist networks (Abrahams et al. 2019), etc. For example, in a movie-actor network, two actors who have done a movie together will know each other quite well and are assumed to be connected. It is a somewhat strict definition as many actors know each other but has never done a movie together. But intrinsically, it is not a negative thing. Accurate knowledge is entirely appropriate if it can be implemented consistently. Bipartite or multipartite graphs are the best mode of drafting these networks where nodes of one set are not connected but can have similar interests, ways of working, compositions, etc. To collaborate these networks is a good idea to identify relations among entities of the same set and communities of entities that are not directly connected but exhibit a relationship with each other. More information is available in the collaborative networks comparing the simple networks that they have built, telling us whether entities cooperated. For example, remarkable information is the number of movies each pair of actors has worked together, as well as the number of additional actors who have done one or more movies with these actors. This information is to assess the strength of collaborative relations. Firstly, a movie with multiple actors has the probability that more actors know each other and spend time together. Then actors with more movies together know each other better comparing the actors with whom they have one or two movies. For both scenarios, the strength of the ties matter. This concept was get introduced by Newman (2001) that defined the strength of the tie between entities and projected the bipartite weighted graph as one-mode or unipartite.

(Michael J. Barber 2007) Barber suggested a specific modularity measure for bipartite networks to identify the community structure. Various algorithms get developed for this purpose, like label propagation (Barber and Clark 2009), BRIM (Liu and Murata 2010), etc. The Greedy network algorithm (Blondel et al. 2008) performs in many cases is used by various accepted software modules to address extensive community detection issues. Although, it would seem rare that the existing methods explore the Louvain algorithm for community detection of bipartite networks since challenges are very high. Firstly, when nodes of different sections get combined as super nodes, their identities will not retain, that will be the outcome if the initial stage of the Louvain algorithm is considered directly. For example, to add several consumers to a section of consumers with a related buying behaviour makes sense the addition of both consumers and products in one partition for buying behaviour group looks ambiguous. Then, an improved Louvain method called BiLouvain (Zhou et al. 2018) adaptable to bipartite networks was introduced that results in balanced networks. But due to resolution problems, it is not suitable for some applications having a low modularity score that becomes negligible to address the strength of the connection between nodes inside the community compared to the nodes outside the community. In these cases, one-mode projection (Wu et al. 2017; Zweig and Kaufmann 2011) of bipartite networks
onto one of its node’s sets can condense the information in better-understood networks.

### 1.2 Motivation

All modular techniques like the Louvain method and Girvan-Newman method, have a certain degree of constraint but are still common in analysing major real-life problems. These algorithms are acceptable for modularity optimization in unipartite networks in one way or another but are found insubstantial or restricted in the case of bipartite networks. The detection of one-mode communities in the dynamic bipartite network at different timestamps is also limited. The techniques like the BiLouvain Method improvised the bipartite modularity but it ignores the strength of the relationship i.e., weight of edges between entities of a different set. To capitalize and utilize bipartite networks at best, a new model is proposed in the present research to optimize the modularity and formation of one-mode communities in bipartite networks in response to the challenges faced due to the two-mode nature of networks.

### 1.3 Major contribution

An innovative model is proposed to optimize the modularity and formation of one-mode communities in bipartite networks in response to the challenges faced due to the two-mode nature of networks. This model integrates the one-mode projection approach with modularity analysis. There are two significant steps in our model. Firstly, the bipartite network is projected as a unipartite network using a collaborative method applied on weights of edges, then aggregate the nodes in communities using the Louvain method as in unipartite networks. One-mode or unipartite projection of a bipartite network is the process of building a network between the nodes on one set based on their associations with nodes on the other side. It is a popular way to study bipartite networks as this process makes the network much dense with more information that can be retrieved or analysed. In the proposed generic model, a collaborative projection is performed on a weighted bipartite graph onto one of its node sets. Then used the clustering technique to identify communities hierarchically. Finally, the division with the maximum modularity or the appropriate number of application communities is selected. Three datasets are analysed to assess the proposed model. Out of the three datasets used two are practical datasets whereas one is computer-generated to check the generosity of the applied study. The findings demonstrate that the proposed model detects more dense communities in two-party networks with an improved modularity score of networks.

The rest of the paper is as follows: Sect. 2 consists of preliminaries and definitions required for the study, Sect. 3 covers the framework of the proposed model, which covers the collection of datasets, pre-processing and feature selection of desired data, generation of dynamic bipartite network and its projection in unipartite form. It also covers modularity analysis and evaluation of the model. Section 4 covers the result and discussions. The last section covers the conclusion, limitations, and future work.

### 2 Preliminaries

In this section, we explain specific vital definitions and concepts that will be used throughout the article to help understand the topic better.

1. **Time Series of Graph (Lahiri and Berger-Wolf 2008):** For a given sequence of T graphs $G = \{G_1, G_2, ..., G_T\}$ with $G_t = (V_t, E_t, W_t)$, where $1 \leq t \leq T$, $V_t$ is the vertex-set, $E_t$ is the edge-set and $W_t$ (weight on edges) is the weight-set of the graph at a timestamp $t$. We define $G$ as a time series of graphs which can be transformed into weighted dynamic network.

2. **Bipartite Weighted Network (Sinclair et al. 2004):** Weighted bipartite graph is defined as $G (L, R, W, E)$ where $L$ and $R$ are two distinct sets of vertices such as $L \neq R$ and $E$ is the edge list containing tuples $(l_i, r_j, w_{ij})$, where $w_{ij} \in W$ is the weight between nodes $l_i \in L$ and $r_j \in R$.

3. **Bipartite Clique (Wang et al. 2008):** A bipartite graph $G (L, R, W, E)$ is called complete bipartite graph or biclique if each node $l_i \in L$ are connected to all nodes $r_j \in R$ such as $|G| = n \times m$ where $|L| = n$ and $|R| = m$.

4. **Bipartite Clustering Coefficient:** The bipartite clustering coefficient is a measure of local density of connections defined as Eq. 1 (Latapy et al. 2008):

$$C_u = \frac{\sum_{v \in N(N(u))} C_{uv}}{N(N(u))}$$

where $N(N(u))$ are the second order neighbours of $u$ in graph $G$, excluding $u$, and $C_{uv}$ is the pairwise clustering coefficient between nodes $u$ and $v$.

5. **Collaborative weighted projected graph:**

The collaboration weighted projection is the projection of the bipartite network $B$ onto the specified nodes with weights assigned using Newman’s collaboration model given as Eq. 2 (Newman 2001):

$$W_{uv} = \sum_k \frac{\delta_k^u \delta_k^v}{k_w - 1}$$
where $u$ and $v$ are nodes from the bottom bipartite node-set, while $k$ is a node from the top bipartite node-set. The degree of node $k$ in the bipartite network is represented by the number $kw$ and $\delta_{w}^{v}$ is 1 if node $u$ is connected to node $k$ in the original bipartite graph, and 0 otherwise.

6. Modularity

Modularity: Modularity is a measure to find the strength of dividing the network into modules. In a network with a high degree of modularity, the connections between nodes in the modules are dense, and between nodes of different modules are sparse. For a weighted network, the value of modularity depends on the edge weight connecting the nodes and is in the range $(-0.5$ to 1). That measures the relative density of edges inside communities for that outside communities (Thakur and Gupta 2017). Modularity of bipartite networks can be optimised to maximum with homogenous communities using one mode projection of networks (Arthur 2020).

7. Model integration

All these approaches: one-mode projection, community detection, and clustering, are widely used methods in bipartite networks. But it has been observed that when used individually, these methods do not provide significant results and retrieval of the meaningful information from bipartite networks. One-mode projection makes the network denser but, in some cases, gives ambiguous or weak results. For community detection algorithms, the modularity score comes very little in bipartite networks, which leads to a very sparse relationship between entities of the same community. The clustering coefficient of small bipartite networks is also not a significant measure to find the strength of clusters (Heer et al. 2020). Thus, an integrated model gets developed to optimize the results for both modularity and clustering. Both approaches integrate to achieve better results: increased modularity, strong communities, and increased clustering coefficient of nodes. Figure 1 shows the proposed model, which is an integration of collaborative unipartite projection of bipartite network based on the weight of edges, and techniques for finding the clustering coefficients of nodes, detecting communities, and modularity analysis of the network.

3 EMBM algorithm of bipartite weighted networks

Based on the dynamic bipartite networks, a model is proposed for optimizing bipartite modularity using model integration as an ensemble model for bipartite modularity (EMBM algorithm). Figure 2 represents the framework of the EMBM algorithm. It is mainly composed of the five steps: Firstly, the dataset is collected and pre-processed for feature selection and removal of undesired data in real datasets. Then the second step is to generate a dynamic bipartite network and visualize it for better understanding.

The third step is to project bipartite data in the unipartite network by collaborating nodes based on the weight of the edges. In the fourth step, modularity of the network, community formation, and the average clustering coefficients get calculated at different timestamps. In the last step, the proposed algorithm gets evaluated by comparing results with the existing algorithms. All these steps can be explained in detail as follows:

3.1 Data

For better results, several procedures were conducted before constructing a dynamic network for visualization and analysis. Initially, related datasets are discovered and examined for desirable characteristics. After that, unwanted data gets removed by pre-processing and feature selection.

Fig. 1 The model integration framework for EMBM

Fig. 2 The flow diagram of the EMBM algorithm
3.1.1 Collection of datasets

To establish the robustness of the proposed method, we perform experiments analysis on two real world dataset and one synthetic dataset. Experimental analysis is performed in python using networkx and for visualization of the results, Gephi is used.

3.1.1.1 COVID-19 dataset For the analyses purpose, Oxford COVID-19 Government Response Tracker (OxCGRT) is customized, that reports the prerequisite for uninterrupted updating, readily functioning and comparable statistics on policy measure. OxCGRT intelligences information on 20 indicators of government response of 186 nations across the globe. The indicators are of three types namely-ordinal, numeric and text. Ordinal pointers are used to quantify policies on a modest gage of sternness or intensity. Although, daily testification of data by the indicators is available, but to preserve the simplicity and clarity of process in our research we concentrate on fortnight data. The aggregated information of four policy indices as: Overall government response index, Stringency index, Containment and health index and Economic support index is accessible in the dataset for readily use. As mentioned above, the present research concentrates on two dimensions—stringency indices and containment and health indices and shows variation in these Responses as value of policy indicator changes over time. All of the indexes are created depending on several distinct policy response indicators. A score is then calculated for every policy indicator by deducting a half-point from the ordinal value if the policy is broad rather than tailored (Hale et al. 2020). These are further escalated by their highest value to provide a score ranging from 0 to 100, with a missing value adding 0. The composite indices are created by averaging the found scores. Both indexes are simply averages of the component indicators’ values.

3.1.1.2 Crime dataset The data is taken from datasets of crime activity from San Francisco that are publicly available (we picked this because of the well-organized data). San Francisco crime data includes 2,142,685 events from January 1, 2003, to November 8, 2017, roughly 456,071 incidents from January 1, 2018, to May 8, 2021. Both datasets are merged to create a combined crime dataset from 2013 to 2021. Then dataset is pre-processed for better analysis that covers discretization of time series and normalization of data (Feng et al. 2018). Then a dynamic network is generated according to the timeline in proper layout for better insight of network. After the generation of dynamic network visualization and analysis techniques are applied to retrieve useful information from the network to meet our objectives.

3.1.1.3 Artificially generated E-commerce dataset To further proves the strength of the model, we create an artificially generated E-commerce weighted network for product and purchaser. The network represents the shopping pattern of the purchasers in a year with monthly timestamps. The relationship between both types of nodes: product and purchaser, shows that the purchaser bought the subproducts under the product category in a particular month. The additional weight on edges represents the number of subproducts purchased by the purchaser of a particular category in a month.

3.1.2 Feature selection and pre-processing

Real-world datasets contain vast data with additional features. Desired features are selected for our study and then pre-process the dataset to remove undesired instances and redundancy. In the Covid-19 dataset, government policies: Testing, Vaccination, International Travel ban, Contact Tracing, and Restriction on gatherings adopted by 11 countries and two response indexes: Stringency and Containment Health, are selected to find the effect of policies on the pandemic. However, the model can extend to numerous other countries and their corresponding policy indicators. In the crime dataset, five crime categories: Disorderly conduct, Robbery, Kidnapping, Larceny/Theft, and other offenses incident in 10 police districts of San Francisco city selected. All the police districts had multiple attributes: number of incident cases, resolved cases, weekend cases, neighbourhood cases, etc. For artificially created e-commerce dataset, 30 purchasers and five categories of products: Beauty Health & Grocery, Fashion & Beauty, Mobiles & Laptops, Home Furnishings, and Electronic Appliances having multiple subproducts selected.

3.2 Generation of dynamic bipartite weighted network

To convert tabular data into a graph format, a set of nodes and edges have to identify. Nodes are the entities that specify the domain of a network, and edges show the relationship between the nodes. Two tables are created for the data laboratory: a node table and an edge table to define the features associated with nodes and edges, respectively.

3.2.1 Node table

It holds node-specific information such as node id, label, timestamp, and attribute list. The timestamp is an extra element that gives the graph dynamics. Node ids are necessary to distinguish between each node. The timestamp includes information about the date and time that each node in the network was active. Table 1 represents the
nodes for all three dynamic networks. For Dynamic Covid 19 network, nodes get divided into two types: Countries and policies adopted by them against the pandemic. There are 11 nodes of countries and five nodes of adopted policies. The timestamp for data aggregation used is fortnight. For Dynamic Crime network, nodes get divided into police districts and crime categories. Ten police district nodes and five crime category nodes get identified. Year-wise aggregated data is used as a timestamp. For the Dynamic E-commerce network, five nodes for product types and 30 purchaser nodes get identified. The monthly purchasing pattern of the purchaser get analysed for these products.

3.2.2 Edge table

An edge in a graph denotes the link between two unique nodes and the relationship between them. Edges can be directed or undirected or can be weighted or unweighted. Weight on edges specifies the strength of the connection between two distinct nodes. Table 2 represents the edge table of all three dynamic networks. Edges are undirected and weighted for all dynamic networks and present between nodes of different types only. This feature helps us to deduce that all of these networks can follow the bipartite pattern. The weighted edge between nodes in the dynamic Covid 19 network shows the ordinal value of policy indicator: the measure of application of policies against the pandemic in different countries on a fortnight basis. In a dynamic crime network, edges between police district and crime category and weight on edges show the number of cases incident in the police district of a particular crime category in a year. In a dynamic E-commerce network, weighted edges present between purchaser and product represent the number of subproducts purchased by the purchaser of the particular product category in a month.

3.2.3 Visualization of dynamic bipartite network

The dynamic network created can be visualized in Gephi (Wajahat et al. 2020), open-source software to visualize and analyse the dynamic networks. Visualization helps to show the effect of variation in nodes and edges based on weight and attributes (Bokhare and Metkewar 2021). For example, it can be used to visualize the effect of variation in policies indices value on government response index values for dynamic Covid-19 Network or effect of variation in crime cases incident in police districts for dynamic Crime network. Visualization in a dynamic E-commerce network helps visualize variation in product sales and its impact on the purchaser. These deviations can be easily viewed by applying the subsequent features like variation in size and colour of the nodes and edges based on their attribute values. Different communities detected within a dynamic network can also get visualized for better understanding.

3.2.3.1 Multi-mode network transformation of network

As all the networks analysing are bipartite, the next step is to visualize the graph in the form of the two-mode pattern to understand the results better of further analysis. Most of the time, such patterns in a network are detected but not visualized. A multi-mode network transformation method (Jaroslav Kuchar 2014) got applied for the projection/transformation of the network as bipartite. Based on the matrix multiplication approach, it allows different types of projections. This process helps project a network in bipartite or multipartite form based on the number of node types or any other attribute.

The resultant graph shows the entities of different sets in different colours. Two entities of similar colour show that they lie in the same set. For the multipartite graph, the

| Table 1 | Node table for (a) Dynamic covid 19 network, (b) dynamic crime network, and (c) dynamic e-commerce network |
|---------|---------------------------------------------------------------------------------------------------|
| Dynamic network | Dataset | Node type 1 | Node type 2 | Timestamp |
| Covid-19 | https://covidtracker.bsg.ox.ac.uk/ | Policies against pandemic | Country | Fortnight |
| Crime | San Francisco Crime dataset 2003-March 2021 | Crime category | Police District | Yearly |
| E-commerce | Artificially generated | Product category | Purchaser | Monthly |

| Table 2 | Edge table for (a) dynamic covid 19 network, (b) dynamic crime network, and (c) dynamic e-commerce network |
|---------|---------------------------------------------------------------------------------------------------|
| Dynamic network | Edge type | Number of edges | Weight on edges |
| Covid-19 | Weighted, undirected | 55 | Policy Indices |
| Crime | Weighted, undirected | 50 | Number of cases of particular crime |
| E-commerce | Weighted, undirected | 150 | Number of subproducts purchased of particular product category in a month |
number of colours will depend upon the number of different sets of entities present in the network. For a bipartite network as the entities are divided into two sets only thus the two colours are there for entities.

Figure 3 shows the result of multi-mode transformation for the bipartite network for Covid 19 and the resultant graph after graph colouring of nodes based on the node type. Nodes of the same set have the same colour, where red nodes representing countries and green nodes represent policies adopted by these countries against pandemics.

Figure 4 shows the Dynamic bipartite crime network where red nodes indicate police districts of San Francisco city and green nodes represents the crime categories.

Figure 5 shows the Dynamic Bipartite E-commerce network for online shopping portal, where red nodes are representing purchasers and green nodes are for product types.

3.3 One-mode projection of bipartite network

After the creation and visualization of dynamic bipartite networks, the next step is the unipartite projection of the network. It not only helps to connect the nodes of a similar type with much density but also in the detection of communities having nodes only of the same set is otherwise not possible in the bipartite pattern. The nodes retain their attributes and are connected in the resulting graph if they have an edge to a common node in the original bipartite graph. A widely approach to project or partition bipartite networks of original networks in a one-mode projection. Consider two nodes, say a and b, in a bipartite network one-mode projection, nodes of the same kind, say a link with one another if they have a common node of the other, say b. For example, in a scientific author network, two authors are considered to be connected if they are working on similar scientific experiments or working in the same lab. While there are certain topological features in the network, the one-mode screening process may simultaneously mean that the converted network becomes denser and more information can be analysed and retrieved with these networks (Zweig and Kaufmann 2011). In most cases, various bipartite networks can decrease to one mode, and a clustering based on this could result in the retrieval of sound information, and the simplicity of this approach may utilize for numerous projections (Bailey et al. 2020a, b).

3.3.1 Weighted collaboration of dynamic network

Distances between entities can get calculated using weighted cooperation graphs. The distance between nodes is equal to the inverse of the weight of their collaborative relationship in this basic computation (Newman 2001). As
a result, if one pair of nodes connects with double strength as another pair, their gap is cut in half. In the Bipartite weighted network, this concept is well implemented to find a collaborative weighted network where the strength of the connection between conjunct entities depends on the weight of edges connected them indirectly (Moshontz et al. 2018). This method may be used to project bipartite networks in unipartite form with collaboration made based on weights of edges. A collaborative weighted approach applied for one-mode projection on all three dynamic networks creates the relationships between two nodes of the same type based on the strength of relations these nodes have with the same node of a different. Figure 6 shows the one-mode projection of dynamic networks. In the dynamic Covid-19 network, country nodes are connected based on their connection with policy nodes in the original network. Policy nodes are also get connected in the same manner. In a dynamic crime network, crime categories are connected based on the number of cases in the same police district. Police districts also get related to each other. In a dynamic E-commerce network, purchasers are related based on their choice of the product purchased, and products get connected according to their quantity purchased by the same purchaser.

3.4 Modularity analysis

Modularity is a metric that assesses how well a network can be divided into modules. A high modularity number indicates the presence of many communities, and each group may then comprise items from a specific category or genre. Modularity score and community structure get analyzed for all three dynamic networks.

3.4.1 Modularity score

Modularity is a measure to find the strength of dividing the network into modules. In a network with a high degree of modularity, the connections between nodes in the modules are dense, and between nodes of different modules are sparse. For a weighted network, the value of modularity depends on the edge weight connecting the nodes and is in the range $(-0.5 \text{ to } 1)$. That measures the relative density of edges inside communities for that outside communities. By finding the modularity of original bipartite networks, using different methods: Louvain method (Blondel et al. 2008) and Girvan-Newman method (2002) of evolution graphs at different timestamps and plotting them, we find that at almost all discrete-timestamps, the value of modularity of the network is on the lower side ($< 0.2$). That concludes that the graph is robustly connected where the
nodes inside the community have a weak connection. The modularity is then got analysed on unipartite projected dynamic networks at different timestamps. It has been observed that the modularity score increased in all three dynamic networks for almost every timestamp. It indicates that the network gets divide into strongly connected modules: nodes within the same module are strongly connected compared to the remaining nodes of the network.

3.4.2 One-mode communities

As communities represent the strength of the connection between nodes, nodes of the same community can treat the same for most decision-making and future planning (Cherifi et al. 2019). Communities are detected at different timestamps for both original networks and then projected networks. It gets noticed that in native networks, the nodes of different types include in the same community. This aggregation is allowable for community detection, but it does not contribute anything if the connection between similar entities has to be analysed. For example, similar strategic planning will not always work for entities of a different type. Then we detect the communities in the unipartite projected version of dynamic networks. The number of communities detected increased with nodes of similar types only get included in the same community. This information is vital to retrieve for the decision-making process for the same entity types within a community. For dynamic covid 19 networks, this information helps to plan similarly for policies of the same community and follow a different approach for policies of different communities. For dynamic crime networks, it aids in formulating similar crime prevention strategies at various levels: tactical, operational, and strategic levels for crime categories of the same community. In the case of dynamic e-commerce networks, it helps in similar arrangements and sale policies for products of the same community to increase their sale.

3.4.3 Calculation of bipartite clustering coefficient

Clustering coefficients represent that when two nodes share anything (a neighbour), they connect with a significantly higher likelihood than two randomly picked nodes (Zhang et al. 2008; Chi et al. 2021). On the other hand, they represent the fact that when two nodes are linked together, they are likely to share neighbours (Zhang and Deng 2021). To put it another way, they capture neighbourhood correlations. Using this perspective, the first idea of clustering coefficient for node pairs can be established as Eq. 3 (Latapy et al. 2008) (in the same set or different set).

\[ CC(u, v) = \frac{|N(u) \cap N(v)|}{|N(u) \cup N(v)|} \] (3)

It captures the overlap between node neighbourhoods: for example, if \( u \) and \( v \) have no common neighbour, \( CC(u, v) = 0 \). If they live in the same area, then \( CC(u, v) = 1 \). When their neighbourhoods partially overlap, the value is somewhere in the middle, closer to 1 when the overlap is substantial compared to their degrees. The clustering coefficient is a measurement of the density of each vertex’s 1.5-degree egocentric network. The clustering coefficient is high for dense linkages between nodes (Heer et al. 2020; Chessa et al. 2014). For example, in a friendship network, if all friends know each other very well, its clustering coefficient will be high. The high value of the average clustering coefficient of dynamic networks helps in decision-making and the same strategic planning of all nodes in the network. The average clustering coefficients get calculated of original networks for communities detected using the Louvain and Girvan-Newman method. The values are negligible in both cases. It indicates that the nodes connect weakly with their neighbours. Then we calculate the average clustering of the communities detected from the unipartite projection of dynamic networks based on our approach. The value of the average clustering coefficient increased to a marginal level. It indicates that nodes get connected with a strong relation to each other. It may be applicable in various scenarios, such as determining the impact of having one common neighbour on the rest of the neighbourhood.

3.5 Algorithm evaluation

To evaluate the ensemble model for bipartite modularity, the unipartite projected networks are analysed for the parameters mentioned in the previous step: modularity score, community structure, and average clustering coefficient at different timestamps. Then these parameters were analysed on original networks to evaluate the performance of the proposed approach compared to the existing algorithms.

4 Results and discussions

After analysing both original dynamic bipartite networks and then projected dynamic unipartite networks (based on our integrated approach), results can easily be compared for all three networks.

4.1 Dynamic covid 19 network

We consider the timestamp of January 2021 for analysis. For the Louvain method, the modularity score comes out as 0.017, 2 communities get detected with the average
clustering coefficient of the network was 0. For the Newman method, the modularity score was $-0.010$, 3 communities get detected, and the average clustering coefficient was 0. For present ensemble method, the modularity score was 0.405, 4 communities get detected, and the bipartite clustering coefficient was 1. It concludes that our proposed method achieves an increase in the modularity and clustering coefficient of bipartite networks. Figure 7 shows the communities detected by using all three methods. The same colour represents the nodes of the same community. It can find that in the Louvain and Newman method nodes of different types are included in the same community as Turkey, United Kingdom, Italy, and Germany include in the same community as vaccination and testing. Whereas in present approach the nodes of similar types only are included in a community such as France and Germany make a community, Brazil is a single node in another community and all remaining countries form another community, all the policies are included in another community. It will be very helpful in making future strategies for countries and policies against the pandemic.

4.2 Dynamic crime network

For a dynamic crime network, the timestamp of the year 2020 is considered. The modularity score of the original dynamic network using the Louvain method was $-0.010$, 2 communities were detected and the average clustering coefficient was 0. The modularity score using the Newman method on the same network was 0.014, 3 communities were detected and the average clustering coefficient was 0. For our proposed method on a one-mode projected crime network, the modularity score was 0.405, 04 communities were detected and the average clustering coefficient value was 1.

Figure 8 shows the communities detected in dynamic crime network using all three algorithms. It can be seen that both in Louvain and Girvan-Newman’s method, crime categories and police districts are forming the same community. It will not help in efficient planning against violence.

But in the EMBM method police districts: Bayview and central are in one community, whereas all remaining locations are in another. This will help security analysts to construct crime prevention planning for Bayview and the central district in a similar manner as compared to other districts in the list. Similarly, Disorderly conduct is the crime category that is unique in its community, and all other crime categories forming different communities. This information may be used to treat disorderly conduct differently from other crime categories for the prevention program.

4.3 Dynamic e-commerce network

For dynamic E-commerce network analysis, the timestamp of June 2020 is considered. The modularity score comes out to be 0.146 using the Louvain method, with two communities get detected, and the average clustering coefficient was 0. For the Newman method, the modularity score was $-0.021$, with three communities identified, and the average clustering coefficient was 0.3. For present ensembled model, the modularity score of the projected graph is 0.466, with four communities were detected, and the average clustering coefficient was 1. Figure 9 shows the communities identified in a dynamic e-commerce network using these three methods. A similar result is found in this network also. In both Louvain and Newman’s algorithms, detected communities have nodes of dissimilar types. In the Louvain method, home furnishings include with purchasers: 3, 6, 13, 24, 25, and 26 in one community. In the Newman method, communities formed include nodes of similar types with a negative modularity score indicates their connection is weak within the community.
In present proposed approach, purchasers: 2, 21, and 23 are in one community, purchasers 1, 18, and 22 are in the second community, and the rest of the purchasers form the third community. This division will help in making policies and strategic planning for purchasers according to their community. Information of products making one community also helps to unique sales plan to increase the sale of products.

4.4 Comparative analysis of proposed approach with existing algorithms

Results induced by different algorithms are compared to represents the effect of the proposed study on the modularity, the average clustering coefficient, and the detection of one mode communities in the dynamic bipartite networks. The observations are summarized both in tabular form and chart representations. Table 3 shows the comparative analysis of all three algorithms: Louvain algorithm, Girvan-Newman algorithm, and EMBM algorithm for parameters: modularity, number of communities, and average clustering coefficient. It concludes that the ensemble model for bipartite modularity performed better on all parameters for each dynamic network comparing existing methods. Figure 10 shows the modularity analysis chart for all dynamic networks representing the scores of different algorithms. It displays that the modularity score gets increased significantly by using our integrated model for the bipartite network.

5 Conclusion and future scope

Many recent studies have shown that network science can evaluates complex real-world systems of different domains. The conducted research is an example of how network science provides a better empathetic of the structure and behaviour of the bipartite networks. The evolutionary nature of the interactions in bipartite networks is examined with the incorporation of timestamps to further investigate the dynamics of the network. The utmost requirement for modularity optimization in real-world applications is the network partition in the most pertinent manner. The present model is candid and supports to compute the gain of bipartite modularity and the one-mode projection of networks in an efficient and simple manner.
The model may get used in a variety of real-world scenarios: including author networks, movie networks, community-based recommendation system design and so on. The suggested method of one-mode projection of bipartite graphs aids in the interpretation of information contained in these graphs into a more easily comprehended network. Finally, with computer-generated bipartite networks it is verified that it works well to create one mode community structures that are clearly defined and result in better modularity.

The proposed technique may also apply to provide community-based recommendations for large bipartite networks. In the future, the work may be stretched by adding the knowledge of this prime and plentiful form of the network by extending the framework to a broad study of multipartite networks in many ways.

**Author contributions** All of the mentioned authors contributed a significant and intellectual contribution and approve for publication.

**Funding** There was no external funding for this research.

### Declarations

**Conflict of interest** There are no conflicts of interest declared by any of the authors.

**Ethical approval** None of the authors conducted any human or animal experiments for this publication.

**Informed consent** All individual participants included in the study gave their informed consent.

### References

Abrahams B, Sitas N, Esler KJ (2019) Exploring the dynamics of research collaborations by mapping social networks in invasion science. J Environ Manage 229:27–37

Arthur R (2020) Modularity and projection of bipartite networks. Phys A Stat Mech Appl 549:124341

Bailey L, Harinam V, Ariel B (2020a) Victims, offenders and victim-offender overlaps of knife crime: A social network analysis approach using police records. PLoS one 15(12):e0242621

Bailey L, Vincent H, Barak A (2020b) Victims, offenders and victim-offender overlaps of knife crime: a social network analysis approach using police records.” PLoS one 15.12: e0242621

Barber MJ, Clark JW (2009) Detecting network communities by propagating labels under constraints. Phys Rev E 80(2):026129
