Spectrum Analysis of Contact Network for Public Policy in a Pandemic

QI XU1
1School of Mathematical Sciences, University Of Electronic Science And Technology Of China, Chengdu, China. (email: 2019110601002@std.uestc.edu.cn)

ABSTRACT In a pandemic, in order to slow down the spread of the virus, protect national health, and maintain the normal operation of economic activities, countries around the world will formulate public policies to limit the number of citizens that can gather. Our research focuses on how to achieve optimal public policy under different conditions. Traditional SIR and SEIR models can well reflect the transmission process and obtain credible prediction results from a macro perspective, but lack the sensitivity of micro data, and cannot assess the risk of epidemic transmission brought by close contacts and sub-close contacts. Based on the Barabási-Albert scale-free network and the Random spanning tree algorithm, we generate a simulated spread network for non-specific infectious diseases. At the same time, we also generate group networks under different gather constraints. The superposition of the two forms a composite contact network. Our research work on the contact network shows that after considering close contacts and sub-close contacts, the public policy optimization problem of slowing the spread of the epidemic can be answered by the spectrum analysis of the contact network. We perform computer simulations and theoretical proofs of this model, and conduct a transmission analysis of its process.

INDEX TERMS Contact network, infectious disease, laplacian matrix, public policy, spectrum analysis, transmission network.

I. INTRODUCTION

In the digital age, governments and scientific research institutions can easily obtain some accurate group contact network data in society, such as public transportation passenger lists, company employee lists, and student registration information in educational institutions. The contact network within each group constitutes a sub-network in the giant relational network of a community or city. From the relationship between the sub-networks, we can extract a lot of useful information. In a pandemic, it has become a challenge for government departments to reasonably limit the connections between these groups and formulate optimal public policies. Combining known epidemic data, analyzing the relationship between these sub-networks to provide a scientific theoretical basis for the government’s public policy formulation has become one of the primary goals of our work.

In a pandemic, countries around the world will take various countermeasures and formulate different public policies in order to slow down the spread of the virus, protect national health, and maintain the normal operation of economic activities. Among them, one of the most extensive measures is to limit the number of people gathered in public places, such as limiting the number of passengers on public transport, limiting the number of people in restaurants, limiting the size of gatherings, etc. Our work focuses on researching and evaluating the social infection risk level under different aggregation restriction levels. This evaluation result can take into account the impact of close contacts and sub-close contacts on the risk level during the spread of the epidemic.

In the past, the main mathematical model used to study the spread of infectious diseases was the SIR or SEIR model. An SIR model is an epidemiological model that computes the theoretical number of people infected with a contagious illness in a closed population over time. The name of this class of models derives from the fact that they involve coupled equations relating the number of susceptible people $S(t)$, number of people infected $I(t)$, and number of people who have recovered $R(t)$ [1]. An SEIR model additionally considers number of people exposed to source of infection $E(t)$. The SIR model consists of a system of three coupled nonlinear ordinary differential equations,

$$\frac{dS}{dt} = -\beta SI$$  \hspace{1cm} (1)

$$\frac{dI}{dt} = \beta SI - \gamma I$$  \hspace{1cm} (2)
where $t$ is time, $\beta$ is the infection rate, and $\gamma$ is the recovery rate [2].

Spectrum Analysis is a mathematical method for evaluating the connection properties between vertices by studying the eigenvalues of the Laplacian Matrix of the Contact Network. For infectious diseases, common models based on analytical mathematics, such as SIR and SEIR models, can well reflect the propagation process and obtain reasonable prediction results from a macro perspective, but lack the sensitivity to micro-accurate data. I.e., these models cannot take into account the close contacts and sub-close contacts of the infected person. Therefore, traditional models cannot assess the impact of public policies on crowd gathering restrictions.

The study of the eigenvalue properties of the Laplacian Matrix makes it possible to focus on the microscopic connections between vertices, due to the sensitivity of the Laplacian Matrix. This means that we can analyze the close contacts of infected nodes as well as the next close contacts. Then, the overall infection risk level of the society can be assessed, so that governments of various countries can formulate reasonable public policies for epidemic prevention and control based on their own social conditions and characteristics.

Governments around the world have imposed different policies and undertaken preventive public health measures, such as social distancing orders, travel restrictions, local or national lockdown, and partial or complete border closures, to control the spread of the pandemic [3]. A better understanding of the underlying processes affecting pandemic dynamics and related infection patterns will help to improve the effectiveness of public health interventions worldwide [4]. Consequently, we propose a method of spectrum analysis that can assess the impact of close and sub-close contacts on risk levels across the community.

The eigenvalues of a matrix $A$ are called its spectrum [5]. In the past, research on spectrum of a matrix has mainly focused on the algebraic connectivity of a graph, and the relationship between the largest eigenvalue and domination number. The algebraic connectivity (also known as Fiedler value or Fiedler eigenvalue) of a graph $G$ is the second-smallest eigenvalue (counting multiple eigenvalues separately) of the Laplacian matrix of $G$ [6]. Clemens Brand and Norbert Seifter proved that there is a close connection between the domination number and the largest eigenvalue [7]. With the development of computer performance and the popularization of informatization, the use of spectral analysis methods to evaluate large-scale complex networks has become an acceptable option.

Governments all over the world were and are still looking for ways to contain the disease and alleviate its disastrous consequences to their people’s health and economy [8]. With rapid advances in epidemic surveillance systems, abundant data of infection are collected. The availability of these data provides unprecedented opportunities to investigate the relationships between different risk factors [9]. Human mobility patterns are primarily attributed to the structure of the spatial network [10]. In a small-scale spatial environment, individuals often follow the road network to visit a set of locations (e.g., workplaces, shopping malls, grocery stores). In a large spatial environment, people tend to travel through a spatial network of airports, railroads, and highways [11]. As such, spatial network characteristics can also be potentially related to the infectious disease spread in the spatial environment. The objectification of the spatial network is extremely contact network. Fast, reliable, and easily accessible clinical assessment of the severity of the disease can help in allocating and prioritizing resources to reduce mortality [3]. The analysis of the contact network can help us understand the transmission network of the disease, and thus obtain valuable information.

To achieve our research goals, we simulated a relational network with at least 10,000 nodes, and the relational network was generated by superimposing the simulated propagation network and the group relational network. We hope that our model can cope with different types of infectious diseases. To adapt to diseases with different spreading characteristics, we used the Barabási-Albert scale-free network and the Random spanning tree algorithm to generate our simulated spreading network, respectively, so that the degree distribution of the nodes showed a power-law distribution and a Possion distribution, respectively. By studying the properties of the Laplacian Matrix spectrum, we hope to reduce the time complexity of the process of analyzing the connections between sub-networks from $O(n^3)$ (considering all possible connections and propagation paths) to $O(n^3)$ (time complexity of computing matrix eigenvalues).

Based on a series of simulation operations and theoretical analysis of complex networks, we believe that after considering close contacts and sub-close contacts, the public policy optimization problem of slowing the spread of the epidemic can be solved by the spectrum analysis of the contact network. In the following text, we give the analysis process and proof, and conduct a diffusion analysis of its process.

II. NETWORK BUILDING AND SIMULATION

Before the theoretical analysis, we established a simulation model of infectious disease transmission and a group generation model. We generated our simulated propagation network using the Barabási-Albert scale-free network and Random spanning tree algorithms, respectively. All networks are finally recorded as an adjacency matrix of an undirected graph. From the specific implementation point of view, we divide the nodes into two categories, namely infected vertices and safe vertices. Whenever a new node is infected, we randomly select a node from safe vertices to join infected vertices.

Then, according to different infection network generation algorithms, we select a node from the infected vertices as the infection source and connect it with the new infected vertex. For the Barabási-Albert scale-free network, we recorded the
degree of each node and the sum of the degrees of all nodes. For each new infected vertex, node \( k \) becomes the source of infection with probability

\[
p_i = \frac{d_i}{\sum_{k=1}^{N} d_k},
\]

where \( d_i \) is the degree of vertex \( i \). For networks generated by the Random spanning tree algorithm, the selection of infection sources is equally possible. Finally, modify the adjacency matrix corresponding to the network to adapt it to the changed network structure.

\[
\Gamma_d = d - \frac{1}{\rho} - d
\]

\[
\rho = \frac{d}{2}
\]

where \( d \) is the degree of vertex \( i \). For networks generated by the Random spanning tree algorithm, the selection of infection sources is equally possible. Finally, modify the adjacency matrix corresponding to the network to adapt it to the changed network structure.

\[
\Gamma_d = d - \frac{1}{\rho} - d
\]

\[
\rho = \frac{d}{2}
\]

In addition to this, in order to make our simulated network realistic, our contact network model is based on the following two assumptions during establishment:

1) Stochastic Consistency Assumption. Although in the real world each node represents a different individual, all nodes in our contact network are considered to be randomly consistent. This means that although each individual in reality has its specific contact information, in our simulated large-scale contact network, this specific contact information is always represented by a randomly generated edge. The establishment of the stochastic consistency assumption has become a reasonable theoretical premise for our subsequent simulation calculation and analysis process.

2) Homogeneity assumption. Because our research focuses on the impact of infectious diseases on the overall structure of the network, different infectious diseases spread at different rates, and we only need to focus on the current state of the network. In addition to this, in order to make our simulated network realistic, our contact network model is based on the following two assumptions during establishment:

- **Stochastic Consistency Assumption**: In the entire model we built, there is no time index. The reason is that our study focuses on the impact of the current infection status on the overall structure of the network. Different infectious diseases spread at different rates, and we only need to focus on the current state of the network. In addition to this, in order to make our simulated network realistic, our contact network model is based on the following two assumptions during establishment:

1. **Stochastic Consistency Assumption**: Although in the real world each node represents a different individual, all nodes in our contact network are considered to be randomly consistent. This means that although each individual in reality has its specific contact information, in our simulated large-scale contact network, this specific contact information is always represented by a randomly generated edge. The establishment of the stochastic consistency assumption has become a reasonable theoretical premise for our subsequent simulation calculation and analysis process.

2. **Homogeneity Assumption**: Because our research focuses on the impact of infectious diseases on the overall network after the spread of infectious diseases across groups, in order to facilitate our model, we treat the probability of spread across groups as a probability parameter \( \rho \). Based on this, the following theoretical analysis was carried out.

**III. DIFFUSION ANALYSIS OF COMPLEX NETWORK**

Epidemiological research has a long history, and a variety of epidemic transmission models have been proposed [12]-[14]. In a typical propagation model, individuals in a population are abstracted into several classes, each of which is in a typical state. Its basic states include S (susceptible), usually healthy state; I (infected), infected state; R (removed, refractory or recovered), removed state (also known as immune state or recovery state). The different contagion models are often named after the transitions between these states. For example, susceptible populations become infected and then recover and become immune, which is called the SIR model.

In our model, the diffusion process between groups is explained by a modified NW small-world model. Newman and Watts first described the diffusion process on a d-dimensional NW small-world network based on the NW small-world network model [15]. Moukarzel made a more specific analysis [16]. We construct the diffusion equation by adapting Moukarzel’s ideas to our diffusion model.

Suppose that starting from the initial infected node, the virus runs at a constant velocity \( v = 1 \) start spreading. The density of shortcut endpoints in the NW small world network is \( \rho = 2\rho \), where \( \rho \) is the probability parameter of adding a new shortcut in the small world network model. To fit our model here, we treat the probability of spread across groups in our model as a probability parameter \( \rho \).

Suppose that this diffusion process is continuous. Therefore, the infection volume of a node in the network \( V(t) \) is a sphere with a radius \( \Gamma_d \) starting from the initial infected node \( \Gamma(t)^d \), where \( \Gamma_d \) is the hypersphere constant in the d-dimensional small-world network. The probability of the infection source spreading across groups is \( \rho \), and the resulting new infection sphere is \( \rho \Gamma_d^d \). Therefore, the average total infection
V(t) is obtained by an integral equation of the form:

\[ V(t) = \Gamma_d \int_0^t \tau^{d-1} \left[ 1 + 2pV(t - \tau) \right] d\tau. \]  

With scale and differential transformation for (5), the linear diffusion equation is obtained

\[ \frac{\partial^d V(t)}{\partial t^d} = 1 + V(t). \]  

Equation (6) will be changed as follows:

\[ \frac{\partial^d V(t)}{\partial t^d} = 1 + V(t - \delta). \]  

In our model, the size \( \delta \) reflects the length of the incubation period of different infectious diseases, and the corresponding solution is

\[ V(t) = \sum_{k=1}^{\lceil \delta \rceil} \frac{(t - k\delta)^d}{(dk)!}, \quad d = 1, 2, 3, \ldots \]  

This theoretical result matches well with the network model we studied. In our model, most of the propagation process analysis is not applicable due to the absence of a time index. But in the model of Moukarzel and Yang, the transmission rate is constant velocity \( v = 1 \), which corresponds to the process of computer program running. In addition, we treat the probability of spread across groups in our model as a probability parameter \( \rho \), which fits our model well.

**IV. EIGENVALUES OF LAPLACIAN MATRIX**

The Laplacian matrix of an undirected graph \( G \) is a real symmetric matrix \( L \), where

\[ l_{ij} = \begin{cases} -1, & \text{if edge } e_{ij} \text{ exists,} \\ 0, & \text{otherwise.} \end{cases} \quad (i \neq j) \]  

and

\[ l_{ii} = -\sum_{k=1, k \neq i}^{n} e_{ij}. \]  

I.e., the degree of vertex \( i \) is denoted by \( l_{ii} \). The eigenvalues of \( L \) are listed in ascending order

\[ \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \ldots \leq \lambda_n. \]  

\( L \) is positive semidefinite and \( \lambda_1 = 0 \).

Let \( W \) be a space such that \( \{W | x^T x = 1, x^T e = 0 \} \). According to Courant-Fischer min-max theorem [19], we have

\[ \lambda_2 = \min_{x \in W} x^T Lx. \]  

The numerical size of \( \lambda_2 \) is very critical, also known as Algebraic connectivity. If \( \lambda_2 = 0 \), then graph \( G \) is disconnected. Furthermore, if a graph \( G \) has \( n \) connected subgraphs, then we have (proof in appendix)

\[ \lambda_1 = \lambda_2 = \lambda_3 = \ldots = \lambda_n = 0. \]  

Therefore, it is necessary to consider the number of 0-eigenvalue firstly. In theory, the number of zero eigenvalues represents the number of connected subgraphs. In the real world, when a transmission chain occurs between two different groups, this means that the number of connected subgraphs decreases. The number of zero eigenvalue is denoted by \( z \). From a macro perspective, if we consider all nodes to be equivalent, then we can treat the transmission process as random. Theoretically, We can calculate the expected value of \( z \) under different epidemic spread scenarios. If every \( k \) nodes form a group, then the percentage of safe nodes is \( k(z - 1)/n \).

Consider 10,000 nodes, set the group limit to 10, and use the random spanning tree model to get the simulation results. The theoretical expectation of the ratio of safe nodes can be calculated using probability. As shown in Figure 3, we compare the simulation results with the theoretical expectations.
curve of the results obtained by the Barabási-Albert scale-free network model or the random spanning tree model is theoretically the same. Because the probability of spread across groups is the same in both cases. We simulated different groups under the limited number of people. The transmission process of the Barabási-Albert scale-free network model and the random spanning tree model are shown in Figure 4 respectively.

After considering zero eigenvalues, we convert all eigenvalues of the Laplacian matrix of the relationship network generated using random spanning tree into a mesh graph. For the largest eigenvalue $\lambda_n$, there is a close connection between the domination number $\gamma(G)$ and $\lambda_n$. It’s proved that

$$\lambda_n < n - \left[\frac{\gamma - 2}{2}\right],$$

(14)

if $\gamma(G) > 3$ [7].

Since our transmission process is randomly generated by the model, it does not make sense to consider a specific dominating set as well as the largest eigenvector. Observing the remaining eigenvalues, a remarkable rule is that these eigenvalues exhibit a ladder feature.

In order to further obtain its change characteristics, we selected the 90% smaller eigenvalues of the Laplacian matrix of the corresponding network during the epidemic spread simulated by random spanning tree to draw a profile. Figure 5 intuitively demonstrates this property.

The eigenvalues are in ascending order, and the ladder feature gradually disappears. Let $L_s$ be a Laplacian submatrix of network $G$, the eigenvalues of $L_s$ will remain constant in that of $L_G$ while there is no edge between group $s$ and others. I.e., if

$$L_s v_s = \lambda_i v_s,$$

(15)

then,

$$L_G v'_s = \begin{pmatrix} L_s & A \\ B & \ldots & 0 \\ \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} v_s \\ 0 \\ \vdots \end{pmatrix} = \lambda_i v'_s,$$

(16)

while the elements of matrix $A$ and $B$ are all zero. The disappearance of the eigenvalue ladder feature is caused by the increase of infected persons. It is reflected in complex networks, that is, the number of edges between each group increases, which changes the eigenvalues of the original group. The edges between different groups will affect the eigenvalues, so the changes in the eigenvalues can reflect the changes in the contact network to a certain extent. By evaluating the ladder features appearing in the eigenvalues, the spread of the epidemic can be indirectly understood. When the number of close contacts and sub-close contacts gradually increased, the eigenvalue curve gradually became smooth. Therefore, our next work needs to evaluate the ladder feature of the eigenvalue curve.

We know the number of 0 eigenvalues $\gamma$. Now we can indirectly evaluate the contact relationship between all nodes by evaluating the ladder feature. It should be noted that the premise of this idea is that the homogeneity assumption of the model holds. This means that we ignore the contacts that exist inside the safe group. This assumption is reasonable, because in the real case, we only need to track the impact of infected individuals on the group. So far, how to evaluate the ladder feature has become the main content of our next work.

V. EVALUATION OF RISK

Evaluating the ladder feature and getting quantifiable results is one of the goals of our research. After seeing data features, we can first think of using classification algorithms, such as Naive Bayesian Classifier algorithm, Logistic Regress algorithm, decision tree algorithm, K-Nearest Neighbor algorithm, Artificial Neural Network algorithm, etc.

To solve the problem with an efficient and suitable algorithm, the following requirements are necessary:

1) The algorithm logic is simple, easy to implement and highly interpretable.
2) The algorithm needs to estimate fewer parameters.
3) The algorithm has a small error classification rate.
4) When the number of attributes is relatively large, the classification effect is not affected or is less affected.
5) For a huge amount of data, the time complexity and space complexity of the calculation are small enough.

Since the numerical distribution of the eigenvalues of a specific Laplacian matrix is deterministic and complex. Each eigenvalue is related to all nodes in a specific connected subgraph, so our challenges are as follows:

1) Eigenvalues are difficult to trace. For a large-scale complex network, it is difficult to determine which sub-matrix of the connected sub-graph a particular eigenvalue comes from. Because the essence of the ladder feature is generated by the invariance of the connected subgraph, with the change of the connected subgraph, the ladder feature gradually disappears.
2) Uncertainty in the number of eigenvalues at a level in the Ladder feature. Because as the connected subgraph changes, the eigenvalues of the corresponding submatrix change. The eigenvalues of the changed matrix may migrate to other levels, which makes the number of eigenvalues at different levels continue to change. The above phenomenon will affect the subsequence analysis.

Combining data characteristics and analysis purposes, we use an algorithm that does not perform classification to evaluate different Laplacian matrices. Here we use an improved linear regression model to analyze the data. Before the calculation, we first analyze the numerical distribution of the eigenvalues. Through model calculation, the number of eigenvalues presenting the ladder feature is limited. For the network models generated by the Barabási-Albert scale-free network and Random spanning tree, the larger eigenvalue sizes of the Laplacian matrices of these two models are related to the degree distribution of nodes. In order to unify the evaluation interval and avoid the influence of larger eigenvalues on the evaluation results, our evaluation interval covers
FIGURE 4. Safe vertices rate in two cases. We can see Barabási-Albert network on the left, and random spanning tree model on the right.

FIGURE 5. Ascending order eigenvalues of Laplacian matrix of merged contact network with 10000 vertices by 30 members per group.

70% of the smaller eigenvalues. The reason for choosing this interval is that the larger eigenvalue reflects the nature of the dominating set, and we focus on the information of close contacts and sub-close contacts in the real world, i.e., those nodes within the group that are close to infected nodes are the ones we want to focus on.

To reasonably quantify the ladder feature, a more intuitive method is to use linear regression to calculate residual sum of squares (RSS). The results of such calculations are intuitive and interpretable, and we demonstrate this well using actual calculation results. The operation results are close to the ideal results, so we analyze the transmission network generated by the Barabási-Albert scale-free network and the Random spanning tree algorithm. Since the difference between the eigenvalues of the two transmission networks is mainly reflected in the larger part, the results of the two are similar in the linear regression model calculated with 70% of the smaller eigenvalues. Finally, we decided to use the RSS of the linear regression model as an indicator for evaluating close and sub-close contacts. Intuitively, the more obvious the ladder feature, the less the connection between the connected subgraphs, and the larger the RSS. As the ladder feature fades away, the RSS decreases, reflecting the decrease in the number of safe vertices. Figure 7 and Table 1 show the RSS values at different infection rates and at different aggregation limits.

At this point, we can make a reasonable evaluation based on the RSS table. For example, when considering 10,000 nodes, when the infection rate is 1%, the RSS value of the simulated network corresponding to the group limit of 15 people is 216.97. This result was similar to the 10-person group limit when the infection rate was 5% to 6% (RSS values were 208.36 and 222.31, respectively). This means that the risk level for the 10,000-person community is similar in both cases. Then, according to this method, the local government can formulate the optimal public policy according to the medical level and economic and trade situation of its own society. On this basis, the infectious disease model has also been extended. The results of the evaluation method for close contacts and sub-close contacts proposed in this paper can also be one of the basis for subsequent policy formulation and situation prediction.
In fact, in a pandemic, the challenges facing governments and residents go beyond infectious disease and economic loss. Around the world, logistics disruptions, shortages of medical resources and living materials due to the pandemic are also real. Therefore, when faced with a pandemic, the government’s public policy must be commensurate with the actual scale of disease transmission and the corresponding level of risk. In addition, through simulation analysis in advance to correctly understand the scale of disease transmission and social risk level, it will also help the government to rationally allocate living materials and medical resources. In conclusion, we hope our work can contribute to the fight against infectious diseases for mankind.

VI. CONCLUSION

In this paper, we propose a method that can assess the impact of close and sub-close contacts on risk levels across the community. Based on the Barabási-Albert scale-free network and the Random spanning tree algorithm, we generate a simulated spread network for infectious diseases. After stacking the group network, we performed spectral analysis on the laplacian matrix of the generated network. Our work shows that, after considering close contacts and sub-close contacts,
the public policy optimization problem of slowing the spread of the epidemic can be answered by profiling the contact network. We perform computer simulations and theoretical proofs of this model, and conduct a propagation analysis of its process. We believe that the conclusions obtained in this paper can provide a basis for making public policy and predicting the state of disease transmission.

Our model also has some shortcomings and needs to be continuously improved in the follow-up work. In the process of simulation and analysis, with the increase of the model scale, the space for storing the laplacian matrix of the transmission network continues to increase, and the time required to calculate the eigenvalues also increases. Considering the need to reduce storage space, after analysis, the matrix that needs to be stored in the model is a sparse matrix. Therefore, we can introduce a compressed storage algorithm for sparse matrices to improve the utilization of storage space. I.e., store the laplacian matrix in a compressed format, such as Compressed Sparse Row format(CSR), Block Row Sparse format(BSR), etc. Considering reducing the computing time, in the follow-up work, we can parallelize the serial code and use GPU or multi-core CPU for calculation, thereby improving the operation efficiency.

APPENDIX
PROOF OF SPECTRUM ANALYSIS
A. MIN-MAX THEOREM
Let $M$ be a Hermite matrix of order $n$, where $M$ has $n$ real eigenvalues corresponding to $n$ linearly independent eigenvectors $v_1, v_2, v_3, \ldots, v_n$. The eigenvalues are listed in ascending order

$$\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \ldots \leq \lambda_n,$$

and $v_1, v_2, v_3, \ldots, v_n$, constitute a set of orthonormal basis for $n$-dimensional space $\omega$.

We can get a $k$-dimensional subspace $U$ and the $n - k + 1$-dimensional subspace

$$\text{span}\{v_k, v_{k+1}, v_{k+2}, \ldots, v_n\}$$

from space $\omega$. The intersection between $U$ and

$$\text{span}\{v_k, v_{k+1}, v_{k+2}, \ldots, v_n\}$$

is not null, since the sum of dimensions $\geq n$. Because we can get a vector $x$ from the intersection such that

$$\|x\| = x^H x = 1,$$

$$x = \sum_{i=k}^{n} \alpha_i v_i,$$

and

$$\sum_{i=k}^{n} \alpha_i^2 = 1.$$

Then,

$$x^H M x = \sum_{i=k}^{n} \alpha_i^2 v_i^H M v_i = \sum_{i=k}^{n} \lambda_i \alpha_i^2 \geq \lambda_k,$$

since the orthogonality of the eigenvectors. Similarly, we can also get a vector $x$ from the intersection between a $n - k + 1$-dimensional subspace $V$ and the $k$-dimensional subspace

$$\text{span}\{v_1, v_2, v_3, \ldots, v_k\}$$

such that

$$\|x\| = x^H x = 1,$$

$$x = \sum_{i=1}^{k} \alpha_i v_i,$$

and

$$\sum_{i=1}^{k} \alpha_i^2 = 1.$$

Then,

$$x^H M x = \sum_{i=1}^{k} \alpha_i^2 v_i^H M v_i = \sum_{i=1}^{k} \lambda_i \alpha_i^2 \leq \lambda_k.$$

For (23) and (28), we have

$$\lambda_k^* = \min_{\text{dim}(U)=k} \max_{x \in U, \|x\|=1} x^H M x.$$
D. PROOF OF ALGEBRAIC CONNECTIVITY

For matrix $L$, suppose
\[ \lambda_2 = 0, \]
then, since
\[ \lambda_1 = 0, \]
the multiplicity of 0 the eigenvalue of Laplacian Matrix $L$, is at least 2. To ensure no loss of generality,
\[ \text{Rank}(L) \leq n - 2. \]

Let $G$ be an oriented graph with $n$ vertices and $m$ edges. We have an incidence matrix $D_{n \times m}$ from oriented graph $G$, where
\[ d_{ij} = \begin{cases} 1, & \text{if edge } e \text{ start from vertex } i \\ -1, & \text{if edge } e \text{ end with vertex } i \\ 0, & \text{else.} \end{cases} \]

Perform a row transformation on $D$, then,
\[ D_{\text{row}} = \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{pmatrix}, \]
where $C_i$ is the incidence matrix from each connected subgraph $G_i$ and $G_i$ has $n_i$ vertices. We have
\[ \text{Rank}(C_i) = n_i - 1. \]

There is no intersection among different sets of edge from different connected subgraphs. Therefore, the row vectors of $C_i$ are linearly independent. Then,
\[ \text{Rank}(D) = \text{Rank}(D_{\text{row}}) = \sum_{i=1}^{\psi} (n_i - 1) = n - \psi. \]

The Laplacian matrix $L$ of graph $G$ is denoted by
\[ L = DD^T. \]

Then, we have
\[ n - 2 \geq \text{Rank}(DD^T) = \text{Rank}(D) = n - \psi, \]
\[ \psi \geq 2. \]

That means graph $G$ is disconnected. The same approach also works for the converse and the corollary. Therefore, the conclusion is also true for undirected graph.

REFERENCES

[1] Weisstein, Eric W. "SIR Model." From MathWorld–A Wolfram Web Resource. https://mathworld.wolfram.com/SIRModel.html
[2] Weisstein, Eric W. "Kermack-McKendrick Model." From MathWorld–A Wolfram Web Resource. https://mathworld.wolfram.com/Kermack-McKendrickModel.html
[3] M. Nicola, Z. Alsafi, C. Sohrabi, A. Kerwan, A. Al-Jabir, C. Iosifidis, et al., "The socio-economic implications of the coronavirus pandemic (COVID-19): A review”, Int. J. Surg., vol. 78, pp. 185-193, Jun. 2020.
[4] N. Sapkota et al., "The Chaotic Behavior of the Spread of Infection During the COVID-19 Pandemic in the United States and Globally," in IEEE Access, vol. 9, pp. 80692-80702, 2021, doi: 10.1109/ACCESS.2021.3085240.
[5] Golub, G. H. and Van Loan, C. F. Matrix Computations, 3rd ed. Baltimore, MD: Johns Hopkins University Press, p. 310, 1996.
[6] Weisstein, Eric W. "Algebraic Connectivity." From MathWorld–A Wolfram Web Resource.
[7] Brand, Clemens, and Seifter, Norbert. "Eigenvalues and domination in graphs." Mathematica Slovaca 46.1 (1996): 33-39. <http://eudml.org/doc/31832>.
[8] A. A. Al-Shargabi and A. Selmi, "Social Network Analysis and Visualization of Arabic Tweets During the COVID-19 Pandemic," in IEEE Access, vol. 9, pp. 90616-90630, 2021, doi: 10.1109/ACCESS.2021.3091537.
[9] S. Zhang, S. Yang and H. Yang, "Statistical Analysis of Spatial Network Characteristics in Relation to COVID-19 Transmission Risks in US Counties," 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 2278-2281, doi: 10.1109/EMBC46164.2021.9629892.
[10] B. Jiang, J. Yin and S. Zhao, "Characterizing the human mobility pattern in a large street network", Physical Review E, vol. 80, no. 2, pp. 021136, 2009.
[11] T. Rahman et al., "Development and Validation of an Early Scoring System for Prediction of Disease Severity in COVID-19 Using Complete Blood Count Parameters," in IEEE Access, vol. 9, pp.
[12] Bailey N T J. The Mathematical Theory of Infectious Diseases and Its Applications. New York: Hafner Press.1975.
[13] Anderson R M, May R M. Infectious Diseases in Humans.Oxford: Oxford University Press, 1992.
[14] Diekmann O, Heesterbeek JAP. Mathematical Epidemiology of Infectious Disease:Model Building. Analysis and Interpretation. New York: John Wiley & Son publisher, 2000.
[15] Newman MEJ, Watts D J. Scaling and percolation in the small-world network model. Phys. Rev. E, 1999, 60: 7332-7342.
[16] Moukarzel C F. Spreading and shortest paths in systems with sparse long-range connections. Phys. Rev. E, 1999, 60(6): R6263.
[17] Yang X S. Fractals in small-world networks with time-delay. Chaos, Solitons and Fractals. 2002, 13: 215-219.
[18] Li C G, Chen G. Local stability and Hopf bifurcation in small-world delayed networks. Chaos, Solitons and Fractals. 2004, 20: 353-361.
[19] Friedler, Miroslav. "Algebraic connectivity of graphs." Czechoslovak Mathematical Journal 23.2 (1973): 298-305. <http://eudml.org/doc/12723>.

Qi Xu is an undergraduate student at School of Mathematical Sciences, University Of Electronic Science And Technology Of China. He will receive a double undergraduate degree in 2023, major in Mathematics and Applied Mathematics and Computer Science and Technology. He is interested in further studies in the field of mathematics, and devotes in the research work that combines computer engineering and applied mathematics.