Research article

Dynamic items delivery network: prediction and clustering

Mokhammad R. Yudhanegara, Sapto W. Indratno*, RR.Kurnia N. Sari

Keywords: Mathematics Spectral Dynamic network Predictive distribution

ABSTRACT

Items delivery companies generally use a model to minimize delivery costs. From a mathematical perspective, the model is an objective function that involves constraints. Meanwhile, from a practical point of view, these constraints include aspects that affect item delivery, for example, delivery zones, number of delivery vehicles, vehicle capacity, trip routes, etc. However, the models built so far have not paid attention to changes in road density. This aspect can result in a nonoptimal delivery model, which results in not a minimum delivery cost. For this reason, this paper discusses how to divide zones using the clustering method and predict changes in the shipping zone of a dynamic network using predictive distribution. So, the model can work optimally if the delivery zones and delivery strategies are suitable.

1. Introduction

At present, items delivery service companies are competing to get the best model for the delivery process. In general, the problem structure in the items delivery process is modeled in an objective function to minimize distance, delivery time, and total delivery costs. For example, in the Traveling Salesman Problem (TSP), the model is used to find the shortest route in items delivery. Then the Capacitated Vehicle Routing Problem (CVRP), the model is used to find the shortest routes by paying attention to the delivery vehicle's capacity. For both TSP and CVRP, the ultimate goal is to minimize delivery costs (Dror and Trudeau, 1990; Reinelt, 2012).

Recently, many researchers have developed TSP and CVRP models to improve the items delivery according to existing problems. One of the most sophisticated models is the model of items delivery using an online algorithm design (Zhang et al., 2018). The model is online scheduling of order picking and delivery with multiple zones and limited vehicle capacity. The objective function of this model is to minimize the makespan and the total delivery cost.

The illustration of the items delivery is presented in Figure 1. In this paper, we do not specifically discuss the objective function of the model. However, we illustrate that by knowing the right items delivery zone, a model will have optimal results.

This paper's contribution is to provide input on the delivery optimization models, it’s an one of Zhang et al. (2018) for setting multiple zones through dynamic network clustering. Why a dynamic network? Because it is assumed that the network weight structure (in this case is the density of road sections) is always changing, so the network conditions not to be static. For example, these changes are caused by congestion due to accidents, hour, broken bridges, holidays, etc.

The problems of a delivery network can be overcome by using network clustering. Events that change the weight of a network result in clusters/subnets (zones) changes. So we need a distribution to describe the dynamic situation of the item delivery network. Thus, a predictive distribution function is necessary to predict the resulting zone. We illustrate the problems into a framework of the solving approach in Figure 2.

Determining the number of clusters of data is an important problem (Mooi and Sarstedt, 2011; Yudhanegara and Lestari, 2019). The clustering method used must meet the requirements. At the beginning of the study, the clustering of networks used the Markov cluster method. Still, it was not suitable because the number of clusters could not be determined (naturally occurring), making it difficult to control. In some instances, ineffective clusters were also found, such as unbalanced cluster members, for example, too few members of the one cluster while the other cluster were too large (Yudhanegara et al., 2020a).

The spectral bisection method can solve cluster management's problem simply based on the Laplacian matrix originating from a connected network. The formation of clusters produced is based on eigenvectors with the second smallest eigenvalues (Chung, 1997), called the Fiedler vector, which corresponds with nodes on the network. A connected network is a necessary condition in the spectral bisection algorithm. Because of the network is not connected, resulting in the Laplacian matrix obtained has an algebraic multiplicity of the eigenvalues \( \lambda = 0 \) as...
much as network components, so \( \lambda_2 = 0 \). In contrast, the spectral bisection method must be the second smallest eigenvalue \( \lambda_2 \neq 0 \) (Elsner, 1997). If a network is found that is not connected, then the clustering is done separately from each network that is connected.

The items delivery network can be represented as a simple network (Yudhanegara et al., 2020a). A network consists of a group of nodes connected by edges (Taha, 2016). Furthermore, network \( G = G(V, E) \) is a set of pairs \((V, E)\) where \( V \) is a set of finite and non-empty nodes \( V = \{v_0, v_1, \ldots, v_n\} \), and \( E \) is a set of pairs between two nodes, called edge, \( E = \{e = (v_i, v_j) : v_i, v_j \in V, i \neq j\} \).

Nodes \( v_i \) and \( v_j \) are said to be adjacent if there is an edge between \( v_i \) and \( v_j \), or there is a pair \((v_i, v_j) \in E\) where \( v_i, v_j \in V \) (Diestel, 2017). \( G = G(V, E, g) \) is called a weighted network \( G \) on \( V \), which the weight is represented by \( (V, E, g) \) the function \( g \) that maps pairs of each element \( V \) to non-negative real numbers. For example, \( G(V, E, g) \) is a weighted network with the number of nodes \( |V| = n \), the adjacency matrix of \( G \) is denoted \( M_G \) with the entry in the provisions of Eq. (1) as follows

\[
(M_G)_{ij} = \begin{cases} g_e & \text{if } e = (v_i, v_j) \in E, \ i \neq j \\ 0 & \text{elsewhere}. \end{cases} \tag{1}
\]

The term of network clustering in applied mathematics is a network partition, i.e., the division of the network into more than one subnet. To partition the network \( G(V, E, g) \) into \( h \) disjoint subsets, the sum of nodes in \( h \) subsets should be equal while the sum of the edges weights between the subsets is minimized. For example, to obtain \( h \) partition \((V_1, V_2, \ldots, V_h)\) of \( V \) i.e.

\[
\bigcup_{i=1}^{h} V_i = V, V_i \cap V = \emptyset, i \neq j \tag{2}
\]

such that

\[
\sum_{\forall v_i, v_j \in V, i \neq j} g_e = (v_i, v_j) \in E, i \neq j \tag{3}
\]

minimized among all possible partitions of \( V \). Eq. (3) is also referred to as cut-size. So, the central concept of network clustering is finding the smallest subset of edges when cut, separating the network into separate subsets. The process is called searching for an edge separator.

In general, a network can be grouped or partitioned into subnets consisting of several similar characteristics (Newman, 2004). Furthermore, this group is referred to as a network cluster. Network cluster is a partition of a set of nodes, i.e., each node is inserted into one cluster. The linkages between nodes and similarity characteristics can form a cluster that describes a group of nodes that interact with each other (Yudhanegara et al., 2020b).

There are many edges in a, but between clusters, there are fewer edges (Van Dongen, 2000). A network is said to have a cluster structure if the network nodes can be easily grouped into clusters so that each set of nodes is densely connected within the cluster and the gap between the clusters (Newman, 2004; Newman and Girvan, 2004). In other words, nodes pairs are more likely to be connected if they are members of the same cluster, and less likely to be connected if they are different clusters (Emmons et al., 2016).

Figure 1. Illustration of items delivery process.

Figure 2. The framework of the solving approach.
In this paper, the simulation of the clustering method is done through an illustrative example of 129 nodes on a weighted network. Nodes represent locations, while edges represent roads or paths that connect between locations. The simulation shows that clustering uses spectral bisection on a dynamic item delivery network, providing information for item delivery companies to consider the types of vehicles and items to be sent at specific locations. Through this procedure, cluster changes can be identified quickly and easily based on recent network.

Spectral bisection has been widely applied as a clustering method on various types of networks (Uruschel and Zikatanov, 2014). However, in this case, it is believed to be the first application to identify cluster changes in dynamic items delivery network with two clusters and four clusters outputs. The spectral bisection method can be used on extensive networks, so it is useful in planning and designing a global items delivery network.

The concept, plot, and algorithm of the spectral bisection method are discussed in this paper’s Materials and methods section. The result and discussion section discuss the clustering process and results. The results of the study are briefly described in the Conclusions section. The sources used in this article can be seen in the Reference section.

2. Materials and methods

2.1. Laplacian matrix

Suppose that G is a weighted network, and \( M_G \) is an adjacent matrix of G with size \( n \times n \). Then \( D_G \) is a diagonal matrix of G with the entry in Eq. (4),

\[
(D_G)_{ij} = \begin{cases} \sum_{t=1}^{n} (M_G)_{ij}, & i = j \\ 0 & \text{otherwise} \end{cases}
\]  

(4)

Next, \( L_G = D_G - M_G \) is a Laplacian matrix of G with an entry

\[
(L_G)_{ij} = \begin{cases} -g_{ee}, & (v_i, v_j) \in E, \ i \neq j \\ \sum_{e \in E(i)} g_{ee}, & (v_i, v_j) \in E, \ i = j \\ 0 & \text{otherwise} \end{cases}
\]  

(5)

Based on the properties of the \( L_G \) (Elsner, 1997), i.e., real, symmetric matrices, all eigenvalues \( \lambda_i \), \( i = 1, 2, \ldots, n \) are non-negative real numbers, where \( \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n \), and there are \( n \) vectors unit \( \phi_i, \ i = 1, 2, \ldots, n \), which are orthogonal to each other.

Furthermore, \( \phi_i \) is an eigenvector such that

\[
L_G \phi_i = \lambda_i \phi_i, \ i = 1, 2, \ldots, n.
\]  

(6)

The orthogonal, it is mean \( \phi_i^T \phi_j = 0 \), \( i \neq j \), \( v_i, v_j \in V \) and the unit vector \( \phi_i^T \phi_i = 1, \ v_i \in V \). To find the eigenvalue and eigenvector from the Laplacian matrix obtained by Rayleigh quotient (Chung, 1997; Spielman, 2015) namely

\[
\phi_i^T L_G \phi_i = \phi_i^T D_G \phi_i - \phi_i^T M_G \phi_i = \lambda_i \phi_i^T \phi_i = \lambda_i,
\]  

(7)

so \( \phi_i \) is an eigenvector corresponding to the eigenvalue \( \lambda_i \).

2.2. The second smallest eigenvalue

Based on the properties of the \( L_G \) (Elsner, 1997), that \( \phi_1 = 1 = (1, 1, \ldots, 1)^T \) is the eigenvector with the smallest eigenvalue \( \lambda_1 = 0 \) with the algebraic multiplicity is the number of components of the network G. So, specifically \( \lambda_2 \neq 0 \) if and only if the G is a connected network. Furthermore, numerically to obtain the second smallest eigenvalue \( \lambda_2 \) of matrix \( L_G \) is

\[
\min \{ \lambda_i \mid \phi_i^T \phi_i = 1, i = 2, 3, \ldots \}
\]  

(8)

(Pothen et al., 1990). From Eq. (6), so we have \( L_G \phi_2 = \lambda_2 \phi_2 \), where \( \phi_2 \) is the eigenvector which corresponds to the second smallest eigenvalue \( \lambda_2 \), it is called Fiedler vector.

2.3. Algorithm of spectral bisection

The spectral bisection method is a clustering method by partitioning the network into two clusters. The partition of the cluster refers to the Fiedler vector with the algorithm (Elsner, 1997; Van Driessche and Roose, 2006), i.e.

1. Determine the Laplacian \( L_G \) of G.
2. Find the Fiedler vector \( \phi_2 \) from the second smallest eigenvalue \( \lambda_2 \).
3. Calculate the median \( m_{\phi_2} \) of all components \( \phi_2 \).
4. Select \( V_1 = \{ v_i \in V : \phi_2 < m_{\phi_2} \} \) and \( V_2 = \{ v_i \in V : \phi_2 > m_{\phi_2} \} \) and if some components are equal to \( m_{\phi_2} \), then distribute the appropriate nodes so that they are balanced.

As a guarantee that the resulting subnet is connected, it uses Elsner’s theorem.

Theorem 1 (Elsner, 1997), let G be a connected network and let \( \phi_2 \) be its Fiedler vector. For any given real number \( r \geq 0 \), define \( V_1 = \{ v_i \in V : \phi_2 \geq r \} \), the subnet induced by \( V_1 \) is connected. Similarly for a real number \( r \leq 0 \), the subnet induced by \( V_2 = \{ v_i \in V : \phi_2 \leq r \} \) is also connected.

Proof
Based on Theorem (Fiedler, 1975), let G be a finite connected network, every edge \( (v_i, v_j), i, j \in \mathbb{N} \) of which a positive number \( c_{ij} \). Let \( y = (y_i) \) be a characteristic valuation of G. For any \( r \geq 0 \), let \( M(r) = \{ v_i \in \mathbb{N} : y_i + r \geq 0 \} \) then the subnet \( G(r) \) induced by \( G \) on \( M(r) \) is connected.

Denote by B the symmetric matrix \( (b_{ij}), i, j \in \mathbb{N} \) define by \( b_{ij} = c_{ij} \) if \( i \neq j \) and \((i, j) \in E, b_{ij} = 0 \) if \( i = j \) and \((i, j) \notin E, b_{ij} = -\sum_{j \neq i} b_{ij} \). Since \(-\langle Bx, x \rangle = \sum_{(i,j) \in E} c_{ij}(x_i - x_j)^2 \), we have \( B = -L_G \) where \( L_G \) is Laplacian matrix.

On the other hand, the off-diagonal part of B being nonnegative, \( B + \sigma I \) is nonnegative for sufficiently large \( \sigma \). The eigenvectors of \( L_G \) are identical with those of \( B + \sigma I \) and to the second smallest eigenvalue of \( L_G \) corresponds the second largest eigenvalue of \( B + \sigma I \). Based on Corollary (Fiedler, 1975), \( y + r \) where \( y \) is the vector of characteristic valuation of G, has the property that the submatrix of \( B + \sigma I \) with indices in \( M(r) \) is irreducible. Thus, the subnet \( G(r) \) is connected.

Corollary (Fiedler, 1975), let \( A \) be an \( n \times n \) nonnegative irreducible, symmetric matrix with eigenvalue \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \). Let \( \phi_1 = (\phi_{1i}) \) be an eigenvector corresponding to \( \lambda_1 \) and \( \phi_2 = (\phi_{2i}) \) corresponding to \( \lambda_2 \). Then for any \( \alpha \geq 0 \), submatrices \( A(M_n) \) is irreducible where \( M_n = \{ i \in \mathbb{N} : \phi_{2i} > \alpha \phi_{1i} \} \).

3. Results and discussion

3.1. Predictive distribution

Let \( X = (X_1, X_2, \ldots, X_m) \) be a random vector that stating the number of vehicles that might pass each road at \( \theta \)th day, where \( X_2 \) is a random variable stating the number of vehicles pass \( \theta \)th road at \( \theta \)th day. Then assuming \( X_2 \) follows a multinomial distribution with the parameter \( \theta \) (probability), \( X_i \sim \text{Mult}(\theta_1, \theta_2, \ldots, \theta_m, n) \), the probability mass function (pmf) can be shown in Eq. (9) as follows:

\[
p(x; \theta) = \frac{n!}{\prod_{i=1}^{m} x_i!} \prod_{i=1}^{m} \theta_i^{x_i}
\]  

(9)
where \( \mathbf{x} = [x_1, x_2, \ldots, x_n]^{T}, \) \( m \geq n \in \mathbb{N}, \) \( \mathbf{\theta} = [\theta_1, \theta_2, \ldots, \theta_m]^{T}, \) dan \nabla \sum_{i=1}^{m} \theta_i = 1, \theta_i > 0, \) \( i = 1, 2, \ldots, m. \)

Suppose that we have the \( k \) samples of random vector \( \mathbf{X}. \) Furthermore, to find the maximum parameter estimator \( \hat{\theta}_i, \) it uses maximum likelihood estimation (MLE) method (Hogg et al., 2005), by first forming the likelihood function of Eq. \( (9), \) namely

\[
L(\theta) = \prod_{i=1}^{k} \left( \frac{n!}{\prod_{i=1}^{m} x_i !} \prod_{i=1}^{m} \theta_i^{x_i} \right)
\]

(10)

where \( \sum_{i=1}^{m} \theta_i = 1, i = 1, 2, \ldots, m \) and \( \sum_{i=1}^{m} x_i = n, \) \( \forall \mathbf{t} = 1, \ldots, k. \) So the log-likelihood function

\[
\ln L(\theta) = \ln \left( \prod_{i=1}^{k} \left( \frac{n!}{\prod_{i=1}^{m} x_i !} \prod_{i=1}^{m} \theta_i^{x_i} \right) \right)
\]

\[
= \sum_{i=1}^{k} \ln n! - \sum_{i=1}^{k} \sum_{t=1}^{m} \ln x_t ! + \sum_{i=1}^{k} \sum_{t=1}^{m} x_t \ln \theta_t,
\]

and by using Lagrange multiplier, the following objective function is obtained

\[
H(\theta, \gamma) = \sum_{i=1}^{k} \ln n! - \sum_{i=1}^{k} \sum_{t=1}^{m} \ln x_t ! + \sum_{i=1}^{k} \sum_{t=1}^{m} x_t \ln \theta_t + \gamma \left( 1 - \sum_{i=1}^{k} \theta_i \right)
\]

(11)

where \( \sum_{i=1}^{m} \theta_i = 1, i = 1, 2, \ldots, m \) and \( \sum_{i=1}^{m} x_t = n. \) Furthermore, if the parameters \( \theta \) and \( \gamma \) are seen as variables, the next step is to determine the maximum estimator by differentiating the objective function \( H(\theta, \gamma) \) with respect to \( \theta \) and \( \gamma. \)

Parameter estimator \( \hat{\theta} \) from multinomial distribution is obtained by maximizing \( H(\theta, \gamma) \) using the differential method, namely

\[
\frac{\partial H(\theta, \gamma)}{\partial \theta_i} = 0,
\]

(12)

So, we obtain for \( i = 1 \to \hat{\theta}_1 = \frac{\sum_{i=1}^{m} x_t}{\gamma}. \)

for \( i = 2 \to \hat{\theta}_2 = \frac{\sum_{i=1}^{m} x_t}{\gamma}, \)

for \( i = m \to \hat{\theta}_m = \frac{\sum_{i=1}^{m} x_t}{\gamma}, \)

generally,

\[
\hat{\theta}_i = \frac{\sum_{i=1}^{m} x_t}{\gamma}, \quad i = 1, 2, \ldots, m.
\]

(13)

Then estimate the parameter \( \gamma \) by applying

\[
\frac{\partial H(\theta, \gamma)}{\partial \gamma} = 0,
\]

(14)

to get

\[
\sum_{i=1}^{m} \theta_i = 1.
\]

(15)

We substitute Eqs (13) into (15), to obtain

\[
1 = \sum_{i=1}^{m} \theta_i = \frac{1}{\gamma} \sum_{i=1}^{m} x_t
\]

It follows that

\[
\sum_{i=1}^{k} n = \gamma
\]

which gives

\[
k \gamma = \sum_{i=1}^{k} x_t / \theta_i
\]

Therefore

\[
\gamma = \sum_{i=1}^{k} x_t / \theta_i
\]

(16)

Because the network is dynamic, it is necessary to predict the density of each road for the future through the predictive distribution. And then, the conditional probability distribution for observation \( x_{t+1} \) given observations \( x_1, x_2, \ldots, x_t \) is

\[
p(x_{t+1} | D) = \int p(x_{t+1} | \theta) f(\theta | D) d\theta
\]

(17)

where \( D = (x_1, x_2, \ldots, x_t). \) The \( f(\theta | D) \) is a posterior distribution, i.e.,

\[
f(\theta | D) = \frac{f(D | \theta)}{\int_{\Omega} f(D | \theta) d\theta}
\]

(18)

\( f(D, \theta) \) are joint probability function of \( X \) and \( \Theta, \) where

\[
f(D, \theta) = p(D | \theta) p(\theta | f(\alpha) / p(\alpha))
\]

\[
= \frac{f(\theta, \alpha) p(D | \theta)}{f(\alpha) p(\theta)}
\]

\[
= f(\theta | m \sum_{i=1}^{m} p(x_i | \theta)),
\]

(19)

where \( \alpha = [\alpha_1, \alpha_2, \ldots, \alpha_m] \) (Blei et al., 2003). For \( \theta_1, \theta_2, \ldots, \theta_m \sim \text{Dir}(\alpha_1, \alpha_2, \ldots, \alpha_m), \) where \( \text{Dir} \) is Dirichlet distribution, and \( x_1, x_2, \ldots, x_m \sim \text{Mult}(\theta_1, \theta_2, \ldots, \theta_m) \), then the posterior distribution is

\[
f(\theta | D) \sim f(D, \theta)
\]

\[
f(D, \theta) = f(\theta | m \sum_{i=1}^{m} p(x_i | \theta))
\]

\[
\propto \prod_{i=1}^{m} \theta_i^{x_i} \prod_{i=1}^{k} \prod_{i=1}^{m} \theta_i^{x_i}
\]

\[
= \prod_{i=1}^{m} \theta_i^{x_i} \prod_{i=1}^{k} \prod_{i=1}^{m} \theta_i^{x_i}
\]

\[
= \prod_{i=1}^{m} \theta_i^{x_i} \prod_{i=1}^{k} \prod_{i=1}^{m} \theta_i^{x_i}
\]

(20)

where \( \alpha_i = \alpha_i + \sum_{i=1}^{k} x_i, \) then \( f(\theta | D) \sim \text{Dir}(\alpha_i, \alpha_2, \ldots, \alpha_m). \) So that the predictive distribution is

\[
p(x_{t+1} | D) = \int p(x_{t+1} | \theta) f(\theta | D) d\theta
\]
Figure 3. Maps of Bandung City as the Item Delivery Network. 
(Source: google.co.id/maps/place/Bandung).

\[
E[X] = \sum_{x=0}^{n} x p(x) = \sum_{x=0}^{n} x \frac{n!}{\prod_{i=1}^{m} x_i!} \prod_{i=1}^{m} \gamma(x_i + a_i) / \Gamma(n + \sum_{i=1}^{m} a_i) = n\theta(1) = n\theta,
\]

Then for dirichlet, \( \theta \sim \text{Dir}(\alpha_1, \alpha_2, \ldots, \alpha_m) \), letting \( a_0 = \sum_{i=1}^{m} a_i \) and

\[
E[X_i] = \frac{\alpha_i}{\alpha_0}.
\]

\[
\sum_{x=0}^{n} x p(x) = \sum_{x=0}^{n} x \frac{n!}{\prod_{i=1}^{m} x_i!} \prod_{i=1}^{m} \gamma(x_i + a_i) / \Gamma(n + \sum_{i=1}^{m} a_i) = n\theta(1) = n\theta,
\]

Then for dirichlet, \( \theta \sim \text{Dir}(\alpha_1, \alpha_2, \ldots, \alpha_m) \), letting \( a_0 = \sum_{i=1}^{m} a_i \) and

\[
E[X_i] = \frac{\alpha_i}{\alpha_0}.
\]
\begin{align*}
\sum_{i=1}^{m} \theta_i &= 1, \text{ for } \theta_1, \text{ we get } \\
E[\theta_1] &= \int \cdots \int \left( \prod_{i=1}^{m} \Gamma(\alpha_0) \right) \prod_{i=2}^{m} \theta_1^{\alpha_i-1} \frac{1}{\Gamma(\alpha_0 + 1)} d\theta_1 \cdots d\theta_m \\
&= \frac{\Gamma(\alpha_0) \Gamma(\alpha_1 + 1) \prod_{i=2}^{m} \Gamma(\alpha_i) \prod_{i=1}^{m-1} \theta_i^{\alpha_i-1} \frac{1}{\Gamma(\alpha_0 + 1)}}{\Gamma(\alpha_1) \prod_{i=2}^{m} \Gamma(\alpha_i) \prod_{i=2}^{m} \theta_i^{\alpha_i-1}} \\
&= \frac{\Gamma(\alpha_0) \Gamma(\alpha_1 + 1) \prod_{i=2}^{m} \Gamma(\alpha_i) \prod_{i=1}^{m-1} \theta_i^{\alpha_i-1} \frac{1}{\Gamma(\alpha_0 + 1)}}{\Gamma(\alpha_1) \prod_{i=2}^{m} \Gamma(\alpha_i) \prod_{i=2}^{m} \theta_i^{\alpha_i-1}} \\
&= \frac{\Gamma(\alpha_0) \Gamma(\alpha_1 + 1) \prod_{i=2}^{m} \Gamma(\alpha_i) \prod_{i=1}^{m-1} \theta_i^{\alpha_i-1} \frac{1}{\Gamma(\alpha_0 + 1)}}{\Gamma(\alpha_1) \prod_{i=2}^{m} \Gamma(\alpha_i) \prod_{i=2}^{m} \theta_i^{\alpha_i-1}} \\
&= \frac{\alpha_1}{\alpha_0},
\end{align*}

Using the same arguments, we obtain \(E[\theta_i] = \frac{n}{\alpha_i}, i = 2, 3, \ldots, m\). Let \(\alpha_0 = \sum_{i=1}^{m} \alpha_i \) and \(\theta_1 = \frac{n}{\alpha_0}, i = 1, 2, \ldots, m\), then the expected number of times the outcome \(i\) was observed over \(n\) trials (Eq. (19)) is

\[
E[X_i] = n \theta_i = n \frac{\alpha_i}{\alpha_0}.
\]

### 3.2. Algorithm for recursive of spectral bisection

The algorithm of recursive of spectral bisection, i.e.,
Figure 6. Network clustering in 1st day until 30th day with two zones.

Figure 7. Network clustering in 1st day until 30th day with four zones.
1. Determine the Laplacian matrix $L_G$ of network G.
2. Find the Fiedler vector $\varphi_2$ from the second smallest eigenvalue $\lambda_2$.
3. Calculate the median $m_{\varphi_2}$ of all components $\varphi_2$.
4. Select $V_1 = \{i \in V : \varphi_2 < m_{\varphi_2}\}$ and $V_2 = \{i \in V : \varphi_2 > m_{\varphi_2}\}$ and if some components are equal to $m_{\varphi_2}$, then distribute the appropriate nodes so that they are balanced.
5. Determine the Laplacian matrix $L_{subG}$ of subnetwork G.
6. Determine cluster members according to stages 2, 3, and 4.

Through this algorithm, clustering is carried out in stages. The first stage is clustering a network into two clusters, then clustering from each cluster, so that four clusters are obtained from a network.

3.3. Simulation

In this simulation, the output generated from network clustering is two clusters and four. So that the clustering process with the algorithm for recursive of spectral bisection is done by repetitive division. There are 1738665 vehicles passing through roads at any $t^{th}$-time in Bandung city (source: jabar.bps.go.id). Suppose the location is represented as nodes, and the road is represented as an edge, hereinafter referred to as the items delivery network. Then the clusters of the items delivery network are determined by using an algorithm for recursive of spectral bisection.

Subsequently identified changes in cluster members at the $t^{th}$ day, $t = 1, 2, 3, \ldots$. The network G consists of 129 nodes ($V = \{v_0 = 0, v_1 = 1, v_2 = 2, \ldots v_{128} = 128\}$) and 260 edges ($E = \{e_1, e_2, \ldots e_{260}\}$). In the network display, nodes are labeled 0, 1, 2, ..., 128, and the edges are labeled as the number of vehicles passing through the side/road as in Figures 4 and 5. For the pmf of $X_i \sim \text{Mult} (\theta_1, \theta_2, \ldots, \theta_{260})$, 1738665 is

$$p(x; \theta) = \frac{1738665}{\prod_{i=1}^{260} \theta_i^{x_i}} \prod_{i=1}^{260} \theta_i^{x_i} \quad (21)$$

where $\sum_{i=1}^{260} \theta_i = 1, i = 1, 2, \ldots, 260$ and $\sum_{i=1}^{260} x_i = 1738665$. The parameter $\theta$ is generated randomly from the standard uniform distribution (U(0,1)).

For programming, we use python 3.8 software. In the output, clustering results are given two network images, which are clustered two clusters and four clusters. Items delivery network is taken from the map of Bandung city, see Figure 3. The network structure can be seen in Figures 4 and 5.

Furthermore, the results of network clustering on the first day can be seen in Figures 4 and 5. Meanwhile, other clusters on other days can be seen in Figures 6 and 7.

Figures 6 and 7 are historical data to predict the next network. From the historical data, we get the parameter values. The parameter values are used to predict the next network. The values of the parameter estimation $\hat{\theta}_i; i = 1, 2, \ldots, 260$ are obtained from Eq. (16) with $k = 30$ and $n = 1738665$. The results are in Table 1.

Based on predictive distribution, by using Eq. (19) we have network clustering in 31$^{th}$ day, see Figures 8 and 9.

3.4. Multinomial goodness-of-fit test

To determine how good the fit between the frequency of occurrence of observations in an observed sample and the expected frequencies obtained from the hypothesized distribution, we assume apply the following hypothesis test:

$H_0$: The distribution is not different from what is expected, $H_1$: The distribution is different from what is expected.

Table 1. The values of the parameter estimation.

| $\theta_i$ | $E[\theta_i]$ | $E[\theta_i^2]$ | $E[\theta_i^3]$ | $E[\theta_i^4]$ |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| $0.0055202764$ | $0.0003764512$ | $0.0000437951$ | $0.0000048759$ | $0.0000050959$ |
| $0.0044048457$ | $0.0002388804$ | $0.0000037926$ | $0.0000023859$ | $0.0000024739$ |
| $0.0037403782$ | $0.0001520173$ | $0.0000019719$ | $0.0000014869$ | $0.0000015842$ |
| $0.0052639646$ | $0.0001190487$ | $0.0000014687$ | $0.0000011486$ | $0.0000011556$ |

Here the chi-squared test for a goodness-of-fit test is used. A goodness-of-fit test between observed and expected frequencies is based on the quantity

$$\chi^2 = \sum_{i=1}^{m} \frac{(x_i - \hat{x}_i)^2}{\hat{x}_i} \quad (22)$$

where $\hat{x}_i = E[X_i] = n \theta_i$, $a_i = \sum_{j=1}^{n} x_i^j$, $0.0037403782\times n \theta_i$, and $m \theta_i = 1$.

From goodness-of-fit test for observed of 31$^{st}$ day and the expected value based on Eq. (21), we have results in Table 2.

As the computed $p$-value is greater than the significance level alpha $0.05$, one cannot reject the null hypothesis $H_0$. So the distribution is not
It means that the delivery zones used are based on the last data (30th day), then we have not to do the new optimization.

4. Discussion

The purpose of this simulation is to analyze the behavior of cluster members if the number of vehicles that pass-through roads (represented by the edge on the weighted network) changes every time. By setting the parameters $\theta$ and $n$ on the multinomial distribution, and generating data at different days ($t = 1, 2, ..., 30$), the clustering output is obtained in Figures 6 and 7. From the clustering results, it can be seen that each cluster member is able to move from one cluster to another.

Table 2. Chi-squared test sample of 30th day and the expected value.

|                      |     |
|----------------------|-----|
| Chi-square (Observed) | 293,928 |
| Chi-square (Critical) | 297,538 |
| DF                   | 259  |
| p-value              | 0,067 |
| alpha                | 0,05  |

Each clustered image using the algorithm for recursive of spectral bisection displayed two images, the first output with two clusters (zones), and the second output with four clusters (zones). The aim is to make a difference in the cluster results, which the we can analyze as needed.
In the simulation of the density of roads that follow this multinomial distribution, a predictive function is given to get the weight of edges. If historical data are given, the results can be obtained using Eq. (19). This case indicates that the network is dynamic. So that we can further analyze cluster changes produced. One of the benefits of the simulation results of item delivery network clustering can determine the sum of delivery costs in a particular zone.

See Figure 8, the network is partitioned into 4 clusters. For example, zone 1 (purple nodes), is a region with very dense road conditions, so that the journey taken by vehicle drivers will experience a slowdown. Because of the slowdown, this results in vehicles requiring much fuel. On the other hand, because the delivery items sent cannot exceed the specified time limit, so more delivery vehicles are needed in this zone 1. Furthermore, in Figure 8, zona 2 (yellow nodes) is a region whose road conditions are not as dense as the zona 1. To anticipate this problem, delivery service companies may have to combine shipments for zona 1 and zona 2 to minimize costs incurred for items delivery.

5. Conclusion

In the face of dynamic network cases with the density of roads that follow a multinomial distribution, to predict the network is done through predictable distribution. In this case, the prediction of the number of vehicles passing each road in the next time $t+1$ if given a data set, then $X_{t+1}|X_t, X_{t-1}, \ldots, X_0 \sim p(\cdot|X_0; D)$. In a network with a large number of nodes, clustering with a recursive of spectral bisection method with $2^c, c \in \mathbb{N}$, cluster outputs are certainly more varied than the two or four cluster outputs. Thus, the predictions and clustering presented in this paper are very useful in the process of item delivery for users, especially companies in the field of item delivery.

In the future (open problems), it will apply the same method to data from different companies to identify whether there is a statistical problem or not. So that if any problems are found, it will solve them immediately and publish the results. It can be helpful for those in need.

Declarations

Author contribution statement

M. R. Yudhanegara: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

S. W. Indratno: Performed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

R. R. Kurnia N. Sari: Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement

Sapto Wahyu Indratno was supported by Peneltian, Pengabdian Masyarakat, dan Inovasi ITB (PPMI) ITB (grant number 197508041999031003).

Data availability statement

Data included in article/supplementary material/referenced in article.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2021.e06934.

Acknowledgements

The authors would like to thank P2MI Institut Teknologi Bandung, Indonesia for research funding, and the anonymous reviewers for providing constructive comments to improve the earlier version of the manuscript. The first author would like to thank LPDP, through the scheme of BUDI-DN, for providing the scholarship.

References

Blei, P.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. J. Mach. Learn. (3), 993–1022.
Chung, F.R.K., 1997. Spectral Graph Theory. American Mathematical Society, Providence, Rhode Island.
Diestel, R., 2017. Graph Theory, fifth ed. Springer, New York.
Dror, M., Trudeau, P., 1990. Split delivery routing. Nav. Res. Logist. 37, 383–402.
Elsner, U., 1997. Graph Partitioning. Technische Universität Chemnitz, Chemnitz.
Emmons, S., Kobourov, S., Gallant, M., Borner, K., 2016. Analysis of network clustering algorithms and cluster quality metrics at scale. PloS One 11 (7), 1–18.
Fiedler, M., 1975. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. Czech. Math. J. 4, 619–633.
Hogg, R.V., McKean, J.W., Craig, A.T., 2005. Introduction to Mathematical Statistics, sixth ed. Pearson, USA.
Mooi, E., Sarstedt, M., 2011. A Concise Guide to Market Research. Springer, Berlin.
Newman, M.E.J., 2004. Detecting community structure in networks. Eur. Phys. J. B 38, 321–330.
Newman, M.E.J., Girvan, M., 2004. Finding and evaluating community structure in networks. Phys. Rev. E 69 (2), 026113-1–026113-15.
Pottho, A., Simon, H.D., Liu, K.P., 1990. Partitioning sparse matrices with eigenvectors of graphs. SIAM J. Matrix Anal. Appl. 11 (3), 430–452.
Reinelt, G., 2012. TSPLIB: a traveling salesman problem library. ORSA J. Comput. 3, 376–384.
Spilman, D.A., 2015. Spectral Graph Theory. http://www.cs.yale.edu/homes/spelman/561.
Taha, H.A., 2016. Operations Research: an Introduction, tenth ed. Pearson, New Jersey.
Urschel, J.C., Zikatanov, L.T., 2014. Spectral bisection of graph and connectedness. Lin. Algebra Appl. 449, 1–16.
Van Dongen, S., 2000. Graph Clustering by Flow Simulation. University of Utrecht, Utrecht.
Van Driessche, R., Roose, D., 2006. Dynamic load balancing with a spectral bisection algorithm for the constrained graph partitioning problem. Int. Conf. High-Perform. Comput. Netw. 392–397.
Yudhanegara, M.R., Lestari, K.E., 2019. Clustering for multi-dimensional data set: a case study on educational data. J. Phys. Conf. 1280 (1-7), 042025.
Yudhanegara, M.R., Indratno, S.W., Sari, R.R.K.N., 2020a. Clustering for items distribution network. J. Phys. Conf. 1496 (1-8), 012019.
Yudhanegara, M.R., Indratno, S.W., Sari, R.R.K.N., 2020b. Clustering for items delivery using rule-k-means. J. Indones. Math. Soc. 26 (2), 185–191.
Zhang, J., Wang, X., Huang, K., 2018. On-line scheduling of order picking and delivery with multiple zones and limited vehicle capacity. Omega. Int. J. Manag. Sci. 79, 104–115.