Research article

A specific type of irregular ring-and-hub network structure and the average shortest distance of its rings

Tao Fu a, Chenguang Li b,∗, Long Wu a, Liling Zou c

a Economics and Management School, Beijing University of Technology, China
b Economics and Management School, North China University of Technology, China
c Faculty of Humanities and Social Sciences, Beijing University of Technology, China

ARTICLE INFO

Keywords:
Dorogovtsev-Mendes model
Ring-and-hub structure
Average shortest distance

ABSTRACT

Ring-and-hub network structure is very common in the real world, while that neighboring rings may sometimes share nodes or line segments makes this structure irregular. In this paper, we modify Dorogovtsev-Mendes model and its subsequent models to analytically estimate the average distance between nodes on the same ring in irregular ring-and-hub networks. In order to observe the accuracy of our modified model, we develop an algorithm to generate irregular ring-and-hub networks by computer. Then, we compare the analytic estimates with the practical values on those computer-generated and real networks. The results show that our modified model actually estimates the average shortest distance of its rings when only straight and U-shape paths between the start and end points are allowed. The accuracy of estimates for innermost several rings can be acceptable.

1. Introduction

With the flourishing of complex network research in the past decades, a series of typical network structure such as small-world [1] and scale-free [2] structure have been brought into the public eye. By scrutinizing them, many of their characteristics and related disciplines have been discovered and applied to different circumstances to solve practical problems. Among them, ring-and-hub structure where a hub takes the central position of the network and is surrounded by a ring containing a number of nodes or multiple such rings seems to be particularly common. This structure could be frequently observed in biological, traffic and social systems. Fig. 1 provides several networks with the ring-and-hub structure discussed by literature [3][4] and this paper, which cover directed and undirected, single-ring and multiple-ring, as well as regular and irregular cases. From these examples we can see that the central hub or area usually provides shortcuts for nodes on the ring around it, and hence may shorten the distance between those nodes. From the same perspective, those inner rings and the hub could also provide shortcuts to shorten the distance of nodes on outer rings. The average shortest distance is a fundamental structure indicator of complex networks, which has a universal relevance to other structure features such as centrality [5], fractality [6] and hierarchy [7], and meanwhile has been widely applied in the studies of random walks [8][9], synchronization [10][11] and complex network information diffusion [12][13]. Especially in papers focusing on ring-and-hub networks [3][4][14][15], to develop the analytical expression for the average shortest distance between nodes on the same ring was the most core and basic work.

Among those ring-and-hub network models, Dorogovtsev-Mendes model (DM model) [3] serves as the start point. This model altogether with its subsequent models [4][14][15] covered many specific cases of ring-and-hub networks in the real world. However, this kind of networks still have a level of diversity, some irregular ones such as examples in Fig. 1(e) and (f) have not yet been examined sufficiently. Fig. 1(e) presents the information diffusion network in a small sanitation team which has three hierarchies. In this system, people often share environmental sanitation information with those who do similar work in the adjacent blocks (they are in the same hierarchy of the network) in order to make their daily work go more smoothly. Meanwhile, information channels between different hierarchies are created to organize and coordinate sanitation work. For some very small blocks, a person may take the roles of both the middle manager and the direct operator, which leads to the irregularity of the ring-and-hub structure. Fig. 1(f) provides the ring road deployment of Baoding (a prefecture-level city in China), and from it we can see that the southern section of the outer ring road coincides with the inner ring road due to lacking of necessary

∗ Corresponding author.
E-mail address: lichenguang@ncut.edu.cn (C. Li).

https://doi.org/10.1016/j.heliyon.2022.e11470
Received 14 July 2022; Received in revised form 3 October 2022; Accepted 3 November 2022

2405-8440/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
space. As far as we know, for many cities all over the world (including Baoding), to construct a completely new and independent outer ring road may not be as economical as to adopt the design that the outer ring road occupies a part of the inner ring road to form a closed loop.

The networks in Fig. 1(e) and (f) mainly reflect a specific type of irregular ring-and-hub structure in which outer rings take use of some nodes or line segments of inner rings. We believe that this type of irregular ring-and-hub structure is even more common than regular ring-and-hub structure due to dozens of reasons not to form independent outer rings. We further decide to narrow our discussion on this structure to analytically estimate the average shortest distance between nodes on the same ring under some constraints, as in previous literature [3][4][11][15], this type of work took up absolute space and served as the basis for deducing other conclusions. It requires us to modify the previous models and correct the defect, and to develop an algorithm to generate irregular ring-and-hub networks by computer, and to give expressions for the average shortest distance between nodes on the same ring in such networks. Accomplishing these tasks also reflects the main contribution of this paper, and the rest of this paper is organized as follows: we first modify the previous models to increase the accuracy of estimates, and then bring in some new parameters to adapt the irregular ring-and-hub structure. In the next section, we examine our inference about the defect of the previous model by simulation on regular ring-and-hub networks, and then we develop an algorithm to generate irregular ring-and-hub networks by computer, and based on them, we observe the accuracy of our modified model. Then, we present an application of our modified model on a real network, and at last, we give the conclusion and prospect.

2. The modified model for the irregular ring-and-hub structure

2.1. The modification of models for the structure containing a single ring

Let us start with DM model [3], which assumes that there are \( n \) nodes on a ring, and each node is connected with two neighbors by edges of unit length. Then each node has a probability \( p \) to be connected with the hub by an undirected edge of length equal to 1/2. If those edges on the ring are oriented in one direction, this network is called directed DM network (see Fig. 1(a)). Otherwise, when all edges are undirected, it is called undirected DM network (see Fig. 1(b)).

Let \( P(l,m) \) denote the probability that the shortest distance between any two nodes on the ring is \( l \), given that their shortest distance along the ring is \( m \) (\( 1 \leq m \leq n-1 \)).

For the directed network, Dorogovtsev and Mendes [3] gave that:

\[
\begin{align*}
P(l < m, m) &= l p^l (1 - p)^{l-1} \\
P(l = m, m) &= 1 - p^l \sum_{i=1}^{m-1} i (1 - p)^{i-1}.
\end{align*}
\]

Based on Eqs. (1), the distribution of shortest distances \( l \) between a pair of nodes on the ring could be:

\[
P(l) = \frac{1}{n-1} \sum_{m=1}^{n-1} P(l, m)
\]

Based on Eq. (2), the average shortest distance \( \bar{l} \) can be written as follows:

\[
\bar{l} = \sum_{l=1}^{n-1} l P(l)
\]

\[
\bar{l} = \frac{1}{n-1} \left( \frac{2 - p}{p} n - \frac{3}{2} + \frac{2}{p} + (1 - p) p (n - 2 + \frac{3}{p}) \right).
\]

For the undirected network and an odd \( n \), they gave

\[
P(l = 1)) = \frac{2}{n-1} (1 + \frac{n-3}{2} p^2)
\]

\[
P(l > 2) = (2 + 4(l - 2)p + 2(l - 1)(2n - 4l - 3)p^2 - 2(2 - l)(2l - 2 - 1)p^3 + l(n - 2l - 1)p^4) \frac{(1 - p)^{l-1}}{n-1}.
\]

Then, Eqs. (4) could be substituted into \( \bar{l} = \sum_{l=1}^{n-1} l P(l) \) to calculate \( \bar{l} \).

Ashton et al. [14] took the hub's congestion cost into consideration. When a shortcut between two nodes on the ring passes through the hub, it may produce a congestion cost \( c \) which could be seen as an additional pass-length, and the total length of the shortcut turns into \( 1 + c \). For the directed model, they had

\[
P(l < c)) = \frac{1}{n-1} \left( \frac{1}{1 + c} + (l - c - 1)p + (n - 1 - l)(l - c)p^2(1 - p)^{l-1} \right)
\]

and

\[
\bar{l} = \frac{1}{p^2(n - 1)} \left( (1 - p)^{n-c} (3 + (n - 2 - c)p) + p(2 - 2c + 2n - (c - 1)(c - n)p - 3) + \frac{c(c - 1)}{2(n - 1)} \right)
\]

The analytic estimations for the undirected network could be obtained in a similar way. Those expressions were even more cumbersome and hence omitted by the authors.
By analyzing DM model [3] and its subsequent models [4][14], we find that their proposers mainly treated the value of \( l \) as an integer explicitly or implicitly. However, \( l \) does take a fractional value when the total length of the shortcut passing through the hub takes a fractional value. Based on that consideration, we design our modified model for DM networks. Let us use \( a \) to denote the length of those edges linking the hub and nodes on the ring. When \( a = 1/2 \), it is the same as the original DM model, and when \( a = 1 \), all edges of the network are of unit length. Let \( C_t = 2a + c \) denote the total length of the shortcut passing through the hub, where \( c \) and other variables remain the same meanings as defined previously. As \( C_t \) may take a fractional value, we let \( Z_c = [C_t] + 1 \), where \( [ ] \) indicates rounding down to the nearest integer.

For the directed network and when \( m < C_t \), we easily obtain

\[
P(l = m,m(m \leq C_t)) = 1.
\]

The case is a little complex when \( m > C_t \), and so we decide to deduce the general expression of \( P(l,m) \) by listing the expressions for small \( l \) and \( m \), just like what they did in Ref. [3] and [14]. We have

\[
P(l = C_t,m(m \leq Z_c)) = p^2;
\]

\[
P(l = Z_c,m(m \leq Z_c)) = 1 - p^2;
\]

\[
P(l = C_t,m(m < Z_c + 1)) = p^2 l
\]

\[
P(l = Z_c + 1,m(m < Z_c + 1)) = 2p^2(1 - p);
\]

\[
P(l = Z_c + 1,m(m < Z_c + 1)) = 1 - p^2 - 2p^2(1 - p);
\]

\[
P(l = C_t,m(m < Z_c + 2)) = p^2 l
\]

\[
P(l = Z_c + 2,m(m < Z_c + 2)) = 2p^2(1 - p);
\]

\[
P(l = C_t + 2,m(m < Z_c + 2)) = 3p^2(1 - p)^2;
\]

\[
P(l = C_t + 2,m(m < Z_c + 2)) = 1 - p^2 - 2p^2(1 - p) - 3p^2(1 - p)^3;
\]

... Thus, the general form of \( P(l,m) \) is

\[
P(l,m) = \begin{cases} 
1 & \text{if } l < C_t,
\end{cases}
\]

\[
P(l = C_t + k,m(k \geq 0),m(m > C_t + k,k \geq 0)) = (k + 1)p^2(1 - p)^k.
\]

\[
P(l = Z_c + k,m(k \geq 0),m(m > Z_c + k,k \geq 0)) = (1 - p)^k + 1 + p + k.
\]

Based on Eqs. (7), we have

\[
P(l = C_t + k,m(k \geq 0),m(m > C_t + k,k \geq 0)) = (k + 1)p^2(1 - p)^k.
\]

\[
P(l = Z_c + k,m(k \geq 0),m(m > Z_c + k,k \geq 0)) = (1 - p)^k + 1 + p + k.
\]

By scrutinizing Eqs. (8), we find that when \( C_t \) takes an integer value, we have \( Z_c = C_t + 1 \). At this moment, \( P(l) \) is the sum of the first and the second sub-equations (when \( l = C_t \)), or the sum of the second and the third sub-equations (\( l > C_t \)). For example, when \( a = 1/2 \) and \( c = 0 \), we have \( C_t = 1 \) and \( Z_c = 2 \). Substitute \( Z_c = 2 \) and \( k = 1 \) into the second sub-equation of Eqs. (8), meanwhile substitute \( k = 1 \), \( Z_c = 1 \) into the third sub-equation, and add them altogether, we could obtain Eq. (2) (from Ref. [3]). Otherwise, if we substitute \( C_t = 1 + c \), \( Z_c = 2 + c \), and \( k = 1 - c \) into the second sub-equation, meanwhile substitute \( k = 1 \), \( Z_c = 1 \) into the third sub-equation, and make the summation, we will get the second sub-equation of Eqs. (5) (from Ref. [14]).

Now based on Eqs. (8), we could calculate the average shortest distance between nodes on the ring as follows:

\[
I = \sum_{l} lP(l) = \frac{1}{2(n-1)p^2}(2(1-p)^{n-1} - (p^2(C_t - Z_c))(n - Z_c) + p^2(C_t + n - 3Z_c - 1) + 3p^2(C_t(p(n - Z_c + 2) - 2) - 4n(p - 1) + p(Z_c - 2)(Z_c + 1 + 8) - 6)).
\]

By substituting \( C_t = 1 \) and \( Z_c = 2 \) into Eq. (9), we obtain Eq. (3), and while we substitute \( C_t = 1 + c \) and \( Z_c = 2 + c \) into Eq. (9), we get Eq. (6). Hence, DM model [3] and its modification by Ashton et al. [14] could be regarded as a simplification of our modified model when \( C_t \) takes an integer value.

For the undirected network, we still list the expressions for small \( l \) and \( m \) to deduce the general expression of \( P(l,m) \). When \( m < C_t \), we still have

\[
P(l = m,m(m \leq C_t)) = 1.
\]

When \( m > C_t \), we have

\[
P(l = C_t),m(m = Z_c)) = p^2;
\]

\[
P(l = Z_c,m(m = Z_c)) = 1 - p^2;
\]

\[
P(l = C_t,m(m = Z_c + 1)) = p^2 l
\]

\[
P(l = Z_c + 1,m(m = Z_c + 1)) = 2p^2(1 - p);
\]

\[
P(l = Z_c + 1,m(m = Z_c + 1)) = 1 - p^2 - 2p^2(1 - p);\]

\[
P(l = C_t,m(m = Z_c + 2)) = p^2 l
\]

\[
P(l = Z_c + 2,m(m = Z_c + 2)) = 2p^2(1 - p);
\]

\[
P(l = C_t + 2,m(m = Z_c + 2)) = 3p^2(1 - p)^2;
\]

\[
P(l = C_t + 2,m(m = Z_c + 2)) = 1 - p^2 - 2p^2(1 - p) - 3p^2(1 - p)^3;
\]

... Thus, the general form of \( P(l,m) \) is

\[
P(l = m,m(m \leq C_t)) = 1
\]

\[
P(l = C_t),m(m > C_t)) = p^2 l
\]

\[
P(l = C_t + k,m(k \geq 0),m(m > C_t + k,k \geq 0)) = (k + 1)p^2(1 - p)^k
\]

\[
P(l = Z_c + k,m(k \geq 0),m(m > Z_c + k,k \geq 0)) = (1 - p)^k + 1 + p + k.
\]

Considering that \( m \) denotes the shortest distance along the ring between a pair of nodes on that ring, here we would like to use \( Z_c \) to denote the maximum of \( m \) for the undirected network, and then we have \( Z_c = (n-1)/2 \) for an odd \( n \), and \( Z_c = n/2 \) for an even \( n \). If \( n \) is relatively large, the probability for each specific value of \( m \) could approximately be \( 1/Z_c \). Then we could calculate \( P(l) \) based on Eqs. (10) as follows:

\[
P(l = C_t),l \in Z) = \frac{1}{Z_c}
\]

\[
P(l = C_t),l \in Z) = \frac{Z_c + 1 - Z_c}{Z_c}
\]

\[
P(l = C_t + 1,m(k \geq 0)) = Z_c + 1 - Z_c - k
\]

\[
P(l = C_t + 2,m(k \geq 0)) = Z_c + 1 - Z_c - k
\]

\[
P(l = Z_c + k,m(k > 0),m(m > Z_c + k,k \geq 0)) = (1 - p)^k + 1 + k + (2 - p) - p^2
\]

\[
P(l = Z_c + 2,m(k > 0),m(m > Z_c + 2,k \geq 0)) = (1 - p)^k + 1 + k + (2 - p) - p^2
\]

\[
P(l = Z_c + 2,m(k > 0),m(m > Z_c + 2,k \geq 0)) = (1 - p)^k + 1 + k + (2 - p) - p^2
\]

Based on Eqs. (11), we have
\[ I = \sum_{i} I(p(i)) = \frac{1}{2p^2(1-p)^2Z_c^2(2-p)^2Z_n-\pi (pZ_c(p-2))(-1-p-pZ_c^2(2-p)2Z_n+p+2Z_c-4+2)} 
+ (1-p)^2Z_c^2(p(1-Z_c)+p(2-p)Z_c+p+2Z_c-4+2) 
- 2 \bigg) + (1-p)^2Z_c^2(p^2Z_c^2(p-Z_c+2)Z_c+2p^2Z_c^3(3Z_c^2-3Z_n-2p^2Z_c^3Z_n+3Z_n+2p) 
+ 2Z_c+1 + 2p(2Z_c-1-Z_c-6Z_n-7) + 8(Z_n+2)) 
- 2(1-p)^2Z_c^2(p^2Z_c(p-Z_c+2)Z_c+2p^2Z_c^3(3Z_c^2-3Z_n-2p^2Z_c^3Z_n+3Z_n+2p) 
+ 2Z_c+1 + 2p(2Z_c-1-Z_c-6Z_n-7) + 8(Z_n+2)) - 8(1-p)^2Z_c^2 + 6(1-p)^2Z_n^2. \]

2.2. The modified model for the structure containing multiple rings

So far we have presented the expressions to calculate the analytic estimates of the average shortest distance between nodes on the ring for both directed and undirected networks consisting of a hub and a single ring. Now let us concentrate on those networks containing multiple rings. For those regular networks consisting a hub and multiple rings where those rings have no interaction with each other (see Fig. 3(c) and (d) for examples), Jarrett et al. [4] proposed a concise and valid method. For a specific ring, its inner rings as well as the hub could be viewed as a whole which acts as a new hub. Think about such a directed network, if we assume that the length of any edge linking different rings (including each edge linking the core and the innermost ring) equal to 1/2 (\(a = 1/2\)), we could calculate the average shortest distance \( I \) between nodes on the innermost ring by substituting \( c = 0 \) into Eq. (6). Then for the ring next to the innermost ring, its new hub is composed of the original hub and the innermost ring, and hence the congest cost \( c \) passing through that new hub will just equal to the average shortest distance between nodes on the innermost ring. Substitute that value into Eq. (6) again, and we will get the average shortest distance between nodes on the ring next to the innermost ring. In this way, we could calculate \( I \) for each ring from inner to outer by Eq. (6).

The main defect of this method lies in that it ignores the impacts of outer rings on the shortest distance between nodes on inner rings. By this method, the estimates of the average shortest distance between nodes on the innermost ring would be always the same no matter whether those outer rings exist or not. However, the existence of those outer rings does provide more path selection and hence would decrease the average shortest distance between nodes on the innermost ring. So here we argue that this method could only be applied to the case where nodes on outer rings could make use of inner rings to shorten their distance, but not vice versa, nodes on inner rings can not use outer rings to do the same work. More exactly, these models estimate the average shortest distance between nodes on the same ring when only straight and U-shape paths between those nodes are allowed. See Fig. 2 for illustration. After we leave the start point A, we could walk along the present ring (A-B) or go to the inner ring next to it (A-F), but once we begin to go inside, we have only one chance to turn the movement direction to outward (e.g. A->F->G->C->D->E, A->F->G->O->J->E, A->F->G->H->I->J->E and A->F->G->C->D->E) and there is no choice to turn the movement direction to inward. This selection of paths rules out the possibility of many repeated shuttles between different rings (e.g. A->F->G->C->D->J->E will be forbidden) and could effectively eliminate the impacts of outer rings on the distance between nodes on inter rings (e.g. A->F->K->E is not allowed). In the real world, such selection of paths indicates that the walker tends not to use outer rings to reduce the total distance and does not like overly complex paths which contain many repeated shuttles between different rings, and thereby has a certain degree of universality.

As our modified model for the irregular ring-and-hub structure would inherit the method of Jarrett et al. [4], we believe that it could work well when only straight and U-shape paths are allowed. Here we would like to use the subscript \( j \) to label all variables belonging to ring \( j \). In this way, \( I_j \) denotes the average shortest distance between nodes on the innermost ring, which could be calculated by Eq. (9) or (12). Then for ring 2, ring 1 and the hub could be viewed as a new hub altogether with a passing distance \( I_1 \). The length of the shortcut between two nodes on ring 2 would be \( Z_2 = 2a + I_1 \), while \( Z_2 \) would be \( 2a + I_1 + 1 \). Substitute \( Z_2, Z_2 \) and other given parameters into Eq. (9) or (12), we could obtain \( I_2 \). Then \( I_3, I_4, \ldots \) could be calculated like this one by one.

For irregular ring-and-hub networks, some additional definitions are necessary. Let us assume that ring \( j \) consists of \( n_j \) nodes, among which \( s_j \) nodes are borrowed from ring \( j-1 \) (those nodes are also intersection nodes of the two rings). For each node on ring \( j \) except those intersection nodes, the probability to exist an undirected edge linking it and a node on ring \( j-1 \) (not an intersection node) is \( q_j \), e.g. for the irregular ring-and-hub network presented in Fig. 3(a), we have \( n_z = 16, s_z = 3 \), and \( q_z = 3/13 \).

As those intersection nodes actually belong to more than one ring, we could split each of them into new nodes on different rings and then add edges with zero length between those new nodes produced by the same intersection node (see Fig. 3(b) for details). This operation could generate an equivalent regular ring-and-hub network for the original irregular one on the condition that the node number of each ring as well as the shortest distance between any two nodes on that ring remain unchanged. By calculating the average shortest distance between nodes on the same ring in the equivalent regular network, we could approximate the corresponding analytic value of the original one. For ring \( j \), Eq (3) and (12) could be converted into
\[ I_{dirij} = \frac{1}{2(1-p_j)p_j} [Z_{cij} - Z_{cij}(n_j - 1) - (1 - p_j)p_jZ_{cij}(n_j)] \\
- Z_{cij} + p_j(2C_{cij} + n_j - 3Z_{ij} - 1 + 3) + (p_j^2C_{cij}(p_j + n_j - 3Z_{ij} - 1 + 3)] + \frac{1}{2}(Z_{cij} - 2)(Z_{cij} + 1) + 8 - 6) \\]

and

\[ I_{undirij} = \frac{1}{2(1-p_j)^2Z_{cij}^2(2 - p_j)} (p_j^2C_{cij}(p_j + n_j - 3Z_{ij} - 1 + 3) + \frac{1}{2}(Z_{cij} - 2)(Z_{cij} + 1) + 8 - 6) \\
- \frac{1}{2}(1-p_j)^2Z_{cij}(p_j + n_j - 3Z_{ij} - 1 + 3) + \frac{1}{2}(Z_{cij} - 2)(Z_{cij} + 1) + 8 - 6) \\
+ (1-p_j)^2(1-p_j) + (1-p_j)^2Z_{cij} + (1-p_j)^2Z_{cij} + 1 + 2p_j(2Z_{cij} \\
- Z_{cij} - 2Z_{cij} - 1 + 2p_j]Z_{cij}(3Z_{ij} - 3Z_{ij} - 2) \\
+ (1-p_j)^2Z_{cij}(13Z_{ij} + 2Z_{cij} + 2Z_{cij} + 2) + p_j(2Z_{cij} \\
+ 1)(2Z_{cij} - 2Z_{cij} + 1) + Z_{cij}(-4Z_{ij} + 4Z_{ij} - 15) + 5Z_{ij} \\
- 7) + 6Z_{cij} - 2Z_{cij} - 8 - 6(1-p_j)^2Z_{cij} + 6(1-p_j)^2Z_{cij}].
\]

Taking the network in Fig. 3 for example, we have

\[ p_2 = (1 - \frac{3}{16}) + \frac{3}{16} + \frac{3}{16} = \frac{3}{8} \quad \text{and} \]

\[ C_{e2} = \frac{2a_jq_j(n_j - s_j)}{q_j(n_j - s_j)} + s_j + \frac{1}{2}I_1. \]

By Substituting Eqs. (15), (16), \( Z_{e2} = [C_{e2}] + 1 \), and \( Z_{e2} = 2/(n_j-1) \) (for an odd \( n_j \)) or \( Z_{e2} = 2/n_j \) (for an even \( n_j \)) into Eq. (13) or (14), respectively, we could compute \( I_{dirij} \) or \( I_{undirij} \).

3. The comparisons between analytic estimates and actual values

3.1. The performance of the previous model on regular ring-and-hub networks

In order to examine our inference on the defect of the method of Jarrett et al. [4], here we apply their model to a very simple directed regular ring-and-hub structure which contains only two 100-node rings (\( n_1 = n_2 = 100 \)) and a hub, and all edges in it are of unit length (\( a = 1 \)). Substitute \( c = 1 \) into their Eq. (6), and we get

\[ I_1 = \frac{2n_1p_1 + (1 - p_1)^2(n_1 - 1) + 3 - 3}{p_1n_1 - 1}. \]

Then substitute \( c = 1 + I_1 \) into Eq. (6), and we get

\[ I_2 = \frac{1}{2p_2(n_2 - 1)} [(1 - p_2)^2(n_2 - 1) + 3 + (n_2 - 3 - I_1)p_2] \]

\[ + p_2(-2I_1 + 2n_2 - I_1(I_1 + 1 - n_j)p_2) + \frac{1}{2}(I_1 + 1). \]

By observing Fig. 4 (a) and (b), we find that when \( p_1 \) is relatively low, the actual average shortest distance between nodes on ring 1 (the innermost ring) is much lower than its analytic value, which hints that at this moment, nodes on ring 1 have to rely heavily on ring 2 to form the shortest path between them, and the higher \( p_2 \) is, the lower the actual value is (the actual value for \( p_2 = 0.6 \) is lower than that for \( p_2 = 0.2 \), when \( p_2 \) is relatively low). However, as \( p_1 \) value increases, the influence of ring 2 on the average distance of ring 1 will become lower and lower, and at this moment the analytic value is very close to the actual value in both Fig. 4 (a) and (b).

According to our inference on the defect of this method, as ring 2 is the most outside ring, the discrepancy between the analytic and actual values of \( I_2 \) mainly comes from the imprecise estimate of its input \( I_1 \). From Fig. 4(c), we can see that the discrepancy is increasing as \( p_2 \) increases. Due to that \( p_1 \) maintains a relatively low value 0.2, the actual value of \( I_1 \) will be decreased by \( p_2 \). The higher \( p_2 \), the greater the discrepancy of \( I_1 \) is, and so is the discrepancy of \( I_2 \). As a comparison, when \( p_1 = 0.6 \) (see Fig. 4(d)), the influence of ring 2 on \( I_1 \) is very small, and hence the analytic values of both \( I_1 \) and \( I_2 \) are close to their actual values.

In a word, the results show that the method of Jarrett et al. [4] may not be suitable for the case where all paths are allowed, and the evolution of the discrepancy also coincides with our inference that the discrepancy of this method mainly comes from its neglect of the impacts of outer rings on the shortest distance between nodes on inner rings.

3.2. An algorithm to generate irregular ring-and-hub networks

Before we compare the analytic estimates and the actual values to observe the accuracy of our modified model, we have to develop an algorithm to generate enough irregular ring-and-hub network samples. Our algorithm could be demonstrated by Fig. 5. This algorithm would generate rings according to the given parameters from inner to outer one by one. For ring \( j \), we first random select \( s_j \) nodes from ring \( j-1 \) (see Fig. 5(a)), and then link them in the same direction as they are linked in ring \( j-1 \) to form the initial ring \( j \) (see Fig. 5(b)). Then we randomly put the remaining \( n_j-s_j \) nodes on that ring one by one until we get a complete ring \( j \) (see Fig. 5(c)). In this way, if two neighboring nodes on ring \( j-1 \) were selected as the intersection nodes and meanwhile no other node was inserted between them, the edge linking them on ring
j-1 would be borrowed by ring j to form itself (see edge AB in Fig. 5(c)). At last, for each node except those intersection ones on ring j, we link it with a randomly selected node (also not the intersection ones) on ring j-1 by probability $q_j$ (see Fig. 5(d)), and we could further prescribe that the selected node on ring j-1 must have not been linked by other nodes on ring j, but it demands that there are enough such nodes on ring j-1. In a word, the irregular ring-and-hub networks generated by this algorithm would cover more diverse cases than that we straightforward select $s_j$ connected nodes from ring j-1 as the intersection nodes to form ring j.

Those solid nodes denote the intersection nodes of multiple rings while hollow nodes only belong to one ring.

By this algorithm, we generate irregular ring-and-hub networks where every ring contains just 100 nodes ($n_j = 100$), and we further let the length of any edge linking different rings equal to 1 ($q_j = 1$), then the computer can easily count the shortest distance between any two nodes on the same ring, and calculate the average.

### 3.3 The accuracy of our modified model on irregular ring-and-hub networks

Finally, let us scrutinize the results on those irregular directed (Fig. 6) and undirected (Fig. 7) ring-and-hub networks which all contain three rings. For the purpose of comparing the accuracy of the analytic estimates for different rings, we let $q_1$, $q_2$, and $q_3$ take the same values, and put those curves for the three rings altogether. Meanwhile, the case where all paths are allowed is placed on the left part of these two figures, while the case where only straight and U-shape paths are allowed is on the right part. As we have announced that our modified model only fits the latter case, the results in Fig. 6(b)(d) and Fig. 7(b)(d) would represent its accuracy.

Here we present the results for the case where all paths are allowed mainly to provide a reference for comparison. From Fig. 6(c) and 7(c), we can see that when $q_j$ ($q_1 = q_2 = q_3$) is between 0 and 1, the gaps between the analytic estimates and the actual values of $l_1$, $l_2$ and $l_3$, are always large. However, their causes are quite different. The discrepancy of $l_1$ mainly comes from the repeatedly mentioned fact that the outer rings (ring 2 and 3) may provide more path selections to shorten the distance between nodes on ring 1, the discrepancy of $l_2$ stems from both the discrepancy of its input $l_1$, and the influence of the outer ring (ring 3), and the discrepancy of $l_3$ is mainly caused by the cumulative discrepancy of its input $l_1$ and $l_2$.

By horizontally comparing the left and right parts of Fig. 6 and 7 (namely, comparing Fig. 6(a) and (b), Fig. 6(c) and (d), Fig. 7(a) and
(b), and Fig. 7(c) and (d)), we find that the accuracy of our modified model for the case where only straight and U-shape paths are allowed is much better than the case where all paths are allowed. Further in Fig. 6(b)(d) and Fig. 7(b)(d), the accuracy of analytic estimates of \( I_1 \) and \( I_2 \) are very good, though the accuracy would decrease from inner rings to outer rings due to the cumulative discrepancy of inputs. Considering the fact that such discrepancy is difficult to be eliminated, we believe that the accuracy of our modified model is acceptable.

4. An application on real network data

Here we would like to present the application of our modified model on a real network, namely, the ring road deployment of Baoding (described by Fig. 1(f)). Actually, this network needs to be adjusted in several steps before fitting the premise of the model, including reasonably selecting nodes on each ring, approximating the distances between neighboring nodes as integers, and so on (see the result of the adjustments in Fig. 8).

From Fig. 8 we can see that ring 1 contains 10 nodes (denoted by seven circles and three diamonds), and so we have \( n_1 = 10 \). Since \( n_1 \) is even but not large, directly using Eq. (14) may cause a large inaccuracy. At this moment, we have

$$
P(m < \frac{n_1}{2}) = \frac{n_1 - 1}{n_1 - 1}.\text{Therefore, Eq. (14) can be adjusted into}$$

$$
\hat{I}_{undist} = \frac{1}{2(n_1 - 1)p_j^2(1 - p_j)^2Z_{c_{ij}}(2-p_j)^2(C_{ij})} (p_j)^{C_{ij}} \right.

\left. \begin{array}{l}
(p_j - 2)(1 - p_j)^3(p_j - 2)p_j - 2\left(n_1(p_j - 2)p_j \right.

\left. + 2 + 2p_j^2 + 2 - 2(p_j - 2)p_j + 6)p_j Z_{c_{ij}} + 2p_j + 8 Z_{c_{ij}}

\left. + 12 - 1)(1 - p_j)^2Z_{c_{ij}}(p_jn_j(p_j - 2) - 2p_j Z_{c_{ij}}

\left. + p_j + 4Z_{c_{ij}} - 6) + 4) + (1 - p_j)^3(-n_j(p_j - 2)(p_j - 2))\right)$$

(17)

The distance between each pair of neighboring nodes on ring 1 is 2 kilometers, and so we choose 2 kilometers as the unit length for this ring. By further observation, we find that those edges between the hub and the nodes on ring 1 are not of equal length. We have to average them and adjust the result according to the unit length. Therefore, we have

$$
c_1 = \frac{1 \times 2 + 1 \times 4 \times 2 + 1.625 \times 2 + 2.563 \times 2 + 1.625 = 4.016, Z_{c_{12}} = 5, p_2 = \frac{7 + 3}{16} = 0.625. \text{By Substituting them into Eq. (17), we get } \hat{I}_{undist} = 3.569.$$ 

The actual average distance between nodes on ring 1 is 4.680 kilometers, while that actual value for ring 2 is 6.146 kilometers. Considering the unit length, we have \( \hat{I}_{true} = \frac{4.680 \times 2}{2} = 2.340, \) and \( \hat{I}_{true} = \frac{6.146 \times 2}{2} = 3.793. \) Comparing these final estimates \( \hat{I}_{undist} = 2.563 \) and \( \hat{I}_{undist} = 3.569 \) with the true values, we can see that their accuracy is relatively good.

5. Conclusion

This paper mainly focuses on the average shortest distance between nodes on the same ring in a type of irregular ring-and-hub network structure. This network structure is very common in the real world. It consists of a hub and multiple rings, one around another, where neighboring rings may share nodes or line segments. We construct the modified analytic model to estimate that average shortest distance between nodes on the same ring based on DM model [3] and the method proposed by Jarrett et al. [4], and apply it to irregular ring-and-hub networks generated by computer and in the real world to observe its accuracy.

The results show that our modified model actually estimates the average shortest distance between nodes on the same ring when only straight and U-shape paths are allowed. The accuracy of estimates for innermost several rings can be acceptable. However, the discrepancy would increase if we take all types of paths into consideration. To develop a more accurate analytic model with relatively concise expressions for the case where all paths are allowed would be the next work in the future.

Declarations

Author contribution statement

Tao Fu: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper. Chenguang Li: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data. Long Wu: Performed the experiments; Contributed reagents, materials, analysis tools or data. Liling Zou: Analyzed and interpreted the data; Wrote the paper.

Funding statement

Dr Chenguang Li was supported by Natural Science Foundation of Beijing Municipality [9222010], Humanities and Social Sciences Youth Foundation, Ministry of Education of the People’s Republic of China [20YJCZH066].

Data availability statement

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

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

[1] D.J. Watts, S.H. Strogatz, Collective dynamics of ‘small-world’ networks, Nature 393 (1998) 440.
[2] A.L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (5439) (2000) 509–512.
[3] S.N. Dorogovtsev, J.F.F. Mendes, Exactly solvable small-world network, Europhys. Lett. 50 (2000) 1.

[4] T.C. Jarrett, D.J. Ashton, M. Fricker, N.F. Johnson, Interplay between function and structure in complex networks, Phys. Rev. E 74 (2006) 026116.

[5] S.N. Dorogovtsev, J.F.F. Mendes, J.G. Oliveira, Degree-dependent intervertex separation in complex networks, Phys. Rev. E 73 (2006) 056122.

[6] C. Song, S. Havlin, H.A. Makse, Origins of fractality in the growth of complex networks, Nat. Phys. 2 (2006) 275.

[7] Z. Zhang, Y. Lin, S. Gao, S. Zhou, J. Guan, Average distance in a hierarchical scale-free network: an exact solution, J. Stat. Mech. Theory Exp. 10 (2009) 1002.

[8] S. Blanco, F. Richard, An invariance property of diffusive random walks, Europhys. Lett. 61 (2003) 168.

[9] A. Baronchelli, M. Catanzaro, R. Pastor-Satorras, Random walks on complex trees, Phys. Rev. E 78 (2008) 011114.

[10] M. Barahona, L.M. Pecora, Synchronization in small-world systems, Phys. Rev. Lett. 89 (2002) 054101.

[11] M. Zhao, T. Zhou, B.H. Wang, G. Yan, H.J. Yang, W.J. Bai, Relations between average distance, heterogeneity and network synchronizability, Physica A 371 (2006) 773.

[12] Z.L. Hu, J.G. Liu, G.Y. Yang, Z.M. Ren, Effects of the distance among multiple spreaders on the spreading, Europhys. Lett. 106 (2014) 18002.

[13] L. Guo, J.H. Lin, Q. Guo, J.G. Liu, Identifying multiple influential spreaders in term of the distance-based coloring, Phys. Lett. A 380 (2015) 837.

[14] D. Ashton, How small is your social life? Exactly solvable small-world networks, https://www.douglashton.net/publications/papers/doug.mphys.pdf.

[15] D.J. Ashton, T.C. Jarrett, N.F. Johnson, Effect of congestion costs on shortest paths through complex networks, Phys. Rev. Lett. 94 (2005) 058701.