A Method for Geodesic Distance on Subdivision of Trees with Arbitrary Orders and Their Applications

Fei Ma\textsuperscript{a,1}, Ping Wang\textsuperscript{b,c,d,2} and Xudong Luo\textsuperscript{e,3}

\textsuperscript{a} School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China
\textsuperscript{b} National Engineering Research Center for Software Engineering, Peking University, Beijing, China
\textsuperscript{c} School of Software and Microelectronics, Peking University, Beijing 102600, China
\textsuperscript{d} Key Laboratory of High Confidence Software Technologies (PKU), Ministry of Education, Beijing, China
\textsuperscript{e} School of Mathematics and Statistics, Lanzhou University, 730000 Lanzhou, CHINA.

Abstract: Tree, as the simplest and most fundamental connected graph, has received considerable attention from a variety of disciplines. In this paper, our aim is to discuss a family of trees of great interest which are in fact divided into two groups. Unlike preexisting research focusing mainly on a single edge as seed, our treelike models are constructed by arbitrary tree $T$. Therefore, the first group contains trees generated based on tree $T$ using first-order subdivision. The other is constituted by trees created from tree $T$ with $(1,m)$-star-fractal operation. By the novel methods addressed shortly, we do capture analytically the exact solution for geodesic distance on each member in tree family of this type. Compared to some commonly adopted methods, for instance, Laplacian spectral, our techniques are much lighter to implement according to both generality and complexity. In addition, the closed-form expression of mean first-passage time ($MFPT$) for random walk on each member is also readily obtained on the basis of our methods. Our results suggest that the two topological operations are sharply different from each other, particularly, $MFPT$ for random walks, and however have likely to show the same function, at least, on average geodesic distance.

Keywords: Geodesic distance, Tree, Subdivision, Fractal, Self-similarity, Random walks.

\textsuperscript{1} The author’s E-mail: mafei1239870@163.com.
\textsuperscript{2} The corresponding author’s E-mail: pwang@pku.edu.cn.
\textsuperscript{3} The author’s E-mail: luoxudong117@163.com.
1 INTRODUCTION

Geodesic distance, conventionally called shortest path length in the language of graph theory, has proved useful in a great variety of areas, for example, discrete applied mathematics, theoretical computer science, biology science and so forth. While this is in essence not a new concept which has a long history, its related researches still keep active in present science community and have received more attention. Included examples have searching information on internet [1], signal integrity in communication networks, disease spreading on relationship networks among individuals [2], navigation in spatial networks [3], to name but a few. Perhaps, one of the most important reasons for this is to intuitively measure information spread and information retrieve on internet. For a pair of vertices \( u \) and \( v \) attempting to interact with one another on a communication network which is also abstractly thought of as a graph denoted by \( G(V, E) \), the much smaller the geodesic distance between them, the much less the time to need. Here, geodesic distance is defined as the edge number of a shortest path connecting this pair of vertices. In addition, this concept is playing a significant role in some newborn disciplines, for instance, complex network.

The recent two decades have seen a bloom of complex network study mainly because of its own right which has been helping us to understand some complex systems around ones. Such complex systems include World Wide Web (WWW), citation networks, metabolic networks, protein-protein interaction networks and predator-prey webs [4]-[8]. Particularly, the portion of complex network study focus mainly on investigating how dynamic and function taking place on complex networks of interest are influenced by the topological structures of their underlying graphs. Among of which, one type of networked models have attracted more attention according to their own specific topological structure like tree. One of such examples is citation networks of scientific papers [9] which is in general considered directed where the vertices represent documents and the directed edges represent citations between them. In this respect, as one of the best studied tool in various fields, traditional random graph models of networks can generate networks of this type that are locally treelike, meaning that all local neighborhoods take the form of trees. To be completely general, a great deal of treelike models have been proposed as research models indirectly predicting some behaviors occurring in many complex networks. Among of which a fraction of tree family become of great concern according to their own some intriguing topological structures including power-law distribution for vertex degree [10], exponential vertex degree distribution [11], fractal feature [12] which can all turn out to be quite prevailing in man-made and natural complex networks.

As known, tree, as the most fundamental and simplest connected graph, has been widely studied in the last hundreds of years. Indeed, there are a surprising number of applications built upon structure
of tree just because it is exactly solvable and owns significant availability. Social networks, for instance, display hierarchical structure such that ones can naturally employ treelike models, the so-called dendrograms, as a graphical representation and summary of the structure of this type \[13\]. Meanwhile, in information science, tree has been playing an important role in search engine at present. It can be worthy mentioning that almost all data structures and a large fraction of algorithms suit for searching information are based on treelike models, such as binary tree \[14\]. Among other things there are many other potential and invaluable applications on the basis of treelike models in coding theory including the best known Huffman tree coding, quedreecoding and so on \[15\]. The last but least, some other interesting topics correlated with treelike models have received consideration attention, such as geodesic distance \[16\], mean first-passage time for random walk \[17\], fractal phenomena \[18\], etc. We, in this paper, propose a family of treelike models and analyze geodesic distance on them.

Roughly speaking, a central problem to answer is to explicitly capture solutions for computation of geodesic distance on treelike models of great interest in view of theoretical value and practical applications. Some related work have been reported in some fields. Therefore, exact solutions for several types of treelike models have been obtained by taking advantage of some typical tools in which the best used is Laplacian spectral and eigenvectors based on laplacian matrix of underlying structure \[19\]. Nonetheless, we here do not adopt methods of this type and instead introduce some novel methods according sufficiently to topological structure of treelike models under consideration. While the core of both our methods is to built up a series of equations in an iterative manner as with other enumeration methods \[20\], they are, in some extent, more convenient to manipulate, at least on treelike models addressed in this paper.

This paper can be organized by the following several Sections. In Section 2, we will introduce some helpful definitions, including first-order subdivision and \((1, m)\)-star-fractal operation, to build up our desired treelike models whose special case has been well studied in some published paper \[20\], and useful notations, for example, surjection and bijection, to smoothly develop our main results. And then, using a novel method for calculating geodesic distance on treelike models proposed in this paper, we in Section 3 derive exact solutions for geodesic distance on two classes of treelike models built up in the above section. Among of them we obtain the closed-form solution for geodesic distance on a special case of the resulting models created by the \((1, m)\)-star-fractal operation in self-similar manner because of its own self-similarity. In such a situation, the self-similar methods are the same as our techniques for calculating what we want. Nevertheless, under more general circumstance, our techniques are more than the former. In addition, to show some other potential applications of our methods addressed here, in Section 4, we make use of connection between random walk and electrical network to derive analytically solutions for
mean first-passage time on the generated treelike models. By the both topological measures, we may state that the two topological operations are significantly different from one another, in particular, about $MFPT$ for random walks, because the end models based on the $(1,m)$-star-fractal operation possess fractal property and the other has no such characters. On the other hand, they have likely to show the same functions, at least, on average geodesic distance. In conclusion, we close this paper by providing some potential applications of our methods as a guide of our future work.

2 DEFINITIONS AND NOTATIONS

Here, we will recall some fundamental definitions and widely adopted notations from graph theory. It is conventional to let symbol $G(V,E)$ denote as a graph where $V$ and $E$ are vertex set and edge set, respectively, and its corresponding order (vertex number) and size (edge number) are denoted by $|V|$ and $|E|$ where symbol $||$ represents the cardinality of a set. In the meantime, we denote by the notation $[1,n]$ an integer set which consists precisely of those integers no more than $n$ and no less than $1$. For more details to see Ref. [21].

Definition 1 Given an arbitrary graph $G(V,E)$, if one inserts a new vertex $w$ to every edge $uv \in E$ then the resulting graph, defined as $G'(V'_1,E'_1)$, is called a first-order subdivision of original graph $G(V,E)$. Put this another way, the first-order subdivision graph can be equivalently obtained from graph $G(V,E)$ by replacing every edge $uv \in E$ by a unique path $uwv$ with length 2 where internal vertex $w$ is in fact that inserted vertex. Henceforth we regard such an operation on each edge of a graph as first-order subdivision and that resulting graph $G'(V'_1,E'_1)$ as first-order subdivision graph of graph $G(V,E)$. It is worth noting that we in this paper focus mainly on discussions about impact from first-order subdivision on geodesic distance of tree $T(V,E)$ which is the simplest yet most important member of graph family. Here, Fig.1(a) illustrates the first-order subdivision on an edge.

For our purpose, it can immediately know by Def.1 that the first-order subdivision graph $G'_1(V'_1,E'_1)$ holds on a couple of equations $|V'_1| = |V| + |E|$ and $|E'_1| = 2|E|$. After applying first-order subdivision until $t$ time steps, the order $|V'_t|$ and size $|E'_t|$ of the first-order subdivision graph $G'_t(V'_t,E'_t)$ will follow a pair of equations

$$|V'_t| = |V| + (2^t - 1)|E|, \quad |E'_t| = 2^t|E|.$$  

(1)

Definition 2 Given an arbitrary graph $G(V,E)$, if one not only inserts a vertex $w$ to every edge $uv \in E$ but also connects $m$ other new vertices $w_i$ ($i \in [1,m]$) to this newly inserted vertex $w$, then
Fig. 1. The diagrams of two types of operations on an edge. First-order subdivision on an edge is shown in panel (a) and the panel (b) shows (1, m)-fractal-operation on an edge where m = 3.

the resulting graph, denoted by $G^*_1(V^*_1, E^*_1)$, is called a (1, m)-star-fractal graph of original graph $G(V, E)$. Equivalently speaking, such an operation can be achieved from graph $G(V, E)$ by directly inserting a star, which vertex $w$ is the central one attached to $m$ leaves $w_i$, to every $uv \in E$ and hence called (1, m)-star-fractal vividly. It is obvious to say that the well known T-fractal can be induced as a result of our (1, m)-star-fractal graph when parameter $m$ is supposed equal to 1 [17]. As before, our aim is to study geodesic distance on each member of (1, m)-star-fractal tree family $T^*_t(V^*_t, E^*_t)$. An example as illustration of (1, m)-star-fractal on an edge connecting two vertices is shown in Fig.1(b) where the newly generated star has 3 leaves.

For brevity, with the help of Def.2, one can find out that (1, m)-star-fractal graph $G^*_1(V^*_1, E^*_1)$ has $|V^*_1| = |V| + (1 + m)|E|$ vertices and $|E^*_1| = (2 + m)|E|$ edges. Similarly, for (1, m)-star-fractal graph $G^*_t(V^*_t, E^*_t)$, its vertex number $|V^*_t|$ and edge number $|E^*_t|$, respectively, obey

$$|V^*_t| = |V| + ((2 + m)^t - 1)|E|, \quad |E^*_t| = (2 + m)^t|E|. \quad (2)$$

The statements above are mainly correlated with an arbitrary graph. However, below will introduce a helpful definition about vertex cover on a path in the jargon of graph theory that is used later to finish our proofs conveniently. This is in essence the classification of vertices on a path. For more detail to see [21], we here just discuss such a problem in the simplest and most fundamental situation.

**Definition 3** Given a path $P$ with $n$ vertices, it is easy to see that these $n$ vertices can be divided into two disconnected vertex sets, without loss of generality, labeled as set $X = \{x_1, x_2, ..., x_{[n/2]}\}$ and set $Y = \{y_1, y_2, ..., y_{[n/2]}\}$. The above classification of vertices is in fact a bipartition of vertex set. More generally, vertices of set $X$ and vertices of set $Y$ can be alternatively arranged on the path $P$ in an appropriate manner such that arbitrary vertex pair $x_i$ and $x_j$ is not connected directly by an edge and similarly for all pairs of vertices $y_i$ and $y_j$. This suggests that either vertex set $X$ or vertex set $Y$ is a minimal vertex cover of path $P$ where $|X| = |Y| + 1$ when $n$ is odd and $|X| = |Y|$ otherwise. The both
vertex sets will be alternatively employed to play a vital role on the growth process of building up our main results in the rest of this paper.

Now let us take a common yet important terminology from analysis mathematics which has been successfully adopted in a great variety of science fields.

**Definition 4** Given two sets $X$ and $Y$ that are all not empty, that is to say, consisting of at least one element each, one can make a **mapping** $f$ from $X$ to $Y$ such that for a provided element $x$ of set $X$ there must be a unique element $y$ belonging to set $Y$ satisfying $f(x) = y$. This can be simply expressed in the following

$$\forall x \in X, \ \exists y \in Y, \ \text{s.t.,} \ f : x \mapsto y$$

where the image set of set $X$ may be set as $f_{X \mapsto Y} = \{y | f(x) = y, \ y \in Y\}$. We in this paper interested in the below two kinds of mappings between sets $X$ and $Y$.

**case 1** If both $|X| > |Y|$ and $|f_{X \mapsto Y}| = |Y|$ hold on true, then this mapping $f$ is considered **surjection**. Besides, for each element $y$ of image set $Y$, if there exist $n$ distinct pre-images $x_i$ ($i \in [1, n]$), i.e., $f^{-1}(y) = x_i$, then the surjection $f$ is considered **$n$-regular**. It is clear to the eye that both sets $X$ and $Y$ follow $|X| = n|Y|$ when surjection $f$ is $n$-regular.

**case 2** If the surjection $f$ under consideration holds $|X| = |Y|$ then it can be thought of as a **bijection**, also called one-one mapping.

For convenience, the compound mapping between two mappings $f$ and $g$ can be expressed as $f \circ g$ mathematically.

So far, we have introduced some helpful definitions and notations used later. As stated above, the topic of this paper focus principally on many discussions correlated to geodesic distance on treelike models of significant interest. In addition, we will take useful advantage of one of intriguing features of tree that any pair of vertices $u$ and $v$ of a tree is connected by a unique path which starts from vertex $u$ to vertex $v$. In fact, the length of such a path is indeed the geodesic distance of this vertex pair. For simplicity and convenience, we denote by vertex pair $< u, v >$ a path whose endvertices are $u$ and $v$. Now, let us turn our insight to problems.

## 3 MAIN RESULTS AND APPLICATIONS

We will in this section show main results with respect to our novel methods and a number of applications which are organized into several theorems, corollaries and applications in form.
**Theorem 1** Given an arbitrary tree $T(V, E)$, the exact solution for geodesic distance $S'_1$ of its first-order subdivision tree $T'_1(V'_1, E'_1)$ is

$$S'_1 = 8S - 2|V|(|V| - 1)$$

in which $S$ is a known expression to geodesic distance of tree $T(V, E)$.

**Proof** Consider an arbitrary tree $T(V, E)$, first, assume that its geodesic distance is equal to $S$. After applying first-order subdivision to each edge of tree $T(V, E)$, the first-order subdivision tree $T'_1(V'_1, E'_1)$ will consist of two different types of vertices, without loss of generality, which are grouped into two disjoint vertex sets $X'$ and $Y'$. Set $X'$ contains the total old vertices of original tree $T(V, E)$ and is in fact set $V$, namely, $x' \in X'$ being the same as $x \in V$. The other set, set $Y'$, is constituted by all the created vertices using the first-order subdivision. Obviously, $Y' = V'_1 - V = V'_1 - X'$. In order to precisely calculate geodesic distance $S'$, it is straightforward to compute three classes of geodesic distances, one for vertex pairs $<x'_i, x'_j>$ of set $X'$, one for vertex pairs $<y'_i, y'_j>$ of set $Y'$ as well the latter for vertex pairs $<x'_i, y'_j>$ between set $X'$ and set $Y'$. To do so, we will in turn accomplish these calculations according to their own complexity.

Case 1.1 For a given vertex pair $<x'_i, x'_j>$ of set $X'$, there must be a bijection $f_1$ between set $X'$ and set $V$ such that $f_1(<x'_i, x'_j>) = <x_i, x_j>$ here vertices $x_i$ and $x_j$ are in set $V$. In fact, such a bijection $f_1$ is self-mapping and so one can write

$$S'_1(1) = 2S$$

where $S'_1(1)$ is the sum of distances of all possible vertex pairs in set $X'$. This is a consequence directly related to the intrinsic nature of first-order subdivision. The left tasks is to look for a reliable relation connecting the future equations to Eq.(4) just because of its own simplicity.

Case 1.2 Distinct with case 1.1 sharply, there indeed exists a self-mapping between set $Y'$ and set $V' - V$ but is no help for addressing our problem. Taking into account results in case 1.1 known to us, the current issue is to build connection to Eq.(4). Therefore, for a given pair of vertices $<y'_i, y'_j>$ belonging to set $Y'$, shown in Fig.2, we may find out a bijection $f_2$ between set $Y'$ and set $X'$ that satisfies our requirement by means of statements in both Def.3 and Def.4. To see why this is, let us pay attention on the both disjoint vertex sets $X'$ and $Y'$. For arbitrary vertex pair $<y'_i, y'_j>$ of set $Y'$, there must be a unique path $P_{y'_i y'_j}$ connecting vertices $y'_i$ and $y'_j$ which is lighted in yellow as shown in Fig.2 (online). At the same time, one certainly derives an extension $P_{x'_i x'_j}$ from path $P_{y'_i y'_j}$ by jointing two edges $x'_i y'_i$
and $x'_i y'_j$ (lighted in blue as shown in Fig.2 online), indicating which there is a mapping $f^*$ between vertex pairs $<y'_i, y'_j>$ and $<x'_i, x'_j>$. Meanwhile, it is not hard to turn out mapping $f^*$ to be bijection according to intrinsic properties among bijection $f_1$, first-order subdivision and tree itself. Thus, the compound function between the candidate $f^*$ and bijection $f_1$ may be designed as our desired bijection $f_2$, i.e., $f_2 = f_1 \circ f^*$. We have

$$S'_1(2) = S'_1(1) - 2\frac{|V|(|V| - 1)}{2}$$  \hspace{1cm} (5)

in which $S'_1(2)$ is geodesic distance of all possible vertex pairs in set $Y'$.

**Case 1.3** The remainder of our problem to answer is to capture the expression of geodesic distance for all possible vertex pairs $<x'_i, y'_j>$ whose vertices are from different sets, set $X'$ and set $Y'$. Along the research line of **case 1.2**, for vertex pair $x'_i$ and $y'_j$, there is also a unique path $P_{x'_i y'_j}$ which can be reduced to another path $P_{y'_j x'_i}$ by deleting an additional edge $x'_i y'_i$ under a similar mapping $f^{**}$ to mapping $f^*$ in **case 1.2**. Therefore, we may create a mapping $f_3 = f_1 \circ f^{**}$ which must be a surjection between vertex pairs $<x'_i, y'_j>$ and $<x'_i, x'_j>$ but not bijection. One of reasons for this is another path $P_{y'_j x'_j}$ may be induced as the path $P_{y'_j y'_j}$ by removing edge $y'_j x'_j$ as well. Armed with the both cases, there are two distinct pre-images under surjection $f_3$, i.e., $<x'_i, y'_j>$ and $<y'_i, x'_j>$, such that $f^{-1}_3(<y'_i, y'_j>) = <x'_i, y'_j>$ and $<y'_i, x'_j>$. Based on Def.4, such a surjection is in principle 2-regular and hence the exact solution for geodesic distance $S'_1(3)$ of all possible vertex pairs $<x'_i, y'_j>$ and $<y'_i, x'_j>$ obeys

$$S'_1(3) = 2S'_1(2) + |V|(|V| - 1)$$ \hspace{1cm} (6)

With the declarations in **cases 1.1-1.3**, Eqs.(4)-(6) together produces an exact solution for $S'_1$ completely equivalent to that of Eq.(3). This completes our proof.

On the basis of Eq.(6), we directly give the solution for geodesic distance $S'_t$ on the first-order
subdivision tree $T_t'(V_t',E_t')$ with omitting detailed computations, as follows

**Corollary 1** After $t$ time steps, the solution of geodesic distance $S_t'$ on the first-order subdivision

$$S_t' = 8^t S - \frac{1}{3} (2^{3t} - 2^t) (|V| - 1) + (2^{2t-1} - 2^{3t-1}) (|V| - 1)^2. \tag{7}$$

**Application 1** After $t$ time steps, the solution of average geodesic distance $\langle S_t' \rangle$ on the first-order subdivision tree $T_t'(V_t',E_t')$ will follow

$$\langle S_t' \rangle = S_t' |V_t'| - \frac{1}{2} - \frac{2^{t+1} S}{3(|V| - 1)} + 1 - 2^t. \tag{8}$$

As known, the most special one member of tree family is an edge. If let the seed of the first-order subdivision tree $T_t'(V_t',E_t')$ be an edge connecting a couple of vertices, then model $T_t'(V_t',E_t')$ will become a path $P_t'(V',E')$ with $2^t + 1$ vertices after $t$ time steps. Therefore, we are able to state the following corollary according to Eq.(7).

**Corollary 2** After $t$ time steps, the solution of geodesic distance $S(t)$ on the first-order subdivision

$$S(t) = \frac{(2t - 1)(2t + 1)}{3}. \tag{9}$$

Equivalently, it is not hard to capture the solution of geodesic distance $S(t)$ on path of such type in the most general manner, i.e. enumeration method, as follows

$$S(t) = \sum_{i=1}^{2^t} \sum_{j=1}^{2^t+1-i} j = \sum_{i=1}^{2^t} \frac{(2^t + 1 - i)(2^t + 2 - i)}{2}. \tag{10}$$

Therefore, Eqs.(9)-(10) here gives a concise proof for one combinatorial identity

$$\sum_{i=1}^{2^t} \sum_{j=1}^{2^t+1-i} j = \frac{(2t - 1)(2t + 1)}{3}$$

based on computation of geodesic distance $S(t)$ on path with length $2^t$. Put this further, one can easily derive the average geodesic distance $\langle S(t) \rangle$ on the first-order subdivision path $P_t'(V',E')$ from Eq.(9).
Application 2 After \( t \) time steps, the solution of geodesic distance \( \langle S(t) \rangle \) on the first-order subdivision path \( P'_t(V'_t, E'_t) \) will follow

\[
\langle S(t) \rangle \approx \frac{2^t + 3}{3} \propto O(|V'_t|).
\] (11)

Taking into account results between Eq.(8) and Eq.(11), we can immediately capture the below theorem which says an interesting phenomenon about influence from the first-order subdivision on average geodesic distance on the first-order subdivision tree \( T'_t(V'_t, E'_t) \).

**Theorem 2** Given an arbitrary tree \( T(V, E) \), the solution for average geodesic distance \( \langle S'_i \rangle \) on its first-order subdivision tree \( T'_t(V'_t, E'_t) \) will follow

\[
\langle S'_i \rangle \approx |V'_t|^{\gamma'} \propto O(D'_i) \tag{12}
\]

where exponent \( \gamma' = 1 \) in the large graph size limit.

By far, the first-order subdivision trees \( T'_t(V'_t, E'_t) \) which are generated by an arbitrary tree considered as a seed all share some features in common including: (i) Each vertex added by first-order subdivision at any time step \( t_i \) \( (1 \leq t_i \leq t) \) into models \( T'_t(V'_t, E'_t) \) has degree 2, (ii) The total number of leaves of models \( T'_t(V'_t, E'_t) \) keeps unchanged [22], and (iii) All models \( T'_t(V'_t, E'_t) \) exhibit homogeneous topological structure. Besides that, hereafter, the rest of this section will discuss another type of treelike models \( T'_t(V'_t, E'_t) \) with inheterogeneous topological structure. These models not only have some similar properties to models \( T'_t(V'_t, E'_t) \) but also inherit some intriguing characters from \((1,m)\)-star-fractal operation unseen in models \( T'_t(V'_t, E'_t) \), particularly, fractal property.

At first, let us start from investigating the simplest form of treelike models \( T'_t(V'_t, E'_t) \) by introducing theorem 3.

**Theorem 3** Given an arbitrary tree \( T(V, E) \), the exact solution for geodesic distance \( S^*_1 \) of its \((1,m)\)-star-fractal tree \( T^*_1(V^*_1, E^*_1) \) is

\[
S^*_1 = 2(m + 2)^2 S - (m + 2)(|V| - 1)(m + |V|) \tag{13}
\]

in which \( S \) is a known expression to geodesic distance on tree \( T(V, E) \).

**Proof** By Def.2, there also are two different groups of vertices, vertex sets \( X' = V \) and \( Y' = V' - V \), and hence we still need to consider three kinds of contributions to computation of geodesic distance \( S^*_1 \) as discussed in the development of theorem 1. While it appears more intractable to obtain closed-form solution for \( S^*_1 \) than the preceding case, the light shed by developing theorem 1 is helpful to accomplish...
our discussion. Therefore, we will again make use of more fine-grained classification method than the foregoing to resolve this computations. Considering that the first-order subdivision is a special case of the \((1, m)\)-star-fractal, we can directly use some existing results from the proof of theorem 1 to consolidate theorem 3 without specific descriptions. Meantime, some adopted notations keep active. Besides, the young vertex set \(Y'\) should be partitioned as \(Y'_1\) and \(Y'_2\) such that \(Y'_1\) is constituted by the total leaf vertices of each newly inserted star and \(Y'_2\) contains the central vertex of each newly created star. To distinguish newborn vertices between sets \(Y'_1\) and \(Y'_2\), each vertex belonging to set \(Y'_2\) remains marked \(y'_i\) and then we label each vertex in set \(Y'_1\) by \(y''_i\). From now on, let us begin with clarifying the correctness of Eq.\((13)\).

**Case 2.1** For an arbitrary vertex pair \(<x'_i, x'_j>\) of set \(X'\), under bijection \(f_1\), the geodesic distance \(S^∗_{1}(1)\) on such type of vertex pairs complies to

\[
S^∗_{1}(1) = 2S.
\]  

**Case 2.2** For an arbitrary vertex pair \(<y'_i, y'_j>\) of set \(Y'_2\), this surjection \(f_2 = f_1 \circ f^*\) will make geodesic distance \(S^∗_{1}(2)\) on such kind of vertex pairs satisfy

\[
S^∗_{1}(2) = S^∗_{1}(1) - 2\frac{|V|(|V| - 1)}{2}. \tag{15}
\]

**Case 2.3** For an arbitrary vertex pair \(<x'_i, y'_j>\) or \(<y'_i, x'_j>\), this surjection \(f_3 = f_1 \circ f^{**}\) will guarantee geodesic distance \(S^∗_{1}(3)\) on such class of vertex pairs obey

\[
S^∗_{1}(3) = 2S^∗_{1}(2) + |V|(|V| - 1). \tag{16}
\]

**Case 2.4** There must be \(|V| - 1\) new stars introduced into original tree \(T(V, E)\) by means of \((1, m)\)-star-fractal. Here we just capture geodesic distances on an arbitrary pair of leaf vertices within the same star but not between two distinct stars, i.e., vertex pair \(<y''_o, y''_o>\) where the first subscript \(o\) represents the central vertex of star attached to the both leaf vertices. Therefore the geodesic distance \(S^∗_{1}(4)\) on all possible leaf vertex pairs of this type follows

\[
S^∗_{1}(4) = (|V| - 1) \left(\frac{2m(m - 1)}{2}\right). \tag{17}
\]

**Case 2.5** We here discuss geodesic distance \(S^∗_{1}(5)\) on all possible leaf vertex pairs \(<y''_{o1}, y''_{o2}>\) in which two vertices come from different stars. As before, in order to accomplish this task, we have to
choose a fresh mapping $f^{4*}$ that bridges between vertex pairs $<y''_{ui}, y''_{vj}>$ and vertex pair $<y'_u, y'_v>$ here both vertices $y'_u$ and $y'_v$ are, respectively, the central of stars to which vertices $y''_{ui}$ and $y''_{vj}$ belong. This anticipated mapping $f^{4*}$ will in essence connect two stars and can be timely verified to be an $m^2$-regular surjection. And then, using that bijection $f_2$ introduced in case 2.2, we can generate a satisfactory surjection $f_4 = f_2 \circ f^{4*}$ in time and say

$$S^{*}_1(5) = m^2 S^{*}_1(2) + \sum_{i=1}^{\lvert V \rvert - 2} 2m^2(\lvert V \rvert - 1 - i).$$  

(18)

**Case 2.6** At the moment, let us pay attention to computation of geodesic distance $S^{*}_1(6)$ on all possible vertex pairs $<x'_i, y''_{o_j}>$ where $x'_i \in X', y''_{o_j} \in Y'_1$ and in some case two subscripts $i$ and $j$ can be equal. Considering such a vertex pair $<x'_i, y''_{o_j}>$ carefully, the first task is to find out a surjection $f^{5*}$ such that $f^{5*}(<x'_i, y''_{o_j}>) = <x'_i, y'_o>$ where vertex $y'_o$ is the central vertex of star including vertex $y''_{o_j}$. And that, combining well proposed surjection $f_3$ in case 2.3, we can create an acceptable surjection $f_5 = f_3 \circ f^{5*}$ connecting vertex pair $<x'_i, y''_{o_j}>$ with $<x'_i, y'_o>$ and so the solution for geodesic distance $S^{*}_1(6)$ is

$$S^{*}_1(6) = mS^{*}_1(4) + m\lvert V \rvert(\lvert V \rvert - 1).$$  

(19)

**Case 2.7** By now, we have successfully achieved the entire computations of geodesic distance on vertex pairs whose one vertex is from set $X'$ and the other belongs to set $Y'$. The issue to answer is to count geodesic distance $S^{*}_1(7)$ on vertex pairs where one vertex is selected from set $Y'_1$ and another one from set $Y'_2$. With the terminologies mentioned above, such a vertex pair can be thought of as $<y'_i, y''_{o_j}>$ in which it is possible that subscript $i$ is the same as $o$ when leaf vertex $y''_{o_j}$ and vertex $y'_i$ are in a common star. For all vertex pairs $<y'_i, y''_{o_j}>$ with $i \neq o$, it is natural to construct an $m$-regular surjection $f^{6*}$ projecting vertex pair $<y'_i, y''_{o_j}>$ to $<y'_i, y'_o>$ and then blurring new surjection $f^{6*}$ with bijection $f_2$ in case 2.2 together provides us with an expectant surjection $f_6 = f_2 \circ f^{6*}$ that is able to be what we hope. On the other hand, as $i$ equal to $o$, the vertex pair $<y'_i, y''_{ij}>$ will be mapped onto an identical vertex $y'_i$ under $m$-regular surjection $f^{6*}$. Even though, the $m$-regular surjection $f^{6*}$ remains available and is still employed. Through the descriptions here, a concise expression of geodesic distance $S^{*}_1(7)$ can be expressed as

$$S^{*}_1(7) = 2mS^{*}_1(2) + m(\lvert V \rvert - 1)^2.$$  

(20)
Fig. 3. The diagram of average geodesic distance $\langle S^* t \rangle$ on the first-order subdivision tree $T^* t(V^* t, E^* t)$ where parameter $m$ is supposed equal to 0, 1, 2, 3, 4, separately.

Plugging Eqs. (14)-(20) into this summarized expression $S^* 1 = \sum_{i=1}^{7} S^* (i)$ and implementing some basic arithmetics together outputs the desirable result as said in Eq. (13). This suggests which theorem 3 is sound.

Similarly, we can immediately capture the solutions for geodesic distance $S^* t$ and average geodesic distance $\langle S^* t \rangle$ on the $(1, m)$-star-fractal tree $T^* t(V^* t, E^* t)$, separately, stated in the next corollary and application due to Eq. (13).

**Corollary 3** After $t$ time steps, the solution for geodesic distance $S^* t$ on the $(1, m)$-star-fractal tree $T^* t(V^* t, E^* t)$ will follow

$$S^* t = 2^t (m + 2)^2 S - (2^t - 1) (m + 2) 2^{t-1} (|V|^2 - 2|V| - 1) - \frac{(m + 1) (|V| - 1)}{2} \times \frac{2^{t+1} (m + 2)^2 2^t - 2(m + 2)^t}{2(m + 2) - 1} \quad (21)$$

in which $S$ is a known expression of geodesic distance on tree $T(V, E)$.

**Application 3** After $t$ time steps, the solution for average geodesic distance $\langle S^* t \rangle$ on the first-order subdivision tree $T^* t(V^* t, E^* t)$ will follow

$$\langle S^* t \rangle \approx \frac{2^{t+1} S}{(|V| - 1)^2} - \frac{2^{t+1}}{2m + 3} \frac{(m + 1) 2^{t+1}}{(|V| - 1)} + 1 - 2^t \quad (22)$$

which is completely consistent with simulation results as plotted in Fig. (3).

As described before, there in fact are considerable differences between the first-order subdivision and the $(1, m)$-star-fractal attributed to crucial influence on topological structure of these models generated by both operations. Some of them will be reported in more detail at the rest of this paper. Nevertheless,
it can be quite evident that the two distinct kinds of treelike models built by the two operations have
similar expression of average geodesic distance, see Eq. (8) and Eq. (22). In another word, Eq. (8) can been
regarded as a special case of results told by Eq. (22) when parameter $m$ is supposed equal to zero. This
implies indirectly that the two operations above share similar function on some topological structure
indices of generated treelike models, at least on average geodesic distance.

Till now, the two families of treelike models are established based on an arbitrary tree $T(V,E)$
using the first-order subdivision and the $(1,m)$-star-fractal, respectively. More general, our results have
answered how to determine analytically an exact solution for geodesic distance on treelike models of such
kinds. Therefore, the consequence published in [20] can be viewed as a special example of our results in
which a single edge is selected as the seed of their models. To keep our work self-contained, we still study
geodesic distance on such type of treelike models and precise expression is shown in corollary 4.

**Corollary 4** After $t$ time steps, the solution of geodesic distance $S(t,m)$ on treelike model $N(t,m)$
will follow

$$S(t,m) = \frac{(m^2 + 2m + 1)2^t + 2m + 3}{2m + 3}(m + 2)^{2t-1} + (m + 2)^t - \frac{(m + 2)^2}{2m + 3}(m + 2)^{t-1}.$$  \hspace{1cm} (23)

where the treelike models $N(t,m)$ is a special case of $(1,m)$-star-fractal tree $T^{*t}(V^{*t}, E^{*t})$ whose seed is
no longer an arbitrary tree but a single edge.

Treelike models $N(t,m)$ of such type in fact have been in-depth studied in many published papers
because they show some interesting structure features including factual phenomena [20]. Although Eq. (23)
can be easy proved by letting the initial conditions $S_0 = 1$ and $|V_0| = 2$ of Eq. (21), we will turn out
Eq. (23) to be correct in another fashion based on self-similarity displayed by treelike models $N(t,m)$.
The reasons why we employ calculation method on the basis of self-similar topological structure have
twofold. The one is that self-similarity is one of most prevailing topological structures of models in nature
and real-life world [23]-[25]. The other is to highlight convenience of our novel methods addressed above
in comparison with the commonly used method which we will show below.

**Proof** To smoothly develop the proof for Eq. (23), we have to describe the development process
of treelike models $N(t,m)$ by utilizing another reconstruction method, shown in Fig.4, where we denote
by $\theta_t$ the center vertex (indigo online) that can be obtained by vertex-merging-operation among the
external vertex $\omega_{i-1}$ of branches $N^i(t-1,m)$ ($i \in [1,m+2]$). Indeed, this reconstruction method shows
self-similar structure of treelike models $N(t,m)$ and further allows us to calculate the exact solution of
Fig. 4. The diagram of treelike model $N(t, m)$ which is in essence constructed by $m + 2$ models $N(t - 1, m)$ using vertex-merging-operation.

gedesic distance $S(t, m)$ analytically, as below

$$S(t, m) = (m + 2)S(t - 1, m) + \Omega_{t,m}$$  \hspace{1cm} (24)

where symbol $\Omega_{t,m}$ represents the total sum of gedesic distance of an arbitrary pair of vertices from two different branches $N^i(t - 1, m)$ ($i \in [1, m + 2]$). Obviously, Eq. \hspace{1cm} (24) can be reorganized in an iterative calculation way as follows

$$S(t, m) = (m + 2)^t S(0, m) + \sum_{i=0}^{t-1} (m + 2)^i \Omega_{t-i,m}$$  \hspace{1cm} (25)

where $S(0, m)$ is the gedesic distance of two vertices connected by the original edge as a seed and in fact equals 1.

To successfully capture the closed-form solution of Eq. \hspace{1cm} (25), the left issues is to answer the expressions of $\Omega_{t-i,m}$ ($i \in [0, t - 1]$). Now, we by definition write

$$\Omega_{t,m} = \sum_{1 \leq i < j \leq m + 2} \Omega_{t,m}^{ij} = \frac{(m + 1)(m + 2)}{2} \Omega_{t,m}^{12}$$  \hspace{1cm} (26)

in which we have made use of self-similar structure among branches $N^i(t - 1, m)$ ($i \in [1, m + 2]$). As before, we can by definition obtain
\[ \Omega_{t,m}^{12} = \sum_{v \in V_{t-1}, v \neq \omega_{t-1}^{1} (\text{or } \neq \theta_{t})} d_{v}^{u} \]
\[ u \in V_{t-1}^{2}, u \neq \omega_{t-1}^{2} (\text{or } \neq \theta_{t}) \]
\[ = \sum_{v \in V_{t-1}^{1}, v \neq \omega_{t-1}^{1} (\text{or } \neq \theta_{t})} (d_{v}^{\theta_{t}} + d_{v}^{\theta_{t}u}) \]  \hspace{1cm} (27)
\[ u \in V_{t-1}^{2}, u \neq \omega_{t-1}^{2} (\text{or } \neq \theta_{t}) \]
\[ = 2(|V_{t-1}^{1}|-1)\Theta_{t-1} \]

Here we define \( \Theta_{t-1} \) as the total sum of geodesic distance between the external vertex \( \omega_{t-1}^{1} \) and vertex \( v \neq \omega_{t-1}^{1} (\text{or } \neq \theta_{t}) \) of branch \( N^{1}(t-1,m) \). We take useful advantage of self-similar structure between branches \( N^{1}(t-1,m) \) and \( N^{2}(t-1,m) \) again. Analogously, \( \Theta_{t-1} \) can be written as

\[ \Theta_{t-1} = \sum_{v \in V_{t-1}, v \neq \omega_{t-1}^{1}} d_{v}^{\omega_{t-1}^{1}} \]
\[ = \Theta_{t-2} + \sum_{j \in [2,m+2]} \sum_{i \in \mathcal{V}_{t-2,i}, i \neq \omega_{t-2}^{1}} (d_{i}^{\omega_{t-2}^{j}} + D_{t-2}) \]  \hspace{1cm} (28)
\[ = (m+2)\Theta_{t-2} + (m+1)(|\mathcal{V}_{t-2}|-1)D_{t-2} \]

where symbol \( D_{t-2} \) is the diameter of treelike models \( N(t-2,m) \) and self-similarity among branches \( N^{i}(t-2,m) (i \in [2,m+2]) \) is again used for simplicity. With the similar calculation to Eq.(25), the closed-form of \( \Theta_{t-1} \) can follow

\[ \Theta_{t-1} = (m+1) \sum_{i=0}^{t-2} (m+2)^{i}(|\mathcal{V}_{t-2}-i|-1)D_{t-2}-i \]
\[ + (m+2)^{t-1}\Theta_{0}. \]  \hspace{1cm} (29)

Substituting both initial conditions \( \Theta_{0} = 1 \) and \( D_{t} = 2^{t} \) into Eq.(29) yields

\[ \Theta_{t} = (m+2)^{t-1} + (m+1)(m+2)^{t-2}(2^{t-1}-1). \]  \hspace{1cm} (30)

Armed with Eqs.(25)-(30), the exact solution of geodesic distance on treelike models \( N(t,m) \) may obey
\[ S(t, m) = (m + 2)^t + (m + 2)^{2t-1} - (m + 2)^{t-1} + \frac{(m + 1)^2}{2m + 3} [2^t(m + 2)^{2t-1} - (m + 2)^{t-1}]. \]  

(31)

By some simple arithmetics, Eq. (31) can be induced as the same outline of Eq. (23) which completes our proof.

Here provides two methods for determining the concise solution for geodesic distance on treelike models \( N(t, m) \). While the both computations are proceeded in an iteration manner, the nature concealed by them is completely different from one another. The two techniques, in some sense, have the same impact on calculation process from the complexity point of view, in particular, an edge as seed. If we go into some other situations, for instance, the seed being assigned as a larger tree \( T(V, E) \) on about tens of vertices, the method based on self-similarity seems to become inadequate but our novel ways addressed in this paper can still be adequately employed to work well where only requirement is to know the geodesic distance \( S \) and vertex number \(|V|\) before carrying out our algorithm. Furthermore, our methods are more light to implement than some universally studied ones built by matrix, such as, Laplacian spectral and eigenvectors of underlying structure. One of most important reasons for this is the sparsity of adjacency matrix corresponding to treelike models of such types. Equivalently, while the total number of entries of adjacency matrix of treelike models in question increases exponentially over time, the nonzero are order of magnitude as vertex number. Therefore, some matrix methods may be used to calculate the exact solution for geodesic distance on treelike models discussed here after executing a number of matrix operations. Sometimes, these such operations appear to be complicated because an arbitrary tree \( T(V, E) \) can be considered as a seed. Facing with situations of such kind, one can choose our algorithms to derive desired consequences because our algorithms have no great dependence on the choice of original trees (seeds). Put this another way, they can be quite competent to deal with such tasks in a reasonable time using present laptops as the seed with up to hundreds of vertices.

As before, the average geodesic distance \( \langle S(t, m) \rangle \) on treelike model \( N(t, m) \) can be readily derived by both definition and Eq. (23) and exhibited in the below application.

**Application 4** After \( t \) time steps, the solution of average geodesic distance \( \langle S(t, m) \rangle \) on treelike model \( N(t, m) \) will follow

\[ \langle S(t, m) \rangle \approx \frac{2^{t+1}(m + 1)^2}{(m + 2)(2m + 3)} \propto O(D) \approx |V|^{\gamma} \]

(32)

where exponent \( \gamma \) is equal to \( \frac{\ln 2}{\ln m + \frac{3}{2}} \).
Fig. 5. The diagram of T-graph(3) that is able to be constructed from three T-graph(2)s by applying vertex-merging-operation among three external vertices (colored in yellow online).

Based on statements from Eq. (22) and Eq. (32), we can immediately arrive at the following theorem which says a more general phenomenon.

**Theorem 4** Given an arbitrary tree $T(V,E)$, the exact solution for average geodesic distance $\langle S_t \rangle$ on the $(1,m)$-star-fractal tree $T^*_t(V^*_t,E^*_t)$ is

$$\langle S_t \rangle \approx |V_t|^{\gamma} \propto O(D_t)$$  \hspace{1cm} (33)

where exponent $\gamma = \frac{\ln 2}{\ln (m+2)}$ in the limit of large graph size.

In addition, the best studied case of $(1,m)$-star-fractal tree $T^*_t(V^*_t,E^*_t)$ is in practice the T-graph($t$) which can be created from a single edge as a seed by using $(1,1)$-star-fractal until $t$ time step. Technically, the T-graph($t$) may also be reconstructed from three preceding T-graph($t-1$s $(t \geq 1)$ by vertex-merging-operation, plotted in Fig. 5. Apparently, let parameter $m$ be equivalent to 1 in Eq. (23) and then one may write the next corollary.

**Corollary 5** After $t$ time steps, the solution of geodesic distance $S_t$ on the T-graph($t$) will follow

$$S_t = 3^t + 2t + 5 \times 3^{t-1} - \frac{3^{t+1}}{5}.$$  \hspace{1cm} (34)

At the same time, it is not difficult to directly obtain the fifth application.

**Application 5** After $t$ time steps, the solution of average geodesic distance $\langle S_t \rangle$ on the T-graph($t$) will follow

$$\langle S_t \rangle \approx 2^t \approx |V_t|^{\gamma}, \quad \gamma = \frac{\ln 2}{\ln 3}$$  \hspace{1cm} (35)

where $|V_t|$ is the total number of vertices of the T-graph($t$).

It is worth noticing that the $(1,m)$-star-fractal trees $T^*_t(V^*_t,E^*_t)$ all display some same topological structure properties with each other. These such properties include: (i) Each central vertex added by
(1, m)-star-fractal at any time step $t_i \ (1 \leq t_i \leq t)$ into models $T^*_t(V^*_t,E^*_t)$ has degree $m + 2$, (ii) the total number of leaves of models $T^*_t(V^*_t,E^*_t)$ keep successively changed over time $t_i$, and (iii) each member of model family $T^*_t(V^*_t,E^*_t)$ exhibits homogeneous topological structure. As mentioned above, (1, m)-star-fractal trees $T^*_t(V^*_t,E^*_t)$ just share property (iii) with first-order subdivision trees $T'_t(V'_t,E'_t)$ and show sharply different appearance from the latter due to another two properties, in particular, property (i). Hence, there is significant difference between the both classes of treelike models from the geodesic distance point of view. Except for these difference addressed here, there should be other potential ones only because of distinct topological structure between them which will still wait to unveil in the days to come.

4 POTENTIAL APPLICATIONS

Consider that the above discussions corresponding to average geodesic distance on treelike models under consideration are thought of as some evident results by directly applying our methods, then this section will demonstrate a kind of potential applications by means of our methods, i.e., explicitly determining precise solutions of mean first-passage time (MFPT) for random walks on our treelike models. On the one hand, we indeed obtain desired expressions in perfect agreement with the results previously reported. On the other hand, our methods can perform much better to handle such problems in a general environment than some widely used manners by diminishing redundant computations. To organize the outline of our work narrated below in a self-contained manner, we have to revisit some notations and terminologies relevant to random walks on graph (network).

As the discrete-time representative of Brownian motion and diffusive processes, random walk, which is put first forward by Einstein and Smoluchowski, have proven useful in a wide range of distinct applications. Therefore, random walk keeps quite active at present and still attracts more attention [27]-[32]. As well known, the random walk describes an ideal situation in which a walker (particle) has no information of the underlying graph $G(V,E)$ and just chooses uniformly at random one vertex of its neighbor set to move on, called the unbiased Markov random walks as well. The random walk of this type can be depicted by using Markov chains [33]. For a walker performing random walk on a graph $G(V,E)$, the most fundamental and significant issue to solve is to analytically determine the first-passage time $FPT$ from source (start vertex) to trap (destination vertex).

In the last several decades, the random walks performed on a graph $G(V,E)$ with a single trap have been extensively studied [34, 35] and hence some interesting results have been reported which reveal
some scaling relations and dominating behavior on graph. Nevertheless, more and more researchers think that in some real environment each vertex of graph $G(V, E)$ has very likely to be selected to serve as a trap. It has been proved that the location of traps and underlying topological structure both strongly affects the behavior of random walks. To sufficiently consider effects from the both, mean first-passage time ($MFPT$) can be a commonly adopted measure to describe the efficiency of random walks with the perfect trap uniformly allocated at all vertices on graph $G(V, E)$. The smaller the value $MFPT$ is, the higher the efficiency is, and vice versa. Respective of the obvious importance and ubiquity of random walks themselves, we will study random walks on our models in more detail using electrical network $G^\otimes(V^\otimes, E^\otimes)$ which can been obtained from its underlying graph $G(V, E)$ by placing a unit resistance on its every edge $uv \in E$.

Given a graph $G(V, E)$ of interest, the $MFPT$ is by definition equal to the average over $FPT$s of all pairs of vertices $u$ and $v$ in $V$. While the $FPT$ for any vertex pair can be expressed on the basis of the fundamental matrix corresponding to graph $G(V, E)$, the fundamental-matrix method for calculating the $MFPT$ for random walk on graph $G(V, E)$ is to calculate the inversion of number $|V|$ of matrices with cardinality $(|V| - 1) \times (|V| - 1)$. This evident implies that such type of method is just adequately employed for small graphs and but becomes prohibitively difficult to calculate the quantity for some other graphs with thousands of vertices. To overcome this cumbersome problem related to the fundamental matrix method when facing with large-size graphs, one can make use of another candidate introduced in [36] which is in essence a class of method based on the matrix of graph $G(V, E)$. Although, the issue above to address can easily be induced to calculate the pseudoinverse of the Laplacian matrix $L$ of graph $G(V, E)$ now, which guarantees us to compute the $FPT$ between arbitrary pair vertices $u$ and $v$ directly from the inversion of a single $|V| \times |V|$ matrix. The entry $l_{uv}$ of Laplacian matrix $L$ follows

$$l_{uv} = \begin{cases} -1, & \text{an edge connects vertex } u \text{ to } v \\ k_u, & u = v \\ 0, & \text{otherwise} \end{cases} \quad (36)$$

In the language of matrix theory, the Laplacian matrix $L$ can be compacted as $L = Z - A$ where symbol $A$ is the adjacency matrix of corresponding graph $G(V, E)$ and $Z = diag(k_1, k_2, ..., k_{|V|})$. And then, the pseudoinverse of the Laplacian matrix $L$ is

$$L^* = \left( L - \frac{EE^\top}{|V|} \right)^{-1} + \frac{EE^\top}{|V|} \quad (37)$$

where vector $E = (1, 1, 1, ..., 1_{|V|})^\top$. If let $FPT_{uv}$ denote as the first time token by a walker on graph
In a graph \( \mathcal{G}(V, E) \) to arrive at vertex \( v \) from its source vertex \( u \), then one is able to write

\[
FPT_{uv} = \sum_{i=1}^{\vert V \vert} (l_{ui}^* - l_{uv}^* - l_{vi}^* + l_{vv})l_{ii}
\]  

(38)

here \( l_{ij}^* \) is entry of the matrix \( L^* \) well defined above and \( l_{ii} \) the \( i \)-th entry of the diagonal of the Laplacian matrix \( L \). With the help of knowledge prepared above, the mean first-passage time (MFPT) may be expressed as

\[
MFPT = \frac{1}{\vert V \vert(\vert V \vert - 1)} \sum_{u \neq v \in \mathcal{V}} \sum_{v=1}^{\vert V \vert} FPT_{uv}.
\]  

(39)

By far, Eq.(39) says that the problem of answering MFPT may be reduced to calculate the entries of the pseudoinverse matrix \( L^* \) and since its complexity becomes much lighter than the previous case because of only requirement for inverting a \( \vert V \vert \times \vert V \vert \) matrix. Nevertheless, as shown above, the total number of vertices of each model studied in this paper increases exponentially over time step \( t \), for large \( t \) this matrix-based technique still becomes too headaches to obtain an exact formula for MFPT just because of the need on time and computer memory. Fortunately, the topological structure of each of our models and the relationship effective resistance to \( FPT \) together help us to accomplish analytically the computation of MFPT [39]. As an illustrative example, the effective resistance \( R_{uv} \) between a pair of vertices \( u \) and \( v \) in a electrical network \( \mathcal{G}^\otimes(\mathcal{V}^\otimes, \mathcal{E}^\otimes) \) can be not hard transformed to calculate the \( FPT_{uv} \) between the same vertex pair on its corresponding underlying graph \( \mathcal{G}(V, E) \), i.e., \( R_{uv} = (FPT_{uv} + FPT_{vu})/2\vert E \vert \). In practice, the expression of numerator, \( FPT_{uv} + FPT_{vu} \) is customarily viewed as the commute time \( C_{uv} \) between vertices \( u \) and \( v \), defined as \( C_{uv} = FPT_{uv} + FPT_{vu} \). In other words, \( R_{uv} \) is denoted by \( C_{uv}/2\vert E \vert \).

Armed with the statements stressed above, Eq.(39) may be reorganized as

\[
MFPT = \frac{1}{\vert V \vert} \sum_{u \neq v \in \mathcal{V}} \sum_{v=1}^{\vert V \vert} R_{uv}
\]  

(40)

here we already take equality \( C_{uv} = C_{vu} = 2\vert E \vert R_{uv} \) for any couple of vertices \( u \) and \( v \). The preceding Eq.(39) further tells us that the problem of how to determine the mean first-passage time (MFPT) for random walk on graph \( \mathcal{G}(V, E) \) can be immediately switched to calculate the effective resistance \( R_{uv} \) on its corresponding electrical network \( \mathcal{G}^\otimes(\mathcal{V}^\otimes, \mathcal{E}^\otimes) \)

Given a general graph \( \mathcal{G}(V, E) \), the complexity of effective resistance computation of its corresponding electrical network \( \mathcal{G}^\otimes(\mathcal{V}^\otimes, \mathcal{E}^\otimes) \) is still to invert a \( \vert V \vert \times \vert V \vert \) matrix which is in some extent the same that of determining the MFPT as said in Eq.(39). This appears to not lighten our workload. However,
we want to stress that Eq. (40) is versatile for any graph $G(V,E)$ of interest. For some graphs with specific topological structure, this transformation addressed in Eq. (40) has most possible to be reduced to work well. For instance, as the simplest and most fundamental connected graph, tree highlights the convenience and significance of transformation of such type by supporting a fact that the effective resistance of arbitrary two distinct vertices is completely equivalent to the geodesic distance between this vertex pair. Since then, we are allowed to use the lights shed by Eq. (40) to find out rigorous expression of $MFPT$ for each member of our models mainly because of its own treelike structure.

From now on, let us divert insights into discussing the mean first-passage time ($MFPT$) on two types of treelike models, $T'_t(V', E'_t)$ and $T^*_t(V^*_t, E^*_t)$ using Eq. (40). The closed-form solutions for $MFPT$ for random walk on them are reported in the next theorems in form.

**Theorem 5** Given an arbitrary tree $T(V,E)$, the solution for the mean first-passage time $MFPT'_t$ on its first-order subdivision tree $T'_t(V', E'_t)$ will follow

$$MFPT'_t \approx \frac{2^{2t+2}S}{(|V| - 1)} - \frac{2^{2t+2}}{3} + (2 - 2^{2t+1})(|V| - 1).$$  \hspace{1cm} (41)

For a tree $T(V,E)$ with finite number of vertices, according to Eq. (11), Eq. (41) may be asymptotically expressed as

$$MFPT'_t \approx 2^{2t} |V| \propto O(V'^{\lambda'_t})$$  \hspace{1cm} (42)

where exponent $\lambda'$ is equal to 2 in the limit of large graph size. This means that for first-order subdivision tree $T'_t(V', E'_t)$ the $MFPT'_t$ is markable power function correlation with its order. The more the vertex number, the more the $MFPT'_t$. At the moment, it is apparent to find out an equality $\lambda' = 1 + \gamma'$. By analogously computation as Eq. (41), one can have

**Theorem 6** Given an arbitrary tree $T(V,E)$, the solution for the mean first-passage time $MFPT^*_t$ on its $(1,m)$-star-fractal tree $T^*_t(V^*_t, E^*_t)$ will follow

$$MFPT^*_t \approx \frac{4(4 + 2m)^tS}{(|V| - 1)} - \frac{4(4 + 2m)^t(m + 1)}{2m + 3} + (2 - 2^{t+1})(2 + m)^t(|V| - 1).$$  \hspace{1cm} (43)

As before, for the large value of time $t$ and finite-size tree $T(V,E)$, the $MFPT^*_t$ will have an asymptotical relationship with its order $|V^*_t|$, as follows

$$MFPT^*_t \approx (4 + 2m)^t |V| \propto O(V^*^{\lambda^*_t})$$  \hspace{1cm} (44)
where exponent $\lambda^* = 1 + \gamma^*$ in the large graph size limit. This shows that the $MFPT_t^*$ for random walk on \((1,m)\)-star-fractal tree $T^*_t(V^*_t, E^*_t)$ grows as a power-law function of its order $|V^*_t|$ which is in perfect agreement with the published result for such kind of fractals in [20].

To distinguish in-depth how topological structures of the both classes of treelike models affect mean first-passage time for random on them, we here need to review some prominent topological structure properties of the both. As known, as the special case of the \((1,m)\)-star-fractal tree $T^*_t(V^*_t, E^*_t)$, the T-graph is a fractal with the fractal dimension $d_f = \frac{\ln 3}{\ln 2}$ and the random-walk dimension $d_w = \frac{\ln 6}{\ln 2} = 1 + d_f$. In the meantime, the spectral dimension of T-graph is $\tilde{d} = \frac{2d_f}{d_w} = \frac{\ln 9}{\ln 6} < 2$, suggesting which a random walk on it is persistent [40]. Similarly, our model $T^*_t(V^*_t, E^*_t)$ has the fractal dimension $d_f^* = \frac{\ln 6}{\ln 2(m + 2)}$, the random-walk dimension $\lambda^* = \frac{\ln 2(m + 2)}{\ln 2} = 1 + d_f$ and the spectral dimension $\tilde{d}^* = \frac{2d_f^*}{\lambda^*} = \frac{\ln 9}{\ln 2(m + 2)} < 2$.

Therefore, a random walk on our model $T^*_t(V^*_t, E^*_t)$ is persistent also. By contrast, treelike models $T'_t(V'_t, E'_t)$ do not show fractal phenomena and so have a lack of such rich properties.

5 CONCLUSION AND DISCUSSION

To summarize, based on two different types of operations, i.e., first-order subdivision and \((1,m)\)-star-fractal, we generate two classes of treelike models whose seed can not be limited to a single edge connecting a pair of vertices but be an arbitrary tree $T(V, E)$. This implies, in some sense, our work covers the case of a single edge. In addition, we propose a family of novel and useful enumeration methods for determining the exact solution for geodesic distance on each member of treelike models under consideration. In comparison with some commonly used methods for such problems, for instance, Laplacian spectral and eigenvalue, our techniques not only provide what we want, but also reduce a significant amount of computations, i.e., diminishing the demand of redundant time and space memory when considering this type of treelike models with thousands of vertices or more. At the same time, thanks to the special topological structure of models $N(t,m)$ as a special case of the \((1,m)\)-star-fractal tree $T^*_t(V^*_t, E^*_t)$, we derive an exact formulas for its geodesic distance and average geodesic distance using self-similar method, respectively. This is in strong agreement with these results using our techniques addressed in this paper. This further suggests the convenience and correctness of our methods.

To highlight the potential applications of our methods, by connection between random walk and electrical network, we study the mean first-passage time ($MFPT$) for random walks on treelike models built here. Indeed, we obtain the precise expression of $MFPT$ for a walker allocated on each treelike model in question. While a part of these consequences have been captured in some published papers in
another general manner as in [20], this does not erase our contribution. As above, two of reasons for this are to have a chance to choose an arbitrary tree as a seed to create our desired models and to cut off a heavy number of computation needs.

However, we want to express that our work is only a tip of the iceberg and our methods are just applied to treelike models under consideration to analytically obtain exact solutions for some quantities, for instance, (average) geodesic distance as well mean first-passage time as reported in this paper. On the other hand, we believe that the lights shed by our technique can be helpful to direct future related research. Meantime, referring to Ockham’s razor we would like to point out that it is important and necessary to develop some new and professional methods for answering some special cases when people attempt to find out more general ways for addressing some kind of scientific issues. Along this research pathes of such type, we will continue to study much more general models [25] and also wish to witness some other applications of the thoughts reflected by our work in different science areas in the next future.

ACKNOWLEDGMENT

The research was supported by the National Key Research and Development Plan under grant 2017YFB1200704 and the National Natural Science Foundation of China under grant No. 61662066.

References

[1] S. Angelopoulos, D. Arsenio and C. Durra. Theoretical Computer Science. 670 (2017) 11-22.
[2] J.Y. L, C. Yang, C.J. Fu, Y.C Gao and H.C Yang. Chaos 28 (2018) 113116.
[3] W. Huang, S.Y. Chen and W.L. Wang, Physica A. 393 (2014) 132-154.
[4] L.P. Wackett. Environmental Microbiology. 21 (2019) 1511-1512.
[5] M. Golosovsky, S. Solomon. Phys. Rev. E. 95 (2017) 012324.
[6] J. Falk, L. Bronstein, M. Hanst, B. Drossel, and H. Koepppl. J. Chem. Phys. 150 (2019) 024106.
[7] M. Galiceanu, A.S. Reis, and Maxim Dolgushev. J. Chem. Phys. 141 (2014) 144902.
[8] U, Brose, P. Archambault, A.D. Barnes, et al. Nature Ecology Evolution. 3 (2019) 919-927.
[9] J.P. Casasnovas, M. Gerlach, N. Aguirre and L.A.N. Amaral. Nature Human Behaviour. (2019).
[10] F. Ma, P. Wang and B. Yao. (Submitted). arXiv:1904.11314v1.
[11] X.D. Luo, F. Ma and W.T. Xu. (Submitted).
[12] J.H. Peng, G.A. Xu. J. Chem. Phys. 140 (2014) 134102.
[13] A. Clauset, C. Moore, and M.E.J. Newman. Nature. 453 (2008) 98-101.
[14] S.D. Palma, P.L. Lanzi. IEEE Transactions on Games. 10 (2018) 317-332.
More generally, the ratio $\varpi_t$ of the leaf number $|L_t|$ of model $T_t(V_t, E_t)$ and its order $|V_t|$ will tend to be zero for the large value of time $t$, i.e., $\varpi_t = \lim_{t \to \infty} \frac{|L_t|}{|V_t|} \to 0$.

Generally speaking, the ratio $\varpi_t^*$ of the leaf number $|C_t^*|$ of model $T^*_t(V^*_t, E^*_t)$ and its order $|V^*_t|$ will tend to be some constant dependent on parameter $m$ of the (1, $m$)-star-fractal for the large value of time $t$, i.e., $\varpi_t^* = \lim_{t \to \infty} \frac{|C_t^*|}{|V_t^*|} \to \frac{m}{m+1}$. Obviously, the minimal value for $\varpi_t^*$ can be arrived at point $m = 0$ which is in perfect agreement with that of [22]. The upper bound of $\varpi_t^*$ is 1 as $m \to \infty$. Meantime, there is a relationship close connecting ratio $\varpi_t^*$ with exponent $\gamma^*$, namely, $\gamma^* = \frac{\ln 2}{\ln (1 + \varpi_t^*) - \ln \varpi_t^*}$, for the large value of parameter $m$.

Fei Ma receives the M.S degree in the College of Mathematics and Statistics from Northwest Normal University, Lanzhou, CHINA. He is currently working toward the PhD degree in the School of Electronics
Engineering and Computer Science, Peking University, Beijing, CHINA. His current research interests include Graph theory with applications, Random walks, Complex network, and Privacy-preserving. He is a Reviewer of Mathematical Reviews.

**Ping Wang** received his doctorate degree in Computer Science from the University of Massachusetts, USA in 1996. He is currently a professor at Peking University, China. Dr. Wang has authored or co-authored over 80 papers in journals or proceedings such as IEEE TDSC, ACM CCS, IEEE ICWS and IEEE TII, USENIX Security, etc. His research interests include Internet of Things, Distributed Computing, and Information Security. He is an IEEE senior member.

**Xudong Luo** is now studying for the Ph.D degree in the School of Mathematics and Statistics, Lanzhou University, CHINA. His interests include Infinite dimensional dynamical systems and Synchronization in complex system.