TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214
DOI: 10.26599/TST.2018.9010040

On the Mathematical Nature of Wireless Broadcast Trees

Fulu Li, Junwei Cao*, Chunfeng Wang and Kui Wu

Abstract: Trees are arguably one of the most important data structures that are widely used in information theory and computing science. For example, different number of intermediate nodes in wireless broadcast trees may have great impact on energy consumption of each node that are typically equipped with limited power supplies in a wireless sensor network, which may eventually determine how long the given wireless sensor network can last. Thus, it is of great importance to have a deep understanding of the mathematical nature of wireless broadcast trees. In this paper, we give a new proof of Cayley's famous theorem for counting labeled trees. A distinct feature in this proof is that we purely use combinatorial structures instead of constructing a bijection between two kinds of labeled trees as almost all of the existing proofs do [1,8,9]. Another contribution of this work is the presentation of a new theorem on trees based on the number of *intermediate* nodes in the tree. To be best of our knowledge, it is the first time that a tree enumeration theorem based on the number of *intermediate* nodes in the tree has been presented.

Key words: Energy Efficiency; Tree Theorem; Broadcast Trees; Wireless Transmissions

1 Introduction

Notably, trees are arguably one of the most important data structures that are widely used in information theory and computing science. From B^+ trees that are used in almost every major database system's indexing and queries, to decision and strategy trees that are used in probabilistic inferences such as the one used in Google's AlphaGo system that defeated human champions in recent human-machine grand challenge, trees are almost the most-sought utilities to depict and represent a variety of choice-making strategies. For example, different number of *intermediate* nodes i.e., the non-leaf nodes except the root node, in wireless broadcast trees may have great impact on energy consumption of each node that are typically equipped with limited power supplies in a wireless sensor network, which may eventually determine how long the given wireless sensor network can last. It is of great importance to have a deep understanding of the mathematical nature of wireless broadcast trees.

• Fulu Li, and Junwei Cao are with the Research Institute of Information Technology, Tsinghua University, Beijing, P. R. China.
• Chunfeng Wang is with the Qian Xuesen Laboratory of Space Technology, China Academy of Space Technology, Beijing, P. R. China.
• Kui Wu is with the Department of Computer Science, University of Victoria, BC, Canada.
∗ Corresponding address: Junwei Cao, FIT 4-415, Tsinghua University. E-mail: jcao@tsinghua.edu.cn
Manuscript received: 2017-06-26; accepted: 2017-11-30

The studies of trees have captured the imaginations of some of the most prominent minds in history. In 1889, British mathematician Arthur Cayley, who helped found the modern British school of pure mathematics, published his famous theorem on trees. Donald Knuth, a Turing award winner and a professor at Stanford University, and Peter Shor, a winner of Nevanlinna Prize, a professor at MIT and the inventor of Shor's algorithm that is widely regarded as a major breakthrough for quantum algorithm for integer factorization, all explored deep understanding of Cayley's theorem on trees [4,9].

The rest of the paper is organized as follows: we revisit the free tree theorem presented by Cayley in his landmark work in 1889 and present our new results of tree enumeration theorems based on the number of *intermediate* nodes in the tree (see Equation 10) in Section 2. We examine the n^m series and its combinatorial structures in Section 3, which pave the way for our new proof of Cayley's theorem in Section 4. We give a formal proof of the NP-hardness of wireless broadcast problems towards the optimization of certain objectives in Section 5. The conclusions are given in Section 6.

2 Free Tree Theory Revisited

According to Cayley's theory, there are n^{n-2} distinct labeled free trees on n vertices [2]. For if X is a particular vertex, the free trees are in one-to-one correspondence with oriented trees having root X [4]. Therefore, for the case with one root node S , which is also called the source node in wireless broadcast trees [5,6], and N non-root nodes, which are also called destination nodes in wireless broadcast trees, we have $(N+1)^{(N-1)}$ distinct labeled free trees.

In the following, we give another look at Cayley's free tree theory based on the number of *intermediate* nodes in the tree, i.e., the non-leaf nodes except the root node, which are also called *relaying nodes* in wireless broadcast trees. The motivation behind this work is that the number of *intermediate* nodes in the tree often plays an important role in the design of energy-efficient communication protocols in wireless ad hoc networks and in particular on the construction of energy-efficient broadcast trees in multi-hop wireless ad hoc networks [5,6]. As stated at the beginning of this section, we consider a root node S and N non-root nodes. The number of possible trees with zero *intermediate* nodes is given by:

$$R(0) = \binom{N}{N} = 1 \tag{1}$$

Namely, for the case with zero *intermediate* nodes, i.e., all the non-root nodes are leaf nodes, the unique tree is a hub-like tree with every non-root node as direct child node of the root node.

The number of possible trees with one *intermediate* node is given by:

$$\begin{aligned} R(1) = &\binom{N}{1} \times (\binom{N-1}{1} + \binom{N-1}{2} + \\ &\binom{N-1}{3} + \dots + \binom{N-1}{N-2} + \binom{N-1}{N-1}) \\ = &\binom{N}{1} \times \sum_{i=1}^{N-1} \binom{N-1}{i} \end{aligned} \tag{2}$$

The above formula indicates that we first need to pick one among N non-root nodes as the *intermediate* node. Then we will decide which nodes are direct child nodes of the *intermediate* node (the remaining nodes are direct child nodes of the root node S). Notably, the *only* intermediate node has to be the direct child node of the root node, thus the maximum number of direct child nodes of the *intermediate* node in this case is $(N-1)$ and the minimum number of direct child nodes of the *intermediate* node is one.

Similarly, the number of possible trees with two *intermediate* nodes is given by:

$$R(2) = \binom{N}{2} \times \sum_{i=1}^{N-2} (\binom{N-1}{i} \times \sum_{j=1}^{N-1-i} \binom{N-1-i}{j}) \tag{3}$$

Roughly, Equation (3) states that first we need to decide how many choices there exist to pick up the two intermediate nodes among N non-root nodes, then we decide how many of the remaining non-root nodes except those two intermediate nodes are directly reached by one of the intermediate nodes and how many of the remaining non-root nodes are directly reached by the other intermediate node.

By analogy, the number of possible trees with i *intermediate* nodes can be given as follows:

$$\begin{aligned} R(i) = &\binom{N}{i} \times \sum_{k_1=1}^{N-i} (\binom{N-1}{k_1} \times \sum_{k_2=1}^{N-1-k_1-(i-2)} (\binom{N-1-k_1}{k_2} \times . \\ &.. \times \sum_{k_m=1}^{N-1-\sum_{j=1}^{m-1} k_j-(i-m)} (\binom{N-1-\sum_{j=1}^{m-1} k_j}{k_m} \times \\ &... \times \sum_{k_i=1}^{N-1-\sum_{j=1}^{i-1} k_j} \binom{N-1-\sum_{j=1}^{i-1} k_j}{k_i}))...) \end{aligned} \tag{4}$$

Simply put, Equation (4) states the same straightforward philosophy as that of Equation (3). Essentially, the maximum number of direct child nodes of the m^{th} picked relaying node, i.e., $(N-1-\sum_{j=1}^{m-1} k_j - (i-m))$ ($\sum_{j=1}^{m-1} k_j$ is the sum of the directly child nodes of the first $(m-1)$ relaying nodes), guarantees the basic principle for the subsequent *intermediate* nodes that each *intermediate* node must have at least one direct child node. While we will provide a simplified formula later, this seemingly clumsy formula gives a different view and the number be easily calculated by a computer program due to its special structures.

Notably, among all the possible trees with one root node and N non-root nodes, the number of *intermediate* nodes ranges from zero to $(N-1)$. For the extreme case with $(N-1)$ *intermediate* nodes, among N non-root nodes only one node is the *non-intermediate* node, i.e., the leaf node. In theory, all of the possible trees can be categorized into *zero-intermediate-node* trees, *one-intermediate-node* trees, *two-intermediate-nodes* trees, and so forth until $(N-1)$-*intermediate-nodes* trees and they are *mutually exclusive* and *collectively exhaustive*. Therefore, the total number of possible trees according to the classification based on the number of *intermediate* nodes is summarized as:

$$\begin{aligned}
T_{total}(N) = \sum_{i=0}^{N-1} R(i) = R(0) + \sum_{i=1}^{N-1} \binom{N}{i} \times \\
(\sum_{k_1=1}^{N-i} (\binom{N-1}{k_1} \times \sum_{k_2=1}^{N-1-k_1-(i-2)} (\binom{N-1-k_1}{k_2} \times \\
\dots \times \sum_{k_m=1}^{N-1-\sum_{j=1}^{m-1} k_j-(i-m)} (\binom{N-1-\sum_{j=1}^{m-1} k_j}{k_m} \times . \\
.. \times \sum_{k_i=1}^{N-1-\sum_{j=1}^{i-1} k_j} \binom{N-1-\sum_{j=1}^{i-1} k_j}{k_i}))\dots)
\end{aligned} \tag{5}$$

The above analysis provides a novel view on the tree enumeration theory based on the number of *intermediate* nodes, which has crucial impact on some emerging applications, e.g., the power consumption and signal interference analysis in the broadcast protocols in multi-hop wireless ad hoc networks [5,6].

The following table illustrates the relationship between the total number of trees and the number of trees with different number of relaying nodes as dictated in Formula (5). $T(i,N)$ denotes the number of trees with i intermediate nodes and N non-root nodes.

While we will use the clumsy Eq. (4) and Eq. (5) in the final proof of Cayley's formula in Section III, we also give a simplified form of it by virtue of the results on labeled rooted trees with degree sequence given by Goulden and Jackson in [3].

Regarding the number of labeled trees with i intermediate nodes in the tree, we have the following theorem:

Theorem 1.1: $R(i)$ denote the number of labeled trees with i intermediate nodes in the tree, we have

$$R(i) = \binom{N}{i} \sum_{k=1}^{i+1} (\binom{i}{k-1} (-1)^{i+1-k} k^{(N-1)})$$

Proof: following [3], we first define the degree of a vertex v to be the number of edges incident to vertex v, and a sequence $r = (r_1, r_2, \dots)$ of non-negative integers, where r_i is the number of vertices that have degree i , is the type of labeled rooted trees with $N+1$ vertices if and only if

$$\sum r_i = N+1 \qquad \sum i \times r_i = 2N \tag{6}$$

To make it self-contained, let us recall the results on labeled rooted trees with degree sequence in [3].

In [3], it states that the number of labeled **rooted** trees with degree sequence $(r_1, r_2, \dots)$ (i.e., r_j vertices are of degree j , for $j \geq 1$), is

$$\frac{(N+1)(N-1)!(N+1)!}{\prod_{j\geq 1} r_j!(j-1)!^{r_j}} for N \geq 1, \tag{7}$$

where $r_1+r_2+\dots = N+1$, $r_1+2r_2+3r_3+\dots = 2N$.

Thus, the number of labeled trees with degree sequence $(r_1, r_2, \dots)$ (i.e., r_j vertices are of degree j , for $j \geq 1$), is

$$\frac{(N-1)!(N+1)!}{\prod_{j\geq 1} r_j!(j-1)!^{r_j}} for N \geq 1, \tag{8}$$

where $r_1+r_2+\dots = N+1$, $r_1+2r_2+3r_3+\dots = 2N$.

Further, let $R(i)$ denote the number of labeled trees with i intermediate nodes in the tree. We have

$$R(i) = \sum \frac{(N-1)!(N+1)!}{\prod_{j\geq 1} r_j!(j-1)!^{r_j}} for N \geq 1, \tag{9}$$

where the sum is over all degree sequence $(r_1, r_2, \dots)$ such that r_j is a non-negative integer for each j , and $r_1 = N-i$ if the degree of the root node is greater than 1 and $r_1 = N-i+1$ if the degree of the root node is 1, $r_1+r_2+\dots = N+1$, $r_1+2r_2+3r_3+\dots = 2N$.

Further, by virtue of Lagrange Theorem [3], we have

$$R(i) = \binom{N}{i} \sum_{k=1}^{i+1} (\binom{i}{k-1} (-1)^{i+1-k} k^{(N-1)}) \tag{10}$$

Notably, Eq. (5) and Eq. (10) show different aspects on the same counting problem.

3 The n^m Series and its Combinatorial Structures

To give the proof of formula (5) against Cayley's theorem, first let us have a look at the interesting combinatorial structures of the n^m series, which will be used in the proof in Section III. In the following, we give the specific combinatorial structures of n^2 , n^3 and n^4 , the combination of which covers all the new combinatorial structures based on which we will derive the generalized form for n^m.

$$\binom{n}{1}+\underline{\binom{n}{2}\times\binom{2}{1}}=n^2 \tag{11}$$

$$\binom{n}{1}+\underline{\binom{n}{2}\times(\binom{3}{1}+\binom{3}{2})}+\underline{\binom{n}{3}\times\binom{3}{1}\times\binom{2}{1}}=n^3 \tag{12}$$

$$\begin{aligned}&\binom{n}{1}+\underline{\binom{n}{2}\times(\binom{4}{1}+\binom{4}{2}+\binom{4}{3})}\\&\underline{+\binom{n}{3}\times(\binom{4}{1}\times(\binom{3}{1}+\binom{3}{2})+\binom{4}{2}\times\binom{2}{1})}\\&\underline{+\binom{n}{4}\times\binom{4}{1}\times\binom{3}{1}\times\binom{2}{1}}=n^4\end{aligned} \tag{13}$$

We give the proof for the Equation (11), (12) and (13) respectively, based on which we will derive a generalized form for the similar combinatorial structures of n^m.

Proof of Equation (11): We have n distinct balls and we want to pick two balls at random sequentially with replacement. The right-hand side of Equation (11) is the number of ways of picking two balls sequentially from n distinct balls based upon our rules. On the other hand, the possible number of ways of picking two balls at random among these n distinct balls according to the rule could include either picking the same ball twice or picking two different balls. The first component in the left-hand side of Equation (11), i.e., $\binom{n}{1}$, is the number of ways that the two picked balls are the same ball. The second component in the left-hand side of Equation (11), i.e., $\binom{n}{2}\times\binom{2}{1}$, is the number of ways that the two picked balls are different, which means that we can first pick two balls from n distinct balls at one time and then pick one from the two balls as the first-picked ball and the other as the second-picked ball.

Following the same rule as above, we can see that the right-hand side of Equation (12) is the number of ways of picking three balls sequentially from n distinct balls at random with replacement. On the other hand, the possible number of ways of picking three balls among n distinct balls according to the rule could include the following three categories: (a) picking the same ball for three times; (b) picking two different balls with one ball picked twice (two balls are the same among the three picked balls); (c) picking three different balls. The first component in the left-hand side of Equation (11), i.e., $\binom{n}{1}$, is the number of ways that the three picked balls are the same ball. The second component in the left-hand side of Equation (12), i.e., $\binom{n}{2}\times(\binom{3}{1}+\binom{3}{2})$, is the number of ways that there are exactly two different types of balls among the three picked balls (two balls are the same),which means that we first pick two balls from n distinct balls at one time and then there are two placement options for the first to-be-picked ball among the two balls: (a) put it in one of the three slots, i.e., $\binom{3}{1}$, in which case it appears only once (the other one takes the remaining slots and it appears twice); (b) it appears in two of the three slots, i.e., $\binom{3}{2}$, in which case it appears twice (the other one takes the remaining slots and it appears only once). The third component in the left-hand side of Equation (12), i.e., $\binom{n}{3}\times\binom{3}{1}\times\binom{2}{1}$, is the number of ways that the three picked balls are

Table 1 Examples of $T(i,N)$ according to Eq. (5)

i/N	1	2	3	4
0	$\binom{1}{1}=1$	$\binom{2}{2}=1$	$\binom{3}{3}=1$	$\binom{4}{4}=1$
1		$\binom{2}{1}=2$	$\binom{3}{1}\times(\binom{2}{1}+\binom{2}{2})$=9	$\binom{4}{1}\times(\binom{3}{1}+\binom{3}{2}+\binom{3}{3})=28$
2			$\binom{3}{2}\times\binom{2}{1}=6$	$\binom{4}{2}\times(\binom{3}{1}\times(\binom{2}{1}+\binom{2}{2})+\binom{3}{2}\times\binom{1}{1})=72$
3				$\binom{4}{3}\times\binom{3}{1}\times\binom{2}{1}\times\binom{1}{1}=24$
$\sum_i T(i,N)$	$(1+1)^{(1-1)}=1$	$(2+1)^{(2-1)}=1+2=3$	$(3+1)^{(3-1)}=1+9+6=16$	$(4+1)^{(4-1)}=1+28+72+24=125$

different from each other, which means that we first pick three balls from n distinct balls at one time and then pick one from the three balls as the first-picked ball and then pick one more from the remaining two balls as the second-picked ball and the last one as the third-picked ball.

Similarly, we can see that the right-hand side of Equation (13) is the number of ways of picking four balls successively from n distinct balls at random with replacement. On the other hand, the possible number of ways of picking four balls among n distinct balls according to the rule could include the following four categories: (a) picking the same ball for four times; (b) picking two different balls (there are two different types of balls among the four picked balls); (c) picking three different balls (two balls are the same among the four balls); (d) picking four different balls. The first component in the left-hand side of Equation (13), i.e.,$\binom{n}{1}$, is the number of ways that the four picked balls are the same ball. The second component in the left-hand side of Equation (13), i.e., $\binom{n}{2} \times (\binom{4}{1} + \binom{4}{2} + \binom{4}{3})$, is the number of ways that there are two exactly different types of balls among the four picked balls, which means that we first pick two balls from n distinct balls at one time and then there are three possible placement options for the first to-be-picked ball among the two balls: (a) put it in one of the four slots, i.e., $\binom{4}{1}$, in which case it appears only once (the other ball takes the remaining slots and it appears for three times); (b) it appears in two of the four slots, i.e., $\binom{4}{2}$, in which case it appears twice (the other ball takes the remaining slots and it also appears twice); (c) it appears in three of the four slots, i.e.,$\binom{4}{3}$, in which case it appears for three times in the four slots (the other one takes the remaining slot and it appears only once). The third component in the left-hand side of Equation (13), i.e., $\binom{n}{3} \times (\binom{4}{1} \times (\binom{3}{1} + \binom{3}{2}) + \binom{4}{2} \times \binom{2}{1})$, is the number of ways that there are exactly three different types of balls among the four picked balls, which means that we first pick three balls from n distinct balls at one time and then there are two possible placement categories for the first to-be-picked ball among the three balls: (1) it appears only once in the four slots in which case we first pick one slot for this ball, i.e.,$\binom{4}{1}$, then the second to-be-picked ball among the remaining two balls has two choices: (a) it appears only once among the remaining three slots (the last ball takes the remaining slots and it appears twice), i.e.,$\binom{3}{1}$; (b) it appears twice among the remaining three slots (the last ball takes the remaining slot and it appears only once), i.e.,$\binom{3}{2}$. (2) it appears twice in the four slots in which case we first pick two slots for this ball, i.e.,$\binom{4}{2}$, and then we pick one slot from the remaining two slots for the second to-be-picked ball (the last ball takes the remaining one slot), i.e.,$\binom{2}{1}$. The fourth component in the left-hand side of Equation (13), i.e.,$\binom{n}{4} \times \binom{4}{1} \times \binom{3}{1} \times \binom{2}{1}$, is the number of ways that the four picked balls are different from each other, which means that we first pick four balls from n distinct balls at one time and then pick one from the four balls as the first-picked ball and then pick one more from the remaining three balls as the second-picked ball and then pick one more from the remaining two balls as the third-picked ball and the last one as the fourth-picked ball.

Based on the above observations for the combinatorial structures of n^2,n^3 and n^4, we derive a generalized form for the combinatorial structure of n^m as follows:

Table 2 Some decomposition examples of n^m based on Eq. (17).

n^m	$\sum \frac{m!}{\prod_{i=1}^{n} a_i!}$
2^2	1+2+1 = 4
2^3	1+3+3+1 = 8
2^4	1+4+6+4+1 = 16
3^2	1+1+1+2+2+2 = 9
3^3	1+3+3+3+3+3+3+6+1+1 = 27
3^4	1+1+1+4+4+4+4+4+4+6+6+6+12+12+12 = 81
4^2	1+1+1+1+2+2+2+2+2+2 = 16
4^3	1+1+1+1+3+3+3+3+3+3+3+3+3+3+3+3+6+6+6+6 = 64

$$
\begin{aligned}
&\binom{n}{1}+\binom{n}{2}\times(\binom{m}{1}+\binom{m}{2}+...\binom{m}{m-1})+\\
&\binom{n}{3}\times\{\binom{m}{1}\times(\binom{m-1}{1}+\binom{m-1}{2}...+\\
&\binom{m-1}{m-2})+\binom{m}{2}\times(\binom{m-2}{1}+\binom{m-2}{2}+\\
&...+\binom{m-2}{m-3})+...+\binom{m}{m-3+1}\times\binom{2}{1}\}+\\
&...+\binom{n}{m}\times\binom{m}{1}\times\binom{m-1}{1}\times...\times\binom{2}{1}\\
&=n^m
\end{aligned}
\tag{14}
$$

In a more compact form, we have:

$$
\begin{aligned}
&\sum_{i=1}^{m}\binom{n}{i}\times(\sum_{k_1=1,i\geq 2}^{m-1}\binom{m}{k_1}\times...\times\\
&(\sum_{k_s=1,1<s<i-1}^{m-\sum_{j=1}^{s-1}k_j-1}\binom{m-\sum_{j=1}^{s-1}k_j}{k_s}\times...\times\\
&(\sum_{k_{(i-1)}=1}^{m-\sum_{j=1}^{i-2}k_j-1}\binom{m-\sum_{j=1}^{i-2}k_j}{k_{(i-1)}})...)=n^m
\end{aligned}
\tag{15}
$$

By analogy, the Equation (14) can also be interpreted as the number of ways of picking m balls from n distinct balls successively with the pick-and-put-back rule (every time picking one from the same n distinct balls). The left-hand side of Equation (14) indicates the sum of the number of ways of picking exactly i different balls in the m picking trials.

While we will use the clumsy Eq. (14) in the final proof of Cayley's formula in Section III, we give another simplified form of the combinatorial structures of n^m by looking at this from a different angle.

Following [9], we first define $\binom{m}{a_1, a_2, ..., a_n}$ as the coefficient of $x_1^{a_1}x_2^{a_2}...x_n^{a_n}$ in $(x_1+x_2+...+x_n)^m$ and we have

$$
\binom{m}{a_1, a_2, ..., a_n}=\frac{m!}{a_1!a_2!...a_n!}=\frac{m!}{\prod_{i=1}^{n}a_i!}
\tag{16}
$$

Essentially, n^m is the sum of the multi-nomial coefficients of $(x_1+x_2+...+x_n)^m$ where $x_1 = x_2 = ... = x_n = 1$.
Therefore, we have

$$
n^m=\sum\frac{m!}{\prod_{i=1}^{n}a_i!}
\tag{17}
$$

where the sum is over all $(a_1, a_2, ...a_n)$ such that a_i is a non-negative integer for each i and $a_1+a_2+...+a_n = m$.

The following table illustrates some examples of n^m based on the Eq. (17).

4 The Proof of Cayley's Theorem

In this section, we prove that Formula (5), which presents the sum of the number of all possible trees based on the number of *intermediate* nodes in the trees, coincides with Cayley's free tree theorem.

To better understand the procedure of the proof, we first give an example of the proof when $N = 4$. According to Formula (5), we have

$$
\begin{aligned}
&T_{total}(4)=R(0)+R(1)+R(2)+R(3)=\\
&\binom{4}{4}+\binom{4}{1}\times(\binom{3}{1}+\binom{3}{2}+\binom{3}{3})+\\
&\binom{4}{2}\times(\binom{3}{1}\times(\binom{2}{1}+\binom{2}{2})+\binom{3}{2}\times\binom{1}{1})\\
&+\binom{4}{3}\times\binom{3}{1}\times\binom{2}{1}\times\binom{1}{1}
\end{aligned}
\tag{18}
$$

Now note that $\binom{k}{k}=\binom{3}{0}$ for all $1 \leq k \leq 3$.

Substituting these in the right-hand side of Equation (18), we obtain

$$
\begin{aligned}
&T_{total}(4)=\binom{4}{4}+\binom{4}{1}\times(\binom{3}{1}+\binom{3}{2}+\binom{3}{0})+\\
&\binom{4}{2}\times(\binom{3}{1}\times(\binom{2}{1}+\binom{3}{0})+\binom{3}{2}\times\binom{3}{0})\\
&+\binom{4}{3}\times\binom{3}{1}\times\binom{2}{1}\times\binom{3}{0}
\end{aligned}
\tag{19}
$$

Also note that

$$
\binom{4}{4}=\binom{3}{3}
$$

.

Substituting this in the right-hand side of Equation (19), we have

$$
\begin{aligned}
&T_{total}(4)=\binom{3}{3}+\binom{4}{1}\times(\binom{3}{1}+\binom{3}{2}+\binom{3}{0})\\
&+\binom{4}{2}\times(\binom{3}{1}\times(\binom{2}{1}+\binom{3}{0})+\binom{3}{2}\times\binom{3}{0})\\
&+\binom{4}{3}\times\binom{3}{1}\times\binom{2}{1}\times\binom{3}{0}
\end{aligned}
\tag{20}
$$

$$
\begin{aligned}
&T_{total}(4)=\binom{3}{0}\times\\
&(\binom{4}{1}+\binom{4}{2}\times(\binom{3}{1}+\binom{3}{2})+\binom{4}{3}\times\binom{3}{1}\times\binom{2}{1})\\
&+\binom{3}{1}\times(\binom{4}{1}+\binom{4}{2}\times\binom{2}{1})\\
&+\binom{3}{2}\times\binom{4}{1}+\binom{3}{3}
\end{aligned}
\tag{21}
$$

Now we reorganize the right-hand side of

Equation (20) according to the common factors of $\binom{3}{0}$, $\binom{3}{1}$, $\binom{3}{2}$, $\binom{3}{3}$ with *descending* priorities and we have Formula (21) as listed above.

According to the Formula (14) or the specific forms of Formula (11) and Formula (12), we have

$\binom{4}{1} + \binom{4}{2} \times \binom{2}{1} = 4^2$ and

$\binom{4}{1} + \binom{4}{2} \times (\binom{3}{1} + \binom{3}{2}) + \binom{4}{3} \times \binom{3}{1} \times \binom{2}{1} = 4^3$

Substituting these in the right-hand side of Equation (21), we obtain

$$T_{total}(4) = \binom{3}{0} \times 4^3 + \binom{3}{1} \times 4^2 + \binom{3}{2} \times 4 + \binom{3}{3} \tag{22}$$

By the Binomial Theorem, from Equation (22), we have

$$T_{total}(4) = (4+1)^3 \tag{23}$$

Now we give the proof for the general case of Equation (5).

Proof: From Equation (5), we have

$$\begin{aligned}
&T_{Total}(N) = R(0) + R(1) + ... + R(i) + ... + R(N-1) \\
&= \binom{N}{N} + \binom{N}{1} \times \sum_{k=1}^{N-1} \binom{N-1}{k} + ... \\
&\binom{N}{i} \times \sum_{k_1=1}^{N-i} (\binom{N-1}{k_1} \times ... \times \\
&+ \sum_{k_m=1}^{N-1-\sum_{j=1}^{m-1} k_j-(i-m)} (\binom{N-1-\sum_{j=1}^{m-1} k_j}{k_m} \times ... \times \\
&\sum_{k_i=1}^{N-1-\sum_{j=1}^{i-1} k_j} \binom{N-1-\sum_{j=1}^{i-1} k_j}{k_i}))...) + ... \\
&+ \binom{N}{N-1} \times \binom{N-1}{1} \times \binom{N-2}{1} \times ... \times \binom{2}{1} \times \binom{1}{1}
\end{aligned} \tag{24}$$

Now note that

$\binom{k}{k} = \binom{N-1}{0}$ for all $1 \leq k \leq (N-1)$;

Substituting this in the right-hand side of Equation (24), we obtain

$$\begin{aligned}
&T_{Total}(N) = \binom{N}{N} + \binom{N}{1} \times (\sum_{k=1}^{N-2} \binom{N-1}{k} + \binom{N-1}{0}) + \\
&... + \binom{N}{i} \times \sum_{k_1=1}^{N-i} (\binom{N-1}{k_1} \times ... \times \\
&\sum_{k_m=1}^{N-1-\sum_{j=1}^{m-1} k_j-(i-m)} (\binom{N-1-\sum_{j=1}^{m-1} k_j}{k_m} \times ... \times \\
&\sum_{k_i=1}^{N-1-\sum_{j=1}^{i-1} k_j-1} (\binom{N-1-\sum_{j=1}^{i-1} k_j}{k_i} + \binom{N-1}{0}))...) + ... + \\
&\binom{N}{N-1} \times \binom{N-1}{1} \times \binom{N-2}{1} \times ... \times \binom{2}{1} \times \binom{N-1}{0}
\end{aligned} \tag{25}$$

Also note that

$\binom{N}{N} = \binom{N-1}{N-1}$;

Substituting this in the right-hand side of Equation (25), we have

$$\begin{aligned}
&T_{Total}(N) = \binom{N-1}{N-1} + \binom{N}{1} \times (\sum_{k=1}^{N-2} \binom{N-1}{k} + \binom{N-1}{0}) \\
&+ ... + \binom{N}{i} \times \sum_{k_1=1}^{N-i} (\binom{N-1}{k_1} \times ... \times \\
&\sum_{k_m=1}^{N-1-\sum_{j=1}^{m-1} k_j-(i-m)} (\binom{N-1-\sum_{j=1}^{m-1} k_j}{k_m} \times ... \times \\
&\sum_{k_i=1}^{N-1-\sum_{j=1}^{i-1} k_j-1} (\binom{N-1-\sum_{j=1}^{i-1} k_j}{k_i} + \binom{N-1}{0}))...) + ... + \\
&\binom{N}{N-1} \times \binom{N-1}{1} \times \binom{N-2}{1} \times ... \times \binom{2}{1} \times \binom{N-1}{0}
\end{aligned} \tag{26}$$

Now we reorganize the right-hand side of Equation (26) according to the common factors of

$$\binom{N-1}{0}, \binom{N-1}{1}, \binom{N-1}{2}, ..., \binom{N-1}{N-2}, \binom{N-1}{N-1}$$

with *descending* priorities and we have

$$T_{Total}(N) = \binom{N-1}{0} \times \sum_{i=1}^{N-1} \binom{N}{i} \times$$
$$\left(\sum_{k_1=1, i\geq 2}^{N-1-1} \binom{N-1}{k_1} \times ... \times \left(\sum_{k_{(i-1)}}^{N-1-\sum_{j=1}^{i-1} k_j - 1} \binom{N-1-\sum_{j=1}^{i-2} k_j}{k_{(i-1)}}\right)...\right)$$
$$\binom{N-1}{1} \times \sum_{i=1}^{N-2} \binom{N}{i} \times$$
$$\left(\sum_{k_1=1, i\geq 2}^{N-2-1} \binom{N-2}{k_1} \times ... \times \left(\sum_{k_{(i-1)}=1}^{N-2-\sum_{j=1}^{i-1} k_j - 1} \binom{N-2-\sum_{j=1}^{i-2} k_j}{k_{(i-1)}}\right.\right.$$
$$+ ... + \binom{N-1}{N-1} \tag{27}$$

By Equation (15), we have

$$\sum_{i=1}^{m} \binom{N}{i} \times \left(\sum_{k_1=1, i\geq 2}^{m-1} \binom{m}{k_1} \times ... \times \left(\sum_{k_s=1, 1<s<i-1}^{m-\sum_{j=1}^{s-1} k_j - 1} \binom{m-\sum_{j=1}^{s-1} k_j}{k_s} \times ..\right.\right.$$
$$. \times \left(\sum_{k_{(i-1)}=1}^{m-\sum_{j=1}^{i-2} k_j - 1} \binom{m-\sum_{j=1}^{i-2} k_j}{k_{(i-1)}}\right)...) = N^m$$
$$forall 1 \leq m \leq N-1; \tag{28}$$

Substituting this in the right-hand side of Equation (27), we obtain

$$\begin{aligned} T_{Total}(N) &= \binom{N-1}{0} \times N^{(N-1)} + \binom{N-1}{1} \times N^{(N-2)} \\ &+ ... + \binom{N-1}{k} \times N^{(N-1-k)} + ... + \binom{N-1}{N-1} \end{aligned} \tag{29}$$

Now note the following Binomial Theorem:

$$\begin{aligned} (x+y)^n &= \binom{n}{0} x^n + \binom{n}{1} x^{n-1} y + ... \\ &+ \binom{n}{k} x^{n-k} y^k + ... + \binom{n}{n-1} x y^{n-1} + \binom{n}{n} y^n \end{aligned} \tag{30}$$

By the above Binomial Theorem, from equation (29), we have

$$T_{total}(N) = (N+1)^{(N-1)} \tag{31}$$

So far, we have proved the correctness of Formula (5) against Cayley's theorem and essentially we have provided a new proof of Cayley's formula for counting labeled trees by showing that the sum of the number of all possible trees based on different number of *intermediate* nodes in the trees is equal to the number by Cayley's theorem.

5 NP-hardness of Wireless Broadcast Problems

In the following, we present the proof of the NP-hardness of a group of wireless broadcast problems with respect to certain optimization objectives. For example, given a source node S and destination nodes $D_1, D_2, ..., D_N$, we want to establish a broadcast tree, rooted at S and reaching all of the destinations, to achieve the minimum required energy.

Regarding the transmission energy, we have the following theorems.

Theorem 5.1: The power required for a transmitting node, say T, to *directly* reach a set of destination nodes, say $D_1, D_2, ..., D_m$, is determined by the *maximum* required power to reach any of them individually. As mentioned above, we assume omni-directional antennas are used and the required power for a distance of d between the transmitting node and the receiving node is proportional to d^λ . For the sake of brevity, throughout this paper we will use d^λ to stand for the required power for a transmitting distance of d. Let $d_1, d_2, ..., d_m$ stand for the distances from the transmitting node T to the destinations $D_1, D_2, ..., D_m$ respectively. The required power is determined by:

$$p_{req} = \max(d_1^\lambda, d_2^\lambda, ..., d_m^\lambda) \tag{32}$$

Theorem 5.2: The power required for a broadcast tree is the sum of the energy required for each of the transmitting node in the tree. Let $S, T_1, T_2, ..., T_r$ stand for the transmitting nodes for the given broadcast tree. Notably, S is the source node and $T_1, T_2, ..., T_r$ are the relaying nodes. The required power for the broadcast tree with the transmitting nodes of $S, T_1, T_2, ..., T_r$ is given by:

$$p_{tree} = p_S + \sum_{i=1}^{r} p_{T_r} \tag{33}$$

In Figure 1, we show an extreme case of the broadcast tree, where only one destination node is non-relaying node and all of the other $(N-1)$ intended destination nodes are essentially relaying nodes in the tree.

Theorem 5.3: The minimum-energy broadcast (MEB) tree problem is NP-hard even if the broadcast trees are restricted to have exact $(N-1)$ relaying nodes, assuming we have one source node and N intended destination nodes (Notably, the number of relaying nodes for the broadcast trees can range from zero to $(N-1)$).

Proof: First, we reduce the general minimum-energy broadcast tree problem to the MEB optimization problem just for $(N-1)$ relaying nodes cases (see Figure 1 for an example) and then we transform the restricted MEB optimization problem to the well-known NP-complete problem: Traveling Salesman Extension (TSE) [7]. The TSE problem assumes that the inputs are: a finite set $C = \{c_1, c_2, ...c_m\}$ of cities,

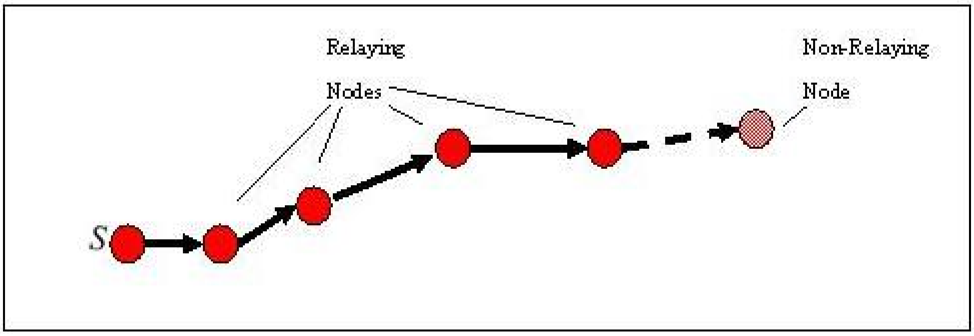

Fig. 1 An extreme case with only one non-relaying destination node

a distance $d(c_i, c_j) \in Z^+$ for each pair of cities $c_i, c_j \in C$, a bound $B \in Z^+$, and a particular tour $\Theta =< c_{\Pi(1)}, c_{\Pi(2)}, ..., c_{\Pi(k)} >$ of k distinct cities from C, $1 \leq k \leq m$. The problem is if Θ can be extended to a full tour $< c_{\Pi(1)}, c_{\Pi(2)}, ..., c_{\Pi(k)}, c_{\Pi(k+1)}, ..., c_{\Pi(m)} >$ [7]. We rephrase the restricted MEB optimization problem for $(N-1)$ relaying nodes case (suppose we have one source node S and N destination nodes) as a TSE problem by the following transformations: each node in MEB optimization problem can be seen as a "city", the energy cost between any two nodes (if one transmits to another), say $d^{\lambda}_{i,j}$ for node i and node j , can be viewed as the distance between each pair of "cities", a particular tour $\Theta =< c_{\Pi(1)} >$ and $c_{\Pi(1)} = S$. Our goal is to find a minimal energy tour from source "city" S to cover all the other N "cities" with exact $(N-1)$ relaying "cities" (see Figure 1 for the illustration). Thus, the problem can be transformed as if we can extend $\Theta =< c_{\Pi(1)} >$ to a full tour$< c_{\Pi(1)}, c_{\Pi(2)}, ..., c_{\Pi(N+1)} >$ such that the total "length", i.e., the total energy cost, is B or less. To here, since the TSE problem is NP-complete, as a consequence, the restricted MEB optimization problem is at least as hard as TSE problem. Further, since the restricted MEB problem is a sub-problem of the general MEB problem, the TSE problem can also be viewed as a sub-problem of the general MEB problem. Obviously, the general MEB problem is NP-hard.

Essentially, the proof of the NP-hardness of the MEB problem can be extended to other wireless broadcast problems with respect to the optimization of other objectives.

6 Conclusion

In the fields of information theory and computing science, trees are arguably one of the most important data structures to depict and represent choice-making strategies such as the one used in Google's AlphaGo system that defeated human champions in recent human-machine grand challenge. Another example is that different number of intermediate nodes in wireless broadcast trees may have great impact on energy consumption of each node that are typically equipped with limited power supplies in a wireless sensor network, which may eventually determine how long the given wireless sensor network can last. Thus, it is of great importance to have a deep understanding of the mathematical nature of wireless broadcast trees. In this paper we present a new theorem for counting labeled trees based on the number of *intermediate* nodes, i.e., the non-leaf nodes except the root node, in the tree and we prove its correctness against Cayley's famous theorem for counting labeled trees. Essentially we provide a new proof of Cayley's formula for counting labeled trees, during the procedure of which we introduce an interesting combinatorial structure of n^m series.

To the best of our knowledge, it is the first time that a proof purely based upon combinatorial structures without constructing a bijection between two kinds of trees has been presented and it is the first time that a new tree enumeration theorem based on the number of *intermediate* nodes, which are also called *relaying* nodes in a wireless broadcast tree, is presented and it is also the first time that an interesting combinatorial structure of n^m series has been explored and proved. Towards the end of the paper, we give a formal proof of the NP-hardness of wireless broadcast problems with respect to the optimization of certain objectives.

Acknowledgment

The authors would like to thank Prof. David M. Jackson at University of Waterloo for his invaluable help on the derivation of Eq. (10).

This work was supported in part by National Natural

Science Foundation of China (grant No. 61472200) and Beijing Municipal Science & Technology Commission (grant No. Z161100000416004).

References

[1] P. J. Cameron, *Combinatorics: Topics, Techniques, Algorithms*. England, Cambs: Cambridge Univ. Press, 1^{st} edition, 1994.

[2] A. Cayley, *A Theorem on Trees*. Quart. J. Math. vol. 23. 1889, pp. 376-378.

[3] I. P. Goulden, D. M. Jackson, *Combinatorial Enumeration*. John Wiley & Sons Press, 1983, pp. 191.

[4] D. E. Knuth, *The Art of Computer Programming: Fundamental Algorithms*. (vol 1, 2nd edition), 1968, pp. 362-363.

[5] F. Li, I. Nikolaidis, On minimum-energy broadcasting in all-wireless networks, In *the proceeding of the 26th Annual IEEE Conference on Local Computer Networks (IEEE LCN '2001)*, Tampa, Florida, USA, Nov. 2001, pp. 1-12.

[6] F. Li, K. Wu, Reliable, distributed and energy-efficient broadcasting in multi-hop mobile Ad Hoc networks, presented at the proceeding of IEEE International Conference on Local Computer Networks and Wireless Local Networks (IEEE WLN '2002), Tampa, Florida, USA, Nov. 2002, pp.0761.

[7] M. Garey, D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman and Company, 1979.

[8] J. W. Moon, Various proofs of Cayley's formula for counting trees, *A Seminar on Graph Theory (F. Harary, Ed.)*. pp. 70-78, 1967.

[9] P. Shor, A new proof of Cayley's formula for counting labeled trees, *Journal of Combinatorial Theory*. Ser. vol.71, no.1, pp. 154-158, 1995.

[10] R. P. Stanley, *Enumerative Combinatorics*. England, Cambs: Cambridge Univ. Press,1999, pp. 16-17, Vol. I and pp. 28-35, Vol. II.

Junwei Cao the corresponding author, achieved his bachelor's degree at Tsinghua University in 1996, master's degree in 1998 and PhD in University of Warwick in 2001. From 2002 to 2006, he worked in the NEC European Laboratory in Germany and the MIT/LIGO laboratory of the United States.Cao has been working at the Tsinghua University since 2006 as Professor and Vice Dean of the Research Institute of Information Technology. He has published more than 200 academic papers and 8 books.

Fulu Li is the Dean of Research Institute of Emerging Technology at Taikang Insurance Group and an an Adjunct Researcher with Research Institute of Information Technology, Tsinghua University, Beijing, P. R. China, he got his PhD from MIT in the USA, Email: lifl26@taikanglife.com.

Chunfeng Wang is a Senior Researcher at Qian Xuesen Laboratory of Space Technology, China Academy of Space Technology, Beijing, P. R. China, he got his PhD from Xian Jiaotong University, Email wangchunfeng@qxslab.cn.

Kui Wu is a Professor at University of Victoria, Canada, he got his PhD from University of Alberta, Canada, Email: wkui@uvic.ca.