Template for Preparation of Manuscripts for
Tsinghua Science and Technology

This template is to be used for preparing manuscripts for submission to Tsinghua Science and Technology. Use of this template will save time in the review and production processes and will expedite publication. However, use of the template is not a requirement of submission. Do not modify the template in any way (delete spaces, modify font size/line height, etc.).
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 widely used in information theory and computing science. Different numbers of intermediate nodes in wireless broadcast trees may exert great impacts on the energy consumption of individual nodes, which are typically equipped with a limited power supply in a wireless sensor network; this limitation may eventually determine how long the given wireless sensor network can last. Thus, obtaining a deep understanding of the mathematical nature of wireless broadcast trees is of great importance. In this paper, we give new proof of Cayley’s well-known theorem for counting labeled trees. A distinct feature of this proof is that we purely use combinatorial structures instead of constructing a bijection between two kinds of labeled trees, which is in contrast to all existing proofs. 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 the best of our knowledge, this work is the first to present a tree enumeration theorem based on the number of intermediate nodes in the tree.

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

1 Introduction

Trees are arguably one of the most important data structures widely used in information theory and computing science. From $B^+$ trees, which are used for indexing and querying in nearly every major database system, to decision and strategy trees, which are used in probabilistic inferences such as that used in Google’s AlphaGo system that defeated human champions in a recent human-machine grand challenge, trees are the most-often-sought utilities to depict and represent a variety of choice-making strategies. Different numbers of intermediate nodes, i.e., non-leaf nodes except the root node, in wireless broadcast trees may exert great impacts on the energy consumption of individual nodes, which are typically equipped with a limited power supply in a wireless sensor network; this limitation may eventually determine how long the given wireless sensor network can last. Thus, obtaining a deep understanding of the mathematical nature of wireless broadcast trees is of great importance.

Studies of trees have captured the interest of some of the most prominent minds in history[1,2,3]. In 1889, British mathematician Arthur Cayley, who helped found the modern British school of pure mathematics, published a well-known theorem on trees. Donald Knuth, a Turing award winner and a professor at Stanford University, and Peter Shor, a winner of the Nevanlinna Prize, a professor at MIT, and the inventor of Shor’s algorithm, which is widely regarded as a major breakthrough for quantum algorithm for integer factorization, all explored Cayley’s theorem on trees to obtain a deeper understanding of the concept[3,4].

The rest of this 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 then examine the $n^m$ series and its combinatorial structures in Section 3, paving the way
for our new proof of Cayley’s theorem in Section 4. We further give formal proof of the NP-hardness of wireless broadcast problems toward the optimization of certain objectives in Section 5, and, finally, present conclusions in Section 6.

2 The Free Tree Theorem Revisited

According to Cayley’s theorem, \( n^{n-2} \) distinctly labeled free trees are present on \( n \) vertices [5]. If \( X \) is a particular vertex, the free trees are in one-to-one correspondence with oriented trees having root \( X \). Therefore, for the case with one root node \( S \), which is also called the source node in wireless broadcast trees [6,7], and \( N \) non-root nodes, also called destination nodes in wireless broadcast trees, we have \((N + 1)^{(N-1)}\) distinct labeled free trees.

In this section, we take another look at Cayley’s free tree theorem based on the number of intermediate nodes in the tree, i.e., 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, in the construction of energy-efficient broadcast trees in multi-hop wireless ad hoc networks [6,7]. 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:

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

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

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

\[
R(1) = \frac{N}{1} \times \left( \binom{N-1}{1} + \binom{N-2}{2} + \binom{N-3}{3} + \ldots + \binom{N-1}{1} + \binom{N-2}{2} \right) \tag{2}
\]

The above equation indicates that one among \( N \) non-root nodes must first be selected as the intermediate node. We then decide which nodes are direct child nodes of the intermediate node (the remaining nodes are direct child nodes of the root node \( S \)). Here, the only intermediate node must be the direct child node of the root node, so 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:

\[
R(2) = \binom{N}{2} \times \left( \sum_{i=1}^{N-2} \binom{N-1}{i} \times \sum_{j=1}^{N-1-i} \binom{N-1-i}{j} \right) \tag{3}
\]

Equation (3) roughly states that we first need to decide how many choices exist to pick up the two intermediate nodes among \( N \) non-root nodes. We then 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:

\[
R(i) = \binom{N}{i} \times \sum_{k_1=1}^{N-i} \left( \binom{N-1-k_1}{k_1} \right) \times \sum_{k_2=1}^{N-1-k_1} \left( \binom{N-1-k_1-k_2}{k_2} \right) \times \ldots \times \sum_{k_i=1}^{N-1-\sum_{j=1}^{i-1} k_j} \left( \binom{N-1-\sum_{j=1}^{i-1} k_j}{k_i} \right) \ldots \tag{4}
\]

Simply put, Equation (4) states the same straightforward philosophy demonstrated in Equation (3). Essentially, the maximum number of direct child nodes of the \( m \)th picked relaying node, i.e., \((N - 1 - \sum_{j=1}^{m-1} k_j - (i - m))\) (where \( \sum_{j=1}^{m-1} k_j \) is the sum of the direct child nodes of the first \( m - 1 \) relaying nodes), guarantees the basic principle for 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 of child nodes can be easily calculated by a computer program because of its special structures.

Among 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:

\[
T_{\text{total}}(N) = \sum_{i=0}^{N-1} R(i) = R(0) + \sum_{i=1}^{N-1} \left( \binom{N}{i} \right) \times \left( \binom{N-1}{k_1} \right) \times \left( \prod_{k=2}^{N-1} \binom{N-1-k_{k-1}}{k_k} \right) \times \left( \binom{N-1-k_{N-1}}{k_{N-1}} \right) \times \left( \prod_{k=1}^{N-1-k_{N-1}} \binom{N-1-k_{k-1}}{k_k} \right) 
\]

(5)

The above analysis provides a novel view on 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 [6,7].

The following table (Table 1) illustrates the relationship between the total number of trees and the number of trees with different numbers of relaying nodes, as dictated in Equation (5). \(T(i, N)\) denotes the number of trees with \(i\) intermediate nodes and \(N\) non-root nodes.

While we will use Equation (4) and (5) in the final proof of Cayley’s theorem in Section 3, we also give a simplified form of the former using the results of labeled rooted trees with degree sequence given by Goulden and Jackson in [8].

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

**Theorem 1.1:** \(R(i)\) denotes the number of labeled trees with \(i\) intermediate nodes in the tree; thus, we have

\[
R(i) = \binom{N}{i} \sum_{k=1}^{i+1} \left( \binom{i}{k-1} \right) (-1)^{i+1-k} k^{(N-1)}
\]

**Proof:** Following [8], we first define the degree of a vertex \(v\) to be the number of edges incident to vertex \(v\) and a sequence \(r = (r_1, r_2, \ldots)\) 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

\[
\sum r_i = N + 1 \quad \sum i \times r_i = 2N \quad \text{(6)}
\]

To ensure self-containment, let us recall the results on labeled rooted trees with degree sequence in [8].

In [8], the number of labeled rooted trees with degree sequence \((r_1, r_2, \ldots)\) (i.e., \(r_j\) vertices are of degree \(j\), for \(j \geq 1\)), is

\[
\frac{(N + 1)(N - 1)!}{\prod_{j \geq 1} r_j !(j - 1)! r_j} \quad \text{for} N \geq 1, \quad \text{(7)}
\]

where \(r_1 + r_2 + \ldots = N + 1, r_1 + 2r_2 + 3r_3 + \ldots = 2N\).

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

\[
\frac{(N - 1)!}{\prod_{j \geq 1} r_j !(j - 1)! r_j} \quad \text{for} N \geq 1, \quad \text{(8)}
\]

where \(r_1 + r_2 + \ldots = N + 1, r_1 + 2r_2 + 3r_3 + \ldots = 2N\).

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

\[
R(i) = \sum_{j \geq 1} \frac{(N - 1)!}{\prod_{j \geq 1} r_j !(j - 1)! r_j} \quad \text{for} N \geq 1, \quad \text{(9)}
\]

where the sum is the over-all degree sequence \((r_1, r_2, \ldots)\) such that \(r_j\) is a non-negative integer for each \(j\), \(r_1 = N - i\) 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, \(r_1 + r_2 + \ldots = N + 1, r_1 + 2r_2 + 3r_3 + \ldots = 2N\).

Further, by virtue of the Lagrange Theorem [8], we have

\[
R(i) = \binom{N}{i} \sum_{k=1}^{i+1} \left( \binom{i}{k-1} \right) (-1)^{i+1-k} k^{(N-1)} \quad \text{(10)}
\]

Equation (5) and (10) show different aspects of the same counting problem.

**3 The \(n^m\) Series and its Combinatorial Structures**

To give proof of Equation (5) against Cayley’s theorem, let us look at the interesting combinatorial structures of the \(n^m\) series, which will be used in the proof in Section 3. In the following, we give the specific combinatorial structures of \(n^2, n^3,\) and \(n^4,\) the combination of which covers all new combinatorial structures we will use as a basis to derive the generalized form for \(n^m\).
\[
\binom{n}{1} + \binom{n}{2} \times \binom{2}{1} = n^2 \tag{11}
\]
\[
\binom{n}{1} + \binom{n}{2} \times \left( \binom{3}{1} + \binom{3}{2} \right) + \binom{n}{3} \times \binom{3}{1} \times \binom{2}{1} = n^3 \tag{12}
\]
\[
\binom{n}{1} + \binom{n}{2} \times \left( \binom{4}{1} + \binom{4}{2} + \binom{4}{3} \right) \times \binom{n}{3} \times \left( \binom{3}{1} + \binom{3}{2} \right) + \binom{n}{4} \times \binom{4}{1} \times \binom{3}{1} \times \binom{2}{1} = n^4 \tag{13}
\]

Proofs of Equations (11), (12) and (13) are provided below. Based on these equations, we will derive a generalized form for the similar combinatorial structures of \( n^m \).

\textbf{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. The possible number of ways of picking two balls at random among these \( n \) distinct balls according to the rules 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 of the two balls as the first-picked ball and designate the other as the second-picked ball.

Following the same rule as above, 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. The possible number of ways of picking three balls among \( n \) distinct balls according to the rules include the following: (a) picking the same ball three times, (b) picking two different balls with one ball picked twice (two balls are the same among the three picked balls), or (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., \( \binom{n}{2} \times \left( \binom{3}{1} + \binom{3}{2} \right) \), 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 have 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 ball takes the remaining slots and it appears twice) or (b) it appears in two of the three slots, i.e., \( \binom{3}{2} \), in which case it appears twice (the other ball 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 different from each other, which means

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
\( \frac{i}{N} \) & 1 & 2 & 3 & 4 \\
\hline
0 & \binom{1}{1} = 1 & \binom{2}{2} = 1 & \binom{3}{3} = 1 & \binom{4}{4} = 1 \\
\hline
1 & \binom{2}{1} = 2 & \binom{3}{1} \times \binom{2}{2} = 9 & \binom{4}{1} \times \left( \binom{3}{1} + \binom{3}{2} \right) = 28 \\
\hline
2 & \binom{3}{2} \times \binom{2}{1} = 6 & \binom{4}{2} \times \left( \binom{3}{1} \times \binom{2}{2} \right) = 72 \\
\hline
3 & \binom{4}{3} \times \binom{3}{1} \times \binom{2}{1} \times \binom{1}{1} = 24 \\
\hline
\hline
\( \sum T(i, N) \) & \( (1 + 1)^{(1-1)} = 1 \) & \( (2 + 1)^{(2-1)} = 1 + 2 = 3 \) & \( (3 + 1)^{(3-1)} = 1 + 9 + 6 = 16 \) & \( (4 + 1)^{(4-1)} = 1 + 28 + 72 + 24 = 125 \) \\
\hline
\end{tabular}
\caption{Examples of \( T(i, N) \) according to Equation (5)}
\end{table}
that we first pick three balls from \( n \) distinct balls at one time, pick one from the three balls as the first-picked ball, pick one more from the remaining two balls as the second-picked ball, and then designate the last ball as the third-picked ball.

Similarly, 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. The possible number of ways of picking four balls among \( n \) distinct balls according to the rules include: (a) picking the same ball 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), or (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., \( \binom{n}{2} \times \left( \binom{4}{1} + \binom{4}{2} + \binom{4}{3} \right) \), is the number of ways that there are two exactly different types of balls among the four picked balls, which means we first pick two balls from \( n \) distinct balls at one time and then have 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) put it 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), or (c) put it in three of the four slots, i.e., \( \binom{4}{3} \), in which case it appears 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., \( \binom{3}{3} \times \left( \binom{4}{1} \times \left( \binom{4}{1} + \binom{4}{2} \right) + \binom{4}{2} \times \binom{4}{3} \right) \), is the number of ways that there are exactly three different types of balls among the four picked balls, which means we first pick three balls from \( n \) distinct balls at one time and then have 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., \( \binom{4}{1} \). In this case, 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., \( \binom{3}{1} \) or (b) it appears twice among the remaining three slots (the last ball takes the remaining slot and it appears only once), i.e., \( \binom{3}{2} \). Alternatively, (2) the ball appears twice in the four slots, in which case we first pick two slots for this ball, i.e., \( \binom{4}{2} \) and then 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., \( \binom{4}{3} \times \binom{1}{1} \times \binom{1}{1} \times \binom{1}{1} \), is the number of ways that the four picked balls are different from each other, which means we first pick four balls from \( n \) distinct balls at one time, pick one from the four balls as the first-picked ball, pick one more from the remaining three balls as the second-picked ball, pick one more from the remaining two balls as the third-picked ball, and then designate the last ball 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>
<thead>
<tr>
<th>( n^m )</th>
<th>( \sum \prod a_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2^1</td>
<td>1+2+1 = 4</td>
</tr>
<tr>
<td>2^2</td>
<td>1+3+3+1 = 8</td>
</tr>
<tr>
<td>2^3</td>
<td>1+4+6+4+1 = 16</td>
</tr>
<tr>
<td>3^1</td>
<td>1+1+1+2+2+2 = 9</td>
</tr>
<tr>
<td>3^2</td>
<td>1+3+3+3+3+3+3+6+1+1 = 27</td>
</tr>
<tr>
<td>3^3</td>
<td>1+1+1+4+4+4+4+4+6+6+6+12+12+12+12 = 81</td>
</tr>
<tr>
<td>4^2</td>
<td>1+1+1+1+2+2+2+2+2+2 = 16</td>
</tr>
<tr>
<td>4^3</td>
<td>1+1+1+3+3+3+3+3+3+3+3+3+6+6+6+6 = 64</td>
</tr>
</tbody>
</table>

Table 2 Some decomposition examples of \( n^m \) based on Equation (17).
The number of ways of picking \( m \) balls is the sum of the number of ways of picking exactly \( n \) balls, simplified form of the combinatorial structures of Cayley’s formula in Section 3, we give another

\[
\binom{n}{1} + \binom{n}{2} \times \left( \binom{m}{1} + \binom{m}{2} + \cdots + \binom{m}{m-1} \right) + \binom{n}{3} \times \left( \binom{m-1}{1} + \binom{m-1}{2} + \cdots + \binom{m-1}{m-1} \right) + \cdots + \binom{m-1}{1} \times \left( \binom{m-2}{1} + \binom{m-2}{2} + \cdots + \binom{m-2}{m-2} \right) + \cdots + \binom{n}{m} \times \binom{m}{1} \times \binom{m-1}{1} \times \cdots \times \binom{2}{1} \right) \tag{14}
\]

\[
= n^m.
\]

In a more compact form, we have:

\[
\sum_{k_1=1}^{n} \binom{n}{k_1} \times \prod_{k_1=1}^{m-2} \binom{m-k_1}{k_1} \cdots \times \binom{m-3}{k_1} \times 2^{m-1} = n^m. \tag{15}
\]

By analogy, 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 Equation (14) in the final proof of Cayley’s formula in Section 3, we give another simplified form of the combinatorial structures of \( n^m \) by looking at this equation from a different angle.

Following [3], we first define \( \binom{m}{a_1, a_2, \ldots, a_n} \) as the coefficient of \( x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} \) in \( (x_1 + x_2 + \cdots + x_n)^m \) and we have

\[
\binom{m}{a_1, a_2, \ldots, a_n} = \frac{m!}{a_1! a_2! \cdots a_n!} = \frac{m!}{\prod_{i=1}^{n} a_i!} \tag{16}
\]

Essentially, \( n^m \) is the sum of the multi-nomial coefficients of \( (x_1 + x_2 + \cdots + x_n)^m \) where \( x_1 = x_2 = \cdots = x_n = 1 \).

Therefore, we have

\[
n^m = \sum_{a_1 + a_2 + \cdots + a_n = m} \frac{m!}{\prod_{i=1}^{n} a_i!} \tag{17}
\]

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

The following table (Table 2) illustrates some examples of \( n^m \) based on Equation (17).

### 4 Proof of Cayley’s Theorem

In this section, we prove that Equation (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 Equation (5), we have

\[
T_{total}(4) = R(0) + R(1) + R(2) + R(3) = \left( \frac{4}{4} \right) + \left( \frac{4}{4} \times \left( \frac{3}{1} + \frac{3}{2} + \frac{3}{3} \right) \right) + \left( \frac{4}{2} \times \left( \frac{3}{1} \times \left( \frac{2}{1} + \frac{2}{2} \right) + \frac{3}{2} \times \frac{1}{1} \right) \right) + \left( \frac{4}{3} \times \frac{3}{1} \times \frac{2}{1} \times \frac{3}{0} \right) \tag{18}
\]

Note \( \left( \frac{k}{k} \right) = \left( \frac{3}{0} \right) \) for all \( 1 \leq k \leq 3 \).

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

\[
T_{total}(4) = \left( \frac{4}{4} \right) + \left( \frac{4}{4} \times \left( \frac{3}{1} + \frac{3}{2} + \frac{3}{3} \right) \right) + \left( \frac{4}{2} \times \left( \frac{3}{1} \times \left( \frac{2}{1} + \frac{2}{2} \right) + \frac{3}{2} \times \frac{1}{1} \right) \right) + \left( \frac{4}{3} \times \frac{3}{1} \times \frac{2}{1} \times \frac{3}{0} \right) \tag{19}
\]

Also note that

\[
\left( \frac{4}{4} \right) = \left( \frac{3}{3} \right).
\]

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

\[
T_{total}(4) = \left( \frac{4}{4} \right) + \left( \frac{4}{4} \times \left( \frac{3}{1} + \frac{3}{2} + \frac{3}{3} \right) \right) + \left( \frac{4}{2} \times \left( \frac{3}{1} \times \left( \frac{2}{1} + \frac{2}{2} \right) + \frac{3}{2} \times \frac{1}{1} \right) \right) + \left( \frac{4}{3} \times \frac{3}{1} \times \frac{2}{1} \times \frac{3}{0} \right) \tag{20}
\]

We reorganize the right-hand side of Equation (20) as

\[
(\frac{4}{1} + \frac{4}{2} \times \frac{3}{1} + \frac{3}{2} \times \frac{3}{1} \times \frac{2}{1}) + \frac{3}{2} \times \frac{4}{1} + \frac{3}{1} \tag{21}
\]
\[
\left(\frac{3}{0}\right), \left(\frac{3}{1}\right), \left(\frac{3}{2}\right), \left(\frac{3}{3}\right) \text{ with descending priority, and we have Equation (21) as listed above.}
\]

According to the Equation (14) or the specific forms of Equations (11) and (12), we have
\[
\left(\frac{1}{1}\right) + \left(\frac{4}{2}\right) \times \left(\frac{2}{1}\right) = 4^2
\]
and
\[
\left(\frac{1}{1}\right) + \left(\frac{4}{2}\right) \times \left(\frac{3}{2} + \frac{3}{3}\right) + \left(\frac{4}{3}\right) \times \left(\frac{3}{1}\right) \times \left(\frac{2}{1}\right) = 4^3
\]
Substituting these in the right-hand side of Equation (21), we obtain
\[
T_{\text{total}}(4) = \frac{3}{0} \times 4^3 + \frac{3}{1} \times 4^2 + \frac{3}{2} \times 4 + \frac{3}{3}
\]
(22)

By the Binomial Theorem, from Equation (22), we have
\[
T_{\text{total}}(4) = (4 + 1)^3
\]
(23)

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

**Proof:** From Equation (5), we have
\[
T_{\text{Total}}(N) = R(0) + R(1) + ... + R(i) + ... + R(N-1)
\]
\[
= \left(\frac{N}{N}\right) + \left(\frac{1}{1}\right) \times \sum_{k=1}^{N-1} \left(\frac{N-1}{k}\right) + ...
\]
\[
\left(\frac{N}{1}\right) \times \sum_{k=1}^{N-1} \left(\frac{N-1}{k}\right) \times ... \times \left(\frac{N-1}{k_m}\right) \times ...
\]
\[
N-1 \sum_{j=1}^{N-1} \left(\frac{N-1}{j}\right) \times ...
\]
\[
+ \left(\frac{N}{N-1}\right) \times \left(\frac{N-1}{1}\right) \times \left(\frac{N-2}{1}\right) \times ... \times \left(\frac{2}{1}\right) \times \left(\frac{1}{1}\right)
\]
(24)

Note that \(\binom{N}{k} = \binom{N-1}{0}\) for all \(1 \leq k \leq (N-1)\);
Substituting this in the right-hand side of Equation (24), we obtain
\[
T_{\text{Total}}(N) = \left(\frac{N}{N}\right) + \left(\frac{1}{1}\right) \times \sum_{k=1}^{N-1} \left(\frac{N-1}{k}\right) + \left(\frac{N-1}{0}\right) + ...
\]
\[
\left(\frac{N}{i}\right) \times \sum_{k=1}^{N-1} \left(\frac{N-1}{k}\right) \times ...
\]
\[
N-1 \sum_{j=1}^{N-1} \left(\frac{N-1}{j}\right) \times ...
\]
\[
+ \left(\frac{N}{N-1}\right) \times \left(\frac{N-1}{1}\right) \times \left(\frac{N-2}{1}\right) \times ... \times \left(\frac{2}{1}\right) \times \left(\frac{1}{1}\right)
\]
(25)

Also note that \(\binom{N}{N} = \binom{N-1}{N} = \binom{N-1}{N-1}\);
Substituting this in the right-hand side of Equation (25), we have
\[
T_{\text{Total}}(N) = \left(\frac{N}{N}\right) + \left(\frac{1}{1}\right) \times \sum_{k=1}^{N-1} \left(\frac{N-1}{k}\right) + \left(\frac{N-1}{0}\right) + ...
\]
\[
+ \left(\frac{N}{i}\right) \times \sum_{k=1}^{N-1} \left(\frac{N-1}{k}\right) \times ...
\]
\[
N-1 \sum_{j=1}^{N-1} \left(\frac{N-1}{j}\right) \times ...
\]
\[
+ \left(\frac{N}{N-1}\right) \times \left(\frac{N-1}{1}\right) \times \left(\frac{N-2}{1}\right) \times ... \times \left(\frac{2}{1}\right) \times \left(\frac{1}{1}\right)
\]
(26)

We reorganize the right-hand side of Equation (26) according to the common factors of
\[
\left(\frac{N-1}{0}\right), \left(\frac{N-1}{1}\right), \left(\frac{N-1}{2}\right), ..., \left(\frac{N-1}{N-1}\right)
\]
with descending priority, and we have
\[
T_{\text{Total}}(N) = \left(\frac{N-1}{0}\right) \times \sum_{i=1}^{N-1} \left(\frac{N}{i}\right) \times
\]
\[
\left(\frac{N-1}{1}\right) \times \left(\frac{N-1}{i}\right) \times \left(\frac{N-2}{1}\right) \times ... \times \left(\frac{N-1}{k_i}\right) \times ...
\]
\[
N-1 \sum_{j=1}^{N-1} \left(\frac{N-1}{j}\right) \times ...
\]
\[
+ \left(\frac{N}{N-1}\right) \times \left(\frac{N-1}{1}\right) \times \left(\frac{N-2}{1}\right) \times ... \times \left(\frac{2}{1}\right) \times \left(\frac{1}{1}\right)
\]
(27)

By Equation (15), we have
\[
\sum_{i=1}^{N} \binom{N}{i} \times \left(\frac{m-1}{k_1}\right) \times ... \times \left(\frac{m-1}{k_{i-1}}\right) \times \left(\frac{N-1}{0}\right)
\]
\[
\times \left(\frac{N-1}{1}\right) \times \left(\frac{N-2}{1}\right) \times ... \times \left(\frac{2}{1}\right) \times \left(\frac{1}{1}\right)
\]
(28)

Substituting this in the right-hand side of Equation (27),
we obtain
\[ T_{total}(N) = \sum_{k=0}^{N-1} \binom{N-1}{k} N^{N-1-k} = \sum_{k=0}^{N-1} \binom{N-1}{k} N^{N-1} \] (29)
Note the following Binomial Theorem:
\[ (x + y)^n = \binom{n}{0} x^n + \binom{n}{1} x^{n-1} y + ... + \binom{n}{n} y^n \] (30)
By this theorem, from Equation (29), we have
\[ T_{total}(N) = (N + 1)^{N-1} \] (31)

Thus far, we have proven the correctness of Equation (5) against Cayley’s theorem and essentially provided a new proof of this theorem for counting labeled trees by showing that the sum of the number of all possible trees based on different numbers of intermediate nodes in the trees is equal to the number determined by Cayley’s theorem.

5 NP-hardness of Wireless Broadcast Problems

In this section, we present 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 that omni-directional antennas are used and that the required power for a distance of \( d \) between the transmitting node and the receiving node is proportional to \( d^\lambda \). For the sake of brevity, throughout this paper we will use \( d^\lambda \) to represent 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:
\[ P_{req} = \max(d_1^\lambda, d_2^\lambda, ..., d_m^\lambda) \] (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 \( S, T_1, T_2, ..., T_r \) represent the transmitting nodes for the given broadcast tree. Here, \( S \) 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:
\[ P_{tree} = P_S + \sum_{i=1}^{r} P_{T_i} \] (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 (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 the \((N-1)\) relaying nodes case (see Figure 1 for an example) and then transform the restricted MEB optimization problem to the well-known NP-complete problem Traveling Salesman Extension (TSE) [9]. The TSE problem assumes that the inputs are: a finite set \( C = \{c_1, c_2, ..., c_m\} \) of cities, 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 \( \Theta = \{\Pi(1), \Pi(2), ..., \Pi(k)\} \) of \( k \) distinct cities from \( C, 1 \leq k \leq m \). The problem is if \( \Theta \) can be extended to a full tour \( \{\Pi(1), \Pi(2), ..., \Pi(k), \Pi(k+1), ..., \Pi(m)\} \) [9].

We rephrase the restricted MEB optimization problem for the \((N-1)\) 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 considered a “city,” the energy cost between any two nodes (if one transmits to another), say \( d_{ij}^\lambda \) for node \( i \) and node \( j \), can be considered as the distance between each pair of “cities,” a particular tour \( \Theta = \{\Pi(1), \Pi(2), ..., \Pi(k)\} \) and \( \Pi(1) = S \). Our goal is to find a minimal energy tour from source “city” \( S \) to cover all 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 \( \Theta = \{\Pi(1)\} \) to a full tour \( \{\Pi(1), \Pi(2), ..., \Pi(N-1)\} \) such that the total “length,” i.e., the total energy cost, is \( B \) or less. Here, since the TSE problem is NP-complete, as a consequence, the restricted MEB optimization problem is at least as hard as the TSE problem. Furthermore, 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 that used in Google’s AlphaGo system that defeated human champions in a recent human-machine grand challenge. Different numbers of intermediate nodes in wireless broadcast trees may exert great impacts on the energy consumption of individual nodes, which are typically equipped with a limited power supply in a wireless sensor network; this limitation may eventually determine how long the given wireless sensor network can last. Thus, obtaining a deep understanding of the mathematical nature of wireless broadcast trees is of great importance. 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 prove its correctness against Cayley’s theorem for counting labeled trees. Essentially, we provide a new proof of Cayley’s theorem for counting labeled trees, during which we also introduce an interesting combinatorial structure of $n^m$ series.

To the best of our knowledge, this work is the first to propose a proof purely based on combinatorial structures without constructing a bijection between two kinds of trees. It is also the first to present a new tree enumeration theorem based on the number of intermediate nodes, which are also called relaying nodes in a wireless broadcast tree, and the first to explore and demonstrate an interesting combinatorial structure of $n^m$ series. Toward the end of the paper, we give 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

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.

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

Fulu Li is the Dean of Research Institute of Emerging Technology at Taikang Insurance Group and 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.