

# A Modified Measure of Covert Network Performance

LYNNE L. DOTY

Marist College

Department of Mathematics

Poughkeepsie, NY

UNITED STATES

lynne.doty@marist.edu

*Abstract:* In a covert network the need for secrecy is at odds with the desire for easy transmission of information. Lindelauf *et al.* [7] have defined several measures that can be used to evaluate the total performance of a covert social network by using a product of individual measures of secrecy and of information transmission. As one of their simplest measures of secrecy, Lindelauf *et al.* use a modified version of the idea of neighbor connectivity. Using this measure, the optimal network structure is a star graph. In this paper we modify the Lindelauf measure of secrecy to include information about the connectedness (not just the order) of the survival subgraph using a measure based on ideas related to Chvátal's toughness parameter. We determine an upper bound on performance of trees and conjecture that a specific class of spider graphs achieves maximum performance. We also describe several opportunities for further research.

*Key-Words:* covert networks, neighbor connectivity, spider graphs, toughness

## 1 Introduction

In a covert network the need for secrecy is at odds with the desire for easy transmission of information. A hallmark of any covert network is the idea that the discovery of one person (one vertex) in the network renders that person and all members of the network known to him useless to the network. This means the closed neighborhood of the vertex is removed from the graph producing what has been called the survival subgraph. Gunther, Hartnell, and Novakowski studied survival subgraphs from the standpoint of their ability to remain connected using neighbor connectivity. Lindelauf *et al.* [7], [8] have defined several measures that can be used to evaluate the total performance of a covert social network by using a product of individual measures of secrecy and of information transmission. They use a standard social network notion (total distance) to measure ease of information transmission. As one of their simplest measures of secrecy, Lindelauf *et al.* use a modified version of the idea of neighbor connectivity. To measure a network's secrecy they calculate for each vertex the ratio of the order of the survival subgraph when this vertex is discovered (subverted) to the order of the original graph, and then sum these ratios over all vertices of the graph. Using this measure, the optimal network structure is a star graph. The discovery of any vertex in the star, however, results in a graph that is totally disconnected or empty. The purpose of this paper is to modify the Lindelauf

measure of secrecy to include information about the connectedness (not just the order) of the survival subgraph. Using this modified measure, cycles perform better than paths or stars, as seems more intuitively appealing. We determine an upper bound for this new performance measure for trees. We conjecture that the optimal tree is a double star—a special type of spider graph—and we give some evidence and analysis to support the conjecture. Hartnell and Gunther [6] have identified these double star graphs as best possible using different covert network measures thereby providing some validation for the usefulness of the new measure introduced in this paper.

## 2 Definitions and Examples

In order to explain these ideas more precisely, we introduce some definitions. All standard notation and terminology can be found in [10]. Suppose  $G$  is a graph with vertex set  $V$ . For any vertex  $u$  of  $V$ ,  $N[u] = \{u\} \cup \{v \in V \mid v \text{ is adjacent to } u\}$  is called the *closed neighborhood of  $u$* . The next two definitions essentially follow Gunther and Hartnell [2], [3] and Gunther, Hartnell, Nowakowski [4]. To *subvert a vertex  $v$  of  $G$*  means to remove all elements of  $N[v]$  from  $G$ . The resulting induced subgraph, called the *survival subgraph*, is exactly the subgraph of  $G$  induced by  $V \setminus N[v]$ . Lindelauf *et al.* use the term *discover* instead of subvert, but the concept is the same. The secrecy of a graph  $G$  is defined as a sum of se-

crecy values for each vertex in the graph. In the scenario where we assume 1) each vertex is equally likely to be discovered or subverted, and 2) whenever a vertex is discovered all its neighbors are detected and therefore must be removed from the graph, the secrecy measure of Lindelauf *et al.* is defined as follows. For any graph  $G$  with  $p$  vertices  $\{v_1, v_2, \dots, v_p\}$ ,

$$\begin{aligned} \widetilde{\text{TS}}(G) &= \sum_{i=1}^p \frac{1}{p} \left( \frac{p - |N[v_i]|}{p} \right) \\ &= \frac{1}{p^2} \sum_{i=1}^p (p - |N[v_i]|). \end{aligned} \tag{1}$$

This secrecy measure is combined with a standard measure for ease of information sharing, the normalized reciprocal of total distance. For any graph  $G$  with  $p$  vertices  $\{v_1, v_2, \dots, v_p\}$ , we use  $d_G(v_i, v_j)$  to denote the distance in  $G$  between  $v_i$  and  $v_j$ , suppressing the subscript whenever the graph being referred to is clear from context. The *total distance* of  $G$ ,  $\text{TD}(G)$ , is defined by

$$\text{TD}(G) = \sum_{v_i} \sum_{v_j} d(v_i, v_j).$$

The *information performance*,  $\text{IP}(G)$ , is defined by

$$\text{IP}(G) = \frac{p(p-1)}{\text{TD}(G)}.$$

Note that information performance is normalized with respect to  $K_p$  and  $0 \leq \text{IP}(G) \leq 1$  with a higher value indicating better performance. Finally, the Lindelauf measure of network performance of graph  $G$  is simply the product  $\widetilde{\text{TS}}(G) \cdot \text{IP}(G)$ . To define a secrecy measure that includes information about the order and number of components of the survival subgraph, we use a modified version of toughness of a graph, an idea first defined by Chvátal [1]. For each vertex  $v_i$  of graph  $G$ , we use  $c_i$  to denote the number of components in the survival subgraph  $G - N[v_i]$ . For the rest of the paper we use the following total secrecy measure of a graph  $G$ .

**Definition 1.** Let  $G$  be a graph with  $p$  vertices  $\{v_1, v_2, \dots, v_p\}$ . The *total secrecy* of graph  $G$ ,  $\text{TS}(G)$ , is defined by

$$\begin{aligned} \text{TS}(G) &= \sum_{i=1}^p \frac{1}{p^2} \left( \frac{p - |N[v_i]|}{c_i} \right) \\ &= \frac{1}{p^2} \sum_{i=1}^p \left( \frac{p - |N[v_i]|}{c_i} \right). \end{aligned} \tag{2}$$

This secrecy measure modifies each term of the sum in equation (1). For the subversion of each vertex, the order of the survival subgraph is replaced by the average order of each component of the survival subgraph. Using this measure we have the following definition of performance of a graph  $G$ ,

**Definition 2.** Let  $G$  be a graph with  $p$  vertices  $\{v_1, v_2, \dots, v_p\}$ . The *performance* of graph  $G$ ,  $\text{PERF}(G)$ , is defined by

$$\begin{aligned} \text{PERF}(G) &= \text{IP}(G) \cdot \text{TS}(G) \\ &= \frac{p(p-1)}{\text{TD}(G)} \cdot \frac{1}{p^2} \sum_{i=1}^p \left( \frac{p - |N[v_i]|}{c_i} \right) \\ &= \left( \frac{p-1}{p} \right) \frac{\sum_{i=1}^p \left( \frac{p - |N[v_i]|}{c_i} \right)}{\text{TD}(G)}. \end{aligned} \tag{3}$$

Note that larger values of performance occur when the expected average order of components in the survival subgraph is larger and when the total distance of the original graph is smaller. This is the measure of performance that will be used for the remainder of the paper.

**Example 1.** To illustrate these definitions we compute the secrecy, information, and performance for three basic graphs on  $p$  vertices: cycle, path, and star. For the cycle,  $C_p$ ,

$$\begin{aligned} \text{TS}(C_p) &= \frac{1}{p^2} \sum_{i=1}^p \left( \frac{p - |N[v_i]|}{c_i} \right) \\ &= \frac{1}{p^2} \sum \frac{p-3}{1} = \frac{p-3}{p}. \end{aligned}$$

For the path,  $P_p$ ,

$$\begin{aligned} \text{TS}(P_p) &= \frac{1}{p^2} \sum_{i=1}^p \left( \frac{p - |N[v_i]|}{c_i} \right) \\ &= \frac{1}{p^2} \left[ 2(p-2) + 2(p-3) + (p-4) \frac{p-3}{2} \right] \\ &= \frac{p^2 + p - 8}{2p^2}. \end{aligned}$$

For the star,  $K_{1,p-1}$ ,

$$\begin{aligned} \text{TS}(K_{1,p-1}) &= \frac{1}{p^2} \sum_{i=1}^p \left( \frac{p - |N[v_i]|}{c_i} \right) \\ &= \frac{1}{p^2} \left[ 0 + (p-1) \frac{p-2}{p-2} \right] = \frac{p-1}{p^2}. \end{aligned}$$

By comparing these formulas, one can see that for all  $p \geq 4$ ,  $TS(C_p) \geq TS(P_p) \geq TS(K_{1,p-1})$ .

The following total distance formulas for the cycle, path, and star are easily calculated:

$$TD(C_p) = \begin{cases} \frac{p^3}{4}, & \text{if } p \text{ is even} \\ \frac{p^3-p}{4}, & \text{if } p \text{ is odd.} \end{cases}$$

$$TD(P_p) = \frac{p(p^2 - 1)}{3} \text{ and } TD(K_{1,p-1}) = 2(p-1)^2.$$

Combining these formulas and recalling that

$$PERF(G) = IP(G) \cdot TS(G) = \frac{p(p-1)}{TD(G)} \cdot TS(G)$$

leads to the following performance measures for the cycle, path, and star:

$$PERF(C_p) = \begin{cases} \frac{4p(p-1)}{p^3} \cdot \frac{p-3}{p} = \frac{4(p-1)(p-3)}{p^3}, & \text{if } p \text{ is even} \\ \frac{4p(p-1)}{p^3-p} \cdot \frac{p-3}{p} = \frac{4(p-1)(p-3)}{p(p^2-1)}, & \text{if } p \text{ is odd;} \end{cases}$$

$$PERF(P_p) = \frac{3p(p-1)}{p(p^2-1)} \cdot \frac{p^2+p-8}{2p^2} = \frac{3(p^2+p-8)}{2p^2(p+1)},$$

$$PERF(K_{1,p-1}) = \frac{p(p-1)}{2(p-1)^2} \cdot \frac{p-1}{p^2} = \frac{1}{2p}.$$

With a bit more algebra it can be shown that for all  $p \geq 5$ ,  $PERF(C_p) \geq PERF(P_p) \geq PERF(K_{1,p-1})$ , a result that seems to align well with other measures of network survivability.

### 3 Main Result

The main result of this paper is the determination of a bound for performance of trees with  $p$  vertices. As we analyze the performance of trees, the following notation and definitions will be useful. We denote the degree of vertex  $v$  in graph  $G$  by  $deg_G(v)$ . We also use  $ts_G(u)$  to abbreviate  $\left(\frac{p-|N[v_i]|}{p}\right)$  and  $td_G(u)$  to abbreviate  $\sum_{v \in V} d(u, v)$ . In all these abbreviations we suppress the subscript whenever the graph being referred to is clear from context. A  $u-v$  twig in a tree is a  $u-v$  path in which  $v$  has degree at least three,  $u$  has degree one and all other vertices on the path have degree two. We call the vertex  $v$  the *base* of the sprig. A  $k$ -twig is a twig of length  $k$ . For any vertex  $v$  with  $deg_G(v) \geq 3$  and which is adjacent to  $deg_G(v) - 1$  leaves, we define a  $v$ -sprig to be the subgraph induced by  $v$  and its adjacent leaves, and we call the vertex  $v$  the *base* of the sprig.

**Observation 1.** The basic leverage for establishing the bound is the observation that for  $M$ , whenever  $\frac{a}{b} \leq M$  and  $\frac{c}{d} \leq M$ , it follows that  $\frac{a+c}{b+d} \leq M$ . For a given tree  $T$ , we look at the terms contributing to  $\sum ts(v)$  and to  $\sum td(v)$  for an individual vertex, or for a group of vertices,  $V'$ , and we show

$$\frac{\sum_{v \in V'} ts(v)}{\sum_{v \in V'} td(v)} \leq M.$$

Then for the entire vertex set,  $V$ , of  $T$

$$\frac{\sum_{v \in V} ts(v)}{\sum_{v \in V} td(v)} \leq M.$$

Another important observation is that for any vertex  $u$  in a tree of order  $p$  with  $c_u \geq 6$ , we immediately know  $ts(u) = \frac{p-|N[u]|}{c_u} \leq \frac{1}{6}p$ . Moreover, for vertex  $u$  in any tree except  $K_{1,p-1}$ , we know  $td(u) \geq p$ . Hence for all vertices with  $c_u \geq 6$ ,  $\frac{ts(u)}{td(u)} \leq \frac{p/6}{p} \leq \frac{1}{6}$ . For  $c_u \geq 3$ , it is straightforward to show that  $\frac{ts(u)}{td(u)} \leq \frac{1}{6}$ , and when  $c_u = 2$ , to show  $\frac{ts(u)}{td(u)} \leq \frac{1}{6} \left(\frac{3p+5}{3p}\right)$ . When  $c_u = 1$ , however, the proofs are more complex, and so are presented separately.

**Lemma 3.** Let  $T$  be a tree of order  $p \geq 10$  that is not  $K_{1,p-1}$  and  $u$  be a vertex of  $T$ . If  $c_u \geq 3$ , then  $\frac{ts(u)}{td(u)} \leq \frac{1}{6}$ . If  $c_u = 2$  and  $d_T(u) \geq 2$ , then  $\frac{ts(u)}{td(u)} \leq \frac{1}{6}$ . If  $c_u = 2$  and  $deg_T(u) = 1$ , then  $\frac{ts(u)}{td(u)} \leq \frac{1}{6} \left(\frac{3p+5}{3p}\right)$ .

**Proof:** Let  $u$  be a vertex of tree  $T$ , not  $K_{1,p-1}$ , of order  $p \geq 10$ . Let  $c_u \geq 3$ . For ease of notation we use  $du$  as an abbreviation for  $deg_T(u)$ . Since all vertices of  $T \setminus N[u]$  are at least distance 2 from  $u$ ,  $td(u) \geq du + 2(p - du - 1) = 2p - du - 2$ . Combining this with the fact that  $ts(u) = \frac{p-du-1}{c_u}$  and recalling that  $c_u \geq 3$ , we have

$$\begin{aligned} \frac{ts(u)}{td(u)} &\leq \frac{p-du-1}{c_u(2p-du-2)} = \frac{1}{c_u} \frac{p-du-1}{2p-du-2} \\ &\leq \left(\frac{1}{c_u}\right) \cdot \frac{1}{2} \leq \frac{1}{6}. \end{aligned}$$

Now let  $c_u = 2$ . If  $du \geq 2$ , then there are exactly 2 vertices that are exactly distance 2 from  $u$ , and at least

$p - du - 3$  that are distance at least 3 from  $u$ . Hence  $td(u) \geq du + 4 + 3(p - du - 3) = 3p - 2du - 5$ . When  $du \geq 2$ , we have

$$\frac{ts(u)}{td(u)} \leq \frac{p - du - 1}{2(3p - 2du - 5)} = \frac{1}{6} + \frac{2 - du}{3} \leq \frac{1}{6}.$$

When  $du = 1$ , the previous equation becomes

$$\begin{aligned} \frac{ts(u)}{td(u)} &\leq \frac{p - 2}{2(3p - 7)} = \frac{1}{6} + \frac{1}{6(3p - 7)} \\ &\leq \frac{1}{6} \left( 1 + \frac{1}{3p - 7} \right) \\ &\leq \frac{1}{6} \left( 1 + \frac{5}{3p} \right) = \frac{1}{6} \left( \frac{3p + 5}{3p} \right), \end{aligned}$$

since we have  $p \geq 3$ . □

We now begin the analysis of  $\frac{ts(u)}{td(u)}$  for a vertex  $u$  with  $c_u = 1$ . Note that for any such vertex  $u$  in a tree,  $T$ ,  $u$  is adjacent to  $deg_T(u) - 1$  leaves and one vertex with degree 2. Hence  $u$  is the base of a sprig and is adjacent to exactly one other vertex of degree 2, or  $u$  is the leaf on a  $k$ -twig with  $k \geq 2$ , or  $u$  is the vertex adjacent to the leaf of a  $k$ -twig for  $k \geq 3$ . These possibilities are shown in Figure 1. We begin by bounding the value of  $\frac{ts(u)}{td(u)}$  for the base of a sprig.

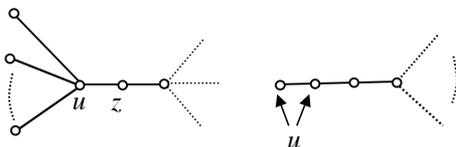


Figure 1: Vertex  $u$  with  $c_u = 1$ .

**Lemma 4.** Let  $T$  be a tree of order  $p \geq 10$  and  $u$  be a vertex of  $T$  that is the base of a sprig and with  $c_u = 1$ . Let  $z$  be the vertex of  $T$  adjacent to  $u$  with  $deg(z) = 2$ . Define  $V' = N[u] \setminus \{z\}$ . Then

$$\begin{aligned} \frac{\sum_{V'} ts(u)}{\sum_{V'} td(u)} &\leq \frac{2p - du - 3}{-2du^2 + (4p - 6)du - p + 2} \\ &\leq \begin{cases} 1/6, & \text{if } du \geq 4; \\ 1/5, & \text{if } du = 3. \end{cases} \end{aligned}$$

where  $du = deg(u)$ .

**Proof:** Let  $u_i, 1 \leq i \leq du - 1$ , be the leaves in  $N[u]$ , so  $V' = \{u_i \mid 1 \leq i \leq du - 1\} \cup \{u\}$ . Then  $ts(u) = p - du - 1$  and  $ts(u_i) = \frac{p - 2}{du - 1}$ . Thus  $\sum_{V'} ts(u) =$

$(p - du - 1) + (du - 1) \frac{p - 2}{du - 1} = 2p - du - 3$ . For the distance calculation, note that since  $c_u = 1$ , there is exactly one vertex with distance 2 from  $u$  (and 3 from each  $u_i$ ). So there are  $p - du - 2$  vertices each of which is at least distance 3 from  $u$  (and distance 4 from each  $u_i$ ). Hence  $td(u) \geq 3(p - du - 2) + du + 2 = 3p - 2du - 4$ , and  $td(u_i) \geq 4(p - du - 2) + 2(du - 1) + 1 + 3 = 4p - 2du - 6$ . Combining these, we have  $\sum_{V'} td(u) = (3p - 2du - 4) + (du - 1)(4p - 2du - 6) = -2du^2 + (4p - 6)du - p + 2$ . Hence

$$\frac{\sum_{V'} ts(u)}{\sum_{V'} td(u)} \leq \frac{2p - du - 3}{-2du^2 + (4p - 6)du - p + 2}. \quad (4)$$

To complete the proof we analyze the quadratic in the denominator. Since  $du \leq p - 2 < p - \frac{3}{2}$ , the minimum value of this quadratic is realized at the minimum value of  $du$ . Substitution into the fraction in equation (4) gives the values listed in the lemma. □

We next establish a bound for the remaining two cases when vertex  $u$  has  $c_u = 1$ , namely when (1)  $u$  is the leaf of a  $k$ -twig for  $k \geq 2$ , and (2)  $u$  is the vertex adjacent to the leaf of a  $k$ -twig for  $k \geq 3$ .

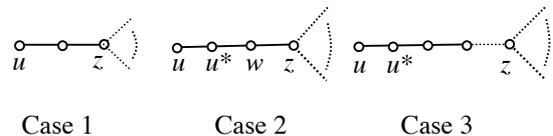


Figure 2: Lemma 5 Cases

**Lemma 5.** Let  $T$  be a tree of order  $p \geq 10$  and let  $u$  be a vertex of  $T$  with  $c_u = 1$  that is not the base of a sprig. Then  $u$  is on a  $k$ -twig with base  $z$ . For  $k = 2, 3$ , let  $V'$  be the set of vertices on the twig, not including  $z$ . For  $k \geq 4$ , let  $V'$  contain the leaf and the single vertex adjacent to the leaf of the  $k$ -twig. Then

$$\frac{\sum_{V'} ts(u)}{\sum_{V'} td(u)} \leq \frac{1}{4}.$$

**Proof:** For vertex  $u$  with  $c_u = 1$  that is not the base of a sprig,  $u$  is on a  $k$ -twig with base  $z$ . We use cases depending on  $k$ . The cases are illustrated in Figure 2. *Case 1:* Assume vertex  $u$  is on a 2-twig. Let  $z$  be the base of the 2-twig, and label the other vertex on the twig  $w$ . Let  $V' = \{u, w\}$ . Again we use  $dz$  to abbreviate  $deg(z)$ . Then  $ts(u) = p - 2$  and  $ts(w) = \frac{p - 3}{dz - 1}$ . Thus  $\sum_{V'} ts(v) = (p - 2) + \frac{p - 3}{dz - 1}$ . For distance, note that there are  $dz - 1$  vertices distance 3 from  $u$  (distance 2 from  $w$ ), and there are  $p - dz - 2$  vertices each of distance at least 4 from  $u$  (at least 3

from  $w$ ). So  $td(u) \geq 4(p - dz - 2) + 3(dz - 1) + 3 = 4p - dz - 8$ , and  $td(w) \geq 3(p - dz - 2) + 2(dz - 1) + 2 = 3p - dz - 6$ . Combining we have  $\sum_{V'} td(v) = 7p - 2dz - 14$ . Hence

$$\frac{\sum_{V'} ts(v)}{\sum_{V'} td(v)} \leq \frac{p - 2 + \frac{p-3}{dz-1}}{7p - 2dz - 14}.$$

Some straightforward algebra shows

$$\begin{aligned} \frac{p - 2 + \frac{p-3}{dz-1}}{7p - 2dz - 14} &\leq \frac{1}{4} \\ \Leftrightarrow \frac{4(p - 3)}{dz - 1} &\leq 3p - 2dz - 6 \tag{5} \\ \Leftrightarrow 0 &\leq -2dz^2 + (3p - 4)dz + p + 18. \end{aligned}$$

The minimum value of the quadratic (for  $3 \leq dz \leq p - 2$ ) occurs at the maximum value of  $dz$ , and substituting  $p - 2$  into the quadratic yields  $p^2 - 9p + 34$  which is clearly nonnegative for all  $p$ . Hence the last inequality in equation (5) is true.

*Case 2:* Assume vertex  $u$  is on a 3-twig. Then  $u$  may be the leaf or the vertex adjacent to the leaf. Since we will group these two vertices we let  $u$  be the leaf and  $u^*$  the vertex adjacent to it. Let  $z$  be the base of the 3-twig, and label the other vertex on the twig  $w$ . Let  $V' = \{u, u^*, w\}$ . In this case,  $ts(u) = p - 2$ ,  $ts(u^*) = p - 3$ , and  $ts(w) = \frac{p-3}{dz}$ . Thus

$\sum_{V'} ts(v) = (p - 2) + (p - 3) + \frac{p - 3}{dz}$ . Analyzing distances as in *Case 1*, we have  $td(u) \geq 5(p - dz - 3) + 4(dz - 1) + 6 = 5p - dz - 13$ ,  $td(u^*) \geq 4(p - dz - 3) + 3(dz - 1) + 4 = 4p - dz - 11$ , and  $td(w) \geq 3(p - dz - 3) + 2(dz - 1) + 4 = 3p - dz - 7$ . Combining we have  $\sum_{V'} td(v) = 12p - 3dz - 31$ . Hence

$$\frac{\sum_{V'} ts(v)}{\sum_{V'} td(v)} \leq \frac{2p - 5 + \frac{p-3}{dz}}{12p - 3dz - 31}.$$

Some straightforward algebra shows

$$\begin{aligned} \frac{2p - 5 + \frac{p-3}{dz}}{12p - 3dz - 31} &\leq \frac{1}{4} \\ \Leftrightarrow \frac{4(p - 3)}{dz} &\leq 4p - 3dz - 11 \tag{6} \\ \Leftrightarrow 0 &\leq -3dz^2 + (4p - 11)dz - 4p + 12. \end{aligned}$$

The minimum value of this quadratic occurs at the extreme values of  $dz$ , and  $3 \leq dz \leq p - 3$ . Substituting each of these values into the quadratic it is easy to verify that the quadratic is nonnegative for all  $p \geq 10$ . Hence the last inequality in equation (6) is true.

*Case 3:* Assume vertex  $u$  is on a  $k$ -twig,  $k \geq 4$ . As in the previous case,  $u$  may be the leaf or the vertex adjacent to the leaf. Since we will group these

two vertices we let  $u$  be the leaf and  $u^*$  the vertex adjacent to it. Let  $z$  be the base of the 3-twig. Let  $V' = \{u, u^*\}$ . In this case,  $ts(u) = p - 2$  and  $ts(u^*) = p - 3$ . Thus  $\sum_{V'} ts(v) = 2p - 5$ . For distances, note that there are  $p - k - 1$  vertices each with distance at least  $k + 1$  from  $u$ . Thus we have  $td(u) \geq (k + 1)(p - k - 1) + \sum_{i=1}^k i = \frac{-k^2}{2} + \frac{2p - 3}{2}k + p - 1$ . Note that the distance between  $u^*$  and the other  $p - 2$  vertices is 1 less than the distance between  $u$  and these other vertices. So  $td(u^*) \geq \frac{-k^2}{2} + \frac{2p - 3}{2}k + 1$ , and we have

$$\frac{\sum_{V'} ts(v)}{\sum_{V'} td(v)} \leq \frac{2p - 5}{-k^2 + (2p - 3)k + p}.$$

Again straightforward algebra shows

$$\begin{aligned} \frac{2p - 5}{-k^2 + (2p - 3)k + p} &\leq \frac{1}{4} \tag{7} \\ \Leftrightarrow 0 &\leq -k^2 + (2p - 3)k - 7p + 20. \end{aligned}$$

Note that  $k \leq p - 2$ , and so the minimum value of the quadratic occurs at the minimum value of  $k$ . We substitute  $k = 4$  and verify that the quadratic is nonnegative for all  $p \geq 10$ . Hence the inequality in equation (7) is true.  $\square$

We are now ready to prove the main result.

**Main Theorem.** *Let  $T$  be a tree of order  $p \geq 6$ . Then*

$$\text{PERF}(T) \leq \frac{p - 1}{4p}. \tag{8}$$

**Proof:** Using the list of trees of order 10 or less in Appendix 3 of [5], we have verified for all trees  $T$  with order  $p$ ,  $6 \leq p \leq 9$  that  $\text{PERF}(T) \leq \frac{1}{4}$  by direct calculation. Now let  $T$  be a tree of order at least 10. We consider vertices, or groups of vertices, according to the value of  $c_u$ . For any vertex  $u$  with  $c_u \geq 2$ , we have  $\frac{ts(u)}{td(u)} \leq \frac{1}{6}$ , by Lemma 3. If  $c_u = 1$  and  $u$  is the base of a sprig, then by Lemma 4,  $\frac{\sum_{V'} ts(v)}{\sum_{V'} td(v)} \leq \frac{1}{5}$ , where  $V'$  is the set of sprig vertices. If  $c_u = 1$  and  $u$  is not the base of a sprig, then by Lemma 5,  $\frac{\sum_{V'} ts(v)}{\sum_{V'} td(v)} \leq \frac{1}{4}$ , for the group of vertices  $V'$  described in the lemma. Using Observation 1, it follows that

$$\text{PERF}(T) = \left(\frac{p - 1}{p}\right) \frac{\sum_{V'} ts(v)}{\sum_{V'} td(v)} \leq \left(\frac{p - 1}{p}\right) \frac{1}{4}.$$

$\square$

### 4 A Family of Almost Optimal Trees

The trees that we conjecture have best performance are spider graphs in which all legs, except possibly one, have length two. Specifically, for all  $p \geq 7$  we define the double star as follows.

**Definition 6.** The *double star* with  $p \geq 7$  vertices, denoted  $DS_p$ , is the graph with vertex set  $V = \{v_0, v_1, \dots, v_{p-1}\}$  and edge set  $E$  where the precise nature of  $E$  depends on the parity of  $p$ . If  $p$  is odd,  $E = \{\{v_0, v_i\}, \{v_i, v_{i+\frac{p-1}{2}}\} \mid 1 \leq i \leq \frac{p-1}{2}\}$ . If  $p$  is even,  $E = \{\{v_0, v_i\}, \{v_i, v_{i+\frac{p-2}{2}}\} \mid 1 \leq i \leq \frac{p-2}{2}\} \cup \{\{v_{p-2}, v_{p-1}\}\}$ .

Thus the double star graph  $DS_p$  has one vertex,  $v_0$ , with degree  $\lfloor \frac{p-1}{2} \rfloor$  together with  $\lfloor \frac{p-1}{2} \rfloor$  paths of length two (2-twigs) if  $p$  is odd, or  $\lfloor \frac{p-3}{2} \rfloor$  paths of length two (2-twigs) and one path of length three (3-twig) if  $p$  is even. See Figure 3.

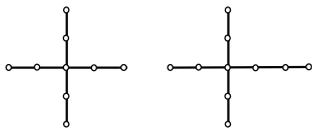


Figure 3: Double stars for  $p = 9, 10$

Hartnell and Gunther [6] have shown, in informal terms, that these double stars provide the best defense against optimal subversion strategies when one considers order of the remaining covert network and the order of the remaining components. Thus our conjecture can be seen as being an extension of their work in the sense that we are measuring the information performance of the network as well as its survivability as measured by secrecy performance.

Before calculating the performance of this graph, we establish useful fact for calculating total distance of any graph.

**Lemma 7.** Let  $G$  be a graph. If  $w$  is a leaf of  $G$ , then  $TD(G - w) = TD(G) - 2td_G(w)$ .

**Proof:** Let  $G$  be a graph with leaf  $w$ , and use  $V$  and  $V - w$  to denote the vertex sets of  $G$  and  $G - w$ , respectively. First we note that since  $w$  is a leaf,  $d_G(v, u) = d_{G-w}(v, u)$ , for all  $u, v \in V - w$ . Thus

$$td_G(v) = \sum_{u \in V} d_G(u, v)$$

$$\begin{aligned} &= \sum_{u \in V-w} d_G(u, v) + d_G(v, w) \\ &= \sum_{u \in V-w} d_{G-w}(u, v) + d_G(v, w) \\ &= td_{G-w}(v) + d_G(v, w). \end{aligned}$$

Using this information we have

$$\begin{aligned} TD(G) &= \sum_{v \in V} td_G(v) \\ &= \left( \sum_{v \in V-w} td_G(v) \right) + td_G(w) \\ &= \left( \sum_{v \in V-w} (td_{G-w}(v) + d_G(v, w)) \right) + td_G(w) \\ &= \sum_{v \in V-w} td_{G-w}(v) + \sum_{v \in V-w} d_G(v, w) + td_G(w) \\ &= TD(G - w) + 2td_G(w), \end{aligned}$$

as desired. □

Now we calculate the performance measures for the double stars and give a lower bound for the measures that is independent of parity of  $p$ .

**Theorem 8.** Let  $DS_p$  be the double star with  $p \geq 7$ .

- (a) If  $p$  is odd,  $PERF(DS_p) = \frac{p^2 - p + 2}{2p(3p - 5)}$ .
- (b) If  $p$  is even,  $PERF(DS_p) = \frac{(p^2 - 2)(p - 1)}{2p(p + 1)(3p - 8)}$ .
- (c) For all  $p$ ,  $PERF(DS_p) \geq \frac{1}{6} \left( \frac{3p + 5}{3p} \right) \left( \frac{p - 1}{p} \right)$ .

**Proof:** For part (a), we assume  $p$  is odd. Note that  $ts(v_0) = 1$  since  $DS_p - N[v_0]$  is  $\frac{(p-1)}{2}$  isolated vertices. Next observe that for  $1 \leq i \leq \frac{p-1}{2}$ ,  $ts(v_i) = 2$  since each component of  $DS_p - N[v_i]$  has exactly two vertices. Finally, for each leaf  $(\frac{p-1}{2} \leq i \leq (p - 1))$ ,  $ts(v_i) = p - 2$  since the survival subgraph is connected. Hence

$$\begin{aligned} p^2 \sum_{i=0}^{p-1} ts(v_i) &= ts(v_0) + \sum_{i=1}^{\frac{p-1}{2}} ts(v_i) + \sum_{i=\frac{p-1}{2}}^{p-1} ts(v_i) \\ &= 1 + \left( \frac{p-1}{2} \right) 2 + \left( \frac{p-1}{2} \right) (p-2) \\ &= 1 + \left( \frac{p-1}{2} \right) p = \frac{p^2 - p + 2}{2}. \end{aligned}$$

So for  $p$  odd,

$$TS(DS_p) = \frac{p^2 - p + 2}{2p^2}. \tag{9}$$

To calculate  $TD(DS_p)$  we partition the vertices in the same way that we did to calculate secrecy.

$$\begin{aligned} TD(DS_p) &= \sum_{i=0}^{p-1} td(v_i) \\ &= td(v_0) + \sum_{i=1}^{\frac{p-1}{2}} td(v_i) + \sum_{i=\frac{p-1}{2}}^{p-1} td(v_i) \\ &= 3\left(\frac{p-1}{2}\right) + \left(\frac{p-1}{2}\right)(2 + 5\frac{p-3}{2}) + \\ &\quad \left(\frac{p-1}{2}\right)(3 + \frac{p-3}{2}7) \\ &= \left(\frac{p-1}{2}\right)(8 + 6(p-3)) \\ &= (p-1)(3p-5). \end{aligned}$$

Hence  $IP(DS_p) = \frac{p(p-1)}{TD(DS_p)} = \frac{p}{3p-5}$ , and so

$$PERF(DS_p) = IP(DS_p) \cdot TS(DS_p) = \frac{p^2 - p + 2}{2p(3p-5)}$$

when  $p$  is odd.

For part (b), assume  $p$  even, then the double star has one 3-twig. So for all calculations we group the vertices on each 2-twig. For secrecy we have

$$\begin{aligned} p^2 \sum_{i=0}^{p-1} ts(v_i) &= ts(v_0) + \sum_{i=1}^{\frac{p-4}{2}} (ts(v_i) + ts(v_{i+\frac{p-2}{2}})) + \\ &\quad ts(v_{\frac{p-2}{2}}) + ts(v_{p-2}) + ts(v_{p-1}) \\ &= \frac{p}{\frac{p}{2}-1} + \left(\frac{p-4}{2}\right) \left(\frac{p-3}{\frac{p}{2}-2} + p-2\right) + \\ &\quad \left(\frac{p-3}{\frac{p}{2}-1}\right) + (p-3) + (p-2) \\ &= \frac{p}{p-2} + \left(\frac{p-4}{2}\right) \left(\frac{2p-6}{p-4} + p-2\right) + \\ &\quad \frac{2p-6}{p-2} + 2p-5 \\ &= \frac{3p-6}{p-2} + (p-3) + \frac{(p-2)(p-4)}{2} \\ &\quad + (2p-5) \\ &= \frac{p^2-2}{2}. \end{aligned}$$

So for  $p$  even,

$$TS(DS_p) = \frac{(p^2-2)}{2p^2}. \tag{10}$$

To calculate  $TD(DS_p)$  when  $p$  is even, we note that  $DS_{p-1} = DS_p - v_{p-1}$ . Since  $v_{p-1}$  is a leaf we use the previous lemma together with the formula for total distance of the double star with an odd number of vertices. First we note that  $td_{DS_p}(v_{p-1}) = 6 + 9\left(\frac{p-4}{2}\right) = \frac{3}{2}(3p-8)$ . From the previous lemma we have

$$\begin{aligned} TD(DS_p) &= TD(DS_{p-1}) + 2ts(v_{p-1}) \\ &= ((p-1)-1)(3(p-1)-5) + 3(3p-8) \\ &= (p+1)(3p-8). \end{aligned}$$

So we have

$$IP(DS_p) = \frac{p(p-1)}{TD(DS_p)} = \frac{p(p-1)}{(p+1)(3p-8)}$$

and

$$\begin{aligned} PERF(DS_p) &= IP(DS_p) \cdot TS(DS_p) \\ &= \frac{(p^2-2)(p-1)}{2p(p+1)(3p-8)} \end{aligned}$$

when  $p$  is even.

For part (c), straightforward algebra verifies that the bound in part (c) is less than or equal to each of the bounds in parts (a) and (b).  $\square$

It is reasonable to conjecture that a best performing tree has only one large degree ( $\geq 3$ ) vertex. More large degree vertices would increase the distance between peripheral vertices. At the same time, if such vertices are not the bases of sprigs, they would contribute smaller (compared to degree 1 and degree 2 vertices) terms to the total secrecy sum. Within this restriction, we can see that the double star has a large number of vertices that contribute the maximum possible term to total secrecy, namely  $p-2$ . This tree however does not have the best secrecy even among spider graphs. A spider graph with legs of length 3 (3-twigs) has approximately  $\frac{2}{3}$  of its vertices that contribute either  $p-3$  or  $p-2$ . Note that once the length of the leg (or twig) increases beyond 3, each additional vertex on the leg contributes only  $\frac{p-3}{2}$ . So it would seem that legs of length 2 or 3 would provide the best secrecy performance. The tradeoff comes, however, when total distance (used to measure information) is considered. Although, as the following proposition establishes, a spider graph with all legs of length 3 has better secrecy than the double star of the same order, the larger total distance (worse information performance) offsets, just barely, the better secrecy performance. The total performance is quite close, and so if secrecy were of higher importance, one could make the case that the spider graphs with legs of length 3 could be a better choice in certain covert network operations.

**Proposition 9.** Let  $S_p$  be the spider graph with order  $p$ . If  $S_p$  has  $p = 3k + 1$  and  $k$  legs of length 3, then  $TS(S_p) = \frac{2(p^2 - 2p + 1)}{3p^2}$  and  $IP(S_p) = \frac{3p}{4(3p - 7)}$ . Further, for  $k \geq 3$ ,  $TS(S_p) > TS(DS_p)$ , but  $PERF(DS_p) > PERF(S_p)$ .

**Proof:** Let  $G$  be the spider graph with order  $p = 3k + 1$  and  $v_0$  be the one large degree vertex. Label the vertices on each leg of the spider, in order, as  $v_{3i+1}, v_{3i+2}, v_{3i+3}$  with  $v_{3i+1}$  adjacent to  $v_0$ , and  $v_{3i+3}$  as the leaf,  $0 \leq i \leq k$ . To calculate  $TS(S_p)$ , note first that  $ts(v_0) = 2$  since each component of  $S_p - N[v_0]$  has exactly two vertices. Next  $ts(v_{3i+1}) = \frac{p-3}{k}$  and  $ts(v_{3i+3}) = p - 2$ . Hence

$$\begin{aligned} p^2 \sum_{i=0}^p ts(v_i) &= ts(v_0) + \sum_{i=0}^{k-1} (ts(v_{3i+1}) + ts(v_{3i+2})) \\ &= 2 + k \left( \frac{p-3}{k} \right) + k(2p-5) \\ &= 2 + (p-3) + \left( \frac{p-1}{3} \right) (2p-5) \\ &= \frac{2p^2 - 4p + 2}{3}. \end{aligned}$$

$$\text{So } TS(S_p) = \frac{2(p^2 - 2p + 1)}{3p^2}. \tag{11}$$

To calculate  $TD(S_p)$  we partition the vertices in the same way that we did to calculate secrecy. Note first that  $td(v_0) = 6k$ , recalling that there are  $4k = \frac{p-1}{3}$  legs on the spider. Next  $td(v_{3i+1}) = 4 + 9(k-1)$ ,  $td(v_{3i+2}) = 4 + 12(k-1)$ , and  $td(v_{3i+3}) = 6 + 15(k-1)$ . Thus

$$\begin{aligned} TD(S_p) &= \sum_{i=0}^{p-1} td(v_i) \\ &= 6k + \sum_{i=0}^{k-1} (14 + 36(k-1)) \\ &= k(6 + 14 + 36(k-1)) \\ &= k(36k - 16) \\ &= \frac{p-1}{3} (12(p-1) - 16) \\ &= \frac{4(p-1)(3p-7)}{3}. \end{aligned}$$

Hence  $IP(S_p) = \frac{p(p-1)}{TD(S_p)} = \frac{3p}{4(3p-7)}$ , and so

$$PERF(S_p) = IP(S_p) \cdot TS(S_p) = \frac{p^2 - 2p + 1}{2p(3p-7)}. \tag{12}$$

when  $p - 1$  is divisible by 3.

To see that  $TS(S_p) > TS(DS_p)$  when 3 divides  $p - 1$ , we compare  $TS(S_p)$  (see Equation 11) with  $TS(DS_p)$  (see Equations 9 and 10). Simple algebra shows

$$\frac{2(p^2 - 2p + 1)}{3p^2} > \frac{p^2 - p + 2}{2p^2}, \text{ for all } p \geq 9, \text{ and}$$

$$\frac{2(p^2 - 2p + 1)}{3p^2} > \frac{p^2 - 2}{2p^2}, \text{ for all } p \geq 8.$$

Finally, to verify that  $DS_p$  has better overall performance than  $S_p$ , we compare Theorem 8(c) with Equation 12 and note that it is easy to verify

$$\frac{1}{6} \left( \frac{3p+5}{3p} \right) \left( \frac{p-1}{p} \right) > \frac{p^2 - 2p + 1}{2p(3p-7)}, \tag{13}$$

for all  $p \geq 12$ . Direct calculation shows that  $PERF(DS_{10}) > PERF(S_{10})$ .  $\square$

It is interesting to note that the advantage in secrecy of these spider graphs over the double stars,  $TS(S_p) - TS(DS_p)$ , asymptotically increases to  $\frac{1}{6}$ , while the performance advantage of the double stars over these spider graphs,  $PERF(DS_p) - PERF(S_p) < \frac{1}{4p}$ , goes to 0 as  $p$  increases. Note that the formula for  $PERF(DS_p)$  (from Theorem 8(c)) in equation 13 drastically underestimates the actual performance of  $DS_p$  for small values of  $p$ . Hence the performance advantage is also underestimated. Table 1 shows secrecy and performance measures for  $DS_p$  and  $S_p$  for selected small values of  $p$  in order to give a more complete understanding of the true relationships.

$p$	TS		PERF	
	$S_p$	$DS_p$	$S_p$	$DS_p$
7	0.4898	0.4490	0.1837	0.1964
10	0.5400	0.4900	0.1761	0.1822
13	0.5680	0.4675	0.1731	0.1787
16	0.5859	0.4961	0.1715	0.1751
31	0.6243	0.4849	0.1688	0.1708

Table 1: Secrecy & Performance for  $DS_p$  and  $S_p$

## 5 Further Research

Since the family of double star graphs does not achieve the bound established in the main result and we have not been able to discover any trees of order at least 7 (the smallest order for which the double star is not simply a path) that perform better than the double stars, we conjecture that the bound can be improved to the performance of the double star of any particular order. A quick review of Lemma 3 and the lower bound for double star performance shows that the any vertex  $u$  with  $c_u \neq 1$  tends to keep the performance of a tree less than that of the double star. As noted in the discussion preceding Proposition 9, the vertices that can help a tree achieve performance larger than  $\frac{1}{6}$  are those with  $c_u = 1$ , and so any attempt to lower the bound to around  $\frac{1}{6}$  should focus on showing that the existence of such vertices forces the existence of other vertices with larger values of  $\frac{ts(u)}{td(u)}$ . In particular, the bound on the contribution from vertices on sprigs and twigs are prime candidates for further analysis.

Another approach to showing double stars are optimal could focus on showing that under certain conditions trees with exactly one vertex of degree greater than 2 perform better than trees with more such large degree vertices.

Once the tree conjecture is settled one could attempt to find high performance graphs by focusing, perhaps, on graphs that have neighbor connectivity at least 2. For these graphs the total secrecy will be high simply because  $c_u = 1$  for all vertices  $u$ . Moreover, whenever  $c_u = 1$  for all vertices in  $G$ ,  $TS(G)$  has a particularly simple form that should make analysis of overall performance easier. Specifically,  $TS(G) = \sum_{i=1}^p \frac{1}{p^2} (p - |N[v_i]|) = \frac{1}{p^2} \sum_{i=1}^p (p - |N[v_i]|) = \frac{1}{p^2} \sum_{i=1}^p (p - deg(v_i) - 1) = \frac{1}{p^2} (p^2 - 2q - p)$ , where  $q$  is the number of edges. This form also suggests that work on characterizing networks with fixed number of vertices and bounds on the number of edges could be fruitful. Memon *et al.* [9] contains several articles focusing on the description of graph models that have been applied successfully to study covert networks. These articles provide a good introduction to applications of graph theory in the study of covert networks.

### References:

- [1] V. Chvátal, Tough graphs and hamiltonian circuits, *Discrete Math.* 5, 1973, pp. 215–228.
- [2] G. Gunther and B. Hartnell, On minimizing the effects of betrayals in a resistance movement, *Proc. 8th Manitoba Conf. on Numerical Mathematics and Computing*, Winnipeg, Manitoba, Canada, 1978, pp. 285–306.
- [3] G. Gunther and B. Hartnell, Optimal k-secure graphs, *Discrete Appl. Math.* 2, 1980, pp. 225–231.
- [4] G. Gunther, B. Hartnell, and R. Nowakowski, Neighbor-connected graphs and projective planes, *Networks* 17, 1987, pp. 241–247.
- [5] F. Harary, *Graph Theory*, Addison-Wesley, Reading, MA 1969.
- [6] B. Hartnell and G. Gunther, Security of underground resistance movements, in *Mathematical Methods in Counterterrorism*, N. Memon, J.D. Farley, D.L. Hicks and T. Rosenorn, eds, Springer Verlag, Wien 2009, pp. 185–204.
- [7] R. Lindelauf, P. Borm, and H. Hamers, The influence of secrecy on the communication structure of covert networks, *Social Networks* 31, 2009, pp. 120–137.
- [8] R. Lindelauf, P. Borm, and H. Hamers, On heterogeneous covert networks, in *Mathematical Methods in Counterterrorism*, N. Memon, J.D. Farley, D.L. Hicks and T. Rosenorn, eds, Springer Verlag, Wien 2009, pp. 215–228.
- [9] N. Memon, J. D. Farley, D. L. Hicks and T. Rosenorn, eds, *Mathematical Methods in Counterterrorism*, N. Memon, J.D. Farley, D.L. Hicks and T. Rosenorn, eds, Springer Verlag, Wien 2009,
- [10] D. B. West, *Introduction to Graph Theory*, second ed. Prentice Hall, Providence, RI 2001.