Assessing Optimality of Covert Networks

1st Nathan Daino
nda@ad.unc.edu

2nd Alex Li

3rd Michael Su

***Abstract*—We investigate the properties of network graphs as representatives of covert communication networks. Building on and verifying portions of previously conducted work, we introduce a new measure that (very) weakly correlates with graph connectivity and minimum edge cut, quantifying the amount of time a bit of information can be sent through a network before network severance by way of an adversary discovering and removing random nodes occurs. To incorporate this new metric into the existing metrics in the network analysis literature, we make use of a three-party Nash bargaining solution. We investigate whether 500,000 iterations of random graph generation is enough to draw conclusions about the maximization of the coupled Efficiency-Secrecy μ metric as introduced in a previously authored paper. Finally, we investigate the practicality of modeling networks with weighted digraphs representative of information flow rates between nodes.**

I. Introduction: What is a Covert Network?

A covert network is a communication model whereby multiple parties, represented in network graphs by nodes, communicate with each other (represented by edges) and wish to evade detection by an adversary. In the analysis of covert networks, the encryption status of a message is sometimes unimportant, as the goal is the prevention of adversary detection of individual parties participating in the network as opposed to the specifics of the messages being conveyed. This is precisely the case in our present investigation.

Nodes in a network can be discovered by an adversary in a probabilistic manner. For example, in Lindelauf, Borm, and Hamers (2008) [1], section 4.1 is devoted to the supposition that the communications of a given node are intercepted at random. 4.2 proceeds with the idea that each node has an exclusive probability of being discovered, uniform across all nodes. The most general form is reached in section 4.3, where each node may have its own detection probability. When nodes are detected, it is assumed that all neighboring nodes, that is to say, all nodes connected to the discovered party by way of an edge, are in turn discovered. While in theory this could be extended an order (or multiple) further, each continued order diminishes the usefulness of the model; if a nodes' connected nodes are discovered and those nodes' connected nodes are discovered, and so on, all nodes will be discovered in a connected component of a graph (Goodrich and Tamassia, 2015) [2]. It follows that if the graph is wholly connected, all nodes are discovered. For this reason, we limit the discovery order to 2; a node and all connected nodes are discovered, but no subsequent nodes are revealed by the original or secondary discoveries.

II. The μ Value

Lindelauf et al. [1] introduce a value called the μ value to balance secrecy and efficiency of communication in covert networks. Primarily of interest will be the formulation for secrecy and efficiency in 4.3, the most general case. We will briefly recap the justification here. First, we assume that the probability of a node being discovered is a function of its importance to the network. This importance is quantified by the number of times a node appears in a random walk through the graph (Lindelauf et al.) [1]. Each node in a graph is also assigned a u_i value dictating the proportion of the graph that is left unexposed upon that node's and its immediate neighbors' discovery. The authors assert that the secrecy of a graph, $S(g)$, can be determined by summing over each graph vertex the product of this unexposed network portion factor and the probability a_i that an adversary discovers the node. The authors note that this definition constitutes an expectation value of the number of nodes left unexposed under an act of surveillance.

The efficiency measure, $I(g)$, is determined to be inversely proportional to the total distance of the graph (the cumulative distances of all path lengths between all pairs of points in a graph) (Goddard, Swart, and Swart) [3]. The authors justify this by noting that the longer a message travels through a network, the higher the probability of discovery. It thus becomes crucial to minimize the travel time of information in the network. The authors normalize this total distance reciprocal by introducing the term $n(n-1)$ to the numerator, after having proven that the total distance of each graph under consideration is greater than or equal to this value.

The assumptions introduced uniquely to section 4.3 are that the probability of discovery is assigned by a distribution gathered from a random walk, whose discrete pieces are dictated by $\pi_i = \frac{d_i+1}{2m+n}$ and whose u_i value is $1 - \frac{d_i+1}{n}$ (Lindelauf et al.) [1]. The secrecy measure is then asserted to be the summation over all vertices of the product of the relevant portion of the probability distribution and the u_i value. The final μ value the authors arrived at is

$$\mu_3(g) = \frac{n-1}{2m+n} \cdot \frac{2m(n-2) + n(n-1) - \sum_{i \in V} {d_i}^2}{T(g)}, \quad (1)$$

Where m and n are the order and size of a graph, respectively, d_i is the degree of a given node, and $T(g)$ is the total distance of the graph (Lindelauf et al.) [1].

While it may have seemed intuitive up until this point to simply find the product of the efficiency and secrecy measures and maximize this product to arrive at an optimal trade off between the two, this is a highly non-trivial result. It turns out that these two quantities must satisfy the tenets of the Nash bargaining solution. It is not our place here to assert that the μ value satisfies these conditions. Rather, we save this discussion for combining secrecy, efficiency, and a metric we call walk length into a tripartite metric based upon others' work of generalizing the Nash bargaining solution to three parties.

III. Verifying Optimal Graphs of Orders 2-10

We presently concern ourselves with the results of section 4.3 of Lindelauf et al. [1]. The authors claim to offer precise results for graphs of orders 2 through 7, and approximately optimal graphs of orders 8 through 10. For each order, we ran our own calculation of the μ value for 500,000 iterations of random connected graphs of a given order, saving the graph with the maximum μ value. The question was then raised as to how to properly compare our supposedly optimal findings with the authors'. The Networkx Python library allows users to compute the edit distance between two graphs. Two graphs' similarity is quantified as a value equivalent to the minimal number of deletions or changes made to edges or nodes to gradually transmute from one graph to the other. ("optimize_graph_edit_distance") [4].

It is worth noting that two techniques are introduced in the paper with the goal of arriving at a maximum-μ valued graph. The first of these technques generates 500,000 graphs of a given order, saving the newest graph if it has the highest μ value as yet seen. The second technique involves starting with a seed graph of the graph order under investigation, and then adds edges gradually if the μ value increases with a given addition (Lindelauf et al.) [1]. In the table below, we summarize our findings for both techniques. Note that graphs are assigned a relative edit distance (REL), which we believe to be a previously unexplored metric of graphs.

Graph Order	Tech. 1 REL	Isomorphic	Tech. 2 REL	Isomorphic
2-5	0.00	Y	N/A	N/A
6	0.25	N	N/A	N/A
7	0.07	N	N/A	N/A
8	0.00	Y	.21	N
9	0.14	N	.14	N
10	0.12	N	.24	N

We perceive the relative edit distance to be a better indicator of graph similarity than raw edit distance because the metric takes into account the size of a graph. An absolute edit distance of 1 between two randomly generated graphs of order 20 seems remarkably more similar than an edit distance of 1 between graphs of order 3, because a smaller percentage of components of the graph need to be changed to transmute one graph into the other. Notice that two graphs with a relative edit distance of zero are called isomorphic; according to UPenn, "Two graphs G1 and G2 are isomorphic if there exists a matching between their vertices so that two vertices are connected by an edge in G1 if and only if corresponding vertices are connected by an edge in G2" ("Graph Isomorphism", 2016) [5]. Informally, isomorphic graphs are those in which a renaming of nodes is sufficient to transmute one graph into another. These graphs are indistinguishable by way of node adjacency relationships, despite possibly having differently appearing shapes as generated by NetworkX.

IV. Introducing the Walk Length

Upon reflection, it was decided that the secrecy-efficiency model was missing a crucial element of covert network behavior; namely, the dependence of networks on single individuals. We wanted to generate graphs that not only optimized efficiency and secrecy, but also prevented an adversary from easily severing the network into two pieces, upon which we assumed for the purpose of convenience that this isolated an individual that was crucial to the operation of the network. The adversary continuously removes random nodes until a network is split into two connected components. We are well aware of the limitations of this metric. It is clear, for example, to see that any separated node is not in general crucial to the operation of a network, and furthermore that this model is more akin to the node discovery model of section 4.1 of Lindelauf et al [1] than section 4.3. Even with these limitations, we felt that this metric had merit in describing the robustness of network operations under the threat model in 4.1.

At first, it was suspected that this walk length metric would be highly correlated with either graph connectivity or a more obscure graph characteristic called the minimum edge cut, which measures the minimum number of edges that need to be removed from a graph remove its connected status ("minimum_edge_cut") [6]. This did not appear to be the case. The walk length metric appears instead to be very weakly correlated with both connectivity and minimum edge cut, at least for graphs of order 20, as figures 1 and 2 suggest.

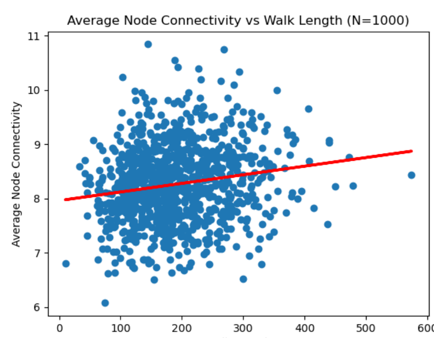

Figure 1: R value ≈.17

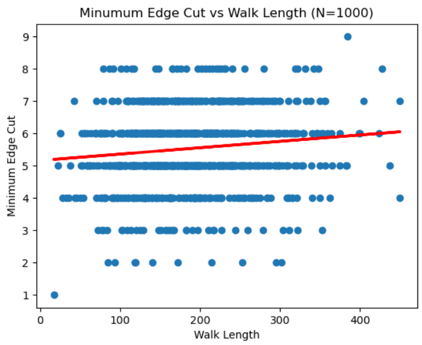

Figure 2: R value ≈.12

V. The Nash Bargaining Solution

Our present goal now shifts to integrating our new walk length metric with the existing μ metric. To ensure that the Nash bargaining solution is an appropriate way to go about combining these three metrics, we must verify that the solution axioms apply in our case. These will be lifted verbatim from Lindelauf et al (2008) [1] and Harsanyi (1959) [7] and addressed one by one:

1) Efficiency; To paraphrase Harsanyi, knowing one solution guarantees there exists no other solution that has higher payoffs for every part. This is desirable in our case because we want to maximize to the fullest extent all three quantities, and so by necessity if there is a solution that universally produces higher payoffs we would immediately elect to take that solution instead, challenging the notion that the first "solution" was a real solution at all.

2) Symmetry; Lindelauf et al. [1] note that if one of the two metrics does not take precedence over the other, that is to say if we value, in this case, all three metrics to the same extent, we may qualify this relationship as symmetry. Indeed, we see no obvious reason to value one metric much more than the other, and although it is fallacious to assume that this by itself means that we should weight each metric the same, it seems intuitive that any of the metrics is equally valuable.

3) Linear Invariance (included here for completeness but an explanation is omitted because we have no new insight to add beyond that of Lindelauf et al. [1] or Harsanyi [7]).

4) Independence of Irrelevant Alternatives; This merely states that if we take a proper subset of randomly generated connected graphs and a solution to the full set exists in this smaller set, then we have necessarily found a solution for the proper subset (Lindelauf et al.) [1]. This makes sense in our case because no graph generated in the subset will be omitted from the full set, all elements of which the full-set solution already dominates.

(5) Completeness; It seems convincing that there exists no solution other than what is afforded by axioms 1 through 4, as Harsanyi [7] necessitates. If a "solution" exists but it is not Pareto efficient, it is worthless because it does not maximize the utility functions for all parties. If the "solution" does not preserve symmetry we would have to justify favoring one metric over another, which as yet we have failed to do. Violating the independence of irrelevant alternatives appears to be a violation of basic set theory, and the violation of linear invariance means that the solution is somehow scale-dependent, which makes little sense in our case given that our scales are largely arbitrarily chosen. A unit of walk length, for example, has little to do with a unit of secrecy.

Particularly exciting for us is that Harsanyi's paper assumes n parties as opposed to the two parties supposed in Lindelauf et al. [1], and so we can be confident that if we are satisfied with these axioms, we have an optimal solution mechanism for our three-party case. Omitting the proofs that can be found in Harsanyi's paper, we arrive at the following n-party solution:

$$\pi = \prod_{i=1}^{n}(u_i - d_i) \tag{2}$$

Here, u_i is the utility function for each metric, and d_i is the disagreement payoff, which Harsanyi characterizes the payoff if any of the parties does not agree to bargaining and acts purely out of self interest. To deal with that case, we have mirrored the supposition of Lindelauf et al. [1] that any instance that solely maximizes one metric's payoff is unfeasible for a covert network. This removes the necessity for disagreement payoffs, as they become 0, and simplifies our solution marginally to the following:

$$\pi = \prod_{i=1}^{n}(u_i) \tag{3}$$

Indeed, we find that the Nash bargaining solution generalizes intuitively to the product of the three metrics, and so we may proceed with multiplying the μ value with the walk length to obtain our new χ value, as we call it.

VI. On the Adequacy of 500,000 Iterations

Both graph generation techniques of Lindelauf et al. [1] rely on 500,000 iterations being enough to draw conclusions about the optimal graph for each order. We wish to determine an adequate sample size for each graph order. Smith [8] provides the following formula for sample size, denoted by n (Gerstman, 2006) [9]:

$$n = \frac{Z^2 \cdot \sigma(1-\sigma)}{\epsilon^2}, \tag{4}$$

Where Z is a desired Z-score, σ is a standard deviation, and ϵ is a desired error margin. Adapted to the information we know, as well as assuming a 95 percent confidence and a 5 percent confidence interval, this formula becomes:

$$n = \frac{1.96^2 \cdot \sigma(1-\sigma)}{.05 \cdot \bar{\mu}}, \tag{5}$$

Where $\bar{\mu}$ is the average of a μ metric. Figure 3 is a chart displaying the relationship between sample size and graph order. The appropriate sample size starts high and rapidly declines, which seems to be the opposite behavior we might expect from a rapidly increasing sample space. An explanation for this may well reside in how the sample sizes were calculated. Notice that a standard deviation is required, but knowing the true standard deviation of μ values for graphs of a given order would require knowing all of the possible connected graphs in the sample space, after which finding the maximum μ-valued graph would be trivial. Because for high orders this is an impossibility, we have instead calculated the standard deviation of 100 randomly sampled graphs of each order (with replacement; another potential confounding factor at the lower graph orders, as the number of samples vastly exceeds the number of possible connected graphs) and used the standard deviation in μ value from these graphs in the formula.

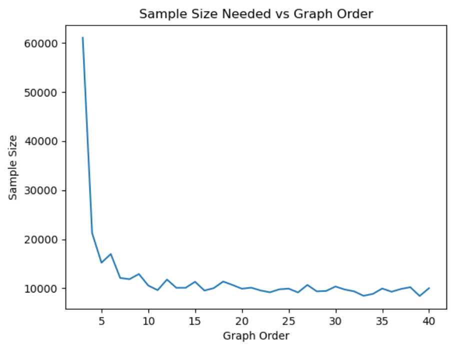

Figure 3: Sample size needed for a 95 percent confidence and 5 percent error margin of the mean μ value for graphs of a given order.

This practice is not without scientific basis. According to Sullivan [10], "Data from the participants in the pilot study can be used to compute a sample standard deviation..."

Why, then, do we stop at a value of 100? This is simply because computation becomes quite expensive after this point. Are these small sample sizes at the higher graph orders artifacts of an inadequate pilot sample size, or are the results truly indicative of the more-than-adequate nature of a sample size of 500,000 for studying graphs' μ values up to order 40? This, we feel, would be a prime candidate for follow-up work with a statistician.

VII. Modeling Information Flow with Weighted Digraphs

We seek to end our investigations with a brief discussion of modeling information volumes. Suppose that we retain the same graph structure we have come to know, whereby nodes are representative of participants in a network and the edges are communication channels. Now suppose that these edges are directed, such that information flows specifically from one node to another over a channel. Suppose furthermore that to each directed edge is assigned a weight corresponding to the amount of information transmitted over a given channel. In reality, it is clear to see that these weights would be time-varying, but we feel that modeling the static case is a valuable stepping stone toward modeling the more general dynamic case.

There is a drawback to modeling information flow in this manner: strictly speaking, flow networks are required to have a source node, from which information flows but into which no information flows, and a sink node, into which information flows but out of which no information flows (Demaine, Karger, and Andoni, 2003) [11]. However, we can adapt some principles of network flows.

It seems natural to assume two different scenarios when modeling information centrality of nodes. The first, which we have taken to calling the "commander" type node, presumes that those with high information outflow to inflow ratios are particularly likely to be central network figures because they are primarily wellsprings of commands and/or information as opposed to primarily collectors of commands and/or information.

In this case, it seems appropriate to assign a probability density distribution that:

1) Is proportional to the outflow/inflow ratio x:

$$y = mx \tag{6}$$

2) Begins at at (0,net_{min}):

$$y = m(x - net_{min}) \tag{7}$$

3) Normalized so total probability is one:

$$\int_{net_{min}}^{net_{min}} m(x - net_{min})\mathrm{dx} = \frac{m(net_{max})^2}{2} - \frac{m(net_{min})^2}{2} + m(\mathrm{net}_{max})(\mathrm{net}_{min} - \mathrm{net}_{max}) \tag{8}$$

$$\frac{m(net_{max})^2}{2} - \frac{m(net_{min})^2}{2} + m(\mathrm{net}_{min})(\mathrm{net}_{min} - \mathrm{net}_{max}) = 1 \tag{9}$$

$$y = \frac{2(x - net_{min})}{(net_{max}^2) - (net_{min}^2) + 2(net_{min})(net_{min} - net_{max})}, \tag{10}$$

Where the net flow of a node is computed by subtracting the sum of its outgoing edges from the sum of its incoming edges, the minimum net flow, net_{min}, is the smallest of all such net flow values across all nodes, and the maximum net flow is similarly evaluated.

This constitutes a remarkably simple technique for differentiating between primary producers of information and primary recipients of information. Nodes can be assigned a measure from the empirical frequency with numbers in the neighborhood of a node's outflow/inflow ratio show up in a random sample from the continuous probability distribution. This naturally normalizes the measure between 0 (pure recipient) and 1 (pure distributor). And so it follows that net flows near 0.5 signify "router" nodes, by which an approximately even outflow to inflow ratio is meant, and those near 1 signify commander nodes. In practice an agreeable threshold of proximity to the actual outflow/inflow ratio will need to be selected. Note that the numbers are merely categorical; no further meaning assigns 1 as "more important" than 0 or the reverse except as needed by a specific application.

There is another factor to be considered. Namely, we may investigate the proportion of inflow and outflow contributed by a node to the total flow of the network:

$$Prob(v_i) = \frac{|O_i| + |N_i|}{\sum_{n}^{|V|} |O_n| + |N_n|}, v_i \in V, \tag{11}$$

Where O_i is an outflow of a particular node, N_i is an inflow to that node. The equation says simply that the probability assigned to a node for discovery is the sum of the absolute values of the outflow and inflow over the summation of all sums of such absolute values for every node, including where $i = n$. It is a single node's share of the total magnitude of network flow. The formula is appealing due to its equal weighting of inflow and outflow.

It can be shown that the summation of probabilities of all nodes is equal to 1. Note that

$$\sum_{i=1}^{|V|} Prob(v_i) = \sum_{i=1}^{|V|} \frac{|O_i| + |N_i|}{\sum_{n}^{|V|} |O_n| + |N_n|},$$
$$v_i \in V = \frac{\sum_{i=1}^{|V|} |O_i| + |N_i|}{\sum_{n=1}^{|V|} |O_n| + |N_n|} = 1 \tag{12}$$

with the property of removing constants from inside the summation attributable to Dawkins (2018) [12].

Further, this result is a result of extending the assumption in Lindelauf et al. [1] that a message that spends twice as long in a network is twice as likely to be discovered. Consider that edge weights are in effect rates of information transmission, and if we observe transmissions over any one unit of time the total information passing into and out of a given node is equal to $|O_i| + |N_i|$.

In this way, it can be shown that we have constructed a measure that proportionately increases with an increase in information-time spent passing through a node. This line of thinking retains Lindelauf's [1] section 4.3 supposition that nodes are discovered, not edges. This model makes an additional important assumption. Consider that the unweighted status of the nodes themselves implies instantaneous pivoting of information from one edge to another. Our model comes very near to approximating this with a uniform (regardless of edge transmission rate) *infinitesimal* time assigned to each time a bit travels through a node. In this manner, we may reconcile information flow through nodes with the idea of time spent in a network being proportional to discovery likelihood. This works because even with infinitesimal transmission times through nodes, we may consider that doubling the rate of information flow doubles the amount of total time any information has spent at the node, and we see that more information passing through a node over a given time period is proportional to the total bit-hours spent at a node, and is subsequently proportional to centrality in the network if we take information discovery to be a proxy for node discovery.

The question then becomes: how do we synthesize these two useful metrics, one of which corresponds to the proportion of information sent versus received, and the other corresponding to the size of the total information flow of a node? There appears to be no bargaining problem since Pareto optimality is not a condition we wish to assume. We must turn to a technique of mathematical economics called goal programming. As prescribed by "hazmat" [13], perhaps the most accessible solution is to take the number of standard deviations of each metric from its respective mean, normalize them, and then take their sum:

$$\frac{Prob_{vi} - \mu_1}{\sigma_1} + \frac{y_i - \overline{y}}{\sigma_2}, \tag{13}$$

(13) Converts the measures to standard units. The equation after normalization of the sum via the softmax function [14] of two standard units is:

$$\frac{e^{\frac{Prob_{vi} - \mu_1}{\sigma_1} + \frac{y_i - \overline{y}}{\sigma_2}}}{\sum_{n=1}^{|V|} e^{\frac{Prob_{vn} - \mu_1}{\sigma_1} + \frac{y_n - \overline{y}}{\sigma_2}}} \tag{14}$$

Where σ_1 is the standard deviation of the connectivity contributions (11) of each node, μ_1 is the average of such contributions, y is equation (10) computed for a given node, $\overline{y}$ is the average value of equation (10) between the net_{min} and net_{max} bounds, and σ_2 is the standard deviation of the y values of all nodes in the graph.

Let us not lose sight of the goal of this section: to establish a probability distribution by way of a metric that weights more central nodes as more likely to be discovered. We accomplished this goal by taking into consideration the volume of information flowing through a node and its ratio of outflow to inflow. Individually these metrics model centrality in probability distributions that may be assigned similarly to 4.3 of Lindelauf et al. Together, they comprise a holistic way of looking at node centrality that includes not just the portion of information flow contributed by a graph but also its particular role (commander, router, or otherwise) in the network. [1].

VIII. Proof of Concept

Excitingly, we are able to present the results of a toy network model to illustrate the utility of our metrics, both separate and combined. Figures 4, 5, and 6 show graphs with node sizes proportional to their $Prob(v_i)$ metric, their y_i metric, and their combined metric, respectively.

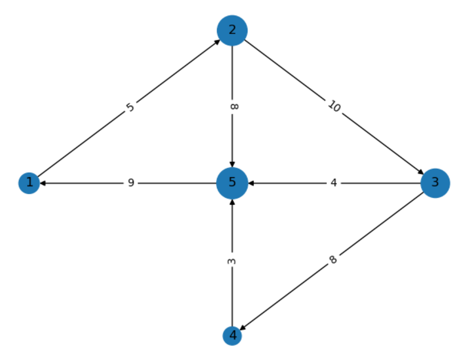

Figure 4: Node sizes proportional to $\frac{Prob_{vi} - \mu_1}{\sigma_1}$

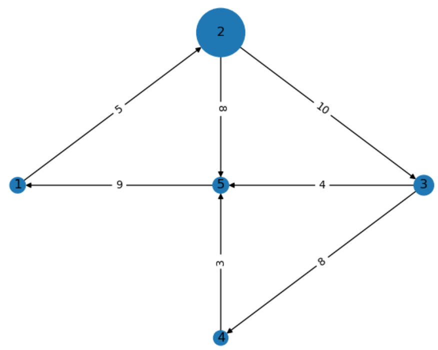

Figure 5: Node sizes proportional to $\frac{y_i - \overline{y}}{\sigma_2}$

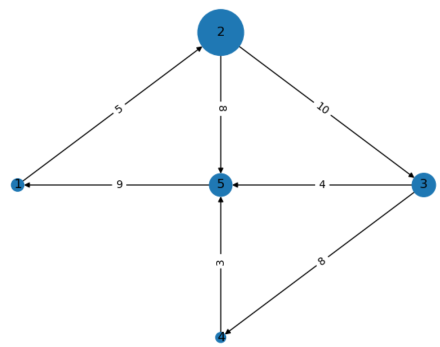

Figure 6: Node sizes proportional to combined metric (13)

Note that the probability distribution assigns weights according to the following priorities:

Outflow/Inflow	Contribution to total inflow, outflow	Probability
High	High	High
High	Low	Medium
Low	High	Medium
Low	Low	Low

This has precisely the desired effect of filtering the strong commander nodes (those with high outflow/inflow ratio and high contribution to the total graph's outflow and inflow) to the highest probability of discovery, followed by those with one strong characteristic, the other, or a blend of two medium characteristic scores. The lowest probability scores are assigned to those nodes that show neither centrality element in a given graph.

It is critical for the future researcher to understand how these graphs were generated. The first two were computed using the softmax function on the left and right halves of (13) separately. The combined-metric graph was computed using (14) directly. However, it is important to note that the exponential base is not e by necessity; rather, it is chosen for convenience. The future researcher can feel free to change this value so long as it stays strictly greater than 1 [15]. The importance of the base remaining greater than one is to ensure that larger exponents lead to larger overall values than those of smaller exponents. The effect of assigning node weights (and subsequently, proportional node sizes) to (14) with a different exponential base is the level to which differences in node size are pronounced. It is possible that some nodes are so large and others so small from a given base choice as to obscure the true structure of the graph in NetworkX. Massaging this exponent base fixes the issue in many cases.

To scale the whole graph at once is similarly trivial. It is achieved by multiplying all node weights by a constant, being sure to divide this constant out when NetworkX viewing is complete and it is time to assign a probability distribution that sums to one.

Another peculiarity of the softmax function is that it does not adequately assign probability distributions to lists of numbers both less than and greater than one in absolute value. To resolve this issue while maintaining the order of priority in probability assignment, all values in (13) were iteratively scaled by 1.1 until each value was greater than one in absolute value. The supplementary file *thetest.py* on Github should elucidate this process.

One byproduct of this generation process is that figures 4, 5, and 6 should not be compared to each other because scaling may vary. They should be used as a tool for the relative size comparison of nodes in the *same* graph.

IX. Implementation into Lindelauf et al.'s Model

We have a new probability distribution based on a new centrality measure. Now we wish to synthesize a new network model based on Lindelauf et al. [1] section 4.3. In fact, most of the work has already been done. The efficiency measure is wholly independent of our new probability distribution. The only thing that changes is $S(g)$, by which we replace the authors' centrality probability distribution with our own:

$$S(g) = \sum_{i \in V} \alpha_i(g) u_i(g) = \sum_{i \in V} \left(\frac{e^{\frac{Prob_{vi} - \mu_1}{\sigma_1} + \frac{y_i - \overline{y}}{\sigma_2}}}{\sum_{n=1}^{|V|} e^{\frac{Prob_{vn} - \mu_1}{\sigma_1} + \frac{y_n - \overline{y}}{\sigma_2}}} \right) \left(1 - \frac{d_i + 1}{n}\right). \tag{15}$$

What, exactly, does this hefty equation represent? Merely the expectation value for the amount of the network remaining after an epoch has passed and nodes (and their neighbors) are discovered according to their respective probabilities.

X. Concluding Remarks

We have presented and checked the validity of a few of Lindelauf et al.'s key results [1]. We have introduced our own measure, the walk length, as a metric to model the optimality of a covert network graph. Furthermore, we have synthesized this metric with the μ metric to generate the χ metric to score networks on their how conducive they are to covert operation. Lastly, we have investigated how to model covert networks where the flow of information is captured and quantified, and we have arrived at a new centrality measure to with which to assign probability distributions to node discovery. We did so by combining scores for the ratio of a node's outflow to inflow and the node's contribution to the total flow of a graph.

One would be justified to ask what contribution this makes, if any, to the theme of privacy enhancing technologies. We posit that this work contributes a great deal to the effort. Consider once more that a covert network, by its nature, attempts to conceal itself from discovery. While encryption plays a part in secret communication, and certainly can be used in conjunction with covert network operation, these networks seek to add to the secrecy by concealing the identities of operatives and preserving the maximum amount of undiscovered operatives. Thus, it is justified to observe the *topology* of the communication network to ensure discovery of all nodes is as difficult as possible. Optimizing these networks equates to optimizing how groups communicate privately while preserving the secrecy of the most of their participants possible.

Let us close with a discussion of the broader implications of this kind of work. What was the justification to study covert networks? Lindelauf et al. [1] offer the explanation that to operate counterterrorist operations that function as covert networks, once must understand their enemy. This alone is an insufficient justification. It does not go far enough to explain the uses of covert networks as a force for good: for organized resistance against tyranny, for control of information flow by militaries to minimize losses in case of messenger interception, and perhaps even in civilian use cases as a way to facilitate communication networks resistant to dependence on single servers and their potential failure.

Does the kind of work embodied in this paper have a place in modern academia, especially after the warnings of Rogaway (2015) in "The Moral Character of Cryptographic Work" [16]? Objections to the work on a moral basis are likely centered on the idea that terror cells may use the research to organize more effectively. To counter this, one must accept an opposing idea to Rogaway's work: the inventor of popularizer of a technique is not responsible for the malicious actions of others using said technique. If we are to assume malicious users to be rational beings capable of self conductance, the blame must be placed squarely on them for implementing a technique, itself inert and harmless until set into motion, in an unscrupulous manner. Rogaway seems at times to confuse the invention of a technique with its application, but the former does not immediately require the latter in a guaranteed cause and effect relationship.

If a paper author is not liable for a paper's malicious uses after publishing, what can be said of assigning blame for an author's malicious intentions? This question seems best left to legal philosophy and is in any case irrelevant since there are no ill intentions on the part of the authors of this paper.

REFERENCES

[1] R. Lindelauf, P. Borm, and H. Hamers, "The influence of secrecy on the communication structure of covert networks," *Social Networks*, vol. 31, no. 2, p. 126–137, 2009.

[2] Goodrich and Tamassia, "Depth-first search." [Online]. Available: https://www.ics.uci.edu/ goodrich/teach/cs260P/notes/DFS.pdf

[3] W. Goddard, C. S. Swart, and H. C. Swart, "On the graphs with maximum distance or k-diameter." [Online]. Available: https://people.cs.clemson.edu/ goddard/papers/extremalDistance.pdf

[4] "Optimize_graph_edit_distance." [Online]. Available: https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.similarity.optimize_graph_edit_distance.html

[5] E. Lazar, "5.2 graph isomorphism," Feb 2016. [Online]. Available: https://www2.math.upenn.edu/ mlazar/math170/notes05-2.pdf

[6] "Minimum_edge_cut." [Online]. Available: https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.connectivity.cuts.minimum_edge_cut.html

[7] J. C. Harsanyi, "17. a bargaining model for the cooperative n-person game," *Contributions to the Theory of Games (AM-40), Volume IV*, p. 325–356, 1959.

[8] S. M. Smith, "Determining sample size." [Online]. Available: https://uncw.edu/irp/ie/resources/documents/qualtrics/determining-sample-size-2.pdf

[9] B. Gerstman, "3: Summary statistics," 2016. [Online]. Available: https://www.sjsu.edu/faculty/gerstman/StatPrimer/sumstats.pdf

[10] L. Sullivan, "Power and sample size determination." [Online]. Available: https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_power/bs704_power_print.html

[11] A. Andoni, "Maximum flows." [Online]. Available: https://ocw.mit.edu/courses/6-854j-advanced-algorithms-fall-2005/resources/max_flow_dff/

[12] P. Dawkins, "Section 7-8 : Summation notation." [Online]. Available: https://tutorial.math.lamar.edu/classes/calci/summationnotation.aspx

[13] Hazmat and user2321, "Combining multiple metrics to provide comparisons/ranking of k objects [question and reference request]," Sep 2015. [Online]. Available: https://stats.stackexchange.com/questions/154888/combining-multiple-metrics-to-provide-comparisons-ranking-of-k-objects-question

[14] B. Gao and L. Pavel, "On the properties of the softmax function with application in game theory and reinforcement learning." [Online]. Available: https://arxiv.org/pdf/1704.00805.pdf

[15] A. de Brébisson and P. Vincent, "An exploration of softmax alternatives belonging to the spherical loss family," Feb 2016. [Online]. Available: https://arxiv.org/abs/1511.05042

[16] P. Rogaway, "The moral character of cryptographic work," Dec 2015. [Online]. Available: https://web.cs.ucdavis.edu/ rogaway/papers/moral.html