Hiding in Multilayer Networks
Supplementary Materials

Marcin Waniek[1,2]**, Tomasz P. Michalak**[2]**, Talal Rahwan**[1]
[1] Computer Science, New York University Abu Dhabi, UAE
[2] Institute of Informatics, University of Warsaw, Poland

Proofs of Complexity Results

Observation 1. *The problem of Multilayer Global Hiding is in P given the degree centrality measure. In fact, for a given problem instance either any A^* that connects $\widehat{v}$ with all contacts is a solution, or there are no solutions at all.*

Proof. Any valid solution to the problem A^* must connect the evader $\widehat{v}$ with all contacts. Therefore, after the addition of A^* the degree centrality of $\widehat{v}$ is $|F|$, while the degree centrality of every contact increases by 1. Hence, the degree centrality ranking in the network does not depend on the choice of layers in which $\widehat{v}$ gets connected with its contacts. □

Theorem 1. *The problem of Multilayer Global Hiding is NP-complete given the closeness centrality measure.*

Proof. The problem is trivially in NP, since after the addition of a given A^* the closeness centrality ranking can be computed in polynomial time.

Next, we prove that the problem is NP-hard. To this end, we show a reduction from the NP-complete problem of *Exact 3-Set Cover*. The decision version of this problem is defined by a set of subsets $S = \{S_1, \ldots, S_m\}$ of universe $U = \{u_1, \ldots, u_{3k}\}$, such that $\forall_i |S_i| = 3$. The goal is to determine whether there exist k pairwise disjoint elements of S the sum of which equals U.

Given an instance of the problem of *Exact 3-Set Cover*, let us construct a multilayer network, $M = (V_L, E_L, V', L)$, as follows (Figure 1 depicts an instance of this network):

- **The set of nodes V':** For every $u_i \in U$ we create a node u_i, as well as 3 nodes $w_{i,1}, w_{i,2}, w_{i,3}$. We will denote the set of all nodes u_i by U, and the set of all nodes $w_{i,j}$ by W. We also create the evader node $\widehat{v}$, the node v', and the following four sets of nodes:
  1. $A = \{a_1, \ldots, a_m\}$;
  2. $B = \{b_1, \ldots, b_{2k+2m}\}$;
  3. $B' = \{b'_1, \ldots, b'_{k+2m+1}\}$;
  4. $B'' = \{b''_1, \ldots, b''_{2k+m-1}\}$.
- **The set of layers L:** For every $S_i \in S$ we create a layer α_i. We also create an additional layer β.
- **The set of occurrences of nodes in layers V_L:** Node $u_j \in U$ appears in layer α_i if and only if $u_j \in S_i$. Node $w_{i,j} \in U$ appears only in layer α_i. The evader $\widehat{v}$, as well as all nodes in A appear in every layer α_i. Node v', as well as all nodes in B, B', and B'' appear only in layer β.
- **The set of edges E_L:** For every node that appears in multiple layers, we connect all occurrences of this node in a clique. For node u_j in layer α_i we connect it with node a_i. In every layer α_i we connect all nodes in A into a clique. Moreover, we connect every node b_i with node v', and connect every node b'_i with node b_i. Finally, we connect every node b''_i with node b'_i.

Now, consider the following instance of the problem of Multilayer Global Hiding, $(M, \widehat{v}, F, c, d)$, where:

- M is the multilayer network we just constructed;
- $\widehat{v}$ is the evader;
- $F = U \cup W$ is the set of contacts;
- c is the closeness centrality measure;
- $d = 1$.

Next, let us analyze the closeness centrality values of nodes in the network. Notice that every node $w_{i,j}$ appears only in a single layer α_i, hence $\widehat{v}$ has to connect with $w_{i,j}$ in layer α_i. Assume that the evader $\widehat{v}$ has connections with nodes in U in exactly x layers, i.e., $x = |\{\alpha_i \in L : \exists u_j (\widehat{v}^{\alpha_i}, u_j^{\alpha_i}) \in A^*\}|$. We then have:

- $c_{clos}(\widehat{v}) = 3k + 3m + \frac{x}{2} + \frac{m-x}{3} \geq 3k + 3\frac{1}{3}m$ as $\widehat{v}$ is a neighbor of $3k$ nodes in U and $3m$ nodes in W, while for any $a_i \in A$ the distance between a_i and $\widehat{v}$ is 2 if $\widehat{v}$ is connected with any u_j in layer α_i and 3 otherwise;
- $c_{clos}(u_i) \leq 1 + m + \frac{3k-1}{2} + \frac{3m}{2} = 1\frac{1}{2}k + 2\frac{1}{2}m + \frac{1}{2} < c_{clos}(\widehat{v})$ as u_i is a neighbor of $\widehat{v}$ and at most m nodes in A, while the distance to all other nodes is at least 2;
- $c_{clos}(a_i) \leq 3 + m - 1 + \frac{1}{2} + \frac{3k-3}{2} + \frac{3m}{3} = 1\frac{1}{2}k + 2m + 1 < c_{clos}(\widehat{v})$ as a_i is a neighbor of 3 nodes from U and all other $m-1$ nodes in A, while the distance to $\widehat{v}$ and all other nodes in U is 2, and the distance to all nodes is W is at least 3;
- $c_{clos}(w_{i,j}) < c_{clos}(\widehat{v})$ as for any other node v we have $\lambda(w_{i,j}, v) = \lambda(\widehat{v}, v) + 1$, since the shortest paths between $w_{i,j}$ and all other nodes go through $\widehat{v}$;

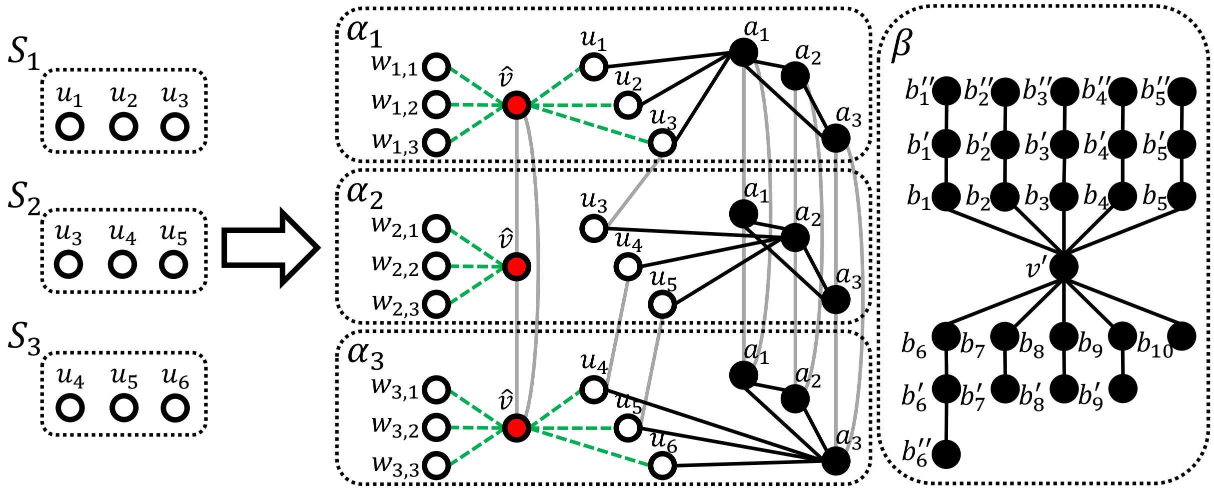

Figure 1: An illustration of the network used in the proof of Theorem 1. Edges connecting occurrences of the same node in different layers are highlighted in grey. The red node represents the evader, while the white nodes represent the contacts. Dashed (green) edges represent the solution to this problem instance.

- $c_{clos}(v') = 2k + 2m + \frac{k+2m+1}{2} + \frac{2k+m-1}{3} = 3k + 3m + \frac{k}{2} + \frac{m-k}{3} + \frac{1}{6}$ as v' is a neighbor of all $2k + 2m$ nodes in B, the distance to all $k + 2m + 1$ nodes in B' is 2, while the distance to all $m - k + 1$ nodes in B'' is 3.

We have shown that all nodes in A, U, and W have smaller closeness centrality than $\widehat{v}$. It is easy to check that v' has greater closeness centrality than all other nodes occurring in layer β. Hence, $\widehat{v}$ is hidden if and only if v' has greater closeness centrality than $\widehat{v}$. This is true when:

$$3k + 3m + \frac{x}{2} + \frac{m-x}{3} < 3k + 3m + \frac{k}{2} + \frac{m-k}{3} + \frac{1}{6}$$

which can be simplified to $x < k + 1$. Since both x and k are in $\mathbb{N}$ this is equivalent to $x \leq k$. Therefore, $\widehat{v}$ is hidden if and only if it has connections with nodes in U in at most k layers.

Now we will show that if there exists a solution to the given instance of the Exact 3-Set Cover problem, then there also exists a solution to the constructed instance of the Multilayer Global Hiding problem. Let S^* be an exact cover of U. In layer α_i we connect $\widehat{v}$ with all nodes from W that occurr in this layer. For every $S_i \in S^*$ we connect $\widehat{v}$ with $u_j \in S_i$ in layer α_i. This way, $\widehat{v}$ becomes connected to all $3k$ contacts from U, since all the sets in S^* are pairwise disjoint.

To complete the proof, we have to show that if there exists a solution A^* to the constructed instance of the Multilayer Global Hiding problem, then there also exists a solution to the given instance of the Exact 3-Set Cover problem. We have shown above that if $\widehat{v}$ is hidden, then it is connected to nodes in U in at most k layers from $\{\alpha_1, \dots, \alpha_m\}$. However, since $\widehat{v}$ must be connected with all $3k$ nodes in U in order for A^* to be a correct solution, then $\{S_i : \exists u_j (\widehat{v}^{\alpha_i}, u_j^{\alpha_i}) \in A^*\}$ is a solution to the given instance of the Exact 3-Set Cover problem. This concludes the proof. □

Theorem 2. *The problem of Multilayer Global Hiding is NP-complete given the betweenness centrality measure.*

Proof. The problem is trivially in NP, since after the addition of a given A^* the betweenness centrality rankings can be computed in polynomial time.

Next, we prove that the problem is NP-hard. To this end, we show a reduction from the NP-complete problem of *Finding k-Clique*. The decision version of this problem is defined by a simple network, $G = (V, E)$, and a constant, $k \in \mathbb{N}$. The goal is to determine whether there exist k nodes in G that form a clique.

Given an instance of the problem of *Finding k-Clique*, defined by k and a simple network $G = (V, E)$, let us construct a multilayer network, $M = (V_L, E_L, V', L)$, as follows (Figure 2 depicts an instance of this network):

- **The set of nodes V':** For every node, $v_i \in V$, we create a node v_i. Additionally, we create the evader node $\widehat{v}$, node a, and the following three sets of nodes:
  1. $B = \{b_1, b_2\}$;
  2. $W = \{w_1, \dots, w_k\}$;
  3. $C = \{c_1, \dots, c_{n+k}\}$.
- **The set of layers L:** We create a layer α, a layer γ, as well as n layers $\beta_1, \dots, \beta_n$.
- **The set of occurrences of nodes in layers V_L:** Node $\widehat{v}$ and node a appear in layer α and all layers $\{\beta_1, \dots, \beta_n\}$. Each node v_i appears in layer α and β_i. Nodes in W appear in all layers $\{\beta_1, \dots, \beta_n\}$. Nodes in B and C appear only in layer γ.
- **The set of edges E_L:** In layer α we create an edge between two nodes $v_i, v_j \in V$ if and only if this edge was present in G. In every layer where a appears we connect it with all occurring nodes from V and W. Finally, we connect every node c_i with both b_1 and b_2.

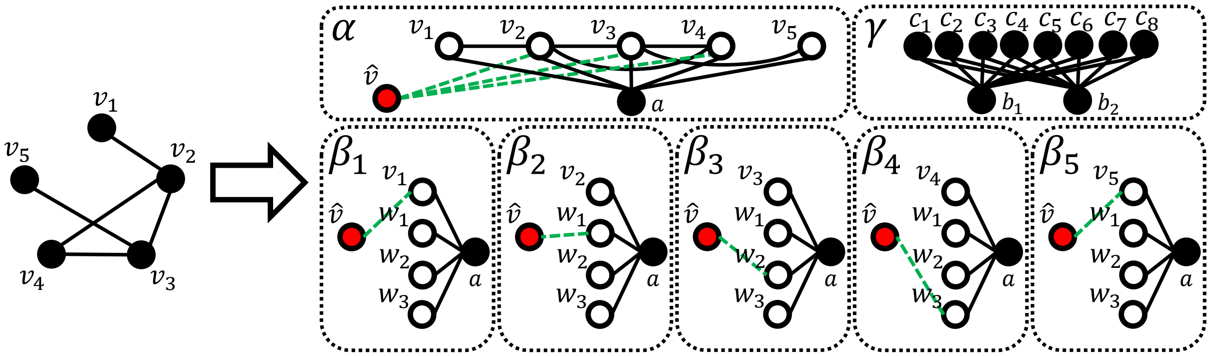

Figure 2: An illustration of the network used in the proof of Theorem 2. The red node represents the evader, while the white nodes represent the contacts. Dashed (green) edges represent the solution to this problem instance.

Now, consider the following instance of the problem of Multilayer Local Hiding, $(M, \widehat{v}, F, c, (d^{\alpha})_{\alpha \in L})$, where:

- M is the multilayer network we just constructed;
- $\widehat{v}$ is the evader;
- $F = V \cup W$ is the set of contacts;
- c is the betweenness centrality measure;
- $d = 2n + 2k + 3$ is the safety margin.

Notice that, since $d = 2n + 2k + 3$, all other nodes must have greater betweenness centrality than the evader in order for $\widehat{v}$ to be hidden. Notice also that the betweenness centrality of every node c_i is $\frac{1}{n+k}$. Moreover, after adding A^* all nodes other than $\widehat{v}$ have non-zero betweenness centrality. If $\widehat{v}$ gets connected to at least two nodes from F that are not connected to each other, then $\widehat{v}$ controls one of at most $n + k - 1$ shortest path between them (other paths can only go through nodes in $V \cup W \cup \{a\}$) and thus the betweenness centrality of $\widehat{v}$ is at least $\frac{1}{n+k-1}$. Therefore, in order to get hidden, $\widehat{v}$ cannot control any shortest paths in the network. This implies that, if $\widehat{v}$ is hidden then all nodes that are connected to $\widehat{v}$ in layer α must form a clique, and also implies that in every layer β, the evader $\widehat{v}$ can be connected to at most one node (otherwise $\widehat{v}$ controls one of the shortest paths between its two neighbors without an edges between them).

Now we will show that if there exists a solution to the given instance of the Finding k-Clique problem, then there also exists a solution to the constructed instance of the Multilayer Global Hiding problem. Let V^* be a group of k nodes forming a clique in G. Let us create A^* by connecting $\widehat{v}$ to nodes from V^* in layer α. Now, we connect every $v_i \in V \setminus V^*$ to $\widehat{v}$ in layer β_i. In the remaining layers from $\{\beta_1, \ldots, \beta_n\}$ (corresponding to elements $v_i \in V^*$) we connect $\widehat{v}$ to all nodes in W. As argued above, for such A^*, the evader $\widehat{v}$ is hidden, hence A^* is a solution to the constructed instance of the Multilayer Global Hiding problem.

To complete the proof we have to show that if there exists a solution A^* to the constructed instance of the Multilayer Global Hiding problem, then there also exists a solution to the given instance of the Finding k-Clique problem. As argued above, in each layer β_i the evader $\widehat{v}$ can be connected to at most one node. Since all k nodes from W appear only in layers from $\{\beta_1, \ldots, \beta_n\}$, the evader $\widehat{v}$ can be connected to at most $n - k$ nodes from V in layers from $\{\beta_1, \ldots, \beta_n\}$. Therefore, $\widehat{v}$ has to have at least k neighbors from V in layer α As shown above, in order for $\widehat{v}$ to be hidden in α, all of its neighbors must form a clique. Hence, the neighbors of $\widehat{v}$ in layer α form a clique in G. This concludes the proof. □

Theorem 3. *The problem of Multilayer Local Hiding is NP-complete given the degree centrality measure.*

Proof. The problem is trivially in NP, since after the addition of a given A^* the degree centrality rankings for all layers can be computed in polynomial time.

Next, we prove that the problem is NP-hard. To this end, we show a reduction from the NP-complete problem of *Exact 3-Set Cover*. The decision version of this problem is defined by a set of subsets $S = \{S_1, \ldots, S_m\}$ of universe $U = \{u_1, \ldots, u_{3k}\}$, such that $\forall_i |S_i| = 3$. The goal is to determine whether there exist k pairwise disjoint elements of S the sum of which equals U.

Given an instance of the problem of *Exact 3-Set Cover*, let us construct a multilayer network, $M = (V_L, E_L, V', L)$, as follows (Figure 3 depicts an instance of this network):

- **The set of nodes V':** For every element, $u_i \in U$, we create a node u_i. We also create $2(m - k)$ nodes $w_1, \ldots, w_{2(m-k)}$. Additionally, we create the evader node $\widehat{v}$ and three nodes a_1, a_2, a_3. We will denote the set of all nodes a_i as A, the set of all nodes u_i as U, and the set of all nodes w_i as W.
- **The set of layers L:** For every $S_i \in S$ we add a layer α_i.
- **The set of occurrences of nodes in layers V_L:** Node $u_j \in U$ appears in layer α_i if and only if $u_j \in S_i$. The evader $\widehat{v}$, as well as all nodes in A and W, appear in all layers.
- **The set of edges E_L:** In every layer we connect every node $u_j \in U$ occurring in this layer to every node in A.

Now, consider the following instance of the problem of Multilayer Local Hiding, $(M, \widehat{v}, F, c, (d^{\alpha})_{\alpha \in L})$, where:

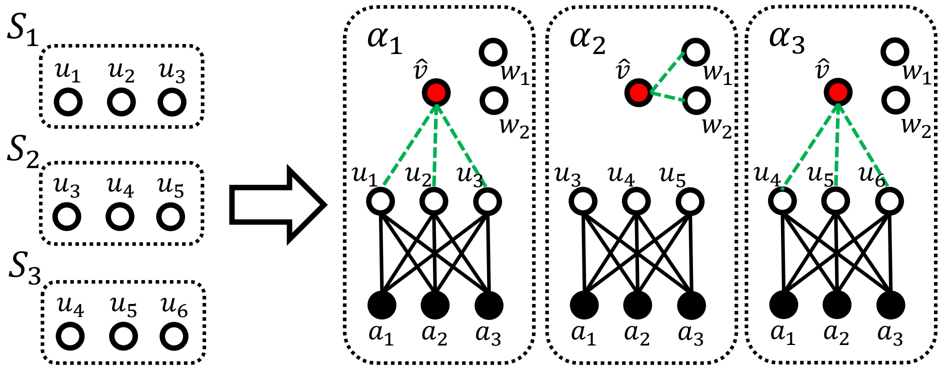

Figure 3: An illustration of the network used in the proof of Theorem 3. The red node represents the evader, while the white nodes represent the contacts. Dashed (green) edges represent the solution to this problem instance.

- M is the multilayer network we just constructed;
- $\widehat{v}$ is the evader;
- $F = U \cup W$ is the set of contacts;
- c is the degree centrality measure;
- $d^{\alpha_i} = 3$ for every $\alpha_i \in L$.

Next, let us consider what are the sets of edges that can be added between the evader $\widehat{v}$ and the contacts F in each layer, so that the evader is hidden. In every layer α_i the nodes in A as well as the nodes $u_j \in S_i$ have degree 3, while all other nodes have degree 0. We can connect $\widehat{v}$ to any two or less contacts and $\widehat{v}$ will still be hidden. If we connect the evader to three contacts, they have to be nodes in S_i (as these are the only nodes that potentially can have degree greater than 3). We cannot connect $\widehat{v}$ to more than three contacts and still have $\widehat{v}$ hidden.

Now we will show that if there exists a solution to the given instance of the Exact 3-Set Cover problem, then there also exists a solution to the constructed instance of the Multilayer Local Hiding problem. Let S^* be an exact cover of U. For every $S_i \in S^*$ we connect $\widehat{v}$ to every $u_j \in S_i$ in layer α_i. This way, $\widehat{v}$ becomes connected to all $3k$ contacts from U, since all the sets in S^* are pairwise disjoint. For every $S_i \notin S^*$ we connect $\widehat{v}$ to two nodes from W in layer α_i (since there are $m-k$ such layers, we can connect $\widehat{v}$ to all $2(m-k)$ contacts from W this way).

To complete the proof we have to show that if there exists a solution to the constructed instance of the Multilayer Local Hiding problem, then there also exists a solution to the given instance of the Exact 3-Set Cover problem. Let x be the number of layers from $\{\alpha_1, \dots, \alpha_m\}$ in which $\widehat{v}$ has at most two neighbors, and let $m-x$ be the number of layers from $\{\alpha_1, \dots, \alpha_m\}$ where $\widehat{v}$ has exactly three neighbors. Since $\widehat{v}$ has to be connected to all $3k + 2(m-k)$ contacts, we have $2x + 3(m-x) \geq 3k + 2(m-k)$, which gives us $x \leq m-k$. However, since $\widehat{v}$ can connect to nodes from W in layer α_i if and only if it connects to at most two nodes in α_i, we also have $2x \geq 2(m-k)$. Hence, we have $x = m-k$, i.e., $\widehat{v}$ is connected with all nodes from W in $m-k$ layers from $\{\alpha_1, \dots, \alpha_m\}$. Therefore, in the remaining k layers from $\{\alpha_1, \dots, \alpha_m\}$, the evader $\widehat{v}$ has to connect to all $3k$ nodes from U. Since the evader cannot connect to more than three nodes in any layer α_i, all these sets of neighbors from U have to be disjoint, thus forming the solution to the given instance of the Exact 3-Set Cover problem. This concludes the proof. □

Theorem 4. *The problem of Multilayer Local Hiding is NP-complete given the closeness centrality measure.*

Proof. The problem is trivially in NP, since after the addition of a given A^* the closeness centrality rankings for all layers can be computed in polynomial time.

Next, we prove that the problem is NP-hard. To this end, we show a reduction from the NP-complete problem of *Exact 3-Set Cover*. The decision version of this problem is defined by a set of subsets $S = \{S_1, \dots, S_m\}$ of universe $U = \{u_1, \dots, u_{3k}\}$, such that $\forall_i |S_i| = 3$. The goal is to determine whether there exist k pairwise disjoint elements of S the sum of which equals U.

Given an instance of the problem of *Exact 3-Set Cover*, let us construct a multilayer network, $M = (V_L, E_L, V', L)$, as follows (Figure 4 depicts an instance of this network):

- **The set of nodes V':** For every $u_i \in U$ we create a node u_i. In addition, we create the nodes $w_1, \dots, w_{2(m-k)}$ and $a_1, \dots, a_{2(m-k)}$. Finally, we create the evader node $\widehat{v}$ and 5 nodes $c_1, \dots, c_5$. We will denote the set of all nodes a_i by A, the set of all nodes c_i by C, the set of all nodes u_i by U, and the set of all nodes w_i by W.
- **The set of layers L:** For every $S_i \in S$ we add a layer α_i.
- **The set of occurrences of nodes in layers V_L:** Node $u_j \in U$ appears in layer α_i if and only if $u_j \in S_i$. The evader $\widehat{v}$, as well as all nodes in A, C, and W, appear in all layers.
- **The set of edges E_L:** In all layers we connect every node w_i with the node a_i, and we create edges $(c_1, c_2), (c_1, c_3), (c_1, c_4), (c_4, c_5)$.

Now, consider the following instance of the problem of Multilayer Local Hiding, $(M, \widehat{v}, F, c, (d^\alpha)_{\alpha \in L})$, where:

- M is the multilayer network we just constructed;
- $\widehat{v}$ is the evader;

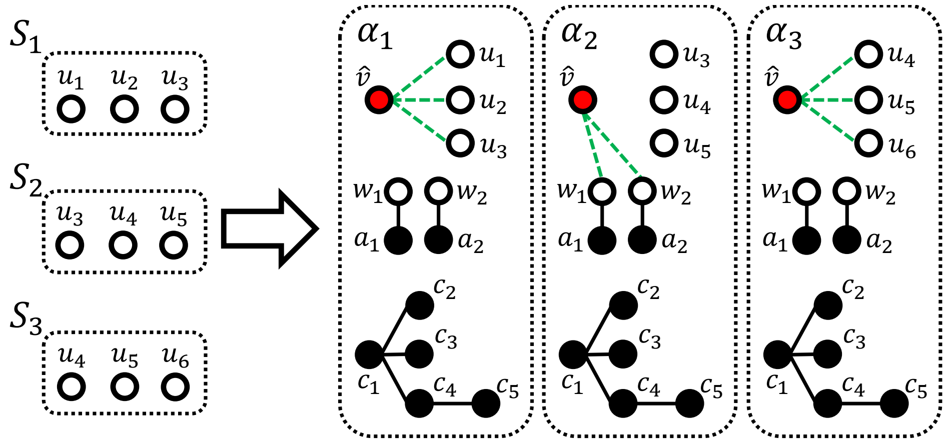

Figure 4: An illustration of the network used in the proof of Theorem 4. The red node represents the evader, while white the nodes represent the contacts. Dashed (green) edges represent the solution to this problem instance.

- $F = U \cup W$ is the set of contacts;
- c is the closeness centrality measure;
- $d^{\alpha_i} = 1$ for every $\alpha_i \in L$.

Next, let us consider what are the sets of edges that can be added between the evader $\widehat{v}$ and the contacts F in each layer, so that the evader is hidden. Notice that closeness centrality of the node c_1 is $3\frac{1}{2}$ and it is not affected by the edges added to $\widehat{v}$. Assume that we connect node $\widehat{v}$ with x nodes from U and y nodes from W. We then have the following (for easier comparison we express the centrality values as fractions with the common denominator 6):

- $c_{clos}(\widehat{v}) = x + \frac{3y}{2} = \frac{6x+9y}{6}$;
- $c_{clos}(w_i) = \frac{x}{2} + \frac{5y}{6} + \frac{7}{6} = \frac{3x+5y+7}{6}$ if $w_i \in N(\widehat{v})$;
- $c_{clos}(c_1) = \frac{7}{2} = \frac{21}{6}$;

No other node can have greater closeness centrality than $\widehat{v}$. We can connect $\widehat{v}$ with at most two of any of the contacts, as node c_1 will still have greater closeness centrality. If we want to connect $\widehat{v}$ with three contacts, these contacts have to be nodes from U. If $x + y = 3$ and $y > 0$, or if $x + y > 3$, then the closeness centrality of $\widehat{v}$ is the highest in the network, meaning that $\widehat{v}$ is not hidden.

Now we will show that if there exists a solution to the given instance of the Exact 3-Set Cover problem, then there also exists a solution to the constructed instance of the Multilayer Local Hiding problem. Let S^* be an exact cover of U. For every $S_i \in S^*$ we connect $\widehat{v}$ to every $u_j \in S_i$ in layer α_i. This way, $\widehat{v}$ becomes connected to all $3k$ contacts from U, since all the sets in S^* are pairwise disjoint. For every $S_i \notin S^*$ we connect $\widehat{v}$ to two nodes from W in layer α_i (since there are $m - k$ such layers, we can connect $\widehat{v}$ to all $2(m - k)$ contacts from W this way).

To complete the proof, we have to show that if there exists a solution A^* to the constructed instance of the Multilayer Local Hiding problem, then there also exists a solution to the given instance of the Exact 3-Set Cover problem. Let z be the number of layers from $\{\alpha_1, \ldots, \alpha_m\}$ where $\widehat{v}$ has at most two neighbors, and let $z - x$ be the number of layers from $\{\alpha_1, \ldots, \alpha_m\}$ where $\widehat{v}$ has exactly three neighbors. Since we have to connect $\widehat{v}$ to all $3k + 2(m - k)$ contacts, we have $2z + 3(m - z) \geq 3k + 2(m - k)$, which gives us $z \leq m - k$. However, since $\widehat{v}$ can connect to nodes from W in layer α_i if and only if it connects to at most two nodes in α_i, we also have $2z \geq 2(m - k)$. Hence, we have $z = m-k$, i.e., $\widehat{v}$ connects to all nodes from W in $m-k$ layers from $\{\alpha_1, \ldots, \alpha_m\}$. Therefore, in the remaining k layers from $\{\alpha_1, \ldots, \alpha_m\}$, the evader $\widehat{v}$ has to connect with all $3k$ nodes from U. Since the evader cannot connect to more than three nodes in any layer α_i, all these sets of neighbors from U have to be disjoint, thus forming a solution to the given instance of the Exact 3-Set Cover problem. This concludes the proof. □

Proofs of Approximation Results

Theorem 6. *The Maximum Multilayer Global Hiding problem can be solved in polynomial time.*

Proof. For a given $k \in \mathbb{N}$ it is possible to connect the evader with k contacts if and only if $\min(k, |\{v \in F : |N(v)| = k + |N(\widehat{v})|\}|) + |\{v \in V : |N(v)| > k + |N(\widehat{v})|\}| \geq d$. It is because the only nodes that count towards satisfying the safety margin are those that already have degree greater than $k + |N(\widehat{v})|$, or the contacts that have degree k and their degree will be increased to $k + |N(\widehat{v})| + 1$ when they are connected with the evader (notice that since we are adding k edges, there can be at most k such contacts). □

Theorem 7. *Maximum Multilayer Global Hiding problem given the closeness centrality cannot be approximated within $|F|^{1-\epsilon}$ for any $\epsilon > 0$, unless P=NP.*

Proof. In order to prove the theorem, we will use the result by Zuckerman (2006) that the *Maximum Independent Set* problem cannot be approximated within $|V|^{1-\epsilon}$ for any $\epsilon > 0$, unless $P = NP$ (notice that the Maximum Independent Set problem is equivalent to the Maximum Clique

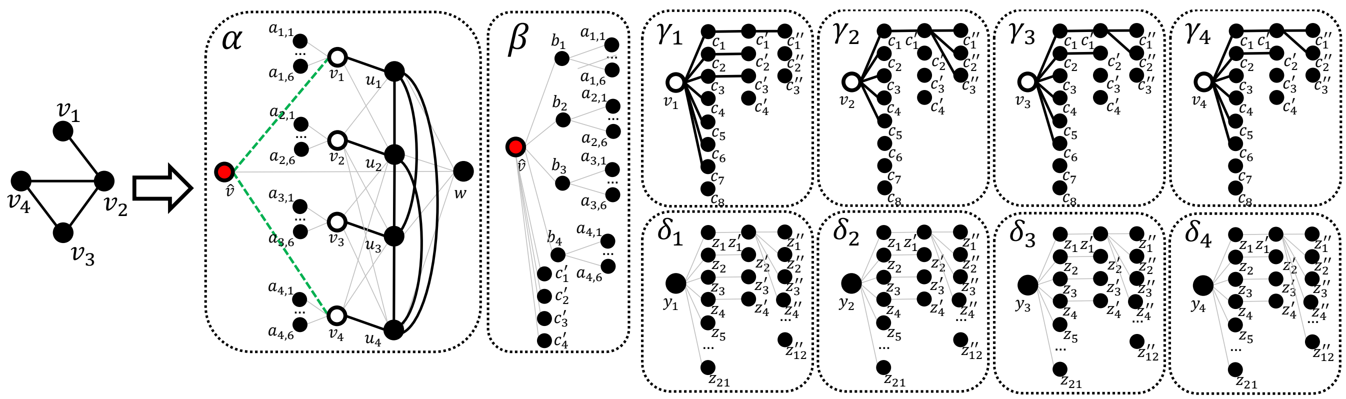

Figure 5: An illustration of the network used in the proof of Theorem 7. The red node represents the evader, while the white nodes represent the contacts. Dashed (green) edges represent the optimal solution to this problem instance.

problem on a complementary network). The Maximum Independent Set problem is defined by a simple network, $G = (V, E)$. The goal is to identify the maximum (in terms of size) group of nodes in G that are independent, i.e., they do not induce any edges.

First, we will show a function $f(G)$ that based on an instance of the problem of Maximum Independent Set, defined by a simple network $G = (V, E)$, constructs an instance of the Maximum Multilayer Global Hiding. In what follows we will assume that $n \geq 4$ (the problem can be easily solve in constant time for $n < 4$).

Let a multilayer network, $M = (V_L, E_L, V', L)$, be defined as follows (Figure 5 depicts an instance of this network):

- **The set of nodes** V'**:** For every node, $v_i \in V$, we create a node v_i, a node u_i, a node y_i, six nodes $a_{i,1}, \ldots, a_{i,6}$, and a node b_i. We will denote the set of all nodes v_i by V, the set of all nodes u_i by U, the set of all nodes y_i by Y, the set of all nodes $a_{i,j}$ by A, and the set of all nodes b_i by B. Additionally, we create the evader node $\widehat{v}$, a node w and six sets of nodes $C = \{c_1, \ldots, c_{3n-4}\}$, $C' = \{c'_1, \ldots, c'_n\}$, and $C'' = \{c''_1, \ldots, c''_{n-1}\}$, $Z = \{z_1, \ldots, z_{5n+1}\}$, $Z' = \{z'_1, \ldots, z'_n\}$, and $Z'' = \{z''_1, \ldots, z''_{3n}\}$.
- **The set of layers** L**:** We create layers α, β, n layers $\gamma_1, \ldots, \gamma_n$, and n layers $\delta_1, \ldots, \delta_n$.
- **The set of occurrences of nodes in layers** V_L**:** Layer α contains occurrences of nodes $\{\widehat{v}, w\} \cup A \cup V \cup U$. Layer β contains occurrences of nodes $\{\widehat{v}\} \cup A \cup B \cup C'$. A given layer γ_i contains occurrences of nodes $\{v_i\} \cup C \cup C' \cup C''$. A given layer δ_i contains occurrences of nodes $\{y_i\} \cup Z \cup Z' \cup Z''$.
- **The set of edges** E_L**:**
  - **In layer** α**:** for every pair of nodes $v_i \in V, u_j \in U$ we create an edge (v_i, u_j) if and only if the edge (v_i, v_j) was present in G. For every node u_i we create edges (u_i, v_i) and (u_i, w). For every node $a_{i,j}$ we create an edge $(a_{i,j}, v_i)$. We also create an edge between every pair of nodes $u_i, u_j \in U$. Finally, we create an edge $(\widehat{v}, w)$.
  - **In layer** β**:** for every node $a_{i,j}$ we create an edge $(a_{i,j}, b_i)$. We also create an edge $(\widehat{v}, b_i)$ for every node b_i and an edge $(\widehat{v}, c'_i)$ for every node c'_i.
  - **In a given layer** γ_i**:** for every node c_j we create an edge (v_i, c_j) if and only if $j \leq 3n - 4 - |N_G(v_i)|$ (i.e., we connect v_i with $3n - 4 - |N_G(v_i)|$ first nodes from C). For every node c'_j we create an edge (c'_j, c_j) if and only if $j \leq n - |N_G(v_i)|$ (i.e., we connect $n - |N_G(v_i)|$ first nodes from C' with their C counterparts). For every node c''_j we create an edge (c''_j, c'_1) if and only if $j \leq |N_G(v_i)|$ (i.e., we connect $|N_G(v_i)|$ first nodes from C'' with the node c'_1).
  - **In a given layer** δ_i**:** for every node z_j we create an edge (y_i, z_j). For every node z'_j we create an edge (z'_j, z_j). For every node z''_j we create an edge (z''_j, z'_1) if and only if $j \leq n + 2i$ (i.e., we connect $n + 2i$ first nodes from Z'' with the node z'_1).

To complete the constructed instance of the problem let:

- $\widehat{v}$ be the evader;
- $F = V$ be the set of contacts;
- c be the closeness centrality measure;
- $d = n$ be the safety margin.

Hence, the formula of the function f is $f(G) = (M, \widehat{v}, F, c, d)$. Let A^* be the solution to the constructed instance of the Maximum Multilayer Global Hiding problem. The function g computing corresponding solution to the instance G of the Maximum Independent Set problem is now $g(A^*) = \{v_i \in V : (\widehat{v}^\alpha, v_i^\alpha) \in A^*\}$, i.e., the nodes forming the independent set are the contacts that the evader is connected to.

Now, we will show that $g(A^*)$ is indeed a correct solution to G, i.e., that the nodes form an independent set. Let $x_i = |N_G(v_i) \cap g(A^*)|$, i.e., the number of neighbours of v_i in G connected to $\widehat{v}$. Let us compute the closeness centrality of all nodes in the network and compare it with the closeness centrality of the evader:

- $c_{clos}(M,\widehat{v}) = 2n+1+|A^*|+\frac{6n+n}{2}+\frac{n-|A^*|}{3} = 5\frac{5}{6}n+\frac{2}{3}|A^*|+1$;
- if $v_i \in g(A^*)$ then $c_{clos}(M,v_i) = 3n+4+\frac{2n-|N_G(v_i)|+|A^*|-x_i-1}{2}+\frac{n+6|N_G(v_i)|+5|A^*|-5x_i-6}{3}+\frac{6n-6|N_G(v_i)|-6|A^*|+6x_i}{4} = 5\frac{5}{6}n+\frac{2}{3}|A^*|+1\frac{1}{2}-\frac{2}{3}x_i$;
- if $v_i \notin g(A^*)$ then $c_{clos}(M,v_i) = 3n+3+\frac{2n-|N_G(v_i)|}{2}+\frac{n+6|N_G(v_i)|}{3}+\frac{6n-6|N_G(v_i)|-6}{4} = 5\frac{5}{6}n+1\frac{1}{2}$, hence $c_{clos}(M,v_i) < c_{clos}(M,\widehat{v})$ for $|A^*|>0$;
- $c_{clos}(M,w) = n+1+\frac{n}{2}+\frac{6n}{3} = 3\frac{1}{2}n+1 < c_{clos}(M,\widehat{v})$;
- $c_{clos}(M,u_i) \leq 2n+\frac{6n}{2} = 5n < c_{clos}(M,\widehat{v})$;
- $c_{clos}(M,a_{i,j}) \leq 2+\frac{n+7}{2}+\frac{2n-2}{3}+\frac{6(n-1)}{4} = 2\frac{2}{3}n+3\frac{1}{3} < c_{clos}(M,\widehat{v})$;
- $c_{clos}(M,b_i) = 7+\frac{n-1}{2}+\frac{(n-1)6}{3} = 2\frac{1}{2}n+4\frac{1}{2} < c_{clos}(M,\widehat{v})$;
- $c_{clos}(M,c_i) \leq n+1+\frac{4n-6}{2}+\frac{n-1}{3} = 3\frac{1}{3}n-2\frac{1}{3} < c_{clos}(M,\widehat{v})$;
- $c_{clos}(M,c'_i) \leq n+1+\frac{3n-1}{2}+\frac{8n-4}{3}+\frac{n-1}{4} = 5\frac{5}{12}n-1\frac{1}{12} < c_{clos}(M,\widehat{v})$;
- $c_{clos}(M,c''_i) \leq 1+\frac{n-1}{2}+\frac{n}{3}+\frac{3n-5}{4}+\frac{n-1}{5} = 1\frac{47}{60}n-1\frac{19}{20} < c_{clos}(M,\widehat{v})$;
- $c_{clos}(M,y_i) = 5n+1+\frac{n}{2}+\frac{n+2i}{3} = 5\frac{5}{6}n+1+\frac{2}{3}i$;
- $c_{clos}(M,z_i) \leq n+1+\frac{8n}{2}+\frac{n-1}{3} = 5\frac{1}{3}n+\frac{2}{3} < c_{clos}(M,\widehat{v})$;
- $c_{clos}(M,z'_i) \leq 3n+1+\frac{n}{2}+\frac{5n}{3}+\frac{n-1}{4} = 5\frac{5}{12}n+\frac{3}{4} < c_{clos}(M,\widehat{v})$;
- $c_{clos}(M,z''_i) = 1+\frac{3n}{2}+\frac{n}{3}+\frac{5n}{4}+\frac{n-1}{5} = 3\frac{17}{60}n+\frac{4}{5} < c_{clos}(M,\widehat{v})$.

Notice that the only nodes that can have greater closeness centrality than the evader are the nodes in V that the evader is connected to and the nodes in Y. For a given node y_i it has a greater closeness centrality score than the evader if and only if the evader is connected to less than i nodes from V (i.e., when $|A^*|<i$), as:

$$c_{clos}(M,\widehat{v}) - c_{clos}(M,y_i) = \frac{2}{3}(|A^*|-i).$$

Hence, since the safety margin is $d=n$ and the only other nodes that can have greater closeness centrality than the evader are the nodes in V that the evader is connected to. Hence, every node in V that the evader connects to must have greater closeness centrality than the evader, in order for the safety margin to be maintained (there are exactly n nodes in Y, and every edge additional edge in A^* causes the evader to get greater closeness centrality than one of the nodes in Y). However, node $v_i \in V$ that the evader is connected to has greater closeness centrality than the evader if and only if $x_i = 0$, i.e., no neighbors of v_i are connected to $\widehat{v}$, as:

$$c_{clos}(M,\widehat{v}) - c_{clos}(M,v_i) = \frac{2}{3}x_i - \frac{1}{2}.$$

This implies that, if $\widehat{v}$ is hidden then all nodes from V that are connected to $\widehat{v}$ must form an independent set.

Therefore, the optimal solution to the constructed instance of the Maximum Multilayer Global Hiding problem is returning nodes from V forming in G an independent set of the maximum size. Hence, the optimal solution corresponds to the optimal solution to the given instance of the Maximum Independent Set problem.

Now, assume that there exists an approximation algorithm for the Maximum Multilayer Global Hiding problem with ratio $|F|^{1-\epsilon}$ for some $\epsilon > 0$. Let us use this algorithm to solve the constructed instance $f(G)$, acquiring solution A^*. and consider solution $g(A^*)$ to the given instance of the Maximum Clique problem. Since the size of the optimal solution is the same for both instances, we obtained an approximation algorithm that solves Maximum Independent Set problem to within $|V|^{1-\epsilon}$ for $\epsilon > 0$. However, Zuckerman (2006) shown that the Maximum Independent Set problem cannot be approximated within $|V|^{1-\epsilon}$ for any $\epsilon > 0$, unless $P = NP$. Therefore, such approximation algorithm for the Maximum Multilayer Global Hiding problem cannot exist, unless $P = NP$. This concludes the proof. □

Theorem 8. *Both Maximum Multilayer Global Hiding and Maximum Multilayer Local Hiding problems given the betweenness centrality cannot be approximated within $|F|^{1-\epsilon}$ for any $\epsilon > 0$, unless P=NP.*

Proof. In order to prove the theorem, we will use the result by Zuckerman (2006) that the *Maximum Clique* problem cannot be approximated within $|V|^{1-\epsilon}$ for any $\epsilon > 0$, unless $P = NP$. The Maximum Clique problem is defined by a simple network, $G = (V, E)$. The goal is to identify the maximum (in terms of size) group of nodes in G that form a clique.

First, we will show a function $f(G)$ that based on an instance of the problem of Maximum Clique, defined by a simple network $G = (V, E)$, constructs either an instance of the Maximum Multilayer Global Hiding or an instance of the Maximum Multilayer Local Hiding.

Let a multilayer network, $M = (V_L, E_L, V', L)$, be defined as follows (Figure 6 depicts an instance of this network):

- **The set of nodes V':** For every node, $v_i \in V$, we create a node v_i and a node a_i. Additionally, we create the evader node $\widehat{v}$ and the set of nodes $C = \{c_1, \ldots, c_{n+2}\}$.
- **The set of layers L:** We create only a single layer α.
- **The set of occurrences of nodes in layers V_L:** All nodes occur in layer α.
- **The set of edges E_L:** In layer α we create an edge between two nodes $v_i, v_j \in V$ if and only if this edge was present in G. We also create an edge (v_i, a_i) for every v_i, and an edge between every pair a_i, a_{i+1}. Finally, for every node $c_i \in C : i \leq n$, we create edges (c_i, c_{n+1}) and (c_i, c_{n+2}).

To complete the constructed instance of the problem let:

- $\widehat{v}$ be the evader;

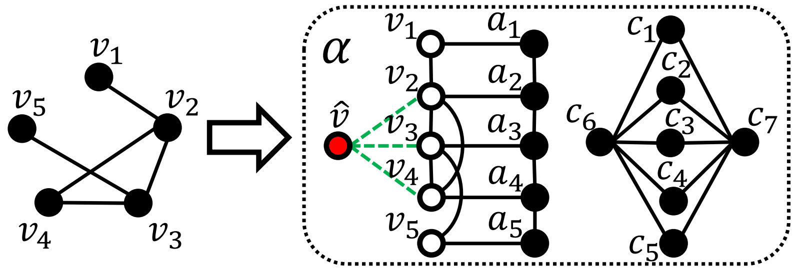

Figure 6: An illustration of the network used in the proof of Theorem 8. The red node represents the evader, while the white nodes represent the contacts. Dashed (green) edges represent the optimal solution to this problem instance.

- $F = V$ be the set of contacts;
- c be the betweenness centrality measure;
- $d = 3n + 2$ be the safety margin in the global version;
- $d^\alpha = 3n + 2$ be the safety margin in the local version.

Hence, the formula of the function f is $f(G) = (M, \widehat{v}, F, c, d)$ for the global version of the problem and $f(G) = (M, \widehat{v}, F, c, (d^\alpha)_{\alpha \in L})$ for the local version of the problem. Notice that since network M has only one layer, both problems are equivalent. In the following we will focus on the global version of the problem.

Let A^* be the solution to the constructed instance of the Maximum Multilayer Global Hiding problem. The function g computing corresponding solution to the instance G of the Maximum Clique problem is now $g(A^*) = \{v \in V : (\widehat{v}, v) \in A^*\}$, i.e., the nodes forming the clique are the contacts that the evader is connected to.

Now, we will show that $g(A^*)$ is indeed a correct solution to G, i.e., that the nodes form a clique. Notice that, since $d = 2n + 2k + 3$, all other nodes must have greater betweenness centrality than the evader in order for $\widehat{v}$ to be hidden. Notice also that the betweenness centrality of every node c_i for $i \leq n$ is $\frac{1}{n}$. Moreover, after adding A^* all nodes other than $\widehat{v}$ have non-zero betweenness centrality. If $\widehat{v}$ gets connected to at least two nodes from F that are not connected to each other, then $\widehat{v}$ controls one of at most $n - 1$ shortest path between them (other paths can only go through nodes in V) and thus the betweenness centrality of $\widehat{v}$ is at least $\frac{1}{n-1}$. Therefore, in order to get hidden, $\widehat{v}$ cannot control any shortest paths in the network. This implies that, if $\widehat{v}$ is hidden then all nodes that are connected to $\widehat{v}$ must form a clique.

Therefore, the optimal solution to the constructed instance of the Maximum Multilayer Global Hiding problem is returning nodes from V forming a clique of maximum size. Since the structure of connections between the nodes V is the same as in the network G, the optimal solution corresponds to the optimal solution to the given instance of the Maximum Clique problem.

Now, assume that there exists an approximation algorithm for the Maximum Multilayer Global Hiding problem with ratio $|F|^{1-\epsilon}$ for some $\epsilon > 0$. Let us use this algorithm to solve the constructed instance $f(G)$, acquiring solution A^*. and consider solution $g(A^*)$ to the given instance of the Maximum Clique problem. Since the size of the optimal solution is the same for both instances, we obtained an approximation algorithm that solves Maximum Clique problem to within $|V|^{1-\epsilon}$ for $\epsilon > 0$. However, Zuckerman (2006) shown that the Maximum Clique problem cannot be approximated within $|V|^{1-\epsilon}$ for any $\epsilon > 0$, unless $P = NP$. Therefore, such approximation algorithm for the Maximum Multilayer Global Hiding problem cannot exist, unless $P = NP$. This concludes the proof. □

Empirical Analysis

In this section we describe network datasets used in our simulations in the main paper, and we present simulation results for real-life-networks.

Network Datasets

In our experiments we use both randomly generated and real-life multilayer networks. As for the randomly generated ones, we use the following standard network generation models:

- *Random graphs*, generated using the Erdős-Rényi model (Erdős and Rényi 1959). We denote by ER(n, k) a network with n nodes with expected degree of k.
- *Small-world* networks, generated using the Watts-Strogatz model (Watts and Strogatz 1998). We denote by WS(n, k) a network with n nodes, an average degree of k, and a rewiring probability of $\frac{1}{4}$.
- *Scale-free* networks, generated using the Barabási-Albert model (Barabási and Albert 1999). We denote by BA(n, k) a network with n nodes, with k edges added with each new node. The size of the initial clique is k.

To construct a multilayer network with n nodes and l layers using a network generation model $X \in \{ER, WS, BA\}$, we perform the following steps:

1. We create the set of node occurrences such that, for every node v and every layer α, the node v occurs in α with probability p_O. If, at the end of this process, v does not occur on any layer, then we create one occurrence of v in a layer chosen uniformly at random; this ensures that every node occurs in at least one layer.
2. For every layer, α, we generate a network $X(|V^\alpha|, k)$ whose set of nodes consists of all the nodes that occur in α from the previous step.

Table 1: Characteristics of the considered datasets.

Dataset	$\lvert V\rvert$	$\lvert L\rvert$	$\lvert V_L\rvert$	$\lvert E_L\rvert$
FF-TW-YT	574	3	1722	5681
Provisional IRA	937	5	1570	3398
Lazega law firm	71	3	211	2051
CS Aarhus	61	5	224	948

3. For every two occurrences of the same node, we create an inter-layer edge between them with probability p_C. This results in a network with diagonal couplings.

The network created using these steps will be denoted by $X^l(n,k)$. In our experiments, we use networks where $p_O = p_C = \frac{1}{2}$ and where the number of layers is $l = 3$ (the choice of 3 layers was inspired by the work of Gera et al. (2017)). Additionally, we use the following real-life networks:

- *FF-TW-YT* dataset (Celli et al. 2010) consisting of a network of connections between 1722 individuals that have an account in each of the following social media site: Friendfeed, Twitter, and YouTube; each of these sites is represented by a separate layer in the network.
- *Provisional IRA* dataset (Gill et al. 2014) consisting of the network of connections between members of the Provisional Irish Republic Army in the period between 1970 and 1998. The network consists of 937 nodes and 5 layers, where layers correspond to contacts between organization members in different time periods.
- *Lazega law firm* dataset (Lazega 2001) consisting of connections between attorneys working for a US corporate law firm. The network consists of 71 nodes and 3 layers, where each layer corresponds to a different type of relationship, i.e., friendship, professional cooperation and mentorship.
- *CS Aarhus* dataset (Magnani, Micenkova, and Rossi 2013) consisting of connections between the employees of the Computer Science department at Aarhus. The network consists of 61 nodes and 5 layers, where each layer corresponds to a different type of relationship between employees, e.g., Facebook friendships, co-authorship of papers, *etc.*

Table 1 presents detailed characteristics of the considered datasets.

Simulation Results of Real-Life Networks

Figure 7 presents results of our simulations for real-life networks. The heuristics and the setting of our simulations are described in the main paper.

References

Barabási, A.-L., and Albert, R. 1999. Emergence of scaling in random networks. *science* 286(5439):509–512.

Celli, F.; Lascio, F. M. L. D.; Magnani, M.; Pacelli, B.; and Rossi, L. 2010. Social Network Data and Practices: the case of Friendfeed. In *International Conference on Social Computing, Behavioral Modeling and Prediction*, Lecture Notes in Computer Science. Springer Berlin Heidelberg.

Erdős, P., and Rényi, A. 1959. On random graphs i. *Publ. Math. Debrecen* 6:290–297.

Gera, R.; Miller, R.; Saxena, A.; MirandaLopez, M.; and Warnke, S. 2017. Three is the answer: combining relationships to analyze multilayered terrorist networks. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, 868–875. ACM.

Gill, P.; Lee, J.; Rethemeyer, K. R.; Horgan, J.; and Asal, V. 2014. Lethal connections: The determinants of network connections in the provisional irish republican army, 1970–1998. *International Interactions* 40(1):52–78.

Lazega, E. 2001. *The collegial phenomenon: The social mechanisms of cooperation among peers in a corporate law partnership*. Oxford University Press on Demand.

Magnani, M.; Micenkova, B.; and Rossi, L. 2013. Combinatorial analysis of multiple networks. *arXiv preprint arXiv:1303.4986*.

Watts, D. J., and Strogatz, S. H. 1998. Collective dynamics of small-world networks. *nature* 393(6684):440–442.

Zuckerman, D. 2006. Linear degree extractors and the inapproximability of max clique and chromatic number. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, 681–690. ACM.

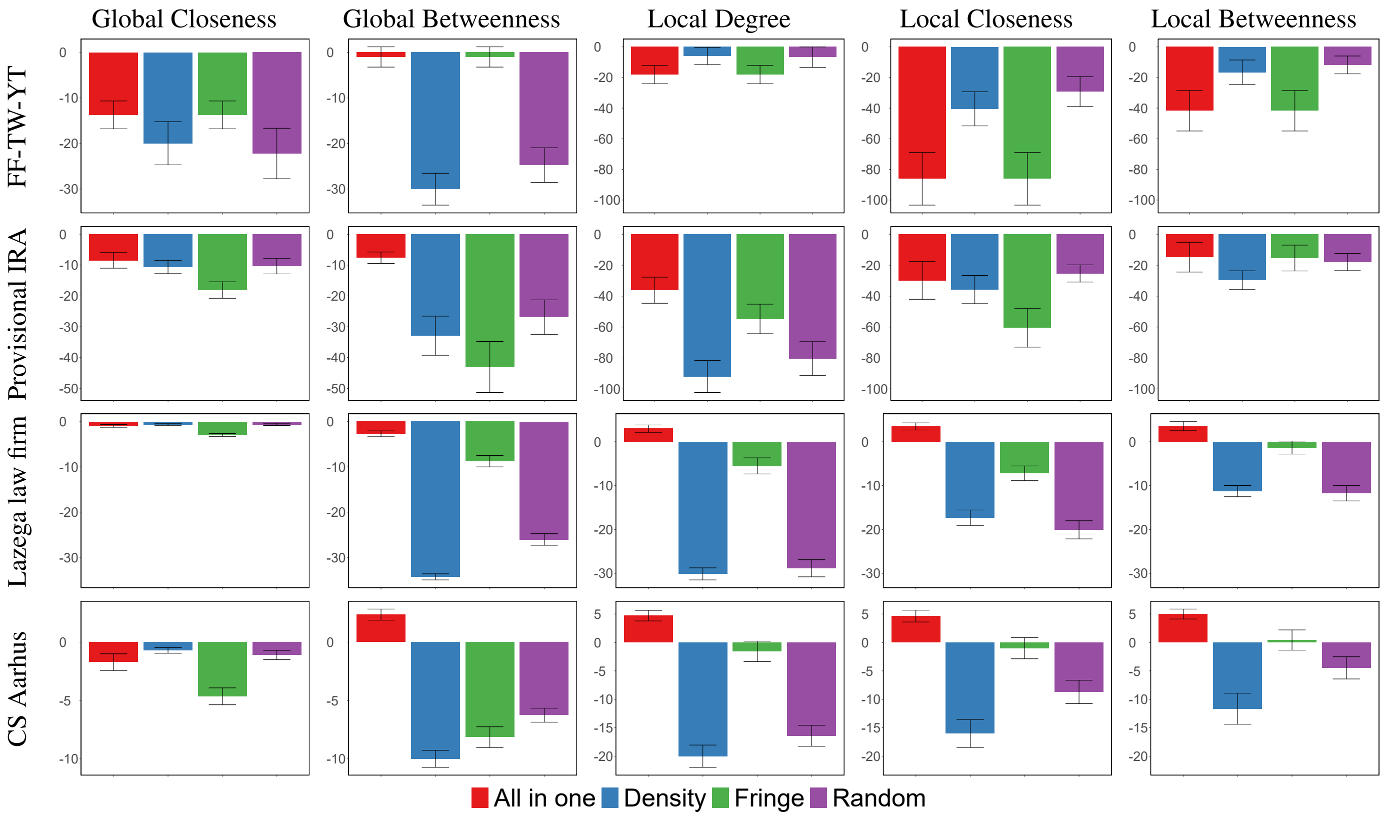

Figure 7: Given different centrality measures and different networks, the figure depicts the average change in centrality ranking of 10 different evaders as a result of execution of different hiding heuristics. For the randomly generate networks the experiment is repeated 100 times, with a new network generated each time. Error bars represent 95% confidence intervals.