4. PROPERTIES

In this section, we discuss some key properties of Attachment Centrality, and show that it is closely related to the notion of *minimal paths* between nodes. The following theorem constitutes the cornerstone of our analysis.

THEOREM 3. *Adding an edge $\{v, u\}$ to a graph G affects only the Attachment Centrality of nodes lying on a minimal path between v and u.*

PROOF. Recall that the value of a coalition, S, in the game (V, f^*_G) depends solely on the number of the nodes in S and the number of connected components in S (see Equation (14)). Now, let $G = (V, E)$ be an arbitrary, incomplete graph, and let $v, u \in V$ be two nodes such that $\{v, u\} \notin E$. Finally, let G' be the graph that results from adding $\{v, u\}$ to G, i.e., $G' = (V, E \cup \{\{v, u\}\})$. Next, we analyse how the marginal contribution of some node $w \in V \setminus \{v, u\}$ to a coalition S differs between G and G':

- Suppose that $\{v, u\} \not\subseteq S$, or that $\{v, u\} \subseteq S$ and both v and u belong to the same component in $G[S]$. Either way, edge $\{v, u\}$ does not affect the value of S and $S \cup \{w\}$: $f^*_G(S) = f^*_{G'}(S)$ and $f^*_G(S \cup \{w\}) = f^*_{G'}(S \cup \{w\})$. So, the marginal contribution of w in G is the same as in G';
- On the other hand, suppose that $\{v, u\} \subseteq S$ and that v and u belong to different component in $G[S]$, namely C_v and C_u. In this case, we have two possibilities: Either w is connected to both C_v and C_u or not. If it is not connected to both, then:

$$f^*_{G'}(S \cup \{w\}) - f^*_{G'}(S) = (f^*_G(S \cup \{w\}) + 2) - (f^*_G(S) + 2) = f^*_G(S \cup \{w\}) - f^*_G(S),$$

  meaning that the marginal contribution of w in G is the same as in G'. In contrast, if w was connected to both C_v and C_u, then w would unite the two in $G[S \cup \{w\}]$ but not in $G'[S \cup \{w\}]$, because in G' the nodes in S are sufficient to unite the two components; they no longer need w to do that.

To summarize, we have shown that the marginal contribution of w in G can be different than in G' only when all of the following three conditions are met: (1) $\{v, u\} \subseteq S$; (2) v and u belong to different components in $G[S]$; (3) w is connected to both C_v and C_u. Importantly, however, if those three conditions are met, then w must be on some minimal path between v and u. To see why this is case, consider a minimal path p_1 between v and w in C_v, and another minimal path p_2 between u and w in C_u. Since C_v and C_u are not connected in G, merging p_1 and p_2 results in a minimal path between v and u that goes through w.

So far, we have shown that the marginal contribution of w in G can be different than in G' only when w is on some minimal path between v and u. Finally, from (1) and (13), we know that the Attachment Centrality of w is a weighted average of the marginal contributions of w in the game (V, f^*_G). Thus, by adding $\{v, u\}$, the Attachment Centrality of w may change only when w is on some minimal path between v and u. This concludes the proof. □

The above theorem leads to a series of corollaries that provide additional insights on how the Attachment Centrality measures the role of a node in connecting the network. We divide our analysis into two parts: the first focuses on nodes with high connectivity (such as *cut vertices*), whereas the second part focuses on nodes with low connectivity (such as leafs).

High connectivity: Given a *cut vertex*, $v \in V$ (i.e., a node that connects disjoint parts of the graph), we will show that the Attachment centralities of the nodes in each part are not influenced by the other parts. This, in turn, implies that the Attachment Centrality of v is simply the sum of its Attachment Centrality computed for each part separately.

THEOREM 4. *Let G be a connected graph, and let v be a node removal of which breaks G into k disjoint components consisting of the following sets of nodes: $C_1, \ldots, C_k$. Then,*

$$A_v(G) = \sum_{i \in \{1,\ldots,k\}} A_v(G[C_i \cup \{v\}]).$$

Furthermore, for every $i \in \{1, \ldots, k\}$, and every $u \in C_i$:

$$A_u(G) = A_u(G[C_i \cup \{v\}]).$$

PROOF. First, let us focus on an arbitrary node $u \in C_i$ for some $i \in \{1, \ldots, k\}$. Now let us remove from G an edge $\{w, w'\} \in E$ such that $w, w' \in C_j \cup \{v\}$ for some $j \neq i$. Note that a minimal path between w and w' cannot contain nodes from C_i. Based on this, Theorem 3 implies that the removal of $\{w, w'\}$ from G does not affect the Attachment Centrality of node u. By removing every such edge one by one, we eventually end up removing from G every edge outside $G[C_i \cup \{v\}]$ without affecting the Attachment Centrality of node u. Based on this, as well as Locality, we have that $A_u(G) = A_u(G[C_i \cup \{v\}])$.

Now, let us turn our attention to node v, and let us start by computing the Attachment centrality of v in each of the following subgraphs separately: $G[C_i \cup \{v\}] : i \in \{1, 2, \ldots, k\}$; we find that:

$$A_v(G[C_i \cup \{v\}]) = 2|C_i| - \sum_{u \in C_i} A_u(G[C_i \cup \{v\}]).$$

Since we already proved that $A_u(G[C_i \cup \{v\}]) = A_u(G)$ for every $u \in C_i$, we get:

$$\sum_{i \in \{1,\ldots,k\}} A_v(G[C_i \cup \{v\}]) = 2(|C|-1) - \sum_{u \in C \setminus \{v\}} A_u(G) = A_v(G)$$

which concludes the proof. □

The next corollary concerns bridges, i.e., edges the removal of which increases the number of connected components in the graph. The corollary follows from Theorem 4 and it is based on the observation that both ends of a bridge are cut vertices.

COROLLARY 1. *Removing a bridge decreases the Attachment Centrality of both its ends by 1, and does not affect the Attachment Centrality of other nodes.*

Interestingly, according to the above corollary, the connectivity role played by a bridge is attributed solely to is two ends. Furthermore, the fact that a node v is an end of a bridge does not influence in any way the Attachment centralities of the nodes connecting v to the rest of the network.

Whereas Corollary 1 focuses on *edges* whose removal increases the number of connected components, Theorem 5 focuses on *cliques* whose removal increases the number of connected components (such cliques are known as *cut cliques*).

THEOREM 5. *Let G be a connected graph. If a set of nodes $K \subseteq G$ forms a clique in G, and the removal of K breaks G into k disjoint components consisting of the sets of nodes: $C_1, \ldots, C_k$, then for every $v \in K$:*

$$A_v(G) = \sum_{i \in \{1,\ldots,k\}} A_v(G[C_i \cup K]) - (k-1)A_v(G[K]).$$

Furthermore, for every $i \in \{1, \ldots, k\}$ and every $u \in C_i$:

$$A_u(G) = A_u(G[C_i \cup K]). \quad (19)$$

PROOF. First, let us focus on an arbitrary node $u \in C_i$ for some $i \in \{1, \ldots, k\}$. Analogously to the proof of Theorem 4, we argue that any edge between two nodes in $C_j \cup K$ for some $j \neq i$ does not affect the Attachment Centrality of u. This, in turn, implies the correctness of (19).

Now let us turn our attention to an arbitrary node $v \in K$, and let us analyse the marginal contribution of this node to an arbitrary coalition $S \subseteq V \setminus \{v\}$. Without loss of generality, let $K(G[S]) = \{C_1, \ldots, C_l\}$ be the components of S, and assume that v is connected to the first m components, where $1 \leq m \leq l$. Following the definition of f^*_G (i.e., Equation (14)) the marginal contribution of v to S equals $2m$. Every connected component C_i for $i \leq m$ either contains all elements $K \cap S$ or is a subset C_j for some j. Thus, whenever S contains at least one element of K, then node v gets 2 for this component k times instead of 1. S contains at least one element of K with the probability $1 - \frac{1}{|K|}$. With the same probability node v has non-zero (and equal 2) marginal contribution in a clique of nodes K. This concludes the proof. □

Low connectivity: This part focuses on nodes with almost no connectivity role. The first corollary concerns *leafs*.

COROLLARY 2. *The Attachment Centrality of a leaf equals* 1. *Furthermore, removing a leaf decreases the Attachment Centrality of its neighbor by 1, and does not affect the Attachment Centrality of any other node in the graph.*

PROOF. Let v be a leaf, and let u be its only neighbor. Furthermore, let $S \subseteq V$ be the set of nodes comprising the component that contains both v and u. The presence of the edge $\{v, u\}$ increases the "*profit*" of $S \cup \{v\}$ by 2, since $f^*(S \cup \{v\}) - f^*(S) - f^*(v) = 2(|S|) - 2(|S| - 1) - 0 = 2$. Furthermore, since we know from the proof of Theorem 2 that the Attachment Centrality satisfies the *Component Efficiency* requirement, we know that the profit of 2 must be divided wholly among the nodes of the graph. However, according to Theorem 3, the edge $\{v, u\}$ only affects the Attachment Centrality of nodes lying on a minimal path between v and u, and since v is a leaf, the edge $\{v, u\}$ only affects the Attachment Centrality of v and u. In other words, the profit must be divided *wholly* between v and u. Finally, according to Fairness, this profit must be divided *equally* between v and u. This implies the correctness of the theorem and concludes the proof. □

So far, we discussed a type of nodes that plays a relatively-small connectivity role, namely a leaf. The following theorem focuses on yet another such type—a node whose set of neighbors forms a clique. The reason such a node, v, plays a relatively-small connectivity role is that v does not appear on any minimal path between any two nodes $u, w \in V \setminus \{v\}$.

THEOREM 6. *Given a node, v, whose set of neighbors, K, forms a clique, the Attachment Centrality of v equals $\frac{2|K|}{|K|+1}$. Furthermore, removing v decreases the Attachment Centrality of each of its neighbors by $\frac{2}{|K|(|K|+1)}$, and does not affect the Attachment Centrality of any other node in the graph.*

PROOF. Let $G = (V, E)$ be a graph in which a node, $v \in V$, has a set of neighbors, K, that forms a clique. It suffices to prove:

$$A_u(G) = \begin{cases} \frac{2|K|}{|K|+1} & \text{if } u = v, \\ A_u(G[V \setminus \{v\}]) + \frac{2}{|K|(|K|+1)} & \text{if } u \in K, \\ A_u(G[V \setminus \{v\}]) & \text{otherwise.} \end{cases}$$

Node v has non-zero marginal contribution to coalition S, equal 2, if and only if S contains at least one of its neighbors. Based on the permutation interpretation of the Shapley value, this happens with the probability $\frac{|K|}{|K|+1} \rightarrow 1$. All his neighbors benefit in a marginal contribution from v (have a greater marginal contribution to coalition with v, than without him) only if S does not contain any other neighbor. Such marginal contribution happens with the probability $\frac{1}{|K|(|K|+1)}$. Since other nodes do not appear on the minimal path between v and his neighbors in G', all edges of v can be removed without the change in their value. □

5. ALGORITHM AND APPLICATION

As an application, we focus on the identification of key terrorists in covert organisations. In particular, we analyse of the terrorist network responsible for the 2004 attacks on Madrid trains. The reasons behind our choice of the application and the network are twofold. Firstly, it has been recently argued that connectivity plays a crucial role in identifying the key members of terrorist networks [13, 15]. Secondly, the Madrid network is relatively big and, thus far, has never been analysed with a centrality index of this kind.

The *Madrid network* consists of 70 nodes and 243 edges. The size of the network makes it impractical to compute the existing connectivity-based centrality indices. In more detail, the computation involves enumerating all induced connected subgraphs of the network. Unfortunately, even the state-of-the-art algorithm for this purpose [22] takes over 100 seconds to compute the Myerson value for sparse network with only 36 nodes. Furthermore, the running time grows exponentially with the size of the network; every additional node nearly doubles it. To address this challenge, we use techniques introduced in the previous section to narrow down the set of nodes for which the Myerson value has to be calculated.

The original Madrid network [18] contains 6 isolated nodes. From Normalization, we know that the Attachment Centrality of each of these nodes is 0. We also know from Locality that those 6 nodes can be removed without affecting the Attachment Centrality of others. Furthermore, we observe that the Madrid network contains 8 leaf nodes. From Corollary 2, we know that every such node has an Attachment Centrality of 1, and can easily be removed from the network (since the corollary specifies the impact of this removal). Moreover, from Theorem 6 we know that every node whose set of neighbors, K, forms a clique has an Attachment Centrality of $\frac{2|K|}{|K|+1}$, and can easily be removed from the network (since the theorem specifies the impact of this removal on the Attachment Centrality of other nodes). Note that removing nodes results in a chain reaction, meaning that the above rules can be applied repeatedly (e.g., by removing a leaf, some other node might become a leaf, which can then be removed, and so on).

Our algorithm carries out the above process systematically, by finding the *cut-clique decomposition* of a graph. In a nutshell, the cut-clique decomposition of a graph G is a binary tree in which every node t is labeled with a subset $S \subseteq V$. If subgraph $G[S]$ has a cut-clique, then for a (possibly one of many) cut-clique K both his children are labeled according to the decomposition of $G[S]$ using K. Specifically, children l and r are labeled with two subsets $L, R \subseteq S$ such that $L \cap R = K$, and there exists no edge between $L \setminus K$ and $R \setminus K$. Theorem 5 allows for considering those subgraphs independently. By using the algorithm proposed by Tarjan [23], which utilizes the *simple elimination ordering* [19], we know that only one child can be further decomposed. This simplifies our algorithm (see Algorithm 1 for the pseudocode).

The results of our analysis are summarized in Table 1. As can be seen, the Attachment Centrality significantly differs from the standard centrality indices. For instance, let us consider two nodes with the highest number of edges – 1 and 3 – which are positioned

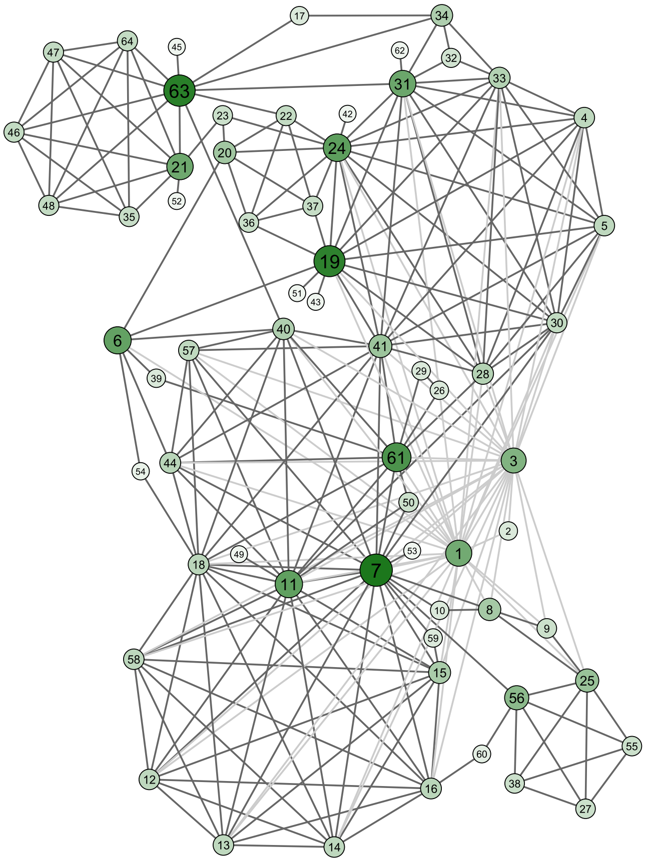

Figure 2: *Madrid network*. The node size reflects the Attachment Centrality (the larger the node the greater its centrality). To highlight the differences even further, the node color is set to reflect the node size (the larger the node the darker the color).

top by Degree and Closeness Centralities. Interestingly, Betweenness Centrality also gives them very high (the second and the third) positions while Attachment Centrality ranks node 1 as the tenth and node 3 even lower. Such a substantial difference between Betweenness and Attachment centralities is quite surprising, given that nodes that connect different parts of the network are more likely to belong to shortest paths than other nodes. However, nodes 1 and 3 have so many neighbors that they are very often parts of the shortest paths in the network, more often than nodes important from the connectivity perspective.

The running time of our algorithm depends on the topology of the graph. In the case of the Madrid network, the largest subgraph for which the Myerson value had to be calculated in line (11) was a subgraph consisting of 26 nodes. The running time on the entire network was 15.01 seconds on a standard desktop PC.

Our analysis revealed a previously unknown aspect of the Madrid network. In particular, we identified several overlapping parts of the network that are almost fully connected, i.e., each part is almost a clique, except for very few missing edges. This new insight confirms the existing belief, that *terrorist networks consist of rather sparsely-connected, highly-dense parts* [24].

6. RELATED WORK

A number of game-theoretic centrality indices have been recently proposed in the literature. In particular, Suri and Narahari [17] proposed an extension of degree centrality defined as the Shapley value of the game $f(S) = |\bigcup_{v \in S} neighbors(v)|$. Michalak et al. [14] considered a number of generalizations of this game. All these measures do not satisfy Fairness and Normalization. Also, if we consider their normalized version then they do not satisfy Locality.

To expose the connectivity role of a node, several authors proposed indices that are based on the Myerson value. Skibski et al. [22] considered several characteristic functions (e.g., $f(S) = |S|^2$ or $f(S)$ = number of edges in $G[S]$) combined with the graph-restrictions from Myerson's model. Depending on the function used, the resulting centrality measures do not satisfy Normalization nor Fairness (note that when the function f used in the Myerson value is based on the graph, Fairness may be violated).

A slightly different model (compared to Myerson's) was proposed by Amer and Gimenez [2], whereby the centrality of a node is the Shapley value of the following function: $f(S) = 1$ if $G[S]$ is connected and $f(S) = 0$ otherwise. This was later expanded by Lindelauf et al. [13] to an arbitrary $f(S)$ when $G[S]$ is connected. The resulting centrality measure does not satisfy Locality, since all centrality indices equal zero in a network with two disjoint parts.

Algorithm 1: Algorithm for the Attachment Centrality

Input: Graph $G = (V, E)$
Output: Attachment Centrality A_v for every $v \in V$
1 *find the simple elimination ordering* π*;*
2 *create a cut-clique decomposition* T *of graph* G *from* π*;*
3 $f \leftarrow$ *function defined as* $f(C) = 2(|C| - 1)$;
4 $t \leftarrow$ *root of T*;
5 **while** t *has children* **do**
6 $(l, r) \leftarrow$ *children of* t *(left one without children)*;
7 $(L, R) \leftarrow$ *labels of* l, r;
8 $K \leftarrow L \cap R$;
9 **foreach** $v \in L$ **do**
10 *calculate* $MV_v(f, G[L])$*;*
11 $A_v \leftarrow A_v + MV_v(f, G[L])$;
12 **if** $v \in K$ **then** $A_v \leftarrow A_v - 2 + \frac{2}{|K|}$;
13 $t \leftarrow r$;
14 **return** A_v *for every* $v \in V$;

Rank	A_v		B_v	C_v	D_v
1^{st}	7	(Imad Eddin Barakat)	63	1	1
2^{nd}	63	(Semaan Gaby Eid)	1	3	3
3^{rd}	19	(Abderrahim Zbakh)	3	41	7
4^{th}	61	(Mohamed El Egipcio)	40	7	11
5^{th}	24	(Naima Oulad Akcha)	7	31	41
6^{th}	11	(Amer Azizi)	31	40	18
7^{th}	6	(Mohamed Chedadi)	24	24	24

Table 1: The seven highest ranked nodes in the Madrid network, according to different centrality indices.

7. CONCLUSIONS

While there were some attempts in the literature to provide theoretical foundations to the standard centrality indices [20, 12, 6], our analysis is the first in the literature that proposes an axiomatization of an index focusing on connectivity.

Acknowledgements: Oskar Skibski and Makoto Yokoo were supported by JSPS KAKENHI Grant (24220003). Tomasz Michalak were supported by the European Research Council under Advanced Grant 291528 ("RACE"). This work was also supported the Polish National Science Center grant DEC-2013/09/D/ST6/03920.

REFERENCES

[1] N. Adler and B. Golany. Evaluation of deregulated airline networks using data envelopment analysis combined with principal component analysis with an application to Western Europe. *European Journal of Operational Research*, 132(2):260 – 273, 2001.

[2] R. Amer and J. M. Giménez. A connectivity game for graphs. *Mathematical Methods of Operations Research*, 60(3):453–470, 2004.

[3] R. Amer, J. M. Giménez, and A. Magaña. Accessibility in oriented networks. *European Journal of Operational Research*, 180(2):700 – 712, 2007.

[4] J. Belau. Consequences of connection failure - centrality and the importance for cohesion. In *Public Choice Society*, 2014.

[5] P. Boldi and S. Vigna. Axioms for centrality. *Internet Mathematics*, 10(3-4):222–262, 2014.

[6] S. P. Borgatti and M. Everett. A graph-theoretic framework for classifying centrality measures. *Social Networks*, 28(4):466–484, 2006.

[7] U. Brandes and T. Erlebach. *Network Analysis: Methodological Foundations*. Springer-Verlag, 2005.

[8] Z. Dezső and A.-L. Barabási. Halting viruses in scale-free networks. *Physical Review E*, 65:055103, 2002.

[9] L. Freeman. Centrality in social networks: Conceptual clarification. *Social Networks*, 1(3):215–239, 1979.

[10] D. Gomez, E. González-Arangüena, C. Manuel, G. Owen, M. del Pozo, and J. Tejada. Centrality and power in social networks: a game theoretic approach. *Mathematical Social Sciences*, 46(1):27–54, 2003.

[11] A. Keinan, C. C. Hilgetag, I. Meilijson, and E. Ruppin. Casual localization of neural function: the Shapley value method. *Neurocomputing*, 58-60(0):215–222, 2004.

[12] D. Koschützki, K. Lehmann, L. Peeters, S. Richter, D. Tenfelde-Podehl, and O. Zlotowski. *Centrality indices. Network analysis*, volume 3418 of *Lecture Notes in Computer Science*, pages 16–61. Springer, 2005.

[13] R. Lindelauf, H. Hamers, and B. Husslage. Cooperative game theoretic centrality analysis of terrorist networks: The cases of Jemaah Islamiyah and Al Qaeda. *European Journal of Operational Research*, 229(1):230 – 238, 2013.

[14] P. T. Michalak, K. V. Aaditha, P. L. Szczepański, B. Ravindran, and N. R. Jennings. Efficient computation of the Shapley value for game-theoretic network centrality. *Journal of Artificial Intelligence Research*, 46:607–650, 2013.

[15] T. P. Michalak, T. Rahwan, P. L. Szczepański, O. Skibski, R. Narayanam, M. J. Wooldridge, and N. R. Jennings. Computational analysis of connectivity games with applications to the investigation of terrorist networks. In *IJCAI*, pages 293–301. AAAI Press, 2013.

[16] R. B. Myerson. Graphs and cooperation in games. *Mathematics of Operations Research*, 2(3):225–229, 1977.

[17] N. Ramasuri and Y. Narahari. Determining the top-k nodes in social networks using the Shapley value. In *AAMAS*, pages 1509–1512, 2008.

[18] J. A. Rodríguez. The March 11 th terrorist network: In its weakness lies its strength. *PhD Thesis*, 2005.

[19] D. J. Rose, R. E. Tarjan, and G. S. Lueker. Algorithmic aspects of vertex elimination on graphs. *SIAM Journal on computing*, 5(2):266–283, 1976.

[20] G. Sabidussi. The centrality index of a graph. *Psychometrika*, 31(4):581–603, 1966.

[21] L. Shapley. A value for n-person games. In H. Kuhn and A. Tucker, editors, *Contributions to the Theory of Games, volume II*, pages 307–317. Princeton University Press, 1953.

[22] O. Skibski, T. P. Michalak, T. Rahwan, and M. Wooldridge. Algorithms for the Shapley and Myerson values in graph-restricted games. In *AAMAS*, pages 197–204, 2014.

[23] R. E. Tarjan. Decomposition by clique separators. *Discrete mathematics*, 55(2):221–232, 1985.

[24] M. Tsvetovat and K. M. Carley. Generation of realistic social network datasets for testing of analysis and simulation tools. Technical report, DTIC Document, 2005.