# Social Network Motifs: A Comparison of Social Groups

Diane Felmlee, Cassie McMillan
Department of Sociology
Penn State University
State College, PA, USA

Don Towsley, Kun Tu
College of Information & Computer Sciences
UMass – Amherst
Amherst, MA, USA

Roger Whitaker
Department of Computer Science
Cardiff University
Cardiff, Wales, UK

Gavin Pearson
Defence, Science and Technology Labs
Porton Down
Salisbury, UK

***Abstract*— Many types of complex networks that occur in nature, such as those from biochemistry, neurobiology, ecology, and engineering, exhibit some of the same, simple, network structural properties, that is, "network motifs". Network motifs refer to recurring, significant patterns of interaction between sets of nodes, and they represent the basic, building blocks of graphs. We examine patterns of motifs in a set of social groups, including the networks of terrorists, nation state alliances, friendship, travel, advice, twitter, and legislative cosponsorship. Findings reveal several common network substructures. Reciprocity of directed ties occurs much more frequently than expected by chance, and approaches the maximum in most social networks. Closed, or transitive, triads are more frequent than in equivalent random graphs, but not in twitter or travel. The motifs of terrorist networks resemble more closely those of close ties, rather than those of isolated cells. Intriguing four node patterns emerge as well. Results suggest that motifs could be used to fill in missing group ties when information is incomplete and to predict network genre from limited information (close ties vs. hierarchy). Motifs also contain the seeds of dynamic change. Networks with high levels of triangulation, for example, would be predicted to close open triads in the future.**

***Keywords—social networks, social groups, network motifs, dyads, triads, tetrads***

## I. Introduction

In the coalition environment, problems related to hostile and extreme external group-behavior may emerge, such as in asymmetric warfare, insurgency and post conflict peace-keeping. A common understanding of how and why groups behave in particular ways is fundamental for military intelligence, informing policy, resource deployment, and wider scenario modeling. However a persistent challenge concerns detecting and understanding the dynamics of partially visible groups. Behavior, interactions and communication may only be visible between particular nodes in the network and at specific points in time, presenting a major obstacle for coherent modeling.

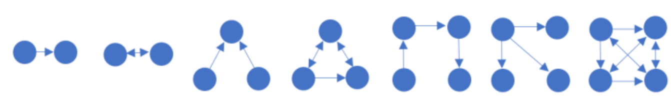
Here we attempt to advance the state-of-the-art by exploring the use of motifs to understand the external networks facing coalition operations, where noise and obfuscation is present. Network motifs refer to recurring, significant patterns of interaction between sets of nodes [1] [2]. They represent the basic, building blocks of graphs. While an external network of interest may not be fully visible to the coalition, motifs represent important atomic sub-structures more likely to be visible, from which inferences can be made. However, research seldom investigates motifs within graphs that capture important human ties and interactions, which represent the focus of our study.

## II. Network Motifs: Technical Approach

Motifs in networks consist of basic representations of small sets of vertices and edges in a graph. They can be used to identify local patterns of structural regularity that occur more often than would be expected in random graphs with the same number of nodes and edges.

Triads, or the (potential) ties connecting subsets of three nodes, are considered to be one of the structural foundations of social networks [3]. By studying triads we can better understand a variety of network phenomena, including transitivity, the tendency for node $i$ to be tied to node $k$ if a tie exists between node $i$ and node $j$ and between node $j$ and node $k$. See sample 2, 3, and 4-node motifs in Figure 1.

Fig. 1. Examples of 2, 3, and 4-node motifs.



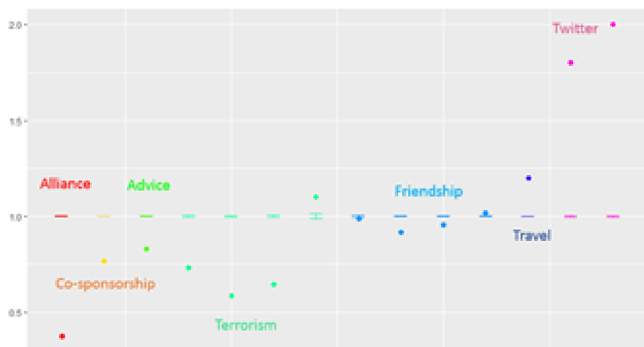
We use simulations of random networks to examine the prevalence of individual, network motifs, by comparing with random networks with the same number of nodes and edges [Mil02]. We examine network motifs in seven types of social networks, including those of 1) nation state alliances, 2) co-sponsorship of legislative bills, 3) organizational advice ties, 4) multiple terrorist groups, 5) multiple friendship groups, 6) airplane travel routes, and 7) negative twitter messaging.

Some of our networks are defined by relationships that are inherently symmetric (e.g., two nations are allies with one another) while others can be asymmetric (e.g., an employee seeks advice from another employee). When we make comparisons between the symmetric and asymmetric networks, we symmetrize the asymmetric networks so that if $a \rightarrow b$ or $b \rightarrow a$, an undirected edge will exist between nodes $a$ and $b$.

## III. Results

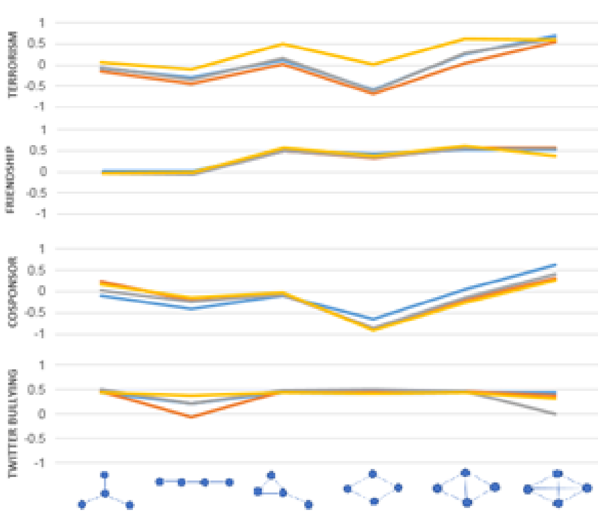
Several motifs appear to be universal in social networks. For example, in all the social networks that contain directed ties, the occurrence of reciprocal dyads is greater than expected in equivalent random graphs (i.e., if *a* → *b*, then *b* → *a*), and it typically approaches the maximum degree of reciprocity possible in the graph [4]. On the other hand, certain types of networks exhibit unique types of 3-node motifs. In particular, twitter communication networks that consist of aggressive, bullying messages and their responses, contain a much higher proportion of intransitive, stressful triads (triad "201") than predicted by chance (i.e., a set of three nodes where *a* --- *b* and *c* --- *b*, but there is no tie between *a* and *c*), and travel networks also contain more than expected. However, the remaining five networks contain fewer than predicted. See Figure 2, below.

Fig. 2. Ratio of number of observed "201" triad motifs to number in random networks (with confidence intervals).




Next we calculate significance profiles for four node motifs [5], in which we compare relative frequencies of motifs in symmetrized networks, as compared to random networks (where a value of zero represents no difference). Several patterns emerge. For example, in all types of networks, the "four clique" motif, which consists of all possible ties between the four nodes, is more common than random. A "straight line" motif, on the other hand, tends to be less frequent than, or no different from, chance. Unique patterns, however, emerge in the overall significance profiles. For example, the terrorist and co-sponsorship graphs display fewer than expected 4 node cycles and greater than expected 4 node cliques. See illustrations of several significance profiles in Figure 3.

Fig. 3. Significance profiles for symmetrized 4 node motifs.




Note that the motif patterns for terrorist groups tend to be more similar to those of positive, cooperative networks than to those of either negative, online twitter communication or to those of hierarchical, organizational advice networks. The four node significance profile for terrorist networks appears to be most similar to that of legislative co-sponsorship ties, for example. These results suggest that interconnections among insurgents are similar to those of close, cooperative connections, providing support for the argument that the terrorist groups studied here developed out of deep ties, as compared to the argument that such groups represented connections among relatively isolated cells.

## IV. Conclusions

In conclusion, we see that similar to biological and physical networks, 2, 3, and 4-node motifs develop patterns that shape the networks of social ties. Some of these subgraphs consistently reoccur in our sample of networks, suggesting a level of universality to specific types of human, micro-level, interactions. Unique patterns emerge as well, which highlights the argument that groups, such as terrorists and those involved in online negative twitter interchanges, tend to develop signature motif patterns that can be used to better understand their interactions. Furthermore, since underrepresented motifs are apt to be much more unstable than those that are overrepresented, in future work these findings can be used to predict change in groups over time as well as the presence of missing tie information.

## V. Acknowledgement

This research was sponsored in part by the U.S. Army Research Laboratory and the U.K. Ministry of Defence under Agreement Number W911NF-16-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

## References

[1] U. Alon, "Network Motifs: Theory and Experimental Approaches." *Nature* 3: 450-461. 2007.

[2] S. S., Shenn-Orr., R. Milo, S. Mangan, and U. Alon, "Network Motifs in the Transcriptional Regulation Network of *Escherichia coli*." *Nature Genetics* 31: 64-68. 2002.

[3] P. Holland and S. Leinhardt, "Local Structure in Social Networks." *Sociological Methodology* 7: 1-45. 1976

[4] B. Jiang, Z. Zhang, and D. Towsley. "Reciprocity in Social Networks with Capacity Constraints." *Proc. KDD* 2015, August 2015.

[5] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon, "Superfamilies of Evolved and Designed Networks." *Science* 303: 1538-1542. March 2004.