Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2019) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

18th Annual Conference on Systems Engineering Research (CSER)

Towards the Design of Artificial Swarms Using Network Motifs

Khoinguyen Trinh, Zhenghui Sha*

Department of Mechanical Engineering, University of Arkansas, Fayetteville, AR 72701, USA

Abstract

Many complex systems evolve as a result of interactions among individual entities whose behaviors cannot be directly controlled. This makes the design of such systems inherently challenging. The objective of this research is to develop a new approach in engineering complex swarm systems with desired characteristics based on the theory of network motifs – subgraphs that repeat themselves among various networks. In recent studies, the discovery of network motifs has presented the ability to determine reoccurring similarities between similar functioning networks that were originally believed to have not shared any characteristics. It is therefore hypothesized that manipulating the types of network motifs within a network can help engineer artificial swarms with improved functionality. In this study, artificial swarm systems have been modeled as a dynamic complex network where each node represents an individual foraging entity and links represent as the communication between entities. Additionally, motif-detecting algorithms have been used to extract subgraphs that reoccur in these complex networks. Our research has shown promising results that reveal a statistically significant correlation between network motifs and the performance of simulated swarm networks. This study contributes as a new approach that can potentially be used in the design and engineering of complex swarm systems.

Keywords: network motifs; complex networks; artificial swarm; network analysis; complex systemdesign

1. Introduction

1.1. Background

The engineering of complex systems has traditionally followed a top-down methodology which creates a framework for the system and adds additional features to meet specific design requirements. This general process is embodied in various existing systematic design methods (e.g., Pahl and Beitz's theory [1]), systems engineering models (e.g., Systems Engineering V [2] and Waterfall model [3]), and system engineering processes adopted by organizations such as NASA [4]. For example, the design of a vehicle system requires the decomposition of high-level requirements (e.g., safety, reliability) into individual units such as an acceleration unit or a braking mechanism which can be further broken down into mechanical components that can be designed and manufactured with pre-existing knowledge and tools.

* Corresponding Author: Tel.: +1-479-575-3422
E-mail address: zsha@uark.edu

© 2019 Khoinguyen Trinh, Zhenghui Sha. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/) Peer-review under responsibility of the scientific committee of the 18th Annual Conference on Systems Engineering Research (CSER)

Over the past decade, there has been a growing interest in developing solutions involving complex systems, where the system-level structure emerges from the behaviors of individual entities and their interactions with each other. Some examples of this approach are the development of swarm robotics to fight forest fires and the use of automated drone system to transport industrial goods. In contrast, the design of these complex multi-agent systems differs from traditional engineering design. Instead, it often draws inspiration from natural swarm, such as ant colonies and bee swarms, where individuals make decisions based on local information. In such systems, there are a large number of individual entities, which are heterogeneous in nature, have private objectives that must be met without compromising system goals. Therefore, engineering such systems towards a desired system-level performance is inherently challenging and maintaining a direct control of the system is nearly impossible. A bottom-up approach that aims at engineering local interactions must be developed, and a better understanding about the relationships between the local network structures and the system-level performance must be obtained.

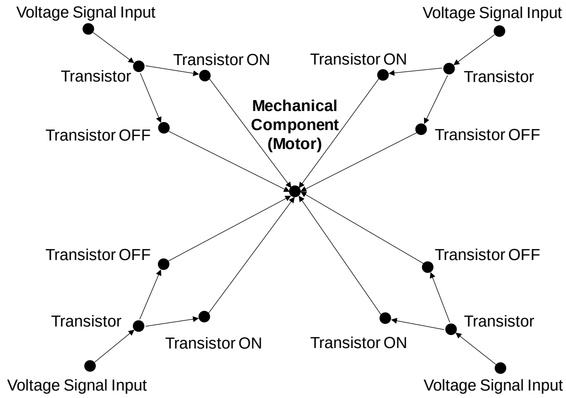


Fig. 1. H-bridge circuit modelled as a network

1.2. Design Methodology Using Network Motifs

In this paper, we present a new approach based on the theory of network motifs to efficiently gain information of how individuals interact with each other and how those interactions would influence the performance of system. Network motifs are defined as reoccurring subgraphs (or patterns) that repeat themselves among various networks [7]. Current research on network motifs has been primarily focused on the development of motif-detecting algorithms to identify reoccurring subgraphs within a network to support the analysis of network topologies [10]. As a result, many non-commercial algorithms, such as *mfinder* and *MAVisto*, have been developed to detect these reoccurring subgraphs [10]. Other motif research has involved the analysis of biological systems to determine similarities across different organisms [5]. Existing research has indicated that network motifs have certain functions to allow for the system to achieve its overall goal. Studies of biological systems, for example, have shown E. coli and yeast to share common motif families that make up both of their entire DNA transcription network. Particularly, one motif families, called auto-regulation motifs, allows E. coli to repress or accelerate the rate of DNA transcription [5].

One of the main problems currently faced with systems engineering is a lack of holistic methodology [11]. Generally, systems engineering focuses on the design of the system's parts rather than the system as a whole. This results in unexpected behaviors of the system which stem from interactions of different subsystems and their parts. We look to bridge this problem through a methodology that correlates complex system characteristics with network motifs. Inspired by existing studies, we hypothesize that the theory of network motifs can be used in designing complex systems. To give an overview on this approach, we use a simple example – the redesign of a traditional H-bridge circuit – as shown in **Figure 1.** The components of this H-bridge circuit can be modeled as various nodes that have connections to other nodes that send various electrical signals. Various subgraphs, such as the one involving the transistor node, can be found repeating themselves within the network [8]. The transistor node is directly responsible for receiving signals from one nodes and sending out new signals to other nodes, making this node crucial in the H-bridge system. The transistor node will receive a signal from voltage signal node that causes the transistor to turn itself 'on' or 'off' to allow for electrical current to flow through it. From the signals that the motor node receives, the motor node will either be in its off state, clockwise-rotation state, or counterclockwise-rotation state. This transistor subgraph is not exclusive to the H-bridge only and can be applied to a plethora of other electronic applications. If we have an in-depth understanding about the correlations between this transistor subgraph and its function within the circuit, then a new system could be engineered by promoting the formation of such subgraphs.

In general, the proposed approach can be summarized in **Figure 2**. First, a complex system is modeled as a complex network in which nodes and links are defined. With an established network structure, key reoccurring subgraphs can be identified by using motif-detection algorithms. Their functions must be analyzed and will eventually be used to modify pre-existing systems. To study the relationship between local network structures and system-level

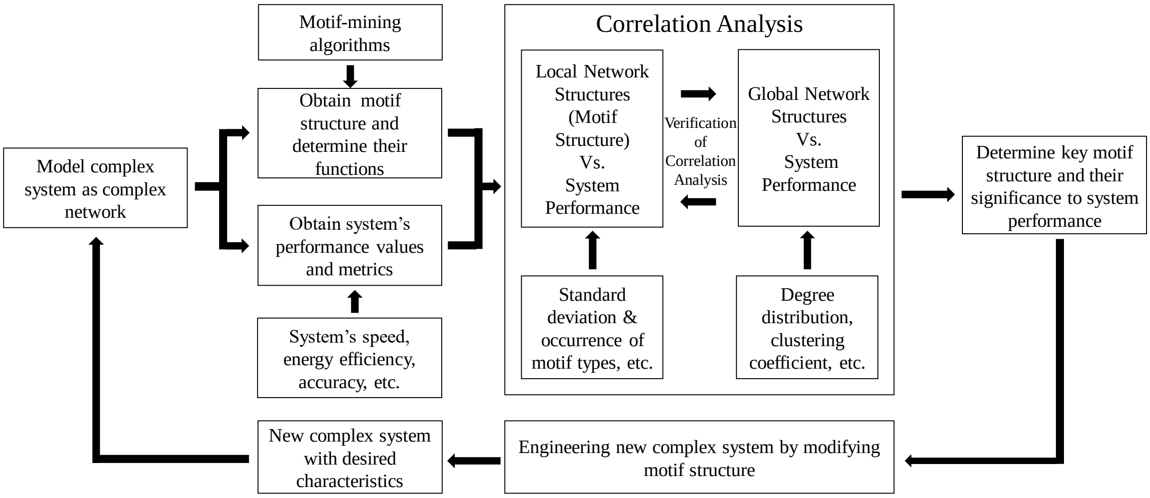


Fig. 2. The proposed approach for engineering complex swarming systems

performance, correlation analysis has to be performed between the extracted network motif data and the metric values that quantify the system performance. Based on the correlation analysis, the most important network motifs, i.e., the ones with the highest correlations, will become potential solutions in engineering the system's performance level. In order to verify such correlations, we propose to perform additional correlation analyses between the network motifs and network-level properties such as the average degree and the degree distribution. This step is necessary in order to verify if a correlation exists between network motifs and local network structure rather than the global network characteristics. After such verifications, those subgraphs can be confirmed in having a significant influence on system performance. It is, therefore, desired to promote the formation of such subgraphs (local structures) to achieve a higher level of performance. Additionally, these subgraphs can be applied to different complex systems to achieve a similar function as the original system.

In the following sections, this approach is demonstrated through a simulation study of a swarm system foraging for food. This swarm system has been selected for this case study because it represents general functions of many swarm systems in which individual entities communicate locally with nearby entities to collectively accomplish predefined tasks. The rest of this paper is structured as following: the description of the complex system of interest and the approach used to analyze the system is presented in Section 2. The results of the analysis and its discussion are presented in Section 3. Finally, we conclude this paper and present our closing insights in Section 4.

2. Case Study

2.1. The Swarm Foraging System and Simulation

A case study has been performed on multiple simulations, modeling a swarm of ants foraging for food. In the swarm, there are 100 entities (or ants) which would search in a confined 12-unit by 12-unit arena for food particles over 10,000 timesteps. Each

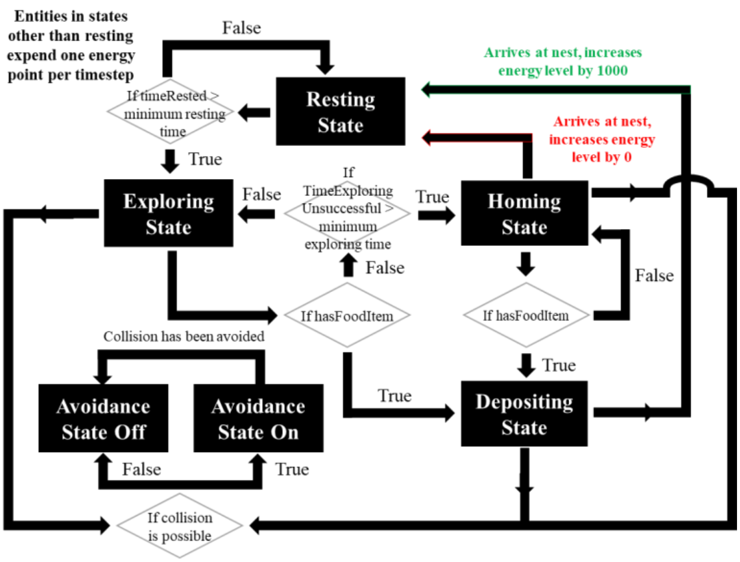


Fig. 3. The model of swarm foraging system

entity has 5 different states (see **Figure 3**): *resting*, *exploring*, *avoidance*, *depositing*, and *homing*.

1) The *resting* state is when entities are in their nest and are inactive.
2) The exploring state is when entities are searching for food particles.
3) The *avoidance* state results during the exploration state and is caused by an entity nearing collision with another entity or an obstacle (arena wall). The entity will then change its vector to avoid collision.
4) The *depositing* state is when an entity has successfully located a food particle and begins returning to the nest with the food particle.
5) The *homing* state results when the entity has failed to locate any food particles after its resting probability has reached a certain value. This state causes the entity to return to the nest.

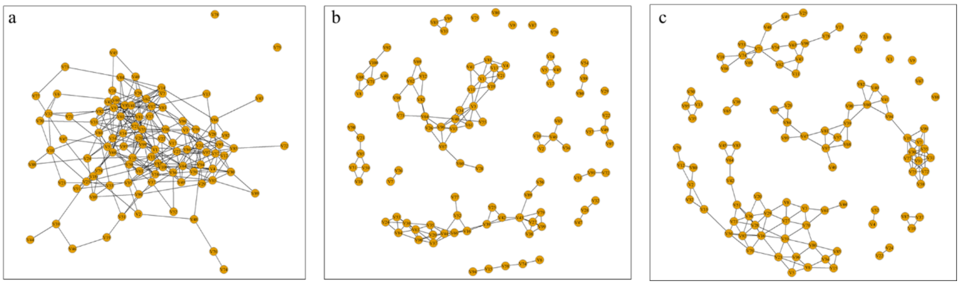


Fig. 4. Three simulations at timestep 10,000 represented by three networks (a) simulation 8, the best performing swarm; (b) simulation 18, the intermediate performing swarm; (c) simulation 13, the worst performing swarm

At the beginning of the simulation, all entities will begin in the nest area, or the gray area, of the arena. Each simulation begins with an energy level of 0 points. Once the simulation begins, all entities switch to the exploring state and search for food particles in the white area of the arena. For each timestep that an entity is moving, one energy point will be expended from the nest; this rule is implemented to account the energy lost (or food consumed) from active entities searching for food. For example, if 100 entities are active (exploring, avoidance, depositing, or homing) during a timestep, 100 energy points will be taken from the nest resulting in a -100-energy point loss for the nest. For each food particle that is brought back to the nest, 1,000 energy points are added to the overall energy level of the nest. As the entities are collecting food, they can communicate with each other by using an RAB (range-and-bearing) sensor. These RAB sensors allow multiple entities to communicate with one another as long as they are within a range of one-unit radius of each other. The information pertaining the position and state of the entity is relayed to other entities. Based on this information, each entity will modify its probability values which determine how the entities behave by changing their state once each value has become high enough.

2.2. Modeling the Complex System as a Complex Network

30 simulations have been produced. Each simulation generates a file that contains 10,000 100 × 100 adjacency matrices that show how the entities communicate with each other during each timestep. In each network, a node represents an ant and a link denotes the communication between a pair of ants. **Figure 4** shows the communication networks at the timestep of 10,000 of three swarms in three representative simulations: simulations 8, 18 and 13. These simulations correspond to the best performing, the intermediate performing and the worst performing swarms respectively. Once the network has been established, various network metrics such as the degree, geodesic distance, eccentricity, betweenness centrality, and clustering coefficient can be immediately obtained. The average values of these metrics of simulation 8 are shown in **Table 1**. For example, the degree metric indicates the average number of links an entity has at a particular timestep. This metric gives insight on how many entities are active and how tightly grouped the entities are for communication. The degree metric can help compare whether the overall number of links plays a stronger role in system performance or if various network subgraphs takes precedence in system performance.

Table 1. Network Properties of Simulation 8

Timestep	Foraging Performance	Degree	Geodesic Distance	Eccentricity	Betweenness Centrality	Closeness Centrality	Edge Betweenness Centrality	Local Clustering Coefficient	Global Clustering Coefficient
1-10000	0.389	3.293	5.261	7.473	0.015	0.017	52.957	0.435	0.534

2.3. Obtaining the Network Motif Data of the Simulations

Among the 10,000 networks in each simulation, 200 networks (one every 50 timesteps) have been selected for further analysis. This sampling decision has been made due to the computational consideration, and the same approach has been applied to all simulations. The edge list of each network has been used as a data input for a motif-detection program called *mfinder* [6]. **Figure 5** shows the occurrence of an example of motif type found in a representative simulation, simulation 8, during the 10,000 timesteps in ten timeframes.

In this study, 15 representative simulations have been selected for the correlation analysis. To select these simulations, the performance level has to be evaluated. The performance level of each simulation is measured by two aspects: the amount of food collected and the amount of energy expended during the entire foraging period. The performance of these simulations is calculated based on Equation (1)

$$P_s = \left|\frac{f * 1000}{e_f}\right|, \tag{1}$$

where f is the total number of food particles collected, e_f is the final energy level at timestep 10,000, and P_s is the performance value of the entire simulation. Based on this equation, the following simulations have been selected: the five best performing simulations (8, 9, 11, 16, 28), the five worst performing simulations (13, 20, 25, 27, 30), and the five most intermediate performing simulations (1, 18, 19, 26, 29).

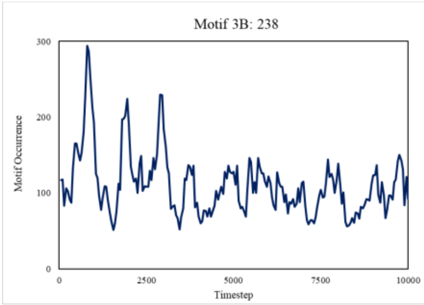


Fig. 5. The occurrence of motif 3B in simulation 8

2.4. Determining Important Network Motif Structures

The average and the standard deviation of network motif occurrence have been calculated across ten different timeframes, i.e., 1-1,000, 1,001-2,000, … 9,001-10,000, for every simulation. The average foraging performance of each simulation has also been calculated across ten different timeframes. This is because there needs to be at least one food particle collected within a selected timeframe in order to evaluate the system performance. Since the swarms in these simulations take over 50 timesteps to collect one food particle on average, all simulations have been divided into 10 separate timeframes to ensure that at least one food particle is collected within each time interval. The average foraging performance per time interval is calculated by Equation (2)

$$P_i = \left|\frac{f * 1000}{e_f - e_i}\right|, \tag{2}$$

where e_i is the initial timestep's energy level, e_f is the final timestep's energy level, f is the number of food particles collected within the timestep frame, and P_i is the performance value of the simulation during the interval of interest. In determining which motif structures are significant, correlation analysis has been performed between the average occurrence of each motif structure and the average performance of each swarm (i.e., each simulation). In addition, the standard deviation of each motif structure has also been correlated to the average performance. This is done because observations showed that some simulations yield a better performance with highly fluctuating occurrence values of network motifs in certain timeframes. Both data types have been correlated by using Pearson's correlation coefficient. In this study, a motif is determined to be significant if a) the correlation coefficients are 0.5 or higher, or b) it appears as one of the simulation's top-10 correlated motif types. These results are discussed in the Section 3.

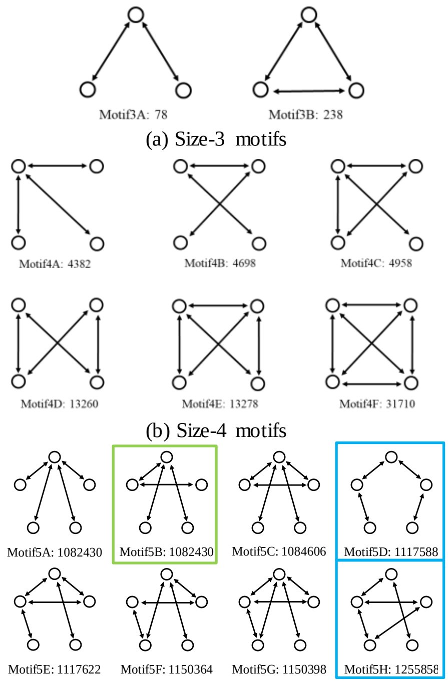
3. Results/Data

3.1. Motif Structure Study

The search of important network motifs has been limited to size-3, size-4, and size-5 motif structures (i.e., the motifs that consist of three, four, and five nodes, respectively). The motifs above the size of 6 nodes have been ignored in this study due to the large computational resources required – the *mfinder* algorithm has to run about 2 hours to analyze a single timestep using a computer configuration of a 64-bit Windows 10 Dell Laptop with an Intel Quad Core i7-7700HQ @ 2.80 GHz processor and 8 GB of RAM.

There are 29 different motif types that have been analyzed: 2 of which are size-3 motif structures, 6 of which are size-4 motif structures, and 21 of which are size-5 motif structures. Each motif structure is named by the following format: motif #L, where # represents the motif structure size and L represents a specific motif structure. For example, the two size-3 motif structures have been named as motif 3A and motif 3B and the motif structures for size-4 motifs have been named motif 4A, 4B, 4C, 4D, 4E, and 4F. **Figure 6(a)** shows the differences between the two types of size-3 motifs. Motif 3A has a centralized node that can freely communicate with two other nodes while the other two nodes cannot relay any information. This differs from motif 3B where all nodes are able to communicate with each other.

Due to the length of paper, we are unable to present the results of the correlation analysis for all the 15 simulation. Instead, we present the top-five swarms (i.e., simulation 8, 9, 11, 16 and 28) as a demonstration of the data. **Table 2** shows an abbreviated table of the sorted correlation coefficients between the average motif occurrence and the performance values of the five best-performing simulations. The top of the table shows motifs that have lower correlation coefficients and the bottom of the table shows higher correlation coefficients. In simulation 8, motif 5K is the least correlated motif type at 0.38 and motif 5N is the highest correlated motif type at 0.76. It has also been observed that certain motifs appear to be highly correlated more often than others. For example, Motif 5B (highlighted green) has a correlation coefficient higher than 0.50 for at least three of the five simulations. Motif 5D and 5H (highlighted blue) appear three times in the top-10 correlated motif types out of the five simulations. This correlation analysis has been applied to all 15 simulations. Based on the criteria of significance aforementioned, the following observations have been acquired:



(c) Sample of Size-5 motifs

Fig. 6. Different types of network motifs

a) At least 10 of the 15 simulations studied shows three motif types that have been valued as significant when correlating the foraging performance to average motif occurrence. In 11 of the simulations (8, 9, 11, 16, 18, 19, 20, 26, 27, 29, and 30), motif 5B has been found to have a correlation coefficient of at least 0.50. The average correlation coefficient of this motif type is 0.525. In 11 of the simulations (9, 11, 13, 16, 18, 20, 26, 27, 28, 29, and 30), motif 5D has been found to be in the top 10 of 29 motif types studied. The average correlation coefficient of this motif type is 0.542. Similarly, motif 5H is in the top 10 of 29 motif types in 10 simulations (11, 13, 16, 18, 20, 26, 27, 28, 29, and 30). The average correlation coefficient of this motif type is 0.522.
b) When correlating the foraging performance to the standard deviation of each motif type, at least 10 of the 15 simulations studied shows two motif-types that have been valued as significant. In 10 of the simulations (1, 8, 9, 11, 16, 18, 20, 26, 27, and 29), motif 5B has been found to have a correlation coefficient of 0.50 or greater. The average correlation coefficient of this motif using the standard deviation metric is 0.544. In 10 of the simulations (1, 9, 11, 13, 16, 19, 20, 26, 28, and 30), motif 5H has been found to be in the top 10 of 29 motif types studied. The average correlation coefficient of this motif using the standard deviation metric is 0.498.

Table 2. Correlation Between Motif Types Occurrence and System Performance of the Best-Performing Simulations

Simulation 8		**Simulation 9**		**Simulation 11**		**Simulation 16**		**Simulation 28**	
Motif5I	0.626	Motif5H	0.484	Motif5T	0.377	Motif5G	0.860	Motif5G	0.280
Motif5L	0.626	Motif5E	0.489	Motif3B	0.396	Motif5B	0.864	Motif5F	0.282
Motif4A	0.627	Motif4B	0.492	Motif5N	0.407	Motif5F	0.869	Motif5B	0.285
Motif5B	0.629	Motif4A	0.497	Motif5M	0.425	Motif3A	0.871	Motif5H	0.288
Motif5F	0.629	Motif5C	0.536	Motif4C	0.434	Motif4C	0.881	Motif4B	0.289
Motif5C	0.636	Motif5D	0.544	Motif4A	0.446	Motif5O	0.882	Motif5J	0.289
Motif5E	0.642	Motif5B	0.549	Motif5H	0.447	Motif5I	0.883	Motif5T	-0.298
Motif5P	0.643	Motif5K	0.554	Motif5E	0.475	Motif5J	0.889	Motif5I	0.300

Motif5J	0.664	Motif5S	-0.524	Motif5A	0.492	Motif4B	0.890	Motif5E	0.306
Motif5O	0.666	Motif5R	-0.583	Motif5D	0.495	Motif5H	0.897	Motif5P	0.307
Motif5G	0.670	Motif5T	-0.583	Motif5B	0.501	Motif5D	0.900	Motif5L	0.310
Motif5M	0.682	Motif5U	-0.602	Motif4B	0.526	Motif5E	0.902	Motif5D	0.323
Motif5N	0.756	Motif5A	0.631	Motif3A	0.549	Motif5M	0.904	Motif5U	-0.345

Our initial results indicated a possibility that the global network structure of a swarm system were likely to play a role in its performance rather than the local network structures of the system. Particularly, it was observed that the degree of a node showed potential relations with how well a system performed. To verify the conclusion that the system's performance level is highly correlated with the local subgraph structures rather than the number of links present within a network, we studied the correlation between the system performance and system's degree distribution.

The complementary cumulative distribution (CCD) of each simulation's degree metric has been analyzed. Similarly, each simulation has been divided up into ten different timeframes. Within each timeframe, the CCD curves have been plotted via a logmarithmic scale for every 200 timesteps (see **Figure 8**). In characterizing a CCD, the standard deviation has been calculated to determine the variability of the degree metric throughout each timeframe, and the slope has been calculated for the sake of completeness [9]. Therefore, the average and the standard deviation of the slopes of all curves' fitting lines have been calculated within each timeframe and correlated to the system performance values. Based on the results shown in **Table 3**, little correlation has been observed between the average slopes, standard deviation of the slopes, and the system performance values. These results therefore conclude that the degree distribution of a network is unlikely to have any effect on simulation performance. This strengthens the conclusion that local network structures are indeed significant to the system performance and thus provides a potential solution to engineering complex systems.

It is shown from the results that motif 5B, 5D, and 5H are significantly influential to the foraging performance of this specific swarm system (see **Figure 7**). Therefore, to engineer swarm systems with a higher level of performance, mechanisms can be designed to promote the formation of those motif types. Further validation studies will be required to reevaluate the newly engineered foraging performance and compare it to original simulations. Additionally, the following observation should be noted: the three motifs that have been valued to be significant had two-dangling nodes – two nodes that do not contain outgoing links except with each other. The only other motif structures that contain two-dangling nodes are motif 3A and motif 4B, but they are valued as important in only 9 of the 15 simulations.

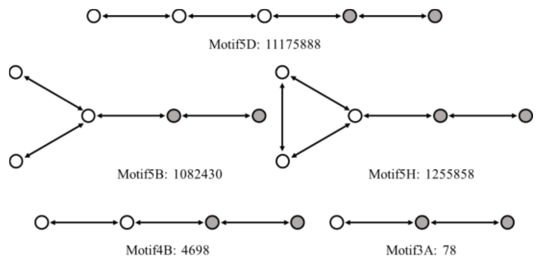


Fig. 7 Motif structures that contain two-dangling nodes

Table 3. Simulation 8 Properties (Motif and Degree Distribution Data)

Time step	Foraging Performance	Size-3 Motif Occur. Avg.	Size-3 Motif S.D.	Size-4 Motif Occur. Avg.	Size-4 Motif S.D.	Size-5 Motif Occur. Avg.	Size-5 Motif S.D.	Degree Dist. (Avg.)	Degree Dist. (S.D.)
1-1000	0.548	587.050	222.783	1977.100	1165.050	6994.900	5303.562	-0.079	0.011
1001-2000	0.368	436.500	197.852	1315.050	943.448	4282.750	4028.140	-0.085	0.014
2001-3000	0.450	500.950	137.584	1620.600	699.578	5577.300	3290.148	-0.077	0.003
3001-4000	0.321	364.800	110.636	977.750	478.131	2803.800	1996.968	-0.094	0.009
4001-5000	0.467	352.850	84.166	932.600	352.614	2636.150	1349.570	-0.089	0.007
5001-6000	0.312	402.600	76.747	1128.400	318.644	3241.050	1133.058	-0.085	0.007
6001-7000	0.253	347.200	42.780	847.050	172.140	2162.800	625.669	-0.090	0.007
7001-8000	0.309	335.550	99.238	827.750	404.755	2118.350	1409.952	-0.093	0.013
8001-9000	0.398	290.150	72.806	704.150	319.134	1897.100	1145.961	-0.095	0.014
9001-10000	0.490	387.250	70.659	1046.300	319.930	2933.850	1257.656	-0.088	0.011

4. Conclusions

This paper introduces a new approach in engineering complex system based on the theory of network motifs. This theory provides a method to determine the function that is associated with a group of nodes or network motifs that are essential to the system structure and performance. By determining these motifs, complex systems with desired features can be designed. Our approach is network-based, thus is general enough to be applied in other complex system as long as they can be modeled as complex networks.

The complex system used in this study are swarm simulations modeling a group of ants foraging for food. These simulations have been created in ARGOs – an experimental swarm simulation software. There are 30 simulations that have been created and 15 of these simulations have been chosen to be used in this study. The results found from this study show a correlation between specific network motifs and swarm foraging performance. The motifs that have been found to have a strong correlation to the foraging performance are motif 5B, 5D, and 5H; these three motifs all share a two-dangling nodes structure. These motifs have been valued as important in at least 10 of the 15 simulations studied.

Based on these results, new simulations can be generated to study the effect of these motif structures. For example, by controlling how often these motif structures appear, we hypothesize that we can manipulate how well we would want for these systems to perform. These motif structures could then be analyzed on a more microscopic scale to determine why these motifs tend to determine the system perform. This has yet to be achieved due to the complexities required to develop the algorithms to engineering motif structures. We will further investigate it in our future studies. This is also the reason that even if a high correlation has been observed, we are conservative to draw any causations and only provide possible explanations for such phenomenon. Additionally, future studies will be used to determine if the results of this correlation analysis are sensitive to changes in context and architectural parameters such as area dimensions and energy unit assignments.

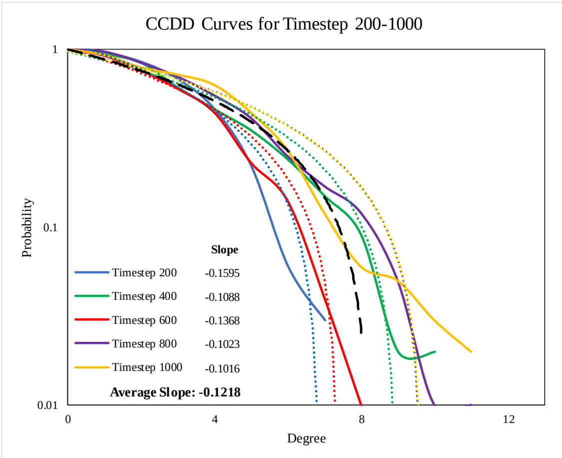


Fig. 8 Sample of a complementary cumulative degree distribution curve set for simulation 8 from timestep 200 to timestep 1000

Acknowledgements

We would like to thank Dr. Mengqi Hu and Zishun Yu from the University of Chicago at Illinois for providing the data, images, and simulation code that was necessary to completing this study. We would also like to thank Laxmi Poudel for helping with revisions of the manuscript.

References

[1] G. Pahl, W. Beitz. (1996) *Engineering Design: A Systematic Approach*. Springer, London, 2nd edition.

[2] D. M. Buede. (2000) *The Engineering Design of Systems: Models and Methods*. John Wiley and Sons, Inc., New York, N.Y.

[3] W. Scacchi. (2001) *Process Models in Software Engineering*. Marciniak, J.J. (ed.). Encyclopedia of Software Engineering, 2nd Edition, John Wiley and Sons, Inc, New York.

[4] NASA. (2007) *NASA Systems Engineering Handbook (NASA/SP-2007-6105 Rev1)*. National Aeronautics and Space Administration, Washington, DC.

[5] N.R. Zabet. (2011) "Negative Feedback and Physical Limits of Genes." *Journal of Theoretical Biology* 284 (1): 82-91.

[6] N. Kashtan, S. Itzkovitz, R. Milo, and U. Alon (2004) "Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs." *Bioinformatics* 20 (11): 1746-1758.

[7] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon. (2002) "Network Motifs: Simple Building Blocks of Complex Networks." *Science* 298 (5594): 824-827.

[8] S. Itzkovitz, R. Levitt, N. Kashtan, R. Milo, M. Itzkovitz, U. Alon. (2005) "Coarse-Graining and Self-Dissimilarity of Complex Networks." Phys. Rev. 71 (1): 016127

[9] National Research Council. (1996) *The Waste Isolation Pilot Plant: A Potential Solution for the Disposal of Transuranic Waste.*, The National Academies Press.

[10] M. Zuba. (2009) "A Comparative Study of Network Motif Detection Tools." *UConn Bio-Grid, REU Summer*.