Efficient Topological Distances and Comparable Metric Ranges

Musad Haque, Waseem Abbas, Abigail Rafter, and Julie A. Adams

Abstract— **Field studies indicate that European starlings in a flock coordinate utilizing a nearest neighbors approach. This model of communication is known in the biological swarm literature as the topological model. Each starling coordinates with a fixed number of nearest neighbors, and this number is referred to as the topological distance. The presented research evaluates the topological model in the context of a swarm robotics task, where a simulated artificial swarm is tasked to search for and go to a goal in the environment. Specifically, experiments are conducted on a high–fidelity, multi-robot simulator to analyze the performance of the tasked swarm as the topological distance is varied. Connections are drawn between the topological model based on starlings and the Delta–disk model, which is widely used in the robotics community to model agent interaction zones. Using graph measures, conditions under which the two models generate comparable networks are presented.**

I. INTRODUCTION

European starlings, *Sturnus vulgaris*, display remarkable coordination while flocking and a single flock can contain up to tens of thousands of individual birds. Ballerini *et al.* [1] analyzed data on the $3D$ positions of starling flocks and report that each individual bird coordinates, on average, with its nearest 6–7 neighbors. This communication model, where individuals interact with a fixed number of nearest neighbors, is referred to as the topological model, and the number of nearest neighbors is known as the topological distance [1, 2].

A “central challenge” in the study of collective behavior of animals is revealing how information flows within the group [3]. A number of communication models have been proposed in the biological swarm literature to describe interactions between individuals, including the metric and visual models. The metric model is directly based on spatial proximity: two individuals interact if they are within a metric range from each other (e.g., [4]). Strandburg–Peshkin *et al.* [3] define the visual model, where an individual only coordinates with swarm members that are within its visual field.

Consider a remotely deployed artificial swarm, capable of implementing the biologically inspired topological communication model. While it is known that starlings interact with their nearest 6–7 neighbors, it is unclear what the “right” topological distance is for a tasked artificial swarm. The distance can depend on the swarm’s size, and to this end, experiments were conducted with a simulated artificial swarm searching for a goal area. The first contribution of this paper is to present the swarm’s performance for all possible topological distances. Maintaining communication links is computationally costly; therefore, an important application of the reported research is to conserve computational resources by automatically tailoring a distance to a given task.

Each of the prescribed biologically-inspired communication models results in an interaction network, where nodes are swarm agents. One question that arises is how to relate communication models to each other by analyzing their resulting interaction networks. The second contribution of this work lays the foundations of relating the topological model to the metric model. Certain graph measures, such as the clustering coefficient, are used to determine similarity, and for each of the measures, an expression for the topological distance is produced in terms of the metric range to show the relationship between these two model parameters. Continuing with the remotely deployed swarm example, if the swarm is capable of selecting between the topological and metric models, and the swarm is stationary, the derived expressions can allow setting the model parameters, when switching from one communication model to the other, without affecting the network, in terms of the measures.

The paper is organized as follows: Section II provides related work. The search for a goal experiments are presented in Section III. Section IV investigates conditions under which the topological and metric models generate comparable networks in terms of certain graph measures. Concluding remarks are drawn in Section V.

II. RELATED WORK

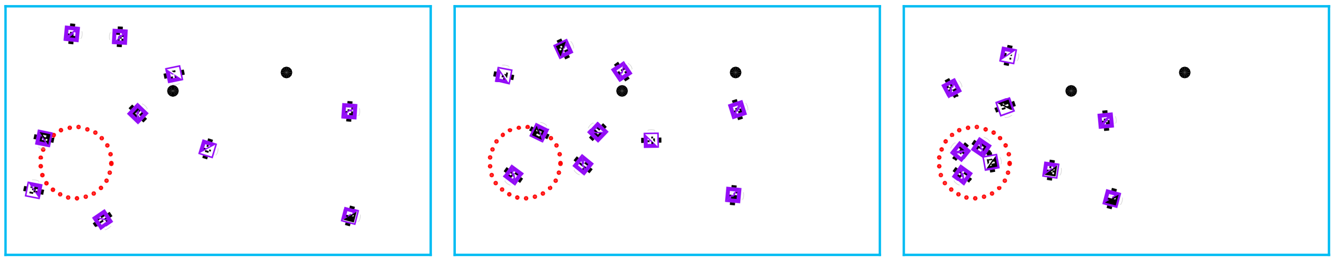
Haque *et al.* [5] compared the performances of metric, topological, and visual models for simulated artificial swarms searching for a goal. The experimental topological distance was restricted to the range 5–8 in order to mimic what is observed in starlings. The main performance metric was the percentage of agents reaching goal, and for the topological model, no statistically significant difference was found across the four topological distances. The reported search for a goal experiments in this paper focus solely on the topological model and present the performance of simulated swarms for the entire range of feasible topological distances, given a particular size of the swarm. Differences in the experimental design between the two experiments are discussed.

The influence of the topological distance has been studied using both graph- and system-theoretic frameworks. Komareji and Bouffanais [6] studied the relationship between group size and the topological distance for a simulated swarm. The authors showed that there exists a minimum topological distance, dependent on the swarm’s size, that is necessary to maintain the connectedness of the resulting

This material is based upon research supported by, or in part by, the U.S. Office of Naval Research under award #N000141210987.

Authors are with Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN 37212, USA

{musad.a.haque, waseem.abbas, abigail.d.rafter, julie.a.adams}@vanderbilt.edu



(a) $t = 13$ (b) $t = 100$ (c) $t = 200$

Fig. 1. Simulation of a swarm searching for a goal area using the Robotarium simulator. The center of the large circle denotes the location of the goal and black dots denote obstacles. This particular trial used the following parameter settings: $N = 10$, $N_{obs} = 2$, $r_r = 0.18$, $r_o = 0.36$, $r_a = 0.54$, $n_T = 9$.

interaction network. Shang and Bouffanais [7] proved that irrespective of the swarm's size, the rate of convergence of the consensus on agents' headings using a topological distance of 10 yields results close to the all-to-all communication case. System-theoretic analysis has further shown that the topological distance that optimizes a swarm's robustness does not depend on the swarm's size [8]. Robustness was formulated as the swarm's ability to reach a consensus on agent headings in the presence of uncertainty, and it was shown that the optimal topological distance needed to maximize robustness was 7.

III. Search for a Goal

The simulated artificial swarm's objective for this experiment is to discover a goal area in the environment. This task was one of the two tasks evaluated by Haque *et al.* [5] to compare the performances of the metric, topological (with topological distances around 7), and visual models of biological swarm communication. The current focus is to study the influence of the topological distance on the swarm's performance on this task.

A. Experimental design

The MATLAB-based Robotarium simulator [9] was used to conduct the experiments. The use of this simulator marks a departure from the prior experiment [5], and is motivated by the fact that simulator code can be directly deployed onto the Robotarium, a remote, multi-robot testbed consisting of GRITSBots [9]. The simulator, shown in Fig. 1, closely approximates the actual testbed and behavior of the GRITS-Bots. The simulator's environment is 1.20×0.70 and the dimension of the agents are 0.06×0.06.

The primary independent variable is the topological distance (n_T) and the performance metric measures the number of agents that have reached goal as a percentage of the swarm's size (N). The environment includes a number of obstacles, denoted by N_{obs}.

Swarming is achieved in a two-step process [5]. At each iteration t, the topological model assigns n_T neighbors to all the agents, and each agent interacts with those neighbors based on Reynold's repulsion-orientation-attraction rules for boids [10]: 1) veer away from neighbors that are within the repulsion range (r_r), 2) align headings with neighbors that are between the repulsion range and the orientation range (r_o), and 3) head toward the centroid of neighbors that are between the orientation range and the attraction range (r_a).

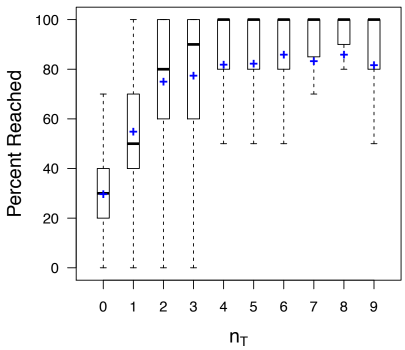
TABLE I
Experimental Design.

Parameter	Values Explored
Number of agents (N)	$\{10, 20, 40\}$
Number of obstacles (N_{obs})	$\{0, 0.1N, 0.2N\}$
Radius of repulsion (r_r)	$\{0.06, 0.12, 0.18\}$
Radius of orientation (r_o)	$\{1.1r_r, 1.5r_r, 2r_r\}$
Radius of attraction (r_a)	$\{1.1r_o, 1.5r_o, 2r_o\}$
Topological distance (n_T)	$\{0, 1, \ldots, N-1\}$

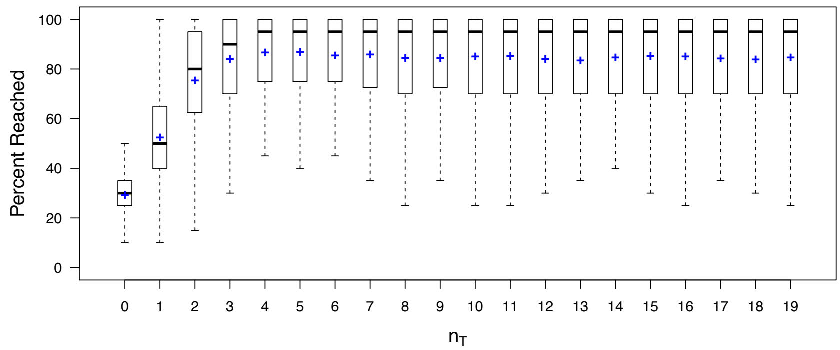
Table I summarizes the independent and dependent variables. An experiment is defined as a particular setting of the tuple (N, N_{obs}, r_r, r_o, r_a, n_T). A simulation run of 750 iterations in t is referred to as a trial. Twenty-five trials were completed for each experiment. The number of trials for the $N = 10$, $N = 20$, and $N = 40$ cases were $1\times3\times3\times3\times3\times10\times25 = 20250$, $1\times3\times3\times3\times3\times20\times25 = 42525$, $1 \times 3 \times 3 \times 3 \times 3 \times 40 \times 25 = 81000$, respectively; thus, the total number of trials was 143775.

Various coordinated movement patterns [11] are simulated for the swarm through the use of the different repulsion-orientation-attraction configurations. Individual agents are not intelligent, in the sense that they are not using any intricate search technique to locate the goal area. Instead, an agent swarms with its neighbors and senses the goal area if it is within r_a from goal's center. An agent can also be informed of the goal's location by an informed neighbor. An agent's effort to beeline toward goal's location is equally weighed with the effort to swarm based on the swarming rules, as was used by Couzin *et al.* [4].

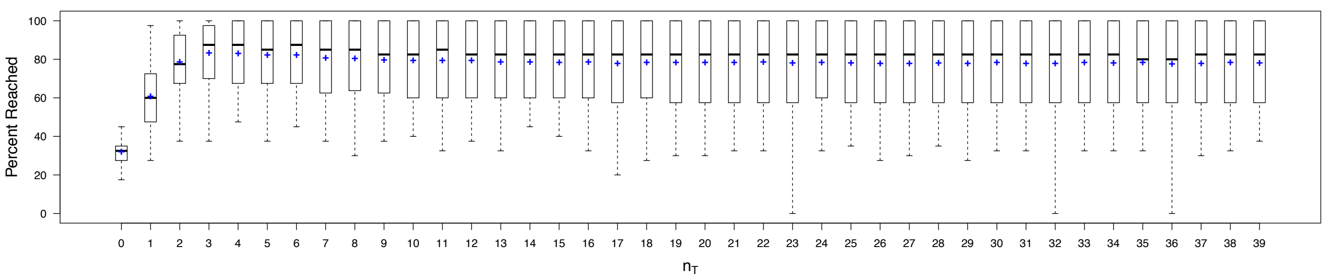
Hypothesis H_{T1} is that the mean performance will increase in the topological distance and plateau beyond a critical value of n_T. H_{T1} is based on the fact that Komareji and Bouffanais [6] proved that for a swarm where agents move by aligning headings with neighbors, there exists a minimum topological distance necessary to maintain the swarm's connectedness. The dynamics of the swarm simulated in this work are based on a more involved interaction scheme of repulsion, orientation, and attraction. Moreover,



(a) $N = 10$



(b) $N = 20$



(c) $N = 40$

Fig. 2. The performance data for the three swarm sizes. The horizontal line inside a box plot represents the median and the cross represents the mean.

TABLE II

VALUES OF $n_T^\star$ AND n_T' BY THE NUMBER OF SWARM AGENTS.

N	$n_T^\star$	n_T'
10	8 (0.800N)	4 (0.40N)
20	5 (0.250N)	3 (0.15N)
40	3 (0.075N)	2 (0.05N)

informed agents do not break from the swarm and dart toward the goal's location. The results presented in the next section demonstrate whether a critical distance also exists for the search for a goal task, where agent-agent interactions go beyond striving for a consensus on headings.

Hypothesis H_{T2} is that the critical distance for the designed experiment's performance measure (percentage of agents reaching goal), if it exists, will not be an absolute value like the one identified for robustness [8], but will depend on the swarm's size [6].

B. Results

The results of Anderson-Darling tests indicated that the performance was normally distributed for $N = 10$ ($A = 288.59$, p-value < 0.001), $N = 20$ ($A = 382.46$, p-value < 0.001), and $N = 40$ ($A = 424.89$, p-value < 0.001).

The distribution of performance data using box plots is shown in Fig. 2. The highest **mean** performance for the three cases were 85.78% for $N = 10$, 86.87% for $N = 20$, and 83.29% for $N = 40$. The topological distance that achieved the highest mean is denoted by $n_T^\star$, as presented in Table II.

An analysis of variance (ANOVA) determined that n_T had a significant impact on performance for $N = 10$ ($F(9,20241)=575.10$, $p<0.001$), $N = 20$ ($F(19,42506)=654.20$, $p<0.001$), and $N = 40$ ($F(39,80961)=95.01$, $p<0.001$). Further, Fisher's LSD tests found that for each case there exists a critical topological distance, denoted by n_T', such that there is no statistically significant difference between the performances achieved by $n_{T,1}$ and $n_{T,2}$, if $n_{T,1}, n_{T,2} \geq n_T\prime$. Table II shows the values n_T' for each of the three cases.

The **median** performances for $N = 10$ is 100% for all $n_T \geq 4$, as shown in Fig. 2(a). The value of n_T' is also 4 for this case (see Table II). This result indicates that the median performance follows a pattern similar to the mean, and generates similar performances for all $n_T \geq n_T'$.

The **interquartile ranges** of the performance data, the range of values between the first and third quartile, are also similar for all $n_T \geq n_T'$. The third quartiles of the performance data for the case of $N = 20$ (see Fig. 2(b)) have a value of 100% (i.e., all the swarm agents are reaching goal) for all topological distances greater than 3, which is also the value of the critical topological distance.

C. Discussion

Hypothesis H_{T1} was supported: There exists a critical topological distance, beyond which, communication with more neighbors does not yield any statistically significant improvement in the percentage of agents reaching goal. Thus, the notion of a critical topological distance described by Komareji and Bouffanais [6], where agents align their headings with neighbors, exists for the search for a goal setting as well, with a more complex movement law described by a repulsion-orientation-attraction scheme. The implication of finding the critical topological distances for a swarm robotics tasks allows artificial swarms to conserve computational resources by only maintaining links to the nearest n'_T neighbors.

A critical topological distance does not depend on the swarm's size in terms of coordination [1], rate of convergence, [7], and robustness [8]. The critical topological distance reported in this work for the percentage of agents reaching goal depends on the swarm's size, and as such, H_{T2} was supported. The results suggest that the critical distance's value is exponentially decreasing in N. The critical topological distances ($n'_T = 2$ for $N = 40$) help explain why there was no statistically significant difference in performance across the topological distances of $5 - 8$ explored for 50, 100, and 200 swarm agents [5].

There are task-specific conditions where the use of the topological model is preferred over the metric model [5]. However, the metric model, which is known in the multi-agent robotics community as the Δ-disk model (e.g., see [12]), is widely used due to its ability to characterize sensor range limitations. An open research question is how can the two models be compared at a more fundamental level? What is the relationship between the two model parameters, namely, the metric range and the topological distance? An analysis of these parameters, presented in the next section, guides the selection of a topological distance, given a metric range, or vice versa, in a way that both model parameters produce interaction networks comparable to each other. The results are based on simulated static agents (or nodes) in an attempt to find a relationship between the two model parameters. Understanding the relationship between the model parameters for dynamic swarm agents remains an open research question.

IV. Comparisons of Models

There can be multiple approaches for comparing the metric and topological models. A graph-theoretic approach is adopted, in which the values of the metric range (r_M) and the topological distance (n_T) are identified for which the corresponding graphs provide an identical value for a particular graph measure. Three graph measures are considered: size of the minimum dominating set, edge-connectivity, and clustering coefficient. The three measures are widely used to indicate centrality, connectedness, and clustering.

Numerical evaluations are based on the simulation of N nodes distributed uniformly at random in a 10×10 area. The values of the graph measures are obtained for different values of r_M and n_T. Each point on the plots in Figs. 3, 4, and 5 is an average of fifty randomly generated instances.

A. Minimum dominating set

A central issue in monitoring, controlling, or spreading information within a network is determining a small subset of nodes through which every node can be accessed in the most efficient manner. Thus, the goal is to find the most influential or dominating nodes within the network. Along with many other centrality notions, the concept of a *dominating set* of nodes captures this idea.

Definition 4.1: (*Dominating set*) In a network modeled by a graph $G(V, E)$, a subset of nodes $D \subseteq V$ is a dominating set if every node in V is either included in a dominating set, or is adjacent to a node in a dominating set.

The significance of dominating sets of nodes within networks has been exploited in many different contexts, including network controllability, optimal resource allocation, and social influence propagation (e.g., see [13]–[15]). The goal is to determine a dominating set of nodes of the minimum cardinality, which is an NP-hard problem. Efficient approximation algorithms can compute the minimal dominating sets in a distributed manner [16].

The number of nodes in the minimal dominating sets are plotted as functions of r_M and n_T for various network sizes N (see Fig. 3). Then, to relate metric and topological models in the context of minimal dominating sets, the values of r_M and n_T that give the minimal dominating sets of the same sizes in the corresponding graphs are calculated. The results are illustrated in Fig. 3(c). Using extensive simulations, an approximate relationship between r_M and n_T is found. Given r_M, the value of n_T that produces the minimal dominating set of same size, is approximately given by:

$$n_T \approx 2.95(N/A)(r_M)^{1.85} \tag{1}$$

Note that the above relation holds for the graphs obtained, as per the metric and topological models by distributing N nodes uniformly at random in an area A.

B. Edge-connectivity

Connectivity is a primary attribute of any networked system. Various metrics have been proposed in literature to quantify the level of inter-connectedness between nodes; however, the notions of *edge-connectivity* (or *node-connectivity*) are widely used in this regard.

Definition 4.2: (*Edge-connectivity*) A graph $G(V, E)$ is k-edge-connected if there does not exist a set of $k - 1$ edges whose removal disconnects the graph. Edge-connectivity is the maximum k for which the graph remains k-edge-connected.

There exists a path between any two nodes, in a k-edge connected graph, even if $k - 1$ edges are removed. A graph with a higher edge-connectivity is more resilient to edge failures or removals. The edge-connectivity as a function of r_M and n_T is illustrated in Figs. 4(a)-(b). As with the dominating sets, r_M and n_T are analyzed by fixing the edge connectivity. Fig. 4(c) illustrates the relationship between r_M and n_T given the same edge-connectivities. Using extensive simulations, an empirical relationship is obtained:

$$n_T \approx 2.15(N/A)(r_M)^2 \quad (2)$$

C. Clustering coefficient

Based on the network structure, nodes may tend to group together. Clusters can arise such that nodes are densely connected within the clusters and relatively few interactions exist between different clusters. The *clustering coefficient* is a widely-used measure to quantify clustering within a network (e.g., see [17, 18]).

Definition 4.3: (*Clustering coefficient*) Let u be a node within a network with N_u neighbors, and E_u be the number of edges that exist between the neighbors of u. The clustering coefficient of a node u, denoted by $C(u)$, is

$$C(u) = \frac{2E_u}{N_u(N_u - 1)} \quad (3)$$

The total number of possible edges that can exist between the neighbors of u is $N_u(N_u - 1)/2$. The clustering coefficient is the average of clustering coefficients of all the nodes, i.e.,

$$C(G) = \frac{1}{N}\sum_{u \in G} C(u). \quad (4)$$

The clustering coefficient of a node with less than two neighbors is considered to be 0. The directed nature of links is not considered [3]. The clustering coefficient as a function of r_M and n_T is shown in Fig. 5. Again, a relationship between r_M and n_T is explored, i.e., for a given r_M in the metric model, what is the value of n_T in the topological model that results in the same clustering coefficient. The approximate relationship between r_M and n_T is given by:

$$n_T \approx 3.5(N/A)(r_M)^{1.5} \quad (5)$$

D. Discussion

The rates at which the graph measures vary with the respective model parameters are different for the metric and topological communication models. Moreover, for each of the three network properties – centrality, connectedness, and clustering – there exist a relationship between the parameters of the two models. The relationships for all three measures are captured by a power function of the form $n_T = c(r_M)^p$. However, constants c and p are different for each of these measures. As such, obtained relationships between r_M and n_T remain specific to their associated network properties, and only generalize in form.

V. Conclusions

Experiments were conducted using the Robotarium simulator where an artificial swarm, using the topological communication model and a repulsion-orientation-attraction interaction rule was tasked with searching for a goal area. It is shown that there exists a topological distance, dependent on the swarm's size, beyond which no statistically significant improvement is gained in the percentage of a swarm's agents reaching the goal.

Using extensive simulations of stationary nodes uniformly distributed at random, networks formed by the topological and metric models were compared to each other in terms of three graph measures: the dominating set, edge-connectivity, and clustering coefficient. Expressions were derived, one for each measure, to tie the metric range and the topological distance such that the two model parameters produced comparable networks.

Combined together, these two contributions serve to aid artificial swarm design by conserving computational resources through the critical topological distance and guiding how to set model parameters when selecting between the topological model, based on European starlings, and the metric model, which is widely used in the swarm-robotics community.

Acknowledgment

We thank Daniel Pickem, Paul Glotfelter, and Magnus Egerstedt for facilitating our use of the simulator.

References

[1] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic, "Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study," *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1232–1237, 2008.

[2] N. Abaid and M. Porfiri, "Fish in a ring: spatio-temporal pattern formation in one-dimensional animal groups," *J. of the Royal Society Interface*, vol. 7, pp. 1441–1453, Jun. 2010.

[3] A. Strandburg-Peshkin, C. R. Twomey, N. W. Bode, A. B. Kao, Y. Katz, C. C. Ioannou, S. B. Rosenthal, C. J. Torney, H. S. Wu, S. A. Levin, and I. D. Couzin, "Visual sensory networks and effective information transfer in animal groups," *Current Biology*, vol. 23, no. 17, pp. R709–R711, 2013.

[4] I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin, "Effective leadership and decision-making in animal groups on the move," *Nature*, vol. 433, pp. 513–516, Feb. 2005.

[5] M. Haque, C. Ren, E. Baker, D. Kirkpatrick, and J. Adams, "Analysis of swarm communication models," in *Proc. 22nd European Conf. on Artificial Intelligence*, Aug. 2016.

[6] M. Komareji and R. Bouffanais, "Resilience and controllability of dynamic collective behaviors," *PLoS one*, vol. 8, no. 12, 2013.

[7] Y. Shang and R. Bouffanais, "Influence of the number of topologically interacting neighbors on swarm dynamics," *Scientific Reports*, vol. 4, Feb. 2014.

[8] G. F. Young, L. Scardovi, A. Cavagna, I. Giardina, and N. E. Leonard, "Starling flock networks manage uncertainty in consensus at low cost," *PLoS Computational Biology*, vol. 9, no. 1, p. e1002894, 2013.

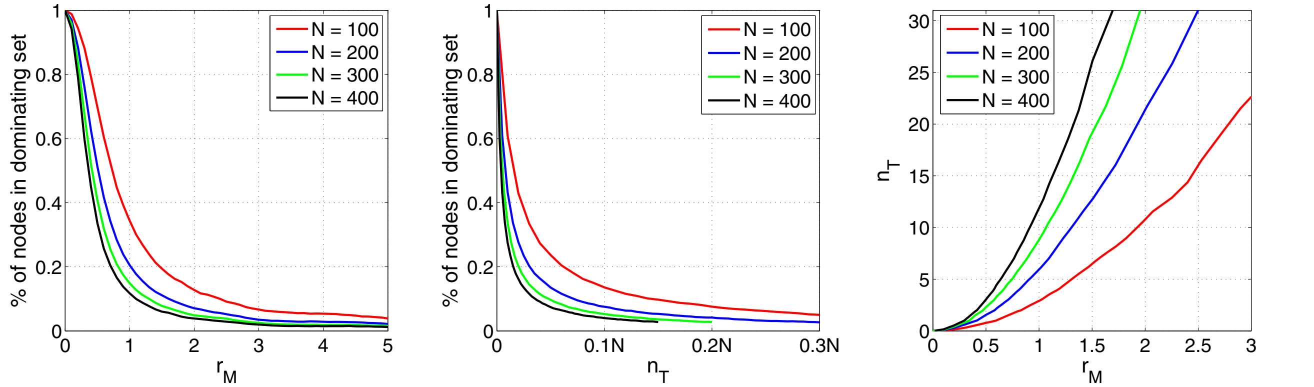
[9] D. Pickem, L. Wang, P. Glotfelter, Y. Diaz-Mercado, M. Mote, A. Ames, E. Feron, and M. Egerstedt, "Safe, remote-access swarm robotics research on the robotarium," *Computing Research Repository*, vol. abs/1604.00640, 2016.

[10] C. Reynolds, "Flocks, herds and schools: A distributed behavioral model," *Computer Graphics*, vol. 21, pp. 25–34, Jul. 1987.

[11] I. D. Couzin, J. Krause, R. James, G. D. Ruxton, and N. R. Franks, "Collective memory and spatial sorting in animal groups," *J. of Theoretical Biology*, vol. 218, no. 1, pp. 1–11, 2002.

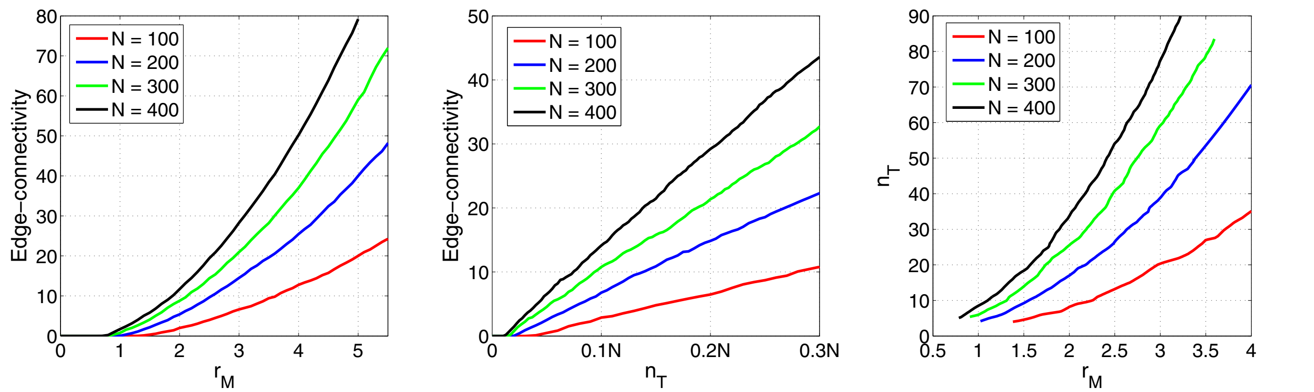
[12] M. Ji and M. Egerstedt, "Distributed coordination control of multiagent systems while preserving connectedness," *IEEE Transactions on Robotics*, vol. 23, pp. 693–703, Aug. 2007.

[13] N. J. Cowan, E. J. Chastain, D. A. Vilhena, J. S. Freudenberg, and C. T. Bergstrom, "Nodal dynamics, not degree distributions, determine the structural controllability of complex networks," *PloS one*, vol. 7, no. 6, p. e38398, 2012.



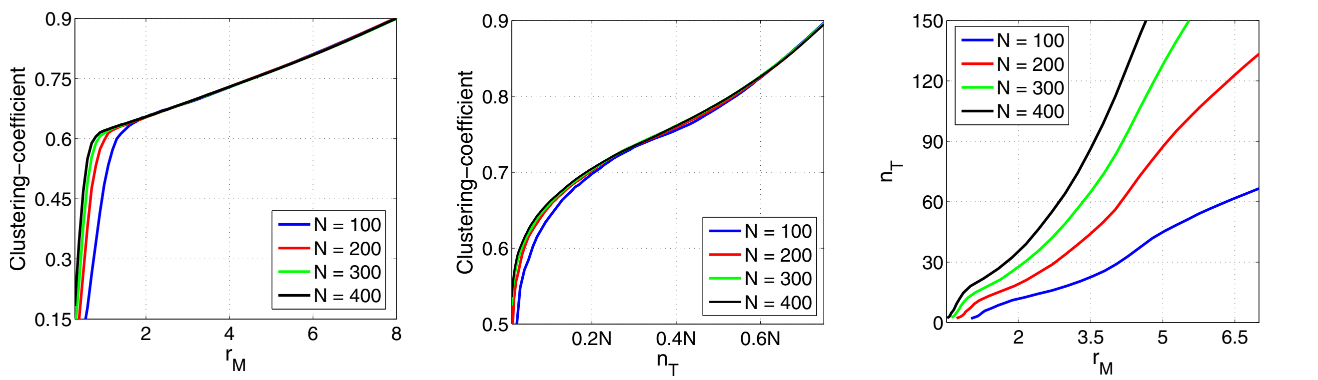
(a) (b) (c)

Fig. 3. (a)-(b) The minimal dominating set as a function of r_M and n_T. (c) n_T as a function of r_M for minimal dominating sets of same cardinalities.



(a) (b) (c)

Fig. 4. (a)-(b) Edge-connectivity as a function of r_M and n_T. (c) n_T as a function of r_M for the same connectivities.



(a) (b) (c)

Fig. 5. (a)-(b) Clustering coefficient as a function of r_M and n_T. (c) n_T as a function of r_M for the same clustering coefficients.

[14] L. L. Kelleher and M. B. Cozzens, "Dominating sets in social network graphs," *Mathematical Social Sciences*, vol. 16, no. 3, 1988.

[15] W. Abbas and M. Egerstedt, "Distribution of agents in heterogeneous multiagent systems," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pp. 976–981, IEEE, 2011.

[16] T. W. Haynes, S. Hedetniemi, and P. Slater, *Fundamentals of domination in graphs*. CRC Press, 1998.

[17] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world'networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[18] M. E. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.