

Representing Multi-Robot Structure through Multimodal Graph Embedding for the Selection of Robot Teams

Brian Reily¹, Christopher Reardon², and Hao Zhang¹

Abstract—Multi-robot systems of increasing size and complexity are used to solve large-scale problems, such as area exploration and search and rescue. A key decision in human-robot teaming is dividing a multi-robot system into teams to address separate issues or to accomplish a task over a large area. In order to address the problem of selecting teams in a multi-robot system, we propose a new multimodal graph embedding method to construct a unified representation that fuses multiple information modalities to describe and divide a multi-robot system. The relationship modalities are encoded as directed graphs that can encode asymmetrical relationships, which are embedded into a unified representation for each robot. Then, the constructed multimodal representation is used to determine teams based upon unsupervised learning. We perform experiments to evaluate our approach on expert-defined team formations, large-scale simulated multi-robot systems, and a system of physical robots. Experimental results show that our method successfully decides correct teams based on the multifaceted internal structures describing multi-robot systems, and outperforms baseline methods based upon only one mode of information, as well as other graph embedding-based division methods.

I. INTRODUCTION

Because of their robustness and flexibility, multi-robot systems are being increasingly researched and used in large-scale applications, such as disaster response [1], search and rescue [2], and area exploration [3]. Accomplishing complex operations using multi-robot systems typically requires the division of a system into effective teams that are capable of achieving multiple separate tasks simultaneously or accomplishing mission objectives over a large area [4]. However, as the number of robots in a system increases, it becomes cognitively more difficult for humans to understand and command [5].

Successful division of a large multi-robot system into effective teams is a particularly difficult problem as the size and complexity of the interaction between robots increases. At scale, the complexities of internal relationships become difficult for human operators to conceptualize and integrate [6]. Figure 1 provides an illustration of how robots can appear organized in physical space, but also contain hierarchical relationships or communication capabilities within a system that are more difficult to perceive. These relationships are further complicated by robot interactions with obstacles

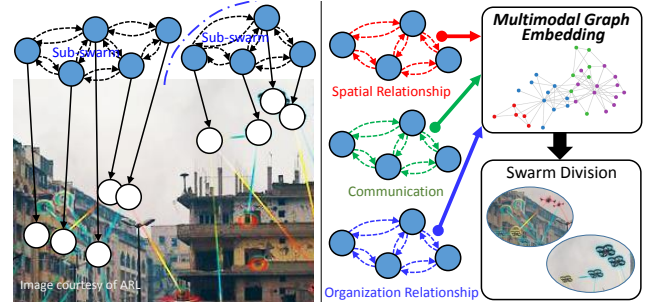


Fig. 1. A motivating example of automatic multi-robot division and our solution based on multimodal graph embedding. In real-world multi-robot systems, robots typically have a variety of relationships, such as spatial relationships, communication connectivity, and organization hierarchy. Our proposed multimodal graph embedding approach can integrate multiple relationship graphs and identify effective multi-robot teams.

and the surrounding environment. When combined, these challenges result in robotic system states that are both difficult for a human operator to accurately perceive and for a system to display [7]. Thus, effective representations of multi-robot structure are critical for understanding and selecting team divisions.

Due to the importance of multi-robot teams, division approaches have been widely studied in the literature. A commonly used paradigm approached the division of a multi-robot system based solely on spatial locations of members, assigning territories to each robot [8], [9]. These methods have the drawbacks of relying on spatial information only and assigning agents to individual territories, without the ability to identify teams or grouped divisions. Another paradigm views multi-robot division as a task allocation problem, e.g., based on task sequencing [10], [11], game theory [?], [12], and Markov chains [13]. These techniques require explicit knowledge about the tasks, such as task sequences, requirements, or priorities, as well as knowledge of opponent behaviors, which may not be available to a human operator. In addition, methods were also implemented that are inspired by biological swarms such as ants [14], [15] or wasps [16]. These techniques often assume simple agents, often without the ability to communicate, and are not capable of creating a complete representation of multi-robot structure by addressing multiple relationship modalities.

To address the selection of multi-robot teams, we propose a novel multimodal graph embedding approach to encode diverse relationships of multiple robots as graphs and integrate these multiple graphs into a unified representation that is applied to divide a multi-robot system into teams. We model each internal relationship of the robots using a

*This work was partially supported by ARL DCIST CRA W911NF-17-2-0181, ARO W911NF-17-1-0447, and NSF CNS-1823245.

¹Brian Reily and Hao Zhang are with the Human-Centered Robotics Laboratory at the Colorado School of Mines, Golden, CO, 80401, USA. email: {breily, hzhang}@mines.edu

²Christopher Reardon is with the U.S. Army Research Laboratory in Adelphi, MD, 20783, USA. email: christopher.m.reardon3.civ@mail.mil

directed graph as an information modality. Given a set of member relationships, we construct multiple graphs that are applied as the input to our approach. Then, we propose a new multimodal Katz index to integrate multiple graphs of robot relationships and embed them into a unified representation for each robot. Then, the constructed representation is used to identify multi-robot teams through unsupervised learning. Our multimodal graph-embedded robot division approach is capable of fusing diverse robotic relationships and identifying team divisions without requiring explicit knowledge of tasks.

This paper has two major contributions.

- 1) First, we propose a novel multimodal graph embedding approach that formulates robot team selection as a graph-based embedding and clustering problem, models multiple asymmetrical relationships of robots using directed graphs, and fuses them into a unified representation to identify robotic teams.
- 2) Second, we perform a comprehensive evaluation of the proposed method in the scenarios using expert-defined team formations, large-scale simulated multi-robot systems, and teams of physical robots. We validate that our multimodal graph embedding approach is effective in identifying robotic teams, and obtains superior performance over existing graph-based division methods.

II. RELATED WORK

In this section, we review previous approaches to multi-robot division, representation of multi-robot systems as graphs, and existing graph embedding methods.

A. Multi-Robot Division

Dividing multi-robot systems to accomplish multiple objectives simultaneously has been approached previously from three main perspectives: spatial-based division, task allocation, and biologically inspired methods.

Spatial-based division was implemented by assigning robots to their own territories [8], [9]. It was typically formulated as an optimization problem, where a group of robots diffuses to optimize coverage of an area [17]. This spatial-based perspective has two main drawbacks. First, it relies solely on the spatial locations of robots, omitting other information available that could be utilized to better divide the system. Second, it assigns members to areas individually, and is unable to divide a multi-robot system into teams, each still capable of working together.

Task allocation was implemented for multi-robot division in a variety of ways. Explicit assignment was designed based on sub-goals [18], or the temporal relationships between tasks [10], [11]. Markov chains were also developed to schedule task assignments [13]. Agents were designed to negotiate with each other for sub-tasks [19], and game theory was applied to find optimal task assignments [?], [12]. Multi-robot division based on task allocation has the major drawback of requiring explicit information about the task. This can range from requiring task dependency information in order to schedule sequences, task composition in order

to assign sub-tasks, or task priorities and requirements in order to determine the best robot to assign the task to. Many approaches also assign tasks to individual robots, as opposed to tasking groups to accomplish a single task.

Methods inspired by biological swarms were also designed for multi-robot division, inspired by how insects divide into sub-groups. Methods were introduced based on ant foraging [15] and division of labor in wasp colonies [16] and ant colonies [14]. Biological approaches generally assume swarm agents are extremely limited in their ability to sense, communicate, and process information. Previous bio-inspired methods cannot incorporate multimodal relationships of swarm members.

B. Graph Representations of Multi-Agent Systems

Many previous works used graphs to represent multi-agent systems, by representing agents in the team with vertices and relationships between agents as edges. Control laws were first applied as edges between agents, where these laws controlled spacing between agents and were applied to construct group trajectories around obstacles [20]–[22]. Graphs that described multiple nearest neighbors through control laws were applied in [23] and [24]. Communication channels between robots were represented graphically in [25]. Graph rigidity was used in [26], [27] in an optimization-based method to perform split and rejoin maneuvers around obstacles.

Methods to find communities in graphs are mainly based on analyzing edges between vertices. Cut-based methods cut the graph to form the best two clusters, and include min-max cuts [28] and normalized cuts [29]. Communities based on the probability of new links were used in [30], later extended with the idea of modularity [31], [32], which constructs a segmentation metric and uses this to create divisions. These community-finding algorithms are unable to operate on multiple graphs, and often cannot divide a graph into an arbitrary number of communities but instead rely on the structure of the graph to define the number of communities.

C. Graph Embedding

Graph embedding is the representation of graph structure in vector space, allowing typical machine learning techniques to be applied [33]–[35]. Techniques including node2vec [36] employ random walks, representing nodes based upon their transition probabilities to and from other nodes. Methods based on matrix descriptions of graphs, such as adjacency matrices and similarity matrices, include structure preserving embedding [37], graph factorization [38], and High-Order Proximity preserved Embedding (HOPE) [39]. Most graph embedding techniques focused on single graphs representation, e.g., to encode social networks or academic citation networks, without the ability to integrate multiple relationships. The effectiveness of graph embedding for multi-robot division has also not well studied in the literature.

III. APPROACH

We propose a new multimodal graph embedding approach to integrate multiple directed graphs that encode relationships

in a multi-robot system into a unified representation that is used to identify multi-robot teams.

A. Multimodal Robotic Structure Embedding

Given a specified relationship (e.g., communication connectivity) of a multi-robot system including N members, we represent a relationship among robots as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_1, \dots, v_N\}$ denotes the set of vertices, each corresponding to a robot, and $\mathcal{E}$ denotes the set of directed edges between these vertices. The direction and weight of each edge $e_{ij} = (v_i, v_j) \in \mathcal{E}$ depends on the type of relation the graph is representing.

In real-world deployment, robotic members within a system typically have multiple various relationships (e.g., spatial relationships, communication connectivity, and organization hierarchy). When multiple robotic relationships are available, we encode the system with M graphs, where $\mathcal{G}_m$ is the graph describing the m -th relationship between robots. Each graph $\mathcal{G}_m$ is described by an adjacency matrix $\mathbf{A}_m \in \mathcal{R}^{N \times N}$, where each element a_{ij} is the weight of the edge connecting vertex v_i to vertex v_j . The proposed method is able to represent both directed (e.g. communication from one member to others) and undirected (e.g., spatial distances) relationships. In the case of representing undirected relationships, the weight of edges satisfies $a_{ij} = a_{ji}$.

To represent graphs in vector space (i.e., graph embedding), the Katz index [40] has shown to be a promising method that has been used to address real-world graph problems. The Katz index represents the similarity of a pair of vertices given a graph, by summing the weights of edges along all paths between the two vertices to represent how closely connected the two vertices are. Given a decay parameter $\alpha > 0$ that weights paths based on their length L , the Katz index can produce a similarity matrix $\mathbf{S}$ to describe a graph:

$$\mathbf{S} = \sum_{l=1}^L \alpha^l \mathbf{A}^l \quad (1)$$

This equation can be rewritten as

$$\mathbf{S} = (\mathbf{I} - \alpha \mathbf{A})^{-1} - \mathbf{I} \quad (2)$$

where $\mathbf{I}$ is the identity matrix.

To achieve our objective of encoding multiple graphs and embedding them into a single vector representation, we propose a new multimodal Katz index that is able to take multiple graphs as the multimodal input modalities and form a single similarity matrix $\mathbf{S} \in \mathcal{R}^{N \times N}$ that integrates the information of all graphs. To do this, we implement an adjacency matrix $\mathbf{A}_m$ for each graph $\mathcal{G}_m$, where $m = 1, \dots, M$, and introduce a weight w_m for each $\mathbf{A}_m$ describing the importance of the relationship encoded by $\mathcal{G}_m$, where $1 = \sum_{m=1}^M w_m$. Values of w_m can be specified by human experts to incorporate prior knowledge and preferences, or estimated by hyperparameter selection methods (e.g., grid search). We then construct the multimodal similarity matrix

Algorithm 1: Graph Embedding for Multi-Robot Div.

Input : $\mathcal{G}_m = (\mathcal{V}, \mathcal{E})$ and C

Output: $\mathbf{X}$ and $\mathbf{x}_n^{(c)}$

- 1: **for** $m \leftarrow 1$ **to** M **do**
 - 2: Construct adjacency matrix $\mathbf{A}_m$ from $\mathcal{G}_m$.
 - 3: **end**
 - 4: Calculate similarity matrix $\mathbf{S}$ by Equation (3), integrating each $\mathbf{A}_m$ based on its weight w_m .
 - 5: Perform SVD to obtain left singular vectors $\mathbf{U}$ and right singular vectors $\mathbf{V}^T$.
 - 6: Construct multi-robot representation matrix $\mathbf{X}$ from the first K columns of $\mathbf{U}$ and the first K columns of $\mathbf{V}$.
 - 7: Randomly select C cluster centroids Φ .
 - 8: **repeat**
 - 9: Assign each robot representation $\mathbf{x}_n$ a label $\mathbf{x}_n^{(c)}$ based on the nearest cluster centroid ϕ_c .
 - 10: Update cluster centroids Φ based on the labels.
 - 11: **until** Labels have converged;
 - 12: **return** $\mathbf{X}$ and $\mathbf{x}_n^{(c)}$
-

$\mathbf{S}$ that embeds information of all graphs as follows:

$$\mathbf{S} = \left(\mathbf{I} - \alpha \sum_{m=1}^M w_m \mathbf{A}_m \right)^{-1} - \mathbf{I} \quad (3)$$

Proposition 1: $\left(\mathbf{I} - \alpha \sum_{m=1}^M w_m \mathbf{A}_m \right)$ is invertible when α^{-1} is greater than the largest eigenvalue of $\mathbf{A}^*$, where $\mathbf{A}^* = \sum_{m=1}^M w_m \mathbf{A}_m$.

Proof: For $(\mathbf{I} - \alpha \mathbf{A}^*)$ to not be invertible, $\det(\mathbf{I} - \alpha \mathbf{A}^*) = 0$. This is equivalent to $\det(\mathbf{A}^* - \alpha^{-1} \mathbf{I}) = 0$. This determinant is only zero if α^{-1} is equal to the largest eigenvalue of $\mathbf{A}^*$, denoted as e_1 . When $\alpha < \frac{1}{e_1}$, the inverse is valid. ■

In order to create a lower-dimensional representation than $\mathbf{S}$, which is necessary when embedding big graphs of a large-scale multi-robot system, we perform Singular Value Decomposition (SVD) [41] to factor $\mathbf{S}$:

$$\mathbf{S} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \quad (4)$$

The columns of $\mathbf{U}$ are the left singular vectors of $\mathbf{S}$, the rows of $\mathbf{V}^T$ are the right singular vectors, and $\mathbf{\Sigma}$ is a diagonal matrix whose entries are the singular values. As $\mathbf{S}$ is a square $N \times N$ matrix, all of these are $\in \mathcal{R}^{N \times N}$ when a full SVD is performed. The columns of $\mathbf{U}$ are also the eigenvectors of $\mathbf{S} \mathbf{S}^T$, and similarly the rows of $\mathbf{V}^T$ are the eigenvectors of $\mathbf{S}^T \mathbf{S}$.

To further reduce the dimensionality of the representation, given a desired dimensionality D , we propose to use the first K left singular vectors and the first K right singular vectors to approximate $\mathbf{S}$, where $K = D/2$, half of the desired dimensionality. As we only use the first K singular vectors on each side, we can perform a reduced SVD of $\mathbf{S}$. This results in $\mathbf{U} \in \mathcal{R}^{N \times K}$ and $\mathbf{V} \in \mathcal{R}^{N \times K}$.

Finally, we construct the final representation matrix $\mathbf{X} \in \mathcal{R}^{N \times D}$ for all N robots by combining $\mathbf{U}$ and $\mathbf{V}$:

$$\mathbf{X} = (\mathbf{U}, \mathbf{V}) \quad (5)$$

where each row $\mathbf{x}_n \in \mathcal{R}^D$ of $\mathbf{X}$ represents the n -th robot, which is then used for division.

B. Selection of Multi-Robot Teams

Given the multimodal graph embedding results, we propose to perform selection of multi-robot teams through unsupervised learning (i.e., clustering). Specifically, given $\mathbf{X} = \{\mathbf{x}_1; \dots; \mathbf{x}_N\}$, and given the number C of robot teams the human operator needs to accomplish the mission, the goal is to assign each robot a cluster label $\mathbf{x}_n^{(c)}$ that indicates it belongs to team c , thus dividing the robotic system into C total teams.

This unsupervised learning problem can be formulated by:

$$\arg \min_{\mathbf{x}_n^{(c)}} \sum_{n=1}^N \|\mathbf{x}_n^{(c)} - \phi_c\|^2 \quad (6)$$

where $\Phi = \{\phi_1, \dots, \phi_C\}$ are the centroids of all divisions, initially chosen at random. The formulated optimization problem can be solved iteratively. First, labels are assigned to each robot $\mathbf{x}_n$. Then, each centroid ϕ_c is recalculated based on the robots assigned to that cluster. This process is repeated until the assigned cluster labels $\mathbf{x}_n^{(c)}$ converge.

The algorithmic implementation of both multimodal multi-robot structure embedding and multi-robot division is summarized in Algorithm 1.

TABLE I

COMPARISON OF ACCURACY AND SILHOUETTE SCORES IN SCENARIO I

Method	Accuracy	Avg. Silhouette Score
GN [31] (Spatial)	39.22%	-0.012
GN (Connectivity)	81.37%	0.434
GN (Hierarchy)	62.75%	0.354
LLE [42] (Spatial)	44.12%	0.138
LLE (Connectivity)	44.44%	0.082
LLE (Hierarchy)	39.22%	-0.031
HOPE [39] (Spatial)	85.62%	0.455
HOPE (Connectivity)	50.33%	0.018
HOPE (Hierarchy)	39.22%	0.167
HOPE (Concatenated)	93.14%	0.518
Our Approach	96.73%	0.680

IV. EXPERIMENTS

To evaluate the performance of the proposed multimodal graph embedding approach for selection of multi-robot teams, we conduct experiments using robotic systems at multiple scales. We evaluate the accuracy of multi-robot team selection and score the quality of the robot clusters. We also compare with previous graph-based division techniques.

A. Experimental Setup

Our approach was experimentally evaluated and validated in two different scenarios. First, we evaluate on *Scenario I: Expert-defined Team Formations*. This experiment is based on a set of expert-defined graphs that represent relationships

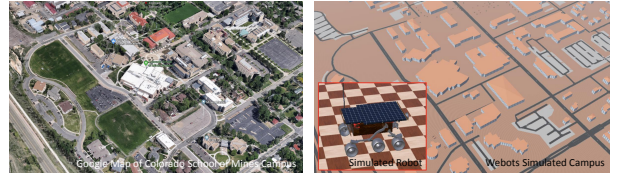
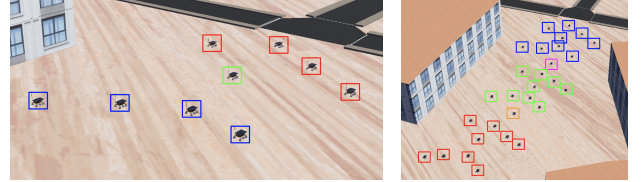


Fig. 2. Experimental setups in a robot-assisted urban search and rescue scenario used to evaluate our approach in Scenario I and Scenario II. The left figure depicts the Google satellite map of the real campus environment of the Colorado School of Mines. The right figure illustrates the simulated environment and a simulated robot in the Webots robot simulator.



(a) Squad Column

(b) Platoon Column

Fig. 3. Qualitative results of multi-robot team selection in Scenario I. In this scenario, the multi-robot system is approaching an intersection where division into teams may be necessary to effectively search down branching roads or to secure the intersection. Figure 3(a) displays the *squad column*, with colored bounding boxes identifying the correct division into three teams [43]. Figure 3(b) displays the *platoon column*, displayed with the correct division into five teams.

in field operations teams at two scales. Second, we evaluate on *Scenario II: Simulated Large-Scale Multi-Robot Systems*. We generate large-scale multi-robot systems to demonstrate that our method scales beyond the size of small groups. These simulated systems do not have ground truth teams, and instead our identified teams are scored by a clustering metric. We demonstrate our approach can be deployed in real-world robotic applications by deploying some of these simulated robotic systems as physical TurtleBot robots.

Simulations of the expert-defined teams and the large-scale multi-robot systems were performed with the Webots robot simulator [44] in an application of area exploration. We created a simulated environment of the Colorado School of Mines campus for the experiments. Figure 2 displays the real campus environment, its simulated Webots map and the simulated robot used. In the experiments on all scenarios, three types of graphs are used as input information modalities, describing spatial positions, known communication connectivity, and a set hierarchy. For quantitative evaluation, we present the clustering accuracy when dividing the expert-defined teams and utilize silhouette scores [45] for all robotic divisions. Silhouette scores rate the quality of a clustering, with values closer to 1 being better and values closer to -1 being worse.

To validate the superior performance of our approach, we compare it with previous methods: (1) *Girvan-Newman* (GN) [31], a single-modality graph community finding algorithm; (2) *Local Linear Embedding* (LLE) [42], a single-modality graph embedding approach; (3) *High-Order Proximity preserved Embedding* (HOPE) [39], a single-modality graph embedding approach; and (4) *Concatenated Combination HOPE*, which applies the HOPE graph embedding on each

TABLE II
SILHOUETTE SCORES FOR SCENARIO I

Columns beginning with ‘GN’ use the Girvan-Newman algorithm [31]. Columns beginning with ‘LLE’ use Local Linear Embedding [42]. Columns beginning with ‘HOPE’ use High-Order Proximity preserved Embedding [39]. Columns also note whether the embedding represents the spatial graph, the connectivity graph, the hierarchy graph, or a concatenated combination. Our approach integrates all three graphs. k indicates the number of teams for that row.

Formation	k	GN sp.	GN co.	GN hi.	LLE sp.	LLE co.	LLE hi.	HOPE sp.	HOPE co.	HOPE hi.	HOPE cc.	Our Approach
Platoon Column	3	0.001	0.434	0.293	0.206	0.053	-0.113	0.542	0.101	0.225	0.341	0.763
Platoon Column	5	0.065	0.460	0.411	0.241	0.079	-0.170	0.347	-0.128	0.008	0.572	0.762
Platoon Vee	3	-0.099	0.569	0.164	0.206	0.129	-0.064	0.569	0.019	0.324	0.447	0.779
Platoon Vee	5	-0.089	0.317	0.283	0.200	0.246	-0.217	0.313	0.034	0.144	0.634	0.805
Platoon Wedge	3	-0.039	0.568	0.264	0.201	0.187	-0.070	0.579	0.028	0.254	0.465	0.785
Platoon Wedge	5	0.004	0.304	0.362	0.232	0.276	-0.149	0.367	0.035	0.125	0.626	0.807
Squad Column	2	0.042	0.466	0.466	0.068	-0.027	0.107	0.466	0.088	0.123	0.466	0.396
Squad Column	3	-0.012	0.323	0.467	0.028	-0.014	0.064	0.483	0.273	0.114	0.467	0.604
Squad File	2	0.036	0.423	0.423	-0.024	0.000	0.114	0.423	0.077	0.102	0.423	0.325
Squad File	3	-0.026	0.454	0.429	0.023	-0.031	0.083	0.283	-0.044	0.081	0.454	0.530
Squad Line	2	0.054	0.494	0.494	0.049	-0.007	0.147	0.494	-0.015	0.147	0.494	0.420
Squad Line	3	-0.054	0.328	0.514	0.014	-0.047	0.075	0.514	-0.044	0.124	0.514	0.651

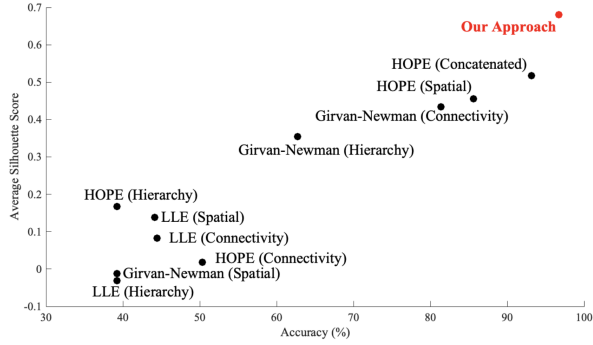


Fig. 4. Quantitative experimental results on the relationship of accuracy and silhouette score, which shows a linear relationship between higher accuracy and higher silhouette scores. Our multimodal graph embedding approach achieves both the highest accuracy and the highest silhouette score, which outperforms the previous methods.

individual graph and concatenates these vectors from three separate graphs into a single representation.

B. Results on Expert-Defined Team Formations (Scenario I)

We first evaluate our approach on the expert-defined team formations known as the *platoon column*, *platoon wedge*, and *platoon vee* and the *squad column*, *squad file*, and *squad line*, based on the field operations teaming protocol in [43]. This protocol contains correct, expert-defined sub-divisions for these formations. Platoon formations incorporate three squads and two separate leadership agents. Squad formations incorporate two teams and one separate leadership agent. Figure 3 displays the *squad column* and *platoon column* in the Webots simulator, with correct sub-divisions labeled.

The spatial relationships, communication capabilities, and structured hierarchy of these formations are defined by the field operations teaming protocol. We encode each of these relationship modalities as a separate graph in order to compare our approach to previous methods. Table I reports the clustering accuracy for our approach versus these baseline approaches. Out of a possible 306 agents in the six different formations, our approach clusters 96.73% of the agents correctly. The second best is the concatenated combination of HOPE embeddings, clustering 93.14% of agents correctly, showing that extending existing graph embedding methods

to leverage multiple information modalities can significantly improve their performance from using a single modality.

Table II displays the silhouette scores for each sub-division of each formation. Our approach achieves the highest score, with an average silhouette score of 0.680. Again, second highest is the concatenated combination of HOPE embeddings, scoring 0.518. We note that our approach achieves its best results on the platoon formations, which contain over three times as many agents as the squad formations. This suggests that our approach’s performance will extend to larger multi-robot systems. There also exists a linear relationship between clustering accuracy and silhouette scores, visualized in Figure 4. This validates the silhouette score as a metric for dividing the simulated robotic systems in Scenario II, where the ground truth divisions are unknown.

For these expert-defined teams, we also evaluated the effect of simulated sensor noise, where robots no longer have exact knowledge of where their robotic teammates are. Figure 5 shows both the effect on clustering accuracy of adding up to 50% error into the distances between robots and the effect of this noise on silhouette score. Our approach, able to utilize information from other modalities, is robust to this error, declining in accuracy only slightly for the smaller squad formations and maintaining 100% accuracy on the larger platoon formations. Our approach is similarly consistent in maintaining a high silhouette score despite the simulated sensor noise. The HOPE embedding, based on the spatial graph, is affected significantly. On the squad formations, this method declines from an average accuracy of 76.47% with no sensor noise to 61.77% with 50% noise. A similar decline in performance occurs with the platoon formations, from 88.24% to 72.95% accuracy. The HOPE embedding also produces clusters with lower silhouette scores as the simulated sensor noise increases.

C. Results on Simulated Large-scale Multi-Robot Systems (Scenario II)

To evaluate the effectiveness of our approach on a larger scale, we simulated larger multi-robot systems in Webots. These large-scale systems have the same relationship modalities as the expert-defined teams. Spatial relationships are calculated from their simulated physical locations, com-

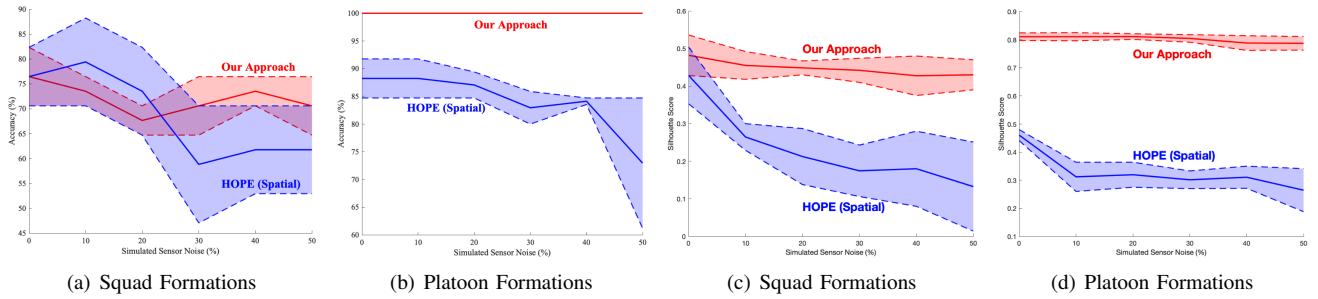


Fig. 5. Experimental results showing the effect of simulated sensor noise when determining the team assignments of robots on both clustering accuracy and silhouette score. Figure 5(a) and Figure 5(b) show the effects on clustering accuracy on the squad and platoon formations respectively, compared to the HOPE embedding method [39]. Figure 5(c) and Figure 5(d) show the effects on the silhouette score metric for the squad and platoon formations, again compared to the HOPE embedding method based solely on spatial information.

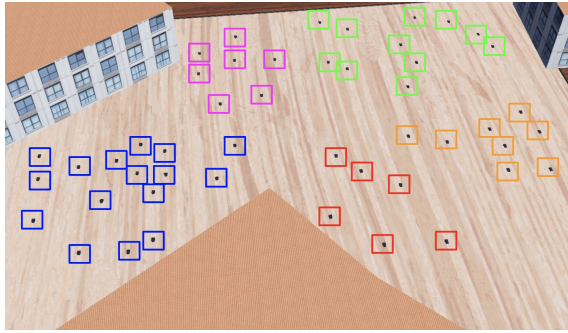


Fig. 6. Qualitative results on multi-robot team selection over a system of 50 simulated robots, with bounding boxes indicating the five teams identified by our approach.

TABLE III
SILHOUETTE SCORES FOR SCENARIO II

Bots	k	LLE s.	LLE c.	LLE h.	HOPE s.	HOPE c.	HOPE h.	HOPE cc.	Ours
10	3	0.19	0.17	0.02	0.34	0.44	0.19	0.35	0.68
10	4	0.22	0.15	0.02	0.34	0.39	0.17	0.39	0.66
10	5	0.23	0.10	0.03	0.32	0.37	0.07	0.40	0.58
20	3	0.34	0.30	-0.02	0.44	0.35	0.27	0.44	0.70
20	4	0.41	0.36	-0.01	0.45	0.44	0.21	0.49	0.68
20	5	0.45	0.32	-0.01	0.45	0.34	0.15	0.52	0.65
30	3	0.36	0.23	-0.01	0.43	0.19	0.28	0.44	0.69
30	4	0.45	0.29	-0.02	0.46	0.30	0.22	0.48	0.70
30	5	0.47	0.27	-0.01	0.47	0.22	0.19	0.51	0.67
40	3	0.36	0.17	-0.00	0.44	0.10	0.27	0.44	0.68
40	4	0.45	0.20	-0.03	0.47	0.12	0.23	0.48	0.71
40	5	0.49	0.20	-0.02	0.46	0.11	0.20	0.49	0.66
50	3	0.36	0.12	-0.02	0.42	0.03	0.27	0.42	0.70
50	4	0.45	0.13	-0.03	0.46	0.02	0.23	0.47	0.71
50	5	0.48	0.15	-0.03	0.46	0.03	0.21	0.49	0.66

munication capabilities are based on each simulated robot being capable of short-range, line-of-sight communication to nearby robots, and hierarchical relationships were defined based on heuristics from field operations teams (i.e., robots near the center of the overall system are higher in the hierarchy than robots on the perimeter). Figure 6 displays a simulated system of 50 robots in Webots.

We repeatedly generated these large multi-robot systems and performed team selection based on each evaluated approach. Table III shows results for simulated systems of 10, 20, 30, 40 and 50 robots divided into three, four and five teams, with comparisons to existing graph embedding approaches. Each combination of size and number of clusters

was run 100 times. As seen in the experimental results in Table III, our approach achieves the highest silhouette score of 0.676. This is consistent with our score of 0.680 on the expert-defined teams, where our approach achieved the highest clustering accuracy, suggesting that our identified sub-divisions of large-scale simulated multi-robot systems outperform the divisions identified by other methods.

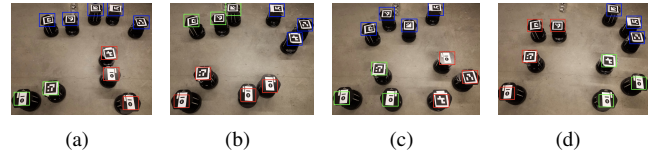


Fig. 7. Overhead views of four of the Turtlebot multi-robot systems. Bounding boxes indicate which team each Turtlebot is assigned to.

To demonstrate the ability and potential of our approach to work on physical robots in the real world, we implemented systems of 10 physical Turtlebot robots in a proof-of-concept case study. We labeled these 10 Turtlebots with tracking tags, and utilized an overhead camera to track the robots and construct their spatial relationship graph from their positions. Communication and hierarchy graphs were defined as they were in the Webots simulator. We applied our proposed multimodal graph embedding approach to identify teams within the Turtlebot system, with the objective to divide the physical robot system into three teams. Figure 7 illustrates four executions of this as a case study, with bounding boxes identifying the positions and team labelings.

V. CONCLUSION

In this paper, we introduce a novel approach to representing multi-robot structure in order to divide a multi-robot system into teams. The proposed approach represents robotic relationships as directed graphs, and constructs a vector representation from multiple graphs through a novel multimodal graph embedding method. Our approach is able to integrate multiple information modalities to describe and divide a multi-robot system, allowing our approach to create teams that are more comprehensive and effective. We demonstrate the effectiveness of our approach over expert-defined team formations and evaluate our approach on large-scale simulated multi-robot systems and on physical robot teams.

REFERENCES

- [1] M. Erdelj, M. Król, and E. Natalizio, "Wireless sensor networks and multi-uav systems for natural disaster management," *Computer Networks*, vol. 124, pp. 72–86, 2017.
- [2] G. Kruijff, M. Janíček, S. Keshavdas, B. Larochelle, H. Zender, N. Smets, T. Mioch, M. Neerinx, J. Van Diggelen, F. Colas, *et al.*, "Experience in system design for human-robot teaming in urban search & rescue," in *International Conference on Field and Service Robotics*, 2012.
- [3] B. Bayat, N. Crasta, A. Crespi, A. M. Pascoal, and A. Ijspeert, "Environmental monitoring using autonomous vehicles: a survey of recent searching techniques," *Current Opinion in Biotechnology*, vol. 45, pp. 76–84, 2017.
- [4] A. Khamis, A. Hussein, and A. Elmogy, "Multi-robot task allocation: A review of the state-of-the-art," in *Cooperative Robots and Sensor Networks*, pp. 31–51, Springer, 2015.
- [5] G. Bales and Z. Kong, "Neurophysiological and behavioral studies of human-swarm interaction tasks," in *IEEE International Conference on Systems, Man, and Cybernetics*, 2017.
- [6] P. Velagapudi, P. Scerri, K. Sycara, H. Wang, M. Lewis, and J. Wang, "Scaling effects in multi-robot control," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2121–2126, 2008.
- [7] M. Lewis, H. Wang, S.-Y. Chien, P. Scerri, P. Velagapudi, K. Sycara, and B. Kane, "Teams organization and performance in multi-human/multi-robot teams," in *IEEE International Conference on Systems, Man and Cybernetics*, pp. 1617–1623, 2010.
- [8] M. Schneider-Fontan and M. J. Mataric, "Territorial multi-robot task division," *IEEE Transactions on Robotics and Automation*, vol. 14, no. 5, pp. 815–822, 1998.
- [9] I. Rekleitis, V. Lee-Shue, A. P. New, and H. Choset, "Limited communication, multi-robot team based coverage," in *IEEE International Conference on Robotics and Automation*, 2004.
- [10] M. Castillo-Cagigal, A. Brutschy, A. Gutiérrez, and M. Birattari, "Temporal task allocation in periodic environments," in *International Conference on Swarm Intelligence*, Springer, 2014.
- [11] A. Brutschy, G. Pini, C. Pinciroli, M. Birattari, and M. Dorigo, "Self-organized task allocation to sequentially interdependent tasks in swarm robotics," *Autonomous Agents and Multi-Agent Systems*, vol. 28, no. 1, pp. 101–125, 2014.
- [12] I. Jang, H.-S. Shin, and A. Tsourdos, "A comparative study of game-theoretical and markov-chain-based approaches to division of labour in a robotic swarm," *IFAC-PapersOnLine*, vol. 51, no. 12, pp. 62–68, 2018.
- [13] I. Jang, H.-S. Shin, and A. Tsourdos, "Bio-inspired local information-based control for probabilistic swarm distribution guidance," *arXiv preprint arXiv:1711.06869*, 2017.
- [14] H. Wu, H. Li, R. Xiao, and J. Liu, "Modeling and simulation of dynamic ant colonies labor division for task allocation of uav swarm," *Physica A: Statistical Mechanics and Its Applications*, vol. 491, pp. 127–141, 2018.
- [15] T. H. Labella, M. Dorigo, and J.-L. Deneubourg, "Division of labor in a group of robots inspired by ants' foraging behavior," *ACM Transactions on Autonomous and Adaptive Systems*, vol. 1, no. 1, pp. 4–25, 2006.
- [16] E. Bonabeau, A. Sobkowski, G. Theraulaz, J.-L. Deneubourg, *et al.*, "Adaptive task allocation inspired by a model of division of labor in social insects," in *BCEC*, pp. 36–45, 1997.
- [17] H. Li, C. Feng, H. Ehrhard, Y. Shen, B. Cobos, F. Zhang, K. Elamvazhuthi, S. Berman, M. Haberland, and A. L. Bertozzi, "Decentralized stochastic control of robotic swarm density: Theory, simulation, and experiment," in *International Conference on Intelligent Robots and Systems*, IEEE, 2017.
- [18] E. Ephrati and J. S. Rosenschein, "Divide and conquer in multi-agent planning," in *AAAI Conference on Artificial Intelligence*, 1994.
- [19] R. Zlot, A. Stentz, M. B. Dias, and S. Thayer, "Multi-robot exploration controlled by a market economy," in *IEEE International Conference on Robotics and Automation*, 2002.
- [20] J. P. Desai, J. P. Ostrowski, and V. Kumar, "Modeling and control of formations of nonholonomic mobile robots," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 6, pp. 905–908, 2001.
- [21] J. P. Desai, "Modeling multiple teams of mobile robots: a graph theoretic approach," in *IEEE International Conference on Intelligent Robots and Systems*, 2001.
- [22] J. P. Desai, "A graph theoretic approach for modeling mobile robot team formations," *Journal of Robotic Systems*, vol. 19, no. 11, pp. 511–525, 2002.
- [23] Y. Li and X. Chen, "Stability on multi-robot formation with dynamic interaction topologies," in *IEEE International Conference on Intelligent Robots and Systems*, 2005.
- [24] M. C. De Gennaro, L. Iannelli, and F. Vasca, "Formation control and collision avoidance in mobile agent systems," in *IEEE International Symposium on Intelligent Control*, 2005.
- [25] R. Fierro and A. Das, "A modular architecture for formation control," in *IEEE International Workshop on Robot Motion and Control*, 2002.
- [26] R. Olfati-Saber and R. M. Murray, "Graph rigidity and distributed formation stabilization of multi-vehicle systems," in *IEEE Conference on Decision and Control*, 2002.
- [27] R. Olfati-Saber, W. B. Dunbar, and R. M. Murray, "Cooperative control of multi-vehicle systems using cost graphs and optimization," in *IEEE American Control Conference*, 2003.
- [28] C. H. Ding, X. He, H. Zha, M. Gu, and H. D. Simon, "A min-max cut algorithm for graph partitioning and data clustering," in *IEEE International Conference on Data Mining*, 2001.
- [29] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000.
- [30] M. E. J. Newman, "Clustering and preferential attachment in growing networks," *Physical Review E*, vol. 64, p. 025102, 2001.
- [31] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical Review E*, vol. 69, p. 026113, 2004.
- [32] M. E. J. Newman, "Modularity and community structure in networks," *Proceedings of the National Academy of Sciences*, vol. 103, no. 23, pp. 8577–8582, 2006.
- [33] P. Goyal and E. Ferrara, "Graph embedding techniques, applications, and performance: A survey," *Knowledge-Based Systems*, vol. 151, pp. 78–94, 2018.
- [34] H. Cai, V. W. Zheng, and K. Chang, "A comprehensive survey of graph embedding: problems, techniques and applications," *IEEE Transactions on Knowledge and Data Engineering*, 2018.
- [35] P. Cui, X. Wang, J. Pei, and W. Zhu, "A survey on network embedding," *IEEE Transactions on Knowledge and Data Engineering*, 2018.
- [36] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.
- [37] B. Shaw and T. Jebara, "Structure preserving embedding," in *International Conference on Machine Learning*, 2009.
- [38] A. Ahmed, N. Shervashidze, S. Narayanamurthy, V. Josifovski, and A. J. Smola, "Distributed large-scale natural graph factorization," in *International Conference on World Wide Web*, 2013.
- [39] M. Ou, P. Cui, J. Pei, Z. Zhang, and W. Zhu, "Asymmetric transitivity preserving graph embedding," in *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.
- [40] L. Katz, "A new status index derived from sociometric analysis," *Psychometrika*, vol. 18, no. 1, pp. 39–43, 1953.
- [41] G. H. Golub and C. Reinsch, "Singular value decomposition and least squares solutions," *Numerische Mathematik*, vol. 14, no. 5, pp. 403–420, 1970.
- [42] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
- [43] Department of the Army, *U.S. Army Field Manual 3-21.9*, 2017.
- [44] Webots, "http://www.cyberbotics.com." Commercial Mobile Robot Simulation Software.
- [45] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987.