Finding Subgroups of UAV Swarms Using a Trajectory Clustering Method

Kongjing Gu, Ziyang Mao, Mingze Qi, Xiaojun Duan*

College of Liberal Arts and Science, National University of Defence Technology, No.109, DeyaRoad, Changsha, 4100072, China

*xjduan@nudt.edu.cn

Abstract. Unmanned Aerial Vehicles(UAV) swarm is a rapidly developing field, and with it comes the need to identify the swarm based on observations. The problem of trajectory clustering is put forward in the identification of UAV swarms, especially modularized UAV swarms. We propose a new method of Network Integrated trajectory clustering(NIT) to solve the trajectory clustering problem in a fast-changing and chaotic environment which requires a quick response, fault tolerance, and accuracy. The experiment results prove the flexibility and adaptability of the NIT method towards various demands and multi-dimensional data. Moreover, the algorithm proposed based on the method shows priority over the other three trajectory clustering methods(DTW, Fréchet distance, GMM) on the accuracy, and fault tolerance in clustering swarm trajectories. The method raised in this paper is an innovation to both multi-agent systems identification and trajectory clustering methods.

Keywords: Trajectory clustering, UAV swarm

1. Introduction

The research on Unmanned Aerial Vehicle(UAV) swarm has attracted much attention in the field of daily applications such as detection and monitoring with the development of distributed communication and control. UAV swarm with modular structure is one of the main types, which achieves modularization, standardization, assembly, and disassembly while maintaining the stability of its function[1]. DB Worth at al.[2] discussed a framework to develop and test the behaviors of scalable and modular swarms of autonomous UAVs. Matthias et al.[3] proposed a networked UAV swarm to form a defense system chasing the malicious UAV, which can be viewed as an embryonic form of modularized UAV swarm. Meanwhile, it is of great significance to distinguish the communication and cooperation structure of the UAV swarm.

Trajectory clustering contains two parts: similarity measurements and clustering[4]. Similarity measurements are diverse due to the variety of the raw data[5] including Euclidean distance, Hausdorff distance[6, 7], Longest Common SubSequence (LCSS) distance[8, 9], Dynamic time warping (DTW) distance[10–12], Fréchet distance[13, 14], etc. Meanwhile, the clustering methods are classified into five categories[15]: partition-based method, hierarchy-based method, density-based method, grid-based method, and model-based method.

As the UAV swarm has its own features, the problem of finding the subgroup structures of the UAV swarm has been overwhelming the conventional pattern. First, traditional methods are not
suitable when processing trajectories of closely linked and highly cooperated entities like UAV swarms, bird flocking, fish schools, etc. because of their high tolerance of data in conformity and low level of nuance identification ability. Second, mistakes in observations call for a fault-tolerant approach to cope with observation errors caused by intricate trajectories. Thus, an accurate and fault-tolerant clustering method aiming at UAV swarm is required.

This paper proposes a Network Integrated trajectory clustering (NIT) algorithm to achieve quick and accurate identification of the UAV swarm. The NIT algorithm has the ability to distinguish the subtle difference among very similar trajectories and their variation trend, which is a typical feature in the UAV swarm and biological schools with high speed and synchronization. Besides, the NIT algorithm is able to minimize the fault caused by observation error such as inconsistency of trajectories and missing time points as it can freely operate the discrete-time slices rather than a whole trajectory.

The rest of this paper is organized as follows. The new method NIT is presented in Section 2 as well as the problem description to be solved. Section 3 proposes a standard of accuracy evaluation and verifies the NIT method by extensive experiments. Section 4 concludes the paper.

2. Model

This paper is to solve the clustering issue of the modularized UAV swarm with a new framework named NIT to cluster trajectories in which concepts in complex networks are introduced to build the frame in Fig.1. The NIT framework consists of three steps including slice, networking, and decomposition (see Fig.1).

![Fig.1 Schematic diagram](image)

**Time slice:** Time slice which cuts the trajectories according to their time stamp is an intuitive way of improving Euclidean distance. Different from Euclidean distance which calculates the distance of the whole trajectories, the UNIT method calculates the distance of each time separately and integrates the results of each time in the next step. As shown in Fig. 1 the objects we need to deal with are the trajectories $L_i = \{l_{i,0}, l_{i,1}, \cdots, l_{i,N}\}, i = 1, 2, \cdots, N$ that become time slices $h$ by time stamp. Data on each slice could be clustered as a general clustering problem, and the dimension limitation of data $l_{ij}$ is extended. After applying ordinary clustering methods on each slice, the clustering results of each time slice is obtained, which is $c_i^j = \{c_i^j, c_i^j, \cdots, c_i^j\}$, such that $s_t = \bigcup_{i=1}^N c_i^j$, and $c_i^j \cap c_i^j = \emptyset$, for $i \neq j$.

**Networking:** The next step is to synthesize results on each slice reasonably and the idea of a complex network is borrowed. Assuming that results on each slice produce one single-layer network,
the problem of synthesizing time slices turns into overlaying networks corresponding to each slice. Constructing a network according to the clustering results of each time slice is divided into two steps. The first step is to produce a single-layer network $g_t$ by giving a fully connected network among nodes in the same cluster on a time slice, and keeping the remaining nodes disconnected. The second step is to build a network $G_t = g_1 \circ \cdots \circ g_{\tau} \circ g_1$.

**Decomposition:** The information is mainly contained on the links of the integrated network. Therefore, clustering results should be decided by the weight distribution. The higher the weight is, the closer the connection of the two nodes is. And there exist a lot of community detection methods that could help to decompose the network, such as modularity Q[16], Louvain algorithm[17], and Label Propagation Algorithm[18], etc.

**Determine the number of clusters:** First to find the optimal cluster number on each slice with the most common standards such as the silhouette method. Then integrate these optimal numbers $k_t, t = 1, 2, \cdots, \tau$ into one cluster number $k$ throughout all trajectory. Then

$$k = \max_{i=1,2,\cdots,\tau} k_i, \text{ for } \#k_i > \mu \sum_{i=1}^{\tau} (\#k_i)$$

(1)

where $\mu$ is a ratio to control the qualified cluster number. This definition aims to pick up stable enough results and prefers the larger cluster number among selectable solutions.

### 3. Performance
To illustrate the effectiveness of the NIT frame and the algorithms proposed based on it. We apply the algorithms in a simulation-example of UAV swarm based on potential function, with contrast to the other three clustering trajectory methods, namely Fréchet’s distance, DTW, and Gaussian Mixture Model (GMM; [19, 20]). It turned out that our algorithm has an advantage in accuracy and fault-tolerance.

**Simulation settings:** This paper uses a potential function[21, 22] to form a UAV swarm, which is capable of building various formations as well as diverse communication structures. Without loss of generality, we assume the UAV swarm contains $N = 40$ entities forming 5 subgroups (each consists of 8 fully connected UAVs), and one of the UAVs in each subgroup is selected as a key node connecting with other key nodes.

**Accuracy measurement:** Given the clustering result of the UAV trajectory, the accuracy rate is composed of two parts. The first part is the overall number of correctly clustered entities, which is defined by

$$A_{\tau} = \begin{cases} 
\sum_{j=1, \ldots, k} \max_{j=1, \ldots, k} \#(C_i \cap C_j), & \text{for } k > k_0; \\
\sum_{j=1, \ldots, k} \max_{j=1, \ldots, k} \#(C_i \cap C_j), & \text{others.}
\end{cases}$$

(2)

where $k$ and $k_0$ are respectively cluster number and true subgroup number, and $C_i$ is the clustering set, $C_j$ the true UAV subgroup. The second part is defined based on the ratio between the cluster number $\$k\$ and the real subgroup number $k_0$, which is
\[
A_i = \begin{cases} 
1 - \frac{k_0 - k}{k_0}, & \text{for } k < k_0; \\
1 - \frac{k - k_0}{k}, & \text{others.}
\end{cases}
\]

which punishes the deviation of the cluster number (e.g. Fig.2).

We hereby derive the accuracy rate as below

\[
A = A_i A_r.
\]

Assuming the subgroup number is 5, the left accuracy changes with the cluster number.

**Performance of the NIT method:** Experiments are repeated 30 times with a swarm of 40 entities each time. The swarm consists of 2 subgroups in the first 10 times, 4 in the middle 10 times, and 5 in the last 10 times. The trajectories are 4-dimensional containing 2-dimensional positions and 2-dimensional velocities. The accuracy of the NIT algorithm is shown in the left plot of Fig.3. Results show that \( A_i \) is usually stable and high, while \( A_r \) fluctuates a lot, which means that the NIT algorithm could accurately group the similar trajectories of individuals together, and the estimation of \( k \) often determines the performance of the NIT algorithm.

Compare the method with three kinds of common trajectory clustering methods (Frechet’s distance, DTW and GMM) to compare with the NIT method. The average accuracies are listed as below:

| Clusters | NIT  | DTW  | Frechet | GMM  |
|----------|------|------|---------|------|
| 2        | 0.98 | 0.80 | 0.67    | 0.92 |
| 4        | 0.52 | 0.37 | 0.53    | 0.50 |
| 5        | 0.55 | 0.51 | 0.44    | 0.36 |
Fig. 3 The average accuracies

Left plot is the accuracy of NIT algorithm. The line of color light blue on the diagram represents the total accuracy \( A \), the green one \( A_i \), and the grey one \( A_r \). The right plot is the accuracy of each algorithm. Blue line represents NIT, grey represents DTW, dark blue represents Fréchet Distance, and light blue represents GMM.

The changing trends of each algorithm under various situations are shown in the right plot of Fig. 3 which indicates the superiority of the NIT method over the other three algorithms.

The method in this paper is more suitable for clustering trajectory and more sensitive to nuance, while DTW and other clustering methods are less sensitive to time-delay information due to their scaling principle. Moreover, the NIT algorithm is suitable for trajectories with high-dimensional components, while the other three methods have many more limitations.

4. Conclusion
Extensive experiments validate its effectiveness as well as its superiority to three existing trajectory clustering methods (DTW, Frechet distance, GMM) on accuracy. To the best of our knowledge, our work is the first place to design a framework of trajectory clustering that comprehensively considers the identification of the UAV swarm according to their trajectories. The main contributions of this study are as follows:

- We propose a new algorithm named NIT for trajectory clustering. The clustering results indicate the effectiveness and superiority of the accuracy and stability of our algorithm.
- We formulate the clustering problem as three sub-problems, clarifying the key issues when distinguishing subgroups of the UAV swarm, which are solved by the proposed scheme. It shows the scheme introduced in this paper is practical and effective in complex trajectory clustering.
- We conduct extensive experiments comparing our NIT algorithm with the other three typical algorithms, i.e. DTW, Fréchet distance, GMM, to verify the superiority of the NIT algorithm. The experimental results prove its advantages on accuracy.

Acknowledgments
This work was supported by the National Natural Science Foundation of China (No. 11771450) and the National Numerical Wind Tunnel Project (NNW2019ZT7-B23).

References
[1] Chen X, Tang J and Lao S 2020 Review of Unmanned Aerial Vehicle Swarm Communication Architectures and Routing Protocols Appl. Sci. 10 p3661
[2] Worth D B, Woolley B G and Hodson D D 2017 SwarmSim: a framework for modeling swarming unmanned aerial vehicles using hardware-in-the-loop J. Def. Model. Simul.
[3] Matthias R B, Gr’egoire D, Pascal B, Dren G, Himadri P and Mike P G 2017 Defending against intrusion of malicious uavs with networked uav defense swarms 2017 IEEE 42nd Conf. on Local Computer Networks Workshops (LCN Workshops) (Piscataway: IEEE) pp103–111
[4] Guan Y, Penghui S, Jie Z, Daxing L and Canwei W 2017 A review of moving object trajectory clustering algorithms Artif. Intell. Rev. 47 pp123–144
[5] Michail V, George K, and Dimitrios G 2002 Discovering similar multidimensional trajectories Proc. 18th Int. Conf. on data engineering (Piscataway: IEEE) pp673–684
[6] Jinyang C, Rangding W, Liangxu L and Jiatao S 2011 Clustering of trajectories based on hausdorff distance 2011 Int. Conf. on Electronics, Communications and Control (ICECC) (Piscataway: IEEE) pp1940–1944
[7] Xiaohai H E, Qizhi T and Mingliang G 2013 Trajectory Classification Based on Hausdorff Distance and Longest Common SubSequence J. Electr. Inf. Technol 35 pp784-790
[8] Claus R 2000 Efficient computation of all longest common subsequences Scandinavian Workshop on Algorithm Theory (Berlin: Springer) pp407–418
[9] Michail V, Marios H, Dimitrios G and Eamonn K 2006 Indexing multidimensional time-series VLDB J 15 pp1–20
[10] Kruskall J and Liberman M 1983 The theory and practice of sequence comparison J. Time warps, string edits, and macromolecules (London: Addison-Wesley) pp1-44
[11] Lei C, M T O and Vincent O 2005 Robust and fast similarity search for moving object trajectories Proc. of the 2005 ACM SIGMOD Int. Conf. on Management of data pp491–502
[12] Jingren T, Hong C, Yang Z and Hongliang G 2018 Structured dynamic time warping for continuous hand trajectory gesture recognition Pattern Recognit. 80 21–31
[13] Eiter T and Mannila H 1994 Computing discrete Fréchet distance, Christian Doppler Laboratory for Expert Systems, TU Vienna, Austria. Tech. Rep. CD-TR 94/64
[14] Khoshaein V 2014 Trajectory Clustering Using a Variation of Fréchet Distance Diss. Université d’Ottawa/University of Ottawa
[15] Jiawei H, Jian P and Micheline K 2011 Data mining: concepts and techniques (Amsterdam: Elsevier)
[16] Jinyin C, Lihong C, Yixian C, Minghao Z, Shanqing Y, Qi X and Xiaoniu Y 2019 Ga-based q-attack on community detection IEEE Trans. Comput. Soc. Syst. 6 pp491–503
[17] Ghosh S, Halappanavar M, Tumeo A, Kalyanaraman A, Lu H, Chavarria-Miranda D, Khan A and Gebremedhin A 2018 Distributed louvain algorithm for graph community detection 2018 IEEE Int. Parallel and Distributed Processing Symposium (IPDPS) (Piscataway: IEEE) pp885-895
[18] Garza S E and Schaeffer S E 2019 Community detection with the Label Propagation Algorithm: A survey Physica A 534
[19] Gaffney S and Smyth P 1999 Trajectory clustering with mixtures of regression models Proc. of the fifth ACM SIGKDD Int. Conf. on Knowledge discovery and data mining pp63-72
[20] Wei W, Feng X, Hansong N, Zhikui C, Zhiguo G, Xiangjie K, and Wei W 2020 Vehicle trajectory clustering based on dynamic representation learning of internet of vehicles IEEE trans. Intell. Transp. Syst.
[21] Rimon E and Koditschek D E 1992 Exact robot navigation using artificial potential functions Departmental Papers (ESE) 323
[22] Rossi F, Bandyopadhyay S, Wolf M and Pavone M 2018 Review of multi-agent algorithms for collective behavior: a structural taxonomy IFAC-PapersOnLine 51 pp112-117