

Aerial Targets Clustering with Unknown Cluster Number

Wangwang Zhou^{1,*}, Peiyang Yao¹ and Shuai Wei²

¹Information and Navigation College, Air Force Engineering University, Xi'an 710077, China

²Unit 95910 of the PLA, Jiuquan 735000, China

*Corresponding author

Abstract—Aiming at the uncertainty of enemy target cluster number in aerial targets clustering process, an improved affinity propagation(AP) algorithm is adopted to determine the target cluster number and clustering result. Firstly, Between-Within Propagation(BWP) index is used as the measure of cluster effectiveness, and Bisection method is used to determine the optimal number of clusters;. Secondly, scanning of preferences subspace corresponding to the optimal cluster number, and determine the clustering result. Finally, the effectiveness and stability of the improved affinity propagation algorithm are verified by experiments.

Keywords—target clustering; affinity propagation; cluster number; BWP index

I. INTRODUCTION

Target clustering is one of effective ways to reduce the amount of information on the battlefield. By clustering aerial targets, the battlefield situation can be simplified and concise. The enemy's combat units are organized and assembled according to certain rules, and then they work together to complete the same combat mission. Therefore, targets performing the same combat mission will have similar motion states, such as distance l , azimuth angle θ , heading angle μ , speed v , height h et al. Based on these maneuvering characteristics, aerial targets can be clustered[1].Reference [2-6] use Chameleon algorithm, hierarchical clustering and K-means algorithm to cluster targets. These methods all achieve nice result of clustering, but they all need to specify the number of clusters before clustering. In actual combat, the number of enemy target clusters is often unknown.

The Affinity Propagation (AP)[7] algorithm determines the clustering center through the information transfer between data, and does not need to specify the number of clusters and the initial clustering center. It has been widely used in face recognition, text mining, feature gene extraction, image segmentation and other fields. AP algorithm can get better clustering results, but it is easy to produce more clusters. Reference [8] determine the number of clusters by the index of Silhouette[9], while [10] uses the Gap [11] index to determine the number of clusters. Because the bias parameter P in a certain range corresponds to the same number of clusters, the same cluster number will correspond to different clustering results. Therefore, the clustering results in [8,10] are contingent.

In the current paper, we use AP algorithm to cluster aerial targets based on their maneuver data. We use bisection method

to search for the clustering results corresponding to all cluster numbers, and use the BWP[12] index value as the cluster validity measure to determine the optimal cluster number. After that, we determine the optimal clustering result by traversing the bias parameter subspace corresponding to the optimal clustering number. Finally, the simulation results show that the proposed method can effectively determine the target cluster number, and the clustering accuracy is better than methods proposed in [8] and [10].

II. AFFINITY PROPAGATION ALGORITHM

The AP algorithm performs clustering based on the propagation of neighbor information. The AP algorithm treats all samples as potential class representatives and has the same probability of being selected as a class representative. That is, $s(i,i)=P$. Each sample establishes its own similarity $s(i,j)$ to other samples, $s(i,j) = -\|x_i - x_j\|, (i \neq j)$. The AP algorithm needs to continuously search for two important information in order to select the proper class representatives in the samples: the attractiveness r and the belonging degree a . $r(i,k)$ represents the extent to which the class x_k is suitable for x_i 's class representation point. $a(i,k)$ denotes the extent to which x_i selects x_k as the class representative point. For the sample x_i , calculate the sum of the attractiveness $r(i,k)$ of the other samples and the belonging degree $a(i,k)$, and the largest sample x_k is the class representative of x_i . The degree of attraction $r(i,k)$ and the degree of belonging $a(i,k)$ are updated as follows:

$$r^{(t)}(i,j) = \lambda r^{(t-1)}(i,j) + (1-\lambda) \left\{ s(i,j) - \max_{j' \neq j} [a^{(t-1)}(i,j') + s(i,j')] \right\} \quad (1)$$

$$a^{(t)}(i,j) = \lambda a^{(t-1)}(i,j) + (1-\lambda) \min_{i' \neq j} \left\{ 0, r^{(t-1)}(j,j) + \sum_{i' \neq i, j' \neq j} \max [0, r^{(t-1)}(i',j)] \right\} \quad (2)$$

$$a^{(t)}(j,j) = \lambda a^{(t-1)}(j,j) + (1-\lambda) \left\{ \sum_{i \neq j} \max [0, r^{(t-1)}(i,j)] \right\} \quad (3)$$

Where t is the number of iterations and λ is the damping factor, $0.5 < \lambda < 1$.

III. AERIAL TARGETS CLUSTERING BASED ON IMPROVED AFFINITY PROPAGATION ALGORITHM

A. Determination of Bias Parameter P

In the affinity propagation algorithm, the diagonal element $s(i, i)$ of the similarity matrix is the bias parameter P . It can be known from (1), (2), and (3) that when the bias parameter P is larger, the obtained $R(k, k)$ is larger, and then $A(i, k)$ is also larger, which results in the AP algorithm outputting more cluster numbers. Since P in a certain range corresponds to the same cluster number, the same cluster number will correspond to different clustering results, as shown in Figure. 1, so it is necessary to determine the optimal bias parameter P by searching.

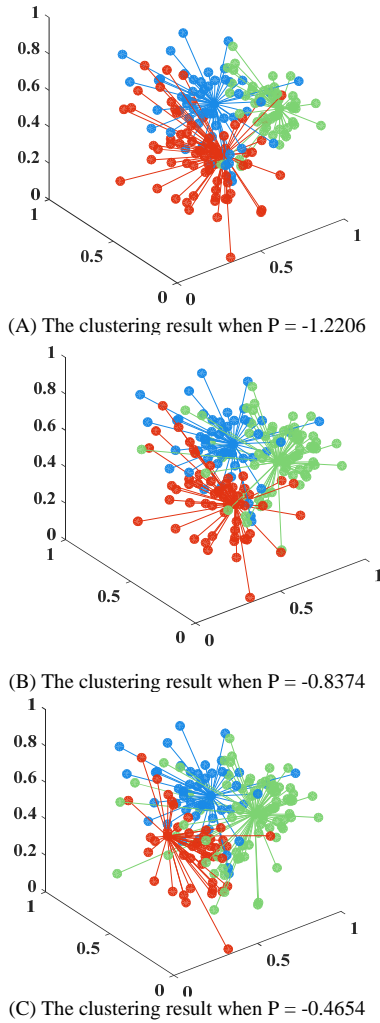


FIGURE 1. DIFFERENT CLUSTERING RESULTS OF WINE UNDER THE CLUSTER NUMBER OF 3

This paper designs two stages of rough search and fine search to determine the best bias parameter P . Firstly, a rough search is used to obtain the bias parameter P subspace

corresponding to the optimal clustering number K , and then the optimal bias parameter P is obtained by traversing the P subspace.

P 's value interval $[p_{min}, p_{max}]$, Wang[13] and other studies have obtained:

$$\begin{aligned} P_{min} &= \min s(x_i, x_j) \\ P_{max} &= \max s(x_i, x_j) \end{aligned} \quad (4)$$

The number of clusters ranges from $[2, k_{max}]$, and k_{max} is generally takes $\text{int} \sqrt{N}$.

Search process of value P .

- Search the value range of P based on Bisection method to obtain $P(i)$, $i=1, 2 \dots k_{max}-1$ corresponding to the cluster number $[2, k_{max}]$.
- Calculate $P(i)$, $i=1, 2 \dots k_{max}-1$, correspond to the BWP index value of the clustering result to determine the optimal clustering number k_{opt} .
- Using Bisection method to search the bias parameter subspace $[p(k_{opt}-1), p(k_{opt})]$, determine the lower bound pn of the parameter subspace corresponding to the optimal cluster number k_{opt} . Using Bisection method to search the bias parameter subspace $[p(k_{opt}), p(k_{opt}+1)]$, determine the upper bound px of the parameter subspace corresponding to the optimal cluster number k_{opt} .
- Traverse the preference parameter subspace $[pn, px]$ in steps β . Calculate the BWP index value of the corresponding clustering result to determine the best bias parameter pb .

B. APBTM Algorithm

Combining Bisection and traversal method with AP algorithm to get Affinity Propagation Based on Bisection and Traversal Method (APBTM) algorithm. The APBTM algorithm steps are as follows:

- Set the search step length β and the calculation precision ε .
- Build a similarity matrix.
- Search value interval $[p_{min}, p_{max}]$ corresponding to bias parameter P based on Bisection method.
- If all $P(i)$ $i=1, 2 \dots k_{max}-1$, corresponding to the cluster number $[2, k_{max}]$ are searched, go to next step; otherwise go to previous step.
- Calculate BWP index value corresponding to different clusters. Use BWP index value to determine the optimal number of clusters k_{opt} .
- Search bias parameter subspaces $[p(k_{opt}-1), p(k_{opt})]$ and $[p(k_{opt}), p(k_{opt}+1)]$ to determine the upper bound

pn and lower bound px of preferences subspace corresponding to the optimal cluster number.

- Traverse the parameter subspace $[pn, px]$ by step size β , and calculate the BWP index value of clustering result to determine the best bias parameter pb .
- Output clustering result corresponding to the optimal bias parameter pb .

IV. EXPERIMENT AND RESULTS

A. Cluster Effect Comparison Experiment

Based on the data sets Pid, Bcw, Wine and Model1, the clustering accuracy is compared with the adAP algorithm proposed in [8] and the APBWMMP algorithm proposed in [10]. The specific results are shown in Table 1.

TABLE I. COMPARISON OF CLUSTERING ACCURACY AMONG APBWMMP, ADAP AND APBTM

Data set	Algorithm	Cluster Number	Best Cluster Number	Accuracy
Pid	APBWMMP	2	2	62.11%
	adAP	2	2	73.57%
	APBTM	2	2	85.42%
Bcw	APBWMMP	2	2	86.70%
	adAP	2	2	95.57%
	APBTM	2	2	95.63%
Wine	APBWMMP	3	3	70.94%
	adAP	3	3	89.33%
	APBTM	3	3	91.57%
Model1	APBWMMP	2	2	89.00%
	adAP	2	2	86.00%
	APBTM	2	2	89.47%

All three algorithms can find the same optimal number of clusters as the actual number of clusters. Compared with the other two algorithms, APBTM algorithm can obtain better clustering accuracy.

B. Target Clustering Experiment

The data set Data is the characteristic value of the motion attributes such as azimuth, distance, horizontal speed, heading angle and altitude of 80 sorties. It is divided into four functional groups: attack, penetration, investigation and surveillance. Some data are shown in Table 2. The data set Data is clustered based on the APBTM algorithm. The clustering results are shown in Figure 2, Figure 3 and Table 3.

TABLE II. TARGET MOTION ATTRIBUTE EIGENVALUES AND GROUP CLASSES

Target	Θ (mile)	L (km)	V (m/s)	M (°)	H (km)	Class
X ₁	820	280	250	200	60	investigation
X ₂	2300	210	300	320	40	attack
X ₃	828	281	245	201	65	investigation
X ₄	2350	215	320	322	42	attack
X ₅	830	282	255	200	63	investigation
X ₆	825	283	250	204	61	investigation
X ₇	2200	150	300	156	50	attack
X ₈	4000	110	300	50	35	penetration
X ₉	2800	260	220	260	80	surveillance
X ₁₀	4050	120	280	51	36	penetration
...						

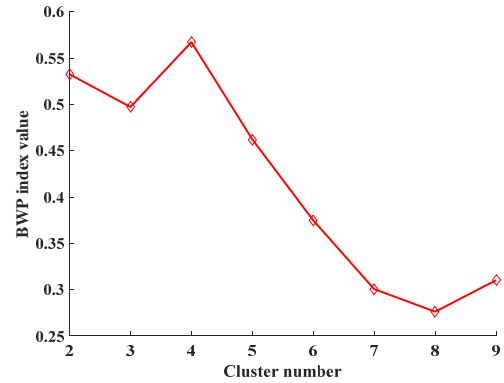


FIGURE II. THE RELATIONSHIP BETWEEN THE NUMBER OF CLUSTERING DATA AND BWP INDICATORS

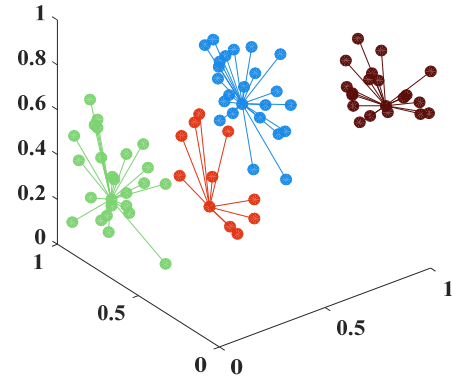


FIGURE III. TARGET CLUSTERING RESULTS

TABLE III. DATA'S CLUSTERING RESULTS BASED ON APBTM ALGORITHM

Data set	Samples Number	Cluster Number	Best Cluster Number	Accuracy
Data	80	4	4	98.75%

Based on the BWP index and the APBTM algorithm, the same number of aerial target clusters as the real cluster number can be obtained, and the accuracy is high.

V. CONCLUSION

In this paper, the rough search and fine search are designed for the problem of aerial targets clustering with unknown cluster number. The search results are evaluated by BWP indicators to confirm the best clustering results. The experimental results show that the clustering of air targets based on APBTM algorithm can confirm the target cluster number and obtain better clustering results.

REFERENCES

- [1] M. J.A. Berry, and G. S. Linnoff, *Data Mining Techniques*, New York: John Wiley and Sons Ltd, 1997.
- [2] Y De Ping, Z Juan Yi, S Hao Shan, and L Ning, "Target grouping algorithm based on multiple combat formations," *Computer Science*, vol. 43, pp. 235-238, February 2016.
- [3] L Ji Jun, L Gang, and W Yan Feng, "The research on multi-target grouping on the sea battlefield based on the clustering-matching method," *Sichuan Ordnance Journal*, vol 34, pp. 104-106, February 2013.
- [4] W Wen Long, H Wen Bing, and M Xiang Hai, "Research on multi-target grouping on submarine battlefield," *Fire Control & Command and Control*, vol. 38, pp. 63-68, November 2013.
- [5] L Zhen Zhen, Z Ce, and W Wei Ping, "Algorithm for operational object grouping in situation assessment based on hierarchical clustering," *Journal of Projectiles, Rockets, Missiles and Guidance*, vol. 29, pp. 203-211, June 2009.
- [6] Z Xu Liang, Z Hong Jun, Q Xiu Li, W Wen Bo, W Ze, et al, "A clustering method of land battlefield maneuvering targets based on improved K-means algorithm," *Information Technology*, vol 3, pp. 128-131, March 2016.
- [7] B J Frey, and D Dueck, "Clustering by passing messages between data points," *Science*, vol. 315, pp. 972-976, February 2007.
- [8] W Kai Jun, Z Jun Ying, L Dan, Z Xin Na, and G Tao, "Adaptive affinity propagation clustering," *Acta Automatica Sinica*, vol. 33, pp. 1242-1246, December 2007.
- [9] S Dudoi, and J Fridlyand, "A prediction-based resampling method for estimating the number of clusters in a dataset," *Genome Biology*, vol. 3, pp. 1-21, June 2002.
- [10] T Dan, and Z Zheng Jun, "New optimized affinity propagation clustering algorithm," *Journal of Computer Applications*, vol. 37, pp. 258-261, June 2017.
- [11] R Tibshirani, G Walther, and T Hastie, "Estimating the number of cluster in a dataset via the gap statistic," *Journal of the Royal Statistical Society*, vol. 63, pp. 411-423, February 2001.
- [12] Z Shi Bing, X Zhen Yuan, and T Xu Qing, "Method for determining optimal number of clusters based on affinity propagation clustering," *Control and Decision*, vol. 26, pp. 1147-1152, August 2011.
- [13] W Chang Dong, L Jian Huang, S Ching Y, and Z Jun Yong, "Multi-exemplar affinity propagation," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 35, pp. 2223-2237, September 2013.