

Tactical intention recognition of aerial target based on XGBoost decision tree

WANG Lei, LI Shi-zhong

(College of Mechatronic Engineering, North University of China, Taiyuan 030051, China)

Abstract: In order to improve the accuracy of target intent recognition, a recognition method based on XGBoost (eXtreme Gradient Boosting) decision tree is proposed. This paper adopts relevant data and program of python to calculate the probability of tactical intention. Then the sequence intention probability is obtained by applying Dempster-Shafer rule of combination. To verify the accuracy of recognition results, we compare the experimental results of this paper with the results in the literatures. The experiment shows that the probability of tactical intention recognition through this method is improved, so this method is feasible.

Key words: tactical intention recognition of target; XGBoost (eXtreme Gradient Boosting) decision tree; Dempster-Shafer combination rule

CLD number: E91

Document code: A

Article ID: 1674-8042(2018)02-0148-05

doi: 10.3969/j.issn.1674-8042.2018.02.008

0 Introduction

In modern war, predicting enemy target intention accurately, assessing the current situation in time and providing a reliable basis to take countermeasures at the next moment, play an important role in control of the war situation. In the past, conflict analysis, markov analysis^[1], Bayesian network technology^[2] etc were presented. In 2017, Cao S Y proposes an improved method of target intent recognition for high-dimensional data similarity^[3].

The process of target intention recognition mainly includes the following steps: ① collecting data from the sensors, such as altitude, speed, distance etc; ② processing the data and extracting the feature; ③ analyzing the target intent according to the recognition mechanism; ④ obtaining the results.

In the military environment, the target intention recognition is based on the relevant data, such as weapon equipment, operational knowledge manual, military expert's experience, map and meteorological data etc. However, it is difficult to obtain and process these data to meet the requirements. In order to construct an intention recognition model successfully, it is necessary to go on a great deal of theoretical researches and engineering practices.

The difficulty of target tactical intention

recognition is to identify intention precisely in the situation based on misinformation^[4]. In order to solve this problem, a recognition method based on XGBoost (eXtreme Gradient Boosting) algorithm^[5] is presented in this paper. As a parallel boosting decision tree, XGBoost has the advantages of fast running speed, good effect and processing large scale data. XGBoost decision tree is compared with method in the Ref. [3]. And the result shows that the accuracy of target true intention is greatly improved by using the XGBoost method.

1 Fundamental of XGBoost decision tree

XGBoost is an implementation of the Gradient Boosting machine algorithm, and it can do parallel computation, which makes the recognition real-time. The traditional gradient boosting decision tree (GBDT) algorithm only uses first order derivative information. When training the n -th tree, it needs to use the remnants of former $n-1$ trees. Therefore, it is difficult to achieve distributed computation. In this paper, square loss function is adopted, and the regular term is added to the loss function to obtain the optimal solution, which is useful for balancing the loss function, decreasing the complexity of the mode and avoiding overfitting.

Received date: 2018-01-03

Corresponding author: WANG Lei (736833825@qq.com)

1.1 Defining objective function of XGBoost

The objective function is defined as

$$F^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}^{(t-1)} + f_t(x_i)) + \Omega(f_t) + \text{const}, \quad (1)$$

where

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2, \quad (2)$$

where const is constant; γ is number of leaf nodes; ω_j is weight of nodes; λ is parameter of the L2 regularization term that controls the weight of the model complexity.

Replacing $l(y_i, \hat{y}^{(t-1)} + f_t(x_i))$ with the square loss function, $F^{(t)}$ is obtained approximately as

$$F^{(t)} = \sum_{i=1}^n (y_i - (\hat{y}^{(t-1)} + f_t(x_i)))^2 + \Omega(f_t) + \text{const} = \sum_{i=1}^n \left[l(y_i, \hat{y}^{(t-1)} + g_i f_t(x_i)) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) + \text{const}, \quad (3)$$

where $g_i = \partial_{\hat{y}^{(t-1)}} (\hat{y}^{(t-1)} - y_i)^2 = 2(\hat{y}^{(t-1)} - y_i)$, $h_i = \partial_{\hat{y}^{(t-1)}}^2 (y_i - \hat{y}^{(t-1)})^2 = 2$.

1.2 Defining complexity of decision tree

Defining $f_t(x) = \omega_{q(x)}$, $\omega \in R^T$, $q: R^d \rightarrow \{1, 2, \dots, T\}$, substituting Eq. (2) into Eq. (3) and removing the constant item, the following result is obtained as

$$F^{(t)} \approx \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) = \sum_{i=1}^n \left[g_i \omega_{q(x_i)} + \frac{1}{2} h_i \omega_{q(x_i)}^2 \right] + \gamma T + \lambda \frac{1}{2} \sum_{j=1}^T \omega_j^2 = \sum_{i=1}^n \left[\left(\sum_{i \in I_j} g_i \right) \omega_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) \omega_j^2 \right] + \gamma T, \quad (4)$$

where ω is the leaf weight of tree; q is the structure of tree. The instance set in leaf j is defined as $I_j = \{i | q(x_i) = j\}$ and $I_j \in I$.

Let $G_i = \sum_{i \in I_j} g_i$, $H_j = \sum_{i \in I_j} h_i$, so Eq. (4) can be simplified as

$$F^{(t)} = \sum_{j=1}^T \left[\left(\sum_{i \in I_j} g_i \right) \omega_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) \omega_j^2 \right] + \gamma T = \sum_{j=1}^T \left[G_j \omega_j + \frac{1}{2} (H_j + \lambda) \omega_j^2 \right] + \gamma T. \quad (5)$$

Since the derivative of $F^{(t)}$ with respect to ω_j equals to 0, the general result is

$$\omega_j^* = -\frac{G_j}{H_j + \lambda}. \quad (6)$$

Substituting Eq. (6) into Eq. (5), F is obtained as

$$F = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T. \quad (7)$$

The smaller the value of F is, the better the structure of the tree is.

1.3 Building XGBoost decision tree

For every leaf node, the greedy method is used to add a segmentation to split the leaf node into two child nodes from the depth 0 of the decision tree to depth n. After adding the segmentation point, the function is considered as

$$G_{\text{ain}} = \frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G)^2}{H + \lambda}, \quad (8)$$

where $\frac{G_L^2}{H_L + \lambda}$ means score of the left child leaf node, while $\frac{G_R^2}{H_R + \lambda}$ is score of the right child leaf node, and $\frac{(G)^2}{H + \lambda}$ is score obtained without segmentation.

There are 6 kinds of features including altitude, speed, distance, relative heading angle and the situation of the search radar and attack radar. In this paper, the target intention is classified as attack, penetration, scout and retreat.

For each feature, the nodes are sorted according to the features values. Then each leaf node is split. For every split step, all the segmentation schemes are traversed to find the optimal segmentation point. The pseudo-code is shown as follows:

Input: I , instance set of current node

Input: m , feature dimension

$G_{\text{ain}} \leftarrow 0$

$G \leftarrow \sum_{i \in I} g_i$, $H \leftarrow \sum_{i \in I} h_i$

for $k=1$ to m do

$G_L \leftarrow 0$, $H_L \leftarrow 0$

for j in sorted (I , by x_{jk}) do

$G_L \leftarrow G_L + g_j$, $H_L \leftarrow H_L + h_j$

$G_R \leftarrow G - G_L$, $H_R \leftarrow H - H_L$

score $\leftarrow \max(\text{score}, \frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{G^2}{H + \lambda})$

end

end

Output: Split with max score

The tree is considered as a function which maps the features to intent. Then the probability of intent is

calculated by using logistic regression method.

1.4 Data standardization

Due to the inconformity of the data's units, it is necessary to standardize the data to 0–1. The method adopted in this paper of Min-Max is

$$X = \frac{x - \min_A}{\max_A - \min_A}, \quad (9)$$

where \max_A is the maximum value, while \min_A is the minimum value.

1.5 Dempster-Shafer combination rule

It is possible to provide evidences for multiple sources for a given frame of discernment. It is necessary that all sources should be independent. To combine belief functions with Dempster-Shafer, the basic probability assignment is used. The combination rule's numerator is as

$$m(A) = \begin{cases} \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1 - K}, & A \neq \emptyset, \\ 0, & A = \emptyset \end{cases} \quad (10)$$

where $K = \sum_{B \cap C = \emptyset} m_1(B)m_2(C)$.

The whole process is shown in Fig. 1.

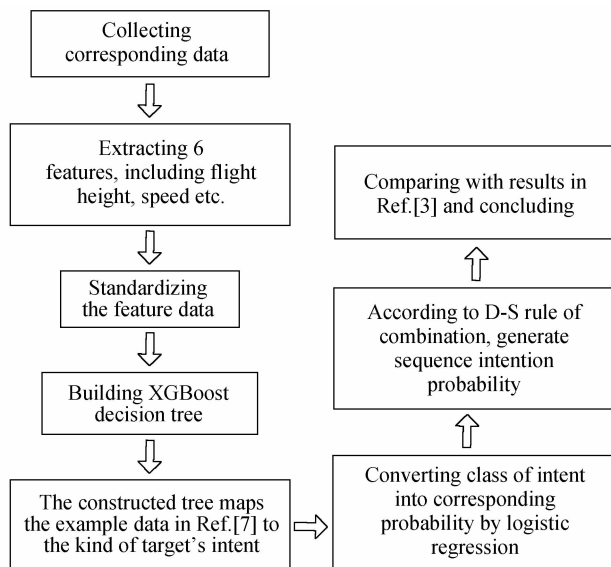


Fig. 1 Whole process of intention recognition

2 Experiment and results

Part of the data is assumed in Table 1, and the data is collected at different time by the sensor.

In the experiment, the objective intention space set by expert and its standard values^[7] including height, speed, distance, relative heading angle, situation of

the search radar and the attack radar are adopted to classify data of the Table 1 into four kinds of intentions including attack, penetration, retreat and scout.

Table 1 Data collected by sensor

Height (m)	Speed (m/s)	Distance (km)	Relative heading angle (°)	Search radar on ((1) or not (0))	Guidance radar on ((1) or not(0))
400	300	200	0	1	1
400	300	200	0	0	1
150	280	300	0	0	0
3 000	250	300	0	1	0
200	250	200	180	0	0
2 500	250	300	0	1	0

Since the standard value has a continuous or discrete range, each row in Table 1 is classified according to its range respectively. The classification result is shown in Table 2.

Table 2 Classification results

Classification	Height (m)	Speed (m/s)	Distance (km)	Relative heading angle (°)	Search radar (on (1) or not (0))	Guidance radar (on ((1) or not (0))
Attack	400	300	200	0	1	1
Attack	400	300	200	0	0	1
Penetration	150	280	300	0	0	0
Scout	3000	250	300	0	1	0
Retreat	200	250	200	180	0	0
Scout	2 500	250	300	0	1	0

The decision tree is constructed by the method introduced in section 1.3 and maps the example data to target intent. Then the probability of enemy target intent is calculated by logistic regression. Based on the python package provided by XGBoost, this paper uses the interface of XGBClassifier to build the XGBoost decision tree. And the whole process is programmed with python. The first column is considered as label data, and other columns as train data. Before the data of Table 2 used as the input data, the first column data needs to be quantified. Therefore, intentions of attack, penetration, scout, retreat are encoded as 0, 1, 2 and 3, respectively. And other column data is standardized according to section 1.4. Part of the codes are listed as follows. Num_class represents classified categories, while max_depth is the maximum depth of tree, and other parameters are default.

```

# load the train data, label data and target data;
train=table 2's data except for the first column
label=the first column data of table 2
target=example data in Ref. [7]
# use interface XGBClassifier to build the XGBoost decision

```

```

tree;
XGBoost=XGBClassifier(
    num_class=4,
    max_depth=6,
    min_child_weight=1,
    gamma=0,
    objective='multi:softprob')
# train the XGBoost decision tree;
XGBoost.fit(train, label)
# predict the probability of target intent;
yprob=XGBoost.predict(target).reshape(target.shape[0],
4)

```

The calculated results (yprob) are shown in Table 3. And t_0, t_1, t_2, t_3 represent arbitrary time series, as in Ref. [7].

Table 3 Probability distributions

Time	Target	Attack	Penetration	Scout	Retreat
t_0	01,02,03	0.229 3	0.315 2	0.227 7	0.227 7
t_1	01	0.180 8	0.183 7	0.456 0	0.179 5
	02,03	0.201 6	0.420 5	0.188 9	0.188 9
t_2	01	0.180 8	0.183 7	0.456 0	0.179 5
	02,03	0.459 4	0.181 5	0.179 6	0.179 6
t_3	01	0.180 8	0.183 7	0.456 0	0.179 5
	02,03	0.459 4	0.181 5	0.179 6	0.179 6

According to Dempster-Shafer rule of combination, the intent distribution probability from t_0 to t_3 in Table 3 is sequentially synthesized firstly. Then the sequential intention result is obtained in Table 4. And the maximal probability represents the real target intent at current time. According to the comparison with Table 3 in Ref. [3], it can be found that the true intent recognized by the presented method is the same as method in Ref. [3] from t_0 to t_3 .

Table 4 Target sequence intention probability

Time	Target	Attack	Penetration	Scout	Retreat
t_0	01,02,03	0.229 3	0.315 2	0.227 7	0.227 7
t_0-t_1	01	0.169 9	0.237 2	0.425 4	0.167 5
	02,03	0.174 6	0.500 5	0.162 4	0.162 4
$t_0-t_1-t_2$	01	0.103 0	0.146 1	0.650 2	0.100 8
	02,03	0.349 7	0.396 0	0.127 2	0.127 2
$t_0-t_1-t_2-t_3$	01	0.015 1	0.074 5	0.823 5	0.050 2
	02,03	0.577 4	0.258 3	0.082 1	0.082 1

To test the performance of XGBoost, the comparisons with method in Ref. [3] are shown in Figs. 2 and 3. Fig. 2 presents the target 01's intent probability, and intent probabilities of target 02 and 03 are shown in Fig. 3. The full line and dotted line

represent XGBoost and Ref. [3]'s method, respectively.

It can be observed that the trend of intent is consistent. Fig. 2 shows that target 01's scout intent trend rises from time t_0 to t_3 , while others decrease. The probability of scout is 0.823 5 at t_3 , and it is higher than 0.525 1 of method in Ref. [3]. As to Fig. 3, it is obvious that the probabilities of penetration increases from t_0 , while the probabilities of other three intentions decrease. However, the probability of attack intention increases and reaches 0.577 4 from t_1 to t_3 , which is higher than 0.384 6 of method in Ref. [3].

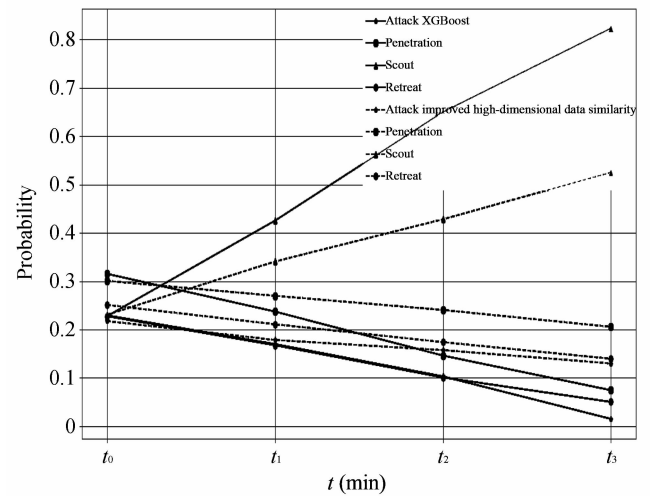


Fig. 2 Intention probability of target 01

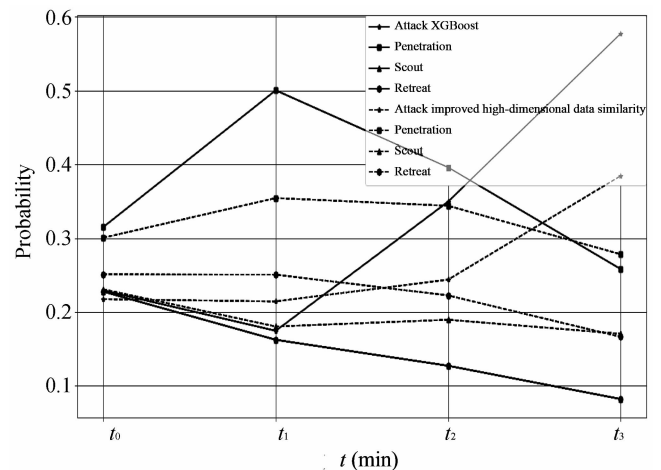


Fig. 3 Intention probabilities of targets 02 and 03

For further comparison, Tables 5 and 6 are listed. The second columns of the Tables 5 and 6 represent the target's real intention probabilities calculated by XGBoost, and the third column is the Ref. [3]'s method. It is obvious to know that the accuracies have increased 29.48% and 19.28% at t_3 , respectively.

Table 5 Comparison of target 01's real intention probability

Time	XGBoost	Ref. [3]'s method	Improved accuracy (%)
t_0	0.315 2	0.300 7	1.45
t_0-t_1	0.425 4	0.340 8	8.46
$t_0-t_1-t_2$	0.650 2	0.428 2	22.20
$t_0-t_1-t_2-t_3$	0.823 5	0.525 1	29.48

Table 6 Comparison of real intention probabilities of targets 02 and 03

Time	XGBoost	Ref. [3]'s method	Improved accuracy (%)
t_0	0.315 2	0.300 7	1.45
t_0-t_1	0.500 5	0.354 5	14.60
$t_0-t_1-t_2$	0.396 0	0.344 0	5.20
$t_0-t_1-t_2-t_3$	0.577 4	0.384 6	19.28

3 Conclusion

In this paper, a method based on XGBoost to predict the target intent is presented. To test the performance of this method, comparisons with the method presented in Ref. [3] are carried out based on the same data in Ref. [7]. Since the target intent has been recognized in Ref. [7], the goals of this paper and Ref. [3] are to improve the accuracy under the condition of ensuring true recognition. And the experiment result shows that our method performs

better than method in Ref. [3] in accuracy of recognizing intent, which implies that XGBoost method can provide valuable solution to evaluate the situation of modern war.

References

- [1] Yuan Z J, Xu Z G, Deng S H. The model for predication of fighting intending under sequential game. *System Engineering Theory and Practice*, 1997, 7(7): 72-78.
- [2] Li W S, Wang S B. Situation assessment based on bayesian networks. *System Engineering Theory and Practice*, 2003, 25(4): 480-483.
- [3] Cao S Y, Liu Y A, Xue S. Target tactical intention recognition method of improved high-dimensional data similarity. *Transducer and Microsystem Technologies*, 2017, 36(5): 25-28.
- [4] Yao Q K, Liu S J, He X Y. Research and prospect of battlefield target operational intention recognition. *Journal of Command and Control*, 2017, 3(2): 127-131.
- [5] Chen T, Guestrin C. Xgboost: A scable tree boosting system. In: *Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.
- [6] Yang F B, Wang X X. D-S evidence theory conflict evidence synthesis method. Beijing: National Defense Industry Press, 2010.
- [7] Sun Y L, Bao L. Study on recognition technique of targets' tactical intentions in sea battlefield based on D-S evidence theory. *Ship Electronic Engineering*, 2012, 32(5): 48-51.

基于 XGBoost 决策树的空中目标战术意图识别

王 磊, 李世中

(中北大学 机电工程学院, 山西 太原 030051)

摘 要: 为了提高空中目标战术意图识别的准确度, 提出了一种基于 XGBoost 决策树的目标意图识别方法。以 python 为开发工具, 基于文献中的数据得到战术意图识别概率, 并依靠 Dempster-Shafer 证据合成理论得出贯序意图概率。为了验证识别结果的准确性, 将其与文献中的实验结果进行对比。结果表明, 该方法可提高目标真实意图的识别概率, 故该方法是可行的。

关键词: 目标战术意图识别; XGBoost 决策树; Dempster-Shafer 证据合成理论

引用格式: WANG Lei, LI Shi-zhong. Tactical intention recognition of aerial target based on XGBoost decision tree. *Journal of Measurement Science and Instrumentation*, 2018, 9(2): 148-152. [doi: 10.3969/j.issn.1674-8042.2018.02.008]