Network selection algorithm based on decision tree in heterogeneous wireless networks

Wang-hong Li, and Zhu Qi*

Jiangsu Key Lab of Wireless Communications, Nanjing University of Posts and Telecommunications,
Nanjing 210003, China

Abstract. A network selection algorithm based on Decision Tree is proposed to solve the problem. Users can select the appropriate network according to their service characteristics and requirements when they decide which network to access. First, we get the training data under the Interactive Class service from the synergetic algorithm which can be used for training set. The network attributes are used for attribute set. And then we can choose the attribute with the largest information gain as the division attribute after the discretization of continuous features by the bisection method. Keep going this step recursively, we can finally get a decision tree with high generalization ability by which we can make the network selection. Simulation results show that the algorithm we proposed is simple and effective and demonstrate the effectiveness of our scheme in improving the quality of service according to the user requirements under the Interactive Class service.

1 Introduction

Today, there are lots of wireless networks with different characteristics around the users, such as wireless local area network (WLAN) with high data transmission rates but limited coverage and Cellular networks with wide coverage but high prices. Therefore, it is an important research challenge of wireless communication to design an efficient network access scheme to provide the users with high quality of service.

A number of achievements have been made in the study of network selection algorithms. Reference [1] proposes a network selection algorithm based on group decision and utility function, which can guarantee the quality of service under different kinds of services. Reference [2] puts forward an improved multi-attribute decision network access selection algorithm, trying to reduce the times of the hand off process. A novel network selection scheme, combined with Technique for Order Preference by Similarity to Ideal Solution(TOPSIS) and Grey Relational Analysis(GRA), is put forward to make the network selection in Ref. [3]. Reference [4] computes the attributes weights with Analytic Hierarchy Process (AHP) and Entropy Weight method (EW) first and then decide which network to access by TOPSIS which can reduce blocking probability without improving the network load. Reference [5] puts forward a network selection scheme based on unsupervised learning decision tree. However, it is too simple to give the algorithm performance analysis.

* Corresponding author: zhuqi@njupt.edu.cn

Also, the unsupervised learning algorithms have poor classification accuracy compared with supervised learning.

This paper proposes a novel network selection algorithm based on the decision tree method in supervised learning. First, we get the training data under the Interactive Class service based on Ref. [6] and think of the network attributes as the attribute set. Second we use them as the input of the proposed algorithm and take the bisection method to discretize the continuous attributes. Then we choose the attribute that maximizes the information gain as the segmentation feature and keep going this step recursively to get a decision tree with high generalization ability by which we can make the network selection. The algorithm we proposed is simple and effective, which can improve the quality of service according to the user requirements.

2 System model

The UMTS has wide coverage, low transmission rate, low packet delay, low packet jitter and high price. The WLAN has small coverage, larger transmission rate than UMTS, lower price and larger delay and jitter than UMTS. The transmission rate provided by the WiMAX is about the same as WLAN and the price is higher compared with WLAN. Users are randomly distributed in the areas covered by these networks and they need to decide which network to access according to their service requirements.

The network selection problem can be seen as a multi-classification problem. In the heterogeneous wireless network environment with multi networks coexist, the users need to classify the candidate networks according to their requirements and then select the network which satisfy the demands most. This paper takes the available data rate a_1 , peak data rate a_2 , packet delay a_3 , packet jitter a_4 , packet loss a_5 and cost per bit a_6 under the Interactive Class service into account to study the network selection problem base on decision tree.

This paper takes the decision tree method to make the network selection, which uses the given training data $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ and the attribute set $A = \{a_1, a_2, \dots, a_n\}$ as the inputs of the algorithm, keeps choosing the optimum feature a_i ($1 \leq i \leq n$) from the n network attributes recursively, and then with which we can split the training data to get a mapping from the attribute value to the network classification. Finally, we get the decision tree T , with which we can classify the networks that are not in the training data. The mapping relationship can be expressed as:

$$f : X \rightarrow Y \quad (1)$$

where $X = \{x_1, x_2, \dots, x_m\}$ represents m training samples, $Y = \{1, 0\}$ represents the classification result and number one means the good network while number zero means the bad, (x_k, y_k) represents the value of the network attributes and the classification result for the given sample data k , a_i represents the i -th attribute of the network.

3 Network selection algorithm based on decision tree

Decision tree is a kind of common machine learning algorithm, which can learn a model with high generalization ability from a given training set to classify the new instances. The advantages of the proposed algorithm are easy to be read, indicating that it is easy to make the manual analysis and it only needs to be built once and used repeatedly. The maximum

depth of each decision does not exceed the depth of the decision tree. As a result, the proposed algorithm is simple, effective and is of lower complexity compared with other algorithms. The detailed steps of the decision tree algorithm include obtaining training data, feature selection, decision tree generation and so on.

3.1 Obtain the training data

By learning the given training data, the decision tree obtains a mapping relation model, reflecting the relationship between the network attribute value and the network classification result, with which to make the following network selection. Therefore, it is reasonable to hope that the classification ability of the given training data performs well, in order to improve the classification accuracy of the model to accommodate the networks that are not in the training data. The ability of the learning model to adapt the new samples is called generalization. A model with high generalization can be well applied to the whole sample space. So, the training data should be able to reflect the characteristics of the sample space well despite the fact that the training set is usually just a small sample of the sample space. Generally, the more the training data is, the more likely the learning model can reflect the characteristics of the sample space well. This paper obtains the training data based on Ref. [6] to analyze the process of establishing a decision tree.

3.2 Feature selection

The step of feature selection prefers to choose the feature with high classification ability for the training data, namely, to decide which attribute of the network has the best ability to compartmentalize, in which way we can improve the efficiency of the decision tree learning. If there is no significant difference between the result of the classification using a feature and the result of the random classification, then the feature is said to have no classification ability. Therefore, it is important to choose an appropriate feature and we use information entropy to select the feature.

3.2.1 Information entropy

Information entropy is a measure of the uncertainty of random variable, which can be used to measure the purity of the sample set, represents the proportion of each kind of network in the sample. We assume that Y represents the classification result of the network and p_k represents the proportion of the k -th kind network in the training data, that is:

$$P(Y = k) = p_k, k = 0, 1 \quad (2)$$

where p_1 represents the proportion of the good network and p_0 represents the proportion of the bad network, so the information entropy of the classification result Y is defined as:

$$H(Y) = -\sum_{k=0}^1 p_k \log p_k \quad (3)$$

And we define $0 \log 0 = 0$. When the base of the logarithm is 2, the unit of entropy is *bit*. It can be seen from the definition that the small the $H(Y)$ is, the more pure the training data is.

3.2.2 Conditional entropy

On the assumption that the discrete network attribute a has V possible values $\{a^1, a^2, \dots, a^V\}$, we use attribute a to split the training data D to generate V branch nodes and the n -th branch node includes the networks in which the value of the attribute a equals to a^v , denoted by D^v . Taking the fact that different branches contain a different number of networks into account, we assign different weight $\frac{|D^v|}{|D|}$ to the branch nodes, indicating that the more networks in the branch, the greater the impact of the branch nodes. Then we can get the conditional entropy $H(D|a)$ obtained by splitting the training data with attribute a , which is defined as:

$$H(D|a) = \sum_{v=1}^V \frac{|D^v|}{|D|} H(D^v) \quad (4)$$

3.2.3 Information gain

Information gain represents the reduction of the uncertainty degree of the classification result Y after we get the knowledge of the feature a . The information gain of the feature a for training data D is defined as the difference between $H(D)$ and $H(D|a)$, denoted by:

$$g(D, a) = H(D) - H(D|a) \quad (5)$$

The difference is also called the mutual information and the network attributes with large information gain have stronger classification ability.

3.2.4 Deal with the continuous attributes

The discussion above is aimed at dealing with the discrete attributes in the decision tree. However, the network attributes mentioned above are continuous and the number of continuous attributes is no longer limited. Therefore, the nodes can not be divided according to the values of the continuous attributes and we use bisection method to discretize the continuous attributes, which is used in the C4.5 decision tree algorithm.

Given a training data D and the continuous attribute a with the assumption that there are n different values of the attribute a in the training data D , we sort these values from small to large, denoted by $\{a^1, a^2, \dots, a^n\}$. We can divide the training data D into D_t^- and D_t^+ based on the divide point t and D_t^- includes the networks whose value of the attribute a is smaller than t and D_t^+ is on the contrary. It is obvious that the arbitrary value of t between the interval $[a^i, a^{i+1}]$ produces the same result of the division for the adjacent attribute value a^i and a^{i+1} . As a result, we can formulate the $n-1$ candidate partition points as:

$$T_a = \left\{ \frac{a^i + a^{i+1}}{2} \mid 1 \leq i \leq n-1 \right\} \quad (6)$$

where we use the midpoint of the interval $[a^i, a^{i+1}]$ to be the candidate partition point and then we can study these partition points like discrete attribute values. We need to select the

optimal partition point to divide the training data. As a result, we should modify the formula (5) as follows:

$$g(D,a)=\max_{t\in T_a} g(D,a,t)=\max_{t\in T_a} (H(D)-H(D|a,t)) \tag{7}$$

where $g(D,a,t)$ represents the information gain of the training data based on the partition point. And then we should choose the partition point with the largest $g(D,a,t)$.

3.3 Generate the decision tree

Now, we use the given training data $D=\{(\mathbf{x}_1,y_1),(\mathbf{x}_2,y_2),...,(\mathbf{x}_m,y_m)\}$ under the Interactive Class service and the attributes set $A=\{a_1,a_2,...,a_n\}$ to generate the decision tree in figure 1.

Fig. 1. Decision tree under the Interactive Class service.

4 Performance analysis

Assume that the simulation scenario is integrated by three kinds of networks, namely UMTS, WLAN and WiMAX. There are two networks with different parameters of each type and six networks in total. Supposing that the users range of movement is within the coverage of six networks.

In order to make the dynamic network selection, the Markov chain is used to represent the change of network attributes parameters. Assume that the number of Markov states of the six network attributes are 10, 10, 20, 20, 30 and 10, respectively. To evaluate the performance of the algorithm we proposed, the group decision algorithm is chosen to make a comparison.

It is illustrated in figure 2 that the average handoff number per unit time of the proposed algorithm and group decision change as the transposition probability P under the Interactive Class service. It is shown that the algorithm we proposed reduces the average handoff number most. This is because we tend to select the network with high stability degree based on the model we have learned and the optimal network is likely to remain the best at the next moment. When the performance is the same, the decrease in handoff number can effectively reduce the power consumption of the terminal, release the network instability caused by frequent handoff and improve the quality of user service significantly. Therefore, the proposed algorithm outperforms the group decision algorithm proposed in ref. [7].

Fig. 2. Handoff Number.

Fig. 3. Throughput.

Figures 3, 4, and 5 illustrate the simulation results that the throughput, packet delay and packet loss of the proposed algorithm and the group decision algorithm change as the transposition probability p under Interactive Class service. As can be seen from figures 3 and 4, the throughput and packet delay of the proposed algorithm perform best. This is because some Interactive Class services set higher requirements in delay and throughput. However, the demand of packet loss is low. So the packet loss of the proposed algorithm is worse than the group decision in figure 5. Taken together, the proposed algorithm outperforms the group decision algorithm in the Interactive Class service on the whole.

Fig. 4. Packet delay.

Fig. 5. Packet loss.

5 Conclusions

This paper proposes a network selection algorithm based on decision tree algorithm which allows the users to decide which network to access according to their service characteristics and requirements. Simulation results demonstrate that, in the Interactive Class service, the throughput and the packet delay of the proposed algorithm perform best while the packet loss is worse. Also, the proposed algorithm reduces the average handoff number most.

Acknowledgments

This work is supported by National Natural Science Foundation of China (61571234 , 61631020)

References

1. Luo R, Zhao S, Zhu Q. Network Selection Algorithm Based on Group Decision Making for Heterogeneous Wireless Networks [C]. IEEE International Conference on Communications (2017)
2. Gao X E, Li K Q. Research on Heterogeneous Network Access Selection Algorithm Based on Improved Multiple Attribute [J]. Computer Science, **44** (2017)
3. Anhua Peng, Zhiming Wang. GRA-based TOPSIS decision-making approach to supplier selection with interval number [C]. Control and Decision Conference (2011)
4. Jiang Y, Gao H Y. Heterogeneous Network Access Selection Algorithm Based on Real-time multi-task service [J]. Computer Applications and Software, **33** (2016)
5. Wang Y, Zhang K. Decision tree based unsupervised learning to network selection in heterogeneous wireless networks [C]. Consumer Communications and NETWORKING Conference. IEEE (2011)
6. Lina Zhang, Qi Zhu. Multiple attribute network selection algorithm based on AHP and synergetic theory for heterogeneous wireless networks[J]. Journal of Electronics(China), **31** (2014)
7. Zheng Shi, Qi Zhu. A network selection algorithm based on multiple attribute decision making and group decision making for heterogeneous wireless networks[J]. The Journal of China Universities of Posts and Telecommunications, **19** (2012)