

[Article ID] 1003- 6326(2001) 01- 0128- 04

## Decision tree method applied to computerized prediction of ternary intermetallic compounds<sup>①</sup>

HUANG Xin(黄欣), YANG Jie(杨杰), YE Chen-zhou(叶晨洲), CHEN Nian-yi(陈念贻)  
(Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University,  
Shanghai 200030, P. R. China)

**[Abstract]** Decision tree method and atomic parameters were used to find the regularities of the formation of ternary intermetallic compounds in alloy systems. The criteria of formation can be expressed by a group of inequalities with two kinds of atomic parameters  $Z_i$  (number of valence electrons in the atom of constituent element) and  $R_i/R_j$  (ratio of the atomic radius of constituent element  $i$  and  $j$ ) as independent variables. The data of 2238 known ternary alloy systems were used to extract the empirical rules governing the formation of ternary intermetallic compounds, and the facts of ternary compound formation of other 1334 alloy systems were used as samples to test the reliability of the empirical criteria found. The rate of correctness of prediction was found to be nearly 95%. An expert system for ternary intermetallic compound formation was built and some prediction results of the expert system were confirmed.

**[Key words]** ternary intermetallic compound; decision tree; computerized prediction

**[CLC number]** TG 146

**[Document code]** A

### 1 INTRODUCTION

Ternary intermetallic compounds are common alloy phases in many alloys of nonferrous metals. Some ternary intermetallic compounds such as  $\text{BF}_{14}\text{Nd}_2$  have been found to be very useful functional materials<sup>[1,2]</sup>. Hence the prediction and exploration of new ternary intermetallic compounds are concerned by metallurgists and materials scientists. Although the method of estimation of the chemical affinity between two metallic elements have been proposed by Miedema and one of the present authors, and these methods are rather useful for binary intermetallic compound prediction<sup>[3,4]</sup>; these criteria cannot be simply extended to predict the formation of ternary intermetallic compounds since the formation of ternary intermetallic compound depends not only on the chemical affinity between its constituent elements, but also on whether the ternary intermetallic compound or the corresponding binary intermetallic compounds are more stable.

In our previous works, an extended Miedema's model using four kinds of atomic parameters (Miedema's electronegativity  $\chi$ , electronic density in Wagner-Seitz cell ( $n_{\text{ws}}$ )<sup>1/3</sup>, atomic radius  $R$ , and the number of valence electrons in the atom of constituent element  $Z$ ) has been found to be an appropriate semi-empirical model for description of the regularity of ternary intermetallic compound formation<sup>[5,6]</sup>. Since there are three elements involved in ternary intermetallic compound formation, 12 independent variables must be used to build the mathematical model. In this work, however, it has been found that only 5

independent variables related to two kinds of atomic parameters (atomic radius and number of valence electrons) and their functions are already enough to describe the regularities of ternary intermetallic compounds, by using decision tree method.

### 2 METHOD OF COMPUTATION

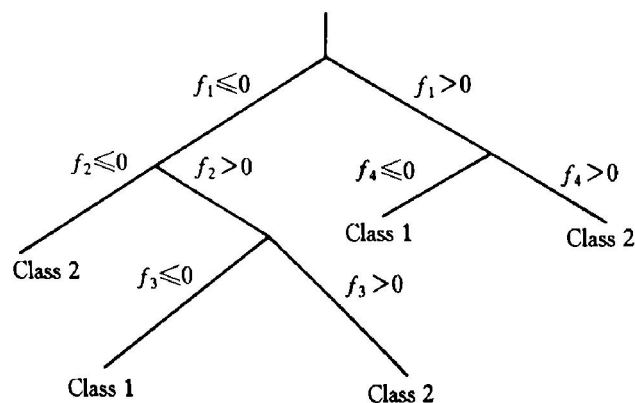
Decision tree method is one of the most popular methods of data processing dealing with complicated, multivariate problems<sup>[7,8]</sup>. In this method a logic tree is built to classify or predict samples belonging to two or more classes. In our work, the ternary compound-forming alloy systems are defined to be the samples of class 1, while that without ternary compound formation are defined as the samples of class 2. The whole sample is tested by some criteria through several steps. So that a sample is classified by tracing a path from the "root" to the "leaf" of the logic tree. For example, the systems consisting of two transition elements and one nontransition element can be classified by the logic tree illustrated in Fig. 1.

In the computation of decision tree, the logic tree should be optimized to make it to be the simplest one. Gini Index<sup>[7]</sup> is used as the criterion of "goodness" of the criteria of classification. When a training set is split into two subsets  $T_L$  and  $T_M$ , we define

$$\text{Goodness} = \left( \frac{|T_L|}{n} \times \text{Gini}L + \frac{|T_M|}{n} \times \text{Gini}M \right) / n$$

$$\text{Gini}L = \sum_{i=1}^k \left( \frac{L_i}{|T_L|} \right)^2$$

$$\text{Gini}M = \sum_{i=1}^k \left( \frac{M_i}{|T_M|} \right)^2$$



**Fig. 1** Typical decision tree

where  $|T_L|$  ( $|T_M|$ ) is the number of samples on the left (right) of a split at a tree node  $T$ ,  $n$  is the number of samples at node  $T$ ; and  $L_i$  ( $M_i$ ) is the number of samples in class  $i$  on the left (right) of the split;  $k$  is the number of classes.

In this work, linear equations with following independent variables are used as criteria for ternary intermetallic compound formation:

$Z_1$  is number of valence electrons of first element;  $Z_2$  is number of valence electrons of second element;  $Z_3$  is number of valence electrons of third element;  $R_2/R_1$  is atomic radius ratio between the second and first element;  $R_3/R_1$  is atomic radius ratio between the third and first element.

The definition of  $Z$  for nontransition elements is self-evident. The definition of  $Z$  for transition elements in this work is the same as in our previous works<sup>[5,6]</sup>, i. e.,  $d$  electrons in the unfilled next outermost shell are also considered as valence electrons because these  $d$  electrons also take part in metallic bond formation through  $spd$  hybridization. The values of atomic radii are given by Teatum<sup>[9]</sup>.

In the computation of this work, the sample is a system consisting of three elements. In order to make the computation results self-consistent, the order of these three elements is defined as: 1) for the system with two transition elements and one nontransition element, the nontransition element is defined as the first element; 2) for the system with two nontransition elements and one transition element, the transition element is defined as the first element; 3) the order between nontransition elements or transition elements is arranged by descending numbers of valence electrons, or by descending values of Miedema electronegativity if the number of valence electrons is the same.

The program of decision tree method was developed by the Institute of Pattern Recognition of Shanghai Jiao Tong University. The computation of this work was performed by a 586 type of microcomputer. The data of ternary compound formation, used as training set, were quoted from Villar P Data Bank

of Ternary alloy systems and our previous papers<sup>[3]</sup>.

### 3 RESULTS OF COMPUTATION

#### 3.1 Systems consisting of three nontransition elements

For the systems of nontransition elements, the criterion of ternary compound formation is as

$$f(1) = 2.154 Z_1 + 2.155 Z_2 - 3.334 Z_3 + 0.356(R_2/R_1) - 7.222(R_3/R_1) + 0.241$$

$$f(2) = 0.649 Z_1 - 0.587 Z_2 - 0.183 Z_3 - 0.513(R_2/R_1) + 0.518(R_3/R_1) - 0.504$$

$$f(3) = 3.532105 Z_1 + 10.23 Z_2 - 3.532105 Z_3 + 9.044(R_2/R_1) - 6.086(R_3/R_1) - 7.064 \times 10^5$$

$$f(4) = 0.012 Z_1 - 0.845 Z_2 - 2.587 Z_3 + 3.347(R_2/R_1) + 2.364(R_3/R_1) - 1.027$$

If  $f(1) < 0$ ,  $f(2) < 0$ , then the sample should belong to class 2, i. e. when no ternary compound was formed.

If  $f(1) < 0$ ,  $f(2) \geq 0$  and  $f(3) \geq 0$ , then the sample should belong to class 1, i. e. it should form ternary intermetallic compound.

If  $f(1) < 0$ ,  $f(2) \geq 0$  and  $f(3) < 0$ , then the sample should belong to class 2.

If  $f(1) \geq 0$ ,  $f(4) < 0$ , then the sample should belong to class 2.

If  $f(1) \geq 0$ ,  $f(4) \geq 0$ , then the sample should belong to class 1.

116 ternary alloy systems not included into the training set were used as test samples to test the reliability of the above-listed criterion. It was found that the rate of correctness of prediction is 94.0%.

#### 3.2 Systems of three transition elements

For the systems of this type, the criterion of ternary intermetallic compound formation is as

$$f(1) = 1.290 Z_1 - 0.192 Z_2 - 2.771 Z_3 + 0.560(R_2/R_1) + 0.020(R_3/R_1) - 0.207$$

If  $f(1) \geq 0$ , then the sample should belong to class 1.

Otherwise, it should belong to class 2.

375 ternary alloy systems not included into the training set were used as test samples to test the reliability of above-listed criterion. It was found that the rate of correctness of prediction was 98.4%.

#### 3.3 Systems of one transition element and two non-transition elements

For the systems of these type, the criterion of ternary intermetallic compound formation is as

$$f(1) = -0.067 Z_1 - 0.272 Z_2 - 0.198 Z_3 +$$

$$0.313(R_2/R_1) - 0.342(R_3/R_1) + 0.875$$

$$f(2) = 0.531Z_1 - 0.977Z_2 - 0.636Z_3 - 4.448(R_2/R_1) - 2.145(R_3/R_1) + 6.863$$

$$f(3) = -0.484Z_1 - 0.211Z_2 + 0.282Z_3 + 3.449(R_2/R_1) - 2.499(R_3/R_1) + 1.579$$

If  $f(1) \geq 0$ , then the sample should belong to class 2.

If  $f(1) < 0$ ,  $f(2) \geq 0$ , then the sample should belong to class 1.

If  $f(1) < 0$ ,  $f(2) < 0$  and  $f(3) \geq 0$ , then the sample should belong to class 1.

If  $f(1) < 0$ ,  $f(2) < 0$  and  $f(3) < 0$ , then the sample should belong to class 2.

149 ternary alloy systems not included into training set were used as test samples to test the reliability of the above-listed criterion. It was found that the rate of correctness of prediction was 93.3%.

### 3.4 System of one nontransition element and two transition elements

For the systems of this type the criterion of ternary intermetallic compound formation is as

$$f(1) = -0.670Z_1 + 0.100Z_2 + 0.458Z_3 - 1.866(R_2/R_1) + 0.159(R_3/R_1) + 0.969$$

$$f(2) = -0.889Z_1 + 1.017Z_2 + 1.843Z_3 + 1.073(R_2/R_1) + 2.880(R_3/R_1) - 13.08$$

$$f(3) = -0.485Z_1 + 1.281Z_2 - 14.40Z_3 + 204.3(R_2/R_1) + 108.7(R_3/R_1) - 330.6$$

$$f(4) = -45.13Z_1 + 177.8Z_2 - 144.2Z_3 - 314.5(R_2/R_1) - 324.9(R_3/R_1) + 1359.6$$

If  $f(1) < 0$ ,  $f(2) < 0$ , then the sample should belong to class 2.

If  $f(1) < 0$ ,  $f(2) \geq 0$  and  $f(3) < 0$ , then the sample should belong to class 2.

If  $f(1) < 0$ ,  $f(2) \geq 0$  and  $f(3) \geq 0$ , then the sample should belong to class 1.

If  $f(1) \geq 0$ ,  $f(4) \geq 0$ , then the sample should belong to class 2.

If  $f(1) \geq 0$ ,  $f(4) < 0$ , then the sample should belong to class 1.

493 ternary alloy systems not included into training set were used as test samples to test the reliability of the above-listed criterion. It was found that the rate of correctness of prediction was 97.2%.

## 4 EXPERT SYSTEM FOR TERNARY INTERMETALLIC COMPOUND PREDICTION

As mentioned above, there are still nearly 5% prediction results are wrong. For example, the pre-

diction for Cs-Mo-Pr, Cs-Nb-V, Cs-Nb-Zr, Cs-Pr-V systems gives wrong results. It is interesting to note that these samples with wrong prediction results are located in several small regions in the multi-dimensional space spanned by the atomic parameters used in this work (such as  $Z_i$  and  $R_i/R_j$ ). If one of these systems is added into the training set, all other systems will give correct prediction results. So that an expert system using more known samples as training set can give higher rate of correctness by "leaving-one method". Based on this idea, an expert system called "Ternary compound predictor" built. It is a module of a software for alloy phase prediction. Fig. 2 shows an interface of this software. You can input the chemical symbols of three metallic elements and click "predict", it will give you the answer whether the ternary intermetallic compound will be formed or not. For example, this expert system predicts that Ge-Pd-Ho, Ge-Ni-Ho, Ge-Pt-Ho, Ge-Ni-Lu and Ge-Ir-Ho systems all form ternary intermetallic compounds and that Al-Ni-Re system does not form ternary intermetallic compound. These prediction results have all been confirmed by very recently published experimental results<sup>[11, 13]</sup>.

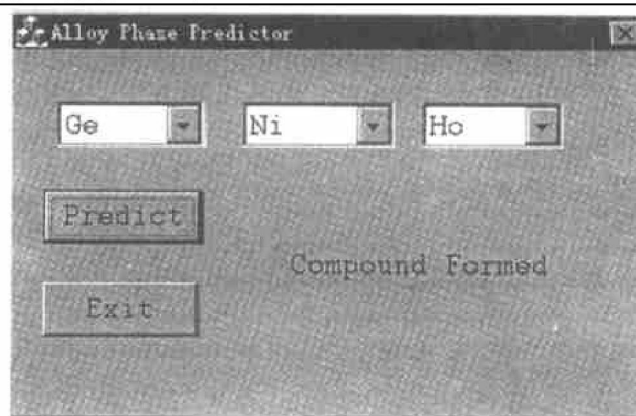


Fig. 2 Interface of "Alloy phase predictor"  
(Example of new alloy phase prediction)

## 5 DISCUSSION

It is generally accepted that three factors are the dominating factors affecting the formation and stability of alloy phases: 1) electrochemical factor: large difference of electronegativity ( $\chi$ ) leads to charge transfer and partial ionicity which stabilize the chemical bond between different kinds of atoms; 2) geometrical factor: certain radius ratio ( $R_i/R_j$ ) between different atoms favors some kinds of crystal lattices of alloy phases; 3) energy band factor: the number of valence electrons ( $Z_i$ ) is a factor influencing the band energy of alloy phases. So it is reasonable to use electronegativity, atomic radius ratio and the number of valence electrons as the atomic parameters to find the regularities of ternary intermetallic compound formation. Actually we have finished research works by

this way, and the results are rather good<sup>[10]</sup>. In this work, however, it has been found that using only the atomic parameters related to  $Z_i$  and  $R_i/R_j$  are already enough to build the mathematical model for ternary compound prediction. From the standpoint of computation methods, to use fewer atomic parameters is of course an advantage. The fact that  $Z_i$  and  $R_i/R_j$  are more generally accepted atomic parameters than  $\chi$  because they have clearer physical meaning, which makes this advantage still stronger.

By nonlinear regression, we have found that  $\chi$  is actually an approximate function of  $Z$  and  $R$ . Partial least square method (PLS) gives the following equation:

$$\chi = 6.190 + 0.0686Z - 1.783R + 9.965(Z/R)$$

The correlation coefficient between the Miedema's values of  $(n_{ws})^{1/3}$  and the values calculated by the equation is 0.926. Therefore some function of  $Z$  and  $R$  can be used to describe the influence of electrochemical factor without using the atomic parameter  $\chi$ . By similar regression method, another Miedema's atomic parameter  $(n_{ws})^{1/3}$  is also proved to be an approximate function as

$$(n_{ws})^{1/3} = 2.370 + 0.0179Z - 0.710R + 3.681(Z/R^3)$$

The correlation coefficient between the Miedema's values of  $(n_{ws})^{1/3}$  and the values calculated by the equation is 0.917.

## ACKNOWLEDGEMENT

This work is financially supported by the National High Tech Plan (863- 511- 945- 005), and we are grateful to Dr. P. Villars due to his cooperation and the data from his data bank.

## [ REFERENCES ]

- [ 1 ] ZHAO Sheng-xu, LI Shun-pu, ZHAO De-qian, et al. Solidification process of  $\alpha(\text{Al})\text{-Mg}_2\text{Ge}$  pseudobinary eutectic alloys [ J ]. Trans Nonferrous Metal Society of China, 1998, 8(1): 14- 17.
- [ 2 ] YANG Xiao-guang, ZHANG Xiao-bin, SHU Kang-ying, et al. Electrochemical properties of Zr-V-Ni system hydrogen storage alloys [ J ]. Trans Nonferrous Metal Society of China, 1998, 8(3): 408- 411.
- [ 3 ] Miedema A R. The electronegativity parameter of transition metals [ J ]. J Less Common Metals, 1973, 32: 117.
- [ 4 ] CHEN Niarr-yi. Parametric Functions of Chemical Bonds [ M ], (in Chinese). Scientia Sinica, 1975, 4(5): 399.
- [ 5 ] CHEN Niarr-yi, LU Wei-cong, CHEN Rui-liang, et al. Regularities of formation of ternary intermetallic compounds [ J ]. J of Alloys and Compounds, 1999, 289: 120- 133.
- [ 6 ] CHEN Niarr-yi, LI Chong-he, LIU Gang, et al. On the formation of ternary alloy phases [ J ]. J of Alloys and Compounds, 1996, 245(1/2): 179.
- [ 7 ] Breiman L, Friedman J H, Olshen R A, et al. Classification and Regression Trees [ M ]. Belmont: Wadsworth, 1984.
- [ 8 ] Quinlan J R. Decision trees and decisionmaking [ J ]. IEEE Transaction on Systems, Man and Cybernetics, 1990, 20(2): 339- 346.
- [ 9 ] Pearson W B. The Crystal Chemistry and Physics of Metals and Alloys [ M ]. New York: Wiley-Interscience, 1972.
- [ 10 ] CHEN Niarr-yi. Software package "Material Designer" and its application in materials research [ A ]. Proceeding of Intelligent Processing of Materials and Manufacture [ C ]. Hawaii, USA, 1999. 15- 20.
- [ 11 ] Kozerlyn G M, Bodak O I, Pavlyuk V V, et al. The Lu-Ni-Ge system [ J ]. J of Alloys and Compounds, 1999, 291: 201.
- [ 12 ] Cornish L A and Witcomb M J. A metallographic study of Al-Ni-Re phase diagram [ J ]. J of Alloys and Compounds, 1999, 291: 145.

( Edited by LONG Huai-zhong )