A novel distributed network database mapping scheme analysis

Liangong Song*, Ke Han

North China University of Water Resources and Electric Power, Zhengzhou, Henan, China

Received 1 July 2014, www.cmnt.lv

Abstract

This research is based on distributed client-server model, and intends to add lacking fields in the target database by mapping in order to complete the database for subsequent analysis and tab the maximum economic value of a single database. This research takes China Industry, Commerce and Service Census Database (TTICSCDB) as the target database, and China Technology Innovation Survey Database (TTTISDB) as the assistant database. Then this research will employ techniques to mapping common fields from the assistant database to the target database and evaluate the accuracy of mapping. We will prove from the statistical perspective the simulation and accuracy of extension field of the network database proposed in this research. This network database will present its advantages in many ways.

Keywords: database, mapping, statistics, expanding

1 Introduction

In the age of knowledge economy, it is important to make sure of the accuracy of the information from the database [1-3]. In the course of data mining, one cannot get accurate or useful knowledge without correct or accurate fields in the database [27, 28]. Therefore, the extensibility and reliability of a database are significant [1, 4-7]. With the development of the internet, database has ever become independent from the computer system and developed into a giant and meticulous system group. Thus, this research is based on distributed client-server model [8-11], and intends to add lacking fields in the object database by mapping in order to complete the database. This system architecture includes five parts. First, distributed network database. Second, automation pre-processing mechanism. Third, create module relation of two network database. Fourth, map the module. Fifth, threshold statistics and testing mechanism. The first part will present the whole distributed network database. The second part will interpolate missing values and correct error values, and unify field and format. The third part will do statistics of the index field to verify the reliability of the network database. The fourth part will base relation on the index field and apply it to four disaggregated model namely, linear network, probabilistic neural network, radial basis function network and back-propagation neural network. The fifth part will create a statistics and testing mechanism and set up its threshold and select the best approach and the mapping module with the high accuracy from the four disaggregated model, and map the field to the object database.

The rest of the paper is organized as: Section 2 will introduce the whole network database, its extensibility and limitations. Section 3 will introduce the background information of two database. Section 4 will account details of the automation pre-processing mechanism. Section 5 will account practical method of establishing module relation. Section 6 will compare four ways of mapping and analyse their practical result which includes accuracy and related efficacy analysis. Section 7 will present the conclusion of the re-search and discuss further improvement. Section 8 will conclude the contribution of this research.

2 An overview on system architecture

2.1 LIMITATIONS OF THE DATABASE

Field for mapping should be equipped with effective index. Otherwise, the disaggregated model will lack its accuracy and the extension field will lack its reliability. Matrix of the object database and the assistant database should be the same as this research adopts Mann-Whitney U test [12-14] to judge their similarity. If the matrix are not the same, it may result in a disordered corresponding relationship that the mapping module cannot apply to the object database.

2.2 DISTRIBUTED NETWORK DATABASE

Distributed network database is more and more important with the popularity of the internet. The distributed network database designed by us is easily established on other software database. Object can be taken as a basis of data processing and data exchange but real data exchange is operated through communication hardware in the way of bit stream transmission in which the object and Bit stream switch to each other under the coordination of Agent, DNDM and Operation System.

***Corresponding author** e-mail: slg@ncwu.edu.cn

2.3 AUTOMATION PRE-PROCESSING MECHANISM

Processing missing values: This research will employ interpolating method to compensate missing fields. If there is too much missing in the data or the field (this research define it as 30%), then the data or the field with missing values will be deleted.

Processing error values: As database cannot avoid error values when transferring documents or inputting data, error values should be judged while taking definition of the domain value into consideration. Regularity and other related characteristics are of error values are expected to find out to return to true values. If error values cannot be returned to true values, then they shall be considered as missing values.

Switch of field and format: as the investors, purpose, time and other factors of two database vary, the domain name or the definition of domain value vary as well. As a result, the switch of domain value is needed to reach the same definition of two database.

Switch of variables: as this research targets at categorical variables, continuous variables in the assistant database should be switched to categorical variables to meet the basis assumption.

2.4 CREATE MODULE RELATION

To avoid the complicated computing of large number of data by way of Mann - Whitney U test and the deviation resulted from sampling, draw from the target database m set of samples so that the number of samples is close to that of the assistant database. Then select the pre-determined indices (refer to $S_I'(i)$ and $S_I(i)$ respectively) when pre-processing the target database and the assistant database by way of comparing fields, and apply Mann-Whitney U test of virtual-index to samples and data of the assistant database to see if their data distribution are the same and if they are qualified to be taken as real virtual-index.

To avoid deviation of sampling, we should consider every set of samples [29]. Thus, this research designs a variable $Temp_m$, representing the different distribution between $S_I'(i)$ and $S_I'(i)$ in set m of samples. When $\text{P-value}>\alpha, Temp_m = 1$ representing zero difference in distribution between $S_I'(i)$ and $S_I(i)$; otherwise, $Temp_m = 0$ representing different distribution between $S_I'(i)$ and $S_I(i)$. Finally, calculate $Temp_t = \sum_{m=1}^{k} Temp_m$, and k represents the number of set of samples. Set θ, if $Temp_t \geq \theta, 0 \leq \theta \leq k$, and θ represents the minimum customized number of set of samples that accept H_0 , then $S_I'(i)$ and $S_I(i)$ can be taken as virtual-index. The purpose of setting θ lies in the effort to make sure that when the number of samples is smaller than the matrix, the error judgment results from the failure of representing matrix by samples can be lowered by imp roving the value of θ.

2.5 MAPPING MODULE AND THRESHOLD STATISTICS AND TESTING MECHANISM

Then, apply virtual-index of the module relation to mapping module. A mapping module includes selecting target filed, operating classification model, and analysing accuracy and executing mapping.

In the course of processing, select mapping field from the assistant database and input it into the target database, then take the selected mapping field as the output layer and the tested virtual-index VI of the assistant database as the input layer, establish classification model by way of linear network, probabilistic neural network, radial basis function network and back-propagation neural network. Then select the classification model with the highest accuracy as the mapping model and test its accuracy [30, 31].

If it has enough accuracy, then replace VI in the mapping model by virtual-index VI' of the target database, and take VI' as the output layer of the mapping model, the result of which shall be taken as the value of the extension field $MI'(i)$ for the purpose of extension. After abovementioned operation, every time when going through the mapping module, one mapping field can be mapped and a new extension field in the target database can be produced so that the purpose of database extension is reached.

The database will be based on China Industry, Commerce and Service Census Database (TTICSCDB) as the target database with the data capacity of 186,550 and China Technology Innovation Survey Database as the assistant database with the data capacity of 3,356. As for efficacy measurement index, we will take common field in two database as mapping fields and compare the accuracy of true value with that of mapping value as a measurement of efficacy. As TTICSCDB and TTTISDB only share common in “number of employers”, “occupation”, “whether to conduct research and development” and “turnover in 1990”, thus this research will take two of the four fields as index and the other two to test simulation. Finally, this research will take different fields as mapping fields to conduct the experiment twice to test the accuracy, and analyse simulation of extension fields produced therefrom to see advantages and disadvantages of four classification models.

The target database is based on TTICSCDB in 1990. TTICSCDB has a data capacity of 186,550 and 99 field variables. The assistant database is based on TTTISDB in 1990. TTTISDB has a data capacity of 3,365 and 189 field variables. By pre-processing, there are four common fields, namely, “number of employers”, “occupation”, “whether to conduct research and development” and "turnover in 1990". The background of two databases is shown in Table 1:

TABLE 1 Experimental data

	The target database	**The assistant database**
Name	TTICSCDB	TTTISDB
Data capacity	186,550	3,356
Number of field	99	189
Common fields	Number of employers Occupation Turnover in 1990 Whether to conduct research and development	

3 Automation pre-processing mechanism

To begin with, deal with missing values and error values. Interpolate fields with a missing value of less than 30% by Hot-deck imputation [15-18] and correct error fields. If error fields cannot be corrected, then deal with them as missing values. After this operation, the number of fields is reduced to 102 from 189, and 87 of them are deleted because of over-missing.

Next, deal with common fields. As common fields can be taken as virtual-index in this research for measuring the simulation of the fields, process each of the four common fields, erase deviation and increase the accuracy of data so that the data type and domain value of the two database are the same. The processing method is shown in Table 2:

TABLE 2 Processing common fields

	Type of domain value		**Processing method**
	TTICSCDB	*TTTISDB*	
Number of employers	Record the actual number of employers.	Divide the number of employers into five classes, and use 1-5 to represent.	Reclassify the number of employers in TTICSCDB on the basis of five classes of TTTISDB and use 1-5 to represent.
Occupation	Divide all enterprises in China into 540 classes, and record them respectively.	Divide all enterprises in China into 18 classes, and record them by specific numbers.	Reclassify 540 enterprises in TTICSCDB into 18 classes defined in TTTISDB.
Turnover in 1990	Record an actual turnover in 1990 with the unit of thousand Yuan.	Record an actual turnover in 1990 with the unit of thousand Yuan.	According to the standard of turn-over in Standards for small and medium enterprises by Ministry of Economics, enterprises with a turnover below 100 million Yuan are small and medium enterprises, and will be recorded a "1", enterprises with a turnover of above 100 million Yuan will be recorded a "2".
Whether to conduct research and development	Record an actual expenses for R&D in 1990 with the unit of thousand Yuan.	If it has, record a "1", otherwise, record a "0".	If expenses for R&D in TICSCDB is above 0, it means R&D is conducted in this enterprise, and record "1", otherwise, record "0".

4 Setting up the modular

Only four fields are commonly shared by TICSCDB and TTISDB. To evaluate simulations of the expanding fields mapped by the schema of this research, numerical value and the number of types of the four fields are measured. Two fields, "number of staff" and "occupation" which have most types respectively, are prepositioned as index fields. The other two fields, "turnovers in 1990" and "whether to develop" are positioned as searching fields to measure the closeness of mapping value and true value and whether the simulation is effective.

Researching the TTISDB used the measurement of stratified sampling based on number of staff of the company. Thus in order to avoid the impact of this measurement, the research process is designed based on the proportion of the staff sampled in TTISDB, with 5 times stratified sampling of staffs in TICSCDB. Five sets of samples are collected with about 3,300 samples in each set. This research assumes that the prepositioned fields ("number of staff" and "occupation") are index fields, thus, these two fields of the two databases will be tested through Mann-Whitney U to prove whether they can be deemed as objective index fields. However, every set of samples of TICSDB contains 3,300 samples, which only accounts for 2% of the total number. θ is set as 4 to resolve the problem that the deficiency of sample number may result in the failure of representing the whole and the misinterpretation of the index. The outcome of the test is that the $Temp_t$ numbers of the two fields are 5, which demonstrates that these two fields can be regarded as indexes.

4.1 RESEARCH STEPS

In the following research process, "turnovers in 1990" and "whether to develop" are positioned as searching fields, while "number of staff" and "occupation" are prepositioned as index fields. The research focuses on the classification model accuracy, sample accuracy and TICSDB accuracy. Measuring the accuracy will be used to determine whether the simulation of the schema is of high rate and whether there exist the absolute advantages and disadvantages of the accuracies of these four fields. In the Experiment 1, "whether to develop" in the assistant database is applied as the mapping fields, while "number of staff" and "occupation" is used as indexes. The input layer of classification model is "number of staff" and "occupation", while the output layer is "whether to develop". Four kinds of classification models, linear networks (LN), radial basis function networks (RBF) [20, 21], probabilistic neural networks (PNN) [22-24], back propagation neural networks (BNN) [25,26] are measured to test the model accuracy $P_C(X)$, sample accuracy $P_S(X)$ and mapping accuracy $P_M(X)$. The data of accuracy P(X) measured will be compared to the mapping data with the equations below:

$$P_i(X)=\begin{cases}1, & if \quad X_i^{true}=X_i^{mapping} \\ 0, & if \quad X_i^{true}\neq X_i^{mapping}\end{cases}, \tag{1}$$

$$P(X)=\sum_{i=1}^{n} P_i(X)/n, \tag{2}$$

where X_i^{true} represents the true value of data *i* in the field X. $X_i^{mapping}$ represents the mapping value of data *i* in the field X. $P_i(X)$ represents the consistency of true value and mapping value of data *i* in the field X. If the outcome is 1, they are consistent. 0, not; P(X) represents the ration of true value and mapping value of data *i* in the field X.

Thus, the accuracy of the classification models $P_C(X)$ represents the accuracy of the model, which are tested. The accuracy of the true value and estimating value of tested data. The sample accuracy $P_S(X)$ represents the consistency of true value and map-ping value after the model is mapped using classification models, which have been tested. Mapping accuracy $P_M(X)$ is the consistency of true value and mapping value after the mapping of target database TICSCDB based on the tested classification models as map-ping models. It is also the measurement to tell whether simulation is effective.

In Experiment 2, "turnover in 1990" in the assistant database is applied as the mapping fields, while "number of staff" and "occupation" is used as indexes. The input layer of classification model is "number of staff" and "occupation", while the output layer is "turnover in 1990". Four kinds of classification models, linear networks (LN), radial basis function networks (RBF) [20, 21], probabilistic neural networks (PNN) [22-24], back propagation neural networks (BNN) [25,26] are measured to test the model accuracy $P_C(X)$, sample accuracy $P_S(X)$ and mapping accuracy $P_M(X)$. The definitions of, and are the same as they are defined in experiment 1.

4.2 EXPERIMENT 1: SET "TO DEVELOP OR NOT" AS THE FIELD FOR MAPPING

Set "To develop or not" as MI, and "The number of the staff" and "category of employment" as VI, and randomly draw 50% bases from the database as training data, the other 50% as test data, and respectively apply four methods, namely PNN, LN, BNN, RBF for classification model training.

According to the diagrams showed above, each model has its own numbers of input layers, hidden layers and output layers. The data are tabulated for clearer compare into Table 3 below:

TABLE 3 Number of Neurons of the Classified Model for Experiment 1

	Number of Input Layer	Number of Hidden layer 1	Number of Hidden Layer 2	Number of Hidden Layer 3
PNN	23	1678	2	2
BNN	23	6	-	1
LN	23	-	-	1
RBF	23	27	-	1

Then map the four models according to the sample and TICSCDB and calculate the accuracy $P_S(X)$ of the sample, mapping accuracy $P_M(X)$ and overall accuracy by comparing the databases with the true values in the original database. The consequence is showed in the Figures 1 and Figure 2.

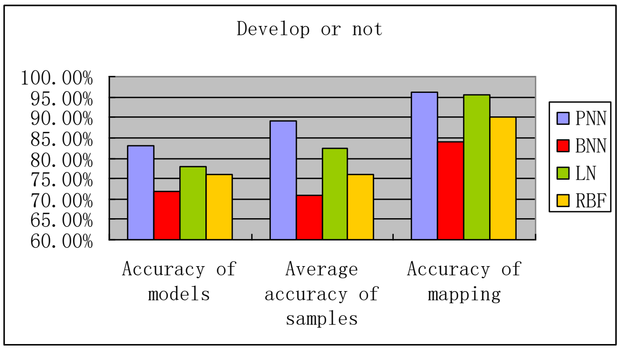

FIGURE 1 Bar chart for the consequence of Experiment 1

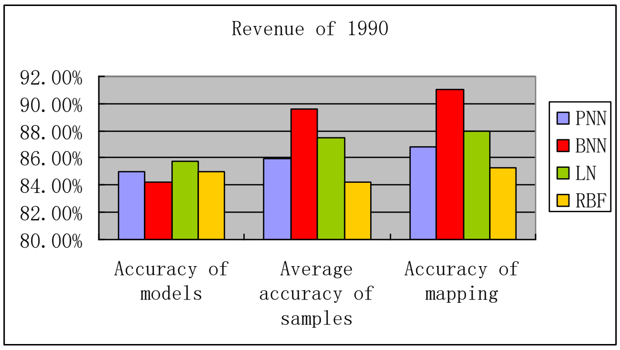

FIGURE 2 Bar chart for the result of Experiment 2

Conclusions of Experiment 1:

1) The ranking consequences of the accuracy $P_C(X)$ and mapping accuracy $P_M(X)$ of a classification model of the four classification models are the same, thus it is feasible to use the accuracy $P_C(X)$ of a classification model to assess whether the classification model can be used for judging which classification model is a mapping model.

2) According to the consequence, all the mapping accuracy are over 84.6%, which indicates that the similarity between the expanded field and the true values

is very high and also proves that the expanded fields produced under the structure of this system possesses a very high simulation.

3) In the four classification models, PNN fits the Experiment 1 for the most, while BNN does the least.

4.3 EXPERIMENT 2: SET "THE REVENUE OF 1990" AS THE FIELD FOR MAPPING

Set "the Revenue of 1990" as MI, and "The number of the staff" and "category of employment" as VI, and respectively apply four methods, namely PNN, LN, BNN and RBF for classification model training. When conducting the classification model training, in order to avoid the influence on the consequence of the accuracy $P_C(X)$ of the classification model caused by dealing with lost data, as well as to allow the material to wait for the computation so as to avoid inaccuracy caused by interpolations, the missing data of the field "the revenue of 1990" are ignored. And finally use the 2984 data of TTISDB for the training of classification models, and randomly draws 50% bases from the auxiliary database as training data, the other 50% of the auxiliary database as test data for the training of classification methods, and map the four models according to the sample and TICSCDB (this process is similar to that of Experiment 1. The accuracy is showed in Figure 3.

Conclusions of Experiment 2:

1) In Experiment 2, the difference of accuracy $P_C(X)$ between each two classification models is not as big as that of the Experiment 1. The difference between the lowest accuracy and the highest accuracy is only 0.4%. Thus, the best classification model for mapping model cannot be determined through Experiment.

2) The rank of the accuracy $P_S(X)$ and mapping accuracy $P_M(X)$ of the samples of the four classification models are the same. Thus it proves that the distribution of the sample is similar to that of parent population, and what caused the different ranking of mapping accuracy and the classification model accuracy of the different classification methods is the error between the target database and auxiliary database.

3) The difference in the rank of accuracy $P_C(X)$ and mapping accuracy $P_M(X)$ of classification model of the four classification models is a problem deserves further study and discuss.

5 Discussion

5.1 DISCUSSION ON ACCURACY

According to the two experiments above, the mapping accuracy for the four classification methods are showed in Figure 3.

No matter apply, which classification model as the mapping model, the accuracy is all very high (more than 84.6%), which proves that the system put forward in this study is a good structure for database expansion, and the expanded fields estimated by the system structure, which is put forward in this study, possess a very high simulation.

If order the four models in the two experiments by the mapping accuracy from the highest to the lowest.

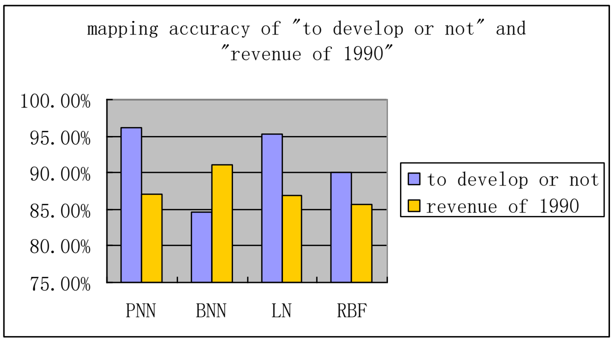

FIGURE 3 Bar chart of mapping accuracy

For Experiment 1, the classification model with the highest accuracy is PNN, up to 96.1%. But in Experiment 2, the accuracy of PNN is only 86.8%, ranked the last but one. While in Experiment 2, the classification model with the highest accuracy is BNN, up to 90.9%. But in the Experiment 1, the accuracy of BNN is only 84.6%, ranked last in accuracy of the four classification models.

Thus, there is no absolute inferior or superior in the four methods and there are different databases fit for each. So the process for the highest mapping accuracy cannot be simplified into a single model. And building several classification models and choosing the fittest model for mapping is a good method.

It's also found in the study that when the accuracy of classification model of different classification models are similar, the fittest classification model for mapping, because the error between two databases might be bigger than the difference between classification models. Thus, how to choose the mapping model effectively is also a problem deserves to study and discuss.

5.2 LIMITATIONS OF THE STUDY

The model does not fit for non-category data: as mentioned repeatedly in this study, the system structure and the methods applied in this study are data aiming at category field and does not fit non-category data.

The difference of indexes between two databases cannot be valued: the index between two databases is verified by Mann-Whitney U. Although the testing power can be controlled by adjusting level of significance α, the influence caused by the difference of P-value cannot be measured after being judged as an index. Thus, it is not sure to ignore the difference between index or not.

6 Conclusion

The four conclusions of this study are as followed:

1. To use the available resources for mutual expansion: by expansion of available databases, the data of an

enterprise can be effectively used and to avoid unsatisfactory subsequent analysis caused by missing of some field.

2. To reduce the cost of reinvestigation of database: by the method put forward in the study, the needed field can be expanded from another database of the enterprise, which can spare the time and cost of reinvestigation.

3. Databases with the same parent population can interflow so to make the subsequent analysis more efficient. By the technique in this study, the databases with the same parent population can interflow. So it will be easier to apply several databases in one analysis, and the inadequacy of relative fields needed frequently happened before in single database analysis can be avoided.

4. It enables the enterprise to know the situation of rivalries by their own databases: this study is feasible for all the databases with the same parent population. The custom group of competitive enterprises is similar. If an enterprise needs to know the information of its rivalry, it can draw some of the characteristics of customs of the rivalry. Then map the field of its own database into the database gained by investigation. Thus the overall characteristic of custom group and the direction of management and other.

Information of the rivalry can be gained effectively, and the plan to cope with it will be easy to make.

In the future, it will improve according to the following directions:

1. To expand its applied range: the method put forward in this study only aims at expansion of field of category data and cannot be applied to assist the field of continues data, text data, time series data, etc. Thus, in order to make is applicable for expansion of fields of different kinds of database, studies on other methods of fields expansion is an important direction in the future.

2. To increase the accuracy of classification model: how to increase the accuracy of classification model is also an important direction of study in the future. It can be improved by modifying the existing theory or citing the improved theories of other researchers, or even creating new methods for classification only if it can increase the overall accuracy of the system.

3. To increase the complexity of the index between two databases: if significance level can be tripped and do not need to factitiously examine whether the significances of the two fields are the same beforehand. Then one-to-many indexes can be built according to the distribution of values among several fields, thus to build between two databases a one-to- many index, for example: $f(\mathrm{x}_3) = f(\mathrm{y}_1) + f(\mathrm{y}_7) + c$, x_i indicates the field of the auxiliary database, y_i indicates the fields of the target database, c is a constant.

4. Add reference value for accuracy: it might be feasible to indicate the accuracy of classification model of the mapping model applied in each mapping field. Although there can be errors among indexed and the conformity between two databases, it offers the user an objective accuracy for subsequent analysis so as to judge whether this system deserves to apply.

References

[1] Ramakrishnan R, Gehrke J 2000 Database management systems *Osborne/McGraw-Hill*

[2] Gollub J, Ball C A, Binkley G, Demeter J, Funkelstein D B, Herbert J M, Hernandez-Boussard T, Jin H, Kaloper M, Matese J C, Schroeder M, Brown P O, Botstein D, Sherclock G 2003 The Stanford Microarray Database: data access and quality assessment tools *Nucleic Acids Research* **31**(1) 94-6

[3] Miles M B, Huberman A M 1994 Data management and analysis methods *Sage Publications*

[4] Agrawal R, Imielinski T, Swami A 1993 *IEEE Transactions on Knowledge and Data Engineering* **5**(6) 914-25

[5] O'Leary D E 1993 The impact of data accuracy on system learning. *Journal of Management Information Systems* **9**(4) 83-98

[6] Palmer D C 1998 Method for determining database accuracy *US Patent* No 5 787 443

[7] Dierickx I, Cool K 1989 Asset stock accumulation and sustainability of competitive advantage *Management science* **35**(12) 1504-11

[8] Özsu M T, Valduriez P 2011 Principles of distributed database systems *Springer*

[9] Bernstein P A, Goodman N 1981 Concurrency control in distributed database systems *ACM Computing Surveys (CSUR)* **13**(2) 185-221

[10] Hector G C 1983 Using semantic knowledge for transaction processing in a distributed database *ACM Transactions on Database Systems (TODS)* **8**(2) 186-213

[11] Bell D, Grimson J 1992 Distributed database systems *Addison-Wesley Longman Publishing Co Inc.* Boston MA USA

[12] Ruxton G D 2006 The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test *Behavioral Ecology* **17**(4) 688-90

[13] Hart H L A 1963 Law, liberty and morality Stanford CA USA *Stanford University Press*

[14] Rosner B, Grove B 1999 Use of the Mann-Whitney U-test for clustered data *Statistics in Medicine* **18**(11) 1387-400

[15] Rao J N K, Shao J 1992 Jackknife variance estimation with survey data under hot deck imputation. Biometrika **79**(4) 811-22

[16] Sande I G 1983 Hot-deck imputation procedures *Incomplete data in sample surveys* **3** 334-50

[17] Andridge R R, Little R J A 2010 A Review of Hot Deck Imputa-tion for Survey Non-response *International Statistical Review* **78**(1) 40-64

[18] Kim J K, Fuller W 2004 Fractional hot deck imputation *Biometrika* **91**(3) 559-78

[19] Chen W-K 1993 Linear Networks and Systems: Algorithms and Computer-Aided Implementations *World Scientific*

[20] Chen S, Cowan C F N, Grant P M 1991 *IEEE Transactions on Neural Networks* **2**(2) 302-9

[21] Park J, Sandberg I W 1991 Universal approximation using radial-basis-function networks *Neural computation* **3**(2) 246-57

[22] Specht D F 1990 Probabilistic neural networks *Neural networks* **3**(1) 109-18

[23] Specht D F 1990 *IEEE Transactions on Neural Networks* **1**(1) 111-21

[24] Specht D F 1992 Enhancements to probabilistic neural networks *International Joint Conference on Neural Networks* **1** 761-8

[25] Hecht-Nielsen R 1989 Theory of the back propagation neural network *International Joint Conference on Neural Networks* **1** 593-605

[26] Wythoff B J 1993 Backpropagation neural networks: a tutorial *Chemometrics and Intelligent Laboratory Systems* **18**(2) 115-55

[27] Frawley W J, Piatetsky-Shapiro G, Matheus C J 1992 Knowledge discovery in databases: An overview *AI magazine* **13**(3) 57-70

[28] Fayyad U M, Piatetsky-Shapiro G, Smyth P 1996 Advances in knowledge discovery and data mining *MIT Press* Cambridge MA 1-34
[29] Rubin D B, N Schenker N 1986 Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. *Journal of the American Statistical Association* **81**(394) 366-74
[30] Streit R L, Luginbuhl T E 1994 *IEEE Transactions on Neural Networks* **5**(5) 764-83
[31] Topchy A P, Lebedko O A, Miagkikh V V, Kasabov N K 1997 Adaptive training of radial basis function networks based on cooperative evolution and evolutionary programming *International Conference on Neural Information Processing and Intelligent Information Systems* 253-8

Authors

Liangong Song, born in May, 1975, Xinjiang, China

Current position, grades: MSc, lecturer at North China University of Water Resources and Electric Power in China.
University studies: Bachelor of engineering at China Agricultural University in 1999, MSc in Information Management Study at Beijing Forestry University in 2005
Scientific interests: theory and application of computer, principle and application of database, information management
Publications: 8.

Ke Han, born in November, 11 1980, Henan, China

Current position, grades: MSc, lecturer at North China University of Water Resources and Electric Power in China.
University studies: Bachelor of engineering at Henan Normal l University in 2002, MSc in Information Management Study at Chinese Academy of Sciences in 2005.
Scientific interests: information management system, theory and application of computer
Publications: 11.