Acta Technica 62 (2017), No. 7A, 1599–1610

Statistics on the new-generation agricultural labour force transfer trend in China based on the maximum entropy

Zhaoxi Wu[1]

Abstract. In order to enhance the effectiveness of the prediction on the transfer trend of the new-generation agricultural labour force in China, a statistical method for predicting the transfer trend of the new-generation agricultural labour force in China based on the maximum entropy has been raised. Firstly, conduct comprehensive analysis from the perspectives of labour force, cultivated land and agricultural mechanization, fully consider the important impetus of the urban-rural income gap for labour force transfer and establish the prediction model of the rural labour force transfer; secondly, learn the rural labour force transfer trend model and make predictions by aid of the statistical method based on the maximum entropy, meanwhile, improve the traditional sample estimation model aiming at its large deviation based on the weight probability moment algorithm; finally, verify the effectiveness of the new-generation rural labour force transfer trend prediction algorithm in China through the analysis on the numerical model and empirical example.

Key words. Maximum entropy statistics, Labour force, Transfer trend, Prediction, New-generation.

1. Introduction

With the rapid reform of the household registration system in China, the large improvement in the level of productive forces and the constant adjustment in the industrial structure, more and more farmers have escaped from the agricultural production and become the surplus rural labour force. Predicting the number of the surplus rural labour force correctly and transferring the surplus rural labour force effectively are of great significance for solving the issues of issues of agriculture, farmer and rural area, optimizing the resource allocation, keeping the sustainable and

[1]Henan Academy of Agricultural Sciences, Zhengzhou 450002, China

healthy development of rural economy, maintaining the social stability and enhancing the living standard of people.

Firstly, most researches on the rural labour force transfer concentrate on the empirical analysis on large number of survey data and policy suggestion currently but lack of suitable economic theory model for making logic and rigorous theoretical explanations on the transfer rule of the rural labour force. Secondly, the current algorithms used for predicting the transfer number of the rural labour force include: point estimation, internal estimation, variance analysis, regression analysis, Markov chain method, GM (1,1) and evolutionary game theory. The above-mentioned algorithms all take the previously-transferred historical data related to the labour force as the base period data of the model for simulation and make no consideration of factors affecting the labour force transfer in the process of establishing the prediction model, in addition, the algorithms are the methods based on pure mathematics, which can not satisfy the demands of the industrial structure adjustment and sustainable development of agriculture. Finally, the prediction on the rural labour force transfer is the regression problem for obtaining the prediction model based on analyzing the sample data, which occupies a small number of only dozens to hundreds. Therefore, it is needed to fully consider the various influencing factors of the rural labour force transfer in the process of establishing suitable rural labour force transfer prediction model for quantification to be the input of the model and make the number of labour force transferred practically in previous years be the output of the model.

The paper conducts comprehensive analysis on the factors such as labour force, cultivated land and agricultural mechanization in order to enhance the prediction precision of the new-generation rural labour force transfer trend in China, fully considers important impetus of the urban-rural income gap for labour force transfer and establishes the prediction model of the rural labour force transfer; besides, it learns the rural labour force transfer trend model and make predictions by aid of the statistical method based on the maximum entropy, meanwhile, improves the traditional sample estimation model aiming at its large deviation based on the weight probability moment algorithm.

2. Rural labour force transfer prediction model

The rural labour force system is a subsystem in the society,which is composed of three basic factors such as person, land and agricultural machinery. The factors influencing the number of the rural labour force occupy a large number, including the quantitative factors and qualitative factors. The author conducts comprehensive analysis from the perspectives of labour force, cultivated land and agricultural mechanization: when the number of the population and the cultivated land as well as the agricultural mechanization level change, the number of the rural labour force will be affected, and the urban-rural income gap is the impetus for the rural labour force flowing into the cities. However, with the increasing in the seeded area of the commercial crop, more labour forces are needed in rural areas,and the number of the rural labour force will decrease accordingly. Therefore, it is needed to select the population, cultivated land, agricultural mechanization level, urban-rural income gap and the seeded area of commercial crop to the the set of influencing factors. Wherein, the agricultural mechanization level is represented by the total power of farm machinery.

The various influencing factors of the rural labour force are not linearly dependent with the various statistical indicators, therefore, the nonlinear regression support vector machine is adopted in the modelling process thus to map the data to the high-dimensional characteristic space based on a nonlinear mapping and establish a linear regression function in the space, namely:

$$f(x) = \omega^T \phi(x) + b\,. \tag{1}$$

In the formula, $\phi(\cdot)$ represents the nonlinear mapping, ω represents the weight and b represents polarization. In accordance with the structure risk minimization principle, the $f(x)$ of best approximation should make the risk function minimum:

$$j = \frac{1}{2}\left\|\omega\right\|^2 + C\sum_{i=1}^{n} L\left(f\left(x_i\right), y_i\right). \tag{2}$$

In the formula, C represents the penalty factor, and the insensitive dissipation function $L(\cdot)$ can be presented as:

$$L\left(f\left(x_i\right), y_i\right) = \left|y_i - \{\omega, \phi\left(x_i\right) - b\} - \varepsilon\right|\,. \tag{3}$$

In the formula, ε represents the minimum allowable error. Introduce the slack variables ξ_i and ξ_i^*, the formula (2) will become:

$$j = \frac{1}{2}\left\|\omega\right\|^2 + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right). \tag{4}$$

The constraint conditions are:

$$\begin{cases} \omega^T\phi\left(x_i\right) + b_i - y_i \leq \varepsilon + \xi_i^*, \xi_i^* \geq 0\,, \\ \quad y_i - \omega^T\phi\left(x_i\right) - b_i \leq \varepsilon + \xi_i; \xi_i \geq 0\,. \end{cases} \tag{5}$$

In accordance with the duality principle, lagrangian multiplier method and kernel function technology, the minimum risk function in the formula (4) can be transformed into the following quadratic programming problem:

$$\min j = \sum_{i=1}^{n} y_i\left(\alpha_i^* - \alpha_i\right) - \varepsilon\sum_{i=1}^{n}\alpha_i^* - \frac{1}{2}\sum_{i,j=1}^{n}\left(\alpha_i^* - \alpha_i\right)\left(\alpha_j^* - \alpha_j\right)K\left(x_i, x_j\right)\,. \tag{6}$$

The constraint conditions are:

$$\sum_{i=1}^{n}\left(\alpha_i^* - \alpha\right) = 0; \alpha_i^*, \alpha_i \in [0, C]\,, i = 1, K, n\,. \tag{7}$$

3. The maximum entropy weight probability moment

3.1. The maximum entropy method

The concept of entropy is firstly raised from the thermodynamics to be used for measuring the system instability, which provides a wide application prospect for the engineering and theory. If the parameter x is defined as the random variable in the space R, and its probability distribution $P(X = X_k) = p_k, k = 1, 2, 3, \cdots, n$, the entropy calculation of the parameter x will be [11∼12] :

$$H = -\sum_{k\geq 1} p_k \ln p_k \,. \tag{8}$$

Similar to the concept of the above-mentioned discrete variable, if the variable x is defined as the random variable in the space R, it will be in the continuous form and its probability density will be $f(x)$, and the entropy calculation of the variable x will be as follows:

$$H = \int_R f(x) \ln f(x) dx \,. \tag{9}$$

In the formula (9), if $f(x) = 0$, $f(x) \ln f(x) = 0$.

In accordance with the definition for the maximum entropy, the minimum deviation distribution model of the variable x in the space R can be defined as:

$$\begin{cases} \max H = -\displaystyle\int_R f(x) \ln f(x) dx \\ s.t. \displaystyle\int_R x^n f(x) dx = \mu_n, x = 0, 1, 2, \cdots, N \end{cases} \tag{10}$$

In the formula (10), N represents the order of the largest sample probability moment; μ_n represents the sample probability moment; m represents the relevant data sample number in the training process; x_i represents the variable i. The sample probability moment can be defined as:

$$\mu_n = \frac{1}{m} \sum_{i=1}^{m} x_i^n \,. \tag{11}$$

$$c_n = \frac{1}{m} \sum_{i=1}^{m} (x_i - \mu_1)^n \,. \tag{12}$$

In the formula (11∼12), the parameter c_n represents the central moment of n-order, which can realize the mutual conversion in its original origin moment μ_n with the form of binomial conversion. In accordance with the lagrange method[13∼14], solve the probability density based on the maximum entropy principle by aid of the traditional variational method:

$$f(x, \lambda_0, \lambda_1, \cdots, \lambda_n) = \exp(\lambda_0 + \sum_{n=1}^{N} \lambda_n x^n) . \tag{13}$$

In the formula (13), the parameters including $\lambda_i, i = 0, 1, \cdots, N$ represent the i-order moment constrained in lagrange method. In accordance with the formula (13), the probability density $f(x)$ can be calculated after determining $\lambda_i, i = 0, 1, \cdots, N$.

3.2. Weight probability moment

If the variable is x defined in the space R, the accumulation function related to it has the distribution form $F(x) = P(X \leq x)$ and the inverse function $x = x(F)$ of the subfunction can be obtained, the weight probability moment of the above functions can be calculated as follows[15]:

$$M_{i,n,k} = \int_0^1 [x(F)]^i F^n (1 - F)^k dF . \tag{14}$$

In the formula (14), the subscripts i, n and k are real numbers. It is known that the weight probability moment is the expanded evolution based on the traditional classic moment which can be taken as a special situation of the weight probability moment, wherein, two kinds of probability moments are usually adopted:

$$\text{Type 1: } \alpha_k = M_{1,0,k} = \int_0^1 x(F)(1 - F)^k dF : \tag{15}$$

$$\text{Type 2: } \beta_n = M_{1,n,0} = \int_0^1 [x(F)] F^n dF . \tag{16}$$

Rank the samples and get the orders $x_1 \leq x_2 \leq \cdots \leq x_{m-1} \leq x_m$, and then the forms of a_k and b_n will be shown as follows:

$$a_k = \frac{1}{m} \sum_{i=1}^{m} \left\{ \begin{bmatrix} m-1 \\ k \end{bmatrix} x_i \middle/ \begin{bmatrix} m-1 \\ k \end{bmatrix} \right\} . \tag{17}$$

$$b_n = \frac{1}{m} \sum_{i=1}^{m} \left\{ \begin{bmatrix} i-1 \\ n \end{bmatrix} x_i \middle/ \begin{bmatrix} m-1 \\ n \end{bmatrix} \right\} . \tag{18}$$

In the formula (17∼18), $k = 0, 1, \cdots, m-1$, $n = 0, 1, \cdots, m-1$, $\begin{bmatrix} i \\ n \end{bmatrix} = \frac{i!}{(i-n)!n!}$.

Because the parameter estimation and calculation strategy based on the weight probability moment presents little sensibility in the small sample application, it shows higher robustness compared with the classic moment, which realizes the zero-error evaluation on the sample rule.

3.3. Weight probability moment and inverse cumulative distribution function

If the variable x in the research is non-negative and random, the parameter β_n can be taken as the inverse cumulative moment of distribution model, and the following relational expression can be obtained:

$$E[X^n] = \int_R x^n f(x)dx = \int_0^1 [x(u)]^n du\,. \tag{19}$$

In the formula (19), $du = dF(x) = \frac{f(x)}{\int f(x)dx}$. In accordance with the parameter model β_n, a certain similarity correlation can be established between $x(u)$ and u and $f(x)$ and x combined with the formula (19), and the following conversion and mapping model will exist:

$$dT(u) = \frac{x(u)du}{\int_0^1 x(u)du} = \frac{x(u)du}{\beta_0}\,. \tag{20}$$

Based on the parameter dT, the weight probability moment β_n can be changed and deduced in its form:

$$\beta_n = \beta_0 \int_0^1 u^n dT(u), n = 1, 2, \cdots, N\,. \tag{21}$$

In the formula (21), β_0 represents the sample average. β_n/β_0 represents the n-order moment of $x\,(u)$ in the inverse accumulative process.Similarly, the inverse accumulative n-order moment form of the function $x(1-u)$ can be defined as α_k/α_0. When $\beta_0 = 1$, β_n represents the n-order inverse accumulative distribution of $x(u)$. In summary, the inverse accumulative maximum entropy function distribution form can be shown as follows:

$$\begin{cases} maxH = -\int_0^1 x(u)\ln x(u)du \\ s.t. \int_0^1 u^n x(u)du = b_n, n = 0, 1, 2, \cdots, N \end{cases} \tag{22}$$

For the model (22), the solution is:

$$x(u, \lambda_0, \lambda_1, \cdots, \lambda_n) = \exp(\lambda_0 + \sum_{n=1}^{N} \lambda_n u^n)\,. \tag{23}$$

It can be known from comparing the formula (10) and the formula (23) that if the original definitional domain interval is changed and the constraint in the model (23) is same to the classic moment in the form, the formula (18) will present the probability density, the constraint will represent the weight probability moment and the formula (18) will present the inverse accumulative form of the distribution function.

4. Parameter estimation

4.1. Computational formula

Firstly, determine the above-mentioned multiplier parameters $\lambda_i, i = 0, 1, \cdots, N$, and the result can be obtained as follows in accordance with the equivalence relation between the sample average and the zero moment presented in the inverse accumulative function distribution:

$$\int_0^1 \exp(\lambda_0 + \sum_{n=1}^{N} \lambda_n u^n) du = b_0 \,. \tag{24}$$

In the formula (24), multiply the factor $e^{-\lambda_0}$ at the same time in the two sides:

$$\int_0^1 \exp(\sum_{n=1}^{N} \lambda_n u^n) du = e^{-\lambda_0} b_0 \,. \tag{25}$$

Conduct derivative operation aiming at λ_n:

$$e^{-\lambda_0} b_0 \frac{\partial \lambda_0}{\partial \lambda_n} = \int_0^1 u^n \exp\left(\sum_{n=1}^{N} \lambda_n u^n\right) du \,. \tag{26}$$

And the result shows that:

$$\frac{\partial \lambda_0}{\partial \lambda_n} = -\frac{b_n}{b_0} \,. \tag{27}$$

$$\frac{\partial \lambda_0}{\partial \lambda_n} = \frac{\int_0^1 u^n \exp\left(\sum\limits_{n=1}^{N} \lambda_n u^n\right) du}{\int_0^1 \exp\left(\sum\limits_{n=1}^{N} \lambda_n u^n\right)} \,. \tag{28}$$

The result can be obtained in accordance with the formula (27∼28):

$$b_n = b_0 \frac{\int_0^1 u^n \exp(\sum\limits_{n=1}^{N} \lambda_n u^n) du}{\int_0^1 \exp(\sum\limits_{n=1}^{N} \lambda_n u^n) du} \,. \tag{29}$$

The formula (29) represents the N-group equation established based on the parameters $\lambda_i, i = 0, 1, \cdots, N$.In order to facilitate the numerical calculation, modify the formula (24) as follows:

$$Q_n = 1 - b_0 \frac{\int_0^1 u^n \exp(\sum_{n=1}^{N} \lambda_n u^n) du}{b_n \int_0^1 \exp(\sum_{n=1}^{N} \lambda_n u^n) du} . \tag{30}$$

Solve the optimization objective:

$$\min Q = \sum_{n=1}^{N} Q_n^2 . \tag{31}$$

Based on the Jacoby calculation strategy, formulate the computer software code to conduct iterative solution for the formula (31) and know $\lambda_i, i = 0, 1, \cdots, N$.

4.2. Algorithm steps

Step 1: Rank the samples related to educational training;
Step 2: Rank the different-order samples of the weight probability moment in accordance with the historical educational sample data, and set the 6-order moment to be the upper limit in order to guarantee the consistency;
Step 3: Establish the objective (31) to be the model residual and set its initial value;
Step 4: Call related program to optimize the algorithm;
Step 5: Identify the targeted convergence, transfer to step 7 if the convergence can be realized and transfer to step 6 on the contrary;
Step 6: Conduct recalculation based on other initial values and then transfer to step 4;
Step 7: Calculate the value of $\lambda_i, i = 0, 1, \cdots, N$, and output the final result;
Step 8: End the calculation.

5. Experiment analysis

5.1. Numerical examples

Take the Pareto distribution as the tested numerical value objects, verify the feature difference between the classic moment and the weight probability moment in the paper through experiment, and conduct simulation research based on practical examples.

$$F(x) = u = \begin{cases} 1 - \left[1 - c\frac{x}{d}\right]^{\frac{1}{c}}, c \neq 0 \\ 1 - \exp\left(-\frac{x}{d}\right), c = 0 \end{cases} \tag{32}$$

The weight probability moment of the inverse accumulative function of the model shown in the formula (32) presents two kinds of forms, such as the above-mentioned type 1 and type 2. And the result will show that:

$$x(u) = \begin{cases} \dfrac{d}{c}\left[1-(1-u)^c\right], c \neq 0 \\ -d\log(1-u), c = 0 \end{cases} \tag{33}$$

In accordance with the feature of the statistical data, the LaGrange parameter based on the maximum entropy can be calculated by aid of the different values of order k. See figure 1∼2 for the different features between the theoretical calculation data and the statistical data. Wherein, select the confidence of 95% in the theoretical data calculation, namely χ^2 hypothesis testing in essence.

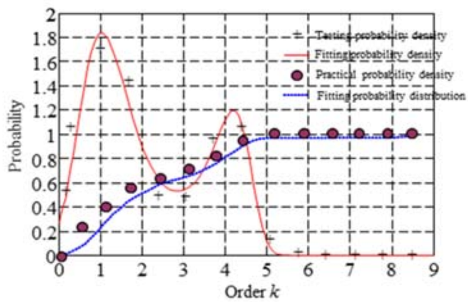

Fig. 1. Probability density function of type 1

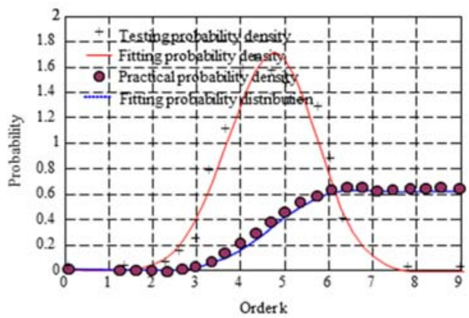

Fig. 2. Probability density function of type 2

5.2. Empirical analysis

Based on the practical situation of the rural labour force transfer in the Ningxia Hui Autonomous Region, conduct simulation testing by aif of the Matlab7.0 software in accordane with the maximum entropy statistics prediction model in the formula (8). Besides, select the data related to the rural labour force transfer from 1990 to 2002 to be the indicator system of the prediction model, wherein, the related data from 1990 to 1997 are taken as the training set for establishing the prediction model and the related data from 1998 to 2002 are taken as the testing set of the prediction model, see form 1 for details.

Form 1. Maximum entropy statistics modeling data

Year	X^1/ten thousand people	X^2/ ten thousand hm^2	X^3/ten thousand W	X^4/Yuan	X^5/ten thousand hm^2	Y
1990	354.29	79.6	191105	676.27	16.5	220.89
1991	359.98	79.8	202687	783.40	17.2	222.96
1992	365.69	80. 1	212126	984.24	16.6	223.88
1993	371.30	80.3	217581	1239.90	17.5	227.09
1994	369. 55	80. 6	228601	1747.20	18.1	228.35
1995	374.43	80.7	241463	1989.48	19.4	230.93
1996	380.43	81.3	255986	1860.82	19.5	230.94
1997	382. 24	127.1	288432	1885.92	19.6	231.10
1998	384.35	127. 5	316183	1889.31	18.6	231.95
1999	387.89	128. 0	377934	2682.20	19.5	236.22
2000	395.09	129.3	380633	3188.10	20.9	236.45
2001	399.29	129. 9	407623	3721.07	22.6	242.19
2002	402.75	129.2	477513	4150.08	26.7	246.95

Conduct normalization processing on the original data, conduct training and testing respectively based on the determined training sets and testing sets, and get all parameters of the maximum entropy statistics regression model. In the testing, select the radial basis function for the maximum entropy statistics kernel function K, besides, adopt the LOO Cross-validation method in the training process in order to obtain the stable model thus to obtain the error rates of various cross validations and adopt the corresponding values of δ and C when the model reaches to the minimum error rate. Meanwhile, the width δ of the radial basis function is 1.256e-3, the penalty factor is 279 and the minimum allowable error ε is 0.0001. See form 2 for the predicted value of the rural labour force transfer prediction model and the forecast error, and the absolute error of the model prediction is 1. 90%∼5.27%. Figure 3 shows the resolute error of the rural labour force transfer prediction model.

Form 2. Comparison among the predicted results of the model

Year	Aetual migration Aetual migration value/ten thousand people	Model forecast value/ Aetual migration value/ten thousand people	Forecast relative error/%
1998	231.95	238.12	2. 66
1999	236.22	240.73	1.09
2000	236.45	231.06	-2.30
2001	242.19	251.08	3. 67
2002	246.95	259.97	5. 27

The average error output from the prediction model is 2.25%, besides, the average relative error is 3.16% and the mean square error is 81900 people. It can be known

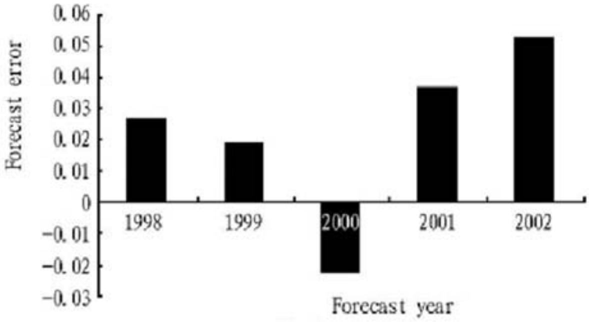

Fig. 3. Predict the relative error

from the predicted result that the prediction accuracy of the rural labour force transfer prediction model based on maximum entropy statistics is higher, which can be used for predicting the rural labour force transfer.

6. Conclusion

Based on the current rural labour force transfer prediction algorithms such as comparison point estimation, interval estimation, variance estimation, regression estimation, Markov chain method, GM (1,1) and evolutionary game theory, the paper puts forward to establish the rural labour force transfer prediction model based on the maximum entropy statistics through analyzing the feature of the maximum entropy statistics theory. In addition, it takes the quantitative indicators influencing the rural labour transfer in the Ningxia Hui Autonomous Region , including population, cultivated area, agricultural mechanization level, urban-rural income gap and seeded area of the commercial crop as the input characteristics of the model, and analyzes the prediction accuracy of the prediction model through simulation testing.

References

[1] W. Pan, S. Z. Chen, Z. Y. Feng: *Automatic Clustering of Social Tag using Community Detection.* Applied Mathematics & Information Sciences *7* (2013), No. 2, 675–681.

[2] Y. J. Zhao, Y. H. Hu, J. J. Liu: *Random Triggering-Based Sub-Nyquist Sampling System for Sparse Multiband Signal.* IEEE Transactions on Instrumentation and Measurement *66* (2017), No. 7, 1789–1797.

[3] X. Du, L. Chen, D. Huang, Z. Peng, C. Zhao, Y. Zhang, Y. Zhu, Z. Wang, X. Li, G. Liu: *Elevated Apoptosis in the Liver of Dairy Cows with Ketosis.* Cellular Physiology & Biochemistry *43* (2017b), No. 2, 568–578.

[4] Y. Y. Zhang, Q. Li, W. J. Welsh, P. V. Moghe, and K. E. Uhrich: (2016) *Micellar and Structural Stability of Nanoscale Amphiphilic Polymers: Implications for Anti-atherosclerotic Bioactivity*, Biomaterials, 84, 230-240.

[5] L. R. Stephygraph, N. Arunkumar, V. Venkatraman: *Wireless mobile robot control through human machine interface using brain signals*, 2015 International Conference on Smart Technologies and Management for Computing, Communication, Con-

trols, Energy and Materials, ICSTM 2015 - Proceedings, art. No. 7225484 (2015), 596–603.

[6] Z. Lv, A. Halawani, S. Feng, H. Li, S. U. Réhman: *Multimodal hand and foot gesture interaction for handheld devices.* ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) *11* (2014), No. 1s, 10.

[7] Y. Z. Chen, F. J. Tang, Y. Bao, Y. Tang, G. Chen: *A Fe-C coated long period fiber grating sensor for corrosion induced mass loss measurement.* Optics letters *41* (2016), 2306–2309.

[8] Y. Y. Zhang, E. Mintzer, and K. E. Uhrich: *Synthesis and Characterization of PEGylated Bolaamphiphiles with Enhanced Retention in Liposomes*, Journal of Colloid and Interface Science *482* (2016), 19–26.

[9] N. Arunkumar, K. M. Sirajudeen: *Approximate Entropy based ayurvedic pulse diagnosis for diabetics - A case study.* TISC 2011 - Proceedings of the 3rd International Conference on Trendz in Information Sciences and Computing, art. No. 6169099 (2011), 133–135.

Received May 7, 2017