Acta Technica 62 (2017), No. 6B, 1205–1216

A method for improving the precision of association prediction based on fuzzy minimax neural network

Mei Li[1]

Abstract. A kind of relevance prediction method for employed population of urban labor productivity in economic increase based on fuzzy min-max neural network was proposed to increase relevance prediction precision for employed population of urban labor productivity in economic increase. Firstly, prediction model of scale of employed population based on labor productivity was selected based on labor productivity of macroscopic conception; secondly, adding new type or deleting an existing type of general fuzzy min-max neural network through contraction process of super-box was proposed. Advantages of general fuzzy min-max neural network were inherited and randomness at the time of categorization for general fuzzy min-max neural network is avoided; finally, effectiveness of algorithm was verified through simulation experiment.

Key words. Urban labor productivity, Employed population, Neural network, Fuzzy, Employed population.

1. Introduction

Labor is undoubtedly a basic production element of economic development, while population scale is basis of labor supply and one of most active non-marketization factors. With increase of productivity and optimization of production way, labor productivity of all industries continuously increases in different degrees and economic aggregate increases continually. Regional economic scale expands and extends with it and production form is transited from extensive style to intensive style of continual development.

Currently, increase of output efficiency and salary level is manifested for production form of Chinese regional economic development, while increase of employment scale is different due to regional development difference; longitudinal deepening and extent expansion of urbanization degree are manifested for space form of regional

[1]School of Finance and Trade, Ningbo Dahongying University, Ningbo, 315175, China

economic development. Current urbanization rate (ratio of urban population to total population) of China is 46%, still lower than average urbanization rate 49% in the scope of the world. At least 30 percentage points shall be increased in reference to average urbanization rate of 78% of high-income country. It is indicated that Chinese urbanization level increases 1.2 percentage points each year in average during 1998-2012 in statistics. China will be close the aim of being a high-income country if increase of one percent each year for urbanization rate in future 30 years can be kept speculated with this development trend, and country population of over 400 million will migrate to urban areas at the same time. Boosted urbanization process of high speed exerts important functions in aspects such as expanding consumption market, optimizing resource configuration, increasing living standard and accelerating economic development, etc. But, continuous expansion of urban scale, highly concentration of urban facility, fast expansion of urban population and rapid increase of urban burden will certainly lead to negative effects such as land shortage, traffic jam, worse public security and environmental population, etc. Besides, plenty of negative externalities will be produced inevitably in resource optimization configuration process advocated and relied by urbanization process, and these negative externalities will act on social environment, natural environment and economic structure in an unbalanced and unsymmetrical way. So, whether matching of hard constraint of resource to urban scale, urban development stage and urban development way, etc is reasonable is of vital influence on the future of a city.

A kind of relevance prediction method for employed population of urban labor productivity in economic increase based on fuzzy min-max neural network was proposed, and advantages of general fuzzy min-max neural network were inherited and randomness at the time of categorization for general fuzzy min-max neural network is avoided, thus precision increase of relevance prediction for employed population was realized.

2. Theoretical model and frame

Labor productivity in traditional meaning is ratio of labor result which the labor creates in certain period to suitable labor consumption amount. Labor productivity level of microscopic enterprises can be measured in two ways, and one is indicated with number of certain products produced by the labor in unit time; the second is indicated with labor time consumed to produce unit product. Macroscopic labor productivity, also called overall labor productivity, refers to product output of each employee in unit time in average calculated according to value index of this product. Extended to industry or area, overall labor productivity refers to mixed average labor productivity in different industries and areas. Output is a quantitative conception of dimension nature, so output value is usually selected as measurement standard in market environment. It can not only eliminate non-conformity of dimension, but also it can realize objective comparison and evaluation. Logic core of prediction model for employed population proposed in this thesis is labor productivity of macroscopic conception. Prediction model of employed population is established for labor productivity. Indication of the most basic productivity is as follows:

$$p_i = \frac{GDP}{Q_l}\,. \tag{1}$$

p_i is labor productivity of macroscopic conception, and Q_l is labor quantity, namely scale of employed population, and GDP is regional economic scale. Elasticity of traditional conception is ratio of change degree for inter-connected two economic indexes in certain period and it is responsive degree of change amplitude for a economic variable to another change amplitude for a economic variable. For example, growth elasticity between a certain area and national economic development speed is:

$$E = \frac{Gdp_{t-1}/Gdp_t}{GDP_{t-1}/GDP_t}\,. \tag{2}$$

Numerator is regional economic growth speed, and denominator is national economic growth speed, and t is time axis. E indicates elasticity between regional economic growth speed and national economic growth speed. Mathematic indication way of elasticity prediction system based on labor productivity, namely prediction model of employed population is as follows:

$$\begin{aligned} Labor_t \sum_{i=1}^{i=3} \frac{GDP_{i,t}}{produotivity_{i,t}} \\ = \sum_{i=1}^{i=3} \frac{\sum_{x=1}^{n} \left(\frac{Gdp_{x,t-1}/Gdp_{x,t}}{GDP_{i,t-1}/GDP_{i,t}}\right) GDP_{x,t,i}}{\sum_{x=1}^{m} \left(\frac{Gdp_{x,t-1}/Q_{x,t}}{GDP_{i,t-1}/Q_{i,t}}\right) Labor_{x,t,i}} \end{aligned} \tag{3}$$

T is time axis, and l indicates three major industries, and x indicates different trades. M and n respectively indicates selected sample number. Horizontal differentiation and longitudinal differentiation are decomposed among enterprise, trade, industry, city and nation by the model, and macroscopic economic growth speed is reached through transmission of elasticity difference step by step of microscopic enterprise. Finally scale of employed population is obtained in combination with change of labor productivity uniformly. Eg is urban and national GDP elasticity, and El is industrial elasticity of labor productivity.

$$Eg_{x,t,i} = \frac{Gdp_{x,t-1}/Gdp_{x,t}}{GDP_{i,t-1}/GDP_{i,t}}\,. \tag{4}$$

$$El_{x,t,i} = \frac{Gdp_{x,t-1}/Q_{x,t}}{GDP_{i,t-1}/Q_{i,t}}\,. \tag{5}$$

Model prediction relies on many basic variables, including national predicted GDP, urban predicted GDP, predicted GDP of all industries, etc. Processing way of time sequence is adopted for prediction of economic growth speed in existing literature, and applications such as ARIMA, exponential smoothing, etc are universal. National predicted GDP and urban predicted GDP, etc in model can be categorized as exogenous variables. They can not only be obtained through methods such as statistical prediction, etc according to time sequence data, but also they can be determined on the nature through policy planning and empirical data. Set urban predicted GDP data as an example, regional economic development relies on national economic development trend, and GDP development goals of different areas and levels are determined and emphasized in documents of various levels such as national economic planning and urban development planning, etc. So, indexes of various levels are decomposed and implemented in steps, and economic goal becomes economic task. At the same time, it becomes development goal of all industries and important guide for industries to absorb capital, technology and labor. So, it conforms to objective reality to summarize it as exogenous variable and include it into model for prediction of employed population under Chinese current economic development way and planning environment of powerful government. In consideration of factors such as ways, intensity and effect, etc of government regulation in the process of current economic development, scale change of urban employed population is controlled by understanding of the government on economic development speed to a great extent. So, government has powerful influence on economic development in terms of current national situation. It is a selection better conforming to reality to conduct converse deduction in phases on economic growth speed through government planning for accompanied mobility of current urban employed population on employment opportunities. So, it is chosen to conversely deduct government goal to achieve national and urban index value of economic growth speed in this thesis.

3. Improved fuzzy min-max neural network model and topological structure

3.1. Model of overlapping

Training sign sample is handled through fuzzy min-max neural network by us firstly in solving the problem of employed population of labor productivity. A kind of situation is described in Fig. 1: super-box needs contraction to eliminate overlapping when two super-boxes are overlapped. Result of contraction process is shown in Fig. 1(b). It is noted that memberships of B and C belonging to type I and type II are respectively u_1 and u_2.

Observe membership change of B and C:

Before the process of contraction: $u_2(B) = 1$, $0 \le u_1(B) < 1$

$u_1(C) = 1$, $0 \le u_2(C) < 1$

After the process of contraction: $u_1(B) = 1$, $0 \le u_2(B) < 1$

$u_2(C) = 1$, $0 \le u_1(C) < 1$

Memberships of B and C belonging to type I and type II respectively change, indicating that contraction process also leads to training error meanwhile and non fuzzy area (namely grey area) is also disturbed similarly. Overlapped compensation nerve cells are added in the training process for improved fuzzy min-max neural network pertinent to overlapping of the type in Fig. 1, thus accuracy of network will be higher.

Overlapped compensation nerve cell is a super-box, equaling to overlapped sec-

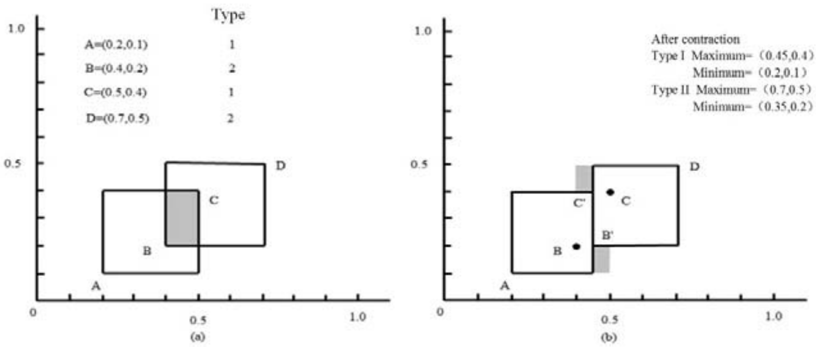

Fig. 1. Problem of overlapping

tion between two super-boxes belonging to different types. Overlapped compensation nerve cell produces two output values corresponding to different types (as Fig. 2(a)). When test set data is in overlapped area, overlapped compensation nerve cell is activated, and activation function is:

$$d_{j_p} = T\left(b_j\left(\mathbf{a}_h, \mathbf{V}_j, \mathbf{W}_j\right) - 1\right) \times \left(-1 + \frac{1}{n}\sum_{i=1}^{n} \max\left(\frac{a_{hi}}{w_{p_{ji}}}, \frac{v_{p_{ji}}}{a_{hi}}\right)\right). \tag{6}$$

$d_{j_p}, p = 1, 2$ is output value of type I and type II, and $\mathbf{a}_h$ is input data. $\mathbf{V}_j, \mathbf{W}_j$ are minimum and maximum of overlapped compensation neural cell, $\mathbf{V}_j = (v_{j1}, v_{j2}, \cdots, v_{jn}), \mathbf{W}_j = (w_{j1}, w_{j2}, \cdots, w_{jn})$

$T(x)$ is threshold function:

$$T(x) = \begin{cases} 1 & x \geq 0\,, \\ 0 & x < 0\,. \end{cases} \tag{7}$$

When test set data is outside overlapped area, there is to avoid activation of equation (1):

$$T\left(b_j(x) - 1\right) = \begin{cases} 1 & b_j = 1\,, \\ 0 & b_j \neq 1\,. \end{cases} \tag{8}$$

$b_j\left(\mathbf{a}_h, \mathbf{V}_j, \mathbf{W}_j\right)$ is membership function[4] of network:

$$b_j\left(\mathbf{a}_h, \mathbf{V}_j, \mathbf{W}_j\right) = \min_{1 \leq i \leq n}\left(\min\left(\left[1 - f\left(a_{hi} - W_{ji}, \gamma\right)\right], 1 - f\left(V_{ji} - a_{hi}, \gamma\right)\right)\right) \tag{9}$$

$\mathbf{V}$ and $\mathbf{W}$ are minimum and maximum of super-box in node layer of$b_1, b_2, \cdots\cdots, b_j$.

$f(x, \gamma)$ is threshold ramp function of double parameters:

$$f(x, \gamma) = \begin{cases} 1 & x\gamma > 1\,, \\ x\gamma & 0 \leq x\gamma \leq 1\,, \\ 0 & x\gamma < 0\,. \end{cases} \tag{10}$$

γ is fuzzy-control parameter [5]:

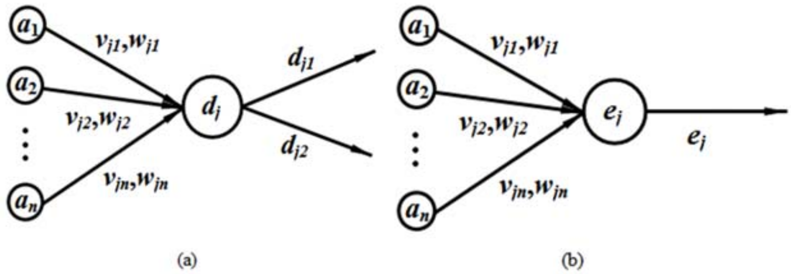

Fig. 2. Overlapped compensation nerve cell and inclusion compensation nerve cell

Inclusion problem of super-box is described in Fig. 3(a), and type II is totally included in type I. This problem is solved through contraction of super-box for fuzzy min-max neural network also. Membership change of D in type I is shown in Fig. 3(b), and $u_1(D) = 1$ before contraction process. $0 \leq u_1(D) < 1$ after contraction process, and contraction process also leads to categorization error.

Partial inclusion problem of super-box is described in Fig. 4(a), and it is shown in Fig. 4(b) after contraction. Maximums of type II and type I of partial inclusion change and membership of B also changes. Non-fuzzy area is also affected.

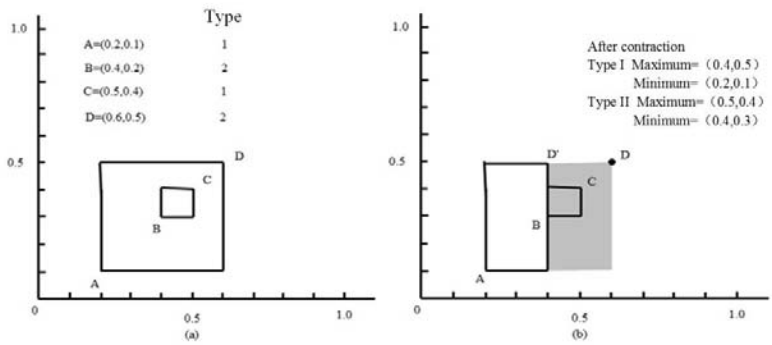

Fig. 3. Problem of total inclusion

Inclusion compensation nerve cell is added to improved fuzzy min-max neural network pertinent to overlapping in the figure. It handles problems of partial inclusion and total inclusion, equaling to overlapped area between two super-boxes of different types. When test set data is in the area, inclusion compensation nerve cell is activated, and activation function is:

$$e_j = -1 \times T\left(b_j\left(\mathbf{a}_h, \mathbf{V}, \mathbf{W}\right) - 1\right) . \tag{11}$$

e_j is output value of inclusion compensation nerve cell, and $\mathbf{V}, \mathbf{W}$ are minimum and maximum of inclusion compensation nerve cell. $\mathbf{a_h}$ is input data, and $T(x)$ is threshold function.

3.2. Topological structure of improved fuzzy min-max neural network

Input mode is an n-dimension fuzzy quantity in form of section for improved fuzzy min-max neural network, namely a n-dimension super-box. Meanwhile, a C_0 type is added in output terminal corresponding to input data with type not indicated. A new type can be added or an existing type can be deleted after adding compensation nerve cell. Specific topological structure of network is shown in Fig. 4:

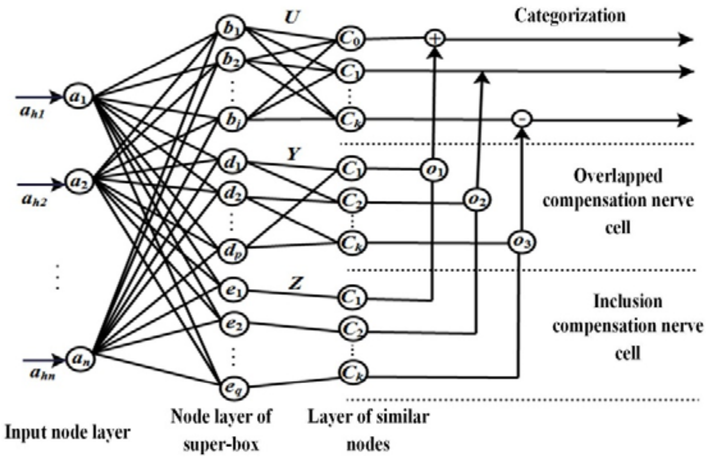

Fig. 4. Topological structure of improved fuzzy min-max neural network

$a_{h1} - a_{hn}$: Input data
$a_1 - a_n$: Input node
$b_1 - b_j$: Node of super-box
$C_0 - C_k$: Similar node
$d_1 - d_p$: Node of overlapped compensation super-box
$e_1 - e_q$: Node of inclusion compensation super-box
$o_0 - o_k$: Overall compensation node
U, **Y**, **Z**: Connection matrix

4. Experimental analyses

4.1. Performance analysis of neural network

The goal of neural network is to offer accurate prediction of employed population according to labor productivity. Problems of feasibility and accuracy are mainly verified in this thesis for improved general fuzzy min-max neural network, and 100 groups of sample data in recent five years are collected. 50 groups of the data are used as data and the other 50 groups of data is inspection data. Final evaluation results of enterprises including A (excellent), B (good), C (median) and D (bad) through training with improved general fuzzy min-max neural network.

There are totally three layers of general fuzzy min-max neural network constructed in this thesis. Number of input neural nerve cell is 24 and number of output nerve cell is 1, and it is achieved that number of suitable specific hidden layer nerve cell is 15 through verification with Kolmogorov theorem. Error between super-box and training data is Fig. 5 upon practical operation, and error reaches minimum when super-box expands in certain scope for improved fuzzy min-max neural network. Higher level is achieved for improved general fuzzy min-max neural network to employed population of labor productivity through the process of expansion and contraction for super-box.

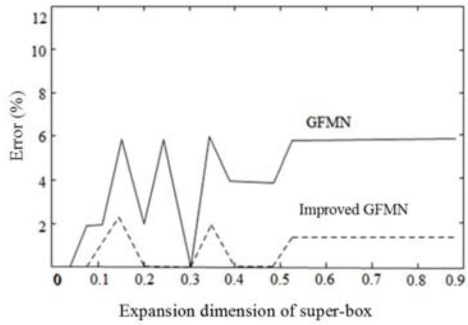

Fig. 5. Error analysis

Test accuracy can be affected through changing dimension of super-box and the aim of adding new type or deleting old type can be achieved through changing its number. Comparison of it and general fuzzy min-max neural network is shown in Fig. 6:

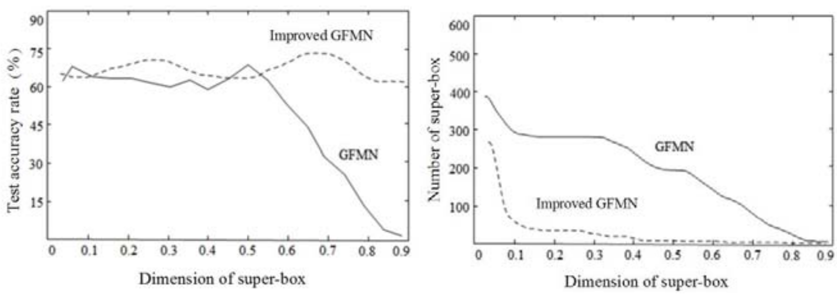

Fig. 6. Test accuracy rate and number of super-box

It can be found in Fig. 7 that general fuzzy min-max neural network of adding new type or deleting old type in this thesis to avoid randomness of general fuzzy min-max neural network at the time of categorization and increase accuracy and efficiency of employed population for labor productivity.

4.2. Relevance prediction of employed population

Sample data is from *China Statistical Yearbook 2013* and *Statistical Yearbook of Selected Cities in 2013*. It is indicated from data in Table 1 that economic growth of selected cities has been stable in recent 20 years with slight fluctuation in the period, such as low points in 21994 and 1996 and high points in 2003 and 2004, but it is of overall high-speed stable development in two digits. Convergent tendency will be stronger in the future due to constraint of marginal decrease rule. So, Elasticity value of economic development of selected cities and national GDP growth rate will be round 1.2 of gradually stability in premise of no major exogenous impact. Economic growth speed of selected cities and national GDP are respectively predicted with ARIMA method adopted and the conclusion is close to elasticity prediction result.

Predicted output value of the three major industries is achieved with constructed elasticity system and all industrial elasticity and output value substituted, thus prediction is conducted on labor productivity and scale of employed population. Prediction data is shown in Table 1.

Table 1. Prediction of related data for employed population

	2012	2013	2015	2020
Output of the primary industry (100 million yuan)	195.08	203.66	220.07	260.10
Output of the secondary industry (100 million yuan)	6502.25	7022.43	8198.55	11725.92
Output of the tertiary industry (100 million yuan)	5314.32	6031.75	7600.16	12575.98
Population of the primary industry (ten thousand)	25.34	23.41	19.81	19.87
Population of the secondary industry (ten thousand)	334.98	331.91	326.15	326.15
Population of the tertiary industry (ten thousand)	227.95	233.08	254.08	246.93
Total employed population (ten thousand)	588.27	589.42	600.03	603.59

5. Conclusion

An elasticity system with comprehensive consideration of variable factors of labor productivity in all industries in micro level and city of macro level and national economic growth speed included was constructed in this thesis on the basis of perspective of labor productivity. “Enterprise-trade-industry-economic entity” was linked through elasticity measurement of horizontal differentiation and longitudinal differentiation, thus scale of dynamic employed population can be predicted. Prediction logic of obtaining optimal solution of “how should be” through many condition constraints was abandoned for it and time sequence data based on objective reality was selected, and scale prediction of future employed population was completed through the hypothesis that there would be no development by leaps and bounds for labor productivity in short term in consideration of substituting of government planning as powerful role in the process of economic development. This prediction model of employed population is applicable to cities with good economic basis, where there is certain difference between employed scale and registered population and employment market has big mobility. Selected cities basically conform to above hypothesis conditions, so it is applicable to predict future employed population scale in perspective of labor productivity.

Several basic judgments are formed in this thesis through scale prediction of employed population in future years for selected cities: firstly, productivity of the secondary industry for selected city is increasing and economic scale is expanding with adjustment of industrial structure, but corresponding employment scale is relatively stable for growth is mainly derived from productivity; secondly, overall scale of the tertiary industry for selected city will expand and employed population will continually increase, but growth speed is mild; actual type and scale of the tertiary industry and labor quantity should be slightly bigger than prediction value, and the secondary industry in transition phase can only be supported and certain production efficiency can only be maintained by moderate scale of the tertiary industry in perspective of economic operation. Thirdly, total amount of overall employed population for selected city tends to be stable. Employed population scale is composed of employed population quantity of the three major industries. Employment opportunity will decrease and employment quantity tends to decrease in stability due to structure adjustment and increase of labor productivity for the secondary industry; labor productivity and employment opportunity will gradually increase due to expanded scale of the tertiary industry. Frequency and amplitude of mobility among industries for labor are increased by frictional unemployment. Small amount of net flow-in of employed population in future periods for selected cities is achieved by employment opportunity and overall employment amount gradually tends to be stable.

References

[1] Pahlavan K, Krishnamurthy P and Geng Y: (2015) *Localization challenges for the emergence of the smart workd.* IEEE Access, 3(1), pp. 3058-3067

[2] Lv, Z., Tek, A., Da Silva, F., Empereur-Mot, C., Chavent, M., & Baaden, M. (2013).: *Game on, science-how video game technology may help biologists tackle visualization challenges.* PloS one, 8(3), e57990.

[3] Shen G U, Jingbiao M A, Hanmuo M A: (2015) *The Way of Chinese Urban-rural Dual Economic Structure Transformation: Based on Dynamic Cycle Evolution of Labor Division*[J]. Journal of Audit & Economics.

[4] Duarte M, Restuccia D (2006)Mei L I, et al.: *The Evolution and Transformation Strategies of Urban-rural Dual Economic Structure in Beijing*[J]. Science Technology & Industry, 2013.

[5] Roumasset J, Lee S H: (2007) *Chapter 52 Labor: Decisions, Contracts and Organization*[M]// Handbook of Agricultural Economics. Elsevier B.V. 2007:2705-2740.

[6] Mion G: (2004) *Spatial externalities and empirical analysis: the case of Italy*[J]. Journal of Urban Economics, 56(1):97-118.

[7] Messina J: (2006) *The role of product market regulations in the process of structural change*[J]. European Economic Review, 50(7):1863-1890.

[8] Zuleta H, Young A T: (2006) *Labor's Shares - Aggregate and Industry: Accounting for Both in a Model of Unbalanced Growth With Induced Innovation*[J]. Social Science Electronic Publishing, 24:112–122.

[9] Biemann T, Fasang A E, Grunow D: (2011) *Do Economic Globalization and Industry Growth Destabilize Careers? An Analysis of Career Complexity and Career Patterns Over Time*[J]. Organization Studies, 32(12):1639-1663.

[10] Weisen Pan, Shizhan Chen, Zhiyong Feng: (2013) *Automatic Clustering of Social Tag using Community Detection.* Applied Mathematics & Information Sciences, 7(2): 675-681.

[11] Yingyue Zhang, Jennifer W. Chan, Alysha Moretti, and Kathryn E. Uhrich: (2015) *Designing Polymers with Sugar-based Advantages for Bioactive Delivery Applications*, Journal of Controlled Release, 219, 355-368.

[12] Yingyue Zhang, Qi Li, William J. Welsh, Prabhas V. Moghe, and Kathryn E. Uhrich: (2016) *Micellar and Structural Stability of Nanoscale Amphiphilic Polymers: Implications for Anti-atherosclerotic Bioactivity*, Biomaterials, 84, 230-240.

[13] Jennifer W. Chan, Yingyue Zhang, and Kathryn E. Uhrich: (2015) *Amphiphilic Macromolecule Self-Assembled Monolayers Suppress Smooth Muscle Cell Proliferation*, Bioconjugate Chemistry, 26(7), 1359-1369.

Received May 7, 2017