Acta Technica *62* (2017), No. 6B, 157–166

Mutual influence between eastern literature and western literature in vitoria period

Yuan Hu[1]

Abstract. One research program on interaction between eastern and western literature based on Mapreduce parallel evolutionary tree algorithm (MRRF) has been proposed to improve the accuracy of interaction analysis between eastern and western literature. First, based on the basic idea of HashRF algorithm, grid construction of MapReduce model of evolutionary tree as well as grid parameter calculation have been realized; second, two-phase MaoReduce process has been adopted to attain distance sub-matrix of Robinson-Foulds as well as joint distance matrix and realize Robinson-Foulds matrix representation of evolutionary tree under MapReduce structure. Finally, the proposed algorithm has been applied to the interaction research on eastern and western literatures. Experimental results have shown that the proposed method can attain analysis results of interaction research between western and eastern literatures more quickly and effectively.

Key words. Victoria, Mapreduce parallel, Eastern and western literatures, Influence research
.

1. Introduction

Literature is an art that reflects objective reality vividly by taking languages as the tools. It includes poems, essays, novels, plays, fables and fairy tales etc. It is an important expression of culture, which expresses internal feelings and reproduces social life in certain period of time and in certain region with different forms (called as genres). As an activity originated from human thinking, literature is greatly affected by human thoughts, values, living environment and social background etc. It can be said that literatures of different countries and different periods of time is with distinctive national color, regional color and era color. With the increasing of communication means and improvement of communication demand, cross-cultural communication has become an inevitable trend in times development and it has

[1]Department of Postgraduates, Shanghai International Studies University, Shanghai, China, 201620

become a current wave. As the subjects of cross-cultural communication, people are both communicators and recipients. Studying eastern and western differences and well realizing mutual communications on the basis of respecting differences are of great significance to study the differences between eastern and western culture. This paper takes differences between eastern and western literatures in cross-cultural communication as main research objects and selects translation versions of different literary works as research carriers to study various literary forms under the influence of eastern and western culture.

In cross-cultural communication, the typical means of literary communication is translation of various literary works. It is not only a cross-cultural behavior, but also a cross-cultural art practice activity, which increases communications among different countries and different people. Rich human geography, traditional code of conduct, national customers, social and cultural phenomena of other countries included in literary works will the delivered to people of other countries through the work of translators. It is a good diagnostic behavior for culture of other countries. In different eastern and western literatures, there are a lot of literary images and literary types with national and regional colors and they are the products under different folk customs in different eastern and western development periods. Translations and elaborations for these images will cause misunderstandings and barriers in cultural communication if do not put emphasis on their cultural differences and aesthetic connotations.

Influence of eastern and western literatures is the greatest motivation to promote regional economic development and the most fundamental factor of productivity development. Eastern and western literatures in one region affect the productivity level and economic development ability directly. It is of important significance to evaluate the influence of eastern and western literatures in one area, one region and even one country on development degree. This paper considers about adopting evolutionary tree algorithm to analyze influence of regional literature, attain relatively accurate analysis results and instruct practice decision.

2. Evaluation index for influence of eastern and western literatures

It is very important to construct a scientific, practical and reasonable evaluation index system to study the influence of eastern and western literatures; it is of great significance to evaluate the influence of eastern and western literatures of one region and one country. There are four categories and 13 indexes for the index system of influence of regional eastern and western literatures, which cover various factors of technology, literature and education etc that promote regional investment, as shown in figure 1. Considering about the education, education expenditure, number of colleges and universities, proportion of people with at least junior college degree and number of students in colleges and universities with over 10,000 students have been selected. Considering about the culture, number of presses, per capita purchasing of books and editions of news papers have been selected. Considering about technology, expenditure, cost for three items of science and technology, number of accepted

patent applications and transaction amount of technology market have been selected. All these factors reflect scientific research level and transformation ability of science and technology in one region comprehensively. In which, expenditure refers to the expenditure spent by one region to increase total amount of knowledge, with which it can create new applications to make systematic and creative activities. It includes three kinds of fundamental research, application research and experimental development. It is usually used to reflect regional science and technology strength and core competitiveness internationally. Cost for three items of science and technology refers to the cost of the country for science and technology expenditure, including development fee for new product, intermediate experiment fee and subsidy for important scientific research. In addition, we also select two indexes of number of internet users and per capita education, culture and entertainment expenditure. The reason for selecting number of internet users is that network is playing an increasingly prominent role in social and economic life in the new period; in particular, it can reflect influencing level of eastern and western literatures in one region comprehensively. As for per capita education, culture and entertainment expenditure, it reflects the regional consumption tendency comprehensively, which explain the comprehensive cultural level of people from a side view.

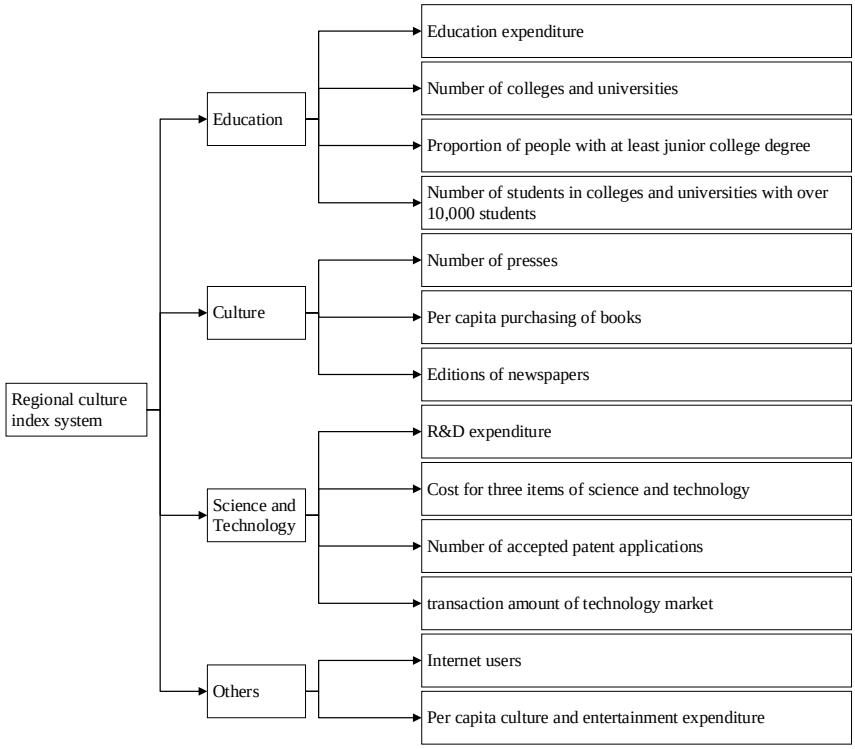

Fig. 1. Evaluation index system to study the influence of eastern and western literatures

Regional culture index system	Education expenditure	Per capita purchasing of books	transaction amount of technology market
Education	Number of colleges and universities	Editions of newspapers	Internet users
Culture	Proportion of people with at least junior college degree	R&D expenditure	Per capita culture and entertainment expenditure
Science and Technology	Number of students in colleges and universities with over 10,000 students	Cost for three items of science and technology	
Others	Number of presses	Number of accepted patent applications	

3. Parallel evolutionary tree algorithm based on Mapreduce

3.1. Algorithm description

Design idea of MRRF algorithm is from Hash table based Robinson-Foulds (HashRF in short) [12]. It is believed that input tree document includes cluster configuration of t trees and $N \times c$. Cluster node will designate physical machine to execute program code. The number of core calculating is the CPU core number of each physical node. For serial execution, $N = 1$, $c = 1$. If on two machines, each machine has four CPU cores, and then $N = 2$, $c = 4$. MRRF algorithm mainly has two steps:

Step 1: (grid construction) make grid division for N nodes and then divide t inputs based on node. Phoenix frame is adopted to make automatic allocation for input of single node and the configuration is $1 \times c$. Grid construction relates to N value.

A. If N can construct perfect square, and then it can adopt n nodes to construct $\sqrt{N} \times \sqrt{N}$ grid.

B. If N can't construct perfect square, set $i = \left\lfloor \sqrt{N} \right\rfloor$ (it indicates rounding through square root). If $N \bmod i = 0$, the calculation node can construct grid $N/i \times i$. If $N \bmod i \neq 0$, and then $i = i - 1$, until it meets $N \bmod i = 0$.

For example, when $N = 4$, grid dimension is 2×2 (as shown in figure 2). When $N = 18$, grid dimension is 6×3.

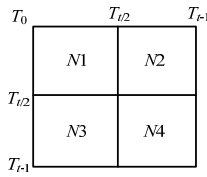

Fig. 2. MRRF grid partition scheme

Step 2: (parameter calculation) make use of OpenMPI to deploy node in grid, and then execute $MRRF\,(p, q)$ algorithm on each node to attain $p \times q$ sub-

matrixes. For example for 2×2 in figure two, sub-matrix of each partition can be calculated as:

$N1$: $MRRFpq\left(T_{0:(t/2-1)}, T_{0:(t/2-1)}\right)$, and then row and column of this region are $\mathcal{Q}1 = \left\{T_0, \cdots, T_{0:(t/2-1)}\right\}$, $P1 = \left\{T_0, \cdots, T_{0:(t/2-1)}\right\}$, it can get $p1 = |P1|$, $q1 = |\mathcal{Q}1|$;

$N2$: $MRRFpq\left(T_{(t/2-1):(t-1)}, T_{0:(t/2-1)}\right)$, and then row and column of this region are: $\mathcal{Q}2 = \left\{T_{t/2-1}, \cdots, T_{t-1}\right\}$, $P2 = \left\{T_0, \cdots, T_{0:(t/2-1)}\right\}$, it can get $p2 = |P2|$, $q2 = |\mathcal{Q}2|$;

$N3$: $MRRFpq\left(T_{0:(t/2-1)}, T_{(t/2-1):(t-1)}\right)$, and then row and column of this region are $\mathcal{Q}3 = \left\{T_0, \cdots, T_{0:(t/2-1)}\right\}$, $P3 = \left\{T_{t/2-1}, \cdots, T_{t-1}\right\}$, it can get $p3 = |P3|$, $q3 = |\mathcal{Q}3|$;

$N4$: $MRRFpq\left(T_{(t/2-1):(t-1)}, T_{(t/2-1):(t-1)}\right)$, and then row and column of this region are $\mathcal{Q}4 = \left\{T_{t/2-1}, \cdots, T_{t-1}\right\}$, $P4 = \left\{T_{t/2-1}, \cdots, T_{t-1}\right\}$, it can get $p4 = |P4|$, $q4 = |\mathcal{Q}4|$.

3.2. Robinson-Foulds matrix

The core of MRRF algorithm lies in the calculation for sub-thread $MRRF\,(p, q)$. If each node has t trees, the subordinate relationship between tree and set of nodes P and $\mathcal{Q}$ can be known through keys within the node grid [13]. In figure 2, row trees of node N_3: $P3 = \left\{T_{t/2-1}, \cdots, T_{t-1}\right\}$, column trees: $\mathcal{Q}3 = \left\{T_0, \cdots, T_{0:(t/2-1)}\right\}$. Assume node N_3 has eight calculation cores and then trees of $P3$ and $\mathcal{Q}3$ will be divided into four files and distributed to 8 calculation cores as inputs.

In $MRRF\,(p, q)$ calculation, compare all the trees in P and all the trees in$\mathcal{Q}$, it meets the relation $P \equiv \mathcal{Q}$. It can realize parallel of node N_i through following two MapReduce processes:

Step 1: (Mapper 1) refers to HashRF algorithm and generate Hash table firstly. The system offers one key for each partition and its value is number of tree TIDs in partition. However, the number of Mapper functions is determined by the calculation cores of specific nodes and then get $\langle key, value\rangle$.

Map process in figure 3 has two Map functions. The first Map function attains following $\langle key, value\rangle$ in total:

$$\begin{cases} \langle key1, value1\rangle = \langle AB\,|CDE\,, (1, T0, T1)\rangle\,, \\ \langle key2, value2\rangle = \langle ABC\,|DE\,, (1, T0)\rangle\,, \\ \langle key3, value3\rangle = \langle ABD\,|CE\,, (1, T1)\rangle\,. \end{cases} \tag{1}$$

The second Map function attains following $\langle key, value\rangle$ in total:

$$\begin{cases} \langle key4, value4\rangle = \langle AB\,|CDE\,, (2, T2)\rangle\,, \\ \langle key5, value5\rangle = \langle BC\,|ADE\,, (2, T3)\rangle\,, \\ \langle key6, value6\rangle = \langle ABC\,|DE\,, (2, T2, T3)\rangle\,. \end{cases} \tag{2}$$

And then, for these $\langle key, value\rangle$ sets, combine based on key value through middle processing.

Step 2: (Reducer 1) After Map1 stage finished, ⟨key, list(value)⟩ will be taken as the input of r Reducer. For each two-fork tree, there are m and treeIDs at most. Each Reducer will combine $O(m)$ list and separate the trees in file 1 and file 2, and then it can form a straight line in global Harsh table.

As shown in the Reduce phase in figure 3, the key lists produced by the first Reducer are ABD|CE and ABC|DE; therefore, the inputs of this reducer are ⟨ABD|CE, (1, T1)⟩and⟨ABC|DE, (1, T0) , (2, T2, T3)⟩, while the outputs are ⟨ABC|DE, (T0||T2, T3)⟩ and ⟨ABD|CE, (T1||)⟩.

The keys lists produced by the second Reducer are AB|CDE and BC|ADE, and then the inputs of this reducer are ⟨AB|CDE, (1, T0, T1), (2, T2)⟩and⟨BC|ADE, (2, T3)⟩, while the output forms are ⟨AB|CDE, (T0, T1||T2)⟩ and⟨BC| ADE, (T3||)⟩, in which "||" is separator of two input files.

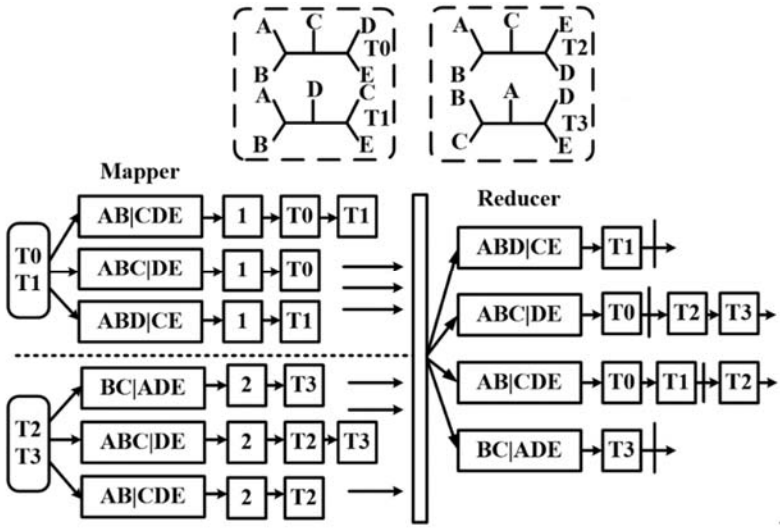

Fig. 3. First phase of MRRF algorithm

Step 3: (Mapper 2) as shown in the Mapper process in figure 4, the list value of key value (AB|CDE) is (T0, T2||T3). Generally, if "||" has u revolutionary tree keys (tree-IDS) on the left and $v$tree-IDS on the right, and then the total comparison times areuv. For example, the comparison times of key value (AB|CDE) is $2 \times 1 = 2$. Compare tree-IDS in the first region and the second region and increase the value of corresponding local similarity matrix. For example, ⟨AB|CDE, (T0, T1|| T2)⟩, add 1 at local similarity matrix positions (T0, T2) and (T1, T2), while ⟨ABC|DE, (T0||T2, T3)⟩, add 1 at local similarity matrix positions (T0, T2) and (T0, T3), as the Mapper process output in figure 4.

Through above Mapper process, key is matrix row ID, while value is corresponding value of matrix row (total times of comparison). For example, in figure 4, the ⟨key, value⟩ of the first Mapper are ⟨T0, (1, 1)⟩ and⟨T1, (0, 0)⟩; the ⟨key, value⟩ of the second Mapper are ⟨T0, (1, 0)⟩ and⟨T1, (1, 0)⟩.

Step 4: (Reducer 2) as shown in the Reduce process in figure 4, Reducer

inputs row keys of similar matrixes, i; combining the output results of Step 3, the first Reducer inputs T0 : (T0, (1, 1), (1, 0)) and the second Reducer inputs T1 : (T1, (0, 0), (1, 0)). The above can be expressed as:

$$\mathrm{T}i: \ \left(\mathrm{T}i, \ \left(v_1^1, v_1^2\right), \cdots, \left(v_{n-1}^1, v_n^2\right)\right) . \tag{3}$$

Reducer operation is:

$$\begin{aligned} &\mathrm{T}i: \ \left(\mathrm{T}i, \ \left(v_1^1, v_1^2\right), \cdots, \left(v_{n-1}^1, v_n^2\right)\right) \rightarrow \\ &\mathrm{T}i: \ \left(\mathrm{T}i, \ \left(v_1^1 + \cdots + v_{n-1}^1, v_1^2 + \cdots + v_n^2\right)\right) . \end{aligned} \tag{4}$$

Execute above operation for each node N_i to get Robinson-Foulds sub-matrix and finally get Robinson-Foulds matrix through combination.

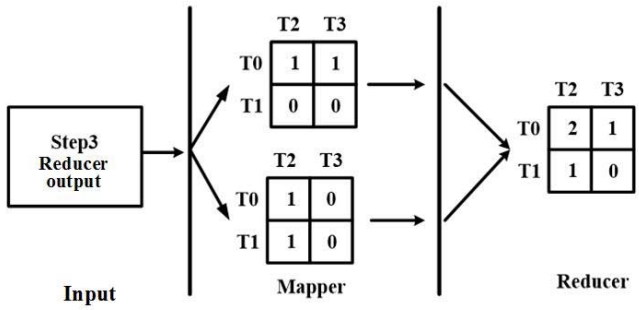

Fig. 4. Second-phase of MRRF algorithm

4. Experimental analysis

Sort out influencing levels of eastern and western literatures in 31 domestic sample regions based on total value of related programs of eastern and western literatures influence in their gross domestic product and make segmentation rating based on the total amount. Items of gross domestic product related to the influence of eastern and western literatures include education, cultural art and radio, film and television, scientific research and comprehensive technical service, health, sports and social welfare etc. The sample values of these items are from 2013 China Statistical Yearbook, see details in table 1.

After establishing the above input and output relation, MATLAB software has been adopted to make sample fitting for this model. The input dimension of this model is 13, output dimension is 1, there are two middle hidden layers, the accuracy is $\varepsilon = 0.0001$, study rate is 0.01. After 23 times of iterations, it can get:

Table 1. Scores of influencing levels of eastern and western literatures in each region

Province	Influence of eastern and western literatures GDP/hundred million Yuan	Score	Province	Influence of eastern and western literatures GDP/hundred million Yuan	Score
Shandong	488,72	5	Jilin		3
Jiangsu	462.67	5	Guangxi		3
Beijing	438.61	5	Chongqing		3
Guangdong	412.82	5	Shanxi		3
Zhejiang	409.32	5	Yunnan		3
Shanghai	364.57	4	Jiangxi		3
Hunan	312.48	4	Xinjiang		2
Liaoning	299.31	4	Shanxi		2
Sichuan	271.57	4	Inner Mongolia		2
Hubei	263.43	4	Guizhou		2
Henan	257.92	4	Gansu		2
Hebei	239.81	3	Hainan		1
Fujian	199.74	3	Qinghai		1
Heilongjiang	171.28	3	Ningxia		1
Anhui	157.72	3	Tibet		1
Tianjin	153.93	3			

$$\omega_1 = \begin{matrix} 5.738 & 2.992 & -3.379 & 0.045 \\ 5.486 & 2.877 & -7.388 & -1.473 \\ -1.261 & 0.625 & -3.663 & -4.369 \\ 2.133 & -2.741 & 1.037 & 0.776 \\ 1.188 & 1.422 & -0.758 & 0.461 \\ 3.316 & -2.765 & 2.939 & -0.999 \\ 1.169 & 1.\ 654 & \text{-}1.683 & \text{-}2\ 258 \\ 7.\ 814 & \text{-}6.538 & 0.676 & 6.084 \\ \text{-}7.278 & & & \end{matrix}$$

$$b_1 = 5.555\ 1.328\ 0.307\,,$$

$$b_2 = -0.031\,,$$

$$\omega_2 = 0.\ 535\ 1.\ 023\ 0.\ 532\,.$$

Through fitting calculation of matlab, this model can well evaluate the influencing level of regional eastern and western literatures. After inputting data of other regions, it can also get satisfactory evaluation results. Therefore, this model can realize the quantitative evaluation for influence of regional eastern and western literatures.

Two simulation comparison charts in figure 5 list the comparison of operation time and comparison of relative acceleration ratio of algorithm under the situation of different evolution numbers in the process of evaluation for eastern and western literatures influencing levels. It can be seen from simulation result that in terms of

operation time, the operation time of MRRF algorithm is less than that of Hash and HashRF algorithm, while HashRF algorithm is faster than Hash algorithm. It indicates that the calculation speed of algorithm with adoption of Robinson-Foulds matrix is faster than that without adoption of Robinson-Foulds matrix; while under the situation of adopting the same Robinson-Foulds matrix, algorithm based on MapReduce is faster than algorithm based on Hash. In terms of acceleration ratio, when the number of evolutionary trees is 2510, HashRF algorithm is 1.274 times of Hash algorithm, while MRRF algorithm is 2.654 times of Hash algorithm. When the number of evolutionary trees is 35634, HashRF algorithm is 1.145 times of Hash algorithm, while MRRF algorithm is 2.867 times of Hash algorithm. The acceleration ratio of MRRF algorithm is the highest among the three algorithms.

5. Conclusions

Evaluation index system for influence of regional eastern and western literature has been constructed scientifically on the basis of deep analysis. MRRF algorithm based on Mapreduce parallel has been adopted to establish related data of 31 provinces and cities and establish evaluation model for influence of regional eastern and western literatures. The verification results are satisfactory and the method is easy to operate. Evaluation for influence of eastern and western literatures in other regions can be made with adoption of this system method. With this evaluation results, it can understand the comprehensive influence level of eastern and western literatures. It can find out shortages through deep analysis and make targeted improvement for influence of regional eastern and western literatures, which is with good reference value for improving influence of eastern and western literatures in this region.

References

[1] X. Han X: *The Mutual Influence on Ancient Literature between China and the West*[J]. Journal of Suihua University (2008).

[2] Y. Finer, F. Jaffer, J. P. Santerre: *Mutual influence of cholesterol esterase and pseudocholinesterase on the biodegradation of dental composites*[J]. Biomaterials, *25* (2004), No. 10, 1787–1793.

[3] F. Gasbarro, F. Rizzi, M. Frey: *The mutual influence of Environmental Management Systems and the EU ETS: Findings for the Italian pulp and paper industry*[J]. European Management Journal, *31* (2013), No. 1, 16–26.

[4] M. A. El-Shirbeny, M. A. Aboelghar, S. M. Arafat, et al.: *Assessment of the mutual impact between climate and vegetation cover using NOAA-AVHRR and Landsat data in Egypt*[J]. Arabian Journal of Geosciences, *7* (2014), No. 4, 1287–1296.

[5] G. W. Pulford: *Taxonomy of multiple target tracking methods*[J]. Radar, Sonar and Navigation, IEE Proceedings, 152 (2005), No. 5, 291–304.

[6] D. S. Amorim, KatiaVitoria, TelmaRossettiFerreira, et al.: *Rede de significa??es: perspectiva para análise da inser??o de bebês na creche*[J]. Cadernos De Pesquisa, *44* (2000), No. 109, 115–144.

[7] D. Erdogmus, K. E. Hild, J. C. Principe, et al. *Adaptive blind deconvolution of*

linear channels using Renyi's entropy with Parzen window estimation[J]. IEEE Transactions on Signal Processing, *52* (2004), No. 6, 1489–1498.

[8] T. Wandmacher, J. Y. Antoine, F. Poirier: *Sibylle, An Assistive Communication System Adapting to the Context and Its User*[J]. Acm Transactions on Accessible Computing, *1* (2008), No. 1, 1–30.

[9] M. Allegra, P. Giorda, A. Montorsi: *Quantum discord and classical correlations in the bond-charge Hubbard model: Quantum phase transitions, off-diagonal long-range order, and violation of the monogamy property for discord*[J]. Physics, *84* (2011), No. 84, 2461–2468.

[10] T. A. Schlacher, A. R. Baco, A. A. Rowden, et al.: *Seamount benthos in a cobalt-rich crust region of the central Pacific: conservation challenges for future seabed mining*[J]. Diversity & Distributions, *20* (2014), No. 5, 491–502.

[11] M. Cuenca-García, F. B. Ortega, J. R. Ruiz, et al.: *More Physically Active and Leaner Adolescents Have Higher Energy Intake*[J]. Journal of Pediatrics, *164* (2014), No. 1, 159–166.

[12] C. Wöhler, P. D'Angelo, L. Krüger, et al.: *Monocular 3D scene reconstruction at absolute scale*[J]. Isprs Journal of Photogrammetry & Remote Sensing, *64* (2009), No. 6, 529–540.

[13] E. Pellegrini: *PSII photochemistry is the primary target of oxidative stress imposed by ozone in Tilia americana*[J]. Urban Forestry & Urban Greening, *13* (2014), No. 1, 94–102.

Received May 7, 2017