

CLASSIFICATION TREE FOR MODELLING THE RELATIONSHIP BETWEEN IT&C AND THE ECONOMIC DEVELOPMENT FOR THE EU-28 MEMBER STATES

Daniela Alexandra CRISAN¹, Mihai Christian CRISAN²,
Oana Teodora GRIGORE³, Razvan BALUTA⁴

Abstract

The relationship between IT&C and the economic development on EU level has been investigated through many approaches. In this paper a new solution based on classification trees is proposed. The classification tree will describe models, templates for certain types of economies in EU-28, based on IT&C indicators. Three types of economies (classes) have been considered: Advanced economies, Developing economies and Modest economies; they have been clustered using their GDP-per capita indicator. The classification tree uses ten IT&C attributes provided by The Global Information Technology Report (GITS) 2015, they are the ten pillars composing the Network Readiness Indicator.

Keywords: IT&C, GDP, NRI, classification tree, CART, data analysis

JEL Classification: C22, J24, M15

This paper has been financially supported by scientific research within the project entitled “PRACTICAL SCHOOL: Innovation in Higher Education and Success on the Labour Market”, project identified as POSDRU/156/1.2/G/132920. The project is co-financed by the European Social Fund through the Sectorial Operational Programme for Human Resources Development 2007-2013. Investing In people!

¹ Associate Professor, PhD, School of Computer Science for Business Management, Romanian-American University, Bucharest, Romania e-mail: crisan.daniela.alexandra@profesor.rau.ro

² Economist, Financial Supervisory Authority, Bucharest, Romania e-mail: crisanmihai@hotmail.com

³ Master student, School of Computer Science for Business Management, Romanian-American University, e-mail: grigoreoanateodora@gmail.com

⁴ Master student, School of Computer Science for Business Management, Romanian-American University, e-mail: razvan2593@gmail.com

1. INTRODUCTION

Classification trees are a particular example of computer trees, used for offering a clear image of the classification process, by making rules easy to understand and interpret. It is a very attractive method both for exploration and prediction of data.

A classification tree is a flowchart structure able to graphically model a specific phenomenon. It partitions a set of objects into classes using *IF <conjunct / hypothesis> THEN <result>* rules.

The components of a classification tree are:

- internal nodes, corresponding to tests / questions (attributes);
- branches, corresponding to all possible answers to the question contained in the father-node (all possible values of attributes);
- leaves (terminal nodes), containing classes / categories / decisions.

There are several methods to build a classification tree, the most common is Classification and Regression Tree (CART) method. It is a data mining method used to build a classification tree using a training set (historical data). The method consists in repeatedly extracting the most significant attribute and splitting the training set into smaller and smaller subsets according to the alternatives (possible values) of the chosen attribute. Although it has its advantages, i.e. it's fast in computational terms, the use of the CART method is debatable because it generates unstable classification trees: minor changes in the learning set may lead to radical changes in the final tree.

Classification trees have been used in various domains, such as artificial intelligence, robotics and control systems, biomedical engineering, financial analysis, manufacturing and production, agriculture, astronomy, and so on.

In this paper, a new approach based on a classification tree is proposed in order to describe the connection between IT&C and the economic performance of the 28 European Union Member States. This is a continuation of the work published in [2,3] where regression analysis is used based on ISOC (computer use of individuals or in enterprises, percentage) as a measure of the IT&C usage, and the GDP-per capita indicator as a measure of the economic development of a country. Some experiments proving that the two indicators are in a strong relationship were described.

Using a classification tree, we propose a new method for describing a model, a template, for every type of economy based on the IT&C development, at structural level. Using the GDP-per capita indicator, the 28 Member States were divided into three groups: Advanced economies, Developing economies and

Modest economies. The three groups are the classes in the classification tree (classes are represented as leaves).

The internal nodes correspond to the ten pillars of the Networked Readiness Index (NRI), a composite indicator measuring the IT&C use at national level computed annually by World Economic Forum.

The branches in the tree correspond to the values of every pillar. For computational reasons, the every pillar range of values was divided into three groups: modest, moderate, and high values in relation with the EU-28 average.

2. NATIONAL INDICATORS FOR IT&C READINESS AND ECONOMIC PERFORMANCE

2.1. The Networked Readiness Index (NRI)

The Global Information Technology Report (GITR) has been published since 2001 at the initiative of the World Economic Forum and its partner INSEAD. The two organizations “*recognized the need for a report such as because of the increasing proliferation of technology and its effects on advancing global competitiveness*” [9, pg. VII].

The Networked Readiness Index (NRI) is a composite indicator published annually in GITR. It measures the ICT revolution through a set of drivers. Its structure has changed in time. In 2015 edition, the NRI is made up of four subindexes and 10 pillars [9] (see Figure 5. The structure of the Networked Readiness Index 2015):

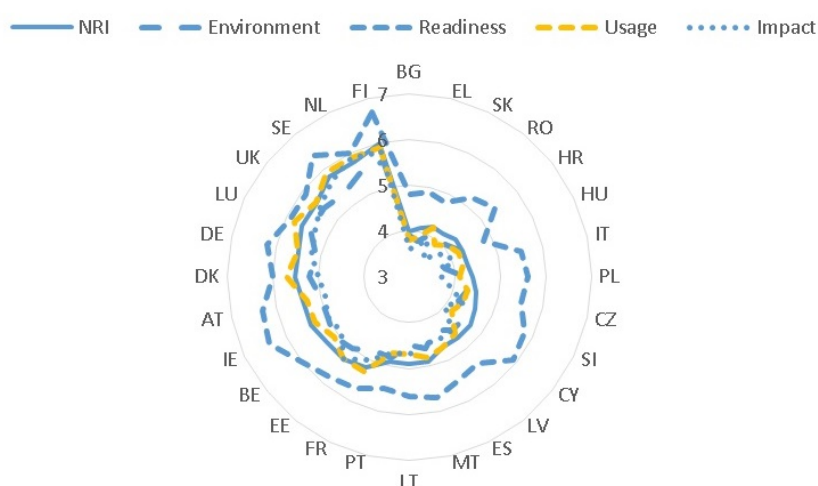
- the “Environment” for ICT is composed of two pillars: (1) Political and regulatory environment and (2) Business and innovation environment and “*assesses the extent to which a country’s market conditions and regulatory framework support entrepreneurship, innovation, and ICT development*” [9, pg. 5].
- the “Readiness” to use ICT is composed of three pillars: (3) Infrastructure and digital content, (4) Affordability and (5) Skills and “*measures the extent to which a country has in place the infrastructure and other factors supporting the uptake of ICTs*” [9, pg. 5].
- the “Usage” of ICT is a composition of three indexes: (6) Individual usage, (7) Business usage and (8) Government usage and “*assesses the extent of ICT adoption by a society’s main stakeholders: government, businesses, and individuals*” [9, pg. 5].
- the “Impact” of ICT: measures, through its two pillars: (9) Economic impacts and (10) Social Impacts, “*the broad economic and social impacts accruing from ICTs*” [9, pg. 6].

The NRI has been computed for 143 economies, but in this paper we'll refer the 28 EU Member States.

Network Readiness Index and its main components for the 28 EU Member States

In this section the structure of the Network Readiness Index for the 28 EU Member States will be presented. The states were considered according to their global performance. In the first graphic, NRI and its four subindexes are represented.

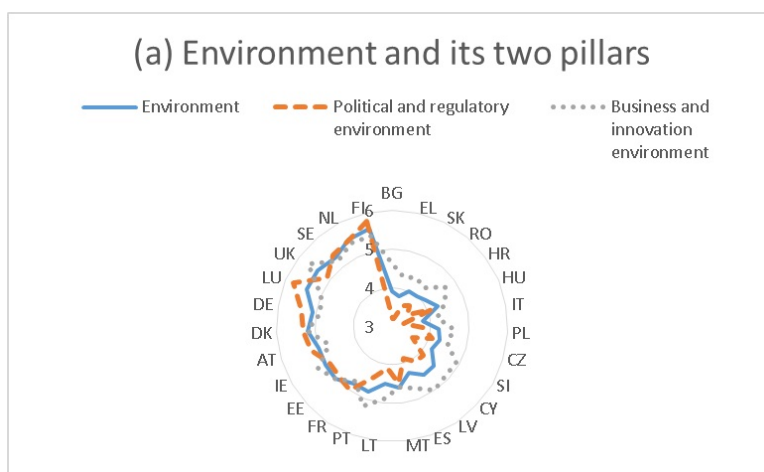
Figure 1. NRI index and its four subindexes

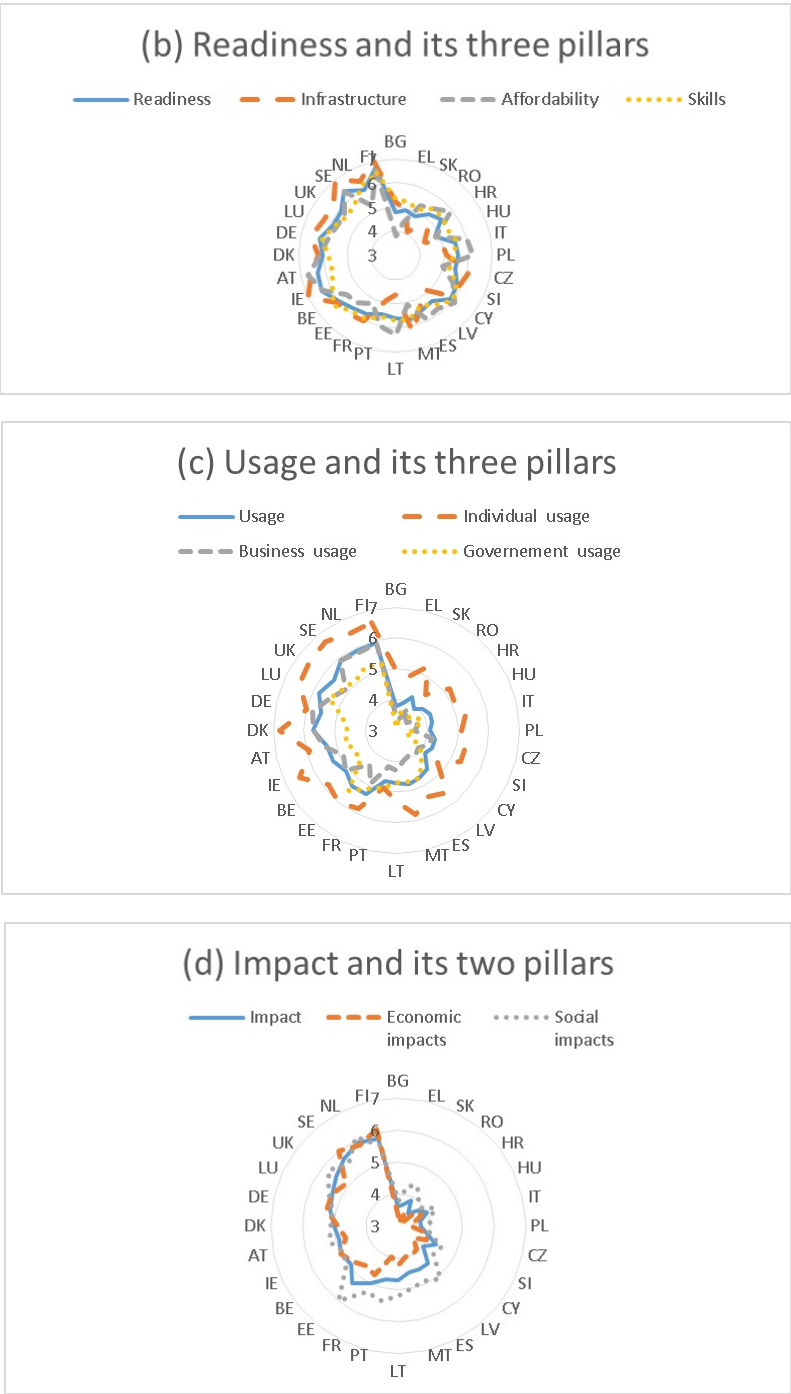


Source: made by the authors using Ms. Excel, based on GITR 2015 data

The next four graphics represent the subindexes, together with their pillars:

Figure 2. The four subindexes together with their pillars





Source: made by the authors using Ms. Excel, based on GITS 2015 data

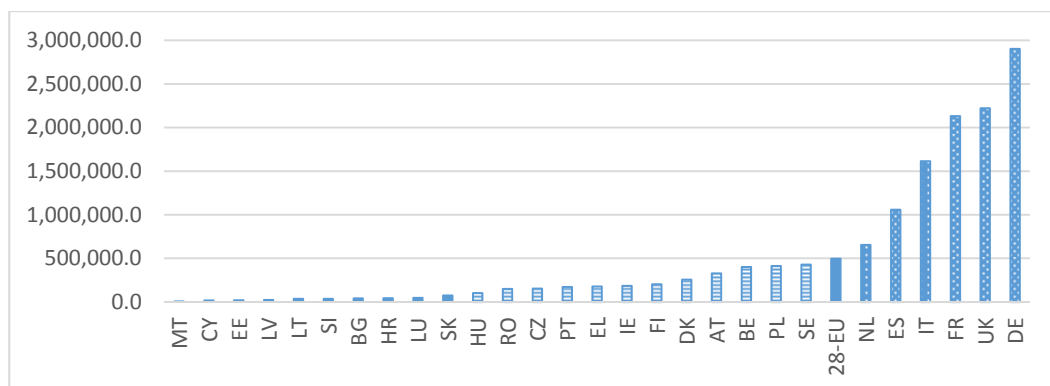
2.2. The GDP-per Capita indicator

The GDP-per Capita indicator has been made available on-line by EuroStat [5]. The most recent data, published in June 2015 includes the 28 EU Member States, along with 3 EFTA Member States (Iceland, Norway, and Switzerland), 5 EU candidate countries (Montenegro, the former Yugoslav Republic of Macedonia, Albania, Serbia and Turkey) and one potential candidate (Bosnia and Herzegovina). For most of them, data is available from 2005.

In our paper we'll refer the 28 EU Member States. In order to perform the “model” of an economy, the 28 countries were grouped into three classes:

- Advanced economies: Netherlands, Spain, Italy, France, United Kingdom, and Germany. Their GDP-per capita values are above the EU-28 average.
- Developing economies: Slovak Republic, Hungary, Romania, Czech Republic, Portugal, Greece, Iceland, Finland, Denmark, Austria, Belgium, Poland, and Sweden. Their GDP-per capita values are between 10%-100% from the EU-28 average.
- Modest economies: Malta, Cyprus, Estonia, Latvia, Lithuania, Slovenia, Bulgaria, Croatia, and Luxembourg. Their GDP-per capita values are less then 10% from EU-28 average.

Figure 3. The 28 EU Member States grouped into three classes: advanced, developing and modest economies according to the GDP-per Capita indicator



Source: made by the authors using Ms. Excel, based on EUROSTAT data

3. CLASSIFICATION TREE FOR THE THREE “ECONOMIC-DEVELOPED-MODELS” OF EU COUNTRIES

In this section, we'll define a classification tree able to describe the characteristics of the “economic-developed-models”. The classes are the three groups of the 28 EU countries: advanced economies, developing economies and modest economies, as defined in the previous section in accordance with their economic performances.

The classification tree uses as attributes the ten pillars composing the Network Readiness Indicator, numbered from 1 to 10. The tree was built using the Classification and Regression Tree (CART) method. The algorithm was implemented into an original application using C#.NET and Visual Studio 10.0 IDE which will be presented somewhere else.

Every attribute (internal node) has values in a specific range. In a similar method to the one used in section 2.2, the 28 Member States were divided into three groups using the attribute's average:

- Group 1 – modest values - less than 0.9% from EU-28 average;
- Group 2 – moderate values - between 0.9% and 1.15% from EU-28 average;
- Group 3 – high values - greater than 1.15% from EU-28 average.

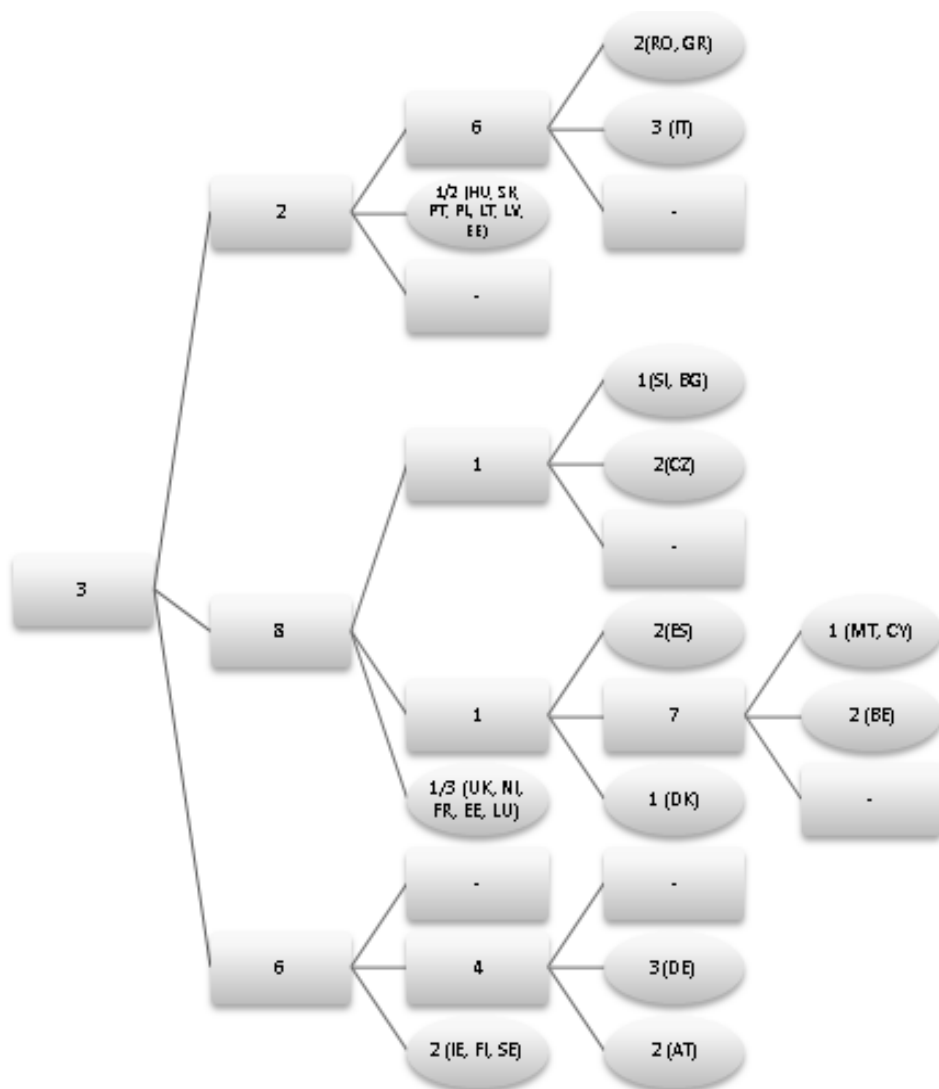
For instance, the Political and regulatory environment pillar has values between 3.2-5.8. The computed EU-28 average is 4.446. The two thresholders are 4.001 and 5.113. There are 9 countries in the first group, having the Political and regulatory environment values less than 0.9% from EU-28 average: Slovenia, Bulgaria, Croatia, Slovak Republic, Romania, Greece, Poland, Spain, and Italy (we considered the economic development order).

12 countries are in the second group, having the Political and regulatory environment values greater than 0.9% from EU-28 average, but lower than 1.15% EU-28 average: Malta, Cyprus, Estonia, Latvia, Lithuania, Hungary, Czech Republic, Portugal, Iceland, Belgium, France, and United Kingdom.

The most performant group is characterized by the highest values, greater than 1.15% from EU-28 average. There are seven Member States: Luxembourg, Finland, Denmark, Austria, Sweden, Netherlands, and Germany.

So, every attribute has three possibilities corresponding to the three groups: modest values, moderate values and high values. In the tree below, the upper branch correspond to the lowest values (group 1), the middle branch correspond to the moderate values (group 2) and the lower branch to the highest values (group 3). If there is no object corresponding to an alternative a minus (-) sign was used. 1 / 2 means that the node contains objects from both 1 and 2 classes.

Figure 4. Classification tree for modelling the relationship between IT&C and the economic development for the EU-28 Member States



Source: own model

Legend:

Key questions (round nodes):	Classes (square nodes)
1. Political and regulatory environment	1. modest economies
2. Business and innovation environment	2. developing economies
3. Infrastructure and digital content	3. advanced economies
4. Affordability	
5. Skills	
6. Individual usage	
7. Business usage	
8. Government usage	
9. Economic impacts	
10. Social Impacts	

The CART method gives us that the most significant attribute for this classification is the third one:

Infrastructure and digital content. For the countries having modest values for the Infrastructure and digital content pillar the second pillar in significance is Business and innovation environment. For countries with moderate values for Infrastructure and digital content, the second most important pillar is Government usage. The most developed in Infrastructure and digital content have, on the second significance scale, the Individual usage pillar.

Traversing in top-down manner the classification tree, we depict the template of the three “economic-developed-models” of the 28 EU Member States: advanced economies, developing economies and modest economies.

Four of the six developed economies are characterised by the next template: Medium values in Infrastructure and digital content and Medium and High values in Government usage. We count here: Netherlands, France, United Kingdom, and Spain.

Another two developed economies follow individual paths: Germany has High values in Infrastructure and digital content and Medium values in Affordability and Individual usage, while Italy has the description: Small values in Infrastructure and digital content and Medium in Individual usage.

The developing economies follow different paths: some are characterized by Modest values in Infrastructure and digital content and Modest and Medium values

in Business and innovation environment. This includes: Hungary, Slovak Republic, Portugal, Romania, Greece, and Poland.

The second category is characterised by Medium values in Infrastructure and digital content and Modest and Medium values in Government usage. We include here the following countries: Czech Republic, Belgium, and Denmark.

High values in Infrastructure and digital content and in Individual usage characterise Iceland, Finland, Sweden, and Austria.

Estonia and Luxemburg, two modest developed countries, prove a very good practice in IT&C, they follow the description of most of the developed countries: Medium values in Infrastructure and digital content and High values in Government usage.

Also Latvia, Lithuania and Croatia tend to follow the more developed economies, they share the same characteristics in IT&C consumption as almost half of the developing countries Slovak Republic, Hungary, Romania, Portugal, Greece, and Poland.

Malta, Cyprus, Slovenia, and Bulgaria have a medium level in IT&C: they feature Medium values in Infrastructure and digital content and Modest and Medium values in Political and regulatory environment.

6. CONCLUSIONS and future work

The objectives of this paper were met: a classification tree able to describe the model of a European Member State was built. The CART method have been used on a training set composed of the 28 EU States cases. We considered as input the IT&C usage level, measured through its ten pillars, and as the output, the economic development level, taking into consideration three classes: modest developing and advanced economies, grouped on the basis of their GDP-Per Capita indicator. The result could be a handy tool, easy to understand and explain for specialists and not only, useful in defining the template for the development level of a country.

A major disadvantage of the CART method is its instability. In a future work we'll investigate how sensitive is the classification tree on the variations of the thresholds used for splitting the attributes' ranges of values (the ten pillars of NRI).

Selected references

- [1] Ben Miller, Robert D. Atkinson, Raising European Productivity Growth Through ICT, The Information Technology and Innovation Foundation, Washington, 2014

- [2] Crisan Daniela Alexandra, Data Analyses of the Connection Between IT&C and the Economic Performance in European Union. Study Case: Romania, Romanian Economic and Business Review – Vol. 10, No. 2, pp. 130-139, Ed. ProUniversitaria, 2015
- [3] Crisan Daniela Alexandra , Stanescu E. D., Usta A., “The relationship between IT&C and the economic development for the EU-28 member state”, Romanian Economic and Business Review (REBE), Vol 10 No 3, 2015 , pp., Ed. ProUniversitaria, 2015;
- [4] European Commission, Pillar I: Digital Single Market – Digital Agenda for Europe, Enterprise and Industry, available on-line: <http://ec.europa.eu/digital-agenda/en/our-goals/pillar-i-digital-single-market>; “Pillar VII: ICT-Enabled Benefits for EU Society – Digital Agenda for Europe (accessed 09.10.2015)
- [5] Eurostat, available on-line: <http://ec.europa.eu/eurostat/> (accessed 15.10.2015)
- [6] Martin Feinberg, Damir Tokic, ITC investment, GDP and stock market values in Asia-Pacific NIC and developing countries, Journal of the Asia Pacific Economy, Volume 9, Issue 1, 2010
- [7] Oxford Economics, Capturing the ICT Dividend: Using Technology to Drive Productivity and Growth in the EU (Oxford Economics / AT&T), available on-line: http://www.corp.att.com/bemoreproductive/docs/capturing_the_ict_dividend.pdf (accessed 20.10.2015)
- [8] The World Bank & ITU, The little data book on Information and Technology, 2013, available on-line: data.worldbank.org/sites/default/files/wdi-2014-book.pdf (accessed 05.10.2015)
- [9] World Economic Forum (2015), “The Global Information Technology Report 2014 - ICTs for Inclusive Growth”, Soumitra Dutta, Thierry Geiger, Bruno Lanvin, Ed. SRO-Kundig, Geneva
- [10] World Economic Forum (2014), “The Global Competitiveness Report 2014–2015”, K. Schwab, Ed Geneva

Figure 5. The structure of the Networked Readiness Index 2015
(made by authors using Ms. Excel and GITS 2015 information)

