CLUSTER ANALYSIS OF INTERNET USAGE IN EUROPEAN COUNTRIES

SIMICEVIC, V[anja]; ZOROJA, J[ovana] & PEJIC BACH, M[irjana]

Abstract: Internet provides a new communication medium and is becoming very influential for many people regardless of age, gender or education. This paper reviews and discusses Internet issues and reports the findings of a cluster analysis. Indicators such as: age, gender, education, Internet usage, on-line shopping and e-government are used to measure the differences between European countries. Data from Eurostat were used for this paper. There were data for 31 European countries which were divided into four clusters. Considering different characteristics of Internet users and different forms of Internet usage the most developed countries are placed in the first cluster. The second cluster consists of developed western-European countries, and the third consists of the southern European countries. The final group consists of the eastern European and candidate countries to the European Union. Conclusion is that there is still a significant difference between European countries according usage of Internet

Keywords: Internet usages, gender, age, education, shop on-line, e-government, European countries, cluster analysis

1. INTRODUCTION

In the last two decades Internet has affected almost every aspect of people's life [4]. Through Internet people communicate, gather and exchange information [14]. Internet as a communication tool and as a source of information [17] has a strong impact on social, economic and cultural structures [15].

The number of people who use Internet continues to increase worldwide. According to 2002 estimates, more than 400 million people use the Internet [3] and eight years later there are 1,9 billion Internet users [21] which means that 28,7% of population uses Internet. Apart from that, differences between different continents and regions are significant. North America has the highest number of users (77,4%), next are Oceania and Australia (61,3%), and then Europe (58,4%). Other regions fall behind, so the share of users is around one third in Latin America (34,5%) and in the Middle East (29,8%). The lowest number of users is in Asia (21,5%) and Africa (10,9%). [2].

There are differences in Internet usage among the European Union countries, candidate countries and countries that are not the European Union countries or its candidates. According to Eurostat, 45% of households in Croatia in 2008 had Internet access whereas the European Union average was 60%. Nevertheless Croatia is ahead of some European Union countries: Bulgaria (25%), Greece (31%), Italy (42%), Ciprus (43%) and Romania (30%) [20]. Estonia and Slovenia have made the biggest progress compared with other 10 countries

that have entered European Union in 2004 and they are very close to EU 15 average [18].

Internet influences the economic growth and productivity of the country [16, 9] because it offers a broad range of possibilities for business [10]. Internet usage mostly depends on the level of development of a country and its investment opportunities in infrastructure. The use of Internet is increasing in many developing countries [10] but there are still significant differences between countries [16] because unequal access to and use of Internet is connected with social and economic state in the particular country. In most developed countries like Canada and USA there are 68,2% of the population that are connected to the Internet and this percentage is less than 2,7% in Africa [2].

Internet usage, particulary in fields of trade and government, is important for social and economic development [19], since Internet has a positive influence on country's productivity [9]. Other factors that also influence the level of Internet usage are the level of education, income, gender and age [7], as well as users' attitude and behaviour and cultural differences [17].

According to data from Eurostat, usage of Internet among three groups of countries (European Union countries, European Union candidate countries and countries that are not the European Union countries or its candidates) has been analysed. The analysis considers the share of population that uses the internet according to different charachteristics (age, gender, education) and reasons for usage (e-government and on-line shopping) in order to determine if there are differences of Internet usage among countries. Focus of the paper is whether the collected data can be successfully used to divide European countries into clusters so that countries with similar charachteristics belong to one particular cluster, and so that the differences between clusters are clear. This can be especially important for managers and policy makers who want to encourage and increase usage of Internet, especially in countries where there is a low rate of Internet usage among population.

The main purpose of this paper is to analyze and determine are there differences in Internet usage among the European Union countries, candidate countries and countries that are not the European Union countries or its candidates on the basis of available date on the use of Internet, different demographic characteristics and reasons for usage. According to collected data and the

goal of the paper we define hypothesis related to the results of this research:

Hypothesis 1: Internet usage, according to different demographic characteristics and reasons for usage, leads to differences among the European Union countries, candidate countries and countries that are not the European Union countries or its candidates.

The purpose of the paper will be reached by using the cluster analysis, a multivariable technique of statistical analysis. The goal of the cluster analysis is to determine relatively homogeneous groups of objects or variables. Data should be grouped so that the perceptions within a cluster are similar, but that clusters are different among themselves. In a cluster analysis a group belonging of objects is not known, nor is the final number of groups [8]. The most important factor in a cluster analysis is to determine the number of clusters. If the number of clusters is high, the differences between particular clusters will be small, and the cluster highly specialised. It is likely that in this case the clusters will be too similar. Consequently, it will not be simple to interpret the results or generalise. If the number of clusters is low, there will be more differences between individual clusters, and an analysis of this kind cannot offer new or valuable information. It is therefore important to know that there are not an exact number of clusters [6].

The paper is organized as follows. At the beginning there is an introduction part about Internet, its influence on economic and society and statistical data about Internet usage in the world. The goal of the paper is stated as well as hypothesis and cluster analysis which will be used in research part of the paper. Research methodology and empirical results are clearly stated in the next two sections. The fourth part of the paper discusses the results of the research and concludes the paper.

2. RESEARCH METHODOLOGY AND SAMPLE CHARACTERISTICS

In order to reach the main purpose of the paper an empirical research was conducted using the European Statistics Database – Eurostat [21].

Data are from section Information Society Statistics and were collected from the household surveys conducted by the Central Bureaus of Statistics (CBS). Data that are used in this analysis are: Internet usage, education, gender, age, e-government and shop on-line and they relate to seven year period, from 2003-2010. However, cluster analysis has been conducted using data from 2010.

There are 31 countries tha are used for the analysis. Countries are divided into three groups. In the first group there are European Union countries: Belgium, Denmark, Germany, Ireland, Greece, Spain, France, Italy, Luxembourg, Netherlands, Austria, Portugal, Finland, Sweden, United Kingdom, Czech Republic, Estonia, Latvia, Lithuania, Hungary, Poland, Slovenia, Slovakia, Bulgaria, Romania, Malta and Cyprus. Second group consist of European Union candidate countries: Croatia and Former Yugoslavian Republic Macedonia. Norwey

and Iceland are in the third group, they are neither European Union, nor European Union candidate countries.

In this research we use demographic data that are classified by gender (all users, male and female), age (aged 16-74) and by educational level (low, medium and high) and we also use data about reasons of using Internet. Percentage of individuals in a European countries that use Internet, that use Internet to made their on-line purchase and that use Internet to interact with public authorities at least once within the last 12 months was investigated. After all data were analysed, cluster analysis was conducted and European countries were classified into four clusters named "I", "II", "III" and "IV".

During the cluster analysis a method of average values was used, and the following are its characteristics: (1) number of groups defined in advance (k), (2) data observation in a multidimensional space that has variable values like axis. To determine the number of clusters the following formula was used [13]:

$$ k \approx \sqrt{\frac{n}{2}} \approx \sqrt{\frac{31}{2}} \approx 3{,}8 \approx 4 \; clusters $$

Using the standard form of the statistical package Statistica, we carried out Generalized EM cluster analysis, the method where number of clusters has to be determined in advance [1].

3. EMPIRICAL RESULTS

Before conducting the cluster analysis it was necessary to adjust and prepare the data. First, data by gender, age and educational level were classified. After that percentage of individuals in a particular country that use Internet, that use Internet to made their on-line purchase and that use Internet to interact with public authorities was investigated. All data were compared to determine the differences between European Union countries, candidate countries and countries that are not the European Union countries or its candidates.

Several conclusions were reached on the basis of the studied data and different analysis. Firstly, there are differences in demographic characteristics of Internet users, as well as differences in ways of Internet usage in different European countries.

According to the data on Internet users there are significant differences between the European Union countries and Croatia. Comparing the European Union countries before and after 2004 when new members joined the European Union, there are no significant differences. In the EU 15 countries the percentage of regular Internet users has an ascending trend in the period from 2003 to 2010; in 2008 this was 60%, whereas in the EU 27 countries it was 56%. In Croatia this procentage was 39%.

If we analyse the regular users according to age, the biggest gap is noticed by users between 55 and 74 years.

The percentage of elder regular users in the European Union in 2007 is 25%, whereas in Croatia it is only 7%.

Comparing the users according to gender, the number of male users is higher than that of female users both in Croatia and the European Union countries. In 2010 there are 52% of male users in the age of 16 to 74 in Croatia, whereas in the EU 15 countries this procentage equals 69%, and in the EU 27 countries it equals 66%. The above data is lower for the female Internet users. In Croatia this was 43%, and in the European Union countries, the percentage was 64%.

Considering the level of education of Internet users in Croatia in 2010 there was 22% of population without or with low formal education in the age from 16 to 74 that used the Internet regularly, whereas this percentage in the countries of the European Union was the double (42% and 39%). Considering regular Internet users in all European Union countries and Croatia it can be seen that differences are the smallest comparing highly educated population. In 2010 the procentage of highly educated Internet users is equal in all European Union countries (89%), whereas in Croatia it is 79%.

On-line shopping in Croatia is almost statistically insignificant. In 2007 and 2008 only 5% of population shopped on-line for private purpose, whereas in 2010 it was 7%. In the European Union coutries there is an ascending trend during the period of the 7 years. In 2010 in the EU 15 countries 35% of population shopped on-line for private purposes, whereas in the EU 27 countries were 29% individuals who shopped on-line.

The use of Internet for public services is not very popular in Croatia, in 2007 only 10% of population used it, and in 2008 this number decreased to 9,3%. Comparing the European Union countries there is a higher percentage of the population that uses the Internet for public services in the EU 15 countries than in the EU 27 countries. Furthermore, in 2008 there was a decrease in all mentioned countries in using the Internet for public services.

This analysis was conducted with the purpose to clasify the European Union countries, the candidate countries and countries that are not the European Union countries or its candidates by means of cluster analysis. On the basis of collected and analysed data and k-means algorithm, six cluster analyses were conducted. Table 2 shows 31 countries defined by clusters which they belong to according to different demographic characteristics (gender, Internet access, age, education) and reasons for using Internet (on-line shopping and e-government). Table 2 shows that all countries can be divided into four groups considering the ranking of individual clusters. Clusters are marked with letters „I", „II", „III" and „IV".

Cluster marked 'I' represents the best cluster and consists of most developed countries. Denmark, Luxembourg, Netherland, Finland, Sweden and two countries that are not European Union countries or its candidates, Norway and Iceland, are in the cluster 'I'. The population of these countries uses the Internet the most.

The other group is a part of the cluster 'II'. Certain countries in cluster 'II' also belong to a third or a fourth

cluster depending on certain criteria. Within the 'II' cluster there are Belgium, Germany, Estonia, France, Austria and the UK. Estonia is one of the ten countries that joined the European Union in 2004, and has progressed in Internet usage since then the most.

Countries in the third group belong to the cluster 'III'. These are mostly countries that joined the European Union in 2004 (The Czech Republik, Ireland, Spain, Latvia, Lithuania, Hungary, Poland, Malta, Slovenia and Slovakia). These are mostly considered developed eastern European countries.

Country	Demographic characteristics and reasons for using Internet					
	Gender	Internet usage	Age	Education	E-government	Shop on-line
EU countries						
Belgium	II	III	IV	II	II	II
Bulgaria	III	IV	III	III	IV	IV
Czech Republic	IV	II	IV	III	III	III
Denmark	II	I	I	II	I	I
Germany	I	II	III	III	III	I
Estonia	II	ND	IV	III	II	III
Ireland	III	IV	IV	II	II	II
Greece	III	IV	II	IV	IV	IV
Spain	IV	IV	III	III	III	III
France	II	II	III	II	II	II
Italy	III	IV	II	IV	IV	IV
Cyprus	IV	III	III	IV	III	III
Latvia	IV	II	IV	III	III	IV
Lithuania	III	III	IV	III	III	IV
Luxembourg	II	I	I	I	I	I
Hungary	III	I	IV	II	IV	IV
Malta	III	IV	III	II	III	II
Netherlands	II	I	I	I	I	II
Austria	I	II	III	II	III	II
Poland	III	ND	III	IV	IV	III
Portugal	III	IV	III	III	III	IV
Romania	III	IV	III	IV	IV	IV
Slovenia	IV	IV	IV	III	II	III
Slovakia	II	II	IV	ND	III	II
Finland	I	I	I	I	II	II
Sweden	I	II	I	I	II	I
United Kingdom	I	II	I	II	II	I
Candidate countries						
Croatia	III	II	III	III	IV	IV
FYR Macedonia	IV	III	III	IV	IV	IV
Other countries						
Iceland	I	II	I	I	I	II
Norwey	I	I	I	II	I	I

Tab. 1. Cluster of each country according to demographic .
characteristics and reasons for using Internet, ND=No Data

In this cluster marked 'IV' are the Mediterranean countries (Greece, Italy, Cyprus and Portugal), the countries that joined the European Union in 2007 (Bulgaria and Romania) and the candidate countries (Croatia and FYR Macedonia). Bulgaria and Romania are the least developed countries of the European Union, and this reflects in the conducted analysis on Internet usage. The economies of Croatia and the FYR Macedonia that are candidate countries also fall behind compared with other European Union countries. The lower level of development in both of the above groups of countries reflected also on the Internet usage in these countries. The Mediterranean countries are affected by the economic crisis that influenced both the entrepreneurship activities on-line as well as other personal activities.

The conclusion of this research is that the use of Internet according to different demographic characteristics and reasons for its use leads to differences

between the European Union countries, candidate countries and countries that are not the European Union countries or its candidates. The conclusion proves the hypothesis of the paper.

4. CONCLUSION

Internet is a prerequisite for the development of today's economies, since a society that has no highly developed skills in information technology cannot be competitive. The research shows differences in Internet usage among the European countries. The greatest differences are between the candidate countries and the countries of the European Union. There are slight differences among the EU 27 countries, but the EU 15 countries has a clear competitive advantage. According to the research conducted by Cilan [5] the countries of EU 15 have the maximum level of Information Society. Research results made by Cilan [5] show that Portugal and Greece have the same Information Society level as the countries that become member states of European Union in 2004, and which correlate with cluster "D". In all European countries the number of male population that uses the internet is higher than the number of female users. Shiu and Dawson [14] made a research on Internet usage in four countries (UK, Germany, Japan, Taiwan) and also reached a conclusion by which there is a male predominance in the Internet usage. Likewise, in all European countries there are more Internet users among the highly educated population, and the differences in usage are smallest between the highly educated (EU 88%, Croatia 78%).

According to Orviska and Hudson [12] Internet access increases significantly with the education. On-line shopping and usage of public services is increasing during the seven years in the countries of the European Union, whereas in Croatia this number is almost insignificant. Based on the research made by Shiu and Dawson [14] young people shop on-line less than older people. In order to use public services on-line, governments have to invest a lot of money in the ICT infrastructure [11] and because of that e-government is better developed in the European Union countries.

The research objective of this paper was to examine the Internet usage among European countries. According to data and cluster analysis the hypothesis was conformed – inhabitants in developed European Union countries use Internet intensively to do shopping on-line and for public services regardless of age, educational level or gender while candidate countries and less developed European Union countries fall behind. Differences among European countries in Internet usage can be explained by factors such as age, gender, cultural structures and economic development. It might be concluded that Internet is more used and accepted in developed European countries. Less developed European Union and candidate countries should take actions to develop policies that will encourage inhabitants and entrepreneurs to use Internet more and that will reduce Internet usage gap.

There are limitations of the paper. First, there were no data for all variables and for all European countries. Second, it is not specify is it Internet used from households, universitites or enterprises. Future research should include more data and the reserach would provide more information. Further studies may explore Internet usage in other countries in order to compare them.

5. REFERENCES

[1] Aczel, D. A., Sounderpandian, J. (2006). Complete Business Statistics (6th Edition). McGraw Hill, New York.

[2] Akman, I., Mishra, A. (2010). Gender, age and income differences in internet usage among employees in organizations. Computers in Human Behavior, Vol. 26, pp. 482–490

[3] Amiel, T., Lee Sargent, S. (2004). Individual differences in Internet usage motives. Computers in Human Behavior, Vol. 20, pp. 711–726

[4] Blackman, C. (2004). Stumbling along or grave new world? Towards Europe´s information society. Foresight, Vol. 6, No. 4, pp. 261-270

[5] Cilan, C. A., Bolat, B. A., Coskun, E. (2009). Analyzing digital divide within and between member candidate countries of European Union. Government Information Quarterly, Vol. 26, pp. 98-105

[6] Dimitriadou, E., Dolničar, S., Wingessel, A. (2002). An examination of indexes for determining the number of clusters in binary data sets. Psychometrika. Vol. 67, No. 1, 137-160

[7] Fisher, Y., Bendas-Jacob, O. (2006). Measuring Internet usage: The Israeli case. International Journal of Human-Computer Studies, Vol. 64, pp. 984-997

[8] Hair, J. F., Jr., Black, W. C., Babin, B. J., Anderson, R. E. (2009). Multivariate Data Analysis (7th Edition). Englewood Cliffs, NJ: Prentice - Hall

[9] Ho, S., Kauffman, R., Liang, T. (2007). A growth theory perspective on B2C e-commerce growth in Europe: An exploratory study. Electronic Commerce Research and Application, Vol. 6, pp. 237-259

[10] Indjikian, R., Siegel, D. S. (2005). The Impact of Investment in IT on Economic Performance: Implications for Developing Countries. World Development, Vol. 33, No. 5, pp. 681-700

[11] Kaliannan, M., Awang, H., Raman, M. (2010). Public-Private Partnerships for E–Government Services: Lessons from Malaysia. International Journal of Institutions and Economics, Vol. 2, No. 2, pp. 207-220

[12] Orviska, M., Hudson, J. (2009). Dividing or uniting Europe? Internet usage in the EU. Information Economics and Policy, Vol. 21, No. 4, pp. 279-290

[13] Parr Rud, O. (2003). Data mining cookbook. Wiley, New York; Chichester

[14] Shiu, E. C. C., Dawson, J. A. (2004). Comparing the impacts of Internet technology and national culture on online usage and purchase from a four-country perspective. Journal of Retailing and Consumer Services, Vol. 11, pp. 385-394

[15] Turlea, G., Olle Sanz, E., Ciupagea, C. (2011). Romania and integration into the European model of e-inclusion: from access to the use of advanced communication services. Romanian Journal of European Affairs, Vol. 11, No. 1, pp. 43-62

[16] Wallsten, S. (2003). Regulation and Internet Use in Developing Countries. Development Research Group, The World Bank, pp. 1-29

[17] Zhang, Y. (2005). Age, gender and Internet attitudes among employees in the business world. Computers in Human Behavior, Vol 21, pp. 1-10

[18] Zon, H. (2005). The variety of information society development paths in central Europe. AI and Society, Vol. 19, No. 3, pp. 309-326

[19] Zott, C., Amit, R., Donlevy, J. (2000). Strategies for Value Creation in E-Commerce: Best Practice in Europe. European Management Journal, Vol. 18, No. 5, pp. 463-475

[20] ec.europa.eu/eurostat (2012). European Statistics database, Accessed on: 2012-04-28

[21] http://www.internetworldstats.com/stats.htm, (2012). Internet World Stats, Accessed on: 2012-03-2