CLASSIFICATION METHODS IN ANALYSIS OF INTERNET USE AND E-COMMERCE ACTIVITIES IN EC COUNTRIES

Zlatko J Kovacic, The Open Polytechnic of New Zealand, School of Information Science and Humanities, Private Bag 31914, Lower Hutt, New Zealand (e-mail: KovZla@topnz.ac.nz)

Keywords: Cluster Analysis, Discriminant Analysis, Internet Diffusion, E-Commerce, EC Countries

Abstract
The Internet has become a new vehicle for bringing global social and economic changes to the world. At the same time Internet diffusion is happening at the fastest rate compared to growth of any other technical innovation in the recent history. Why is the Internet diffusion so important to study? Knowledge-based industries built on new information technologies in developed countries generate more than a half of economic production. Therefore for each country it is very important to understand factors underlying Internet diffusion. Differences in the level of Internet diffusion exist not just between developing and developed countries but also within developed countries. These differences are not caused only by the country's economic power but also because of its telecommunications infrastructure and its penetration, educational level, cultural factors, and so on.

The theoretical framework for Internet diffusion analysis is based on Rogers (1995) diffusion of innovations theory modified for application in telecommunications. This paper explores the Internet diffusion adopting quantitative approach based on a wide range of Internet indicators. The set of indicators used in this paper cover most of the six dimensions of Internet presence in the country mentioned in Wolcott et al. (2001), i.e., connectivity infrastructure, pervasiveness, sophistication of use, organisational infrastructure, sectoral absorption and geographical dispersion.

What makes our approach to Internet diffusion analysis different from other research papers that adopt quantitative approach is the use of multivariate methods of classification such as cluster analysis and discriminant analysis. We have applied hierarchical cluster analysis based on different measures of proximity, testing the number of clusters and checked cluster validity.

Using the analytical framework and data described above we consider the following questions:
(1) How many different patterns/clusters of Internet diffusion could be find among EC countries?
(2) Which Internet indicators contribute the most to separation of these clusters?

The three-cluster solution we have obtained seem quite robust across different clustering methods used. Based on three-cluster solution we have applied discriminant analysis to data set to find which indicators have the most discriminating power in the set of Internet indicators. The result of discriminant analysis shows that differences between EC countries are not caused by the differences in their e-commerce activities or Internet use, but because of the differences in the telecommunication, computing and Internet infrastructure they have.

References
Rogers E.M. (1995) 'Diffusion of Innovations: Modifications of a Model for Telecommunications', in *Die Diffusion von Innovationen in der Telekommunikation*, Stotzer M.-W. and Mahler A. (eds), Berlin: Springer Verlag, pp. 25-38.

Wolcott P., Press L., McHenry W., Goodman S. and Foster W. (2001) 'Framework for Assessing the Global Diffusion of the Internet', *Journal of the Association for Information Systems*, Vol. 2, Article 6, November 2001.
(http://mosaic.unomaha.edu/2001_GDI_Framework_files/article_files/Wolcott_Main.htm)