Exploring the Distance among Websites: A Multiple Correspondence Analysis

Myoung-Jin Lee
Kookmin University, Department of Sociology
861-Chongneung-dong, Seongbuk-gu, 136-702
Seoul, Korea
mjlee@kookmin.ac.kr

1. INTRODUCTION

Despite the facts that research on the internet has greatly expanded with the support of public grants and that general sociological interest in the phenomena is on increase, research on its social implication and micro aspects has not taken a sociological turn. Recently sociologists have been concerned with man in one's role as 'homo faber' on the internet, i.e., creator and user of websites. They have focused on the person-website relations as well as inter-individual patterns of relations. The most prominent perspective for characterizing the internet user and websites is based on the cultural sociologists' idea of person-object relations. They maintain that objects, apart from their use value and exchange value, have 'sign value'. The objects often represent users' specific aspects of personality, lifestyle or conviction (Bourdieu, 1998; Csikszentmihaly and Rochberg-Halton, 1981). People's choices in clothing, leisure, forms of entertainment, etc., are related because their choices reflect and express a certain taste or habitus,

2. THEORETICAL FRAMEWORK

This study concerns on the relations of the internet websites and their users. Our primary objective is to assess the influence of user's characteristics in the choices and formation of websites on The internet. According to the previous theory, men and women pay attention to different objects in the same environment. The theory assumes that how one reacts to objects is often ascribed by gender. Males devote more time to 'objects of action' than to 'object of contemplation' (Csikszentmihaly and Rochberg-Halton, 1981). Age, education, occupation are considered as additional variables in the person and object relations. For example, the young often cherish objects of action, and the old are more concerned with objects of contemplation.

Three questions regarding the influence are addressed: (1) To what extent does the user's characteristics have an influence on the choice of websites? (2) Which characteristic has more impact on the choice of website. (3) How is websites classified?

3. DATA AND METHOD

Multiple correspondence analysis and cluster analysis with log tracking data from 9,051 the internet users in Korea is performed (Intermetrix, 2001). We use 5 variables: top 50 websites, gender, age, education, and occupation. The variable has 50, 2, 4, and 5 levels, respectively. Top 50

websites can be considered a dependent variable whereas the others being independent variables. In order to combine the pairwise relationship between one dependent variable and four independent variables, 50×15 matrix of table is analyzed (See, Greenacre, 1981; Hand, 1981).

4. RESULTS

The results of this study reveal the influence of user's characteristics on the choice of websites, showing a certain level of polarization among top 50 websites. It is obvious that choice of websites is to some degree constrained by gender, age, education, and occupation. The websites with the same contents are often separated from one another. 'iloveschool' is separated from 'damoim' although the two websites have the same purpose to search friends and alumni. The similar result is found among search engine and newspaper websites.

The highest association is found between age groups and the choice of websites, showing that 'damoim' and 'iloveschool' associated with the younger age groups (10s and 20s) and that 'chosun' and 'edunet4u' associated with the higher age groups (30s and 40s). The second highest association is found between education and the choice of websites. Student category is separated from the other occupational categories, associated with music-related and community-related websites (ex, 'soribada', 'sayclub', etc.). The association between gender groups and the choice of websites are relatively low, with females more associate with 'hananet' and males more associated with 'joins' and 'empas'.

The result from the hierarchical clustering analysis shows how top 50 websites are classified. Four different clusters can be separated. Most websites (36 websites) belong to the first cluster. This cluster includes various forms of websites (i.e., newspaper, broadcasting, search-engine, etc). The second cluster includes music-related and community-related websites. The last clusters are 'iloveschool' and 'damoim'. Although these two websites offer similar service, they are clearly separated in terms of users' characteristics, especially in terms of age.

5. DISCUSSION

This investigation of the websites shows a certain degree of spatial dispersion in terms of users' characteristics. The main contrast occurs on the basis of age and education.

REFERENCE

Bourdieu, P. 1998. Distinction: A Social Critique of the Judgment of Taste. Cambridge: Harvard University Press.

Csikszentmihaly, M. and E. Rochberg-Halton. 1981. The Meaning of Things: Domestic Symbols and the Self. Cambridge: Cambridge University Press.

Greenacre, Michael and Leslie G. Underhill, 1981. "Scaling a Data Matrix in a Low-Dimensional Euclidean Space." Douglas M. Hawkins, eds. Topics in an Applied Multivariate Analysis, 183-266. Cambridge: Cambridge University Press.

Hand, D. J. 1981. Discrimination and Classification. New York: John Wiley & Sons.

Intermetrix. 2001. Seminar on The internet Economy. Seoul. [In Korean]