6-2011

ICT Development Index and the Digital Divide

Anteneh Ayanso
Brock University, aayanso@brocku.ca

Danny I. Cho
Brock University, dcho@brocku.ca

Kaveepan Lertwachara
California Polytechnic State University, klertwac@calpoly.edu

ICT Development Index and the Digital Divide

Anteneh Ayanso, Danny I. Cho
Brock University
aayanso@brocku.ca, dcho@brocku.ca

Kaveepan Lertwachara
California Polytechnic State University
klertwac@calpoly.edu

Abstract
The disparity in access to information and communications technologies (ICTs), often referred to as "digital divide", has received significant attention among policy makers and academics around the world. In particular, the need for continuous monitoring of the ICT adoption rates by communities in different economies has been one of the top priorities in various forums. In order to address this need, the International Telecommunication Union (ITU), a United Nation agency, proposed a single, comprehensive ICT Development Index (IDI) in 2009. IDI is designed to consolidate useful information from previous measurement indices. One of its main uses, as illustrated by ITU, is to measure the magnitude of the digital divide and how it is evolving over time. The objective of this paper is to supplement the methodologies applied in ITU's project as well as in previous research in measuring and analyzing the digital divide. Using the data set and the statistical processes applied in the construction of IDI, we employ a cluster-based methodology to analyze the global and regional digital divide and provide additional insights.

Keywords
Digital Divide, ICT Profiles, ICT Development Index, Cluster Analysis.

1. Introduction
The advent of information and communications technologies (ICTs) has offered millions of people around the world access to the unprecedented wealth of information and knowledge. However, while most communities in developed countries are able to reap benefits from these valuable resources, the number of ICT users in developing countries has grown at a relatively slow rate (ITU, 2009). This disparity, often referred to as "digital divide", remains at the forefront of the discussions among policy makers and academics around the world (Bagchi, 2005; DiMaggio et al, 2001; Deichmann et al., 2006; ITU, 2009). Essential among these discussions is the need for continuous monitoring of the ICT adoption rates by communities in different countries. A unified effort to monitor the ICT adoption rate has allowed international agencies to create a benchmark and historical trends in order to measure the progress of ICT adoption, especially in the developing countries. The most widely-cited work on ICT adoption monitoring tool is a set of recommendations set forth at the 2003 World Summit on the Information Society (WSIS) in Geneva, Switzerland (WSIS, 2003). Since then, a number of ICT composite indices have been developed and used by numerous organizations in the international community. In order to improve the monitoring efforts regarding the global digital divide, the International Telecommunication Union (ITU), a United Nation agency, commissioned the

development of a single, comprehensive ICT Development Index (IDI) that is designed to consolidate useful information from previous measurement indices (ITU, 2009). One of the key uses of this composite index that is also demonstrated in ITU's report (2009) is to capture the magnitude of the digital divide and how it is evolving over time.

Past studies have also examined the digital divide among different geographical regions and economies around the world. For example, Wong (2002) studied the digital divide among Asian countries. Buys et al. (2009) studied the disparity in ICT access and adoption among the sub-Saharan African countries. More recently, Ayanso et al. (2010) analyzed the regional and global digital divides by profiling the ICT infrastructure of 192 UN member states. However, the past studies utilized various ICT measurements and methodologies, making direct comparisons and benchmarking extremely difficult. As a result, not only did the development of IDI mark a momentous step in the measurement of the information society, but it also presented a tremendous opportunity for policy makers and scholars to comprehensively assess the digital divide in various settings. The purpose of this research is to supplement the methodology and analysis applied in previous research and provide additional insights into the digital divide using the more coherent and comprehensive data used in the construction of IDI. The next section provides the historical and developmental details of the IDI.

2. The ICT Development Index (IDI)

Developed by the ITU, the ICT Development Index (IDI) represents a single ICT measurement that is designed to capture "the level of advancement of information and communication technologies" in 154 economies worldwide and compares the progress made by these countries between the years 2002 and 2007. The main objective of the index is to provide policy makers with a useful tool to benchmark and assess their information society developments and to monitor progress concerning the digital divide (ITU, 2009). The index was designed primarily to merge previous ICT measurements into a single index. Accordingly, the IDI index consolidates previous measurement indices (i.e., the Digital Access Index (DAI), a measurement of access and usage of ICTs; the Digital Opportunity Index (DOI), a tool used to measure potential benefits of access to ICTs; and the ICT Opportunity Index (ICT-OI) which is designed to track the global digital divide among countries with similar income levels).

In addition, the IDI incorporated a conceptual framework, based on a basic three-stage information society model (readiness-use-impact as shown in Figure 1) and the use of principal components analysis (PCA) to eliminate indicators that have less influence on the index calculation. Envisioned and designed as a composite index, the IDI consists of three main components (or three stages): ICT access, use, and skill indicators. These three components eventually allow scholars and policy makers to make important assessment on each country's progress toward becoming an information society and the impact that ICTs have on the society. Figure 2 outlines the specific variables that make up each of the three component of the IDI as well as the weighing factor of each component in the overall IDI score.

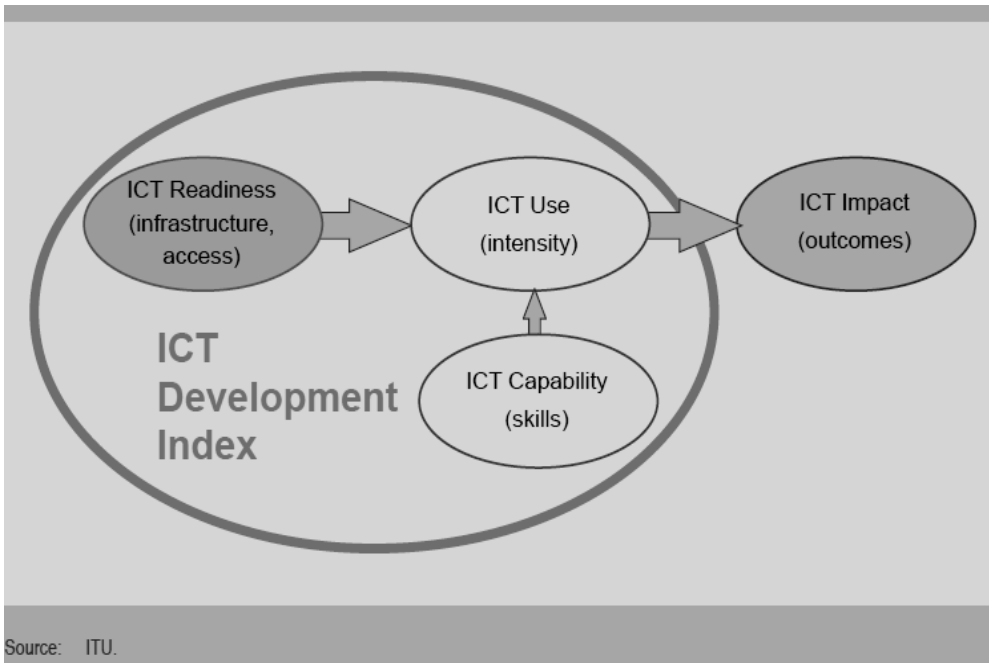

Figure 1: Three stages in the evolution towards an information society
Source: (ITU, 2009)

3. Analyzing the Digital Divide: Methodology

One of the main applications of IDI is to determine the magnitude of the digital divide and monitor how it evolved over time (ITU, 2009). ITU used a four-step approach for their digital divide analysis using IDI scores of 154 economies for the year 2002 and 2007. The 154 economies included in the study account for 97.6% of the total world population in 2007 (ITU, 2009). In this section, we review the method used by ITU in the digital divide analysis and present our cluster-based methodology.

3.1 The ITU Methodology and Analysis

Following the methodology developed by Orbicom (2003), ITU first grouped the 154 economies included in the study into four groups according to each country's IDI score. The two groups with above average IDI score are labeled "high" and "upper", while the other two groups with below average scores are labeled "medium" and "low". Secondly, ITU then computed the average IDI score for each of the four groups. The third step in ITU's analysis involved normalizing the IDI scores from the 2002 data set relative to the average IDI score in 2007. Finally, the normalized data allowed ITU to make a direct comparison and analyze the progress

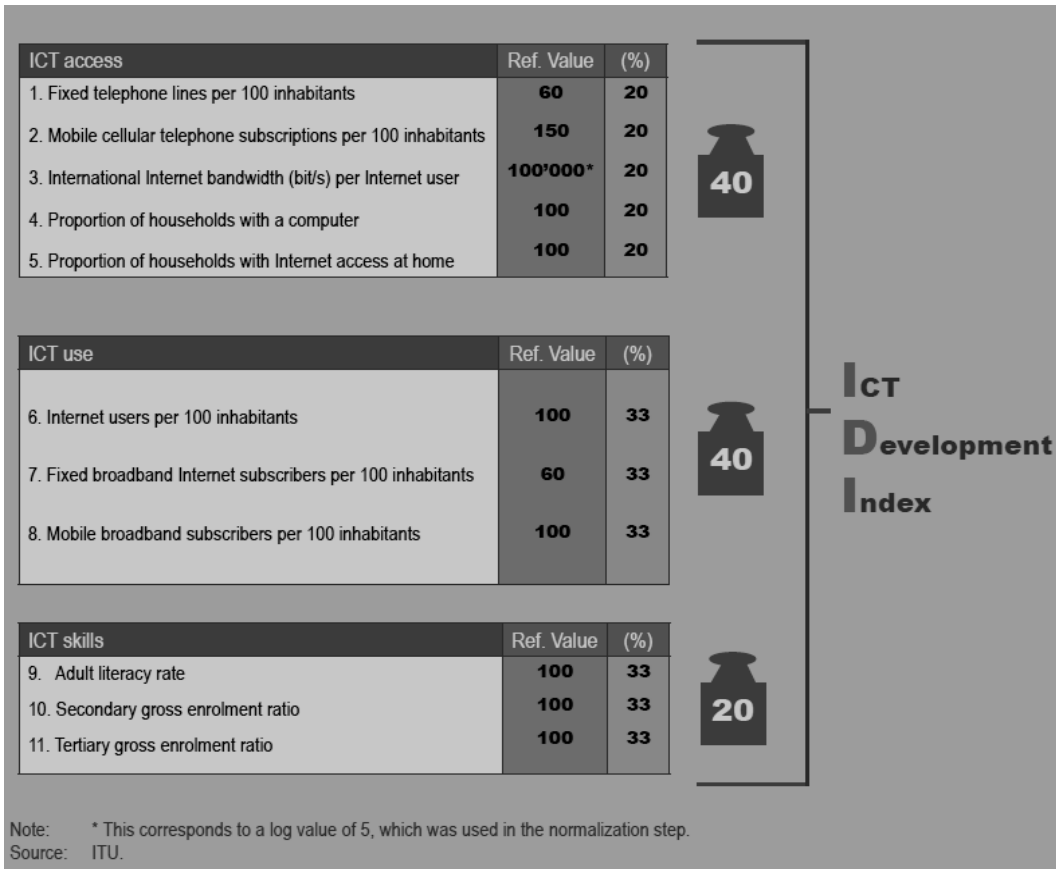

Figure 2: IDI: Indicators, Reference Values for Normalization, and Weighting
Source: (ITU, 2009)

that each country has made between 2002 and 2007. More specifically, ITU used the following steps for the digital divide analysis (ITU, 2009):

Step 1: Grouping the countries according to their index values
The overall 2007 average (i.e., 3.40) was used to divide the 154 economies into four groups, with two groups lying above the average and two groups lying below the average. The average value was placed after the 66th country, which resulted in 66 countries above the average and 88 below. The 66 countries were then classified into two equal groups (High and Upper). The High group has IDI values above 5.29 and the Upper group has IDI values between 3.41 and 5.25. The remaining 88 countries located below the average were also divided into two equal groups (Medium and Low).The Medium group has IDI values between 2.05 and 4.34 and the Low group has IDI values between 0.82 and 2.03.

Step 2: Computing the groups' average IDI values
The average IDI value of each group was computed to conduct further analysis, such as in showing the magnitude and the evolution of the digital divide between groups, and for determining whether the divide is shrinking or widening. The High group has an average IDI score of 4.8 and 6.4 for 2002 and 2007, respectively. Corresponding averages are 2.9 and 4.1 for the upper group; 1.9 and 2.7 for the Medium group; and 1.0 and 1.3 for the Low group.

Step 3: Normalizing the average IDI values
To emphasize the relative nature of the digital divide concept, the overall 2007 average IDI score was used as the reference value. Group averages were transformed into their corresponding normalized values using the 2007 average IDI value for all countries.

Step 4: Computing changes in the digital divide
The normalized IDI scores were then used to illustrate the magnitude of the digital divide between the groups. Changes in the digital divide were computed by subtracting the magnitude of the 2007 digital divide from the 2002 corresponding value. The direction (sign) of the computed values shows the evolution of the digital divide: a negative value indicates a closing divide between the two groups, and a positive value indicates a widening divide.

3.2 Cluster-Based Analysis of the Digital Divide

While the method used by ITU is simple to follow, it may not capture the natural grouping of the countries as well as the movement of the countries from one group to another based on the underlying structure of the data. This is because the four groups were identified using simple heuristics based on the 2007 IDI scores and the overall average value. Our specific objective in this paper is to supplement the analysis and discussion provided by ITU using a cluster-based methodology. We use the data set and the statistical processes applied in the construction of IDI and propose a cluster-based methodology to analyze the global and regional digital divide. Our focus will be on the clustering methodology and the comparative discussion of the results from our analyses, rather than on developing a new list of variables or proposing new normalization and weighting steps. Accordingly, we use the final data used by ITU in the construction of IDI without any modifications.

Cluster analysis is commonly used to organize observed data into meaningful structures or taxonomies that are not known in advance. For example, cluster analysis has been used for a wide variety of problems, such as identifying market segments (Chaturvedi et al., 1997; Dolnicar, 2003), profiling mobile Internet adopters (Okazaki, 2006), and more relevant to this research, profiling access to ICT and utilization using primary ICT indicators which include PCs, Internet, Telephone lines, Mobile phones, and broadband (Ayanso et al., 2010); and profiling the levels of digital development in European Union (Vicente Cuervo & Lo´pez Mene´ndez, 2006). In this research, we use the *Two-Step* clustering algorithm (Zhang et al., 1997; Chiu et al., 2001) to analyze the global and regional digital divide using the IDI values of the 154 economies (ITU, 2009).

The *Two-Step* clustering algorithm is a scalable algorithm designed to handle large data sets with a number of advantages over other traditional clustering methods such as k-means and hierarchical clustering methods. The algorithm is efficient in forming clusters for large data sets by reducing the number of original records into pre-clusters. In the second step, the algorithm merges the pre-clusters using hierarchical clustering, thus avoiding the need for distance computations between all pairs of initial data records. More importantly, the method has an auto-clustering feature which selects the number of clusters based on statistical criterion such as the Bayesian Information Criterion (BIC) or the Akaike Information Criterion (AIC). The algorithm

also allows user-specified number of clusters if user chooses to determine the number of clusters a priori.

In this paper we conduct the following set of cluster analyses and discuss the results:

- Cluster analysis to examine the global digital divide using IDI values for all 154 economies. We use this analysis to supplement the methodology and results presented in ITU's report.
- Cluster analysis to examine the regional digital divide using IDI values within each geographic region (Africa, Americas, Asia, Europe, and Oceania).

In each case above, we use the auto-clustering feature where the number of clusters is automatically determined using the default criterion, Bayesian Information Criterion (BIC). The BIC is a likelihood criterion penalized by the number of parameters in the model as a measure of model complexity (Chiu et al., 2001). BIC is also considered more appropriate than AIC when the goal is exploration rather than prediction (Kuha, 2004).

3.2.1 Analyzing the Global Divide using Auto-clustering

Our first analysis involves all the 154 economies included in ITU and their IDI scores for the 2002 and 2007. For both years, the auto-clustering method identified two clusters which we label here as ICT "leaders" and "followers" for consistency with previous research (Ayanso et al., 2010). Table 1 and Table 2 present the summary statistics for the clusters obtained for the 2002 and 2007 IDI scores, respectively.

Clusters	Number of Countries	Average IDI Score	Minimum IDI Score	Maximum IDI Score
Leaders	37	4.70	3.34	6.05
Followers	117	1.77	0.51	3.3

Table 1: Summary Statistics of Global Clusters for 2002 IDI Scores

Clusters	Number of Countries	Average IDI Score	Minimum IDI Score	Maximum IDI Score
Leaders	45	6.00	4.37	7.5
Followers	109	2.33	0.82	4.16

Table 2: Summary Statistics of Global Clusters for 2007 IDI Scores

Table 1 and Table 2 show that the summary profiles for both groups have increased from the year 2002 to 2007. In addition, the comparison of the cluster formation for 2002 and 2007 shows that a total of eight economies from the followers group in 2002 were later identified in the leaders group in 2007. All the remaining economies remained in their respective group in 2007. The eight economies include four from Asia (i.e., *Bahrain, Brunei Darussalam, Qatar, and United Arab Emirates*) and four from Europe (i.e., *Bulgaria, Croatia, Latvia,* and *Lithuania*).

The rest of the Asian economies that formed the leaders group for both years include *Cyprus, Hong Kong (China), Israel, Japan, Korea (Rep.), Macao (China), Singapore, and Taiwan (China)*. The rest of the European economies that formed the leaders group for both years include *Austria, Belgium*, *Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, and United Kingdom*.

The movement of the eight Asian and European economies from the followers group in 2002 to the leaders group in 2007 indicates a positive ICT development progress of these nations over the years. The increase in the average IDI score for each group also indicates that nations in both groups are making progress, although the ICT gap is not shrinking significantly.

On the other hand, all the countries in Africa are found in the ICT followers group for both 2002 and 2007. This indicates that while the African nations are making some progress, their relative positions in the global profiling remains the same. The groupings of the African nations in the ITU report show that none of the nations belong to the High and Upper groups. However, the ITU methodology classifies some of the African nations in the Medium group and the vast majority of them in the Low group. Unlike the simple heuristic used in the ITU methodology to create these groups, the cluster-based methodology clearly shows that this division between the two groups of the African nations is not really significant. This division, however, becomes significant in the regional context as our regional analysis shows with additional insights in the following section. Similarly, all other countries in the Americas are found in the ICT followers group for both years, with the exception of the *United States* and *Canada*. In Oceania, *Australia* and *New Zealand* are found in the leaders group for both years.

3.2.2 Analyzing the Regional Divide using Auto-Clustering

To analyze the digital divide within each geographic region (i.e., Africa, Americas, Asia, Europe, and Oceania), we conducted cluster analysis using the five regions as individual cases. The cluster analysis for 2002 identified three groups for Africa, and two groups for the Americas, Asia, and Europe. The Oceania region has only four economies represented in the ITU report. As a result, the cluster analysis identified only one cluster for both years. Table 3 and Table 4 present the summary statistics of the clusters obtained from the 2002 and 2007 data, respectively.

In 2002, the cluster results for Africa shows three economies, namely *Libya, Mauritius,* and *South Afric*a, forming the leaders group with the average IDI score of 2.21 (See Table 3). The medium group (followers-1) included ten countries, namely, *Algeria, Botswana, Cape Verde, Egypt, Gabon, Morocco, Namibia, Swaziland, Tunisia, and Zimbabwe* with an average IDI score of 1.57. The rest 29 countries formed a third cluster (followers-2) with an average IDI score of 0.93. The cluster analysis for the year 2007, however, identified two clusters with 11 countries in the leaders group with an average IDI score of 2.5, and 31 countries in the followers group with an average IDI score of 1.26 (see Table 4). These 11 countries include the top three countries and eight of the ten countries identified in the leaders and followers-1 cluster of the 2002 results, respectively. The other two countries, namely, *Swaziland* and *Zimbabwe*, have moved to the followers group in the 2007 cluster results.

Region	Cluster	Number of Countries	Average IDI Score	Minimum IDI Score	Maximum IDI Score
Africa	Leaders	3	2.21	2.08	2.45
Africa	Followers-1	10	1.57	1.29	1.86
Africa	Followers-2	29	0.93	0.51	1.21
Americas	Leaders	2	5.29	5.25	5.33
Americas	Followers	22	2.16	1.05	3.06
Asia	Leaders	11	4.33	3.27	5.83
Asia	Followers	36	1.84	0.89	2.84
Europe	Leaders	13	5.27	4.47	6.05
Europe	Followers	24	3.28	1.92	4.38
Oceania	-	4	3.22	1.05	5.02

Table 3: Summary Statistics of Regional Clusters for 2002 IDI Scores

Region	Cluster	Number of Countries	Average IDI Score	Minimum IDI Score	Maximum IDI Score
Africa	Leaders	11	2.50	1.92	3.45
Africa	Followers	31	1.26	0.82	1.73
Americas	Leaders	2	6.39	6.34	6.44
Americas	Followers	22	2.99	1.27	4.12
Asia	Leaders	12	5.77	4.44	7.26
Asia	Followers	35	2.54	1.23	3.79
Europe	Leaders	19	6.56	5.54	7.5
Europe	Followers	18	4.37	2.73	5.47
Oceania	-	4	4.22	1.14	6.58

Table 4: Summary Statistics of Regional Clusters for 2007 IDI Scores

The cluster results for the Americas show *United States* and *Canada* forming the leaders group with an average IDI score of 5.29 (See Table 3). The rest of the countries formed the followers group with an average IDI score of 2.16. This result for the Americas is the same as the one found in the global analysis and the cluster formation also remained the same for the year 2007.

For Asia, ten counties, namely, *Bahrain, Brunei Darussalam, Cyprus, Hong Kong (China), Israel, Japan, Korea (Rep.), Macao (China), Singapore, and Taiwan (China),* and *United Arab Emirates* formed the leaders group in 2002 with an average IDI score of 4.33 (see Table 3). In 2007, the movement of *Qatar* to the leaders group is the only change observed. The most notable change happened for Europe. In 2002, there were more countries in the followers group than in

the leaders group (24 versus 13). The leaders included *Austria, Belgium, Denmark, Finland, Germany, Iceland, Luxembourg, Netherlands, Norway, Slovenia, Sweden, Switzerland*, and *United Kingdom*. However, in 2007, there were more leaders than followers (19 versus 18). In addition to the leaders found in 2002, six countries from the followers group were identified with the leaders group in 2007. These include *Estonia, France, Ireland, Italy, Malta,* and *Spain*.

Overall, the comparison of the average IDI scores for each cluster group between 2002 and 2007 shows a positive progress for each region.

4. Conclusions

Addressing the digital divide requires effective measurements (ITU, 2009). While simplicity is always a priority to help countries measure and monitor their progress over time easily, simple heuristics sometimes may not capture the reality and allow comparative analysis. Although there are different approaches to grouping countries, clustering possesses both the advantage of simplicity as well as effectiveness in capturing the underlying structure of the data and placing similar countries together and dissimilar countries apart. It provides us with nominal scale attributes that explain the membership of each country in each group (cluster) based on its IDI position. ICT Profiles often exhibit groups of economies with relatively high local density, representing a group of countries that are either advanced or behind in terms of ICT penetration and usage. Accordingly, clustering can be effectively used to find naturally occurring groups that correspond to the modality of the ICT data. Thus, the cluster-based methodology presented in this paper supplements the methodology used by ITU and allows further analysis using the individual components of the composite IDI index. Our future research will examine the global as well as regional digital divide using the three sub-indices (ICT Access, ICT Use, and ICT Skills) of the IDI in order to provide additional insight in terms of the three stages in the evolution towards an information society which represents the conceptual framework employed by ITU to describe the process countries are go through.

References

Ayanso, A., D. I. Cho, K. Lertwachara (2010). "The Digital Divide: Global and Regional ICT Leaders and Followers," *Information Technology for Development*, 16(4), pp. 304-319.

Bagchi, K. (2005). "Factors Contributing to Digital Divide," *Journal of Global, Information Technology Management*, 8(3), pp. 47-65.

Buys, P., S. Dasgupta, T.S. Thomas, and D. Wheeler (2009). "Determinants of a Digital Divide in Sub-Saharan Africa: A Spatial Econometric Analysis of Cell Phone Coverage," *World Development*, 37(9), pp. 1494-1505.

Chaturvedi, A., J.D. Carroll, P.E. Green, and J.A. Rotondo (1997). "A Feature-Based Approach to Market Segmentation via Overlapping K-Centroids Clustering," *Journal of Marketing Research*, 34(3), pp. 370-377.

Chiu, T., D.P. Fang, J. Chen, Y. Wang, and A. Jeris (2001). A Robust and Scalable Clustering Algorithm for Mixed Type Attributes in Large Database Environment, *Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 263–268.

Deichmann, J.I., A. Eshghi, D. Haughton, M. Masnghetti, S. Sayek, and H. Topi (2006) "Exploring Break-Points and Interaction Effects Among Predictors of the International Digital Divide," *Journal of Global Information Technology Management*, 9(4), pp. 47-71.

DiMaggio, P., E. Hargittai, R. Neuman, and J. Robinson (2001). "Social Implications of the Internet," *Annual Review of Sociology*. 27, pp. 307-336.

Dolnicar, S. (2003). "Using Cluster Analysis for Market Segmentation - Typical Misconceptions, Established Methodological Weaknesses and Some Recommendations for Improvement," *Australasian Journal of Market Research*, 11(2), pp. 5-12.

International Telecommunication Union (ITU) (2009). "*Measuring the Information Society –The ICT Development Index*," Place des Nations, CH-1211 Geneva Switzerland, ISBN 92-61-12831-9, available at http://www.itu.int/ITU-D/ict/publications/idi/2009/index.html.

Kuha, J. (2004). "AIC and BIC: Comparison of Assumptions and Performance," *Sociological Methods & Research*, 33(2), pp. 188-229.

Okazaki, S. (2006). "What Do We Know About Mobile Internet Adopters? A Cluster Analysis," *Information & Management*, 43(2), pp. 127–141.

Orbicom (2003). "*Monitoring the Digital Divide…and Beyond*, George Sciadas (Ed.), Montreal (Quebec), Canada, available at http://www.orbicom.uqam.ca.

The World Summit on Information Society (WSIS). (2003). Plan of Action, available at http://www.itu.int/wsis/docs/geneva/official/poa.html

Vicente Cuervo, M. and A. Lopez Menendez (2006). "A Multivariate Framework for the Analysis of the Digital Divide: Evidence for the European Union-15," *Information and Management*, 43(6), pp. 756-766.

Wong, P.-K. (2002). "ICT Production and Diffusion in Asia: Digital Dividends or Digital Divide?," *Information Economics and Policy*, 14(2), pp. 167-187.

Zhang, T., R. Ramakrishnon, and M. Livny (1997). "BIRCH: A New Data Clustering Algorithm and its Applications," *Data Mining and Knowledge Discovery*, 1(2), pp. 141-182.