Spatial Distribution of Intra-Urban Digital Divide in Istanbul

OZBAKIR B.Aysegul
Department of City and Regional Planning
Yildiz Technical University
Istanbul, Turkey
aozbakir@yildiz.edu.tr

KURT KARABULUT Gunes
Department of Electronic and Communication Engineering
Yildiz Technical University
Istanbul, Turkey
gkurt@itu.edu.tr

***Abstract*—**The subject of unequal access to and use of the information technologies started to focus on the concept of the so-called "digital divide" has started to attract attention since late 90s. Before that time, more general concepts such as information inequality, information gap or knowledge gap and computer or media literacy were used. Preliminary digital divide research led to the equivalency of media or technology access with physical access. Currently, the majority of the literature still focuses on physical access. However, since the year 2002, an increasing number of researchers suggest to go 'beyond access', to reframe the overly technical concept of the digital divide and to pay more attention to social, psychological and cultural backgrounds. The goal of this paper is to provide a new methodology to understand the spatial variation of Information and Communication Technology (ICT) use in Istanbul Metropolitan Area using Geographic Information Systems (GIS) and K-means Clustering Algorithm. The research is carried out through a questionnaire employed to 1140 individuals in 10 districts of Istanbul in Turkey with a wide range of socio-economic compositions.

Keywords*— *Digital Divide, ICT Use, GIS, K-means Clustering Algorithm.

I. INTRODUCTION

With the rapid development of information and communication technologies in recent years, a significant gap has occurred between the rural and urban areas in terms of accessibility and the use of such technologies. The subject of unequal access to and use of the information technologies started to focus on the concept of the so-called "digital divide" has started to attract attention since late 90s both in Turkey and the world in academia, private and public sectors. Before that time, more general concepts such as information inequality, information gap or knowledge gap and computer or media literacy were used.

In order to decrease the digital divide, many strategies and action plans are proposed by countries around the world. Being one of these, Digital Opportunity Index (DOI) introduced by International Telecommunication Union (ITU) of United Nations has been frequently used. Although this index is among major techniques that measures the divide, it is also argued that the parameters used to construct the index are same for all countries. However, most of the recent research indicates that *one size does not fit all* due to the geographic, social, economical and cultural differences among countries. For this reason, when ranked according to the results of this index, countries or regions might reveal misleading performance results.

Preliminary digital divide research that started around late 90's in Turkey has led to the equivalency of media or technology access with physical access. Currently, the majority of this research still focuses on physical access. However, since the year 2002, an increasing number of researchers suggest to go 'beyond access', to reframe the overly technical concept of the digital divide and to pay more attention to economic, social, psychological and cultural backgrounds. It can be observed that analyses covering geographical location, socio-economic status and local regulations have been incomplete about the digital divide concept that is rapidly increasing and covering new dynamics in parallel to changing technologies.

The goals of this research are first, to make a cluster based analysis in 10 districts of Istanbul with a wide range of socio-economic compositions to understand different types of digital divide. Second, to map the spatial distribution of these types using Geographic Information Systems (GIS). It is a well known fact that the gap between the children and their families increase due to the "ability to learn technology". For this reason, the paper will make significant contributions to understand different characteristics of digital "natives" and "immigrants".

II. DATA AND METHODOLOGY

This paper attempts to introduce a clustering algorithm supported by GIS to understand different types of ICT user behavior at metropolitan area level. Figure 1 illustrates the methodology used in this research. The objective of the introduced methodology is to identify the spatial relationships between the digital divide categories among different neighborhoods (mahalle[1]). To this end, a detailed questionnaire (with 48 items) has been applied to 1140 individuals with different socio-economic profiles. Within this frame, 10 districts with 31 mahalles having

[1] Mahalle: is the smallest administrative unit in Turkey.

ISSN: 2229-6913 (Print), ISSN: 2320-0332 (Online) © 2016 MID Publications in association with Vidya Publications.

different geographic locations and economic status levels are selected in Istanbul Metropolitan Area (Figure 2).

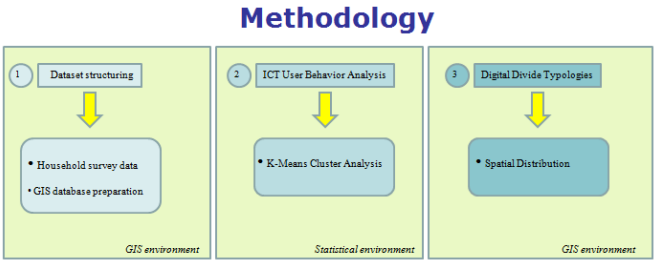

Figure 1: Methodology of the paper

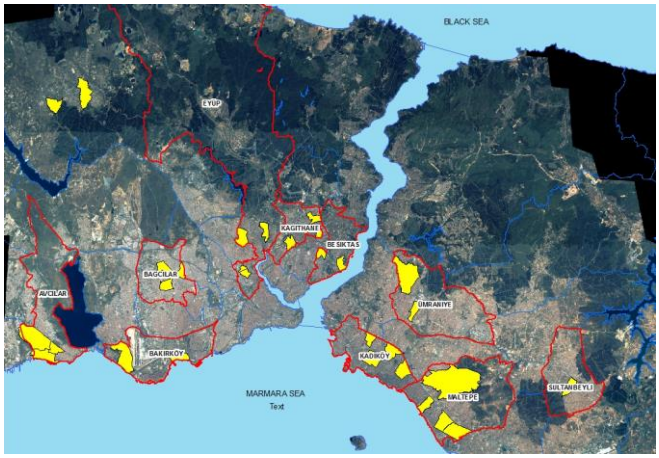

Figure 2: Selected neighborhoods (mahalles)

Table 1 and Table 2 show the ICT service types that are asked to the responders and main categories of questions

Table 1: ICT service types that are asked to the responders.

İnternet	Telephone	Mobile Phone
3G	Home Line (PSTN)	3G
Wi-fi hot spot	Mobile:	2G
xDSL (+wi-fi)	2G	Wi-fi hot spot
Fiber	3G	
WiMAX	VoIP Application	

Table 2: Questionnaire items for Digital Divide Analysis.

Questions for each of the households
Demographic structure: - gender, - age, - place of birth, - mother tongue, - literacy,
Economic structure - occupation status - monthly income
ICT ownership and use: - number of mobile phones, - network type, - invoice type, - computer usage and frequency (hour/day) - internet usage and frequency (hour/day) - place of internet accessibility - mobile phone usage - mobile service - reasons of computer use - reasons of internet use - mobile phone applications
ICT education: - date of learning computer skills (month/year) - date of learning internet skills (month/year) - place of learning computer skills - place of learning internet skills
Expenditure for ICT services: - monthly expenditure for cellular phone

III. CLUSTERING ANALYSIS

Data clustering analyses are empirical steps of classifying various subjects into different clusters with respect to the properties those subjects have. It is also called unsupervised classification, and also described as a method of creating groups of clusters, in such a way that objects in one cluster are very similar and objects in different clusters are quite distinct (Gan *et al*., 2007).Clustering analyses and chosen methodology of each step may also differ from one to another. In this paper a general clustering approach, K-means, will be proposed for the questionnaire data set that contains a mixture of numeric data types.

K–means clustering algorithm is a clustering technique that iteratively calculates cluster centroids with respect to that cluster's new member. In k-means clustering technique, there are 4 main steps to proceed as also given by Tou and Rafael (1974).

Step1: Chose K initial cluster centers $z_1, z_2, \ldots, z_k$,

Step2: At the k^{th} iterative step, distribute the samples (subjects) among the K cluster domains, using below relation:

$$S_i^{(t)} = \left\{ x_p : \left\| x_p - z_i^{(t)} \right\| \leq \left\| x_p - z_j^{(t)} \right\| \forall\, 1 \leq j \leq k \right\} \quad (1)$$

where, x_p is the new member which will be assigned to a clustered group S_i, according to the distance to i^{th} cluster's center and j^{th} cluster's center that is measured on t^{th} step (iteratively).

Step3: From the results of step two, compute new cluster centroids with respect to the below relation:

$$z_j^{(t+1)} = \frac{1}{\left| S_i^{(t)} \right|} \sum_{x_j \in S_i^{(t)}} x_j \quad (2)$$

ISSN: 2229-6913 (Print), ISSN: 2320-0332 (Online) © 2016 MID Publications in association with Vidya Publications.

where the variables are the same explained above.

Step4: If all samples belong to one cluster, stop the procedure, otherwise go to second step.

In K-Means at each step new cluster centroids have been calculated. This illustrates a dynamical approach at creating clusters, where centroids are not calculated until the last step of hierarchical clustering.

Number of clusters: There is no general suggestion on determining exact amount of clusters. According to Clatworthy (2005): "The most straightforward method of deciding on the number of clusters is to examine both the agglomeration schedule and the dendogram. An inconsistent increase in the dissimilarity measure indicates that the clusters joined at that stage were quite distinct, and that the clustering researcher decides what constitutes an inconsistently large increase in the dissimilarity measure."

Besides that, according to the goals of this research, determining number of clusters a-priori seem, it will provide more necessary and meaningful output then expecting a proper amount of clusters by empirical methods. Since the extraction of level of digital gap(s) in the data set is important, three different groups are determined that state digital literacy of their cluster members. These groups are: 1) Digital literates, 2) Digital immigrants and 3) Digital illiterates. Table 3 summarizes the general characteristics of each class.

Table 3: General Characteristics of ICT User Behavior Clusters (*not all items are listed*)

Features	Digital Literates	Digital Immigrants	Digital Illiterates
Use mobile phone	Yes	Yes	No
Determine the model and type of the mobile phone by him/herself	Yes	Yes	No
Use 3G	Yes	No	No
#of apps/reasons for mobile phone use	n > 5	1< n < 4	n <= 1
Connects internet through mobile phone	Yes	No	No
Use computers	Yes	Yes	No
#of apps/reasons for computers use	n > 6	1< n < 5	n <= 1
Use internet	Yes	Yes	No
#of reasons for internet use	n > 6	1< n < 5	n <= 1

IV. SPATIAL DISTRIBUTION OF DIGITAL DIVIDE CLUSTERS

Once the digital divide gaps are clustered, the next step is to explore the spatial distribution of each category among the metropolitan area of Istanbul. Since the power of GIS lies in both its capacity in analyzing and displaying spatial data and bringing data from different sources; a database for the questionnaire respondents has also been structured in this research.

To this end, a basemap of the city that contains buildings, streets and mahalle boundaries is prepared using ArcGIS 10. For each respondent, a unique ID is created and linked to the buildings where the questionnaire is employed. With this spatial database, 1140 individuals are mapped on the metropolitan area. Once the clustering analysis has been completed, the next step is to explore the agglomerations of each class by mahalle level around Istanbul. To do that, geostatistical tools of GIS has been used and three maps for each digital divide category has been created that indicates the number of individuals attached to each class (Figures 3, 4 and 5).

According to the Figures 3, 4 and 5, it is observed that the majority of the "digital literate" groups are spatially located along the southern parts and by the waterfront regions of the city where higher education and income levels are observed (according to the census data). The only outlier within this group exists on the northern part of European side of Istanbul where again gated communities exist with higher education and income levels.

The spatial distribution of "digital immigrants" portrays a different picture than the "literate" group. Since they are mostly composed of individuals who attempt to use ICTs in limited ways, their spatial distribution also reflects a similar pattern around the metropolitan area. The majority of this group locates surrounding the "literate" category; not the waterfront sides but the adjacent mahalles. Furthermore, the "illiterates", in general, spatially locate in outer regions of the city where lower education and income profiles exist.

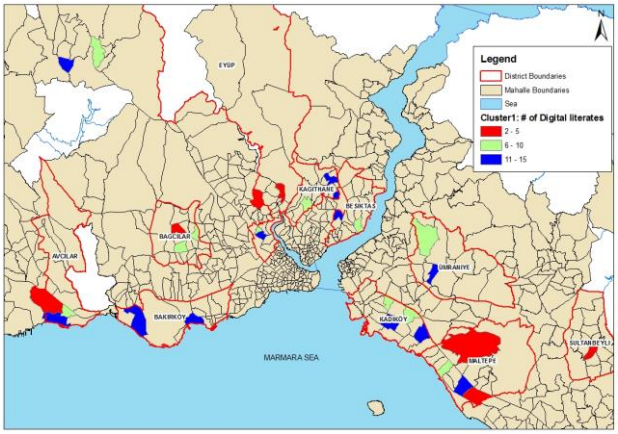

Figure 3: Spatial distribution of "Digital Literates" by mahalle level in Istanbul.

ISSN: 2229-6913 (Print), ISSN: 2320-0332 (Online)

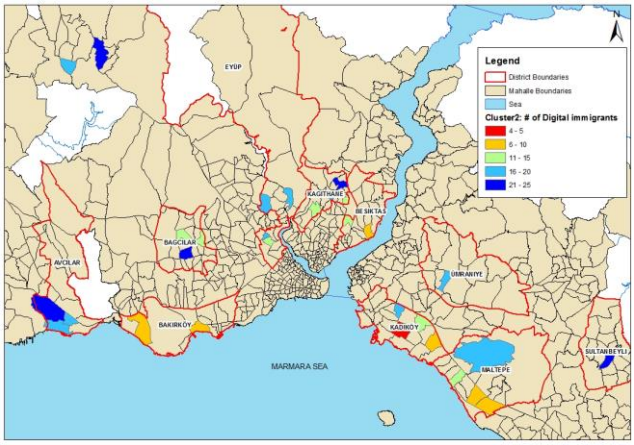

Figure 4: Spatial distribution of "Digital Immigrants" by mahalle level in Istanbul

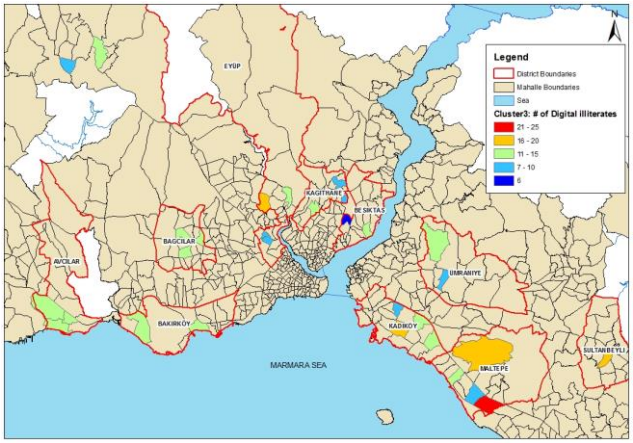

Figure 5: Spatial distribution of "Digital Illiterates" by mahalle level in Istanbul

V. CONCLUSION

The concept of "digital divide" has been an international debate since the last few decades and many countries attempt to define strategical tools and policies to reduce the digital gap between nations, regions or cities. United Nations (UN) has organized the "Millenium 2000 Summit" where actions against the divide were discussed. Similarly, the Okanawa Summit of G8 Countries in 2000 have emphasized that "the gap between rich-knowledge and poor-knowledge societies is also a major determinant of the achievements in economic welfare, education, health, literacy levels. Furthermore, it is also mentioned that the digital gap causes long-lasting effects for differences among nations.

With parallel to the world's awareness on the subject of "digital inequalities", Information and Communication Technologies Authority (ICTA) of Turkey has been founded in 2000 and released its first strategic plan to decrease the digital divide in the country. Currently, the majority of the research practices in Turkey still focus on physical access. Thus, this paper attempts to raise the question of existence of different typologies of the "digital divide" at metropolitan area level. Not only "physical access" but also other socio-economic characteristics are attached to each of the ICT user behavior category. Since this paper is an initial part of a funded project, the future research aims to include more sources of data such as census data and urban data into the clustering algorithms.

ACKNOWLEDGMENT

This work is supported by Yıldız Technical University, 2011-03-02-GEP02

REFERENCES

[1] Bertot J.C., (2003), "The Multiple Dimensions of the Digital Divide: More than the Technology "haves" and "haves not", Government Information Quarterly 20(2), 185-191

[2] Clatworthy, J., Buick, D., Hankins, M., Weinman, J. and Horne, R., (2005), "The use and reporting of cluster analysis in health psychology: A review". British Journal of Health Psychology, 10 pp 329–358.

[3] Dijk V.J.A.G.M., (2006), "Digital Divide Research, Achievements and Shortcomings", Poetics, (34), pp:221-235

[4] Dolnicar V., Vehovar V. And Sicherl P., (2004), "Benchmarking Digital Divide: Definitions Used and Methods Applied", 26th International Conference Information Technology Interfaces ITI, June 2004, Cavlat, Crotia

[5] Gan, G., Chaoqun M. and Jianhong W., (2007), *Data Clustering: Theory, Algorithms, and Applications*, ASA-SIAM Series on Statistics and Applied Probability, SIAM, Philadelphia, ASA, Alexandria, VA, USA.

[6] Gunkel, D., (2003), "Second Thoughts:Toward a Critique of the Digital Divide", New Media and Society 5 (4), pp:499-522

[7] Grubesic T., (2004), "The Demographic Correlates of Broadband Access and Availability in the United States", Telematics and Informatics, (21), 335 – 358

[8] OECD, 2000, The Digital Divide: Diffusion and use of ICT's, Paris, OECD

[9] Şahiner T. and Kurt K.G., 2012, Network Demand Analysis by Clustering by Mobile User Behavior, ICECECE 2012 Conference, Tokyo, Japan

[10] Tou, J. T. and Rafael C. G. (1974), *Pattern Recognition Principles*, Addison-Wesley Publishing Company, Advanced Book Program Reading, Massachusets, USA.

[11] United Nations, "The Digital Divide Report: ICT Diffusion Index 2005", U.N. Conference on Trade and Development, New York and Geneva, 2006