Exploring the Patterns and Trends of Socio-spatial Activities of Architecture Student Community in Istanbul by Data Mining

Ahu Sokmenoglu1, Ceyhun Burak Akgul2
1Istanbul Technical University & TU Delft, 2Vistek-ISRA Vision & Bogazici University, Turkey
1(http://www.akademi.itu.edu.tr/sokmenoglu), 2(http://www.cba-research.com)
1sokmenoglu@itu.edu.tr, 2cb.akgul@gmail.com

Abstract. In this paper, we introduce the process and the outcomes of a graduate-level digital architectural design studio course aiming to explore the patterns and trends of socio-spatial preferences and activities of architecture student community in Istanbul. In the scope of this experimental studio, our aim was to propose data mining as a rigorous methodology for the analysis of socio-spatial problems. Specifically, we designed a detailed set of questionnaires on six different conceptual categories related to socio-spatial activities and personal preferences of architecture and design students. The questionnaires have been completed by 88 student subjects from Istanbul Technical University. The paper provides a descriptive analysis of the collected data from several perspectives.

Keywords. Socio-spatial analysis; data mining, digital architectural design studio.

Introduction

In this paper, we introduce the process and outcomes of a graduate-level digital architectural design studio course aiming to explore the patterns and trends of socio-spatial preferences and activities of architecture student community in Istanbul. The subject of the studio was to gain insight about activity and preference patterns and trends for architecture students by using data mining. Data mining is defined as to extract important patterns and trends from raw data (Witten and Frank, 2005). As such, when applied to discover relationships between spatial and non-spatial attributes, data mining can support the analysis of cities (Gil, Montenegro, Beirao and Duarte, 2009) and lead to novel opportunities in the formulation of new design approaches.

In the following, we begin by presenting our motivation on why we organized such a studio work. Next, we concentrate on the organizational issues and we clarify our methodology. We then focus on the outcomes of the studio by evaluating the student works. We conclude by evaluating the importance of socio-spatial perspective in contemporary urban research.
Motivation

Our main motivation behind this studio was that, as architects and urban designers, we need a deeper understanding of individual preferences and activities occurring in urban space so that we can initiate innovative urban design solutions. Cities, how we think of them, and how we can work strategically to guide their development require deeper analysis in understanding the social and spatial patterns of preferences and activities performed in every-day life. What we propose here is an analysis of the urban space use at the level of individuals by evaluating their social characteristics and preferences. No longer think of urban space as purely physical, and static; instead in the scope of this studio, one of our proposals was to take it as social and relational.

Massey’s (2005) definition underlining three important characteristics of the urban space was inspiring for us:

• Space is the product of interrelations, constituted through interactions from the immensity of the global to the intimately tiny.
• Space is the sphere of possibility of the existence of multiplicity, space is the sphere in which distinct trajectories coexist.
• Space is always under construction, always in the process of being made, never finished, never closed.

To represent the relational and living character of space requires an alternative non-Euclidean imagination of space where detailed characteristics of both places and people should be included in the picture. In fact, many of the contemporary critics of urban research underline that urban geography has not been particularly good at, or interested in, making sense of many of the smaller elements that make up a city. (Thrift & Amin, 2002) According to Thrift and Amin (2002), small elements that give urban life so much of its texture are elements such as day-to-day routines of shopping and household provisioning, the life of public places like parks, sidewalks and shopping malls, the networks of friendship and enthusiasm, and so forth.

In this studio, we are primarily interested in tracing those small intangible elements of places and people that make the space around us. Thus, we are interested in spatial data concerning the information about location, quantity, usage, orientation, dimensions, density, morphological characteristics, topological relations, continuity, connectivity, accessibility, etc. of urban entities, and social data concerning any human related information such as population, gender, cultural and ethical identities, age, education level, occupation, employment status, economical status, personal histories, personal preferences, likes/dislikes, daily activities and paths, political opinions, etc. We aimed to contribute to our knowledge of social processes occurring in urban space, particularly in Istanbul, at the level of individuals.

Within this scope, the basic assignment of the studio was the following:

• Define your questions related to socio-spatial activities of architecture student community in Istanbul.
• Define your concepts, exemplify your instances and select your attributes in terms of socio-spatial activities of architecture student community in Istanbul.

Organization of the Studio and Methodology

We defined two pedagogical goals in the scope of this studio. One was to clearly define what data mining is and to give the student a feeling about what it is capable of, as well as is not, with several examples from a broad spectrum of applications. Our second goal was to use data mining on specific socio-spatial analysis problems to be formulated by students within the context of social and spatial patterns of preferences and activities performed in every-day life. We projected two main phases towards these pedagogical goals.

In the first phase, we introduced the basics of data mining and a battery of statistical tools that can
help the exploration and analysis of urban space. We started by introducing data mining in general; then, in a series of lectures, we elaborated on its applications and related rudimentary concepts (such as instance, attribute and its types, data table, class, rule, etc.), as well as on descriptive statistics. Afterwards, we delved into the realm of sophisticated data mining techniques (Hastie, Tibshirani and Friedman, 2009) ranging from regression analysis and classification to clustering and exploratory data analysis. Several assignments was completed by using the freely available Rapid Miner software [1] in order to get more familiar with data mining processes.

In the second part of the course, we focused on our main problem, that is, the exploration of the effects of social characteristics on the activity patterns of students in the city. First, students were required to design a detailed data collection survey on daily activities, mobility paths, habits, and socio-cultural preferences exclusively of architecture and design students. Their assignment was basically to collect data from interrelated other students on places, objects, and events. Our data-analytic strategy was based on “bottom-up” data collection so that we can use first-hand information gathered from the students.

Surveys were composed of basic quantitative and qualitative questions. Students were required to define their concepts as a group, by exemplifying their instances and attributes in terms of socio-spatial activities of architecture and design student community in Istanbul. We gave them some generic questions in order to trigger the design process of the surveys. A few examples for these questions are as follows:

- How do the students socially use/experience urban space and for what purpose?
- Where and why do they meet?
- Where and what do they eat/drink?
- Where do they have fun?
- Where are their favorite open-spaces?
- What do they feel in specific locations?
- In which proportion do they spend money/time on socio-cultural activities?

These examples were to be augmented by the students with more objective attributes of urban spaces in order to complete the spatial dimension of the socio-spatial analysis.

Next, we concentrated on data preparation, which constitutes an important issue for the success of data mining applications (Witten and Frank, 2005). The raw data gathered from the surveys was cleaned, validated, and stored in terms of tables. This preprocessing was fundamental in constraining the hypotheses that can explain socio-spatial processes generating such data as well as in concretizing the kind of socio-spatial patterns that can be extracted from such data. We then imported the data tables to the RapidMiner software for subsequent data mining.

Familiarized with basic notions of statistics and data mining tools during the first part of the course, students have chosen the methodologies to apply in order to test their hypotheses and tried to get answers to their questions related to socio-spatial patterns and processes in Istanbul.

Finally, we addressed data visualization as it constitutes an important component of data mining analysis in assessing the validity and/or usefulness of the obtained results, which can be, to name a few, in the form of correlation amongst variables, data groupings (clusters), or more complex hypotheses demanding data of a different kind. Visualization is also important for a lucid and useful communication of the results, enabling the users to discuss and explain the logic behind the mining process.

**Evaluation of the Outcomes**

Within the methodological framework explained above, in the second phase of the studio we focused on our main problem that is the exploration of the effects of social characteristics of the students on their activity patterns in the city of Istanbul. Students came with six different conceptual categories prepared by groups of two or three. The six different
conceptual categories were:
1. Personal Information and Identities
2. Social Communication in Virtual and Physical Environments
3. Identity, Perception and Problems of Urban Space
4. Paths from Taskisla (Building of Faculty of Architecture, ITU) through Beyoglu
5. Everyday Activities and their Specific Locations
6. Open Public Space and its Perception

They prepared very detailed data collection surveys for gathering data on their specific conceptual category. We ran these 14-page survey divided into 6 different categories, within Taskisla Community (ITU Faculty of Architecture Students who are studying Bachelor and MSc degree of Architecture, Urban Planning, Interior Design, Landscape Architecture and Industrial Product Design) and we were able to get 88 surveys completed. We have collected approximately 500 attributes in total due to the variety of the questions. Naturally, we did not include all the attributes in data mining but the most interesting ones were selected according to the questions of each group looking for specific relationships and patterns among data. Unfortunately, due to the length limits about the papers, we won’t be able to include all the survey pages and the data tables prepared for data mining but all the visual and textual material is to be found in our blog at https://dads2010.wordpress.com/. We present below each group’s concept and works in detail.

**Group 1: Personal Information and Identities**
This group has concentrated on collecting basic personal data about the students. They prepared five types of question concentrating on;
7. Questions about personal information (gender, age, hometown, department, like/dislike of the department etc.)
8. Questions about living conditions in Istanbul (where and who they are living with, time they spend on daily traveling from home to university, etc.)

9. Questions about their personal interests on social activities (travelling, language, literature, dancing, music etc.)
10. Questions about their knowledge about design tools and their preferences about courses (knowledge on design software, experience and preferences of working in office/on site, preferences on sketching and modeling, preferences about courses of design, theory and practical, etc.)
11. Questions about their interest on design activities (attendance on conferences, workshops, clubs etc.)

Data gathered in this first category is made of 35 nominal attributes and is used by the rest of the groups for matching personal information with the data gathered from the rest of the survey.

**Group 2: Social Communication in Virtual and Physical Environments**
This group has concentrated on collecting data about social communication habits and social trends of students both in virtual and physical environments. They collected data through questions prepared in 3 different subcategories;
12. Questions about the links between physical and virtual environments (How they learn about an activity, how they decide to go, how they invite people joining them and how they buy the tickets or make the reservations)
13. Questions about the social trends of students (Their preferences on magazines, music, movie, reading, art/exhibitions, TV shows, going out trends)
14. Questions about communication preferences of students (usage of social networking, blogs, mobile internet etc.)

Data gathered in the second category is made of 87 attributes most of them in the form of true/false statements. Below in Figure 1 we see a part of the survey pages of Group 2.
Group 3: Identity, Perception and Problems of Urban Space

The third group has concentrated on collecting data about the identity, perception and problems of Istanbul, by asking students to answer their questions by choosing among some visual images. Every image was coded by assigning keywords which made this part of the survey interesting and relatively easier to run. They prepared 14 different questions. Some of them are listed below to give an insight:

15. Questions related to the visual and emotional perception and identity of the city (Such as matching neighborhoods of Istanbul with different architectural materials, with different types of music, identifying the borders of Istanbul, the symbols of Istanbul,)

16. Questions related to the problems of the city (Such as how can the security increase in Istanbul, the most important problems of Istanbul, and the way to increase life quality in Istanbul etc.)

Data gathered in the third category is made of 25 nominal attributes. Below in Figure 2 we see a part of the survey pages of Group 3.

Group 4: Paths from Taskisla through Beyoglu

The fourth group has concentrated on collecting data about daily social activity paths from Taskisla to Beyoglu (the historical and contemporary centre of commerce and entertainment of Istanbul which is at five minutes distance from Taskisla). They prepared maps where the important places are identified and
presented to students for marking. Students were asked to mark the sequence of visiting their favorite places. 154 different place was grouped in 5 different category (eating/drinking places, entertainment places, cafés, cinemas, transition points) and coded by a number and each one added as an attribute to the data table. There were two different maps to fill in, one for week days one for the week-end. Below in Figure 3 we see the map prepared by Group 4. Figure 4 depicts weekdays and weekend paths of the students from Taskisla to Beyoglu, who completed the survey. We generated these figures from the paper maps filled in by student subjects. Specifically, we gridded the map in the horizontal and vertical directions, then we recorded coarse local coordinates of each place displayed. To avoid overlapping of places falling in the same cell due to relatively coarse gridding, we added some jitter on the coordinates thus obtained so that we could have visually pleasing figures. Since the students marked their visiting places in a sequence, it was then easy to generate the paths by drawing the lines connecting the marked places. All these operations have been performed automatically using a script written in MatlabTM 7.0.

**Group 5: Everyday Activities and their Specific Locations**

The fifth group has concentrated on collecting data about everyday activities and the specific locations where they occur. This group sought after identifying

---

**Figure 3**
Pages from the survey prepared by Group 4.

**Figure 4**
Weekday paths of students at the left and Weekend paths of students at the right.
the relation between everyday activities and urban sections. This one also was very attractive and fun to fill in as it mostly consisted of visual images. The survey contains a catalogue of architectural section drawings which are to be matched with the activities such as meeting, chatting, eating/drinking/smoking, studying, and resting. More information such as the reasons of choosing the specific sections, the duration of being there, the frequency of going there was also gathered. In the second part of this category students were asked to “draw themselves as human figures” in the section. This part was unfortunately very difficult to code as a data table but we let the students do it just to enjoy the results and see how the perception of the sectional drawings are changing from one person to another. Below in Figure 5 we see two different examples of “drawing human figures in the given sections”. Data gathered in this category is made of 25 nominal attributes.

**Group 6: Open Public Space and its Perception**

The final group has concentrated on collecting data about open two important public spaces of Istanbul Kadikoy and Taksim in a comparative manner. These two quarters constitute the two main hubs of transport, commerce, and cultural activities in Istanbul. There were eight questions in this part of the survey that can be considered in the following subcategories;

17. Questions about usage (time amount of spent on these places, the frequency of visiting those places, ways and means of transportation)

18. Questions about characteristics (Like/dislike criteria, the ways of improving the quality of these places)

Data gathered in this final category is made of 238 attributes, most of them coded in the form of true/false statements. This part of the survey was the most time-consuming one in terms of transferring into a data table since the sheer size of the attributes.

All of the groups applied initial data mining tests in order to test their hypotheses about the social patterns of students. Unfortunately, due to the time limitations of the studio we allocated most of our time to prepare the surveys, to clean the data coming from the surveys and to properly prepare the database. Therefore we could not allocate enough time to get mature results from the data mining process of the data gathered from the surveys. Hence, the experiments conducted in Rapid Miner are not included in the paper, but they are available in our blog at https://dads2010.wordpress.com/. For further work we plan to enlarge these databases as well as enhance the analysis by data mining.

**Conclusion**

In our paper, we presented the details of our design
studio that aimed to identify patterns and trends related to the socio-spatial activities of architecture student community in Istanbul. The main outcome of this endeavor was a valuable dataset that we plan to analyze in greater detail via data mining in our future work. Data mining is widely applied in many fields of science, engineering and business as a computational methodology of knowledge discovery. Within the scope of this course, we intended to make a case for data mining in the context of social analysis of urban space. Data mining applied in a rich and valuable socio-spatial dataset, can help to extract knowledge about social processes occurring in urban space. This kind of knowledge can further assist architects and urban planners in the design process. We have aimed to put forward a socio-spatial perspective which considers the role played by social factors such as race, class, gender, lifestyle, economics, culture, and politics on the organization of space, which we believe is very enthusiastic in terms of contemporary urban research.

References


Massey, D 2005, For Space, Sage Publications Ltd.

