ITU A|Z • Vol 15 No 1 • March 2018 • 23-38

Interpretation of urban data in the complex pattern of traditional city: The case of Amasya

Pınar ÇALIŞIR ADEM[1], Gülen ÇAĞDAŞ[2]
[1] pinarrcalisir@gmail.com • Department of Architecture, Graduate School of Science Engineering and Technology, Architectural Design Computing Graduate Program, Istanbul Technical University, Istanbul, Turkey
[2] glcagdas@gmail.com • Department of Architecture, Faculty of Architecture, Istanbul Technical University, Istanbul, Turkey

Received: October 2017 • Final Acceptance: November 2017

Abstract
Traditional cities are complex organic systems which have many forms and structures affecting each other spontaneously through time. Today, uncontrolled urbanization activities destroy the wholeness of these organic patterned structures and destroy their local character by top-down planning decisions and designers' personal motives. Thus, for enhancing the continuity in traditional urban patterns, understanding the inner nature of urban patterns is crucial. In this context, the aim of the proposed method is to deal with the huge number of raw data coming from complex traditional cities and analyze them with computational techniques in order to let designers create some basic rules and understanding in terms of the spatial organization of city structures in the early phase of urban design process. These rules can give clues about the essence of the city and guide designers and authorities for better integrated designs with traditional patterns not only in their physical form but also in the social, economic and cultural context. In this sense, the proposed method offers Data Mining algorithms in terms of knowledge discovery in the urban database with the help of Geographic Information System (GIS) technologies.

Keywords
Data mining, Urban patterns, Traditional city.

doi: 10.5505/itujfa.2018.40316

1. Introduction

Recently, modern urban planning approaches have deteriorated the relationship between the newly designed area and the context, which is embodied in traditions and represents a sense of place, as a result of handling the urban structure in parts by considering them unrelated. Besides, top-down planning concepts and schemes have controlled self-organized, unpredictable and slow growth of traditional cities. Aforementioned applications caused cities to become standardized and lose their true essence which gives them their identities. Since profit-base planning decisions cause uncontrolled urbanization, discontinuity in the traditional city structure remains today. Urbanization activities, such as high-density housing and standardized apartments, oversimplify the traditional urban form and create a built environment which is alien and disconnects with local culture, history and life rituals. Thus, the years of natural spatial evolution of traditional cities and the healthy functioning of these environments are damaged.

However, the traditional city is a complex system like a living organism which has many forms and structures affecting each other spontaneously. Even if, we shape specific areas of the traditional city with complete and rigid visions of laws and planning decisions, through time, it reorganizes its parts and transforms with bottom-up forces. Therefore, decomposing and analyzing parts of the city and their relationships are significant to reveal the inner nature and different layers of this urban system. Only with this understanding, we can protect the natural evolution of traditional city patterns and design new harmonious settlements. Collecting as much as data from the city and making sense of them are crucial for making further decision in new design processes. Natural continuity between new and old settlements can be maintained provided that the collected data turned into useful knowledge about existing structures and following periods built on this knowledge.

On the other hand, the built environment supplies designers with vast volumes of data to analyze and interpret. Therefore, most of the time, analysis of urban structures and discovering relationships among parts of the traditional city become superficial, inadequate and depends on designer's personal motives. Furthermore, repetitive patterns, random behaviors and different urban attributes influencing one another in the structure of the city are ignored by designers. So, interferences happen in the natural spatial evolution of traditional cities, and new settlements arise as imitations of old ones or become incompatible with the existing city form.

In this context, this study aims to propose a method to objectively evaluate and analyse a vast number of raw data which are accumulated beneath patterns of complex traditional cities, with computational techniques to let designers develop an understanding regarding the spatial organization of city structures. This knowledge can be turned into basic rules to be used in the early phase of urban design processes, providing designers and authorities with clues about the essence of the city for better-integrated designs with traditional patterns in their physical form. The method can discover the hidden relationships in the structure of the traditional city and let designers to enhance the continuity in the traditional urban patterns by avoiding subjective interpretations and implementations of designers and authorities in the traditional patterns. In this sense, the proposed method offers Data Mining algorithms for automated and computational discovery of knowledge in the urban database with the help of Geographic Information System (GIS) technologies.

2. Complex nature of traditional cities

Traditional cities are complex entities with organic patterns evolved in many years with various determinants. Traditional cities have not only "organic patterns" (Kostof, 1991, p. 43) but also behave like "growing organisms" (Alexander et al., 1987, p. 13) creating complexity. The term "the city with organic patterns" expresses that the city comes into being after a long process without the master plan or the assistance of designers, but takes shape with daily routines of inhabitants and the topography. According to Kostof (1991), organic patterned cities are shaped by local builders, and they are "spontaneous, chance-grown, generated and geomorphic" (Kostof, 1991, p. 43). Narrow streets compatible with the topography, buildings shared similar typologies, and open spaces dispersed randomly around the city throughout the spatial evolution process (Kostof, 1991). On the other hand, Alexander et al. (1987) expressed that traditional cities have always protected their core structure and advance naturally as a whole like "growing organisms" (Alexander et al., 1987, p. 13). Integrity of the city structure is the actual determinant of its future form and capacity for the city growth is unforeseeable. Based on these theories can be asserted that traditional cities are self-organizing growing systems, starting from a particular core structure, such as living organisms. Growth process gradually takes place and the borders of the growth are unpredictable, whereas it is always harmonious with the general city pattern, because it is primarily fed by existing urban structure.

The term 'city as an organism' in this study is more than an abstraction for an analogy of biological form. It is related to traditional surroundings and the reasons of their complex character at the end of their spatial evolution and growth through time under the forces of the natural environment and purposes of their inhabitants. In this context, traditional cities can be seen as complex organizations of parts to create integrity which is more than the sum of its parts (Carmona et al., 2003). Unitary visions of city models, which produce "artificial cities with a hierarchy like a tree" (Alexander, 1965, p. 58), emphasize on the overall appearance of cities and ignore working mechanisms of city components which create the actual complexity. In "natural cities" (Alexander, 1965, p. 58), which have organic patterns, global form and behavior have been generated for many years by local interactions between city parts with self-similar characteristics through a network which is "a semi-lattice" (Alexander, 1965, p. 58). Same with living organisms, during the time, parts of the traditional city cannot comply with the environment disappear or to be forced to change to keep alive. Thus, the evolution of traditional cities through time is the main similarity with living organisms and it maintains natural growth of city patterns based on the core structure of the city. In this process, local masters who shape the city and specialized in traditional techniques create a dynamic/vital structure which can evolve by changing, transforming and adapting itself.

Proposed study sees that traditional cities are organic not just for their organic patterns congruent with the natural environment, but also for their complex nature that is a shared feature of living organisms. Therefore, traditional cities are defined as living organisms and treated as complex systems. Also, the proposed method is designed for understanding the self-organized formative process of the traditional city, shaped by numerous determinants. Thus, in urban design activities, especially in terms of new settlement designs, it is necessary to understand these determinants. In order to achieve that the design area and its surrounding need to be analyzed by both qualitative and quantitative research methods. Later, designers can clearly see repetitive patterns, random behaviors and different factors affecting one another in the structure of the traditional city. However, traditional techniques are not capable of that type of evaluation and interpretation of traditional cities' complex nature. Therefore, we have to propose new methods which are more sensitive in terms dealing with the complex nature of these cities. In this context, we can find various studies regarding to complexity of traditional towns. For instance, Gürbüz et al. (2010) propose fractals to understand the city structures focusing on their geometrical and dimensional features and offer new design alternatives enhancing the continuity of traditional urban patterns. Similarly; Duarte et al. (2006), utilizes shape grammars by considering functional and morphological features of urban structures for preserving the character of traditional urban structures while proposing new design areas to fulfill modern life requirements. Karabağ (2010), in his Ph.D. thesis, offers a methodology for analyzing traditional urban patterns and their relations regarding their architectural features via statistical tools and generate rule sets for an algorithmic design procedure for new design interventions with the results coming from the data analysis phase.

In this article, Data Mining techniques are employed to handle complexity in traditional settings. The method is proposed for analyzing urban data, which can be both categorical and nominal, finding their relationships and constituting an understanding of the organization of traditional urban character through basic formation rules. As we mentioned at the beginning of this chapter; the traditional city produces a massive amount of data based on its complex organic nature and extracting useful information from this complex structure is unfeasible with manual analysis. However, Data Mining can handle a huge amount of raw data in a concise time without the years of training as an analyst and contains various algorithms for analyzing the raw urban data in very different perspectives, such as clustering, frequent pattern mining, classification and so on. Therefore, in this study, we choose to utilize from Data Mining algorithms for knowledge discoveries in the complex traditional urban patterns. Enhancing different Data Mining techniques cannot only help us comprehend urban attributes and their relations from different perspectives but also find anomalies in the traditional patterns. In this way, we can protect city's self-evolved and chance-grown urban structure while new settlements are designed by considering repetitive and unique structures.

3. A method proposal for interpretation of traditional urban data by using data mining techniques

For harmonious new settlements and the protection of traditional city's organic essence, designers must have definite ideas about architecture, social, cultural and economic dynamics of the city in addition to its history and local values. Therefore, the design area and surroundings need to be analysed in depth. Results coming from these analyses may help designers to reveal structures of traditional cities. In order to do this, we have to collect as much as information from cities and find appropriate techniques which can resolve the complex structure of traditional urban systems. In this context, the proposed method utilizes Data Mining algorithms.

Data Mining is a process for finding latent connections and extracting novel and logical inferences from raw data in huge databases (Hand et al., 2001). In other words, "Data Mining is defined as the identification of interesting structure in data" (Fayyad et al., 2002, p. 28). It belongs to a larger context called Knowledge Discovery in Databases (KDD) which "makes sense of the data" (Fayyad et al., 1996, p. 37). Every day, heavy load of information is uploaded into databases, and this raw data cannot be analysed manually. KDD, which arose during the 1980s (Sohtorik, 2016), provides us with computational tools to examine, interpret this data and construct plausible hypotheses for our interest in an automated and objective way. This process of transformation from raw data to useful knowledge allows us to analyse the current situations, make predictions and decisions for the future. Data Mining is the main part of this process which is "the application of specific algorithms for extracting patterns from data" (Fayyad et al., 1996, p. 39). Patterns are specific to the domain from where data are collected. Based on the intentions of the analyst, patterns are evaluated and considered as knowledge according to their "validity, novelty, usefulness and simplicity" (Fayyad et al., 1996, p. 41). In this context, Data Mining contains different mathematical techniques for various tasks, producing patterns from transformed data in databases for further interpretation and evaluation. These tasks can be classification, segmentation and clustering, association, deviations, trends and regression analysis and generalizations (Miller and Han, 2009). Thus, in Data Mining studies, it is crucial to choose appropriate algorithms for the data type, data scale and the aim of the analyst in order to obtain meaningful and reliable results. Looking for patterns in

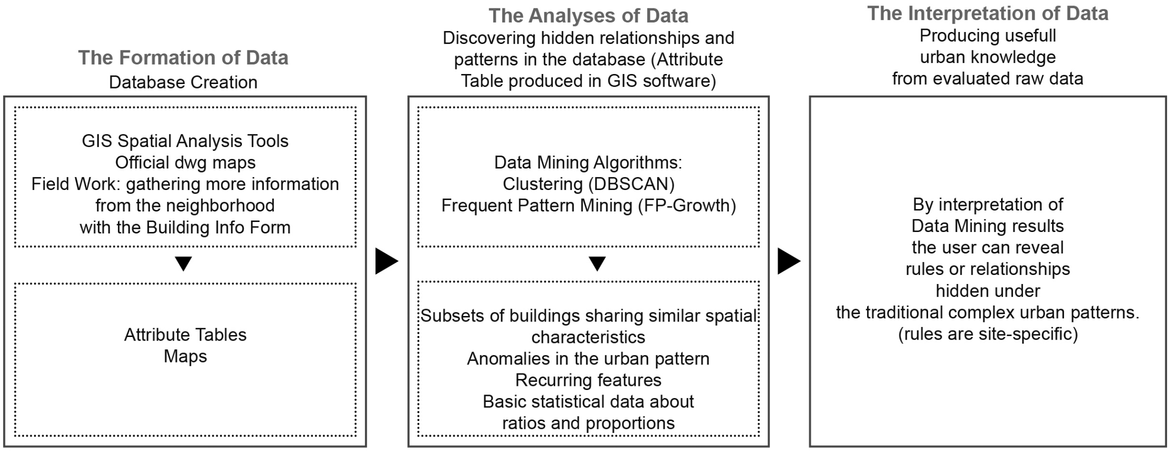

Figure 1. *The phases of the proposed method.*

databases is not unique to Data Mining. Different fields -such as statistics, pattern recognition, and exploratory data analysis- apply various techniques for handling this issue. Data Mining borrows different algorithms from these areas and offers semi or fully automated, easy to use, accessible and practical tools to evaluate large databases and discover useful patterns for users without a significant training as data analysts (Fayyad et al., 2002).

Various studies are utilizing Data Mining techniques in the architectural context. For example, by using Data Mining techniques, Gill et al. (2009) define urban typologies in the street and block level. Reffat (2008), searches patterns in features of contemporary Arabic architecture. Sohtorik et al. (2010), investigate patterns and relationships of urban attributes and propose a Data Mining methodology for new urban interventions in the complex nature of the city (Sohtorik, 2016). Laskari et al. (2008), quantify spatial attributes derived from plan features and finding related patterns of different spaces in the scale of urban blocks. Hanna (2007), defines archetypes from building examples and utilizes them into new design processes. laskari

The proposed method in this article has three consecutive phases as illustrated in Figure 1. The phases are:

- The Formation of Data
- The Analyses of Data
- The Interpretation of Data

Throughout the proposed method, to compile and visualize different sources of data from the traditional city GIS software-ArcMap/ESRI was used. Cleaning digital data was done in AutoCAD/Autodesk. In the analysis phase of this data, RapidMiner –open-source software- was used.

In the first phase of the proposed method, which is called "the formation of data", a database for attributes in the traditional city is created. Attributes, which carry information, are features of urban entities such as urban blocks, buildings, and streets. The urban database should be built in order to collect and store various features of urban entities and include as much as authentic and official data –both quantitative and qualitative- about the existing setting. Therefore, sampling of raw data should be done using various sources such as municipalities and different government agencies. Furthermore, if there is a need for additional information about the urban structure, designers can generate "An Info Form" to increase the data. In this study, GIS software-ArcMap is used for gathering raw data coming from different sources and creating the database as an Attribute table. Also, spatial analyst tools inside ArcMap generate additional information for slope, distance and aspect attributes. Through the database, collected data becomes unique to the neighborhood. From Table 1, we can see that titles in the attribute table which may vary in a wide range of data; such as architectural features, building location, dimensional features, land use information, topography and so forth. Also, data in the attribute table can be classified into different groups. In the present context, we have numerical values and categorical values which are called nominal

Table 1. *Attribute table for the traditional urban database.*

Type	Attribute Name	Type	Attribute Name
nominal	Building Status	**nominal**	Additional_Building
nominal	Construction System	**nominal**	Courtyard Entrance Orientation
nominal	Construction Material	**nominal**	Building Entrance Orientation
nominal	Building Form	**nominal**	View
nominal	Plan Type	**nominal**	Entrance Type
nominal	Roof Form	**nominal**	Entrance Qualification
numeric	Facade Number	**nominal**	Building Position
nominal	Facade Details (front-back-oriel)	**nominal**	Privacy Situation
nominal	Balcony Existence	**numeric**	Slope (Percentage)
nominal	Basement Existence	**nominal**	Aspect
numeric	Parcel Area	**numeric**	Distance to Landmarks
numeric	Building Base Area	**numeric**	Distance to Railway
numeric	Base_Parcel Ratio	**numeric**	Distance to River
nominal	Front Face Direction	**numeric**	Distance to Neighborhood Square
numeric	Ground Floor Access	**numeric**	Distance to City Square
numeric	Hall Type	**numeric**	Number of Floors
nominal	Courtyard Existence	**nominal**	Landuse Basement
nominal	Courtyard Location	**nominal**	Landuse Ground Floor
nominal	Landuse Third Floor	**nominal**	Landuse First Floor
nominal	Landuse Fourth+ Floor	**nominal**	Landuse Second Floor

based on their non-hierarchical structure. Throughout the study, additional attributes can be added or unnecessary data can be eliminated from the table in line with the selection of designers.

The second phase of the method includes Data Mining techniques and is called "the analyses of data". It is the primary phase of the method in which analysts or designers aim to analyse attributes by Data Mining algorithms in the database concerning their latent connections, repeated patterns or unique formations to reveal the complex nature of the traditional surroundings.

If we consider the complex traditional city as a giant database accumulating raw data beneath its spatial organizations, we can use Data Mining techniques to produce useful knowledge by collecting, selecting and evaluating data focusing on our design problems. These methods can help us easily understand the organizational characteristics of urban entities and different algorithms may help us to investigate raw data from the very different point of views in an automated and objective way. In this way, we can detect unfamiliar or dominant patterns, relationships among city elements and interpret these findings to reveal the essence of the traditional city which gives form to it.

In this scope, we applied two different Data Mining algorithms. First one is clustering algorithm which is used to detect subsets of the target data. Clustering is a technique for sparing a dataset into subsets or clusters based on their common features. The primary goal of the clustering is to collect similar objects in the same cluster and put different objects in the separate clusters as much as possible (Han et al., 2011). In the proposed method, the clustering algorithm can help us reveal subsets of the database which contains buildings sharing similar spatial features. Clustering techniques are highly used in image pattern recognition studies, web search techniques, fraud detection and biology disciplines. There are various clustering algorithms according to data scale and data type. In this study, the mixed nature of our dataset and the difficulty of predicting cluster numbers beforehand lead us to use DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm. Shapes of data points inside the database are not regular all the time. Sometimes, these shapes can cumulate in specific places and create arbitrary shapes in the database. DBSCAN algorithm works with large databases, especially with random shapes such as in spatial databases (Ester et al., 1996).

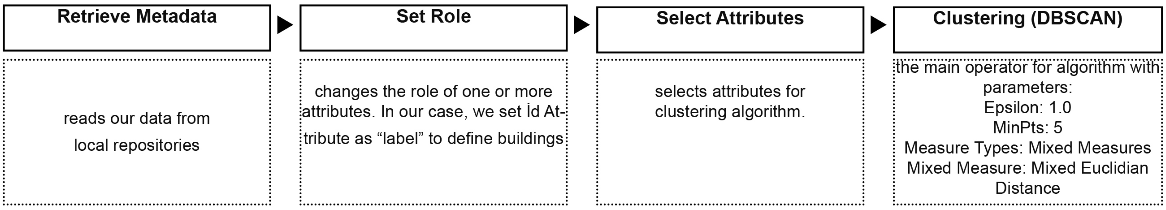

Figure 2. *DBSCAN clustering algorithm process in RapidMiner.*

Association Rule	buy (customer_X, "computer") => buy (customer_X, "software")	[support = 1%, confidence = 50%]
Meaning	1% of ALL customers shopping in the store buy a computer and a software together; 50% of customers PURCHASING COMPUTERS buy also a software from the store.	

Figure 3. *An example of an association rule and how it is defined by the support and the confidence values (Han et al., 2011).*

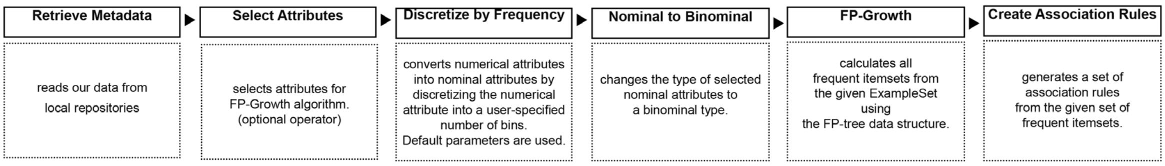

Figure 4. *Frequent pattern mining process in RapidMiner.*

DBSCAN algorithm determines the local regions on which these points concentrate, rather than merely measuring the distances between points (Zaki and Meira Jr., 2014). Overall, the primary strategy of DBSCAN algorithm is to identify clusters according to dense regions of data points which are irregular, such as 'S' or 'oval' shape (Han et al., 2011). Also, this algorithm needs only two user-specified parameters: Epsilon and MinPts. Epsilon determines the neighborhood radius, and MinPts defines the minimum point of numbers inside the neighborhood. Based on these parameters, density is calculated by counting the number of data points inside the neighborhood specified by Epsilon value. DBSCAN algorithm is easy to use with RapidMiner interface. Also, it can be applied both numerical and categorical data as well as the mixed nature of data due to Mixed Euclidean Distance measure type in RapidMiner. From Figure 2, we can see the process of the DBSCAN model inside the RapidMiner software.

The second algorithm used in this study is Frequent Pattern-Growth (FP-Growth) Algorithm. This algorithm finds frequent patterns in the dataset and produces association rules for urban attributes. The algorithm belongs to a broader term - Frequent Pattern Mining which contains different algorithms to find frequent subsets in the given database. Frequent Pattern Mining algorithms are constructive in revealing useful and meaningful patterns in the dataset. The main objective of these algorithms is displaying relationships between objects and their attributes in the database by finding hidden trends and behaviors (Zaki and Meira Jr., 2014). From the results coming from algorithms, we can define Association Rules for target objects and their attributes. In order to do that, two values need to be considered: the first one is the Support value and the second one is the Confidence value. The evaluation of the results and Association Rules can be achieved by considering these two values. From Figure 3, we can see an example of an Association Rule and how it is defined by the Support and the Confidence values (Figure 3).

Besides, FP-Growth Algorithm finds frequent patterns in the dataset and produces association rules for urban attributes and works on an item set by dividing and editing its elements according to a frequency value. Frequency value determines how often an element occurs in an itemset. The algorithm creates a tree structure in order to keep track of subsets, and by doing so, it prevents repeating objects from being held in the memory. In Figure 4, Association Rules model in RapidMiner is illustrated as consecutive steps.

In the last phase of the proposed method, which is called "the interpretation of data", designer or analyst

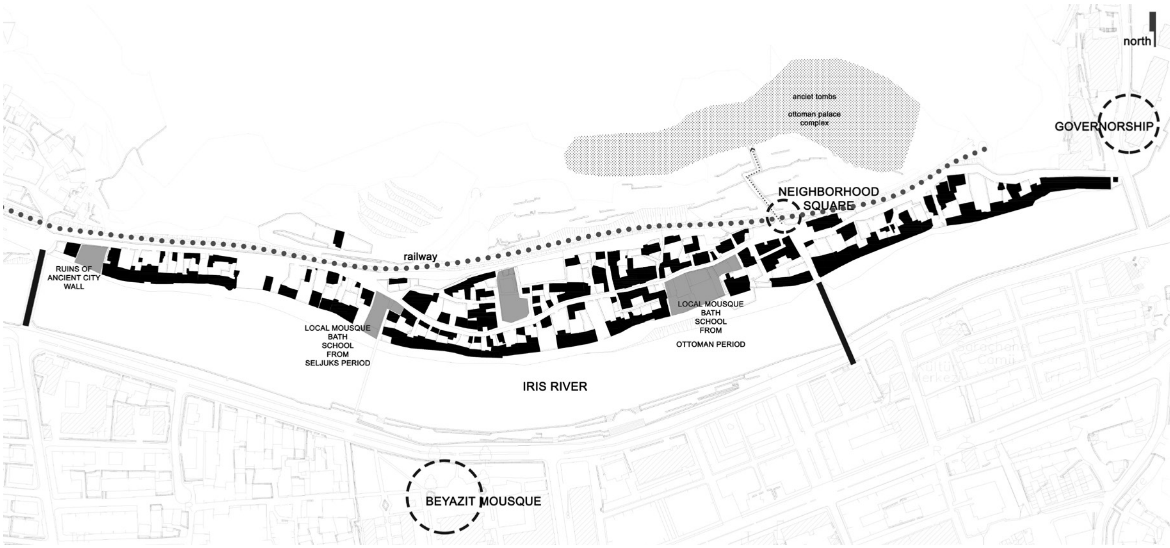

Figure 5. *Hatuniye Neighborhood: Monumental areas and traditional buildings.*

should interpret data mining results and reveal rules or relationships hidden under the traditional urban pattern. These rules or patterns represent the essence of the traditional city and reflect its characteristics. Since rules come from an objective assessment of urban attributes and their relationships, a rationale can be constituted regarding new settlement designs to enhance the continuity of the pattern of the traditional city.

4. Case study: The traditional city of Amasya and Hatuniye Neighborhood

For the implementation of proposed method, historic Hatuniye Neighborhood in Amasya/Turkey was chosen and ArcMap was utilized to create a database for this neighborhood. Amasya is one of the important historical cities in Turkey, which dates back to the 3000 BC according to archaeological excavations in the city center (Özdemir, 1996). The city develops linearly in the valley opened by the Iris River (Yeşilırmak) and is surrounded by sharp mountains on all four sides. Since the Chalcolithic Age; Pontus, Byzantine, Seljuk, and Ottoman dominance reigned over the city. Therefore, the urban configuration of the city is formed as multi-layered. In many quarters of Amasya, current patterns of streets and buildings melt into patterns of the old city so that one cannot be altered without the other. Today, we can witness the historic urban pattern from the Ottoman Empire and houses date back to 300 years ago. But like many traditional settings, the center of the city is under threat of contemporary urbanization which is not sensitive to local climate, topography, and culture and has no respect for the continuity of the traditional pattern as seen in Figure 6. Thus, there is an urgent need for diversifying computational methods, which can handle the complexity of this organic city part, to understand the real essence and organizational characteristics of the traditional city in order to protect the city's self-evolved structure respecting local climate, topography, and culture. Hatuniye Neighborhood, which is one of the historic Neighborhood of Amasya, also called "Inner City" is located along the Iris river and leans its back to the Kırklar Mountain. At the peak of the mountain, Harşena castle, above it 5 Pontic tombs and the urban structure of the neighborhood with the river create "a poetic urban experience" (Bechhoeffer and Yalçın 1991, p. 24). Despite massive destructions throughout the history, this quarter protects itself thanks to its unique location and surrounding city walls, which create barriers with river and mountain position and separate the neighborhood from the rest of the city. From Figure 5, we can see the general characteristic of the neighborhood.

The neighborhood has four bridges, and two of them draw the periphery of the neighborhood. For a general view, the quarter consists of 14 street blocks, 204 parcels, 165 main buildings and 41

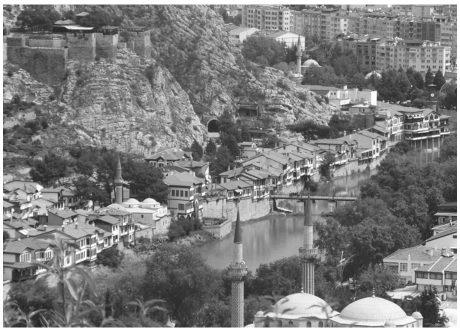

Figure 6. Amasya view with Hatuniye Quarter: Confrontation of the traditional and contemporary pattern.

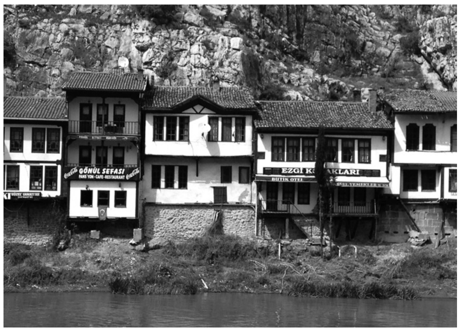

Figure 7. Closer view to Hatuniye Quarter from Iris River: City walls as foundations for buildings.

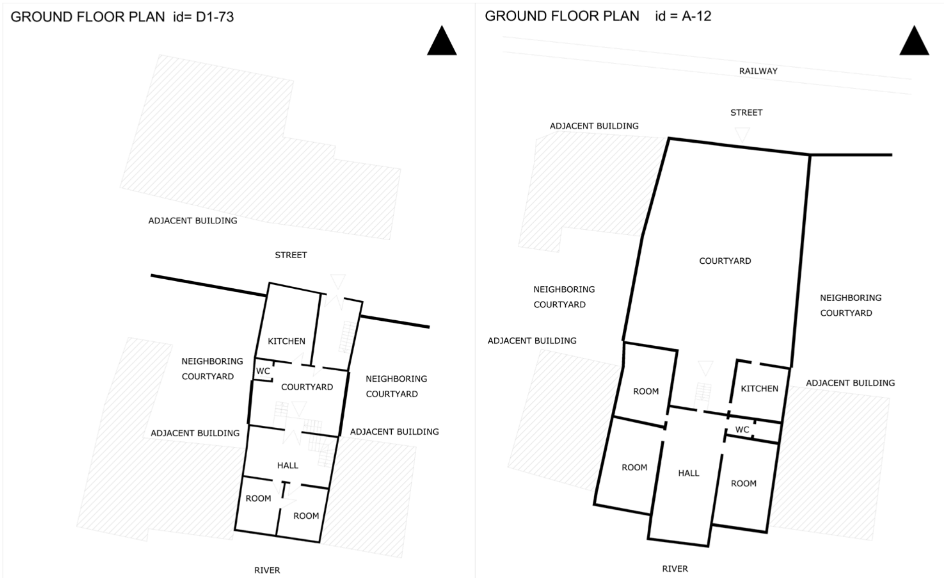

***Figure 8.** Building examples specific to the neighborhood: The building on the left is "Selamlıklı" (Redrawn from Amasya İli, Hatuniye Mahallesi Geleneksel Yerleşim Dokusunun Analizi, Değerlendirilmesi Ve Koruma Geliştirme Önerisi (Unpublished Master Thesis), by Türkoğlu, E.,2006, Ankara: Gazi University; Institute of Science).*

additional buildings in total.

Most of the waterfront houses in the Neighborhood are from the Ottomans in the 19th century. In general, houses were built either in the "bağdadi" or "hımış" techniques. Nevertheless, foundations of waterfront houses were made from cut-stone (Yalçın, 1998). These stone foundations are part of the city walls and riverfront buildings located upon them (Figure 7). In other words, in the ancient times, the city wall was merely a protection for the city. Later, the Ottomans used traces of ancient neighborhood and built their houses upon them. Therefore the city wall did not served as a protection, but a masonry wall for their foundation. We can trace the Ottoman city patterns from the general view of the neighborhood and these patterns reflect great complexity in self-similar, self-organized, and dynamic urban structure which is constantly and slowly change like a living organism. The floor plans in Figure 8 illustrate connections which buildings establish with their neighbors and the public space.

Despite huge construction activities, that ruined most of the traditional buildings in the quarter, there are still traces of Ottoman urban structure through the streets. Therefore, implementation of the proposed method in this quarter may help local designers and authorities to gain a deeper understanding of this unique structure and its underlying principles for future interventions.

5. Implementation of the proposed method in Hatuniye Neighborhood

As we mentioned in Chapter 3, the proposed method subdivided into three steps: "the formation of data"; "the analyses of data" and "the interpretation of data". In the formation phase of the method, we created a database through ArcMap, preprocess and visualize the data for Data Mining algorithms. Firstly, we gathered city information through Amasya Municipality in the format of vector maps. These AutoCAD maps contain official urban data about the neighborhood. Cleaning and selecting useful layers inside maps were done in AutoCAD. However, the collected information from this institution was not enough for more in-depth and reliable operations; therefore, creating a building info form to increase the data was necessary. As illustrated in Chapter 4, floor plans of buildings were collected from Türkoğlu's master thesis (2006) to gain a better understanding of structures in the context of the neighborhood setting. The info form about buildings holds various data which are values of attributes of urban entities previously represented in Table 1. After collecting as much as data, we started visualizing the urban data in ArcMap and matched the data with actual buildings. ArcMap does not only visualize the data in the real space, but also it developed our data with its spatial analyst tools such as slope and aspect operations. Moreover, with Multiple Buffer Ring tool, we classified all buildings, according to their distance to monumental buildings. ArcMap stored all the urban data matched with buildings in the Attribute Table, and we could export this table in Excel format to use in the analysis process with Data Mining.

In the analysis phase of the method, the data table from ArcMap was imported in a RapidMiner as an Excel sheet. For a preliminary action, software analysed all nominal and numeric values, such as parcel and building areas, in terms of maximum, minimum and average values. Based on these analyses, the smallest street block is 66,468 m^2 and the biggest one is 9144 m^2. For the size of parcels, while the smallest one is 21,196 m^2, the biggest one is 949,392 m^2, the average value of parcels is 156,416 m^2 and deviation value is 118.856. Similarly, the smallest building size is 21,196 m^2 while the biggest one is 324,786m^2 with average 83,031 m^2 and the deviation value 53.180. Monumental buildings with large areas were also added into the calculation as we can understand from high value of deviation, therefore, their influence on the average should be considered during the interpretation process. In this way, Rapid Miner can give us statistical results about maximum, minimum and average values for the building envelopes and open spaces for further design studies in order to capture the scale of the existing traditional pattern and avoid discontinuities in the visual integrity of the neighborhood. Also, in terms of nominal values, 146 buildings out of 165 is in the low-slope area (slope value: 0-24%); while only 19 buildings in the high-slope area (24-63%). 137 buildings out of 165 are in the south aspect, while 12 buildings in the southeast and 16 buildings in the west aspect. Only 5 buildings have oriel windows and 15 buildings out of 165 have "selamlık" building (which is an additional room -or building in Amasya case- for male guests in the Ottoman plan layout). 98 buildings do not have any projection on their façade and 89 buildings do not have a hall. 98 buildings out of 165 do not have basement floor and 42 buildings do not have a courtyard. 61 buildings are above ground level with high basement wall. 108 buildings are attached house and 88 buildings have a main entrance from the street to the courtyard, while 53 buildings have direct access to the house from the street. These statistical results can give an idea about the general characteristics of the neighborhood. However, for gaining deeper insight about urban attributes and their relationships, more Data Mining algorithms were applied. In the data table for attributes of Hatuniye neighborhood, there are mainly nominal but also few numerical attributes; so we cannot foresee the number of clusters beforehand. Therefore, DBSCAN algorithm was used to see consistent patterns or anomalies in the urban structure. As we mentioned in Chapter 3, DBSCAN operations need two user-specified parameters: Epsilon and MinPts. In RapidMiner default value for Epsilon is 1.0 when the MinPts val-

Table 2. DBSCAN clustering algorithm. Clusters of Test 1.

Cluster	Dist. to Neig. Square	Additional Building	Num.of Buildings
1	500-400 m	exist	8
2	400-300 m	not exist	13
3	600-500 m	exist	6
4	400-300 m	exist	10
5	300-200 m	not exist	23
6	200-100 m	not exist	29
7	300-200 m	exist	22
8	200-100 m	exist	9
9	100-0	not exist	27
10	100-0	exist	12
		TOTAL	165 (159 + 6 noise data)

ue is 5. We mostly followed MinPts value, unless we did not need more refined or crowded clusters. Low or High number of epsilon value did not give us any feedback; therefore, we set this number to 1.0 in all DBSCAN operations. Additionally, we chose our distance measure, according to our attribute types (the nominal, numerical or mixed type of attributes). First of all, we chose all attributes for DBSCAN algorithm with chosen parameters. But, the model cannot produce any cluster. In order to gain more refined clustering results, we choose some attribute groups which repeat in the neighborhood from a general point of view. Thus, in test 1, we set Epsilon to 1, and MinPts to 5, for revealing attribute relationship between "additional building existence" and "distance to the neighborhood square". Results after of the clustering process can be seen in Table 2 as an example of clustering operations.

Test 1 gave us 11 clusters with Cluster 0 containing six buildings as noise data. According to the results, we can specify that 56 buildings which are 0-200 meters away from the neighborhood square do not have any additional building and the probability of having additional structures increases as buildings become distant from the neighborhood square. In test 2, we wanted to discover the relationship between "additional building existence" and "parcel area" with parameters of test 1. The algorithm created 11 clusters with 19 noise data inside Cluster 0. According to results, there are 45 buildings with no additional structures and their parcel base is between 0-100 m^2. Buildings in the 100-250 m^2 parcels are at similar numbers. In test 3, chosen attributes are "view" and "basement existence" with the same parameters with previous tests. At the end of the algorithm, there were 7 clusters with 1 noise data in Cluster 0. The most important clusters at the end of the operation are Cluster 1 and Cluster 2. Cluster 1 contains 56 buildings with street view and no basement floor. Cluster 2 contains 52 buildings with river view and basement floor. Cluster 3 contains 26 buildings with no basement floor in the river area, and there are 8 buildings with street view and the basement floor in Cluster 4. Cluster 5 contains 6 buildings with basement and Cluster 6 contains 16 buildings with no basement and view of the mountain. As seen from results, in the riverside, there are more buildings with basement floor and buildings with the view of the mountain or street usually do not have a basement floor. In test 4, selected attributes were "view" and "courtyard location". The most important cluster, in this case, is Cluster 2 containing 62 buildings with river view and a front courtyard. Also in cluster 4, there are 10 buildings with an inner courtyard and river view. Only 5 buildings on the riverside do not have a courtyard which is represented in Cluster 3. In test 5, we considered "aspect" and "building position" attributes. According to the results, there were 6 clusters, including Cluster 0 with 8 noise data. In this test, the most important cluster is Cluster 3 with 84 attached houses in the southern aspect. When "building position" was compared to "view" attribute in test 6, results appear in a similar way. Based on the results, in Cluster 3, which is the most crowded one, there are 71 attached houses on the riverside. In test 7, "aspect" and "courtyard" locations were considered. As a result, 7 clusters appeared with 17 noise data in Cluster 0. The most important cluster is Cluster 2 with 77 buildings in the south aspect with a front courtyard. In test 8, we wanted to discover "halltype" and "parcel area" relationship. But when we used the same MinPts value with the previous test, which is 5, we get 13 clusters including Cluster 0 with 57 noise data. Therefore, in test 8 MinPts was set to 3. According to the results, 30 clusters appeared with 23 noise data in Cluster 0. Here, buildings with no hall were detected. Cluster 1 appears with 11 buildings with no hall and 0-50 m² parcel area. Cluster 16 has 29 buildings with no hall and 50-100 m² parcel area. Cluster 3 has 18 buildings with no hall and 100-150 m² parcel area. Cluster 6 contains 10 buildings with no hall and 150-200 m² parcel area. Cluster 9 and 10 has 6 buildings with no hall and 200-250 m² and 250-300 m² parcel area. Cluster 11 has 2 buildings with no hall and 350-400 m² parcel area. Finally, Cluster 24 and Cluster 29 have 2 buildings with no hall and 300-350 and 500-550 m² parcel area. As a result of the clustering studies, we can assert that clustering techniques can help us to see repetitive patterns, anomalies or obvious structures in the urban pattern.

Another Data Mining technique used in this study is the Frequent Pattern Mining which reveals attributes that are frequently used together and create some association rules according to mining results. FP-Growth algorithm is used in these tests as we mentioned in Chapter 3. The results of the first test can be seen in Table 3.

In the first test, we tried to use all nominal attributes for the creation of association rules. By changing support and confidence values we may create a high number of Association Rules. In this case, our support value is 0.5 and confidence value is changing between 0.7 to 0.9. In Table 3, we can see association rules which define urban attributes frequently used together in the dataset. The first column represents premises and the second column represents conclusions. Other columns give us confidence and support values according to attributes relationship between the first two columns. For instance, according to rule. 1 values, we can say that 56% of all buildings in the neighborhood have no additional structures and oriel windows in the south aspect and they have an entrance

Table 3. *Selected association rules for Hatuniye Neighborhood (The first test).*

No	Premises	Conclusion	Support	Confidence
1	plan type=single, aspect = south	oriel window = absent, build. entrance orientation =ns	0,56	0,73
2	slope = group 1(0-24%), aspect = south	plan type=single, balcony details = no balcony	0,55	0,73
3	aspect = south	courtyard existence = with Courtyard	0,61	0,74
4	oriel window = absent, courtyard existence = with Courtyard	building position = attached houses	0,53	0,75
5	plan type=single, build. entrance orientation =ns	roof form = saddle Roof	0,55	0,76
6	courtyard existence = with Courtyard	courtyard location = front	0,57	0,76
7	entrance type = no platform	slope = group 1(0-24%),	0,60	0,89
8	ground floor access = earth level	aspect = south	0,56	0,89
9	ground floor access = earth level	plan type=single, entrance type = no platform	0,56	0,89
10	aspect = south	slope = group 1(0-24%),	0,75	0,90
11	basement existence = no basement	slope = group 1(0-24%),	0,53	0,90
12	aspect = south, courtyard existence = with Courtyard	plan type=single	0,55	0,90
13	roof form = saddle Roof	build. entrance orientation =ns	0,61	0,90
14	front face detail = no projection	oriel window = absent, plan type=single	0,55	0,92
15	slope = group 1(0-24%), building position = attached houses	build. entrance orientation =ns	0,53	0,94

Table 4. *The second association rule test.*

No	Premises	Conclusion	Support	Confidence
1	NumOfFloors=2, distGovernorship= 500-400 m.	buildingBaseArea=50 100m²	0.11	0.63
2	distGovernorship= 700-600 m.	NumOfFloors=2	0.13	0.73
3	distGovernorship= 600-500 m.	NumOfFloors=2	0.12	0.74
4	buildingBaseArea= 50-100m²	NumOfFloors=2	0.37	0.74
5	distGovernorship= 500-400 m.	NumOfFloors=2	0.18	0.81
7	buildingBaseArea= 50-100m², distGovernorship= 500-400 m.	NumOfFloors=2	0.11	0.90

Table 5. *The third association rule test.*

No	Premises	Conclusion	Support	Confidence
2	buildingBaseArea= 50-100m², parcelArea= 50-100m²,	hallType= noHall	0.12	0.64
3	parcelArea= 50-100m²	buildingBaseArea= 50-100m²,	0.18	0.67
4	parcelArea= 150-200m²,	buildingBaseArea= 50-100m²,	0.12	0.67
5	hallType= noHall, parcelArea= 50-100m²,	buildingBaseArea= 50-100m²,	0.12	0.69

from north-south orientation. However, 73% of buildings with no additional structures in the south aspect has an entrance from north-south orientation. For rule no.2, 55% of all buildings in the neighborhood lay on the low slope surface, constructed in the south aspect with no additions and balcony; whereas 73% of buildings constructed in the low slope area and oriented in the south aspect has no additional buildings or balconies. For an overall overview of the rules, we can state that 74% of buildings in the south aspect have courtyards. 75% of courtyard buildings are attached houses with no oriel window, and 76% of them have a front courtyard. 89% of buildings with entrances from the ground level is in the south aspect with no additional buildings and entry platform. 90% of buildings in the south aspect and no basement floor are in the low slope area. Also, 90% of buildings with a courtyard in the south aspect has no additional structures. In test 1, although the algorithm produces more association rules and confidence is more than 70%, most of the time the rules are recurring at different attribute combinations. Therefore, we chose Association Rules for their best fit to our objectives. In the following tests, we tried to filter urban attributes for the creation of more focused association rules. For instance, in the second test represented in Table 4, the selected attributes are: "Building Base Area", "Number of Floors" and "Distance to Governorship". Confidence value, in this case, was changed to 0.6 and the support value was determined as 0.5. So, the algorithm becomes more sensitive concerning finding hidden relationships between urban attributes. According to this test, more focused rules emerged, as seen in Table 4.

From Table 4 we can focus on the rule no. 1 saying that 61% of two-storey buildings 500 meters away from governorship have a base area approximately 50-100 meters square. Also, the rule no. 4 can be useful for revealing an association between "building base area" and "number of floors". As it turns out, 70% buildings with the base area of 50-100 meters square are two-storey buildings.

In test 3 for association rules, the selected attributes were "building base area", "hall-type" and "parcel area". This test gave us rules in Table 5. In these rules, we can concentrate on rule 2 which has more meaning than the others. It says that 65% of buildings which have 50 to 100 meters square parcel and the base area has no "hall/sofa" in their plan schemes.

Interpretation of these results is the final phase of the proposed method. All buildings in the selected area -Hatuniye Quarter was examined by Data Mining algorithms to determine specific characteristics representing the traditional urban setting. These features can be used in new settlement designs in the neighborhood. Thus, in this section, the results coming from Data Mining operations were interpreted to discuss what this method may offer regarding to investigating urban patterns and defining their relations to each other for urban design studies embracing locality. Significant Data Mining results can be interpreted as follows:

Clustering Test 1 and 8: Buildings closer to the neighborhood square usually do not have any additional buildings or hall.

Clustering Test 2: If we have parcels around 0 to 100 m² we do not likely design additional buildings due to small parcel areas.

Clustering Test 3: In the neighborhood, an essential pattern is buildings with the basement floor and river view. Anomalies in terms of basement floor identify new buildings for accommodation which destroy the integrity of the traditional pattern regarding building scale.

Clustering Test 4: In the riverside houses living rooms are usually designed through the river. Therefore the main entrance to the building is on the north and buildings have a front courtyard. Also, only 5 houses do not have a courtyard. So, creating a semi-private area in front of the building is a trend among riverside houses.

Clustering Test 5 and 6: Buildings in the south aspect are usually attached houses. The same pattern is also seen in riverside houses. Therefore, new designs in these areas should follow this pattern.

Clustering Test 7: Similar to test 5 and 6, in the south aspect, most of the buildings have a front courtyard because of the same reasons with clustering test 4. Living rooms face to the south for better river view in addition to natural light.

Assoc. Rules: Test 1. Rule 1: Buildings in the south aspect and without additional buildings are oriented towards North-South.

Assoc. Rules: Test 1. Rule 2: Buildings in the south aspect and low-slope area do not have a balcony or additional buildings.

Assoc. Rules: Test 1. Rule 5: Buildings oriented towards North-South have a saddle roof in the same orientation based on the attached house pattern.

Assoc. Rules: Test 1. Rule 7-8-9-10-11: Buildings with no high foundation walls are usually in the low-slope area and south aspect. Besides, they do not have basement floor.

Assoc. Rules: Test 2. Rule 1-4: Buildings away from governorship usually have smaller base areas. Also, if we design houses between the area 50-100 m², we can use 2-storey to follow the urban pattern scale.

Assoc. Rules: Test 3. Rule 2: If buildings in small parcels and base areas are designed, using the hall in the plan layouts is unnecessary.

These results can be promising for understanding the nature of the neighborhood structure for a start. But still, we need to collect more information and expand our data set to find more intricate relations between urban entities in the scope of further urban pattern explorations which can help us reveal the inner nature of the traditional city shaped through history.

6. Conclusion

The proposed study presented a method based on Data Mining technique for an exploration of urban patterns in the traditional city. Experiments carried out in this paper should be preliminary for further studies; therefore, a small part of Amasya- the traditional Hatuniye Neighborhood- was chosen for the implementation. In the framework of the study, first, a database was prepared in GIS tools. Later, the dataset was used for Data Mining to investigate patterns and relationships among urban entities. Through this, transforming raw data into the data set and into useful knowledge about urban characteristics was aimed. The process of constructing the building database is still takes place, but obtained results in this paper showed that Data Mining presents various useful techniques to analyse raw urban information. For instance, numeric attributes can be classified according to its function, and we can determine lower and upper size limits of urban entities. Also, the repetitive urban patterns can be detected and utilized by designers in the pre-design phases. Number of interpretations can be increased by collecting more data from the city and repeated Data Mining techniques according to the aim of the designers. The proposed study focuses on topographical and morphological attributes in the neighborhood. But in the future studies, demographic, economic and social data can be added to further analyse the traditional city patterns concerning socio-economic and cultural aspects of the city. Besides, because of the context-sensitive nature of the Data Mining results, the proposed method can be easily used in different databases created for different cities. As Bechhoeffer (2001) mentioned, urban culture and history are embodied and frozen, especially in the traditional cities. In order to sustain the continuity of the past, we must protect the physical, cultural and social character of the city. By this way, we can transfer the knowledge from the past and melt the past, present, and future of the city in the same pot to protect the urban culture and enhance local values which creates our lives in the first place. This study can help to constitute a rationale for the new design processes in traditional environments and also sustain the continuation of the urban patterns and characteristics in the city by revealing hidden knowledge of the city form.

References

Alexander, C. (1965). A City is not a Tree. *Architectural Forum*, 122(1-2), 58-62.

Alexander, C., Neis, H., Anninou, A., King, I. (1987). A New Theory of Urban Design. New York; Oxford: Oxford University Press.

Bechhoeffer, W., Yalçın, A.K.(1991). Amasya, Turkey: Lessons in Urbanity. *Architect 40: Architecture in Development.* September (1991), pp. 24-29, London: Concept Media Ltd.

Bechhoeffer, W.(2001). Amasya: The Future of Tradition. In Turgut, H., Kellet, P. (Eds.), Traditional Environments in a New Millennium: Defining Principles and Professional Practice (pp.51-54). Ankara: Yapı Endüstri Merkezi.

Carmona, M., Heath, T., Oc, T., Tiesdell, S. (2003). *Public Places-Urban Spaces: The Dimensions of Urban Design.* Amsterdam; Boston; London: Architectural Press.

Duarte, J., P., Gonçalo, DS., Caldas, L.G. (2006). An Urban Grammar for The Medina of Marakech: Towards a Tool for Urban Design in Islamic Contexts. In Gero, J.,S. (Ed.) Design Computing and Cognition'06 (pp. 483-502). Springer: Netherlands.

Ester, M., Kriegel, H. P., Sander, J., Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In Simoudis, E., Han, J., Fayyad, U. (Eds.) Proceedings Book from *The Second International Conference on Knowledge Discovery and Data Mining* (pp.226-231). AAAI.

Fayyad, U., Uthurusamy, R. (2002). Evolving Data Mining into Solutions for Insights. *Communications of the ACM*, August (2002), vol.45, No.8, 28-31.

Fayyad, U., Piatetsky-Shapiro, G., Symth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, Fall (1996), 37-54.

Gil, J, Montenegro, N., Beirao, J.,N., Duarte, J., P. (2009). On the Discovery of Urban Typologies: Data Mining the Multidimensional Character of Neighborhoods. In Çağdaş, G., Çolakoğlu, B. (Eds.), *Proceedings of the 27th Conference on Education of Computer Aided Architectural Design in Europe* (pp. 269-278). ITU: Istanbul:Turkey.

Gürbüz, E., Çağdaş, G., Alaçam, S. (2010). A Generative Design Model for Gaziantep's Traditional Pattern. In Schmitt, G., Hovestadt, L., van Gool, L. (Eds.), P*roceedings of the 28th Conference on Education of Computer Aided Architectural Design in Europe* (pp. 841-849). ETH Zurich: vdf Hochschulverlag AG.

Han, J., Kamber, M., Pei, J. (2011). *Data Mining:Concepts and Techniques* (3rd. ed.). Amsterdam; Boston; San Francisco; CA: Elsevier; Morgan Kaufmann.

Hand, D., Mannila, H., Smyth, P. (2001). *Principles of Data Mining*. Cambridge, Massachusetts : The MIT Press.

Hanna, S., (2007). Automated Representation of Style by Feature Space Archetypes: Distinguishing Spatial Styles from Generatives Rules. *International Journal of Architectural Computing*, vol.5, issue 1, 1-23.

Karabağ, K. (2010). A Computational Architecture Methodology For Design in Traditional Tissue: The Case Of Kalkan (Unpublished Ph.D. Thesis). METU; Institute of Science, Ankara.

Kostof, S. (1991). *The City Shaped: Urban Patterns and Meanings Through History*. London: Thames and Hudson.

Laskari, A., Hanna, S., Derix, C. (2008). Urban Identity Through Quantifiable Spatial Attributes: Coherence and Dispersion of Local Identity Through the Automated Comparative Analysis of Building Block Plans. In Gero J.,S., Goel, AK, (Eds.), *Design Computing and Cognition'08: Proceedings of the Third International Conference on Design Computing and Cognition* (pp. 615-634). New York: Springer.

Miller, J., H., Han, J., (2009). Geographic Data Mining and Knowledge Discovery: An Overview. In Miller, J., H., Han, J. (Eds.), Geographic Data Mining and Knowledge Discovery (pp. 1-26). London, New York: CRC Press: Taylor and Francis Group.

Özdemir, C. (1996). Amasya. Arkitekt. 9(96). No. 441, pp. 28-35, İstanbul: Nokta Basın A.Ş.

Reffat, M.,R. (2008). Investigating Patterns of Contemporary Architecture Using Data Mining Techniques. In Muyyle, M. (Ed), *Proceedings of the 26th Conference on Education of Computer Aided Architectural Design in Europe* (pp. 601-608). Antwerpen, Belgium: The Higher Institute of Architectural Sciences.

Sohtorik, A. (2016). A Knowledge Discovery Approach to Urban Analysis: Beyoğlu Preservation Area As A Data Mine (Unpublished Ph.D. Thesis). ITU; Institute of Science, İstanbul.

Sohtorik, A. S., Çağdaş, G., Sarıyıldız, S. (2010). Exploring the Patterns and Relationships of Urban Attributes by Data Mining. In Schmitt, G., Hovestadt, L., van Gool, L. (Eds.), *Proceedings of the 28th Conference on Education of Computer Aided Architectural Design in Europe* (pp. 873-881). ETH Zurich: vdf Hochschulverlag AG.

Türkoğlu, E. (2006). Amasya İli, Hatuniye Mahallesi Geleneksel Yerleşim Dokusunun Analizi, Değerlendirilmesi Ve Koruma Geliştirme Önerisi (Unpublished Master Thesis).Gazi University; Institute of Science, Ankara.

Yalçın, A.K.(1998). The Ottoman House Construction in the Old Town of Amasya. In Ireland, S., Bechhoefer, W. (Eds.), The Ottoman House: Papers from the Amasya Symposium (pp.90-94). Ankara: The British Institute of Archaeology; The University of Warwick.

Zaki, M.J., Meira, Jr. W. (2014). *Data Mining and Analysis: Fundamental Concepts and Algorithms*. Cambridge: Cambridge University Press.