Proceedings of the 14th International Space Syntax Symposium

#85.

Decoding Tehran

Exploring Spatial Configuration, Clusters, and types

Sepehr Zhand; University College London; sepehr.zhand.18@ucl.ac.uk

Dr. Tasos Varoudis; University College London; t.varoudis@ucal.ac.uk

Prof. Kayvan Karimi; University College London; k.karimi@ucl.ac.uk

ABSTRACT

Urban systems are intricate and dynamic, arising from a complex interplay between top-down planning, grassroots initiatives, and the influence of a growing population. These systems evolve in response to concurrent trends in overarching processes, planning policies, and migration patterns. Over time, as a city undergoes transformative changes, it adapts to a multitude of conditions that collectively define its unique character. This implies that a city's configuration comprises distinct types intricately connected to a variety of non-spatial factors, encompassing the physical, social, political, and other overarching elements shaping its identity.

This study uses the spatial network of the metropolitan area of Tehran and examines how properties captured through space syntax analysis – specifically Integration and choice - serve to identify types linked to these measures. The fundamental premise of this research is grounded in the idea that the street network's segments and the city's spatial configuration encapsulate its developmental trajectory. Employing unsupervised classification technique at local to global radii, this study uncovers distinct urban types inherently linked to overarching policies and planning strategies. Leveraging the street network model of Tehran's metropolitan area, this research employs syntactic measures to identify clusters, which are then evaluated against various spatial and non-spatial metrics. These identified spatial configuration types exhibit self-similarity and correspond to specific planning and policymaking processes.

In the context of this study, we argue that by exercising a degree of supervision over the number of clusters and their stability, one can employ unsupervised classification techniques over syntactic measures to reveal the physical development patterns of the city. We put this methodological approach forward as a way of typological classification of cities.

doi: 10.36158/979125669032992

KEYWORDS

Spatial network, Space Syntax analysis, unsupervised clustering, K-Means clustering, Contingency map.

1 INTRODUCTION

A fundamental premise within contemporary urban planning and design disciplines is the delivery of predetermined qualities aimed at providing amenities and accommodations suitable for any given period. Consequently, a crucial aspect for improving future planning lies in comprehending how planning mechanisms manifest on the ground. This necessitates an understanding of the realized structures and substructures that emerge in relation to spatial planning and policymaking. However, given the myriad complex systems inherent in the spatial structure of cities, a significant degree of simplification becomes necessary to facilitate a legitimate representation and explanation of these structures and substructures.

Conversely, patterns, structures, and types often signify diverse processes that may lack a uniform defining baseline. For instance, an urban structure with notable local economic significance, such as a niche local market, could hold cultural and political importance beyond its immediate vicinity. This underscores the challenge of understanding a distinct structure amidst a multitude of processes, particularly given conventional typological classification methods. Moreover, it suggests that structures, types, and patterns possess attributes that may exhibit varying and sometimes conflicting structures and substructures. Consequently, an analytical framework that identifies these distinct urban structures and substructures should effectively capture both spatial and social attributes.

Thus, the central inquiry of this paper revolves around exploring the association between conventionally defined urban structures outlined by scholarly research—incorporating historical, social, and economic perspectives—and evidence-based analytically defined urban structures cantered around unsupervised modelling and statistical reasoning. To undertake this investigation, the city of Tehran serves as a case study, with an examination of its history and an assessment of the association between its structure as delineated by socio-political narratives and spatial analysis.

2 THEORY

2.1 Composition of the city

Drawing from explorations into the nature of an ever-evolving entity, this paper tries to address a few issues regarding the ways in which a city can be explained. With profound track record in varied disciplines, the nature of the city has been a wonderment for many scholars that deal with different aspects of it. However, most pre-modern philosophers and theorists based their definition on ideological and isolated ideas that explained urban phenomena incomprehensively. And while during the modernist era, there were more evidential – however rigid – thinking towards cities, the idea of explaining cities in relation to the larger processes and in context gained more currency. On another hand, by the end of the peak of the modernist time, certain recipes about the organization of the city began to develop. Lewis Mumford (1937) in his enlightening paper, initially raised the question of 'what is a city' where from a sociological point of view and in relation to the writings of More, Owen and Stow, he concluded that a city is a related collection of primary groups and associations. Kevin Lynch (1964) approached the question from a rather objective point of view and through proposing a method of observing city elements he suggests that there is characteristic in the composition of urban elements, and further asserts the idea of grouping.

Such ideas gained traction in the modernist era, where the mainstream idea of plannable city was later hardly criticised (Whyte, 1980; Jacobs, 2011; Alexander, 2012) for its failure to accommodate sustainable human life. Same argument was later elaborated on by Rowe and Koetter (1984) and through discussing the ambiguities of modernist planning, they questioned whether the city can be planned. Varzi (2021) revisits the discussion and question the nature of a planned city in light of constant evolution and transformation. From his point of view at any instance a city is a different entity unless you look at them as process.

On this end, it becomes a challenge to highlight and explain such processes and given the multiplicity of dynamical systems that govern and develop the city, it would be even more difficult to examine, explain and explore the potentials of the city. As Bettencourt (2013) explains, city can be explained as an emergent phenomenon resulting from numerous complicated interactions between elements that function concurrently yet it cannot be defined as an organization that can be broken down into these parts. These ideas as well as other iterations of the same concept remarking cities as assemblages (McFarlane, 2011) frame this general discussion in a certain way, and while such questions by nature do not have a direct answer, they provide a framework through which certain evidence can be explored and a perspective into the idea of the city as a complex composition can be tested.

Thus this paper further draws on this idea and tries to test whether the narrative of the composition of the city as a political construct (Fiori, 2019) is supported by the hard evidence from spatial analysis. The next section explains the narrative illustrating the grand image of the composition of the city of Tehran followed by analysis that tests the association between this narrative and analytical representation of the same composition.

2.2 Narrative of the city

Once the idea of the city as a composition of diverse elements is framed, it becomes a challenge to explore these elements, and relate them to the drivers that inform development growth and composition of the city. With regards to large scale research agendas such as understanding the composition of a city that has grown quickly in a context of fluctuating economic, political and social changes. Specific to Tehran, this narrative would be rather useful to review as the city is unique in terms of its development timeline. Situated in a network of strategic routes that connected the city beyond the country, the city grew from a garrison town (Amirahmadi and Kiafar, 1993) to a multi-million resident city in a course of approximately 200 years is unique in terms of the way different dynamical systems have left a footprint in morphology of the city. Therefore, drawing a large picture on the historic processes, one can dissect the composition of the city in relation to these historic developments and explain if the overall dynamics of the growth have generated types that are distinctive. However, evaluating these narratives and its highlights may not be easy, as identifying an indicative variable that relates to both the spatial type and the soft evidence from the dynamics process is not straight forward.

This problem refers back to a more elaborated and commonly discussed problem in urban morphology where the concentration is to find a narrative that explains space and the soft evidence of social, economic and political dynamics at the same time. (Karimi, 2012) Therefore having a look at the grand trends in social, political and economic changes in the country offers some clues into how certain simultaneous social, political and economic processes lead to political and planning decisions for the city. Figure 1 summarizes these trends as well as major incidents – including revolutions and reforms – drawing a large-scale timeline on how these concurrent processes took place in the history of growth of Tehran.

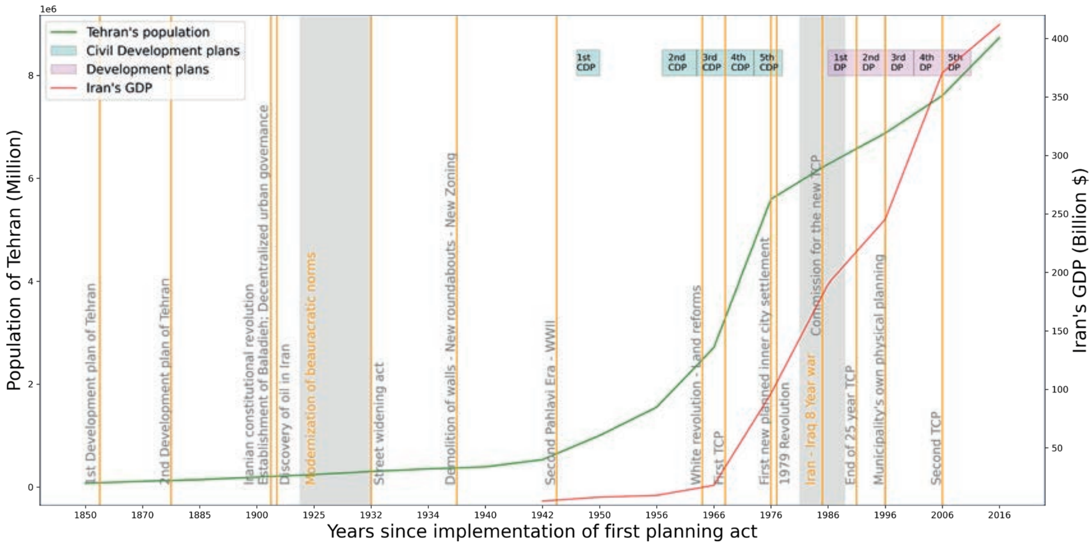

Figure 1 Summary of major political, economic and social timeline in Tehran.

The following subsections, briefly explain the general characteristics of this timeline highlighting major milestones creating periods of homogeneous planning and policy making periods.

2.3 Organic village – up to 1855

The historical development of Tehran, originally a subsidiary of the more prominent city of Rey, served as a refuge from hot climatic conditions. Emerging as an organically grown village with ties to Rey and strategic connections to main commercial routes, Tehran, along with neighbouring villages, formed a network of semi-self-sustaining systems with limited resources. In 1786, amidst internal conflicts and the dissolution of local governance, Aqa Mohammad Khan Qajar declared Tehran as the nation's capital, marking a pivotal transition. By 1855, Tehran had evolved into a walled city, fortified with six gates opening onto major trade routes, strategically positioned to facilitate national accessibility despite resource constraints, notably water scarcity. The city's initial layout exhibited characteristics typical of organic settlements, albeit with interventions and demolitions contributing to notable changes.

The walled city of Tehran at this juncture comprised a fortified royal citadel (Arg), interconnected commercial quarters (the Bazaar), and three distinct residential neighbourhoods (Oudlajan, Sangelaj, and Chal-e-meydan), reflecting both functional and social segregation. Six gates facilitated external access, while primary routes channelled movement and activity, particularly through the Bazaar, linking administrative quarters with the exterior. However, the constrained extent of the walled city hindered natural growth, necessitating informed planning initiatives to accommodate burgeoning populations and enhance infrastructure capacity. These planning efforts not only facilitated population influx but also catalysed new growth patterns.

2.4 First informed planning acts – 1855 to 1891

Upon establishment of the city as the capital of the newly unified nation, Tehran, benefited from its strategic location along the Silk Road, facilitating extensive regional connectivity conducive to effective governance. The city's initial layout reflected an organic growth pattern prior to assuming its capital status, with the walled and gated structure imparting a relatively regular shape. Retaining its organic structure post-inception as the capital, Tehran's gradual development delineated by main accessibility routes underscored the uniformity of its neighbourhoods, each sustaining its civic infrastructure.

August Křziž's 1857 mapping delineated four neighbourhoods within the walled city, each governed from the citadel. Upon solidifying its status as the nation's capital, Tehran experienced a surge in population growth and the demand for expanded governmental and civic infrastructure. Recognizing the restrictive nature of the city walls in accommodating this expansion, a comprehensive 12-year program was initiated in 1868, as outlined by Madanipour (2006), resulting in the transformation of Tehran into a walled symmetrical octagon with Twelve gates, spanning 19 km^2, aimed primarily at enhancing tax collection and regulating urban growth rather than fortifying against external threats. The demolition of the old walls ushered in a geometrically regular layout, replacing them with a grid system that expanded the urban fabric outward from the central old town while preserving the government quarter, Bazaar, and three old neighbourhoods. Planning decisions during this phase primarily focused on city-wide structural adjustments, with selective interventions to accommodate civic and state facilities.

2.5 Isolated planning before WWII – 1891 to 1948

This transformative phase witnessed the establishment of main movement corridors, diverging from the organic formation of thoroughfares, thereby altering traditional commercial routes and fostering competition with the established Bazaar. Concurrently, the introduction of modern governmental and public amenities addressed prevailing hygiene concerns. Local dynamics, governed by a municipal council, drove alterations in the urban fabric during this period, characterized by a national drive towards modernization culminating in the constitutional revolution of 1907. The establishment of a modern Municipality in 1910 signified a shift towards systematic governance, marking a departure from the city's organic growth trajectory. The subsequent political upheavals, including the overthrow of the Qajar dynasty in 1921 and the transition to a constitutional monarchy, heralded a new era of top-down planning and intervention, diverging from the previous bottom-up approach. The socio-political ambiance of the country was further disturbed by the impacts of the World Wars bringing internal political conflicts and famine to the country. However, by then end of the second World War, the second monarch of the Pahlavi Dynasty, supported by the US and the UK began to introduce systematic modernization programmes to the country.

2.6 Immense growth and major reforms – 1948 to 1966

The significant political transformations within the country had profound implications for Tehran's urban form. Stemming from both direct state intervention and shifts in governance structures, these changes were emblematic of the broader modernizing agenda of the newly established Pahlavi monarchy. The decision to dismantle the city walls, symbolizing a departure from traditional fortifications, and the centralization of governance marked pivotal shifts towards modernization (Zad, 2013). During the reign of the first Pahlavi monarch, a fervent drive towards modernization influenced urban planning decisions, characterized by street widening initiatives and the creation of expansive roundabouts, reminiscent of the Haussmanian approach .By 1936, these urban interventions, coupled with broader economic shifts, precipitated the emergence of a new modern social class, prompting demands for innovative urban and housing designs(Chehabi, 1993; Habibi, 2020).

This period heralded the inception of a new growth trajectory for Tehran, marked by the accumulation of employment centres and the imposition of standardized movement corridors, engendering a distinct pattern of expansion beyond the former city walls. This burgeoning modern metropolis accommodated newcomers, facilitated middle-class housing, and fostered grassroots industries. The post-World War II era witnessed a period of growth buoyed by the cessation of hostilities and economic support from the United States (Maloney, 2015, p. 54). However, this prosperity was soon overshadowed by economic instability, characterized by inflation and currency devaluation, exacerbated by political repression following the 1953 coup d'état (Behdad, 1992). These challenges precipitated widespread activism, culminating in the political reforms of 1963 (Maloney, 2015).

2.7 Political revolution – 1966 to 1979

The reforms enacted during this period, particularly the nationalization of agricultural land and the disruption of traditional Iranian feudalism, catalysed rapid rural-urban migration, fundamentally altering the urban landscape, notably in Tehran. Amirahmadi (1993) observes a surge in urban population during the 1960s, fuelled by the privatization of the construction sector and heightened housing demands. Tehran, in particular, experienced exponential growth, accommodating a significant portion of the nation's population while boasting substantial economic output and employment opportunities (Amirahmadi and Kiafar, 1993; Zad, 2013).

This demographic shift witnessed the gradual integration of outlying villages into the urban fabric, coinciding with initiatives to establish industrial zones, notably along the Tehran-Karaj corridor, as part of the initial five-year development plans. Despite a broader national policy aimed at decentralizing industrial hubs, the allure of this corridor continued to attract production units and population influx. The burgeoning capital's tumultuous growth, coupled with geopolitical considerations and alignment with Westernization agendas in response to communist expansion, spurred the formulation of long-term development strategies. Notably, the joint venture between Iranian firm Farmanfarmayan and American firm Victor Gruen culminated in the development of the first Tehran Comprehensive Plan (TCP) in 1968. This landmark plan, as highlighted by Madanipour (2006) identified key urban challenges such as high density, suburban sprawl, pollution, and inadequate infrastructure, attributed largely to rural-urban migration.

The TCP, spanning a 25-year horizon, advocated for linear vehicular development, multicentric zoning, and containment through a green belt to the south and west. Envisioning a city of 5.5 million residents within a 725 km² area, this pioneering plan offered pragmatic solutions rooted in linear logic, setting a precedent for urban growth and management.

2.8 Post revolution modernization – 1979 to present.

The implementation of the Tehran Comprehensive Plan (TCP), with a focus on vehicular movement and the establishment of industrial zones, spurred unprecedented physical growth in the city, leading to the integration of neighbouring villages and the emergence of modern settlements like Shahrak-e-Gharb. However, the plan failed to address the growing population, resulting in urban polarization and migration trends (Madanipour, 2006). The political upheaval culminating in the 1979 revolution disrupted the implementation of the TCP, ushering in a period of under-regulation and populist promises amidst socio-political changes. Subsequently, the informal growth intensified, exacerbated by crises such as war and economic embargoes, leading to demographic shifts and strain on resources (Zebardast, 2006).

Despite criticisms of the TCP's physical bias and association with the former regime's policies, efforts were made to renew planning strategies. A-Tech was commissioned for a new TCP in 1985, aimed at addressing the city's strategic issues, including resource shortages, urban growth pace, pollution, and transport inefficiencies (UPARC, 2000; Madanipour, 2006). The post-Gulf War era saw a resurgence in planning and development initiatives, with a renewed focus on managing growth and aligning with globalizing trends (UPARC, 2000). The ongoing renewal of the TCP since the early 2000s underscores the city's commitment to sustainable growth and adaptation to evolving urban challenges.

Table 1 summarizes the planning acts and programmes that were adopted in each of these periods of time. These can be generally categorized in zoning policies, infrastructure development and land use planning. While initially these took place as a part of isolated planning acts addressing singular issues, the planning system adopted centralized top-down planning during time with comprehensive outreach to different programming.

	Zoning policies	**Infrastructure development**	**Land use planning**
1855	N/A	Removal of the walls and replacing them with movement paths to reinforce commercial activity	Acupuncture interventions with civic establishments and government buildings
1891	Creation of first industrial zones + labour housing schemes	Modern municipality, introduction first top-down planning authority, Widening streets, new large roundabouts	Local changes, no overhaul New housing schemes (neighbourhood units)
1948	New housing neighbourhoods		
1966	First TCP, multi centric zoning. Solid hierarchy	Linear vehicular infrastructure grid, green belt	High density housing and industrial plants
1981	First new inner city planned neighbourhoods, 2 strategic plans	Focus on public transport	
Present	Rigid formal boundary of the city, green belt, certain economic and industrial regions, western regions to accommodate housing and env friendly functions	Tehran renovation organization to address the dilapidation through renewal of housing and infrastructure.	Peripheral to be designated to non-residential, new towns planned.

Table 1 Summary of planning elements in the six major socio-political periods of development in Tehran

The next section explains the dataset and methods used to create and use the models for the quantitative comparison between the soft evidence of planning styles inferred from literary review and unsupervised statistical models identifying urban structures, substructure and types.

3 DATASETS AND METHODS

As laid out, the question in this paper is framed between literary investigation and quantitaive analysis. From this point of view the composition of the city and validity of the types narrated through readings of history and policy is explored against unsupervised methods of classification. Given the theoretical underpinnings of how the city is represented in this project, the core analytical material comes from the spatial network model. The model consists of a segmented road-centre line model derived from Geofabrik outloet of OpenStreetMap (Geofabrik, 2018) constructed through a mitculous process of cleaning and segmenting (Kolovou *et al.*, 2017; Krenz, 2017), and running through DepthmapX (depthmapX development team, 2017) for angular segment analysis at local to global radii (400, 800, 1200, 2000, 5000, 10000, 20000, 40000m).

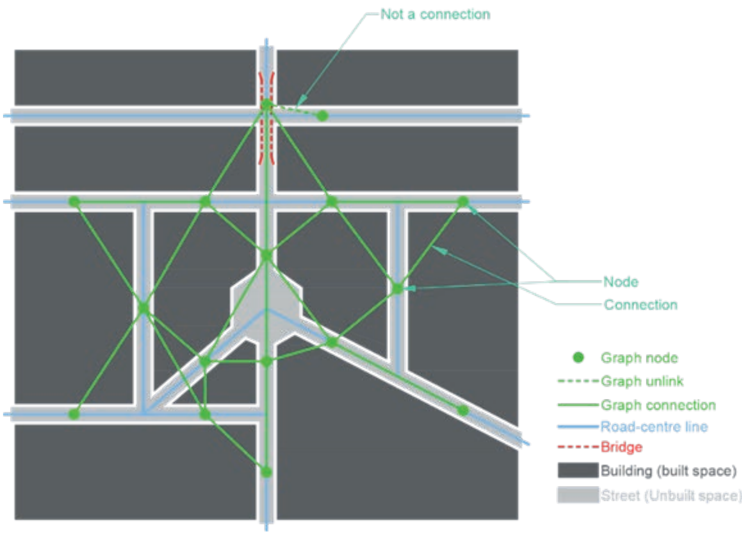

Figure 2 Dual graph constructed from simplified road-centre line model. Adapted from (Zhand, Gomes and Kyriazis, 2022)

This includes centrality measures as well as connectivity and node count – most specifically closeness centrality and betweenness centrality (Sabidussi, 1966; Freeman, 1977) – that measure the extent to which each segment is facilitating movement and activity, and thus relates to the higher order of relations and circles of engagements. The formula for closeness centrality, or integration as its pronounced in space syntax field is as follows:

$$AI_X = \frac{1}{n}\sum_{i=1}^{n} d_\theta(\pi X,\, i)$$

where n is the number of segments and d_θ is the angle between any two segments on the shortest path on segment X (Rashid, 2016, p. 66; Van Nes and Yamu, 2021). Respectively betweenness centrality, or *choice* in space syntax is described as:

$$C_i = \sum_J \sum_K g_{jk}^{(i)lg}\, g_{jk}^{(j<k)}$$

where $g_{jk}^{(i)}$ is the number of shortest paths between segment j and k containing i, and $g_{jk}^{(j<k)}$ is the number of all shortest paths between j and k (Rashid, 2016, p. 64; Van Nes and Yamu, 2021).

Out of the model of the current city – meaning the official extents of the city boundary – areas that correspond to the timeline of the city when major political, economic and/or social changes took place were identified based on literary review. Then each area was ran separately for their centrality measures through DepthmapX and for each added period, the annulus was remarked as the addition to the city between the two major changes in the socio-political overarching processes (See Figure 2). This shapes an index number of *planning style* that is assigned to the segments as well. It is important to note that this selection criteria and modelling is based on the premise that the city has grown in a relatively short period of time and mostly outward, and it can be argued that the conditions of each annulus has remained intact and is a good estimate of the claimed planning style.

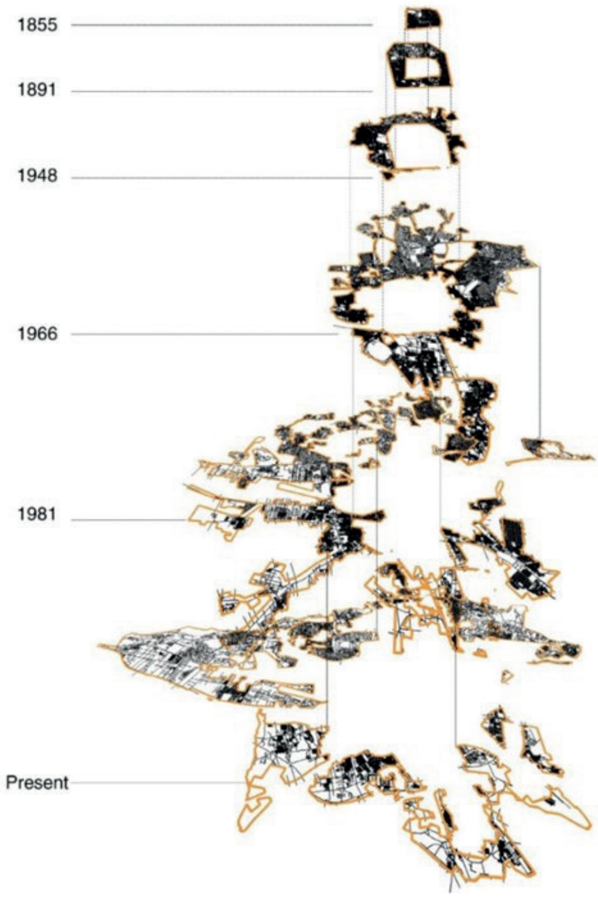

Figure 3 Annuli of the spatial model pertaining to the areas between pivotal moments in the planning history of the city.

Further the model is processed through an unsupervised clustering pipeline to determine whether there are distinguishable types captured through the spatial configuration (Hillier and Hanson, 1984). For this, initially the spatial model dataset – comprised of sixteen labels, integration and choice to the radii mentioned above, and around 600000 features – is processed for its principal component. This firstly to investigate whether the input labels all have considerable loadings to feed the clustering algorithm, and secondly to reduce the dimensions of the dataset to improve the computation of the model. The number of principle components are determined through calculating the eigenvalues (Chatelin, 2012) that use the eigenvectors in the data along which variation is maximized. In this project this is calculated using the Numpy API in python (Harris *et al.*, 2020).

Further the simplified and dimensionally reduced dataset is explored for its potential of identifiable clusters that correspond to distinguishable spatial figuration types. This is investigated in three different ways using clustering coefficient – or silhouette score - (Rousseeuw, 1987), spatial autocorrelation (Griffith, 2005) and Clustergram (Schonlau, 2002; Fleischmann, 2023). The reason for this exploration is to address the fact that such methods developed within computer science field are not fully explored in and in relation to spatial datasets and methods. Therefore, determining the number of clusters solely based on a single method would not explain the validity of clusters number, nor the stability of the clustering solution.

Finally, the association between the identified clusters and the planning styles identified through literature is assessed through Cramér's V test (Bergsma, 2013) which is a measure of association between nominal variables (the clustering index and planning style index). It is an extension of the chi-square test of independence and provides a way to quantify the degree of association between two categorical variables. It is calculated using the formula:

$$V = \sqrt{\frac{\chi^2}{n \cdot \min(k-1, r-1)}}$$

Where χ^2 is the chi-square statistic, n is the total number of observations, k is the number of categories in one variable and r is the number of categories in other variables. Ranging from 0 to 1, where 0 indicates no association and 1 represents a perfect association. The interpretation of Cramer's V (Akoglu, 2018) can vary depending on the field and context of the study. However, the following general guidelines are commonly used:

- A value close to 0 (e.g., Cramer's V < 0.1): This suggests a weak association between the variables.

- A value around 0.1 to 0.15: This indicates a moderate association between the variables.
- A value above 0.25 (e.g., Cramer's V > 0.25): This suggests a strong association between the variables.

These associations will visualized through a contingency map, where not only the association/overlap of difference indexes are illustrated, but also the overall association is also depicted.

4 RESULTS

Initially looking at the processed model, the pairwise correlation between all the labels shows that the normalized angular choice measures (on the left of the heatmap) do not correspond to most other labels. Interesting to see that regardless of the radii of analysis, the NACH is a very collinear measure, suggesting that it can be reduced to lower dimensions without losing much information about the model. (see Figure 3)

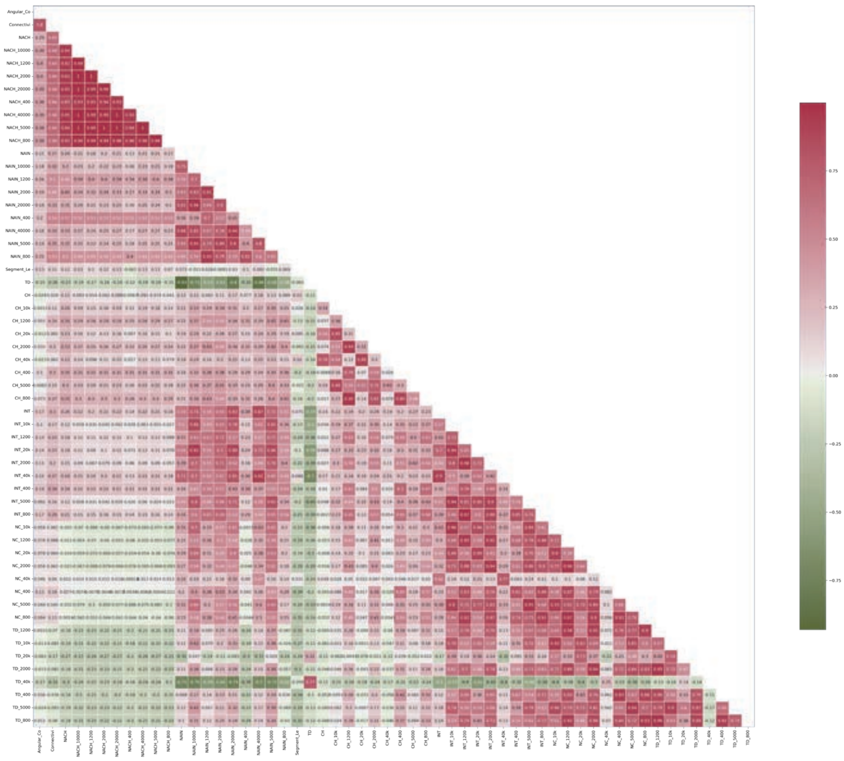

Figure 4 Pairwise correlation heatmap across all centrality metrics and radii.

Looking at this issue further, the colinearity of the normalized values is depicted through four corelations heatmaps across normalized and non-normalized values (See Figure 4). This re-asserts the same finding on how different normalized angular choice (NACH) and choice are in terms of colinearity. While NACH measures are almost identical across measures CH measures lose collinearity as the analysis radius increases. Similar observation with less intensity is valid between normalized angular integration and integrationl. This suggests that both INT and CH (the non-normalized values) are better representative of the multiscalarity of the spatial configuration conditions. Comparable to what Serra (2014) suggests in terms of pipeline for dimensionality reduction, the normalized values with higher multi-collinearity are more suited for PCA analysis, however the degree of collinearity in normalized values may lead to loss in information about the model.

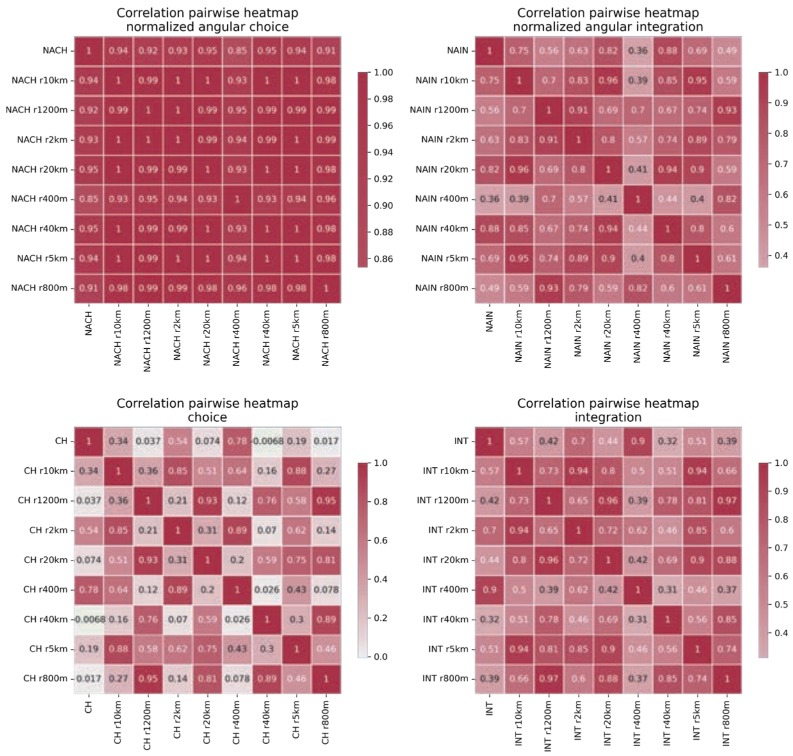

Figure 5 Pairwise correlation heatmap across all radii for four subsets of normalized and non-normalized choice and integration measures.

Evaluating this assumption each subset was analysed for their explained variance in reduced dimension, calculating the eigenvalue for the number of dimensions reduced from nine (the original number of input radii). This analysis suggests that at three components the highest capturing of the minimum number of components are produced without losing much information from the input labels (See Figure 5).

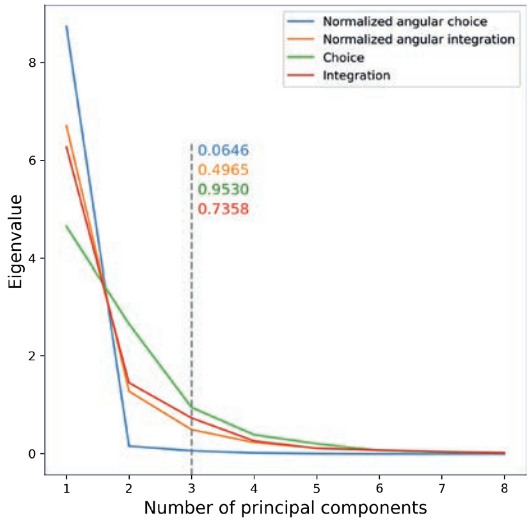

Figure 6 Explained variance (Eigenvalue) of each principal component across all subsets.

This is further examined by looking into the explained variance of the components, at reduced dimension of three. This analysis clearly shows that most of the NACH values are left almost unexplained at two of the three components, suggesting the use of NACH to be redundant as a part of the dataset. Similar observations to an extent are valid for Node Count (NC) and Total Depth (TD), while it is not true for NAIN, INT and CH values. On another hand the previous analysis showed that the NAIN values have a high degree of collinearity and at different radii, they do not provide additional information about the system. This wraps up the input selection from the model and concludes that the use of non-normalized values at local to global scale provides more meaningful variety comparing to other labels. Further, the fact that a single model is being used in the analysis makes the normalization irrelevant at this point.

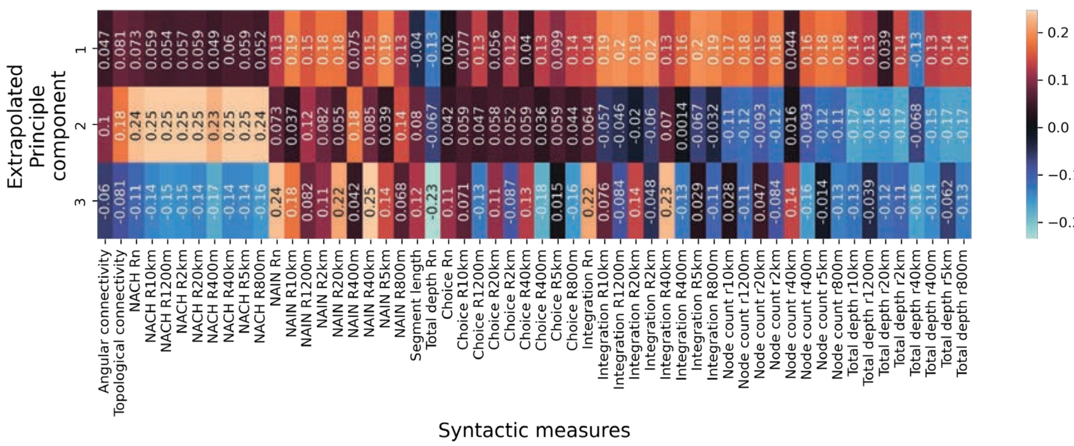

Figure 7 Explained Variance Ratio for all syntactic measures across all radii for 3 principal components.

The next step would be to determine the number of clusters that would translate to meaningful spatial configuration types. While there are numerous ways to determine the number of clusters within the data science domain, the process designed here is an iteration between application of methods, observation, evaluating through other methods and repeat. This is designed to make sure that the number of clusters selected is best suited for the research question. This begins with calculating the clustering coefficients for a rising number of clusters. This analysis shows that at 5 clusters highest explained variance with the minimum number of clusters is achieved (See Figure 7).

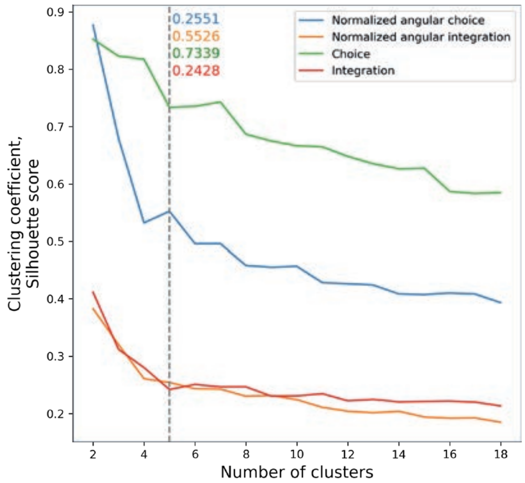

Figure 8 Clustering coefficient calculated for ascending number of clusters using KMeans clustering algorithm.

The plot with 5 clusters highlighted shows that there is a high degree of overlap between the identified types, while distinguishing marginally between the central areas, some peripheric segments, bulk of the planned city and two sets of foreground networks (See Figure 8)

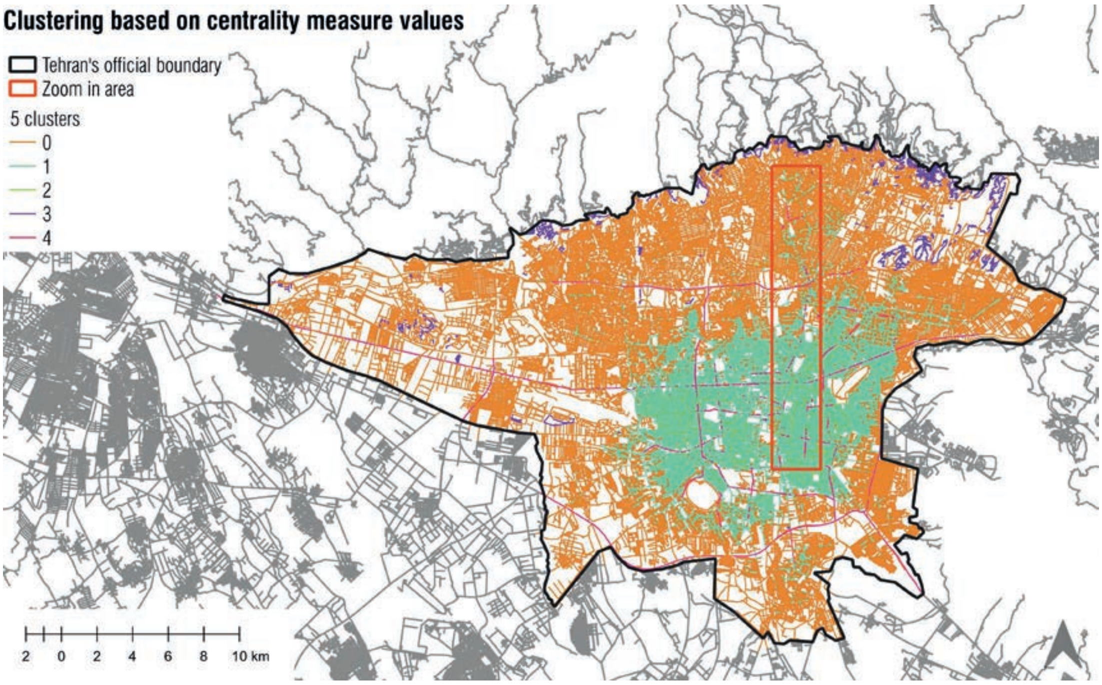

Figure 9 Identified clustered within the formal boundaries of the city of Tehran for K=5

Such large relativeness of the segments within the city – highlighting only two main background types – imply that a high degree of the segments are spatially autocorrelated, meaning that just through the virtue of juxtaposition they have similar configurational properties. While, the output of this process is mathematically sound, this degree of overlap and spatial autocorrelation within the clusters does not provide sufficient answers for the research question posed here. Therefore, the next step would be examining the degree of autocorrelation, as well as the stability of the clusters, to inform the optimum clustering solution.

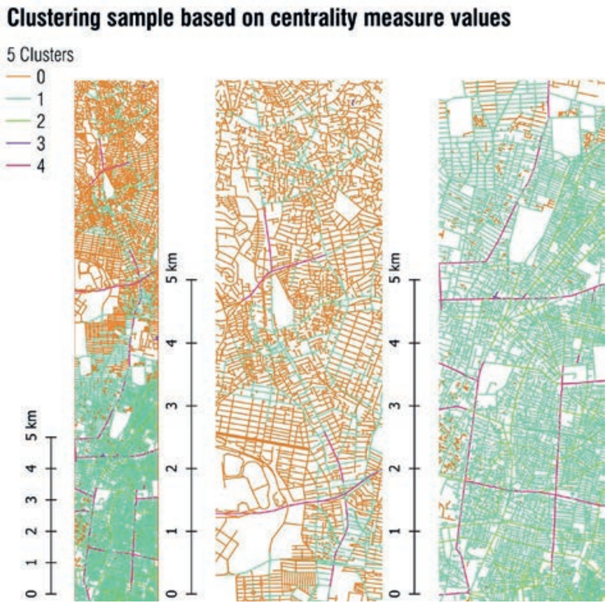

Figure 10 Closeup sample of clustering at K=5

Figure 10 shows the cross section across the five-clustering solution. This closeup illustration further shows the extent of overlap between the identified clusters which reasserts the idea that direct derivative of clustering coefficient analysis may not result in appropriate number of clusters that is required for a niche analysis such as this. Further examining this limitation, a further spatial autocorrelation analysis is performed to investigate the extent to which segments are identified as one cluster, due to their close proximity, and inherited spatial characteristics.

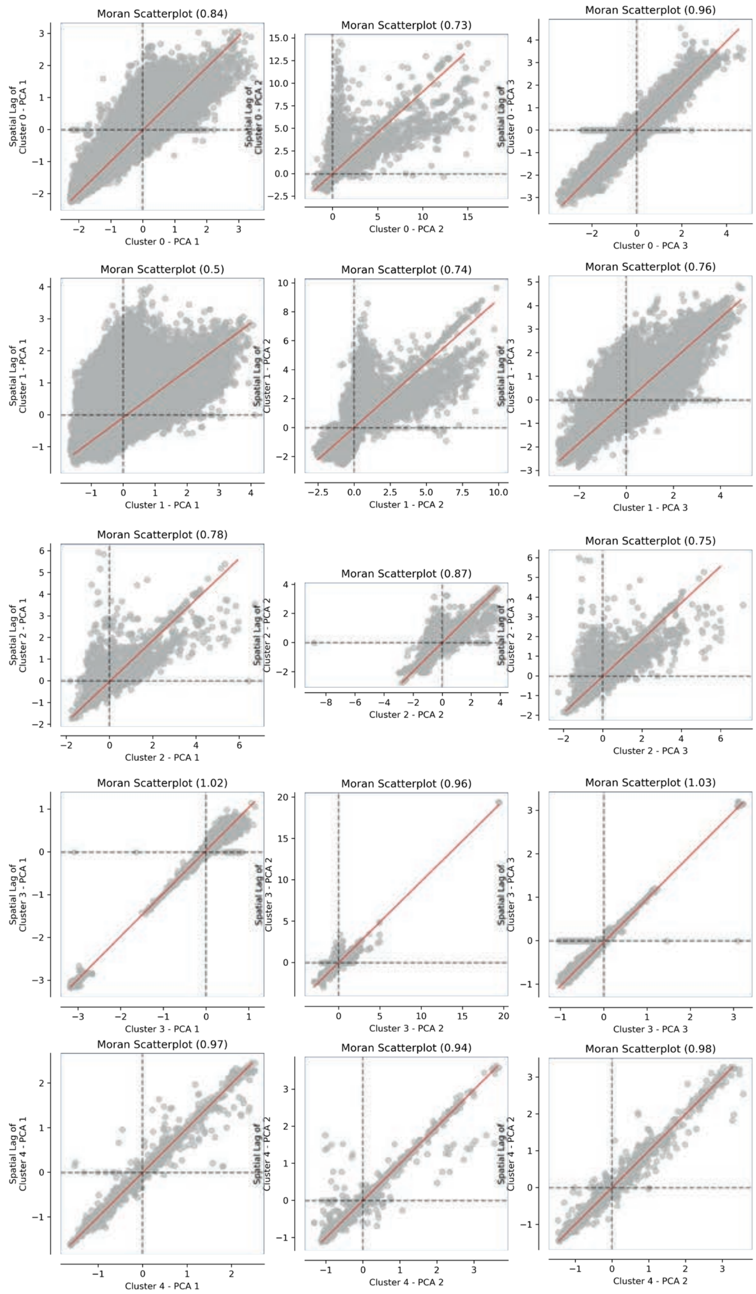

Figure 11 Moran's I spatial autocorrelation across the five clusters calculated for each principal component.

The application of spatial contiguity weights reveals that the two main clusters representing urban areas in the city do not adequately distinguish segments with minimal distance between them. This indicates significant overlap between distinct features within these clusters, suggesting that the spatial characteristics of each cluster are shared. Consequently, the chosen clustering input number (k=5) may not yield the optimal result despite previous considerations. While this doesn't invalidate the mathematical approach of determining the optimal number of clusters using the silhouette score, it prompts further exploration. Due to limitations in ground-truthing through iterative processes, the study delves into examining properties of higher numbers of clusters to potentially achieve more accurate results. This is done through clustergram tool, which illustrates how the clusters divide as the clustering number increases and accordingly shows how stable the clusters are in the process (See Figure 10)

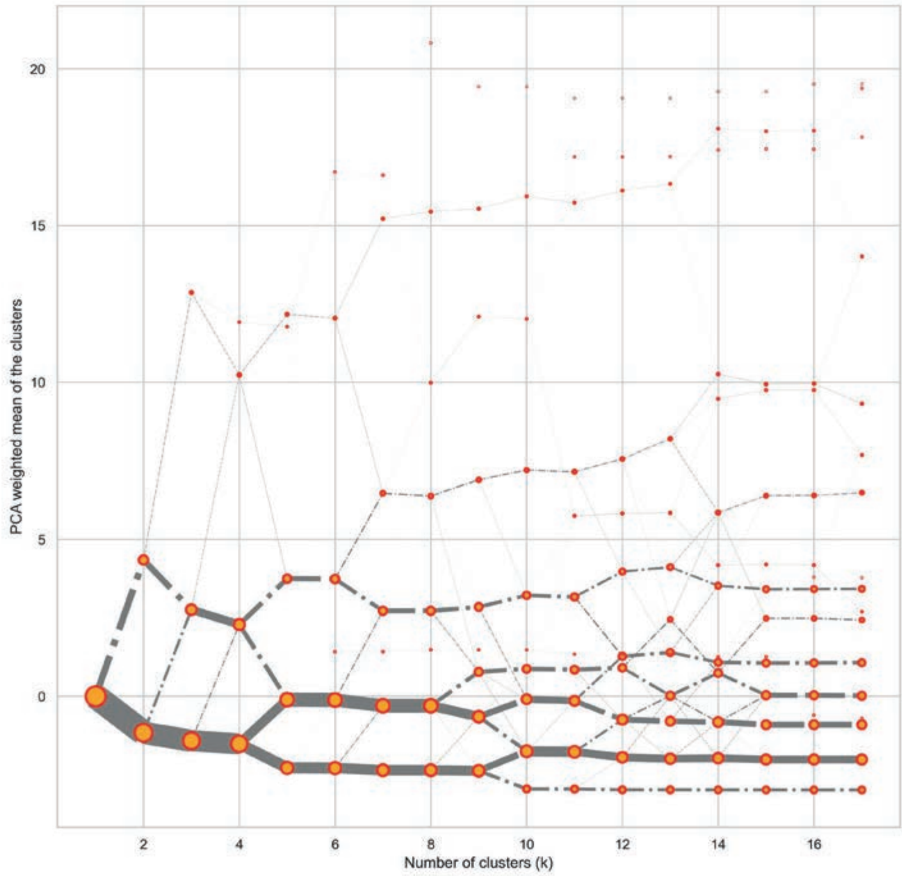

Figure 12 Visualization of PCA weighted mean of the clusters by order of increasing number of clusters.

The analysis highlights that the dataset lacks homogeneity, and as the number of clusters increases, the features may not consistently group together. Notably, meaningful breaks in clustering occur at K>2, K>4, K>8, K>11, and K>15, indicating that not all clusters are equally significant. For instance, between 5⩽K⩽8, there is minimal change in feature distribution, supporting the previous suggestion of an optimal cluster number between 4 and 6. This underscores that not all clusters require detailed examination. Additionally, using spatial configuration as input reveals distinct patterns at K= {3,5,9,12,16}. Figure 11 depicts the evolution of clusters through successive increments in K= {5,9,12,16}, with foreground clusters (coded in red) representing high-value segments for through-movement, and background clusters (coded in green) corresponding to to-movement segments. The analysis of segment proportions and cluster division enhances our understanding of determining an optimal cluster number. While specific clustering numbers remain valid, the analysis output's sufficiency depends on research criteria. The way foreground and background clusters divide can inform the ruling out of certain clustering numbers.

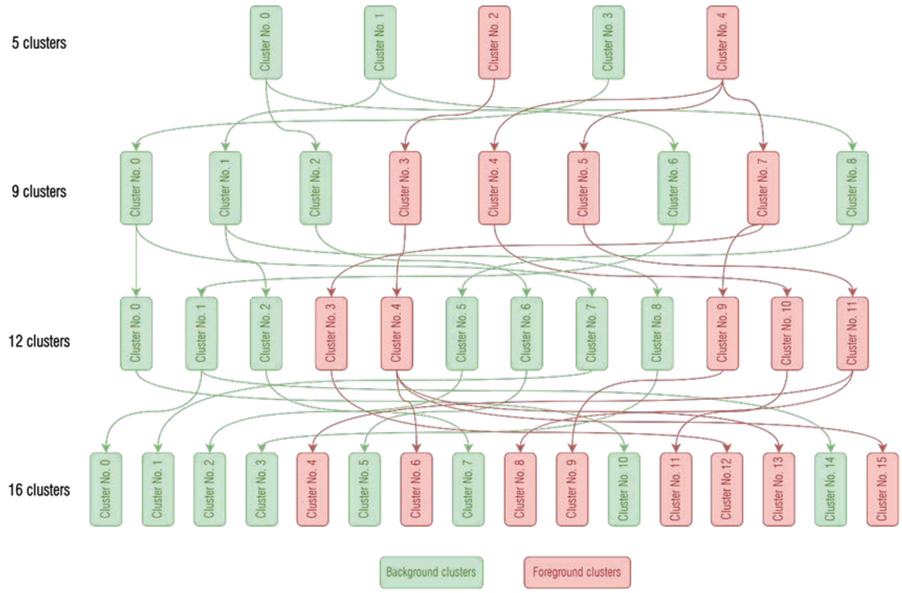

Figure 13 Cluster divergence in increasing clustering number.

With this depiction, at clustering of k= {12,16} there are *relatively* enough background clusters that would translate into significant spatial configuration types. However, between 12 and 16 clustering solutions there is not significant change in the division of background clusters and it’s the foreground clusters that divide into smaller clusters. This suggests that at k=12 clustering solution provides enough information that would translate into types. Given that at k=12 we have the minimum number of background clusters that are big enough to have a logical association with the identified planning styles identified previously.

Accordingly plotting the optimally stable clusters at k=12 (See Figure 12), it is evident that the background clusters shape – Cluster number= {0,1,2,5,6,7,8} – would somehow translate to annuli of the development extents previously discussed. With the same rationale these clusters are labelled by their most apparent attribute, in terms of their relationship to the overall system.

Figure 14 Dividing plots of extrapolated K means clusters at K=12

Table 1 summarizes these clusters with their most apparent trait in relation to other clusters as well as their significant configurational attribute that distinguishes them from the rest of the dataset. The table along with Figure 12 reassert the previous proposition about the stability of the foreground clusters, while offering an insight into the way in which outgrowth of the city with an imposed planning criteria will change the configurational functionality of the clusters. Clusters 10 and 11 are a good example of how the same criteria of vehicular based planned implemented with different intensity results in effectively different configurations. Further, the overall procedure of identifying clusters from a dataset of betweenness and closeness centrality measures, shows how the considerable difference between the order of values in integration and choice essentially denotes the functionality of the segments. While for certain segments through-movement is the main functionality – regardless of radii or metric of analysis – for others it's the to-movement, which highlights the significance of understanding the difference between the two concepts especially when studying the structure of a system that has outgrown.

Cluster No.	Representation	Appellation	Characteristic
Cluster No. 0	Background	Extra urban Natural corridors	Lower integration than the average city. little difference between the local and global integration values. Low node count.
Cluster No. 1	Background	Intra-urban villages	Close integration with the system at the smaller radii. Local functionality.
Cluster No. 2	Background	Planned modern development	Close subset of the overall system, with higher uniformity, and less variability comparing to the city
Cluster No. 3	Foreground	Intra-urban foreground	Significantly higher degree of integration across all radii with high variability
Cluster No. 4	Foreground	Transitory foreground	Higher integration than the average, with the difference decreasing as the radii increases
Cluster No. 5	Background	Transitory background	high integration values similar to cluster no. 4 with slightly lower averages.
Cluster No. 6	Background	Post modern planned city	Lower average integration values, and narrower distribution showing higher self-similarity
Cluster No. 7	Background	City margins	Significantly low integration values, with the difference decreasing as the radii increases.
Cluster No. 8	Background	Premodern planned neighbourhoods	Significantly higher average integration values, while showing homogeneity at smaller scales.
Cluster No. 9	Foreground	Imposed grid	High variance integration values working in between urban and metropolitan scales.
Cluster No. 10 and No. 11	Foreground	Metropolitan foreground	Cluster No. 10 has higher than average integration values, while No. 11 only has higher than average at metropolitan scales

Table 2 identified clusters at k=12

The examination of the relationship between planning style periods and a 12-cluster solution, utilizing Cramer's V coefficient yielding a value of 0.3460, denotes a robust association between these variables (See Figure 14). This indicates a substantial link between the nominal variables under consideration. Furthermore, the exceedingly low p-value of 0.0000 reinforces the statistical significance of this association, suggesting that the observed relationship is highly unlikely to occur randomly. Consequently, the null hypothesis is rejected, affirming a meaningful relationship between the 12-cluster solution and the areas highlighted by periodic study of city planning.

Furthermore, upon closer examination of specific associations within this analysis, nuanced insights into these relationships emerge. The heatmap reveals intricate connections between individual clusters and delineated temporal boundaries. While cluster No. {0,9,10,11} exhibits minimal correlation with policy periods, noteworthy associations are discerned among the remaining datasets, underscoring their significance. Particularly notable is the robust correlation identified between cluster No. 2 – designated as "The planned modern development" – and areas developed between 1948 and 1981. Additionally, a conspicuous linkage is observed between areas developed post the 1979 revolution and cluster No. 6 – termed "The post-modern planned city." Finally, a substantial connection is evident between areas developed preceding the white revolution and cluster No. 8 – labelled "The premodern planned neighbourhoods." These correlations between observed temporal boundaries and unsupervised identified built types imply a discernible influence of top-down planning interventions on urban development, resulting in distinct categories aligned with spatial configuration attributes.

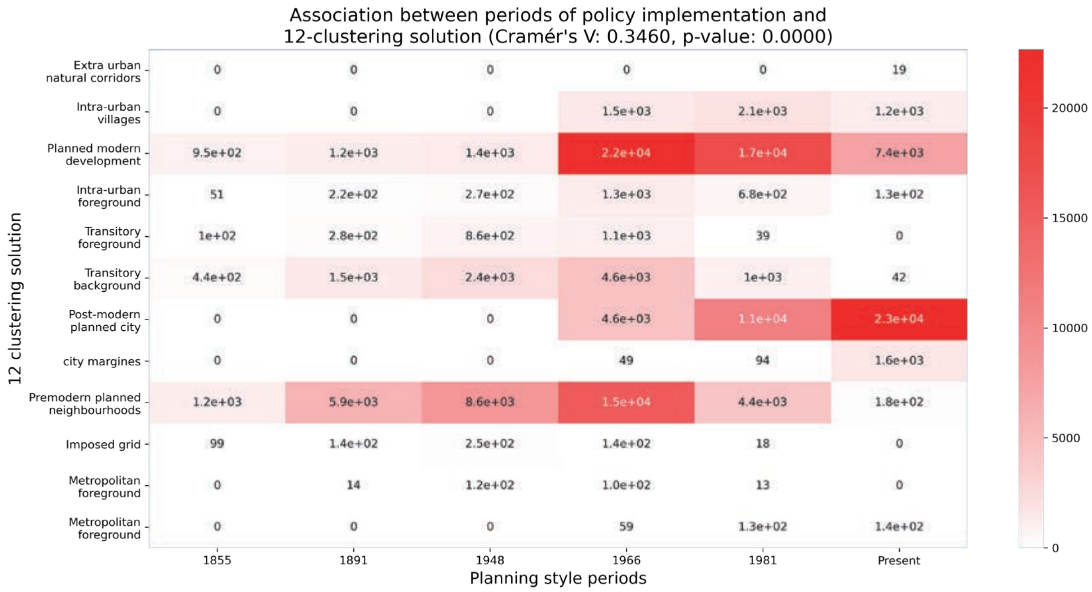

Figure 15 Cramér's V analysis heatmap for 12 clustering solution and periods of planning styles.

Furthermore, upon closer examination of specific associations within this analysis, nuanced insights into these relationships emerge. The heatmap reveals intricate connections between individual clusters and delineated temporal boundaries. While cluster No. {0,9,10,11} exhibits minimal correlation with policy periods, noteworthy associations are discerned among the remaining datasets, underscoring their significance. Particularly notable is the robust correlation identified between cluster No. 2 – designated as "The planned modern development" – and areas developed between 1948 and 1981. Additionally, a conspicuous linkage is observed between areas developed post the 1979 revolution and cluster No. 6 – termed "The post-modern planned city." Finally, a substantial connection is evident between areas developed preceding the white revolution and cluster No. 8 – labelled "The premodern planned neighbourhoods." These correlations between observed temporal boundaries and unsupervised identified built types imply a discernible influence of top-down planning interventions on urban development, resulting in distinct categories aligned with spatial configuration attributes.

5 CONCLUSIONS AND DISCUSSION

This paper emerges from the analytical discussion that tries to explain hard evidence from spatial analysis and soft evidence from urban studies. The framed gap between spatial and social phenomenon (Karimi, 2023) has been dealt is varied discipline dealing with urban studies and planning with documented accomplishments, yet this paper uses methods of computer science disciplines explore definitions of type, urban structure and sub-structure that permeate across the discourses. Accordingly, this exploration draws from the question that looks into the association between the overarching political, economic and cultural processes and the way in which the city is grown in types.

The premise of this question is that the city is a composition of varied structures and sub-structures that may or may not relate to an overarching driving force, and the city materializes to balance the top-down regulatory framework of planning and policy making and bottom-up grass root movement and population growth. However, due to the level of complexity of urban phenomenon it may not be realistically feasible to relate a certain type – or structure – to a singular or a group of driving forces. However, pinning down the major periods of simultaneous conditions may facilitate the idea and explain the associations.

Therefore, this paper looked into the general history of development and growth in the city of Tehran – based on the idea of the city as a composition of types – and highlighted periods of growth, planning and policy making, that informed the development of the city from a small village into a multi-million resident megapolis. This investigation is based on the fact that this process happened through a relatively short period of time, growing outward and with little to no major natural disaster changing the physical arrangement of the city. Although throughout the time certain development policies were undertaken in the city, there were seen as a part of the development and the modern interventions in the old fabric were seen as a substructure itself.

Based on this literary overview that reported six major political, social, cultural, and economic periods that had impact throughout the city and the country. On the other hand, the research used unsupervised analytical methods based on open data and toolkits to exercise the validity of typological distinctiveness. Therefore, using a spatial network model and centrality measures – as outlined and used through space syntax – a dataset was created which correlated with movement and activity at different radii of movement. This was done with the premise that the spatial configuration model captures social structures that resulted from the mentioned overarching dynamics. Therefore, the core task in the research was to find the right number of structures – or clusters in this case – that are concurrently distinct at different radii and correlate to the *planning styles* that governed each of the model's annuli.

Through in iterative process it was shown that the tools need a level of interpretation and direct result from analytical methods – e.g., silhouette score, spatial autocorrelation etc – may not be applicable to the model, yet through the process a meaningful number of clusters emphasized by the unsupervised tools was reached to be used in the clustering analysis. The results showed that there is a significant and robust association between the planning styles highlighted by literature and the clusters identified by the statistical model. This further reasserts the validity of the research question on the relationship between social and spatial structures and gives an example of how the two are shaped in a reciprocal manner.

It can be further argued that the strong association between the two validates the method in observing urban structures yet highlights the question on how the number of clusters should be achieved and against what variables can they be assessed. While providing some examples on how the purely mathematical models may not address the research questions directly, the overall process shows that through the use of spatial configuration models and unsupervised learning mechanisms, urban types at large scale can be extrapolated.

REFERENCES

Akoglu, H. (2018) 'User's guide to correlation coefficients', *Turkish Journal of Emergency Medicine*, 18(3), pp. 91–93. Available at: https://doi.org/10.1016/j.tjem.2018.08.001.

Alexander, C. (2012) '"A City is Not a Tree": from Architectural Forum (1965)', in *The Urban Design Reader*. 2nd edn. Routledge.

Amirahmadi, H. and Kiafar, A. (1993) 'The Transformation of Tehran from a Garrison Town to a Primate City: A Tale of Rapid Growth and Uneven Development', in *Urban Development in the Muslim World*. Routledge.

Behdad, S. (1992) 'Massoud Karshenas, Oil, State and Industrialization in Iran (Cambridge: Cambridge University Press, 1990). Pp. 308.', *International Journal of Middle East Studies*. 2009/01/29 edn, 24(4), pp. 702–704. Available at: https://doi.org/10.1017/S0020743800022534.

Bergsma, W. (2013) 'A bias-correction for Cramér's V and Tschuprow's T', *Journal of the Korean Statistical Society*, 42(3), pp. 323–328. Available at: https://doi.org/10.1016/j.jkss.2012.10.002.

Bettencourt, L.M. (2013) 'The kind of problem a city is: new perspectives on the nature of cities from complex systems theory', *Santa Fe, NM:: Santa Fe Institute* [Preprint].

Chatelin, F. (2012) *Eigenvalues of Matrices: Revised Edition*. SIAM.

Chehabi, H.E. (1993) 'Staging the emperor's new clothes: dress codes and nation-building under Reza Shah', *Iranian Studies*, 26(3–4), pp. 209–233. Available at: https://doi.org/10.1080/00210869308701800.

depthmapX development team (2017) ‘depthmapX’. Available at: https://github.com/SpaceGroupUCL/depthmapX/.

Fiori, J. (2019) ‘The city as a political construct: Rules, Exceptions and Insurgencies’. Available at: https://www.mixcloud.com/dpuucl/dpu-breakfast-jorge-fiori-the-city-as-a-political-construct-rules-exceptions-and-insurgencies/.

Fleischmann, M. (2023) ‘Clustergram: Visualization and diagnostics for cluster analysis’, *Journal of Open Source Software*, 8(89), p. 5240. Available at: https://doi.org/10.21105/joss.05240.

Freeman, L.C. (1977) ‘A Set of Measures of Centrality Based on Betweenness’, *Sociometry*, 40(1), pp. 35–41. Available at: https://doi.org/10.2307/3033543.

Geofabrik (2018) ‘OpenStreetMap data’. Available at: https://download.geofabrik.de/asia/iran.html (Accessed: 6 August 2018).

Griffith, D.A. (2005) ‘Spatial Autocorrelation’, in K. Kempf-Leonard (ed.) *Encyclopedia of Social Measurement*. New York: Elsevier, pp. 581–590. Available at: https://doi.org/10.1016/B0-12-369398-5/00334-0.

Habibi, R. (2020) *Modern Middle-Class Housing in Tehran: Reproduction of an Archetype: Episodes of Urbanism 1945–1979*. Leiden, The Netherlands: Brill. Available at: https://doi.org/10.1163/9789004443709.

Harris, C.R. *et al.* (2020) ‘Array programming with NumPy’, *Nature*, 585(7825), pp. 357–362. Available at: https://doi.org/10.1038/s41586-020-2649-2.

Hillier, B. and Hanson, J. (1984) *The Social Logic of Space*. Cambridge: Cambridge University Press. Available at: https://doi.org/10.1017/CBO9780511597237.

Jacobs, J. (2011) *The Death and Life of Great American Cities*. Available at: https://www.socialcapitalgateway.org/content/book/jacobs-j-1961-death-and-life-great-american-cities-new-york-random-house (Accessed: 30 August 2023).

Karimi, K. (2012) ‘A configurational approach to analytical urban design: “Space syntax” methodology’, *URBAN DESIGN International*, 17(4), pp. 297–318. Available at: https://doi.org/10.1057/udi.2012.19.

Karimi, K. (2023) ‘The Configurational Structures of Social Spaces: Space Syntax and Urban Morphology in the Context of Analytical, Evidence-Based Design’, *Land*, 12(11). Available at: https://doi.org/10.3390/land12112084.

Kolovou, I. *et al.* (2017) ‘Road centre line simplification principles for angular segment analysis’, in *Proceedings-11th International Space Syntax Symposium, SSS 2017*. Instituto Superior Técnico, Portugal, pp. 163–1. Available at: https://discovery.ucl.ac.uk/id/eprint/10042788/ (Accessed: 15 November 2023).

Krenz, K. (2017) ‘Employing volunteered geographic information in space syntax analysis’, in *Proceedings-11th International Space Syntax Symposium, SSS 2017*. Instituto Superior Técnico, Departamento de Engenharia Civil, pp. 150–1.

Lynch, K. (1964) *The Image of the City*. MIT Press.

Madanipour, A. (2006) ‘Urban planning and development in Tehran’, *Cities*, 23(6), pp. 433–438. Available at: https://doi.org/10.1016/j.cities.2006.08.002.

Maloney, S. (ed.) (2015) ‘The Prelude: The Political Economy of Prerevolutionary Iran’, in *Iran’s Political Economy since the Revolution*. Cambridge: Cambridge University Press, pp. 19–72. Available at: https://doi.org/10.1017/CBO9781139023276.002.

McFarlane, C. (2011) ‘Assemblage and critical urbanism’, *City*, 15(2), pp. 204–224. Available at: https://doi.org/10.1080/13604813.2011.568715.

Mumford, L. (1937) ‘What is a City?’, *Architectural Record*, 82(5), pp. 59–62.

Rashid, M. (2016) *The Geometry of Urban Layouts: A Global Comparative Study*. Springer International Publishing. Available at: https://books.google.co.uk/books?id=Iuh6DAAAQBAJ.

Rousseeuw, P.J. (1987) ‘Silhouettes: A graphical aid to the interpretation and validation of cluster analysis’, *Journal of Computational and Applied Mathematics*, 20, pp. 53–65. Available at: https://doi.org/10.1016/0377-0427(87)90125-7.

Rowe, C. and Koetter, F. (1984) *Collage City*. MIT Press.

Sabidussi, G. (1966) ‘The centrality index of a graph’, *Psychometrika*, 31(4), pp. 581–603. Available at: https://doi.org/10.1007/BF02289527.

Schonlau, M. (2002) ‘The Clustergram: A Graph for Visualizing Hierarchical and Nonhierarchical Cluster Analyses’, *The Stata Journal*, 2(4), pp. 391–402. Available at: https://doi.org/10.1177/1536867X0200200405.

Serra, M.L.L.A. (2014) *Anatomy of an Emerging Metropolitan Territory-Towards an Integrated Analytical Framework for Metropolitan Morphology*. Universidade do Porto (Portugal).

UPARC (2000) *Tehran’s Metropolitan Region and its Satellite Cities Plan (Summary)*. Tehran: Urban Planning and Architecture Research Centre of Iran.

Van Nes, A. and Yamu, C. (2021) *Introduction to Space Syntax in Urban Studies*. Cham: Springer International Publishing. Available at: https://doi.org/10.1007/978-3-030-59140-3.

Varzi, A.C. (2021) ‘What is a City?’, *Topoi*, 40(2), pp. 399–408. Available at: https://doi.org/10.1007/s11245-019-09647-4.

Whyte, W.H. (1980) *The Social Life of Small Urban Spaces*. Conservation Foundation.

Zad, V.V. (2013) ‘Spatial Discrimination in Tehran’s Modern Urban Planning 1906–1979’, *Journal of Planning History*, 12(1), pp. 49–62. Available at: https://doi.org/10.1177/1538513212459717.

Zebardast, E. (2006) ‘Marginalization of the urban poor and the expansion of the spontaneous settlements on the Tehran metropolitan fringe’, *Urbanization and the Iranian Revolution*, 23(6), pp. 439–454. Available at: https://doi.org/10.1016/j.cities.2006.07.001.

Zhand, S., Gomes, A. and Kyriazis, A. (2022) ‘Investigating what connectivity means to different social groups in a fast-growing city-The case of Abu Dhabi’.