Armanshahr Architecture & Urban Development, 11(22), 81-94, Spring 2018

ISSN: 2008-5079

EISSN: 2538-2365

Understanding Temporal Human Mobility Patterns in a City by Mobile Cellular Data Mining, Case Study: Tehran City

Abbas Azari[1], Mehdi Mirmoini[2*] and Shadi Mohammadi Oujan[3]

[1] Ph.D. Candidate in Urbanism, Nazar Research Center, Tehran, Iran.

[2] M.A. of Urban Planning, School of Architecture and Environmental Design, Iran University of Science and Technology, Tehran, Iran.

[3] M.A. of Urban Planning, School of Architecture and Environmental Design, Iran University of Science and Technology, Tehran, Iran.

Received 13 March 2018; Revised 22 April 2018; Accepted 13 May 2018

ABSTRACT: Recent studies have shown that urban complex behaviors like human mobility should be examined by newer and smarter methods. The ubiquitous use of mobile phones and other smart communication devices helps us use a bigger amount of data that can be browsed by the hours of the day, the days of the week, geographic area, meteorological conditions, and so on. In this article, mobile cellular data mining is introduced as an emerging approach in analyzing and understanding human mobility patterns, then generic location update is examined as a way to observe and perceive human mobility and movement in cities. This method was examined in Tehran metropolitan area map, the results show that different urban issues can be understood and solved using this huge amount of data like urban transportation, social problems or urban functions. Tehran cellular data analysis shows that it can be recognized as a city in two major parts, the border zone which is mostly the origin of all trips and the central zone which is mostly the destination of all trips and the most visited hotspot of the city during a normal day, also it was concluded that because of low population density in this part of the city and very high human mobility throughout a day, this area would have many social security issues. In the end, taking advantage of more accurate data in cell level was proposed in order to have better and more reliable assumptions about future mobility trends and co-presence patterns.

Keywords: Mobile Cellular Data, Complexity, Spatial Data Mining, Location Update, Human Mobility Patterns.

INTRODUCTION

Cities around the world are growing rapidly and extending their boundaries. This development is at such a rate that by 2050 about 70% (66 percent, estimated by United Nations in World Urbanization Prospects report) of the global population is expected to be living in cities (United Nations, 2014). Cities are considered as complex systems and their rampant growth adds to this complexity, therefore providing comprehensive and systematic approach is essential to create and manage cities with good qualities. As a result of having more complex and dynamic cities, decision and policy making to enhance mobility and movement in cities has become nearly impossible while we are facing new and different problems like, air pollution, social issues, traffic congestion, lack of parking spots, and etc. These problems will be getting worse because the number of vehicles on the road is expected to double to around 2.5 billion by 2050 globally, estimated by The International Energy Agency (IEA, 2012), and the number of world population is increasing radically.

In smart cities, data is the fundamental resource to manage and monitor human mobility and behavior. These data are gathered with various methods. Sensors attached to vehicles or in streets, GPS trackers installed on vehicles, social Media, mobile phones and CC TVs are all tools to generate this kind of data. City planners can

* Corresponding Author Email: p.mirmoini@gmail.com

have accurate estimations on cities' behavior by mining and analyzing these data e.g. predicting different traffic behavioral patterns, finding a way to decrease traffic congestion, locating a new metro station, providing the optimal land-use zoning, managing natural disasters and urban crisis's, understanding co-presence patterns and social interactions and etc. But in smart cities the volume of this data is huge, so they are called "big data". Data mining and analyzing these big data is a process that needs particular methods and tools to maneuver; thus, using data mining to provide a systematic analysis with regard to the complexity of human mobility issues needs the same procedures.

This article is focused on showing the complexity of human behavior in a city and special applications of information and communication technologies in understanding human mobility patterns and urban planning issues associated with this phenomenon, then it will examine big data and data mining and how they can contribute to making better plans and policies in urban contexts. In the end it introduces new applicable methods using cellular mobile network data to understand human mobility patterns and estimate future trends of urban population.

RESEARCH METHODOLOGY

The purpose of this study is to clarify the applications of mobile cellular data mining in understanding temporal human mobility in cities, so after understanding how cellular mobile network operates and some examples of how this huge amount of data can be mined to help solve urban issues, location update cellular data for a major operator in Iran was obtained for further analyzes. Using Geographical Information System, this huge amount of data is mined and mapped in Tehran metropolitan area with a quantitative method as a case study, then the outputs are analyzed qualitatively to connect this information with urban issues like transportation, social problems and urban functions. Generally, the research methodology for this article is quantitative and qualitative; also spatial data mining is analyzed on Tehran city as a case study. Obviously obtaining this data for a Metropolitan area like Tehran is a hard task but fortunately we were able to get the intended data for our research from a major mobile operator in Iran. We hope this work will be helpful for authorities to understand how sharing data can contribute in making cities more livable by having reliable researches and therefore more proper and accurate planning for our cities based on these real-time data.

LITERATURE REVIEW

The topic of using mobile phones in urban planning was a relatively recent debate over the last few years; but in recent years, the large deployment of mobile and wireless technologies has provided new means to understand the dynamics of a city; which has happened for the first time in Real-time Rome project as Ratti et al. (2007) pointed out. The collaboration between Telecom Italia and MIT's Sensible City Laboratory in 2007 explored how researchers might be able to use mobile phone data for an entire metropolitan region to analyze urban dynamics (Ratti et al., 2007). In earlier research on human mobility patterns, Gonzalez and her team observed the trajectory of 100,000 anonymized mobile phone users whose positions were tracked for a six-month period that described the recurrence and temporal periodicity inherent to human mobility (Gonzalez et al., 2008).

In 2009, Kyunghan Lee studied a mobility model for human walks by proposing a new mobility model called SLAW to improve the performance of networking applications and understanding human mobility patterns (Lee et al., 2009). Moreover, a recent study in Northeastern University has used cell phone billing data for 50,000 people in a European country to show that people's travel patterns are extremely predictable (Barabasi et al., 2010). There have been other considerable studies done to explore human mobility and improve the spatial resolution of static population which is counted through the use of census data (Dobson et al., 2000; Bhaduri et al., 2007).

In addition to the facts stated above, in a document written by United Nations Global Pulse, other usages of mobile phone network data were introduced and summarized as: Disaster response, Health, Socio-economics and Transportation (United Nations Global Pulse, 2013). Finally, Google Maps can be referred to as the latest use of mobile phone network, which is nowadays handled by almost all people (http://www.theconnectivist.com).

COMPLEXITY & SYSTEMS THEORY IN CITIES

Half a century or more ago, cities were first formally considered as 'systems' which were defined as distinct collections of interacting entities, usually in equilibrium, but with explicit functions that could enable their control often in analogy to processes of their planning and management. These conceptions treated cities as organized from the top down, distinct from their wider environment (Batty, 2011, p. 1). But as soon as this model was articulated, it was found that cities are more like biological than mechanical systems and the rise of the sciences of complexity, has changed the direction of systems theory from top down to bottom up (Portugali, 2000, p. 66). As a result nowadays, Complexity Theories of Cities (CTC) demonstrate that its urban simulation models can explain the same set of urban phenomena as properties of open and complex systems (Portugali, 2011, p. 105).

Our understanding of cities is being transformed by new approaches from the complexity sciences (Batty, 2011, p. 1). CTC which was a narrow stream of studies two and a half decades ago – written mainly by physicists applying theories from physics – has now become not a flood but an established interdisciplinary research domain engaging urban geographers, planners, urban designers, regional scientists, mathematicians, physicists among others (Portugali, 2009, p. 1). They are promising examples as to how the concepts and methods of mathematics, physics and more generally natural science can be employed in order to achieve a deeper insight into some aspects of the complexity of urban processes (Albeverio et al., 2008).

Furthermore, as Batty argues in his book The New Science of Cities, the various processes that take place in cities, which bring people together to produce and exchange goods and ideas, define a multitude of networks that enable populations to deliver materials and information to support such endeavors. Physical and social networks tend to mutually reinforce one another as they develop (Batty, 2013, p. 48).

Human mobility is one of the urban issues that should be dealt with as a complex phenomenon. Therefore, new models and technologies should be adopted to understand, analyze and forecast this ever-changing dynamic pattern in real time. That is where the emergence of data mining and the use of big data like mobile cellular data contribute to the process of solving the issues of an urban population.

BIG DATA & DATA MINING

Tiny, inexpensive sensors and tags of varied kinds are increasingly being mounted on buildings and infrastructure, carried in moving vehicles, integrated with wireless mobile devices such as telephones, and attached to products. These sensors are harvesting enormous streams of data that can be mined, by means of sophisticated software, to generate detailed, real-time pictures of evolving human activities, needs, and demands within the buildings and neighborhoods of cities (Mitchell & Casalegno, 2008, p. 97). As a result, each of us has become a walking data generator as mobile phones, social networks, package tracking, bank card payments, Global Navigation Satellite Systems (GNSS), public transport ticketing etc. produce torrents of data as a by-product of their operations, generating daily a bit more than 1GB of content per person.

This data production seems to be doubled every couple of years as more data was stored across the internet every second in these days which is more than the entire internet compared to just two decades ago (Semanjski, 2016, p. 2) thus, There is a coincidence between what is now called smart cities and big data, with smartness in cities pertaining primarily to the ways in which sensors can generate new data streams in real time with precise geo-positioning, and how the data bases that are subsequently generated can be integrated so that value can be added (Batty et al., 2012, p. 488).

Characteristics of Big Data are:

- **Volume**: Volume refers to amount of data.
- **Variety:** Variety refers to the data from multiple sources in the form of both structured and unstructured format. Data source is diverse. Gartner mentioned it as dark data. Similar to dark matter in physics, dark data cannot be seen directly, yet it is the bulk of the organizational universe.
- **Velocity:** Velocity deals with the rate at which data flows in from various sources. Big Data flow is massive and continuous.
- **Veracity:** Veracity refers to uncertainty of data. It deals to quality, trustworthiness, accuracy of data (Kumar & Prakash, 2016, p. 12).

The process of examining & analyzing these huge amount of everyday data is seemingly called data mining that helps us improve the way of living especially by enhanced mobility. As Batty et al. describe, by using data mining and network analysis, we are able to create an urban mobility atlas, i.e., a comprehensive catalogue of the mobility behaviors in a city, an atlas that can be browsed (by the hours of the day, the days of the week, geographic area, meteorological conditions, and so on) in order to explore the pulse of a city in varying circumstances, while also observing emerging deviations from normal. To fully realize the idea of an urban mobility atlas for the smart city, there is a need to integrate increasingly richer sources of mobility data, including the data from public transportation, road sensors, surveys and official statistics, social media and participatory sensing, into coherently integrated databases, as well as connecting mobility with socio-economic networks (Batty et al., 2012, p. 490).

THE APPLICATION OF CELLULAR PHONE DATA IN UNDERSTANDING TEMPORAL HUMAN MOBILITY PATTERNS

This study has shown that to overcome the complexity and dynamics of human mobility issues, using new technologies and ideas is inevitable. In the article, review of traffic data estimations extracted from cellular networks, Caceres et al. (2008) mentioned that recent market surveys show that cellular phone penetration levels are reaching to 100% of the population in a large number of countries. So, it is safe to assume that each vehicle will carry at least one phone (Caceres et al., 2008, p. 179). As a result, this huge amount of data extracted from cellular phones, could be used to understand and analyze different attributes of urban population. By mining the big data of cellular networks and with the help of new applications and tools we can obtain accurate road traffic information, estimate traffic congestion, differentiate modes of transportation, observe the flow of traffic, estimate the speed & travel time, identify traffic hotspots, predict human mobility in crises times like earthquake and etc. road traffic could also be analyzed by static sensors and collected from floating car data or GPS technology but these methods have limitations and are not as accurate as mobile cellular data. As Elias et al. (2016) stated, the advantage of such data is that it can be extracted on a national level and therefore it is possible to see connections between different regions at an optimized level of detail. While this is possible on a daily base, it is not possible to track the same user for more than one day, since the user´s identity number changes every day for security reasons (Elias et al., 2016, p. 4479).

Due to the complexity of mobile cellular networks, big data in mobile cellular networks also exhibits several other unique characteristics, which lead to unprecedented challenges as well as opportunities. For instance, to understand behaviors and requirements of mobile users, which in turn allow the intelligent decision making for real-time decision making in various applications (Zhang et al., 2016, p. 2).

In addition to the facts stated above, in a document written by United Nations Global Pulse, other usages of mobile phone network data were introduced. This document examines some of the issues resolved by this kind of data in real life and categorizes them into four different categories while giving some examples for each of them, they are summarized as follows:

• **Disaster response:** Identifying the mobility patterns and population migration during Haiti earthquake to manage the refugees;

• **Health:** Mapping Malaria in Kenya and analyzing the regional travel patterns of million mobile subscribers to map the specific locations to which malaria had a higher probability of spreading;

• **Socio-economics:** Understanding socio-economic indicators in the UK, estimating poverty levels in Cote D'Ivoire with the assumption that higher levels of mobile communications and wider range of calls are a proxy indicator for prosperity;

• **Transportation:** Optimizing transport networks in Abidjan by monitoring citizens' travel routes (United Nations Global Pulse, 2013).

So, it is clear that data collected from cellular networks play an important role in urban planning and management. In this part of the article, first the mobile communication system and its application is described, then the methods of data collection and their usage are examined and finally, data collected from one of major mobile operators in Iran in Tehran Metropolitan Area will be mined and mapped for further analyses.

Mobile Communication System

The general concept of public system architectures for mobile communication systems is a cellular network concept. The cellular networks are built based upon a number of small geographical areas, cells, covering a larger area. Each cell consists of a base station transmitting the information within the limited area. For analytical purposes these areas are usually hexagonal cells, which totally covers an area and are rather similar to real transmission patterns. The figure bellow shows a road segment and the cells covering the area as an example of how a road segment crosses several cells in its path (Gundlegard & Karlsson, 2009, p. 5).

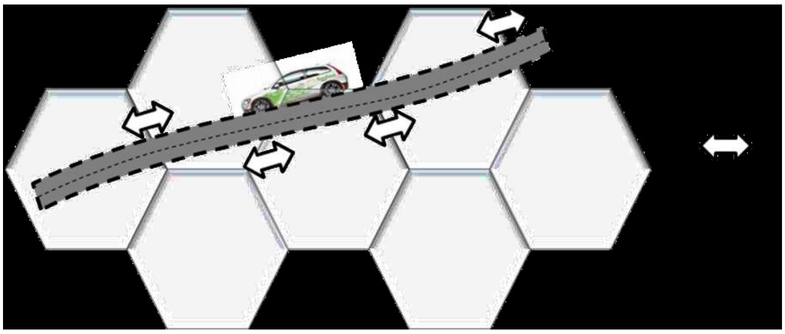

Fig. 1. How a Vehicle Moves from One Cell to Another
(Gundlegård & Karlsson, 2009, p. 5)

The cellular infrastructure is composed of a Core Network (CN) and a Radio Access Network (RAN). The CN is divided in two distinct domains: Circuit-Switched (CS) and Packet-Switched (PS). Mobile terminals can "attach" to the CS for voice call services, to the PS for packet data transfer, or to both. Radio communication occurs between a mobile terminal and a fixed base station serving one or more radio cells (sometimes also called "sectors"). Cells are the smallest spatial entities in the cellular network. The geographic area covered by each radio cell, which often includes or is close to the corresponding base station location, has a radius that varies from a few dozen meters (microcells) up to several kilometers (macro cells) depending whether the area is an urban, suburban, or a country-side area. Depending on the radio bearer, radio cells can be classified as 2G (GSM/EDGE), 3G (UMTS/HSPA) or 4G (Janecek et al., 2012, p. 362).

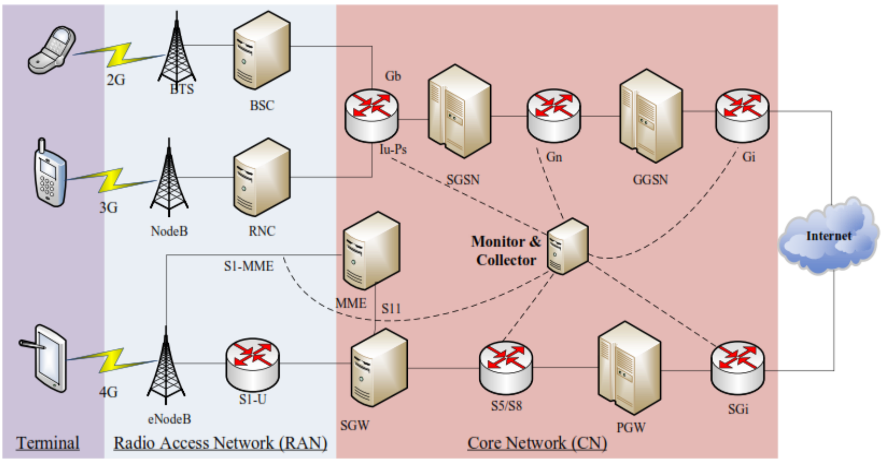

Fig. 2. How CN & RAN Work
(Zhang et al., 2016, p. 2)

The instantaneous position of each terminal is known by the cellular network – and can be inferred from the network signaling data – with different spatial granularity. Cells are grouped into larger logical entities, i.e., Routing Areas (RAs) and Location Areas (LAs) for the PS and CS domains respectively (note that one LA can contain one or more RAs and any RA is always contained within one LA). In order to remain reachable, active and idle terminals always inform the CN whenever they change LA and/or RA. Terminals in active state reveal also cell changes within the RA/LA to the network. In other words, the position of active users is known by the network at the cell level, while the position of idle users is known only at RA/LA level (Janecek et al., 2012, p. 363).

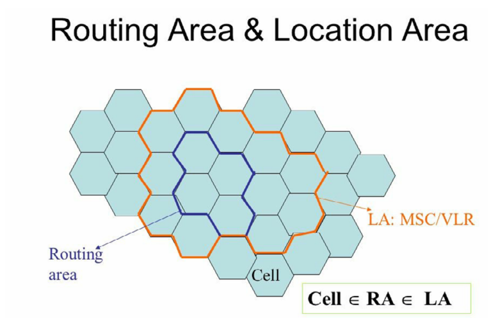

Fig. 3. Relation between Location Area (LA), Routing Area (RA) and Cell
(Permata, 2014, p. 23)

Mobility Management in Cellular Mobile Phone Networks

• **Location Update**

The location update procedure allows a mobile device to inform the cellular network whenever it moves from one location area to the next. Mobiles are responsible for detecting location area codes. When a mobile finds that the location area code is different from its last update, it performs another update by sending to the network, a location update request, together with its previous location, and its Temporary Mobile Subscriber Identity (Ding, 2010).

Various location update strategies are available in mobile computing that is divided into generic location update and periodic location update. In this paper the generic location update is used because it only records events when the terminal (mobile set) moves around.

- Generic Location Update

Whenever a mobile phone moves from one location area to the next while not on a call, a random location update is required. This is also required of a stationary mobile that reselects coverage from a cell in a different location area, because of signal fade. Thus, a subscriber has reliable access to the network and may be reached with a call, while enjoying the freedom of mobility within the whole coverage area (Ding, 2010).

- Periodic Location Update

Each mobile is required to regularly report its location at a set time interval using a periodic location update (Ding, 2010). If after the updating time period, the mobile station has not registered, it is then deregistered.

• **Handover**

Handover is the procedure that transfers an ongoing call from one cell to another as the user's moves through the coverage area of cellular system. The purpose of the handover procedure is to preserve ongoing calls when the mobile station is moving from one cell to another (Khan, 2010, p. 14).

• **Erlang**

The cellular network planning demands a great deal of information, such as market demographics, area to be served or traffic offered. This traffic varies on base station, month, day and even instant of the day. Traditionally, the unit of measurement of the telephone traffic is Erlang. The telephone traffic indicates the amount of voice/data traffic transmitted over a communication channel. This traffic is measured in terms of usage time and depends on the number of communications and their duration. Some examples of traffic data, based on mobile network are as follows:

• **O–D matrix**
• **Volume (traffic flow)**
• **Speed**
• **Travel time**
• **Traffic density** (Caceres et al., 2008, p. 181).

Table 1. Examples of Traffic Data, Based on Mobile Cellular Network

Traffic Data	Source Event	Advantages	Disadvantages
O–D Matrix	Handover or Location Update	Sample Size, Data Collected Directly from the Traffic Stream, Not Time Consuming	Large Areas (Not Detected Intra-area Trips)
Traffic Flow	Handover or Location Update	Work Over a Wide Coverage (Not Limited to Main Roads)	Sample Size, Errors in Densely Populated Areas (Pedestrians)
Speed, Travel Time	Handover	Acceptable Location Accuracy, Report Space-mean Speed, Work Over a Wide Coverage	Multiple Roadway Links within a Cell (Multiple Routes)
Traffic Congestion	Call Volume	Quick Detection	Poor Location Accuracy
Traffic Density	Traffic Intensity Per Cell (Erlang)	Work Over a Wide Coverage (Not Limited to Main Roads)	Sample Size, Errors in Densely Populated Areas (Motionless Users, Pedestrians, ...)

(Caceres et al., 2008, p. 191)

TEHRAN METROPOLITAN AREA CASE STUDY

In this paper, taking advantage of mobile cellular network is proposed to work as a ubiquitous mobility sensor for temporal human mobility analysis and in understanding citizens' behavior patterns in a city. It is shown that many attributes of human mobility can be inferred from anonymized mobile data that are collected from cellular mobile network.

• The Collected Data

As stated before, there are two major collecting methods for mobility information from cellular networks, one of them that only works when the user starts or stops a voice call, data connection or SMS/MMS is called CDR[1] and the other which can observe RA/LA changes of all users, including idle ones is called passive monitoring which is used in this study. The advantage of passive monitoring with location update (LU) method is that it can include larger amount of data and be analyzed in a larger area than the active method but when it comes to a greater scale and smaller area analysis, active method should be used too besides idle method, because it is based on cells which are smaller areas in relation to RA/LA. So, in this way, by monitoring the links between the CN and the RAN, a planner can observe RA/LA changes of all users, including idle ones, resulting in an enormous gain in coverage and estimation accuracy. Our dataset is based on this approach and contains data from CS and PS users from both 2G and 3G cells. Also, in LU method there are two ways to record events, periodic and generic. In this study generic method is used because it only updates the geographical positioning of mobile cells once their location area is changing. In fact, in this way we can have information about the persons that are actually moving and analyze and understand human mobility patterns.

In this study anonymized traces of signaling traffic from the operational 2G/3G network of a major Iranian mobile operator in Tehran Metropolitan area were collected. The data for this study was collected on Monday, 19th December 2016 for the 24 hours of the day. The traces consist of following fields:

- Cell information including the LAC (Location Area Code) and Cell ID, the coordinates, type, beam width, and direction of the base station antenna.
- Information about the event such as the type of event and the input source, i.e., the capturing interface where the signaling message was observed.

• Mapping and Data Mining

There are generally 85 location areas in Tehran metropolitan area and this data was attributed in the Geographic Information System (GIS). Then the map was created by the density of hourly location updates in each LA. The map shows hourly mobility in Tehran and we can see how it is distributed throughout the day and locate regional activity hotspots of Tehran. Fig. 4 shows the overall LAC of Tehran in a normal day. The warmer colored areas, the red ones, have had more mobility and therefore more activity, but the colder colored areas, the blue ones, have had less mobility throughout the day with less activity. The purple colored areas in this map are the congestion charged zones which have been set by the municipality of Tehran to prevent traffic congestion from the Central Business District (CBD) and the great bazaar which is located in region 12. In the smaller congestion charge zone no vehicle is allowed to enter this area from 6:30 in the morning until 5 in the evening unless they pay the price and in the larger congestion zone vehicles are only allowed to enter the area every other day based on their plate number (if it starts with an odd or even number) unless of course the driver pays the charges. But the red area in this map indicates that the CBD is the most visited location in Tehran throughout a normal work day. Unfortunately, in this study different modes of transportation haven't been distinguished, so we can't conclude that setting the traffic rules has been successful or a failure and the CBD is only determined as a certain hotspot in the city. Of course, by using new data mining tools and methods even the transportation modes can be recognized to conclude a more certain idea about the traffic behavior in this part of the city.

Another thing that this map shows is that we can generally divide the city into two parts: The central zone (districts 10, 11, 12, 6 & 7) which have more traffic and mobility and the border zone (districts 1, 4, 8, 13, 15, 22, 21 …) which have much less traffic and mobility. So, the central zone should be a commercial and job oriented area, attracting people which is the destination of trips and the border zone a residential one which is the origin of most intra urban trips.

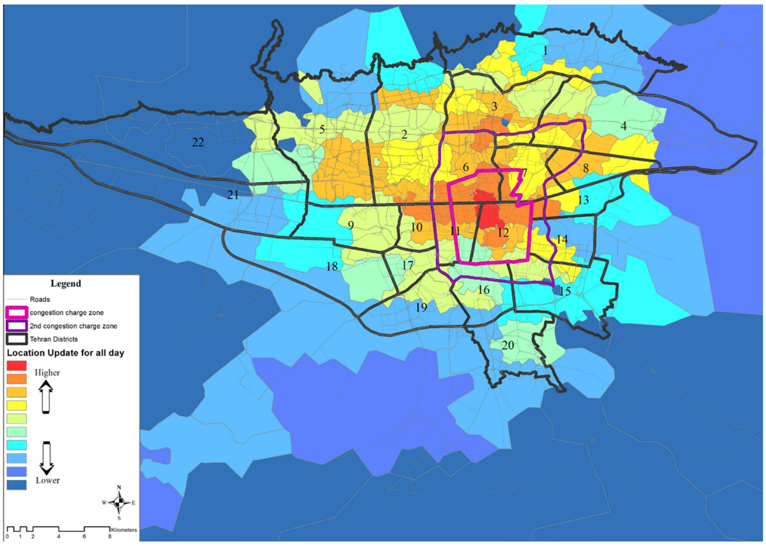

Fig. 4. Tehran Daily Zonal (LAC) Human Mobility Pattern Calculated from Cellular Phone Data with Location Update Method

RESULTS

In this article, human mobility pattern analysis is categorized into three different subjects: Transportation, social issues and city functions. So, the collected data will be analyzed in these subjects respectively.

Transportation

As mentioned before, there are many factors in urban transportation that can be analyzed by cellular data. In this article the general flow of traffic and movement patterns throughout a day in different hours is shown.

Fig. 5 illustrates the traffic flow changes in a day in Tehran Metropolitan area. As it is shown in this chart, 6 PM is the peak hour in this area and traffic congestion is at its highest rate but 5 AM in the morning has the least traffic throughout a normal day.

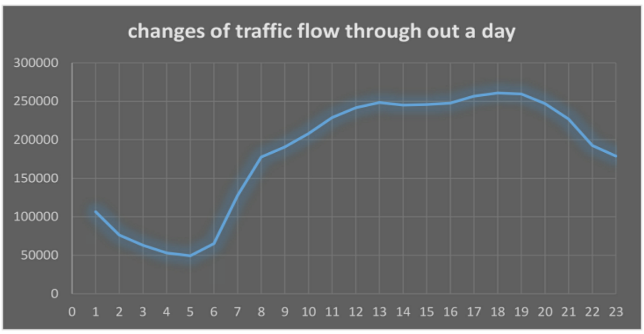

Fig. 5. Changes of Traffic Flow throughout a Day in Tehran Showing Peak Hours

Three-time periods were chosen in this article to analyze due to their certain traffic behavior changes and their importance in urban transportation planning.

Fig. 6 shows the volume of movement in 07:00 to 08:00 AM in the morning which is the time that most of workers go to their workplace or students go to their schools. It is clear that the border zone and the highways near the entrances of Tehran in west, east and south specially the west where Hemmat and Tehran-Karaj highways are located have the most mobility at that time of the day, but the central zone have the least mobility. This behavior illustrates that Karaj city in the west of Tehran with over 1.7 million inhabitants has the most effects to the morning traffic of Tehran than the adjacent cities to the east (Damavand, Pardis…) and the south (Rei, Islamshahr, Varamin). The cities near Tehran are all mostly residential cities and their inhabitants visit Tehran to work on a daily basis and then go back to their own cities in the evening, which is why the population of Tehran is around 8 million at night but over 12 million in the day.

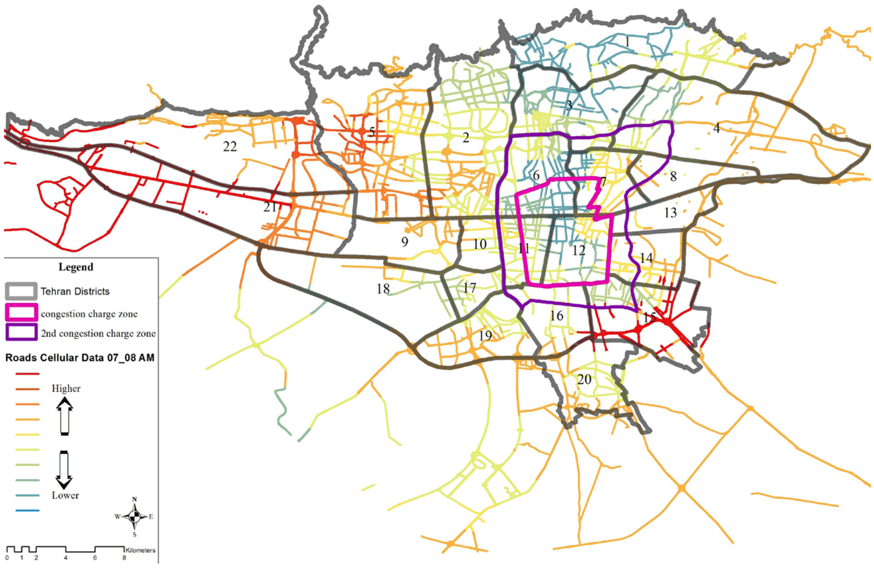

Fig. 6. Tehran Location Updates from 7 to 8 AM Showing Traffic Patterns

Fig. 7 shows the volume of movement in 13:00 to 14:00 PM in the afternoon. At this hour, nearly all of the city has great movement and mobility and the city is very much vibrant, but the central zone has more traffic.

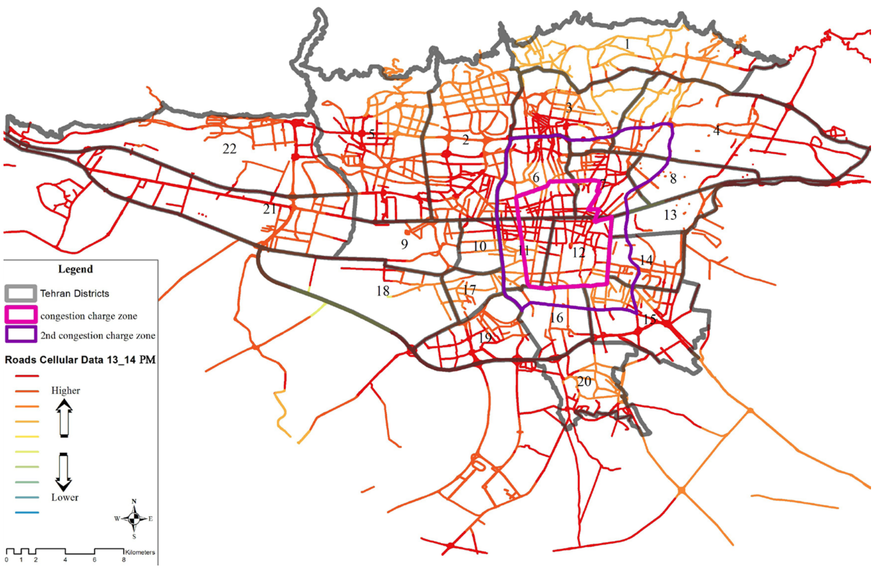

Fig. 7. Tehran Location Updates from 13 to 14 PM Showing Traffic Patterns

Around 5 PM to 8 PM workers start going back to their homes and then there is another peak hour in the city. Fig. 8 shows the volume of movement in 20:00 to 21:00 PM. It is clear that in 8 PM many of the workers have parted their jobs at the central zone for their homes at border zone or adjacent settlements because again the mobility in the periphery is getting higher and lower in the center.

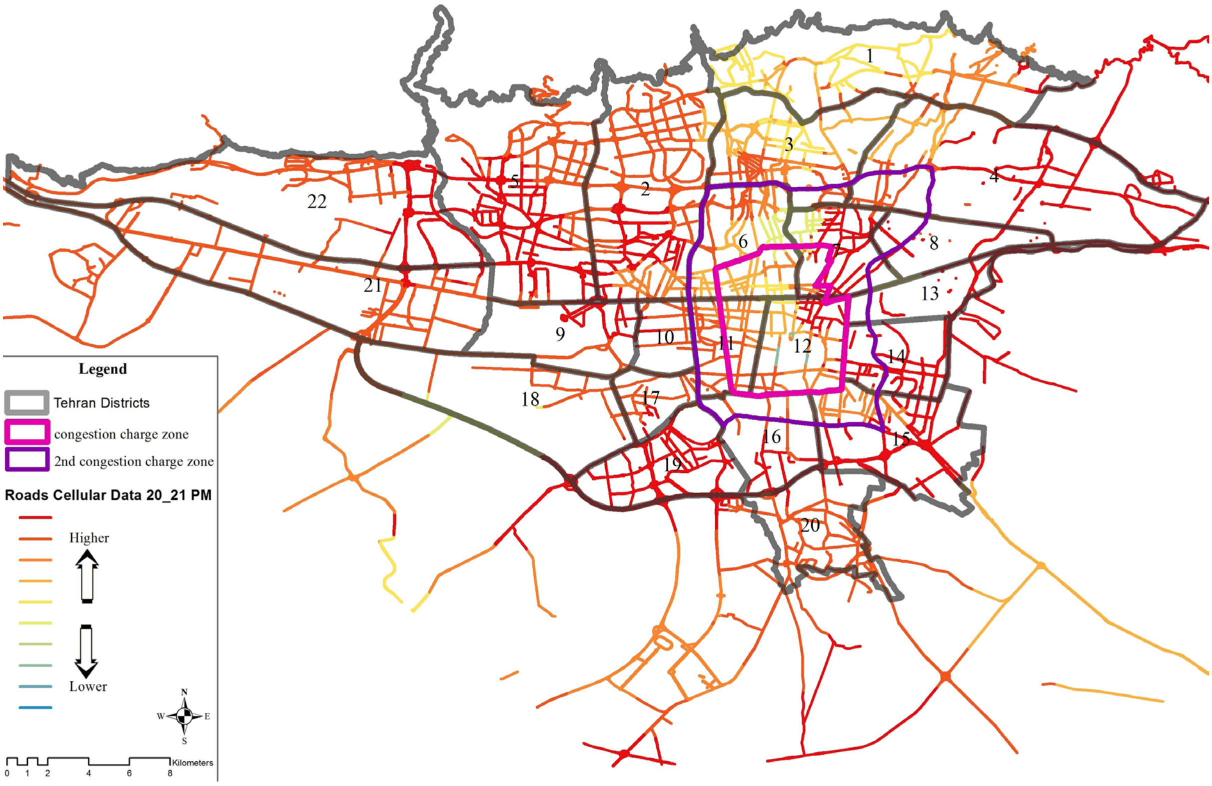

Fig. 8. Tehran Location Updates from 20 to 21 PM Showing Traffic Patterns

By these collected data many transportation issues can be studied, for example the most congested areas and roads in different times of the day can be identified and as a result certain traffic rules can be set to reduce traffic congestion.

Social Issues

There are many social characteristics that are related to human mobility, one of them is social security in a city, which will be analyzed by mobile cellular data in this article. As many theorists like Oscar Newman in "defensible space" theory and Jane Jacobs in the book "life and death of American cities" have noted, crime has a very tight relation with human mobility and social interactions or as they call it "natural control" or "eyes on the streets". These theorists believe that if an urban space is in control of the community and there are people watching these spaces then the criminal wouldn't dare to act criminally. However, researches have shown most criminal acts, like pick-pocketing or rubbery are taken place in crowded places. In this article a human mobility map is combined with the population density map of Tehran city to analyze these theories further.

Fig. 9: Tehran location updates in a day in comparison with density, is a map that shows where the population is living and where the crowded neighborhoods are (the black dots), also the most visited parts of the city by people is shown in colored zones. Although the activity hob of the city is the heart of the city where region 12 is located, this part has a very low population.

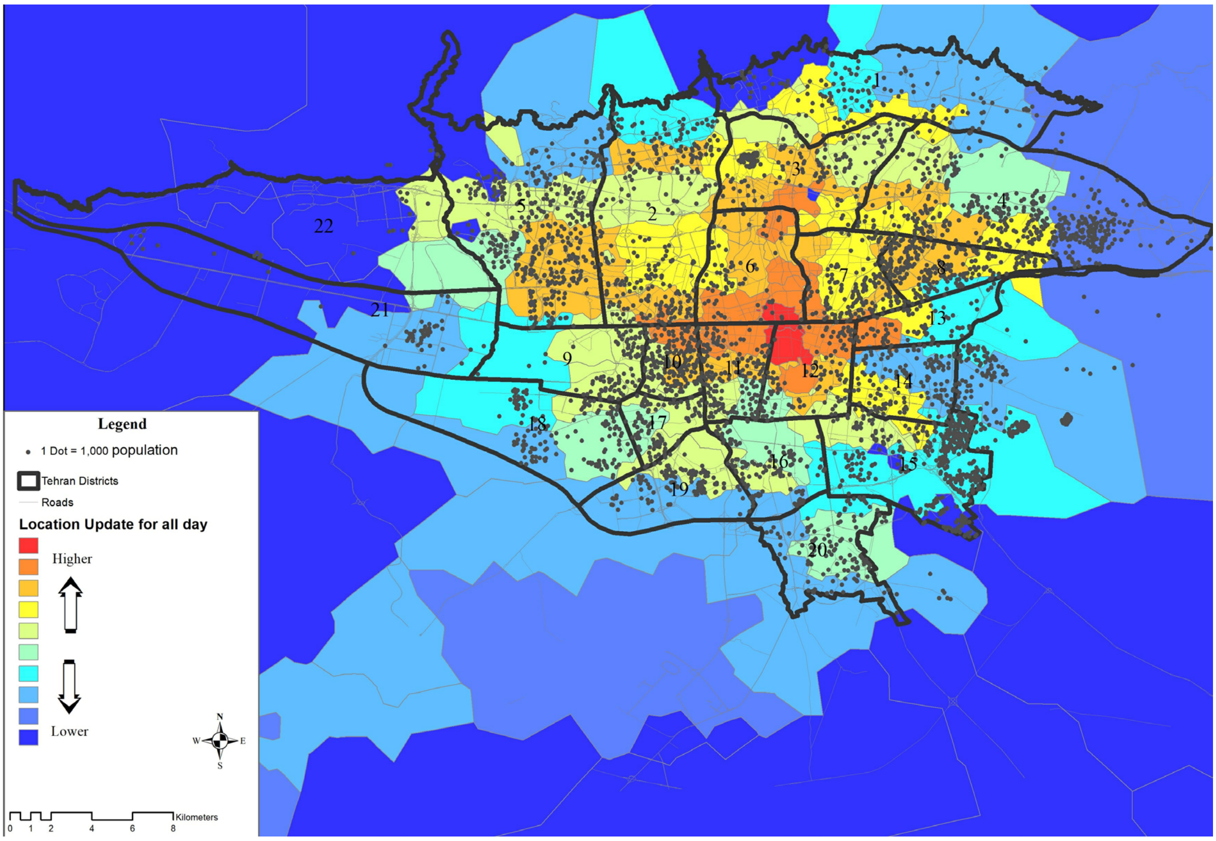

Fig. 9. Tehran Location Updates in a Day in Comparison with Population Density (Extracted from 2006 Census Data)

So, considering the theories mentioned above, this part would probably have a higher rate of crime, because there is no locality, people do not know each other and also there are many people from all over the city that could visit this area, as a result natural surveillance is low and more criminal acts would happen in such a place. Moreover, when this map is analyzed in different hours of the day, for example early morning or late at night (Fig. 10) this assumption gets more real because this region is isolated and haunted in these hours and as a result crime rate would be higher.

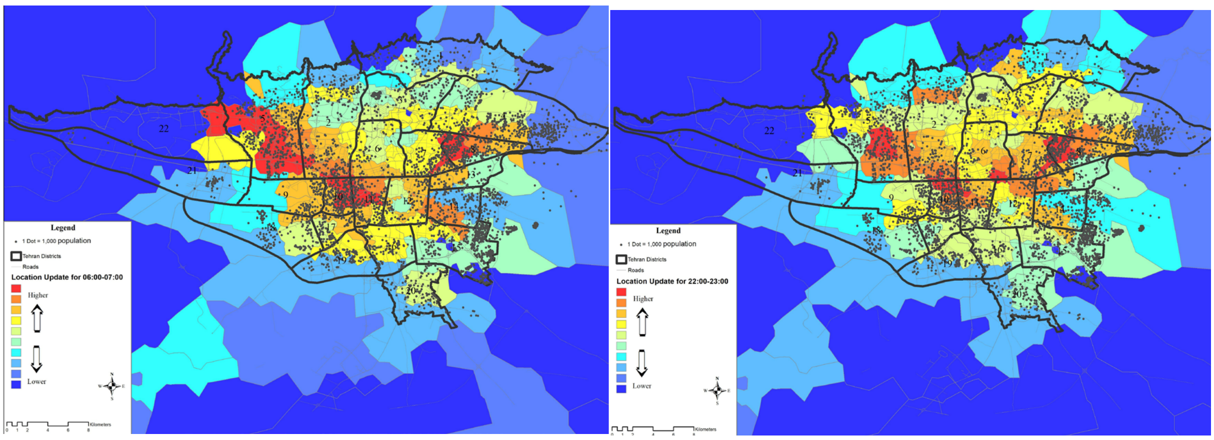

Fig. 10. Tehran Location Updates 6-7 AM and 22-23 PM in Comparison with Population Density
(Extracted from 2006 Census Data)

In support of this assumption, there is an article written by Pour Ahmad et al. (2003) with the title of "Examining the geo-position of crimes in Tehran city" showing that crime rate is very high in region 12 in comparison with other regions.

Of course, there are more social criteria like co-presence or social interactions that can be analyzed with this huge amount of real-time data.

Urban Functions

Another important attribute of a good city is having proper urban functions where land-uses and urban amenities work properly at the service of the citizens. By analyzing cellular mobile data urban planners and managers can understand how these urban functions work and how they could be managed or changed to work optimally. For example, by these data the best place to construct a new land-use can be found or the dispersion of land-uses can be analyzed.

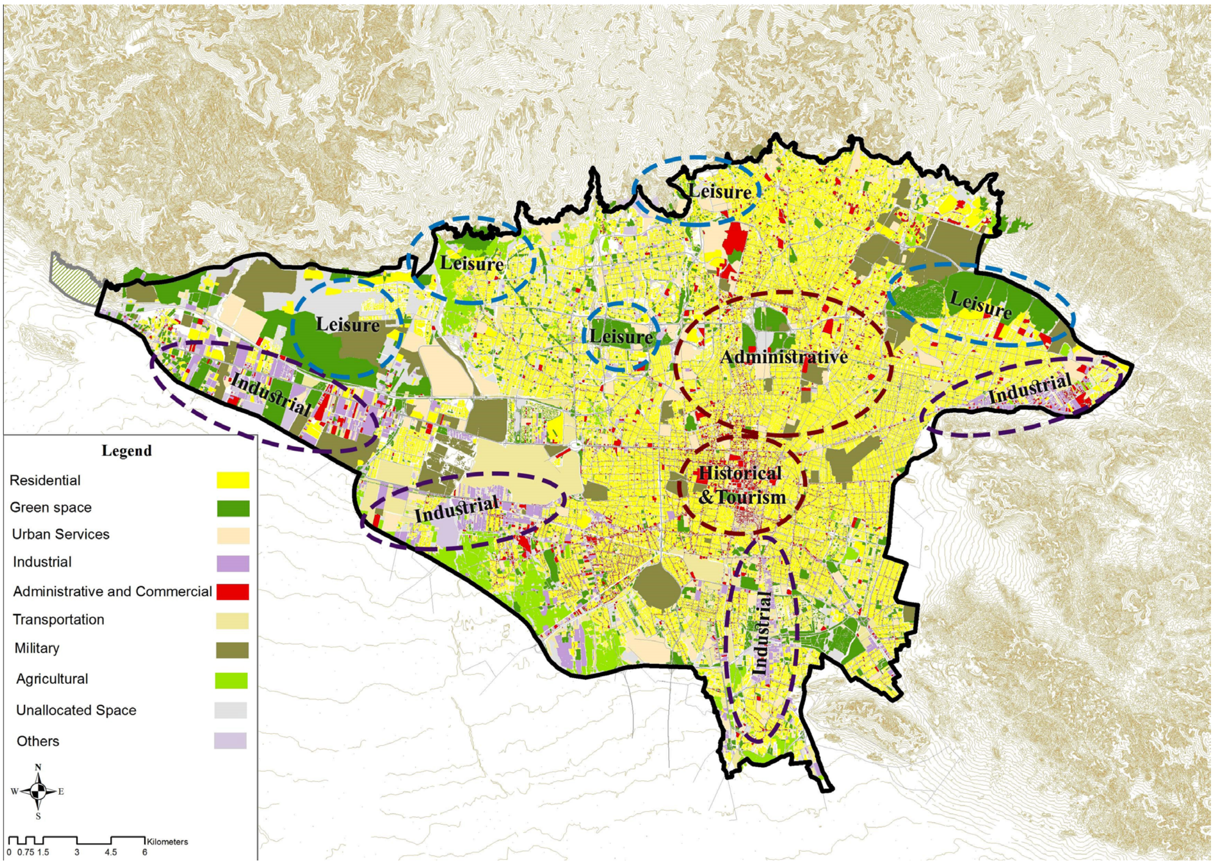

Fig. 11. Tehran Existing Condition Land-use Map in 2006 (Tehran Comprehensive Plan)

Many studies have worked the relation between urban functions and urban mobility or better to say the impacts of land-uses on creating urban trips as destinations of each movement in a city. To sum up, there are three main activities that make every individual travel and move around a city; they are, work, leisure and fulfilling daily/weekly necessities. The land-use map of Tehran is used to understand how these destinations determine where human mobility takes place. As it is shown in Fig. 11, four major zones can be distinguished in Tehran, industrial, leisure, administrative and historical & tourism.

Comparing the land-use map of Tehran (Fig. 11) with location update map (Fig. 4) it shows that in a big city like Tehran and on a normal working day, administrative and commercial land-uses have the most impact on attracting human mobility (of-course this fact could change on weekends in favor of leisure activities), so urban planners and managers should focus on finding the best land-use patterns to facilitate human mobility. An important issue which seems to not have been resolved in Tehran and as a result nowadays this city is having many problems. Unfortunately, centralization of destination land-uses like commercial or administrative buildings in some part of city specially in the center and not dispersing the land-uses properly throughout the city has led to a lack of functionality in urban zoning which would decrease the quality of citizens' life. As a result, many citizens have to take long trips to visit commercial buildings to fulfill their needs or to go to their work-place; this becomes completely clear by observing the flow of human mobility in Tehran throughout a day.

Although these are just illustrations and sketchy examples of how mobile cellular data can help urban experts understand complexity of cities deeper and therefore, plan and design places with higher quality of life. We hope that in the near future, having more accurate and detailed data, helps us improve urban modelling, understand complex human trends and predict human behavior.

CONCLUSION

In this article, first, human mobility was introduced as complex urban phenomenon that needs assessments and analyzes in shorter periods of time to be fully understood, then big data and data mining was examined as a way to study cities more accurately and their applications in understanding temporal human mobility patterns were stated. In the end mobile cellular data analysis was chosen as the best method to understand human behaviors and was used in analyzing Tehran metropolitan area population mobility throughout a day. The concluding remarks of this paper are summarized as follows:

• Human mobility is a dynamic complex phenomenon that in order to be perceived, newer methods should be used.

• Urban human mobility and human behavioral patterns in a city should be analyzed in short term and temporal dimension, to be fully conceived.

• The collected data that is used nowadays to study urban issues by new technologies is often big and hard to cluster and make use of. Data mining science has shed new light on taking advantage of big data in urban issues.

• The use of cellular phone data is a proper way to understand human behavioral patterns and then resolve the problems associated with them.

• Using generic location update method in cellular phone network is suitable to estimate movement and mobility in cities.

• Cellular data can help urban planners and managers resolve many urban issues like transportation, social problems or urban functions.

• Results show that the peak hour of a day in Tehran is from 6 to 7 PM and the least crowded time of a day is from 5 to 6 AM in the morning when nearly all the city is sleep. along with detailed study, the peak hour of even different regions of a city could be determined from the available data.

• Tehran is divided into two major parts, the border zone and the central zone which are the origin and destination of most intra urban trips in the city.

• The CBD and historical context of Tehran where the great bazaar is located is actually the hotspot and the most visited part of the city throughout a day.

• Based on security and crime theories, region 12 in Tehran is a possible place for high crime rate because of being simultaneously isolated from population and the activity hob of the city.

• Centralization of land-uses in Tehran zoning has resulted in a decreased quality of life for citizens who have to travel a long way each day to go to work or do their other daily activities according to Tehran location update map and land-use map.

The analyses presented in this article are only based on mobile data zones (LAC) and showing regional hotspots and general estimation about human mobility in Tehran which are only included in this article to show a small part of what could be done with the use of big data and especially mobile cellular data in understanding and resolving urban complex issues. In future studies with more accurate data and taking advantage of both active and idle networks many more analysis could be done in small scales and smaller areas.

ENDNOTE

1. Call Data Record

REFERENCES

Albeverio, S., Andrey, D., Giordano, P., & Vancheri, A. (2008). *The Dynamics of Complex Urban Systems*, Springer.

Barabasi, A.L., Song, C., Qu, Z., & Blumm, N. (2010). Limits of Predictability in Human Mobility. *Science*, 327(5968), 1018-21.

Batty, M. (2011). Building a Science of Cities. *Cities*, 10(1016), 1-8

Batty, M. (2013). *The New Science of Cities.*, The MIT Press, Massachusetts Institute of Technology.

Batty, M., Axhausen, K.W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Ouzounis, G., & Portugali, Y. (2012). Smart Cities of the future. *The European Physical Journal*, 214, 481-518.

Bhaduri, B., Bright, E., Coleman, P., & Urban, M.L. (2007). LandScan USA: A High-resolution Geospatial and Temporal Modeling Approach for Population Distribution and Dynamics. *Geo Journal*, 69,103–117.

Caceres, N., Benitez, F.G., & Wideberg, J. (2008). Review of Traffic Data Estimations Extracted from Cellular Networks. *Intelligent Transport Systems journal*, 2(3), 179-192.

Ding, J. (2010), *Advances in Network Management*. Taylor and Francis Group.

Dobson, J.E., Bright, E.A., Coleman, P.R., Durfee, R.C., & Worley, B.A. (2000). LandScan: A Global Population Database for Estimating Populations at Risk. *Photogrammetric Engineering and Remote Sensing*, 66(7), 849-857.

Elias, D., Nadler, F., Stehno, J., Krösche, J., & Lindorfer, M. (2016). *Improving Public Transport Planning through Mobile Phone Data Analysis*. Transportation Research Procedia, 4478 – 4485.

Gonzalez, M.C., Hidalgo, C.A., & Barabasi, A.L. (2008). Understanding Individual Human Mobility Patterns. *Nature*, 453, 779–782.

Gundlegård, D.. & Karlsson, M.J. (2009). *Road Traffic Estimation using Cellular Network Signaling in Intelligent Transportation Systems*. Wireless technologies in Intelligent Transportation Systems.

IEA. (2012). *Energy Technology Perspectives*. International Energy Agency.

Janecek, A., Valerio, D., Hummel, K.A., Ricciato, F., & Hlavacs, H. (2012). Cellular Data Meet Vehicular Traffic Theory: Location Area Updates and Cell Transitions for Travel Time Estimation, *Conference on Ubiquitous Computing*, U.S.A, 361-370.

Khan, J. (2003). Handover Management in GSM Cellular System. *International Journal of Computer Applications*, 8(12), 14-24.

Kumar, S., & Prakash, A. (2016). Role of Big Data and Analytics in Smart Cities. *International Journal of Science and Research*, 5(2), 12-23.

Lee, K., Hong, S., Joon Kim, S., Rhee, I., & Chong, S. (2009). *SLAW: A Mobility Model for Human Walks*. In Proceeding, INFOCOM, Rio de Janeiro, Brazil.

Mitchell, W.J., & Casalegno, F. (2008). *Connected Sustainable Cities*. MIT Mobile Experience Lab Publishing.

Permata, Z. (2014). *Generic Packet Radio Service ppt*. slideplayer.info, Accessed on 30nd December 2016.

Portugali, J. (2000). *Self-organization and the City*, Springer.

Portugali, J. (2009). *Complexity Theories of Cities: Achievements, Criticism and Potentials*. An International Conference on Complexity Theories of Cities.

Portugali, J. (2011). *Complexity, Cognition and the City*, Springer.

Pour Ahmad, A., Rahnamaee, M., & Kalantary, M. (2003). Examining the Geo-position of Crimes in Tehran City, *Pajuheshhaie Goghrafiaee*, 44, 81-98.

Ratti, C., Pulselli, R.M., Williams, S., & Frenchman, D. (2007). Mobile Landscapes: Using Location Data from Cell-phones for Urban Analysis. *Environment & Planning*, 33(5), 727–748.

Semanjski, I. (2016). *Transport and Traffic Analytics in Smart Cities*, Conference on Data Analytics, Vencie.

Statistical Center of Iran (2006). Population and Housing Census.

Tehran Municipality (2006). Tehran Master Plan.

United Nations Global Pulse (2013). Mobile Phone Network Data for Development, 1-12.

United Nations (2014). World Urbanization Prospects: 2014 Revision.

Zhang, X., Yi, Z., Yan, Z., Min, G., Wang, W., Maharjan, S., & Zhang, Y. (2016). Social Computing for Mobile Big Data in Wireless Networks, *Computer*, 49(9), 86-90.

www.theconnectivist.com, Accessed on 12/24/2017.