



## MODELING CITY PULSATION VIA MOBILE DATA

Suhad Faisal Behadili <sup>\*1</sup>, Cyrille Bertelle <sup>2</sup>, Loay E. George <sup>1</sup>

<sup>\*1</sup> Computer Science Department, University of Baghdad, Baghdad, Iraq

<sup>2</sup> LITIS, Normandy University, FR CNRS 3638, ISCN, ULH, Le Havre, France



### Abstract:

*In this study, the mobile phone traces concern an ephemeral event which represents important densities of people. This research aims to study city pulse and human mobility evolution that would be arise during specific event (Armada festival), by modelling and simulating human mobility of the observed region, depending on CDRs (Call Detail Records) data. The most pivot questions of this research are: Why human mobility studied? What are the human life patterns in the observed region inside Rouen city during Armada festival? How life patterns and individuals' mobility could be extracted for this region from mobile DB (CDRs)? The radius of gyration parameter has been applied to elaborate human life patterns with regards to (work, off) days for the observed data.*

**Keywords:** *City Pulse; Modelling; Simulation; Radius of Gyration; CDR; Mobile Data.*

**Cite This Article:** Suhad Faisal Behadili, Cyrille Bertelle, and Loay E. George. (2018). "MODELING CITY PULSATION VIA MOBILE DATA." *International Journal of Engineering Technologies and Management Research*, 5(4), 115-122.

## 1. Introduction

The scientists connect between people and their mobiles physically and conceptually, since mobile existence mean physical person existence and mobile spatial behaviour refers to the individual behaviour at that spatial domain, so the type of these data has several aspects with a lot of indications that are reveal human mobility patterns, these patterns are characterized by different time scales and distances, which reflecting the daily circadian rhythm, either long or short distance/time scales to understand the global or local spreading of individuals, epidemics ...etc. [14, 20, 4]. These types of studies are done via analysing the communications data, which are obtained from mobile communication networks, therefore it will be easy for citizens, local administrations and services appliances to be ready for any sudden events with high fluently responding. This would be happened, when well adaptation available to newly and suddenly changes, which could be emerged from city dynamics (people mobility) during grand public events. There are many researches are performed according to these concepts and perceptions, to analyse urban systems of cities, the material mainly were cellular mobile data, SMS, and new mobile applications. The analysis and tracing mobile phone calls makes the geographical representation of people mobility inside cities possible. Therefore, there would be a suitable depicting on future city characteristics and features,

All this could be depicted in real time. As well as, real time processing and analysing for these data could be done, since it is real time data, and came from real infrastructure of the city. Therefore, this will assist greatly in decision making according to people behaviour inside the city, especially when analyse these data in aggregated or disaggregated (individually/collectively), with different kinds analysis patterns [11, 9]. However, ABM is developed to be Multi-Agent System MAS, which adapts more complicated agents' interactions, to be simulated effectively as could as possible [15, 16, 2]. MAS had been proposed to micro-simulate land-use changes, contributing (environment, transport, economic models) to obtain complex urban systems, for instance Ramblas (Veldhuisen et al. 2000), UrbanSim (Waddell 2000, 2002, Waddell et al. 2003), ILUTE (Miller et al. 2004), PUMA (Ettema et al. 2005), ILUMASS (Moeckel et al. 2003, Wagner and Wegener 2007). Also, models of road network uses like ABLOoM (Otter et al. 2001), CityDev (Semboloni 2005), and Kou's artificial urban (Kou et al. 2008) [21].

There are many types of researches that manipulate human mobility from different views, as in research [1], where they proposing novel technique, called Adaptive DAD, to detect dense areas that define the concept of density using the infrastructure provided by cell phone network. They evaluate and validate this approach with real dataset containing CDRs of fifteen million individuals. The adaptive and non-parametric method for identification of dense areas is based on using the ubiquitous infrastructure provided by cell phone network [1]. Data mining methodologies are contributed in representing human mobility, the scientists in research [14] attempt to represent human mobility patterns depending on mobile spatial indication by building human mobility models, and then using data mining methodologies for acquiring the desired knowledge, which indeed are important to explore the whole image of human mobility inside his environment, and the impacts of this relation on urban environment development and the preparedness to facing any unusual events according to predetermined understanding and prepared planning, this research emphasized on important nature of the mobile data, which is the uncertainty and inaccuracy, due to the nature of human life patterns [11], therefore it will be reflected on mobility models and extracted traces for human mobility. Gonzalez et al. explore the modelling of human mobility, and they conclude from their studies, the following [23, 19, 13]:

- 1) The human trajectories clarify high degree of spatio-temporal regularity.
- 2) The individuals are characterized by, non time-dependent travel distances.
- 3) The individuals characterized by high probability of returning to few previously passed locations.
- 4) The individuals' mobility patterns have almost the same patterns in predictability average ratio of 93%, although of the varying in dynamic histories and social features, because of human behaviour regularity.
- 5) The visited locations are connected by mobility trips, constructing movement oriented networks, these networks named mobility motifs, which give view of the daily dynamic activities. These motifs could be assumed as general human mobility characteristics, then to be used in the travel time analysis for urban activity modelling and simulation. The cellular network data have main benefits in revealing human mobility, because cellular device is ubiquitous now, so their data (active/passive) would cover wide range areas, and picked up huge population number, which makes it the most suitable media for human mobility studies [23, 18].

## **2. Materials and Methods**

The emerging visions of knowledge enhance the perception of daily urban pulse developments in the city, which illuminates additional data for a few spaces for example event administration, urban planning, public transportation, and so on. There is clear correlation city functional configuration and human mobility patterns. Despite all highlights of the investigation approaches, there is a need in semantic understandings, which is the appropriate response of why behavior /pattern configure this?

The telecommunication operator gives the gigantic volume of information as CDRs, which contain extremely profitable spatio-temporal data at various levels of granularity (e.g. citywide, statewide, or across the country). This data is pertinent for the telecommunication operator, as well as be the base for a more extensive arrangement of uses with social connotations like people density, commuting patterns, transportation routes, ...etc. and so on. The intense way to the smart cities improvement are the efficient pursuit of spatio-temporal patterns to investigate the CDRs databases and statistical techniques utilization. However, nowadays the obtainable commercial systems of telecommunication operators cannot deal with this form of spatio-temporal processing. One conceivable approach to analyze such patterns is to implement sequential scanning of the all entire database or call records and check them by employing subsequence matching like algorithm against the query pattern [23, 24].

Although, it is computationally costly because of the enormous information to be prepared, and there is no data about the temporal dimension (e.g. Between two given hours or between two given days) or spatial properties (e.g. in a given neighborhood, intersecting given area, or near given spot), which are certified to process the database (e.g. in a given neighborhood, near given spot, or converging given zone), which are considered to process the database [23, 24].

In spite of the fact that, the CDRs have popularity in examination, however it has two constraints, where there are no information when there is no communication activities (inactive mobile phone), and additionally the aggregated spatial information of each tower region (coarse spatial granularity), since the spatial information are not the exact position of active mobile, but rather they define the approximate one within tower coverage area [24].

The Armada DB is the case study data composed of registered records CDRs from an Orange company for mobile communications and services, giving the mobile network activities, which includes 51,958,652 CDRs (entry records) for 615,712 subscribers. The assumption of each mobile represents a human, hence his/her activity, mobility, and then human life patterns, concerning the data of the observed region. The researchers make their efforts to analyze the enormous data, and represent the acquired knowledge on different styles like statistical sheets, statistical diagrams and the more powerful way is the Geographic representation using a GIS science, in order to give more realistic and comprehensive image to the analyzed data, which is in fact characterizing the human behavior patterns in different daily activities [14, 5, 13, 7]. The classical CDRs contain the mobile ID (alias), Cell Id, Cell Positions in coordinates (x, y), towers positions in coordinates (x, y), mobile activities type (Call in/out, SMS in/out, Mobile hang over, abnormal Call halt, normal Call end), date and time of Mobile activity which indeed represents a user activity. However, this data represents individuals occurrence in discrete (irrelevant) mode only, means that any mobile individuals' activity is recorded at (start/end) time, but there is a lack (lost information), which is

supposed to indicate the user's occurrence during inactive case (mobility without any mobile phone activity). There is no data meanwhile the mobile phone is idle, i.e. inactive or doesn't make any communication activities neither calls nor SMS activities. As well as, the available spatial data is only the towers (X, Y) coordinates, hence it would be considered to estimate individuals' transitions from position to another (from tower to tower), hence the positions would be determined approximately with regarding to the tower coverage (signal strength).

Also, the fundamental issue, which is utilized to process individual speed wouldn't be precise, but instead it is evaluated, since it utilizes the distance between every two progressive change positions, that are recorded on CDRs. At that point, the travelled distance would be divided by the time difference of each transition, to obtain the speed of individual's mobility.

The travelled distance is the displacement between two progressive positions in the elapsed time, which the time required to move from position to another. Due to the non-deterministic and discrete nature of these information, and every individual could be disappeared for a moment from the DB records, which makes individual following is unworthy, without significant indications on the individuals' mobility in the city. Along these lines the Therefore the collective behaviour would be the effective approach to analyse and mimic them. Armada DB is a huge one, so it is difficult to handle the aggregate investigated information, where every day has substantial records ranging between 2,715,077-5,661,428 records [25].

### 3. Results and Discussions

The mobile communication networks offer wireless communication services, therefore individuals can communicate in multi locations all over times, which means the acquired data of this ubiquitous sensors would provide dynamic perception on this service users, so individuals behaviour will be sensed "exploited" inside urban environment in high scale level, with regard to the wide urban area under mobile networks cover. The term "in situ" is the opposite of "remote", which is used for the tightly closed sensor or the direct contact sensor with the event/phenomena being sensed. Herein, the mobile phone considered as in situ sensor according to its closeness for both (mobile user and the sensed event), this combination gives user generated traffic inside mobile networks. However, georeferenced social media data could be considered as a proxy for the collective human behaviour, which considered social in situ sensor data [10, 23].

#### 3.1. Individual Trajectory Characteristics

The significant importance of revealing human trajectories enforces the tendency to build the statistical models. Human trajectories have random statistical patterns, hence tracing human daily activities is most challengeable issue, in addition to its importance which is mentioned earlier as urban planning, spread epidemics...etc. In spite of data sources variance (billing system, GSM, GPS), but the common characteristics are the aggregated jump size ( $\Delta r$ ), and waiting time ( $\Delta t$ ) distributions. Where, ( $\Delta r$ ) gives an indication on the covered distances by an individual in ( $\Delta t$ ) for each two consecutive activities, and the ( $\Delta t$ ) is the time spent by an individual between each two consecutive activities [4]. The individual trajectory/mobility considered as microscopic level of mobility abstraction, which is constituted of sequenced coordinates positions along time i.e. the agent displacement in spatio-temporal unit [6, 19, 13, 22].

Cellular data are very heterogeneous, since the users could be very active have so many calls/SMSs or inactive, hence have little usage of the network, so sampling the users would be depending on their activities number [12]. The individual's activities sparseness causes incomplete spatial information, therefore the mobile individual has some general physical characteristics that are useful to compensate the lack of data (when no activity recorded), that could be used to build mathematical model of human mobility patterns, in order to verify the behaviour and life patterns using the most common mobility characteristics, which are (Centre of mass  $c_m$ , Radius of gyration  $R_g$ , Most frequent positions, Principal axes  $\theta$  (moment of inertia), Rotation of user trajectory, Rescale user trajectory according to Standard Deviation  $\sigma_x$ ,  $\sigma_y$ , Compute the spatial density function for each user (agent) [17], and Sample users (agents) using density function and  $r_g$ ,  $\sigma_x = PDF(r_g)$  [12, 8, 9, 3, 19, 13].

### 3.1.1. Center of Mass

It's the most visited positions by individual  $c_m$ , as in the following equations (1, 2):

$$x_{cm} = \sum_{i=1}^n x_i/n \quad (1)$$

$$y_{cm} = \sum_{i=1}^n y_i/n \quad (2)$$

Where  $x_i$  and  $y_i$  are the coordinates of the spatial positions,  $n$  is the number of spatial positions, that are recorded in the CDRs.

### 3.1.2. Radius of Gyration

It is the average of all individual's positions, which is the indication to the area visited by the individual (the travelled distance during time period). In other words it is the individual's daily commute to/from home and work, as could be formulated in equation (3), the probability distribution of  $r_g$  uncovers the population heterogeneity, where individuals traveled in  $P(r_g)$  in (long/short) distances regularly within  $r_g(t)$ , the distribution  $P(r_g)$  produces power law investigated in the aggregated travelled distance distribution  $P(r_g)$ .

$$r_g^a = \sqrt{\frac{1}{n_c^a(t)} \sum_{i=1}^{n_c^a} (\vec{r}_i^a - \vec{r}_{cm}^a)^2} \quad (3)$$

Where  $\vec{r}_i^a$  refers to  $i=1 \dots n_c^a(t)$  Positions recorded for individual  $a$ , and  $\vec{r}_{cm}^a = \frac{1}{n_c^a(t)} \sum_{i=1}^{n_c^a} \vec{r}_i^a$ , which refers to center of mass of the individual's trajectory.

The performed algorithm to compute Radius of Gyration distribution has complexity of  $O(n^2+4n)$ . However, the resulted power law distribution of radius of gyration formulated in equation (4), hence figure (1) elaborates the  $r_g$  distribution for the whole observed period,

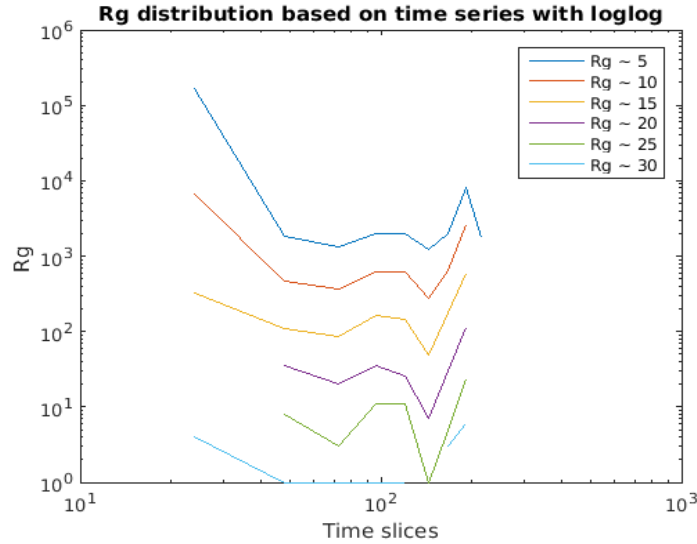


Figure 1: Rg Distribution based on time series with loglog during wholeobserved period

Whereas figure (2) elaborates the  $r_g$  distribution during work days period, and figure (3) elaborates  $r_g$  distribution during off days period.

$$P(r_g) = (r_g)^{\exp(-r_g)} \tag{4}$$

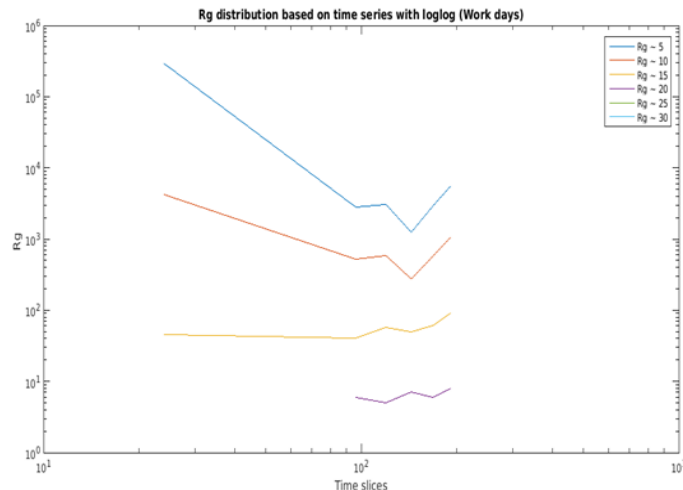


Figure 2: Rg Distribution based on time series with loglog during Work days period



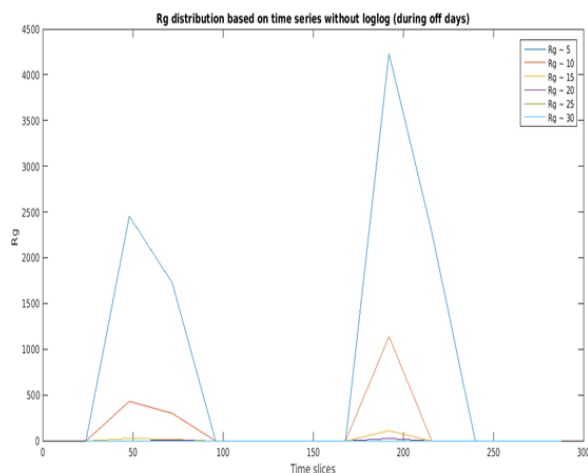


Figure 3: Rg Distribution based on time series with loglog during off days period

#### 4. Conclusions

The radius of gyration is the most common quantity, which is associated with human mobility trajectories, due to its capability in measuring the how far the mass from centre of mass. The concluded results which are investigated show that  $r_g$  gradually increases at the beginning, but it settles down versus time. It has key effect on the travel distance distributions. However the  $r_g$  is considered to be more dedicated feature that capable of characterizing the travel distances ( $\Delta r$ ) of individuals. The distributions showed that the individuals travel activities are almost identical, where the periodic trajectories are invariant. As well as the activities distributions are uncovered the regular patterns and behaviours similarities, during the time evolution of  $r_g$ . The experiments analyzed the relationships between the  $r_g$ ,  $\Delta r$  and the  $r_{cm}$ , the experiments showed that all individuals have similar  $r_g$ , but they are classified generally into two variant classes, which are: working days and off (holiday) days.

#### References

- [1] Alberto Rubioa, Angel Sanchezb, Enrique Frias-Martinez, "Adaptive Non-Parametric Identification of Dense Areas using Cell Phone Records for Urban Analysis," Telefonica Research. Engineering Applications of Artificial Intelligence Volume 26 Issue 1, Pages 551-563, January 2013.
- [2] Armando Bazzani, Massimiliano Capriotti, Bruno Giorgini, Giuseppina Melchiorre, Sandro Rambaldi, Graziano Servizi, and Giorgio Turchetti, "A Model for Asystematic Mobility in Urban Space," Laboratory of Fisica della Citt, University of Bologna, Italy, book The Dynamics of Complex Urban Systems An Interdisciplinary Approach. Physica-Verlag Heidelberg, 2007.
- [3] Balzs Cs. Csja, Arnaud Browetc, V.A. Traagc, Jean Charles Delvennec, Etienne Huensc, Paul Van Doorenc, Zbigniew Smoredae, and Vincent D. Blondelc, "Exploring the Mobility of Mobile Phone Users," 2013.
- [4] Chaoming Song, Tal Koren, Pu Wang, and Albert-Laszlo Barabasi, "Modeling the Scaling Properties of Human Mobility," 2010.
- [5] Christian M. Schneider, Christian Rudloff, Dietmar Bauer, and Marta C. Gonzalez, "Daily Travel Behavior: Lessons from a Week-Long Survey for The Extraction of Human Mobility Motifs

- Related Information,” UrbComp '13 Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing article No. 3, 2013.
- [6] Corina Iovan, Ana-Maria Olteanu-Raimond, Thomas Couronn, and Zbigniew Smoreda, “Moving and Calling: Mobile Phone Data Quality Measurements and Spatiotemporal Uncertainty in Human Mobility Studies,” Springer International Publishing Switzerland, 2013.
- [7] Dani Broitman, and Eric Koomen, “Residential density change: Densification and urban expansion,” Computers, Environment and Urban Systems, 2015.
- [8] Farnaz Fathali, and Hatairatch Charoenkulvanich,”Human-Mobility Modeling for Cellular Networks,” 2013.
- [9] Fosca Giannotti, Luca Pappalardo, Dino Pedreschi, and Dashun Wang, “A complexity Science Perspective on Human Mobility?,” 2012. <http://www.dashunwang.com/pdf/2012-obilitybook.pdf>.
- [10] Gunther Sagl, and Bernd Resch,”Mobile Phones as Ubiquitous Social and Environmental Geo-Sensors,” Chapter. Heidelberg University, Germany, 2014.
- [11] Lin Sun, “Enabling Pervasive Applications by Understanding Individual and Community Behaviors,” PhD thesis, Universite Pierre & Marie Curie -TELECOM SudParis, December 2012.
- [12] Marta C. Gonzalez, Cesar A. Hidalgo, and Albert-Laszlo Barabasi, “Understanding individual Human Mobility Patterns Supplementary Material,” nature, 2008.
- [13] Marta C. Gonzlez, Cesar A. Hidalgo, and Albert-Laszlo Barabasi, “Understanding individual Human Mobility Patterns,” Nature 453, 2009.
- [14] Raubal, Martin, mraubal, Axhausen, KayW., Yuan,Yihong, Goulias, Konstadinos, yriakidis, Phaeton, Liu, Yu, and liuyu, “Characterizing Human Mobility from Mobile Phone Usage,” ETH Research Database, 2014.
- [15] Rawan Ghnemat, Cyrille Bertelle, and Grard H.E. Duchamp, “Agent-Based Modeling Using Swarm Intelligence in Geographical Information Systems,” Innovations’08, Al- Ain: United Arab Emirates 2008, 2008.
- [16] Rawan Ghnemat, Cyrille Bertelle, and Gerard H.E. Duchamp, “Swarm Intelligence Engineering For Spatial Organization Modelling,” 2009, <https://hal.archives-ouvertes.fr/hal-00430562/document>.
- [17] Stat414, “Probability Theory and Mathematical Statistics,” STAT414/415, <https://onlinecourses.science.psu.edu/stat414/node/97>, 2015.
- [18] Thomas Couronné, Valery Kirzhner, Katerina Korenblat, and Zeev Volkovich, “Some Features of the Users’ Activities in the Mobile Telephone Network,” August 2013.
- [19] [Xiao-Yong Yan, Xiao-Pu Han, Tao Zhou, and Bing-HongWang, “Exact Solution of Gyration Radius of Individual’s Trajectory for a Simplified Human Mobility Model,” Physics, 2010.
- [20] Yan Xiao-Yong, Han Xiao-Pu, Zhou Tao, and Wang Bing-Hong, “Exact Solution of the Gyration Radius of an Individual’s Trajectory for a Simplified Human Regular Mobility Model,” Chinese Physical Society and IOP Publishing Ltd, 2011.
- [21] Yikang Rui, “Urban Growth Modeling Based on Land-use Changes and Road Network Expansion,” PhD thesis, Doctoral Thesis in Geodesy and Geoinformatics with Specialization in Geoinformatics Royal Institute of Technology Stockholm, Sweden, 2013.
- [22] Zahedeh Izakiana, Mohammad Saadi Mesgaria, and Ajith Abrahamb, “Automated clustering of trajectory data using a particle swarm optimization,” Computers, Environment and Urban Systems, 2015.
- [23] Zbigniew Smoreda, Ana-Maria Olteanu-Raimond, and Thomas Couronné, “Spatiotemporal Data from Mobile Phones for Personal Mobility Assessment, Transport Survey Methods: Best Practice for Decision Making,” Springer, Emerald, 2013, pp 745-767, 2013.
- [24] Suhad Faisal Behadili, Adaptive Modeling of Urban Dynamics with Mobile Phone Database, Phd., LITIS, Normany University, France, 2016.



---

\*Corresponding author.

E-mail address: suhad.behadili@ scbaghdad.edu.iq