D4D Challenge
Commuting Dynamics 4 Change

R. Lario[*], R. Maestre[†], M. Muñoz[‡], R. Abad [§], A. Martín[¶], J. Gonzalez [‖], I. del Bosque [**], E. Perez [††]

Paradigma Labs Research Group
Paradigma Tecnológico

Geographic Information Systems Unit
Spanish National Research Council

4-Jan-2013

Abstract

Our idea is to use the geolocation data from the antennas processing the mobile phones calls in order to know which sub-prefectures the customers have been getting around. The main goal of our project is developing spatio-temporal models to detect commuting patterns for the different sub-prefectures, including some other factors related to the region and/or time: wealth, development, infrastructure, investment, grants...

By means of GIS technology, we will be able to apply our generated models to the gathered data and to analyze their correlations over the Côte d'Ivoire surface, working with geographical layers: landcover, roads map, railways lines, water sources... Consequently, the reached conclusions from our study will be properly visualized, allowing a better explanation of the facts. With a bigger amount of data gathered for a longer period, more interesting and accurate trends could be discovered, allowing us to calculate associated coefficients.

Our analysis models will provide coherent data to support a correct urban design and will mean a monitoring tool for development, specially related to population dynamics. In the near future, some other measures could be included. For instance, hospitals and police stations locations, their calls rate... Thus, we could know its real use, being able to improve their service to the citizens: dangerous areas, crowded hospitals...

[*]rlario@paradigmatecnologico.com
[†]rmaestre@paradigmatecnologico.com
[‡]rmaestre@paradigmatecnologico.com
[§]rabad@paradigmatecnologico.com
[¶]amartin@paradigmatecnologico.com
[‖]jgonzalez@paradigmatecnologico.com
[**]idelbosque@cchs.csic.es
[††]eperez@cchs.csic.es

1 Introduction and motivation

The Paradigma Labs Research Group core values are conducted by one motivation: "Figure out the dynamic and fuzzy intersection between Humanity and Thechnology". Therefore a Challenge that main goal is the human improve and developt by means of the study of its mobile communications quickly captured our attention.
Our final aim is detecting spatio-temporal patterns in order to obtain an useful knowledge to better manage the country resources. For example, if we could predict the traffic intensity segmented by road, week day and hour, then another secondary roads could be suggested or the budget for the most used ones could be increased. Trough mobile communitacions, a specific user can be tracked along the day, not only by means of "Call comunications" also by applications running on the user terminal like chats, RSS consumers, etc ... The dataset provides by Orange is a sampled one, and only apply "Call communications", however, we belive that with thw whole set of data i.e.: app and call comunications future models could be several kind of dynamics more accurate.

Our experience studing and moedling several kinds of Human Dynamics like ESF project DynCoop-Net(Solana and Alonso, 2012) or Business Intelligence Tracking Tool on Twitter (Marin et al., 2012), has teach us to explore through two main point of view: Space and Temporal components. We belive that in mathematical model related with Human Dynamics must be present this two components. The Temporal component is useful to provides to the final research a tool to go forward and backward in order to get more deeply understand of the dynamic, The Geographical component provides a more high understand related with the human mobility across the time. Therefore our model and the sudy will carry out by this two imain components.

In this paper, we propose a Space-Temporal Model to achieve this goal. A Space Model, because user interactions with geolocated antennas are analyzed and treated, and a Temporal Model because several time windows are used to group this user dynamics. Later, the mix of this two variables (space and temporal) are consumed and visualized by a Geographic Information System. At the begining several results will be showed, however, the tools with the data will be available to the research comunity to study more deeply the studied dynamics.

Our aim i related to the Commuting concept and could be defined as follow: *Commuting* is regular travel between one's place of residence and place of work or full-time study (Wikipedia, 2012), but sometimes its refers to *any regular or often repeated traveling between locations when not work related*. Our first commuting approach is defined like: "Mobility patterns through inferring dynamic users movements grouped by temporal windows".
A dynamic user is defined like an user changing his antenna location within the studied temporal window (i.e.: each temporal window groups the whole user communication into one specific hour). Among these temporal windows, static users have been removed, i.e.: users that do not change their antennas locations within the temporal range. The justification to remove these users comes to focus our study on users that are moving into this temporal windows and perform micro-displacements.

2 Mathematical model

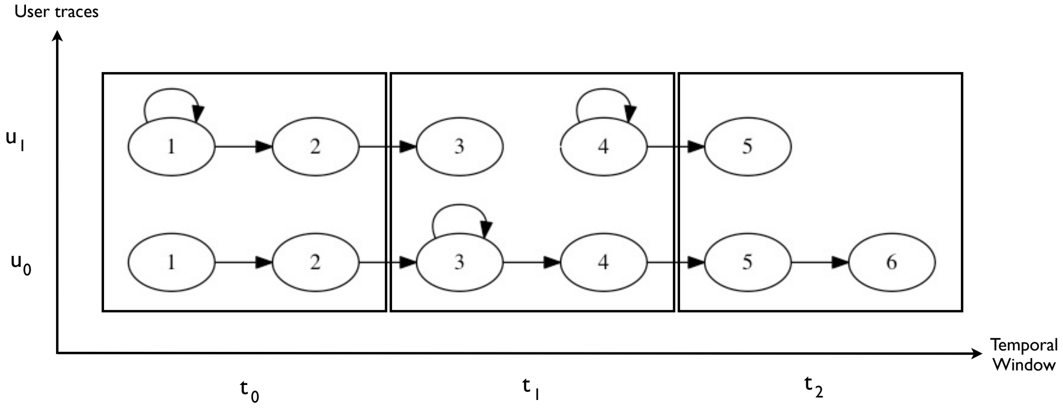


Figure 1: Dynamic and static user patterns

The above figure shows two users u_0, u_1 repesented on the vertical axe, grouped by three time windows t_o, t_1, t_2. A time window is defined as t_n where $n \in \{0, .., 24\}$. Each t_n groups the comunnications traces of the whole set of users in a 60 minutes range.
More formally, an user trace is defined as follows:

$$\vec{T}_{ut} = (p_0, p_1, ..., p_n)$$

where, $p_n \in \mathbb{R}^2$, $n \geq 2$, t is a temporal window range and , u the unique user id.
For instance, $\vec{T}_{00} = (p_1, p_2)$, $\vec{T}_{10} = (p_1, p_1, p_2)$ and so on.

Also, two functions are defined in order to measure the distance (Cook, 2012), given a set of points in spherical coordinates (i.e.: user$_{ut}$):

$$D(p_0, p_1) = \text{acos}(\sin(\phi(p_0^0)) * \sin(\phi(p_1^0)) * \cos(\theta(p_0^1) - \theta(p_1^1)) + \cos(\phi(p_0^0)) * \cos(\phi(p_1^0)))$$

where:

$$\phi(x) = (90 - x) * \frac{\pi}{180}$$

$$\theta(x) = x * \frac{\pi}{180}$$

therefore the function related with distance as defined as follow [result in Km]:

$$U(u, t) = 6373 * \sum_{i=0}^{n-1} D(\vec{T}_{ut}^{\,i}, \vec{T}_{ut}^{\,i+1})$$

The second function is related with the number of antenna connections into a trace, the key point is to count only the dinamic transitions, i.e.: remove the self edges over a given trace as follows:

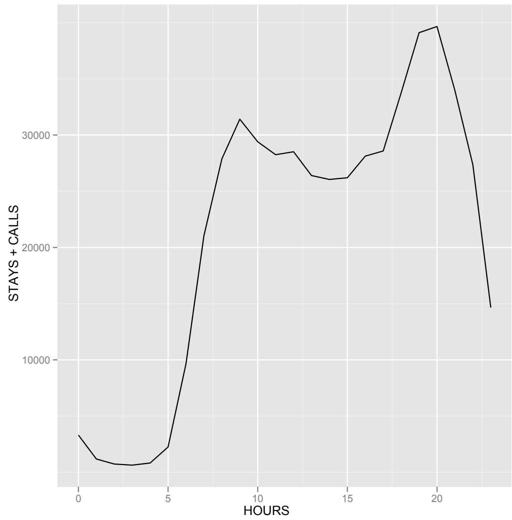
$$\text{S}(p_0, p_1) = \begin{cases} 0 & \text{if } D(p_0, p_1) = 0 \\ 1 & \text{if } D(p_0, p_1) > 0 \end{cases}$$

therefore, the function is defined as follows:

$$N(u, t) = \sum_{i=0}^{n-1} S(\vec{T}_{ut}^{\,i}, \vec{T}_{ut}^{\,i+1})$$

Results

Using the proposed model, the first step is to understand the main component of the commuting dynamic. The result of our approach will emerge from this main component as we can see in the next pictures. This main component model the first logic result and shows the time segments that people have more activite, in call numbers terms.



(a) Total number of calls grouped by Week days

enlazar con que apartir de esta componente temporal, demostramos los picos de lomngitus y un mayor valle central.

Model and Data challenges

to show the last dimension of aour model (the geographical one).
finish to link to whole research work.

Final Conclusions

Several algorithms have been implemented and tested. We observe that there are some main parameters which can be identified, like ρ in order to enhance exploration or to enhance exploitation. Moreover, several approaches have been studied in a real problem through a labeled graph map.
We provide our source code[1] in order to allow researchers and students to check our results and experiment with this algorithms.

[1]https://github.com/yarox/alos

References

John D. Cook. Computing the distance between two locations on earth from coordinates, 2012. URL http://www.johndcook.com/python_longitude_latitude.html. [Online; accessed 4-January-2013].

Oscar Marin, Alejandro Gonzalez, Roberto Maestre, Julio Gonzalez, Marco Martinez, Ruben Abad, and Leonardo Menezes. 15th october on twitter global revolution mapped, 2012. URL http://labs.paradigmatecnologico.com/2011/12/19/15th-october-on-twitter-global-revolution-mapped/. [Online; accessed 23-January-2013].

Ana Solana and David Alonso. Self-organizing networks and gis tools cases of use for the study of trading cooperation (1400-1800). *Journal of Knowledge Management, Economics and Information Technology*, page pp. 402, 2012. ISSN 2069-5934. URL http://www.scientificpapers.org/special-issue-june-2012/.

Wikipedia. Commuting — wikipedia, the free encyclopedia, 2012. URL http://en.wikipedia.org/w/index.php?title=Commuting&oldid=525136196. [Online; accessed 4-January-2013].