# #152

# THE FABRIC OF ENCOUNTER:

## Integration and segregation in the spatiotemporal structure of social networks

VINICIUS M. NETTO
Universidade Federal Fluminense (UFF)
vmnetto@id.uff.br

JOÃO VITOR MEIRELLES
Ecole Polytechnique Federale de Lausanne (EPFL)
joaovitormeirelles@gmail.com

MAÍRA PINHEIRO
Escola Nacional de Ciências Estatísticas (ENCE IBGE)
mairasoares.arq@gmail.com

HENRIQUE LOREA
Universidade Federal do Rio Grande do Sul (UFRGS)
henriquelorealeite@gmail.com

## ABSTRACT

Integrating social and spatial networks will be critical to new approaches to cities as material systems of interaction. In this paper, we propose a way of doing so by focusing on the spatial and temporal conditions of formation of social networks – namely, on 'encounters' as a key social event. Drawing on classic approaches such as Freeman's concept of segregation as 'restriction on contact' and Hägerstrand's time-geography, and recent explorations of social media locational data, we analysed the space-time structure of potential encounters latent in the urban trajectories of agents differentiated by income levels in Rio de Janeiro, Brazil. This approach allows us to estimate agents' urban trajectories examining geographic spatiotemporal positions in tweets, visualise income groups as potentially overlapping class networks, assess spaces of potential encounter and levels of social diversity on the streets. Finally, we discuss our findings and the utility and limitations of this approach in grasping a temporal 'geography of potential encounters' and segregated networks.

## KEYWORDS

Social networks, spatial networks, encounter, segregation, mobility, Twitter locational data.

## 1. INTRODUCTION: LINKING THE SOCIAL CITY TO THE PHYSICAL CITY

> Cities, as Glaeser and Jacobs before him have argued so persuasively, are about 'connections'. . . 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 endeavour (Batty, 2013:30).

Cities are social networks of people and institutions, whose physical organisation allows the exchange of artefacts and information. If linking the social city to the physical city is a major challenge in urban studies (Hillier and Vaughan, 2007; Batty, 2013), the problem of relating social networks to built environments as spatial networks lies at its very centre. In other words, if we are to advance our understanding of the relations of the social city to the physical city, we need to get closer to the very foundational elements of those relations: the city as "sets of actions, interactions, and transactions . . . patterns of flows, of networks of relations, pertaining to both physical-material as well as ethereal movements" (Batty, 2013:9). In this sense, we propose to approach the very material conditions of *formation* of social networks. Social networks are of course formed through opportunities of contact, encounter and interaction as 'social events' in time and space. The city has historically had the role of producing such events.

We will examine in this paper the relationship between forms of social networking and the mobilities of different agents in a city. Closer to a more recent trend in sociospatial studies focused on the positioning of agents in urban space rather than location, our approach explores ways in which the formation of networks is shaped by the potential of encounter within the *trajectories* of bodies in urban space where co-presence is likely to happen – or not. In fact, this aim involves entering an elusive fabric of movement and encounter. This paper develops a method to capture such fabric, using ideas from Linton Freeman's (1978) view of segregation as 'restrictions on contact' to references to the time-geography of Torsten Hägerstrand (1970).

We will develop these ideas in order to achieve a clearer understanding of (i) the role of encounter in the formation of networks, and its opposite, in segregation as a real time phenomenon, and (ii) the role of mobility in the opportunities of encounter, and its relation to social differences, especially in unequal societies. Then we shall (iii) develop a concept of sociospatial networks able to represent movement and potential encounters in time and space; (iv) explore the methodological use of social media locational data to grasp movement and infer potential encounter; (v) apply this framework in an empirical study of networks of trajectories and encounters of different income groups in Rio de Janeiro; and finally (vi) discuss our experiment and findings on sociospatial networks and their levels of superimposition.

## 2. THE ROLE OF ENCOUNTERS IN THE FORMATION OF SOCIAL NETWORKS

Cities may be seen as a fluctuating balance of density, mobility and social connectivity (Bettencourt, 2013). Connectivity is of course highly dependent on how encounters are generated by density and mobility. Encounters can be dispersed in the streets or polarized in places of work, leisure and consumption, at bus stops, subway stations, institutional buildings and so on. These factors may have an impact on our interactions, like sparks to a dense web of daily movements from residential locations. If movement could leave visible traces in space, such web of movement could reveal the potential to encounters and opportunities for social network formation unfolding in urban space.

Mapping these webs of movement in the city where encounter may or not happen is one of the aims of this paper. In fact, the idea of mapping trajectories is far from new. The work of Hägerstrand (1970) was the first systematic attempt to capture trajectories and restrictions spatiotemporal hanging over actions. Although Hägerstrand's approach – fashionable in the early 1980s – has lost interest since then, recent empirical approaches have taken the spirit of that work (e.g. Lee and Kwan, 2011), making use of technologies capable of recording the movement of agents and identify patterns of mobility[1]. We propose to add new layers to this idea, and evaluate how the mobilities of agents shape opportunities of encounter. Webs of movement are of course evanescent features of our effective presence in space. If we could capture at least some of them, we could have a picture of how social groups materialize their actions.

1 A method developed by Gonzales et al (2008) uses an extensive database recorded through mobile phone communication in American cities to map spatial paths, showing that actors have a remarkable tendency to recursivity.

We would like to explore an alternative definition of 'social network' intended to get closer to the web of encounters through which networks may emerge. We consciously opt to not use the concept as an arrangement of agents as in *Social Network Analysis* (SNA)[2]. The SNA tradition focuses on the analysis of phenomena varying from the microstructural, like epidemics or power relations or the spread of information within groups, to the large-scale, such as the small-worlds networks – frequently involving graph theory and a space without physical and temporal dimensions, an abstract space of pure topology. We choose to employ a definition of social network as an open and *potential* set of contacts changing over time – one able to take into account the social positions of the agents and the precise circumstances of time-space where contact may occur. Graphically and mathematically, we do not represent agents by vertices and relationships by links. Instead, we invert this representation, seeing agents as "lifelines" (as in Hägerstrand), with the important factor of the passage of time – which allows us to retain the dynamic property of a social system. The possibility of agents encountering each other is represented by the intersections of the agents' lifelines. Encounters are the vertices, and agents' lifelines, the potential links between them. This non-standard representation is favoured by a principle of homology in which lifelines correspond to urban trajectories, and circumstances of encounter correspond to converging positions (figure 1). This model seeks to *add the temporal and spatial dimensions as inherent dimensions of social networking*, and render the spatiality of encounter more intuitive. In short, it is intended to account for the potential of encounter as a key factor in social network formation.

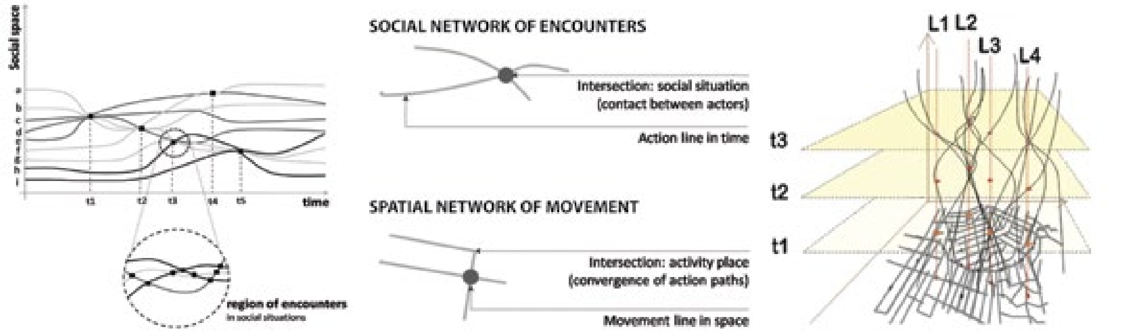


Figure 1 - Principles of homology between networks in time and social space (left), diagrammatic spatial translation (centre), and agent's paths in space-time (right).

Once we map agents' lifelines in space-time (via the trajectories between collected positions), we leave a purely 'social' representation of networks. We are looking at *sociospatial* networks: the places of overlapping trajectories are the connections of potential visual contact and encounter. We do not wish to equalise 'sharing space' with 'encounter as interaction', however. We cannot claim to grasp the passage from *encounter to interaction*, for that would require observations of agents in their actual exchanges, which are outside our methodological scope. Following the work of Goffman (1961), Giddens (1984) and Hillier and Hanson (1984), we understand 'encounter' as the presence of an agent in one's perceptual field in circumstances of co-presence. As the raw material of social life, the importance of encounters can hardly be over-emphasised.

## 3. CONTACT BETWEEN SOCIALLY DIFFERENT AGENTS

An interesting view of the role of encounters in social networks is found in Freeman (1978: 413): "All restrictions on interaction, whether they involve physical space or not, are forms of segregation – in social space." We wish to explore Freeman's view in order to understand the delicate fabric of encounters and interactions that keeps local social systems together. Mapping movement can allow an understanding of the spatiality of presence and absence as active

2 See classic graph theoretical approaches in Freeman (1978) and Wasserman and Faust (1994), among others.

features of social networking and the emergence of segregated networks. In this sense, our approach shares the focus on social spheres and routinized activities found in a number of more recent works (e.g. Schnell and Yoav, 2001; Lee and Kwan, 2011; Selim, 2015).

What is the chance to meet someone from a different social group? We shall first examine a number of conditions of encounter reasonable from a material standpoint.

- First, social contact in cities depends on circumstances of encounter.[3]
- Second, cities are historically produced and spatially structured so as to make social situations in principle accessible. In fact, city-making processes are consistently related to patterns of location and accessibility. In a tradition stemming from Hansen (1959) to Glaeser (2010), approaches in spatial economics have been able to identify agents' preferences and location patterns amidst the apparent randomness of location.
- Third, activity places tend to increase the potential for convergence of agents who share similar interests and mobilities.[4]
- Forth, income may play a role in this process. People with smaller budgets face further restriction in mobility. In turn, limitations in mobility *enhance localism* – the dependency on proximity to enact one's social life (Fischer and Shavit, 1995; Lee et al., 2005). In these cases, the density of encounters would tend to increase especially around home, and agents would tend to use places in the neighbourhood to create and maintain relationships.[5]

Our hypothesis is that similarities in patterns of mobility and appropriation of space (the spaces we are likely to relate to, use or pass by) would lead to increases in the potential density of encounters especially between socially similar agents. In turn, this spatial trend toward both higher levels of homophily and different degrees of connectivity in personal social networks, both generated by differences in income, lifestyles and mobility, may have strong implications for contact between the socially different. We need to clarify how contact is effectively performed spatially, involving circumstances of co-presence and absence. For instance, could *mobility* – and not *proximity* – be the key factor in generating potential co-presence between the socially different? Approaches to mobility that make use of geographic information derived from digital data (say, the data usage of mobile phones) are still restricted to capture spatial patterns of behaviour (e.g., Gonzales et al, 2008) – with little connection with the social conditions of spatial behaviour, such as the influence of income and class.

Mobility and income seems associated in a circle that leads to increases or decreases in the potential to create, maintain and expand personal networks. But how so? If networking depends on situations of encounter, we need to understand how mobility matters in the structure of urban encounters. How (and where) does the potential of encounter between the socially different materialize? In order to answer these questions, we need to examine the *superimposition of trajectories* of different social groups. Income groups related to class, i.e. large-scale groups with common economic features that strongly influence their actions and lifestyles (Giddens, 1993) and other forms of social grouping are shaped by probabilities of encounter.

Space matters here. Even though we do not usually think about it, our daily trajectories constitute the backbone of our encounters and shape the elusive structure of social life in the city. The distance between locations in a city, associated with differences in mobility, income and lifestyle could bring inequalities in the capacity to participate in social situations. Inequalities and incompatibilities in patterns of movement are forms of *disjunction of encounters* – a way of disrupting the possibility of encounters that otherwise could happen. The disjunction of encounters may be especially active among socially different people. Simply put, there would be a greater chance of encountering and networking with those who share similar mobilities.

3 The heart of this idea may be found in Jacobs (1969); see also Giddens (1984), Hillier and Hanson (1984), Bettencourt (2013), and Batty (2013).

4 Bettencourt (2013) has recently theorized the effects of linear paths over the density of encounters.

5 See Marques (2012). Empirical data on transport expenses in Brazil show that higher income groups not only spend more than low-income groups, they spend more than proportionally (POF, 2009).

These ideas begin to portray the complex material fabric of encounters in a city. However, how can we understand in detail its volatile spatiality?

## 4. THE METHODOLOGICAL USE OF TWITTER LOCATIONAL DATA

It seems almost impossible to see the spatiality of the tremendously complex flows of convergences and divergences of our actions and paths in the city. So a key methodological question is how to grasp the panorama of an entire city. The answer is that we can track the movement of a large number of agents exploring the potential of social media locational data. In fact, a number of works has recently emerged using social media locational data in order to extract information of human patterns of movement. On a substantive level, Liben-Nowell et al (2005) related geography and online social networks (a community of bloggers) to find that one-third of relationships are not dependent on geography. Lee et al (2011) examine how the use of mobile communication channels of information affects just-in-time choices in consumption travelling behaviour. On a methodological level, Li et al (2011), Ribeiro et el (2012) and Zielinski and Middleton (2013) developed forms to infer indirect locations from Twitter geotag and timestamp, whereas Veloso and Ferraz (2011) and Takhteyev et al (2012) inferred spatially reliable information through regression models correlating tweet frequencies with real world events. Sakaki et al (2010) filtered georeferenced tweets, whereas Boettcher and Lee (2012) applied density-based spatial clustering.

In the spirit of these works, we conducted an empirical study in the city of Rio de Janeiro. This study is intended as a proxy to the actual dynamic scenario of trajectories of socially differentiated agents. Twitter offers particularly attractive possibilities in this sense, as it makes its metadata bank public through a principle of anonymity and the possibility of inferring characteristics of the spatial behaviour at the individual level, involving potentially large samples – although risks of generalizing from self-selecting users to the populations from which they are drawn must be carefully taken into account (Longley et al, 2015). The set of variables provided by Twitter API includes user IDs along with a spatiotemporal signal, the timestamp and geographic coordinates for each tweet posted by users who opted for having the GPS location in their mobile phones turned on. Considering the relation between tweet location and the actual street network, our study points to an accuracy within 10 meters, adjusted to the street network via shortest distance to the nearer street segment mapped in GIS software.

We collected metadata from tweets with spatiotemporal positions posted in Rio through the official Twitter streaming API between November 12th (0:07:13 am) and 14th (2:36:45) during a period of 56 hours, generating a database of 20,192 users and 333,407 tweets. Then we tested this time frame against a 241 hours database, with 70,403 users and 2,252,348 tweets collected along 18 days, and found a Pearson linear correlation of 0.976 (p-value 2.2e-16) between the datasets regarding the spatial distribution of tweets according to census blocks.

Automatic tweeters posting for commercial purposes (bots), identifiable by the large number of tweets posted from the same position, were also excluded. A statistical analysis of average distances between tweet positions of same users via quantile classification showed that the threshold between (high frequently) short distances and long distances between tweets was 106 meters (step 2). We identified the repetition of location of the first tweet in the morning during the period of observation (first in a sequence of tweets), since the night position brings limitations to the sample regarding tweeting behaviour (Longley, 2015). The first tweet was taken as the origin in the generation of shortest paths to following tweets posted during the day (step 3). Shortest paths between tweet positions within Rio's street network were performed topologically through GIS software, using Open Street Maps (step 4) as predictors of routes between actual positions.[6]

6 Empirical evidence from fields like space syntax and urban network studies suggests that shortest paths are reliable predictors of actual routes (e.g. Hillier et al, 1993).

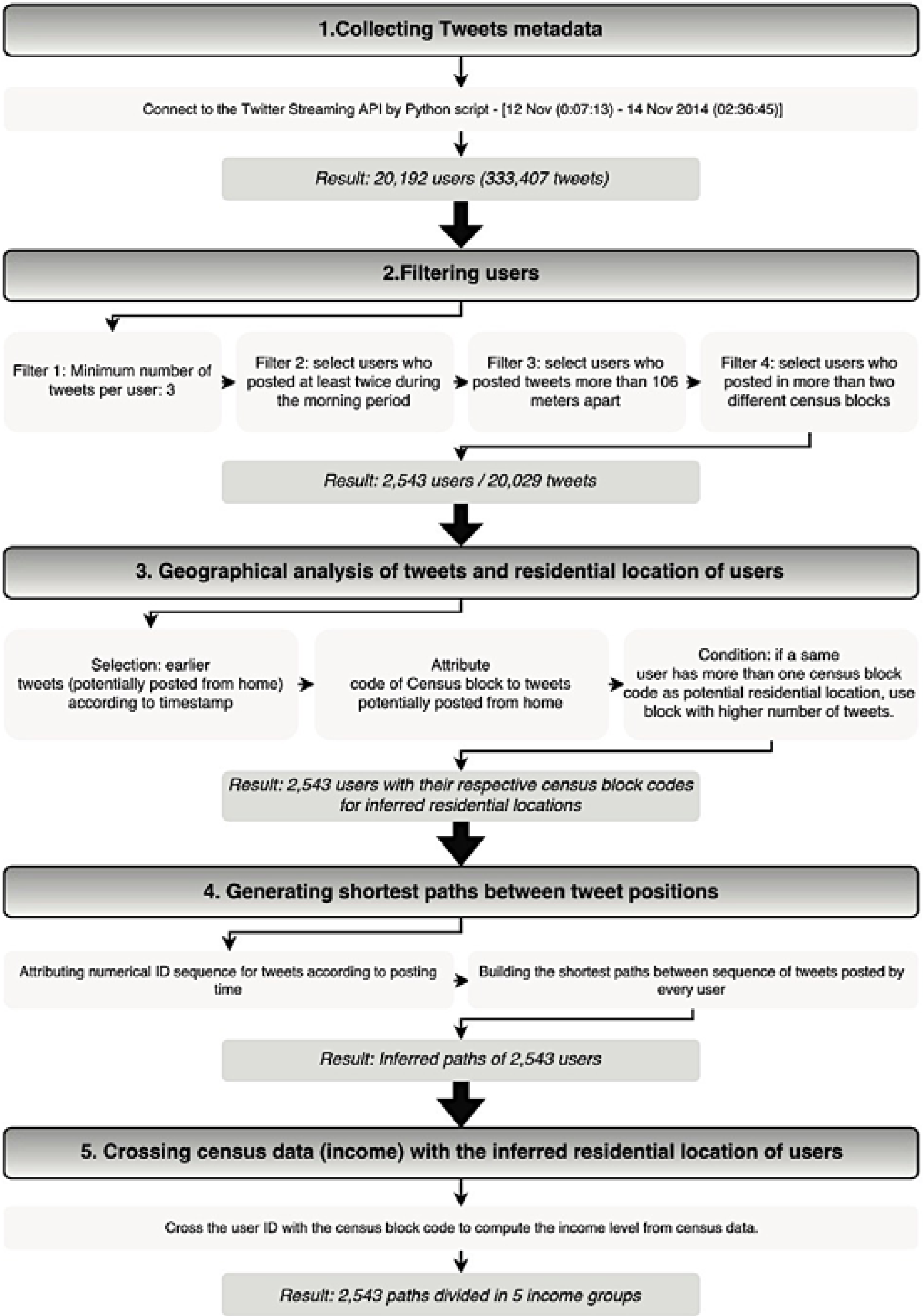


Figure 2 - Methodological procedure

Finally, Twitter users were differentiated according to income (step 5). We assigned income levels to users through a procedure that required crossing residential locations inferred from tweeting behaviour (step 3) with economic data referred to the census blocks (2010 Census, Brazilian Institute of Geography and Statistics, IBGE). The census block unit used to infer income levels was the smallest available[7]. Our methodological procedure to select users according to posting behaviour, spatial behaviour and income levels can be summarized as follows (figure 2):

Of course there are risks of ecological fallacy in interpreting individual users' income from the average of residents from each census block. We statistically assessed such risks looking for the coefficient of variation (CV) for income. The average CV for income within census blocks in Rio is low, about 9.3%, so we could say that there is enough homogeneity of income values between residents of the same block to use average income as a proxy of individual income. We analysed the exponential distribution of income per capita and proposed a classification by quantile which suggested the following levels: less then R$ 750; R$ 750.01 to R$ 1,600; R$ 1,600.01 to R$ 2,500; R$ 2,500.01 to R$ 3,400; R$ 3,400.01 and above[8].

These values were identified as low, lower-middle, middle, upper-middle and high-income users (differentiated by colours in figure 3). For the sake of this experiment, a methodological test involved assessing how representative are inferences about Twitter users in relation to the actual population of Rio de Janeiro. We compared the distribution of the average per capita income in Rio's population from census data and from Twitter users inferred through residential location.

The first histogram (population income – figure 3, bottom left) suggests an exponential distribution with a long tail for higher income values (over R$ 10,000 per month), the same threshold observed for the estimated income distribution of Twitter users. Linear regression brings an adjusted R squared of 0.67, showing that the inferred income distribution of users has a reasonable degree of similarity with the income distribution of the population in general (figure 3, bottom right).

This also suggests that the use of Twitter does not seem to be associated with specific income levels, confirming previous findings about the high penetration rate of Twitter in Brazil (Graham and Stephens, 2012). A geographic analysis in figure 3 shows two readings: residential patterns of income levels (top in figure) and the pattern of distribution of users' estimated location according to income level (below in figure 3).

## 5. NETWORKS OF ENCOUNTER IN TIME AND SPACE: A DIGITAL EXPERIMENT

We may further analyse the spatial and temporal structure of potential encounters between Twitter users. Using as databases the OSM street network and the tweets dataset with timestamps and geo-location to rebuild the shortest paths between consecutive tweets, we also temporalized such trajectories assuming an average speed between tweet locations. We inferred 'potential encounter' as crossing trajectories within a single street segment, of course, and within a 'temporal buffer' of five minutes. In other words, we computed as potential encounter situations where two users were in a same street segment within a 5-minute interval.

7 Within the city of Rio de Janeiro, the census block unit has an average number of 210 households and 616 Inhabitants and a median area of 33,017 squared meters (a considerable variance is found for area).

8 The exchange rate between American Dollars and the Brazilian currency Reais is: U$ 1 = R$ 3.30 in 30th September, 2016.

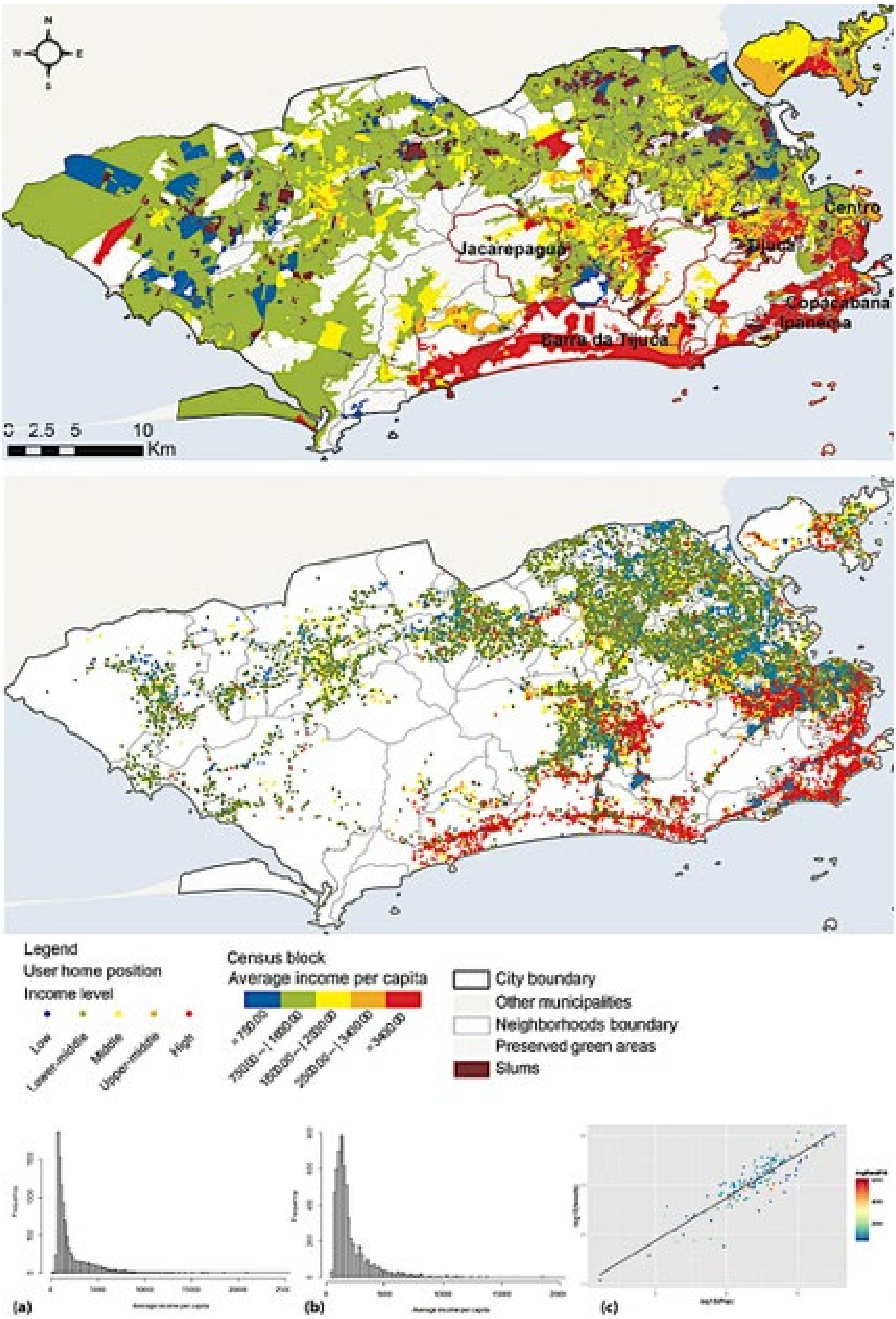


Figure 3 - Income levels in census blocks units (blue to red, top), and estimated locations of Twitter users (below). Below, histograms of average per capita income in Rio's population (left) and Twitter users (centre); regression between users (Y) and population (X) in urban districts (right).

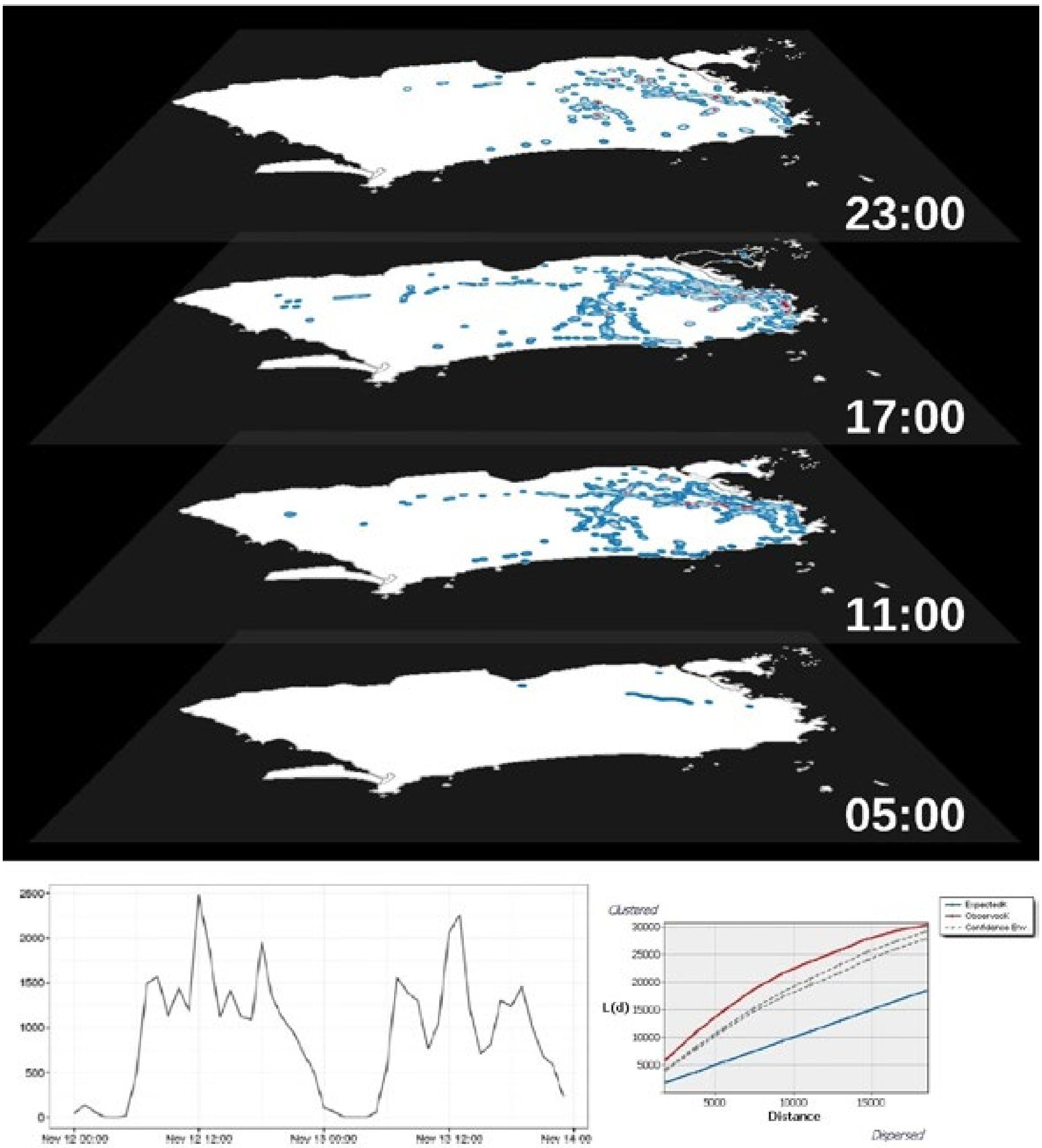


Figure 4 - Space-time prism for the varying intensity of inferred encounters between Twitter users in Rio (top). Number of potential encounters in time (bottom left) and an analysis of their clusterization in space (tK-Function, bottom right).

We may assess the space-time structure of encounters through a time-geography-inspired representation and complementary graphic analyses (figure 4). Not surprisingly, the number of potential encounters peak in the early morning, around midday and around 6pm, to drop considerably in the evening, as agents tend to find themselves in more static positions in space (figure 4, bottom left). In order to understand the spatial pattern of potential encounters, we applied Ripley's K-function to summarize spatial dependencies interactively as clustering or dispersion processes over a range of distances randomly selected. As first stated by Getis (1984) and revisited as an analytic tool by Mitchell (2005), the K-function was calculated as:

$$L(d) = \sqrt{\frac{A \sum_{i=1}^{n} \sum_{j=1,\, j\neq 1}^{n} k_{i,j}}{\pi \times n(n-1)}}$$

Where
d = distance between places of potencial encounters
n = total number of places of potencial encounters
A = total area comprehended by potencial encounters
$k_{i,j}$ = weight (number of potencial encounters in each place).

The blue line in the graph for the K-function (figure 4, bottom right) indicates the average distance between encounter places is randomly selected sets. The red line indicates all observed encounters (in our case, inferred encounters). The X-axis represents the distance between encounter places and the Y-axis represents the average distance between encounter places weighted by the number of encounters. The difference between the lines indicates that observed encounters are clustered. As the curve for the observed K is above the confidence envelope, it is statistically significant.

What does our experiment show about the dynamic of potential encounter as the superimposition of *class networks* as revealed by Twitter users' trajectories in the city? We counted the number of agents' trajectories classified by income level for each street segment (between corners) where there were trajectories. We also quantified the length of streets covered by users with the same metric length of street segments. This information was registered for each agent and accumulated for her/his income groups. Then we calculated the overlapping of income groups, using the number of agents for each group passing through each street segment. Maps in figure 5 show the dominant income group in the streets that make up their trajectories. The criteria for determining visually the dominant presence of a single group over a street segment is 'the group with the higher number of paths overlapped in a street segment provides the colour for that segment' – i.e. once we consider the proportion of income groups in actual numbers, when a group has one person above that percentage, it has dominant presence. Results of the analysis seem to grasp the spatiality of potential encounter along with traces of dynamic segregation. These visual overlaps will be assessed quantitatively below.

We first notice strong evidences of residential segregation related to income. The poorer spread more broadly over the cityscape. Low-income and lower-middle income groups show considerable overlap, but low-income users are dominant in areas farther from the sea and the CBD, located in the East. Landscape amenities related to proximity to the sea are a clear factor in defining land values and higher income location (and both territorial and dynamic segregation) in Rio. Geographical topography adds complexity to residential location patterns in Rio – including the *favelas* scattered in the landscape, allowing the poorer to live in hills near richer areas, closer to the sea and the CBD. Such unusual cityscape enmeshes the lines of movement, as we can see in areas in South Rio, increasing the potential of encountering the socially different. Complexities considered, an overall pattern emerges, as higher income users are more likely to be found in South and Southeast Rio, near to the sea, and a gradual shift in users income levels becomes visible in trajectories as they spread farther, towards North.

We can also assess how isolated is the presence of a single class in Rio's streets, and how the proportion of trajectories income groups share with one another (table 1). Lower-income groups (IG1 and IG2) are much more segregated in their movements across the city, with 19.2% and 29.9% of their trajectories occurring in non-shared streets, respectively. They also show the highest degree of sharing spaces (10.4%). Richer groups (IG4 and IG5) share more of their trajectories with other groups.

The poorer and the richer (IG1 and IG5) share only 0.8% of users' trajectories. IG2 displays a more socially integrative spatial behaviour – but also has a larger share of users (46.7%). IG1 and IG2 trajectories display less social diversity – they are easily the dominant group (i.e. their presence in a particular space is above the average of its proportion in the total number of agents, all groups considered). The fact that IG1 consists of 23.6% of total users and are dominant in 30.5% of streets where they pass through suggests they are more segregated than other groups in their movements.

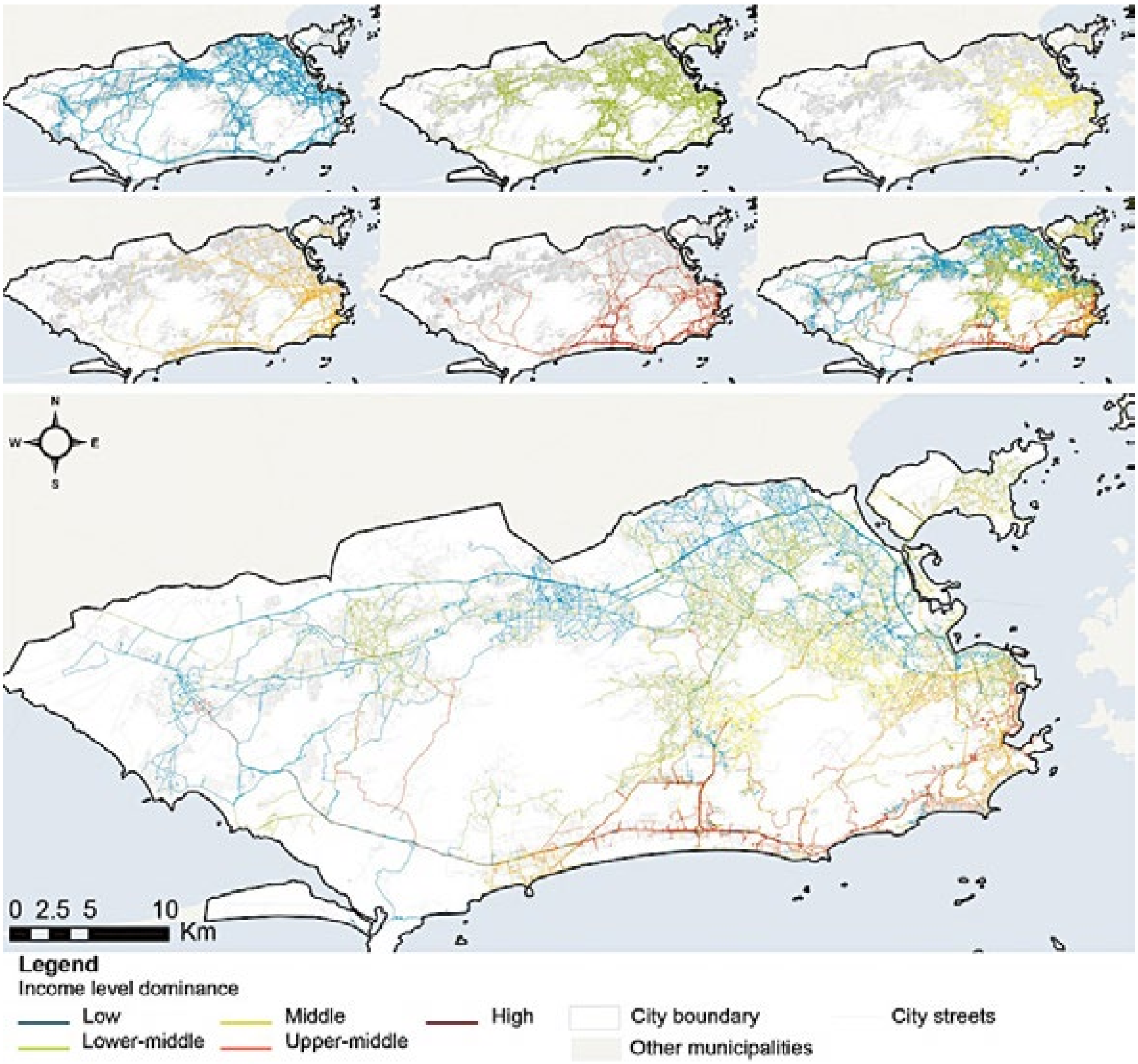


Figure 5 - A picture of segregated networks: blue (low income), green (lower-middle), yellow (middle), orange (middle-upper) and red (high income) groups. The larger map shows the dominant class network.

| | IG1 | IG2 | IG3 | IG4 | IG5 |
|---|---|---|---|---|---|
| IG1 | 19.2% | 10.4% | 0.5% | 0.5% | 0.8% |
| IG2 | | 29.9% | 1.5% | 1.7% | 1.2% |
| IG3 | | | 4.3% | 0.2% | 0.3% |
| IG4 | | | | 4.3% | 0.7% |
| IG5 | | | | | 4.6% |

Table 1 - Matrix of the proportion of streets (regarding the total number of streets) appropriated exclusively by a single income group (italic), and the proportion of streets shared by different income groups in their daily paths.

Where are exactly the spaces that different social groups share? Poorer income users (IG1xIG2) share much more spaces when appropriating the city, mostly in North and West Rio. Pairing the poorer and the rich (IG1xIG5), map show that the world-famous South Rio (Ipanema) is not merely a segregated area marked by the dominating presence of the rich. It is a major area for mutual visibility. Now considering the relationship between *dynamic segregation* and *residential segregation*, how much do other income groups actually pass through territorially segregated areas? We assessed this relation crossing the average income in residential sectors (census blocks) with the average income of the dominant group passing through those areas (table 2). However present in richer sectors, poorer groups (IG1 and IG2) strongly concentrate in poorer areas: 72.1% of IG1 trajectories happen in low-middle income sectors (S2). Richer users (IG4 and IG5) still appropriate rich sectors (S5): 58.58% of IG5 trajectories happen in S5 areas. In turn, middle-income and middle-upper income sectors (S3 and S4) are open to more diverse income groups (table 2).

| | | IG1 | IG2 | IG3 | IG4 | IG5 |
|---|---|---|---|---|---|---|
| Residential Sectors | S1 | 9.1% | 3.9% | 3.6% | 1.7% | 5.2% |
| | S2 | 72.1% | 52.4% | 28.0% | 18.2% | 17.2% |
| | S3 | 13.5% | 30.1% | 29.3% | 15.9% | 7.4% |
| | S4 | 2.0% | 8.4% | 20.5% | 17.5% | 11.4% |
| | S5 | 3.3% | 5.2% | 18.7% | 46.7% | 58.8% |
| | | 100% | 100% | 100% | 100% | 100% |

Table 2 - Proportion of presence of income groups in residential sectors, considering local average income.

What does this pattern of overlapping imply in terms of potential encounters between socially different users? What are the effects of different patterns of spatial trajectory on encounter opportunities? We generated a social network analysis of agents grouped according to income. Links show the number of potential encounters identified between pairs of income groups (vertices in figure 6). The shorter and thicker each link is, the higher the number of encounter between income groups. Encounters are seen as more likely between socially similar agents.

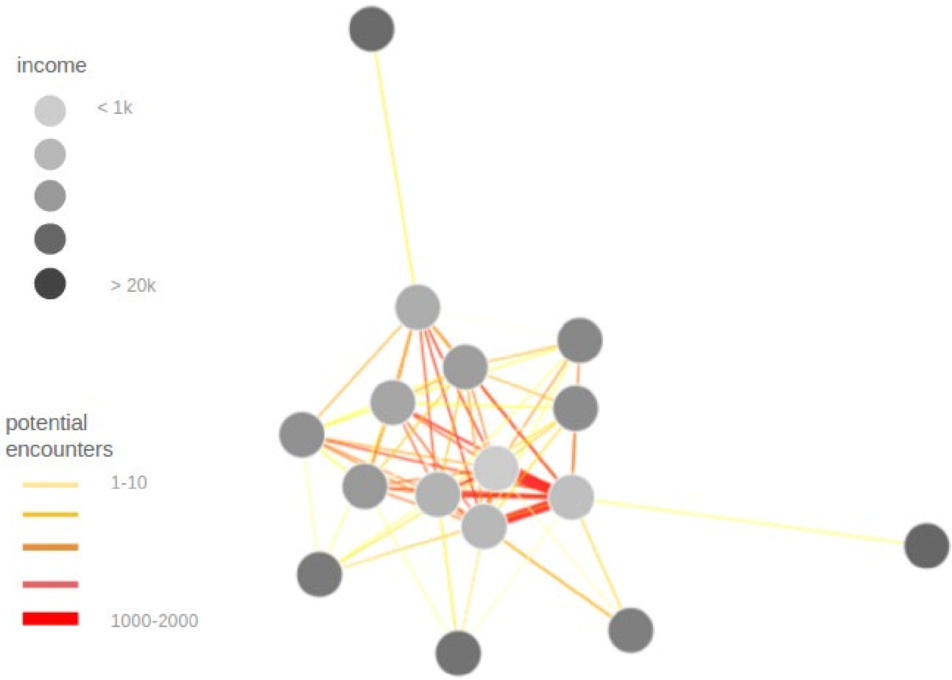


Figure 6 - Intensity of encounters between different income groups.

Finally, where do class networks converge more intensely? What are the streets with more social diversity, where 'the other' is more likely to be seen? We measured *social diversity on the streets*, i.e. the level of superimposition of networks, through Shannon information entropy, calculated as the participation of each class over the total number of agents (spaces with the presence in equal shares of all income groups contain the highest entropy), and associated different entropy levels with colours from blue to red (figure 7). Entropy was calculated for every street segment; intervals were defined via natural breaks.

$$Entropy = -\sum_i P_i log_2(P_i/Pt)$$

where $P_i$ is the total number of users with income i and Pt the total number of users passing by every street segment

There is a small network of socially convergent streets, an interesting superimposition of trajectories of users of all income groups around South Rio (Copacabana and Ipanema) and the CBD (on the East). Spaces of social convergence are to be found in denser, busier areas like the CBD and South Rio, or major centralities like Tijuca and Jacarepaguá, a little to the North. These are the most likely spaces to find socially different agents.

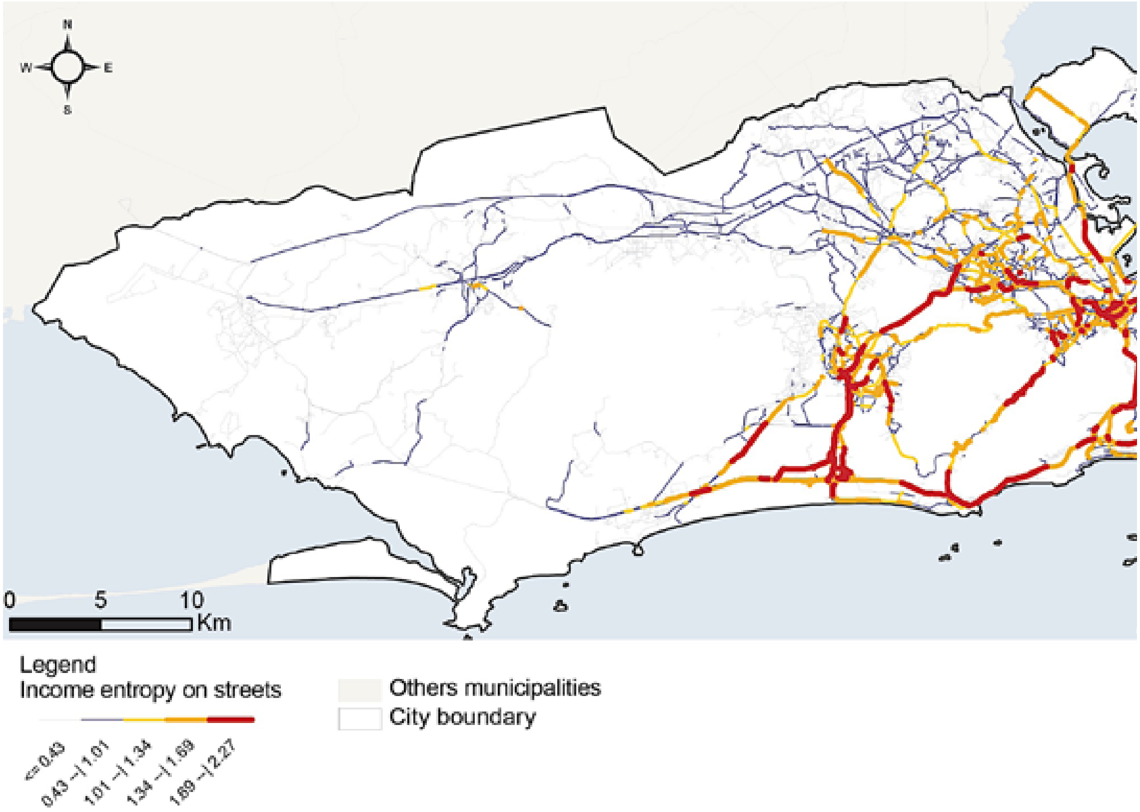


Figure 7 - Entropy map showing in red the spaces with the highest social diversity

## 6. CONCLUSION: SPACE, TIME – AND SEGREGATION – IN SOCIAL NETWORKS

Changing the focus from social networks centred on agents to the networks of 'encounters' performed in daily trajectories as a key spatial and temporal event in the formation of social networks, this approach must be seen as an experiment in identifying a 'geography of overlapped networks' in urban space – and, following Freeman's definition, into segregation as 'restrictions on contact'. Of course such intent poses a number of questions: can poorly overlapping networks be interpreted as segregation? Is 'space sharing' enough to depict social integration? Unlike previous works, our approach is geared to *trace movement*, relate it to patterns according to social differentiation (in this case, based on income) and assess their role in inferred trajectories of agents in urban space. As a proxy to the scenario of potential encounter and segregated networks, this experiment based on Twitter locational data can only show *trends* within the trajectories of a limited number of agents. Nevertheless, it suggests that different patterns of mobility lead to less opportunities for encounter, as seems to be the case between poorer and richer users. If Freeman (1978) is right in asserting that segregation operates through restrictions on contact, (the lack of) space sharing and co-presence are an essential part of the experience of segregation.

Is a study based on Twitter data enough, however? Due to difficulties in generalising conclusions from samples of self-selecting users (Longley et al, 2015), procedures assigning location and income to users must be seen as a *proxy* rather than an actual scenario, as we insist. As such, our study suggests that Twitter data is an invaluable means of *identifying patterns* of movement of agents, with strong possibilities for understanding matters of social integration and equity.

Social media data are also a potential source for generating a precise *'geography of encounters'* in a city, including the temporal dimension – a previously virtually impossible achievement. Graphic and quantitative analyses of overlapping networks seem to add another layer to the understanding of segregation through static maps of income levels and segregated activity or residential location in Rio (in that spirit, compare figures 3 and 5). This is the very purpose of our proposition: to get closer to a temporal geography of potential encounters and segregated networks of movement *and* the public spaces with different potentials for overlapping socially different agents, for the first time monitoring and measuring through locational data spatiotemporal differences in the appropriation of a city by members of different income groups. Neither segregated movement nor potentials for overlapping networks is inferable from income, activity or residential distribution maps alone.

In this sense, our approach suggests that the probability of encounter is impregnated with spatiality, interacting actively with the structure of the street network to generate potentials of convergence and co-presence of social groups. The odds of finding 'the other' seem distributed according to the spatial and temporal frames of action of different groups within a city. The paths of Twitter users suggest greater compatibility between certain users – and, by extension, greater potential for interaction. On the other hand, differences in urban trajectories in daily life may lead to the reduction of opportunities of encounter. This view into the space-time structure of potential encounters also allows a form of bringing to the forefront the complexity of segregation captured as a highly dynamic 'disjunction of encounters' at the level of trajectories of agents, close to Freeman's seminal definition of social segregation as 'restrictions on contact'.

## REFERENCES


Batty M (2013) *The New Science of Cities*. Cambridge, MA: The MIT Press.

Bettencourt LMA (2013) The origins of scaling in cities. Science, 340, 1348-1441.

Boettcher A and Lee D (2012) EventRadar: A real-time local event detection scheme using Twitter stream. In *Proceedings of the IEEE International Conference on Green Computing and Communications*, Besançon, France, 358–67.

Fischer C and Shavit, Y (1995) National differences in network density: Israel and the United States. *Social Networks*, 17.2, 129–45.

Freeman L (1978) Segregation in social networks. *Sociological Methods and Research*, 6.4, 411–429.

Giddens A (1984) *The Constitution of Society*. Cambridge: Polity Press.

Giddens A (1993) *Sociology*. Cambridge: Polity Press.

Goffman E (1961). *Encounters*. Indianapolis: Bobbs-Merrill.

Glaeser E (2010) *The triumph of the city: how our greatest invention makes us richer, smarter, greener, healthier and happier*. New York: Penguim.

Gonzales M, Hidalgo C and Barabási A-L (2008) Understanding individual human mobility patterns. *Nature*, 453, 479-482.

Graham, M. and Stephens, M. (2012) A Geography of Twitter. *WWW document*, http://www.oii.ox.ac.uk/vis/?id=4fe09570.

Hägerstrand T (1970) What about people in regional science? *Papers of the Regional Science Association*, 24.1, 6–21.

Hansen WG (1959) "How acessibility shapes land use." *Journal of the American Institute of Planners*, 25, Issue 2.

Hillier B, Hanson J (1984) *The Social Logic of Space*. Cambridge: Cambridge University Press.

Hillier B, Penn A, Hanson J, Grajewski T and Xu J (1993) Natural movement: or, configuration and attraction in urban pedestrian movement. *Environment and Planning B: Planning and Design 20*(29): 29-66.

Hillier, B; Vaughan, L; (2007) The city as one thing. Progress in Planning , 67 (3) pp. 205-230. 10.1016/j.progress.2007.03.001.

Jacobs J (1969) *The Economy of Cities*. New York: Random House.

Lee R, Ruan D and Lai G (2005) Social structure and support networks in Beijing and Hong Kong. *Social Networks*, 27.3, 249–74.

Lee J and Kwan M-P (2011) Visualisation of socio-spatial isolation based on human activity patterns and social networks in space-time. *Tijdschrift voor Economische en Sociale Geografi*e, 102.4, 468–485.

Lee KH, Lippman A and Pentland A (2011) The Impacts of Just-In-Time Social Networks on People's Choices in the Real World. *The Third International Conference on Social Computing*.

Li W, Serdyukov P, Vries AP, Eickhoff C, and Larson M (2011) The where in the tweet. In Proceedings of the Twentieth *ACM International Conference on Information and Knowledge Management*, Glasgow, Scotland.

Liben-Nowell D, Novak J, Kumar R, Raghavan V and Tomkins A (2005) *Geographic routing in social networks*. PNAS, 102.33, 11623–11628.

Longley PA, Adnan M and Lansley G (2015) The geotemporal demographics of Twitter usage. *Environment and Planning* A 47.2, 465–484.

Marques, EC (2012) Social Networks, Segregation and Poverty in São Paulo. *International Journal of Urban and Regional Research*, 36.6, 958-979.

Ribeiro S, Davis C, Oliveira D, Meira W, Gonçalves,T, and Pappa G (2012) Traffic Observatory: A system to detect and locate traffic events and conditions using Twitter. In *Proceedings of the Fifth International Workshop on Location-Based Social Networks*, Redondo Beach, California: 5–11.

Sakaki, T, Okazaki M., and Matsuo Y (2010) Earthquake shakes Twitter users: real-time event detection by social sensors. In *Proceedings of the Nineteenth International Conference on the World Wide Web*, Raleigh, North Carolina: 851–60.

Schnell I, and Yoav B (2001) The Socio-Spatial Isolation of Agents in Everyday Life Spaces as an Aspect of Segregation, *Annals of the Association of American Geographers*, 91.4, 622-633.

Selim G (2015) The landscape of differences: Contact and segregation in the everyday encounters. *Cities*, 46, 16-25.

Takhteyev Y, Gruzd A, and Wellman B (2012) Geography of Twitter networks. *Social Networks 34*: 73–81.

Veloso A and Ferraz F (2011) Dengue surveillance based on a computational model of spatio-temporal locality of Twitter. In *Proceedings of the Third International Conference on Web Science*, Koblenz, Germany.

Wasserman S and Faust K (1994) *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

Zielinski A and Middleton S (2013) Social media text mining and network analysis for decision support in natural crisis management. In *Proceedings of the Tenth International Conference on Information Systems for Crisis Response and Management*, Baden-Baden, Germany.