

Vincent Blondel, Gautier Krings, Isabelle Thomas

Regions and borders of mobile telephony in Belgium and in the Brussels metropolitan zone

Translation: Jane Corrigan

In this article, we examine the existence of regions and borders in Belgium by means of an original data set of more than 200 million mobile communications and a mathematical method which allows networks to be divided into coherent groups in a natural and automatic manner. A geography of mobile communications is thus proposed based on the relative frequency of communications as well as their average duration. In the case of the relative frequency of communications, we obtain a map of Belgium composed of 17 groups or 'telephone areas' that each contain only adjacent municipalities. On this first map, the area of Brussels is the only one which runs across the linguistic border and covers the three institutional regions of the country. In the case of the average duration of communications, we obtain a group in the north and a group in the south of the country. These two groups concentrate within themselves almost 98% of telephone calls; only 2% of communications are from one group to the other. The group in the south includes the 19 municipalities of Brussels, all of the municipalities in the Walloon region as well as eight municipalities with facilities located in the Flemish region. In particular, all of the municipalities with facilities in the outskirts of Brussels (with the exception of Wemmel) are part of the group in the south of the country.

***Vincent Blondel** is an applied mathematics professor at the Université catholique de Louvain and is visiting professor at the Massachusetts Institute of Technology (Cambridge, USA) where he teaches linear programming. He is a member of the ICTEAM institute at UCL, the Naxys center at FUNDP and of the Laboratory for Information and Decision Systems at MIT. His research is focused on the analysis of complex systems and on various discrete mathematics issues, in particular the analysis of major systems/networks.*

***Gautier Krings** is a civil engineer in applied mathematics, and is working on a doctoral thesis in engineering sciences at UCL under the supervision of Vincent Blondel. He is a member of the ICTEAM institute at UCL. His work is focused on the extraction of information from major systems/networks, and in particular on their geographical distribution and evolution over time.*

***Isabelle Thomas** has a doctorate in geography and is a qualified secondary school teacher (AES). She is Research Director at the Fonds de la Recherche Scientifique and is a professor in the geography department at UCL. Her work is focused in particular on the location of human activities, with special emphasis on statistical and mapping tools as well as on modelling.*

Contacts :

Vincent Blondel, +1 617 253 2141 +1 617 956 2563
vincent.blondel@uclouvain.be, blondel@mit.edu
Gautier Krings, +32 10 47 8041,
gautier.krings@uclouvain.be
Isabelle Thomas, +32 10 472136,
isabelle.thomas@uclouvain.be

Christophe Mincke (senior editor), +32(0)473 21 02 65
mincke@fuw.ac.be



1. Introduction

Origin-destination matrices have fascinated regional science researchers for many decades and have been the object of abundant academic literature studying the impact of distance on the intensity of flow between origin and destination (see for example Carrothers, 1956; Taylor, 1983; Haynes and Fotheringham, 1988; Sen and Smith, 1995; Pumain and Saint-Julien, 2001; Roy, 2004 and Prager and Thisse, 2010). Telephone communications constitute an example of interrelations for which gravitational models are often used (see for example Rietveld and Janssen, 1990, Rossera, 1990 and Palm, 2002). The latter show the significance of language and distance in fixed telephone communications. Our work is in keeping with this, but concerns mobile telephony in Belgium and takes into account the structure of the communication network.

In Belgium, commutes to school or to work, migrations and telephone communications are examples of databases of spatial interrelations which have led to many spatial analyses, and even to proposals for the definition of borders. It is, however, often difficult to compare these borders, as their definitions vary according to the objective, the scale and the criteria used. Let us mention for example the urban areas proposed by Van Hecke *et al.*, 2007, the labour pools (de Wasseige *et al.*, 2000; Lockhart and Vanderhoff, 2009; Verhetsel *et al.*, 2009), the educational areas (Halleux *et al.*, 2007) and the market areas (Mérenne-Schoumacker, 2010). Other studies examine the “thickness” of the linguistic border (see for example Klaassen *et al.*, 1972 and Dujardin, 2001). In most cases, the definition of areas is based on the identification of a centre within a space and on the notion of accessibility to this centre.

The present article is aimed at proposing a spatial structure of Belgium based on the flow of mobile telephone traffic, with no geographical presupposition of distance or centrality. The municipalities are grouped here using a statistical method adapted to large network databases (Newman, 2006; Blondel *et al.*, 2008). The results obtained have been mapped and allow us to consider questions such as: do people telephone those who are close more than those who are far away? Are there border effects within Belgium and what are they? We shall also examine whether telephone habits divide the Belgian territory in an original manner or whether the groups ob-

tained are modelled on urban hierarchy, on job location or on administrative realities (provinces, districts, etc).

The rest of the paper is organised as follows: in Section 2 we briefly describe the data and the method used to identify the groups of municipalities, and then we describe and discuss the results obtained in Section 3 before concluding with a final section.

2. The data and the automatic detection method for groups of municipalities

The database used contains information related to mobile telephone communications of all of the customers of a provider. The database contains 2.6 million customers whose billing municipalities are known (and not their address), as well as the number and duration of calls made and received. In order to preserve their anonymity, each customer is identified by an anonymous code. The database contains more than 200 million communications over a period of six months (between 1 October 2006 and 31 March 2007).

For our analysis, we considered only the mobile communications between the provider's customers; in particular, communications to landlines, international communications and communications to other providers were not considered. SMS text messages, MMS and other 'data' communications were also left out in order to focus only on voice communications. Finally, only reciprocal communications (at least one call in either direction) of a significant importance (at least six calls in six months) were considered. Through these choices, we attempt to obtain a representation of an interpersonal network as much as possible. Unfortunately, it was not possible to distinguish the private calls from the professional calls, which would have enabled us to refine the results.

Each customer is identified (geocoded) according to his or her billing municipality. Given the nature of mobile telephony, it is useful to specify that we did not consider the location of customers at the time of the call. We did not conduct an analysis of the physical mobility of customers: only the billing municipality is used to geocode a customer. We therefore obtain a network in which the information concerning the link between municipality *A* and municipality *B* is provided by all of the communications made by the customers located in *A* to the customers located in *B*.

The *relative frequency* of the *A B* link is defined as the number of communications made by the *A* customers to the *B* customers, divided by the product of the number of customers located in *A* and of the number customers located in *B*. This division relativizes the actual figures and provides a measure that allows us to compare frequencies between pairs of municipalities even if the pairs of municipalities have different sizes and if the market shares of the provider in the municipalities are different.

In addition to the (relative) frequency, we also consider the *average duration* of communications as it informs us of another aspect of the nature of telephone relations. The average duration of the *A B* link is given by the average duration of communications made by the *A* customers to the *B* customers. As there are 589 mu-

municipalities in Belgium, we therefore have two 589×589 matrices: one with the frequency of calls and the other with the average duration of communications.

As an illustration, Figure 1 represents the mapping of the frequency of communications from the municipalities of Arlon (a) and Ostend (b) to all of the Belgian municipalities. We observe that the frequency of communications decreases significantly

with distance, thus confirming well-known properties of telephone communications, and even of most social interactions (see among others Verhetsel *et al.* 2009 for similar maps based on commutes using other methods). At a time when telephone communication rates within Belgium no longer depend on distance and when many economic activities appear to be less dependent on the cost of transport, people still telephone most often those who are geographically close and whom they probably see often.

Based on the two networks described above (*frequency, average duration*), in what follows we identify groups of municipalities in an automatic manner. Our goal is to see if there are groups of municipalities whose frequency of communications is higher or whose average duration of communications is longer. The automatic identification of groups in networks is a subject which has led to many mathematical and algorithmic works in recent years (see for example Fortunato, 2010 for a review). The most widely used measurement to quantify the quality of the division of a network into groups is the notion of 'modularity', introduced by Newman (2006). The *modularity* of a division into groups compares the density of links between the members of a group with that obtained in a random group with the same overall characteristics. A high level of modularity indicates that the density of links within groups is significantly higher than what may be expected in a random division with equivalent characteristics. This measurement does not fix a particular size for the groups, nor does it fix the number of groups.

The search for a division of a network which maximises modularity is a difficult

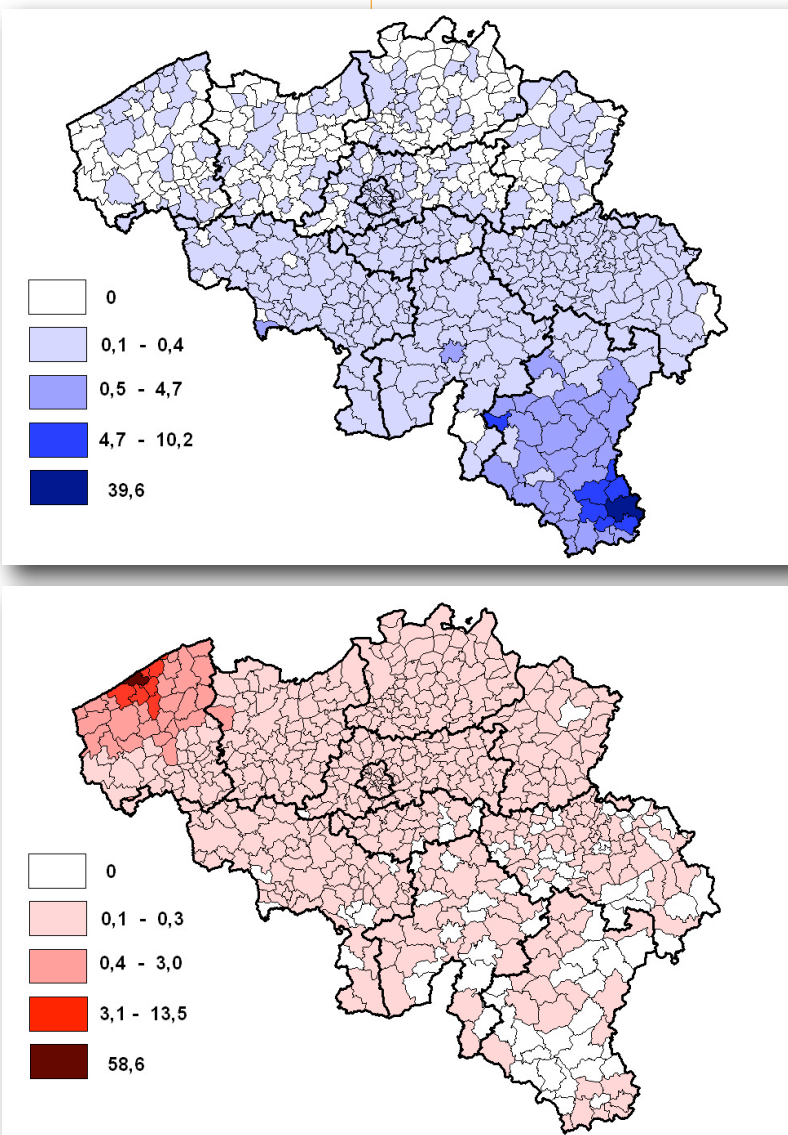


Figure 1: Frequency of communications from the municipalities of Arlon (blue) and Ostend (red).

problem for which several algorithmic methods have been proposed in recent years. A brief description of the method used in this article which is known as the '*Louvain method*' is given below. We invite readers who are not interested in this more technical description to go directly to Section 3, bearing in mind that:

the grouping of municipalities is done with no measurement of centrality or accessibility: only the frequency or the average duration of links between the 589 municipalities are used;

the number of groups and the composition of groups are endogenous: the method provides the optimal number of groups and their composition automatically, with no outside intervention;

the method is well defined mathematically: it attempts to optimise the modularity of the division obtained. Furthermore, the method has been used successfully in many contexts (see Lancichinetti, 2009), and is available in several softwares for the automatic processing of networks, such as the *igraph* library (<http://igraph.sourceforge.net/>) or NetworkX (<http://networkx.lanl.gov/>).

Let us now briefly describe the method; for a more detailed description, we refer the interested reader to Blondel *et al.* (2008) and Aynaud *et al.* (2010). The method uses a greedy algorithm to group the network nodes (in this case, the municipalities) in successive stages and builds a hierarchy of networks. At the lowest level of the hierarchy, all of the nodes first form distinct groups. Then, at each iteration, the method selects a node from a group and adds it to the group which maximises the contribution resulting from modularity if this contribution is positive. When any modification of the joining of a node reduces the modularity, a new network is built whose nodes are the groups from the network one level lower and the process is repeated in this new network.

Note that the results of the method depend on the order in which the nodes are covered. In order to verify the *robustness* of the method with respect to this order, we carried out 100 random permutations of the order of data and then compared the groups obtained. When the criterion used is the frequency of communications, we observe that, on average, 91% of municipalities remain in the same groups. Furthermore, these variations always correspond to modifications of bordering municipalities which go from one group to an adjacent group. When the criterion used is the average duration of communications, we do not observe any variation in the groups obtained. Furthermore, in the latter case, we verify that all transfers of a municipality from one group to another group reduces modularity. These observations appear to be additional proof of the stability and robustness of the two groups obtained in the case of the average duration of communications.

3. Results: division of the Belgian telephone territory

3.1 Division based on the frequency of calls

Figure 2 illustrates the groups obtained based on the *frequency* of telephone communications between municipalities. The colours are of no particular significance and are simply intended to facilitate the reading of the map.

Our main comments may be summarised in four points:

(1) Without having fixed the number of groups or their size, the optimal groups obtained are spatially *balanced*: 17 'telephone areas' composed of 15 to 66 municipalities appear 'naturally'. This result is different from the division in labour pools (47 pools defined by de Wasseige *et al.*, 2000) and, without being identical, resembles the urban hierarchy of Van Hecke *et al.* (2007). To this effect, we have indicated on the map in Figure 2 the regional cities and the major cities as defined in Van Hecke *et al.* (2007). Note that certain telephone areas encompass two cities (for example, the Belgian coast forms a telephone area in itself and groups the cities of Ostend and Bruges; other examples: Hasselt and Genk or Mechelen and Leuven), whilst other telephone areas do not correspond to a 'regional city' as defined by Van Hecke *et al.* (2007) (for example Aalst to the west of Brussels is a telephone area, whereas Aalst is not considered as a 'regional city'; the same is true for the province of Luxembourg).

(2) Surprisingly, the groups of municipalities are always made up of *adjacent municipalities*. As the grouping method does not impose constraints regarding proximity or contiguity of municipalities in groups, the results could have revealed groups composed of separate parts, but this is not the case for the groups obtained.

(3) The linguistic border is followed by the limits of the 'telephone areas', with the exception of the area of Brussels (in red on the map) and the municipalities with facilities Espierre-Helchin, Comines-Warnton, Herstappe and Fourons. Language therefore seems to be a strong barrier in terms of telephone communications: this confirms the former results of Klaassen *et al.* (1972), Rossera (1990) and Rietveld and Janssen (1990). However, it should be noted that the barrier around the German-speaking region is less clearly marked.

(4) The biggest area obtained (66 municipalities) corresponds – not surprisingly – to the biggest city: *Brussels*. Figure 3 presents a zoom-in of Figure 2 centred on Brussels.

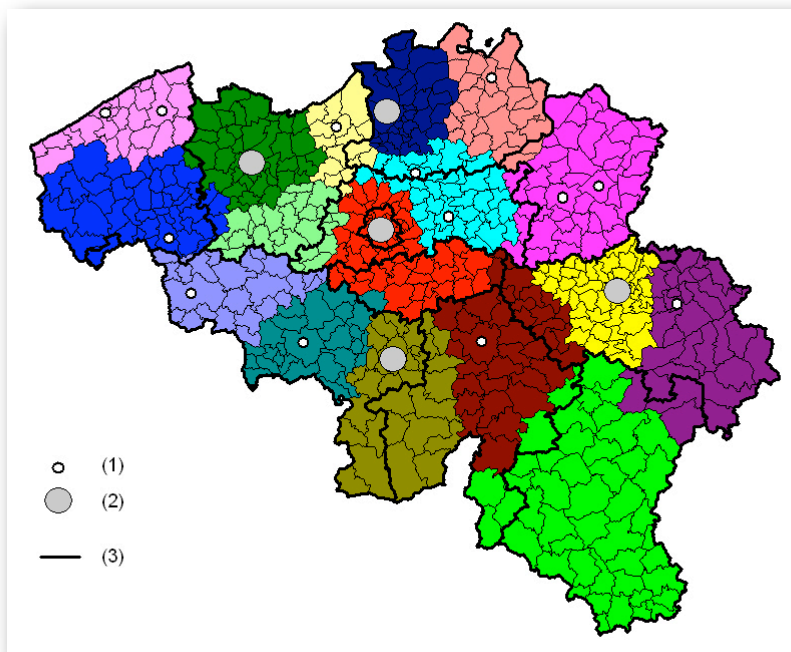


Figure 2: 'Telephone areas' defined based on the frequency of communications between municipalities. We also indicate (1) = regional city (2) major city (definitions from Van Hecke *et al.*, 2007) and (3): provincial borders.

The borders of the urban area of Brussels have been the object of many political and academic debates for many years (see for example the review of the academic question by Dujardin *et al.*, 2007). The borders obtained for the urban area often depend on the criteria used, and there is no consensus in terms of methodology. One thing is certain, however: Brussels, like many other cities in the world, extends beyond its administrative borders, and the Brussels 'telephone area' illustrates this reality once again. The area covers not only the 19 municipalities of the Brussels-Capital Region but also the directly adjacent municipalities regardless of the geographical direction, with however a stronger spatial extension towards Walloon Brabant, which it encompasses entirely with the exception of Nivelles and two municipalities to the far east (Hélécine and Orp-Jauche). This extension towards the south is associated with the history of the suburbanisation of Brussels coupled with a linguistic and socioeconomic reality (see among others Thisse and Thomas, 2010). The Brussels mobile telephone area encompasses municipalities such as Halle,

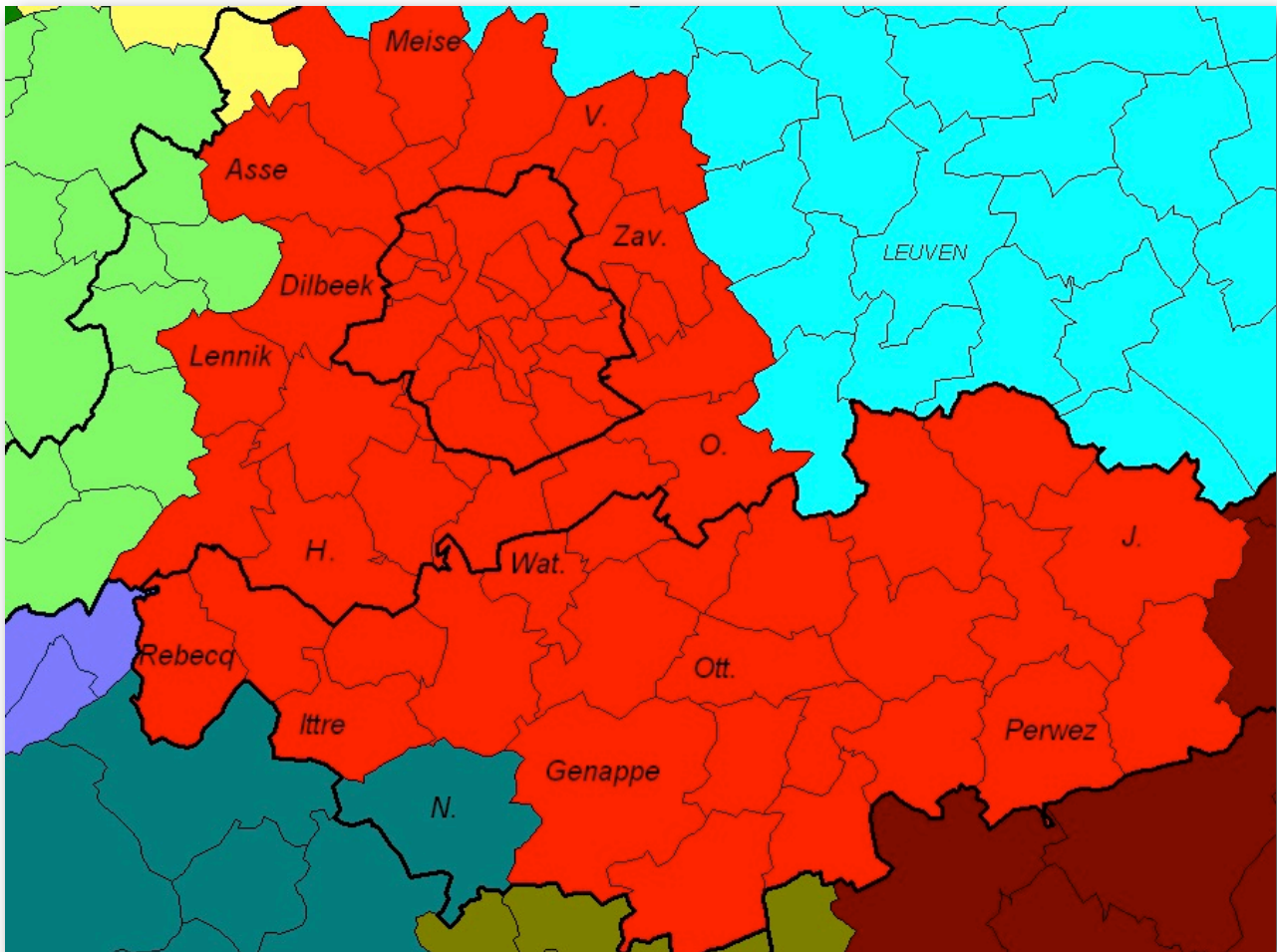


Figure 3: 'Mobile telephone area' of Brussels (zoom-in of Figure 2). Complete list of municipalities in Annexe 1.

Vilvoorde, Zaventem, Tervuren, Braine-l'Alleud, Ottignies-Louvain-la-Neuve, Wavre, Perwez and Jodoigne. However, Leuven is not included and is part of another telephone area with Mechelen (see Figure 2). The Brussels telephone area resembles its urban area: it covers a much bigger area than the 19 municipalities of the Brussels-Capital Region, all around the capital with a stronger spatial extension towards the south.

3.2 Division based on the average duration of communications

The municipalities are grouped here using the same method, according to the *average duration* of communications. The results are illustrated in Figures 4 (national scale) and 5 (a zoom-in on Brussels) and lead to two main commentaries:

(1) the method leads naturally to the constitution of *two groups*: one to the north and the other to the south of the country (Figure 4). Among the more than 200 million communications analysed, only 1.05% are from the group in the north to the group in the south, and 1.04% are from the group in the south to the group in the north. In other words, almost 98% of telephone communications take place between customers within the same group. Let us note that the municipalities in the

German-speaking community do not form a separate group, but are part of the group in the south of the country.

(2) Figure 4 shows that the north-south division follows the *linguistic border* with a few exceptions. Not surprisingly, these exceptions are all municipalities with facilities. With the exception of Wemmel, the municipalities with facilities in the outskirts of Brussels (Drogenbos, Kraainem, Linkebeek, Rhode-Saint-Genèse, Wezembeek-Oppeem) are all grouped with the municipalities in the south of the country (see Figure 5 for a zoom-in). Three other municipalities

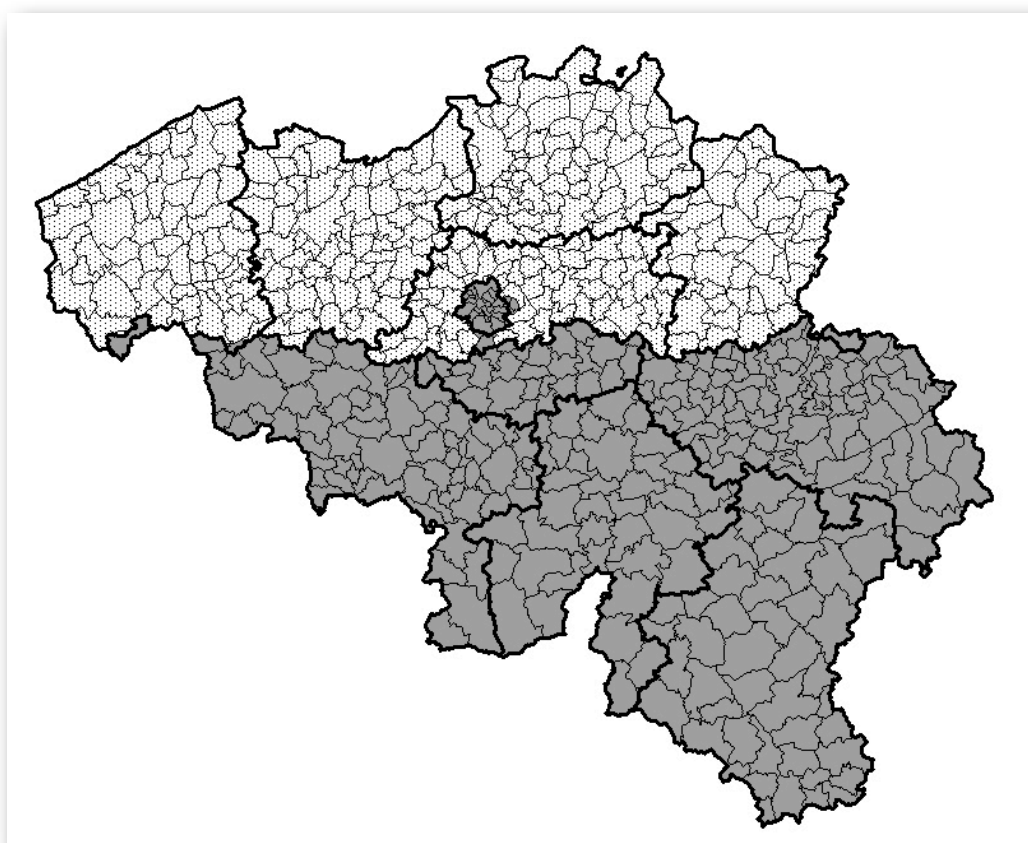
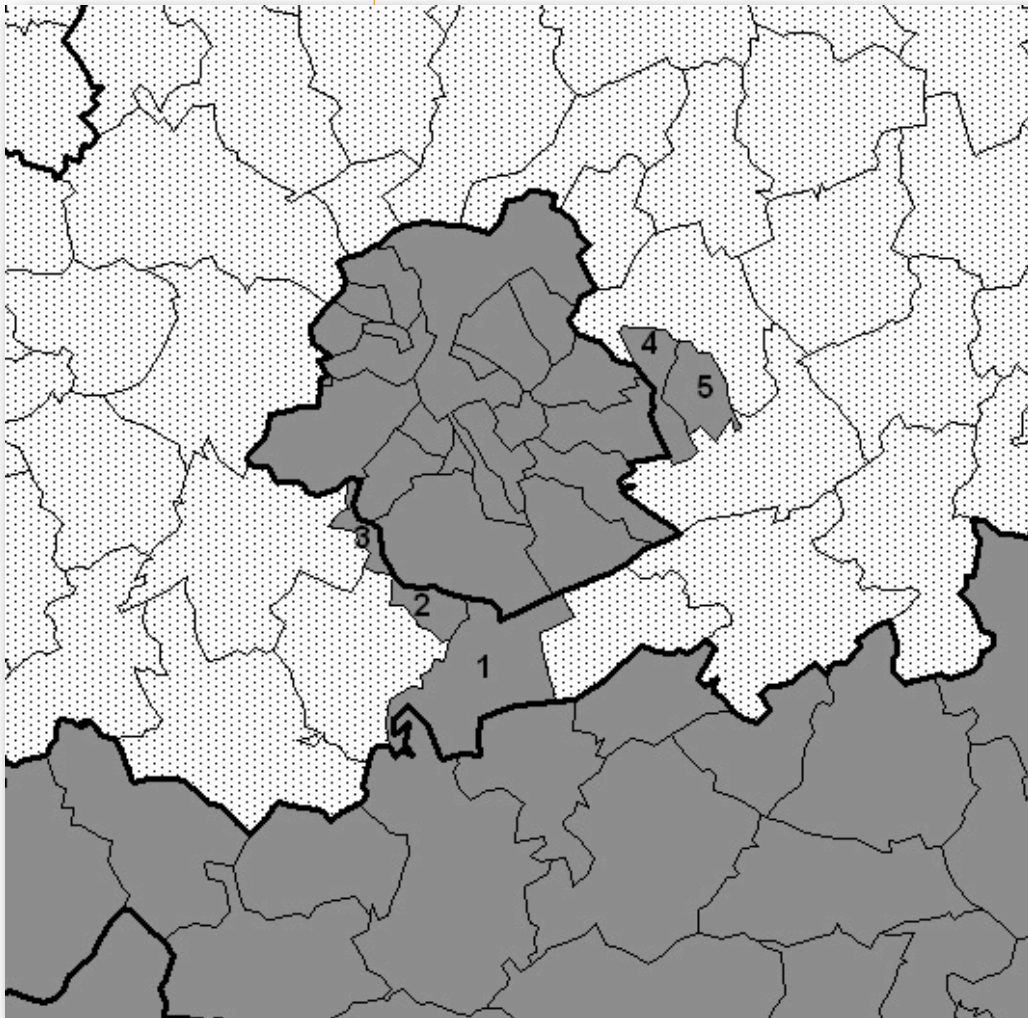


Figure 4: 'Mobile telephone areas' defined based on the average duration of communications.



with facilities in the Flemish region located outside the influence of Brussels are also part of the group in the south of the country: Espierres-Helchin, Fourons and Herstappe. Finally, all of the other municipalities with facilities are associated with the group of their linguistic region: Comines-Warneton, Enghien, Flobecq and Mouscron (group in the south), and Messines, Biévènes and Renaix (group in the north).

Figure 5: Grouping of municipalities based on the average duration of communications. Zoom-in on the Brussels-Capital Region, where 1 = Rhode-Saint-Genèse, 2 = Linkebeek, 3 = Drogenbos, 4 = Kraainem and 5 = Wezembeek-Oppeem.

4. Conclusion

Based on an analysis of more than 200 million telephone communications between more than two million mobile telephone users, the 589 Belgian municipalities are divided into groups of municipalities, or 'telephone areas'. Two analyses were conducted: one based on the relative frequency of communications and the other based on their average duration.

The groups were obtained using a network analysis method which is widely used in the research community on large networks. The method does not consider the geographical distance between the municipalities or their belonging to a community or a region. Nor does it consider the language used for invoicing; only the frequency and the average duration of communications are used. The number of groups and the composition of these groups are provided by the method itself, with no outside intervention. In both cases, we see that all of the groups are composed of adjacent municipalities despite the absence of constraints regarding contiguity.

In the case of the relative frequency of calls, 17 groups appear and resemble the division into urban areas whilst respecting the linguistic border on the whole.

In the case of the average duration of calls, only two major groups appear: one in the north and the other in the south. These two groups concentrate within themselves almost 98% of telephone calls; only 2% of the calls connect municipalities from different groups. The communications therefore appear to be strongly influenced by geographical and psychological proximity (including linguistic and cultural dimensions), thus confirming the spatial studies by Klaassen *et al.* 1972, Rossera, 1990 and Rietveld and Janssen, 1990, and even the general discussions regarding the evolution of communications by Bakis (1995).

In this article we present two maps of telephone areas. These maps are only intended to help us consider our ways of perceiving the Belgian territory, and in particular the Brussels territory. They also shed light on our social and cultural differences which lead to other uses of the territory than our socioeconomic differences (see for example Thisse and Thomas, 2010). The differences between these maps remind us that any division of space depends greatly on the criteria used; the average duration and relative frequency of communications lead to different divisions of space. The borders highlighted in this article reveal important elements of our 'living areas' and social practices, which may be considered in the administrative and geopolitical organisation of Belgium.

Bibliography

- AYNAUD T., BLONDEL V., GUILLAUME J.-L., LAMBIOTTE R. (2010), Optimisation locale multi-niveaux de la modularité, in *Partitionnement de graphe : optimisation et applications*, Traité IC2, Hermes-Lavoisier.
- BAKIS H. (1995), Communication et changement global: un défi politique et culturel *International Political Science Review / Revue Internationale de Science Politique*, 16:3, pp. 225-235
- BLONDEL V., GUILLAUME J.-L., LAMBIOTTE R., LEFEBVRE E. (2008), Fast Unfolding of Communities in Large Networks, *Journal of Statistical Mechanics: Theory and Experiment*, P10008, 12 pp.
- CARROTHERS G. (1956), An Historical Review of the Gravity and Potential Concepts of Human Interaction, *Journal of The American Institute of Planners*, 22, pp. 94-102.
- DE WASSEIGE Y., LAFFUT M., RUYTERS C., SCHLEIPER P. (2000), Bassins d'emploi et régions fonctionnelles. Méthodologie et définition des bassins d'emploi belges, *Service des Etudes et de la Statistiques*, Discussion Paper 2005, 24 pp.
- DUJARDIN C. (2001), Effet de frontière et interaction spatiale. Les migrations alternantes et la frontière linguistique en Belgique, *L'Espace Géographique*, 30, 4, pp. 307-320.
- DUJARDIN C., Thomas I., Tulkens H. (2007), Quelles frontières pour Bruxelles ? Une mise à jour, *Reflets et Perspectives de la Vie Economique*, XLVI : 2-3, pp. 155-176.
- FORTUNATO S. (2010), Community Detection in Graphs, *Physics Reports* 486, pp. 75-174.
- HALLEUX J.-M., RIXHON G., KAMBOTTE J.-M., MÉRENNE-SCHOUMAKER B. (2007) Les navettes scolaires en Belgique: situation en 2001 et évolution 1991-2001. Brussels, SPF Economie, P.M.E., Classes moyennes et Energie, Working Paper.
- HAYNES a K. and FOTHERINGHAM A.S. (1984), *Gravity and Spatial Interaction Models*, Beverly Hills, Sage, 259 pp.
- KLAASSEN L., WAGENAAR S., Van Der WEG A. (1972), Measuring Psychological Distance between the Flemings and the Walloons, *Journal in Regional Science*, 29:1, pp. 45-62.
- KRINGS G., CALABRESE F., RATTI C., BLONDEL V. (2009), Urban Gravity: a model for inter-city telecommunication flows, *Journal of Statistical Mechanics: Theory and Experiment*, L07003 (8 pp.), 2009.
- LAMBIOTTE R., BLONDEL V., De KERCHOVE C., HUENS E., PRIEUR C., SMOREDA Z., VAN DOOREN P. (2008), Geographical Dispersal of Mobile Communication Networks, *Physica A: Statistical Mechanics and its Applications*, 387, pp. 5317-5325.

- LANCICHINETTI A. and FORTUNATO S. (2009), Community Detection Algorithms: A comparative analysis, *Physical Review E* 80, 056117 .
- LOCKHART P. and VANDERMOTTEN C. (2009), *Atlas des dynamiques territoriales. Fiche : Les bassins d'emploi en 2001*. Atlas en ligne, CPDT
<http://www.cpd.be/telechargement/atlas-dt/fiches/Bel-Bassins-emploi.pdf>
- MÉRENNE B. (2010), *Aménagement du territoire et bassins de vie*, Conference at the Maison de l'Urbanisme du Pays de Liège, 9 March 2010
www.maisondelurbanite.org/fichiers_annexes/MerenneB.ppt
- NEWMAN M. (2006), Modularity and Community Structure in Networks, *Proceedings of the National Academy of Sciences of the USA*, 103:23, pp. 8577–8582.
- PALM R. (2002), International Telephone Calls : Global and Regional Patterns, *Urban Geography*, 23 :8, pp. 750-770.
- PRAGER, J.-C. and J.-F. THISSE (2010), *Economie géographique et développement économique*. La Découverte, Collection Repères.
- PUMAIN D. and SAINT-JULIEN T. (2001), *Les interactions spatiales*, Paris, Armand Colin, Cursus-géographie, 191 pp.
- RIETVELD P. and JANSSEN L. (1990), Telephone Calls and Communication Barriers: The Case of the Netherlands, *The Annals of Regional Science*, 24(4), pp. 307-18.
- ROSSERA F. (1990), Discontinuities and Barriers in Communications. The Case of Swiss Communities of Different Language, *Annals of Regional Sciences*, 24, pp. 319-336
- ROY J. (2004) *Spatial Interaction Modelling: A Regional Science Context*, Berlin, Springer, 239 pp.
- SEN A. and SMITH T. (1995), *Gravity Models of Spatial Interaction Behavior*. New York, Springer.
- TAYLOR P. (1983) Distance Decay in Spatial Interactions, *Catmog (Concepts and Techniques in Modern Geography)* 2, 35 pp.
- THISSE J.-F. and THOMAS I. (2010), Bruxelles au sein de l'économie belge : un bilan. *Regards Economiques*, 80 (June 2010), 18 pp. _
- VAN HECKE E., HALLEUX J.-M, DECROLY J.-M., MÉRENNE-SCHOUMACKER B. (2007), *Noyaux d'habitat et régions urbaines dans une Belgique urbanisée*, Brussels, Monographies Enquête Socio-économique 2001, n°9, Brussels, SPF Economie en Politique Scientifique Fédérale. 201 pp.
- VERHETSEL A., VAN HECKE E., THOMAS I., BEELEN M., HALLEUX J., LAMBOTTE J., RIXHON G. & B. MÉRENNE-SCHOUMACKER (2009), *Le mouvement pendulaire en Belgique. Les déplacements domicile-lieu de travail. Les déplacements domicile-école* in Monographies Enquête Socio-économique 2001, n°10, Brussels, SPF Economie en Politique Scientifique Fédérale. 217 pp.

Annex 1: List of municipalities which belong to the 'mobile telephone area' of Brussels

Anderlecht	Linkebeek
Asse	Machelen
Auderghem	Meise
Beauvechain	Merchtem
Beersel	Molenbeek-St-Jean
Berchem-Ste-Agathe	Mont-St-Guibert
Braine-L'Alleud	Ottignies-LLN
Braine-Le-Château	Overijse
Bruxelles	Pepingen
Chastre	Perwez
Chaumont-Gistoux	Ramillies
Court-Saint-Etienne	Rebecq
Dilbeek	Rixensart
Drogenbos	Saint-Gilles
Etterbeek	Saint-Josse
Evere	Schaerbeek
Forest	Sint-Genesius-Rode
Ganshoren	Sint-Pieters-Leeuw
Genappe	Steenokkerzeel
Grez-Doiceau	Tervuren
Grimbergen	Tubize
Halle	Uccle
Hoeilaart	Villers-La-Ville
Incourt	Vilvoorde
Ittre	Walhain
Ixelles	Waterloo
Jette	Watermael-Boitsfort
Jodoigne	Wavre
Koekelberg	Wemmel
Kraainem	Wezembeek-Oppem
La Hulpe	Woluwe-St-Lambert
Lasne	Woluwe-St-Pierre
Lennik	Zaventem