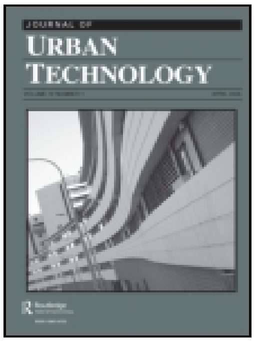

Routledge
Taylor & Francis Group

Journal of Urban Technology

ISSN: 1063-0732 (Print) 1466-1853 (Online) Journal homepage: http://www.tandfonline.com/loi/cjut20

Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones

Rein Ahas , Siiri Silm , Olle Järv , Erki Saluveer & Margus Tiru

To cite this article: Rein Ahas , Siiri Silm , Olle Järv , Erki Saluveer & Margus Tiru (2010) Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones, Journal of Urban Technology, 17:1, 3-27, DOI: 10.1080/10630731003597306

To link to this article: http://dx.doi.org/10.1080/10630731003597306

Published online: 30 Mar 2010.

Journal of Urban Technology, Vol. 17, No. 1, April 2010, 3–27

Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones

Rein Ahas, Siiri Silm, Olle Järv, Erki Saluveer, and Margus Tiru

ABSTRACT *The article introduces a model for the location of meaningful places for mobile telephone users, such as home and work anchor points, using passive mobile positioning data. Passive mobile positioning data is secondary data concerning the location of call activities or handovers in network cells that is automatically stored in the memory of service providers. This data source offers good potential for the monitoring of the geography and mobility of the population, since mobile phones are widespread, and similar standardized data can be used around the globe. We developed the model and tested it with 12 months' data collected by EMT, Estonia's largest mobile service provider, covering more than 0.5 million anonymous respondents. Modeling results were compared with population register data; this revealed that the developed model described the geography of the population relatively well, and can hence be used in geographical and urban studies. This approach also has potential for the development of location-based services such as targeting services or geographical infrastructure.*

Introduction

Population mobility is growing rapidly in the contemporary world. According to the mobility paradigm (Sheller and Urry, 2006), movement has become a phenomenon in its own right in the twenty-first century; employment is more and more mobile, tourism has become a lifestyle, and it is common to own multiple residences and jobs. Rising everyday mobility has been combined with the rapid growth in the use of mobile phones and other mobile communication devices.

In order to assess and analyze the location of individuals and populations in this mobile world, new methods and approaches are needed. Traditional census and population registers are solid sources for long-term processes. For the short term and everyday mobility, more flexible methods such as various registers and indirect databases (Raymer *et al.*, 2007), satellite-based methods (Chen *et al.*, 2006), or modern sensing technologies (Kwan, 2000; Eagle and Pentland, 2005; Shoval, 2007) are needed.

One of the proposed methods for developing such a monitoring tool is mobile positioning or mobile telephone tracking (Mountain and Raper, 2001; Spinney, 2003), also called the social positioning method (Ahas and Mark, 2005) or cellular census (Reades *et al.*, 2007). Mobile positioning is often considered to be a novel and exciting source of information for investigating the spatial dynamics of human society, while at the same time the number of published studies is small

1063-0732 Print/1466-1853 Online.

DOI: 10.1080/10630731003597306

because of problems concerning limited access to such data and privacy issues. A number of interesting works regarding urban studies (Ratti, 2005; Reades *et al.*, 2007; Ahas *et al.*, 2007a), tourism studies (Ahas *et al.*, 2008a), transportation studies (Asakaura and Hato, 2004; Herman, 2006), and GIS (Nurmi and Koolwaaij, 2006; Gartner, 2004) however, have been published in recent years.

There are many methods and approaches that can be used to locate mobile telephones. Technical solutions vary from handset-based systems with special telephone software to satellite navigation and peer-to-peer positioning tools using Bluetooth. The most positive aspect of mobile positioning is that nowadays mobile phones are becoming pervasive in developed and developing countries alike. In order to conduct population and mobility studies based on positioning data, a great deal of preparatory work is necessary, because data needs to be processed and assessed, and the appropriate methods must be developed.

The objective of this paper is to develop a model for locating places that are meaningful to mobile phone users. This is done by using passive mobile positioning data. Meaningful places or meaningful locations (Nurmi and Koolwaij, 2006) are defined as regularly visited places that have meaning for individuals. Technically, they are similar to personal anchor points; home and work anchors are the most common among them. With the rising popularity of the mobile lifestyle and mobile communication, there is an increasing number of regularly visited places, and mobile phones can be used to detect those places. Passive mobile positioning data are a secondary source of geographical data that are automatically stored in the memory files and logs of mobile operators (Ahas *et al.*, 2008b). Here we introduce our model for determining the geographical location of meaningful places using an Estonian example. The modeled results of the research were compared with the data of the Estonian Population Register.

This quantitative-based model is an important step in developing alternative tools to monitor the everyday mobility of the population using various tracking data and to develop location based services (LBS) for mobile phone users. Mobile positioning data are becoming more and more common in geographical studies, and LBS are finding new markets and products around the globe. Therefore, there is growing interest in such "mobile geography" as a current model for anchors, not only from a geographical perspective but also from the information technology (IT) sector for the personalization of mobile services. The issues of privacy and surveillance are important aspects of mobile positioning. The focus of this paper does not allow us to discuss this in depth, but we cover some topics related to passive positioning in our discussion of the data.

Theoretical Framework

Passive Mobile Positioning

Mobile positioning means tracing the location coordinates of mobile phones. There are different frameworks for positioning, for instance handset-based, network-based or GPS-based. In order to locate phones, radio waves are used, and positioning is done using various methods, such as cell ID, triangulation with direction angle, and/or distance from an antenna. Different positioning methods are used because of the different network standards (GSM, CDMA, 3G), and there are different purposes for positioning (Zhao, 2002). In addition, the use of mobile positioning data in geographical studies has different approaches and algorithms (Roos *et al.*, 2002).

Mobile positioning can be divided into passive and active positioning. Active mobile positioning is used for mobile tracking in which the location of the mobile phone is determined (asked) with a special query using a radio wave (Ahas *et al.*, 2007a). Passive mobile positioning is data which is automatically stored in memory or log files (billing memory; hand-over between network cells, Home Location Register, etc.) of mobile operators (Ahas *et al.*, 2008a). The easiest method for passive mobile positioning is "a billing log" that is recorded for called activities. Any active use of a mobile phone (call and SMS messages in and out, GPRS, etc.) is deemed to be call activity. Passive mobile positioning data is normally collected to the precision of network cells. Every cell has a certain geographical coverage area and unique identity code, and, therefore, this method is called Cell ID. Mobile operators can aggregate anonymous geographical data from log files, such as location points or movement vectors, and researchers can use these in surveys for scientific purposes. Issues of privacy and surveillance are very important aspects of any mobile positioning data. Other sources than call activity are also used for passive positioning, such as Erlang of antennae (Reades *et al.*, 2007).

When using passive mobile positioning data, as we do in the current paper, it is important to explain some basic definitions of the topic. The cellular network is based on a set of base stations, which usually have one tower and several directed antennae. The radio coverage of a single antenna forms a network cell; several antennae form a cellular network. Every network cell in a mobile phone network has a unique ID and geographical coordinates, and the location of a phone in the cell can be easily determined for every call activity.

The size of a network cell and all cellular networks is not fixed; the phone normally switches to the closest antenna or the one with the strongest radio coverage or best visibility. If the network is crowded or visibility is disturbed, the phones can be switched not to the nearest station but to any other in the neighborhood. The maximum distance from a handset to an antenna in the GSM network is less than 35 km. There are amplified antennae used in GSM networks in less inhabited or coastal areas that cover greater distances.

Conceptual Framework for Using Mobile Positioning Data

Our objective is to develop a model that recognizes meaningful places visited by single telephones (persons). These places can also be referred to as regularly visited places of personal anchor points such as home and work, etc. It has been noted, however, that with the rising mobility of individuals, the dominance of home and work anchors is decreasing, and people also spend their time in other important locations. Therefore, the term "meaningful places," which is very relevant to mobile positioning-based data, is often used. To use mobile positioning data to model locations of regularly visited places, we must conceptualize the interactions between mobile phones and space: Is the location of a mobile call something special, or it is just another location of another activity? Several researchers have been discussing the location of mobile phones and mobile calls. Most of them have pointed out that a wireless phone is a special place with its own spatial features. Finnish researcher Timo Kopomaa called a mobile phone, "an instrument for maintaining contact: the mobile phone can be viewed as a place" (Kopomaa, 2004, p. 268). He also labeled the mobile phone as a "third place" meaning the third important location after home and work. The third place was defined by Oldenburg (1989) as a place of social interaction with communication networks. The locations of home and work places in mobile networks may be different from their actual geophysical locations. But they are recognizable as regularly visited places in the third place, the mobile telecommunications network.

The conceptualizing of mobile positioning data begins with the need to remember that mobile calls are social events; every "call out" or "call in" involves another person or call partner with a telephone. Therefore, we must remember that each call also depends on the availability of a communication partner. This is why, even if some people wake up early or drive through the night they do not call as often, as their potential call partners are asleep. Indeed, the calls of some are calls with other time zones.

There are many interesting methods for investigating the geographical and social aspects of networks of personal communication. In this study we focus on the location of persons calling someone by phone. In connection with mobile telephones, we must also keep in mind that people either do or do not use phones in certain places and for certain activities. The culture of phone use has not been systematically studied in Estonia, but many countries have learned about the use of phones. For example, automobiles or public transportation are among the most popular places to make calls, and for the same reason, mobile phones are a "modern" source of traffic accidents. As we study the location of calls, it is important to know that people may call during or after certain events (the New Year, after a football match has been won) or at times when emotions are stronger than normal, and those times and locations may not leave a proportional trace in the data.

Because of privacy issues, the massive mobile positioning datasets used in several countries (Ahas *et al.*, 2007a; Reades *et al.*, 2007) do not have many identifications or personal features of the owners of the phones studied. Normally those databases have huge quantities of locations or "dots". Those dots can serve as sources for the modeling of meaningful places for people.

Approaches for Determining Meaningful Places

There are many conceptual approaches to the description and analysis of people's everyday movements. The Swedish geographer T. Hägerstrand (1970) conceptualized geographical movement by including a time dimension in the data model. He defined paths as movement corridors and stations as visited places which are typical for most of humans. Space-time is a dimension we can also use in the conceptualization of the given mobile positioning data, since both chronological and chorological dimensions are exact and vast in the database. The lack of additional personal information and identifications forces us to disregard models and approaches that rely on the individual potentials of travel as mental maps, action space, and perceptual space using this positioning data (Schönfelder and Axhausen, 2004).

The anonymous passive mobile positioning database we use is more applicable with interpretations based on the actual activity spaces concept that describes the structures of realized locational choices for single travelers and is also referred to as micro-geographical activity space (Schönfelder and Axhausen, 2004). This approach has been used by different schools for many decades (Dürckheim, 1932). Activity spaces represent the distribution of places visited and the space that contains those places frequented over a period of time. Activity spaces are geometric indicators of observed or realized daily travel patterns (Axhausen *et al.*, 2002; Schönfelder and Axhausen, 2004; Dijst, 1999). The study of the geometry of activity spaces may be a possible solution for the handling of our passive mobile positioning data. Some examples of possibilities are the theoretical concepts of confidence ellipses (Schönfelder and Axhausen, 2004; Schwarze and Schönfelder, 2001). The confidence ellipse method using the Visar program was successfully used in a study of gender differences in mobile positioning data (Silm *et al.*, 2008). Other GIS-based analytical tools have also been proposed, such as kernel densities or geometrical analyses of locations. The investigation of the geometry of activity spaces is insufficient if there are not any additional data available about individuals and their activities.

This is why the further use of passive positioning data depends on the implementation of the anchor points concept. Anchor points are locations where people regularly stay (Golledge and Stimson, 1997). Besides anchor points, the terms bases (Mitchell and Rapkin, 1954; Kutter, 1973; Vidakovic cit Dijst, 1999) or core stops (Schank and Abelson, 1977) have also been used. Anchor points can be divided into two meanings. "Common anchors" are significant places in the environment that are commonly recognized and used as key components of cognitive maps (Dijst, 1999). "Personalized anchors" are related to a person's activities (e.g., a specific work place or home-base) (Golledge, 1990). The meaning of anchor points is described from different perspectives by different authors. Our interest here is in the GIS-based analysis of huge datasets. For this purpose, there is one relevant concept—"a personal network of usual places," consisting of activity places and the routes between them (Flamm, 2004; Flamm and Kaufmann, 2006). Flamm and Kaufmann determine a certain number of daily life centers among the activity places, i.e., places where individuals usually spend a considerable amount of time and which they consider important in the conduct of their everyday lives; these are typically the home and the workplace. It can be assumed that these daily life centers represent "territorial anchor points" of the personal activity space, and that they are most closely interconnected to other activity places by travel routes. Near daily life centers are clusters of activity places such as grocery stores for daily shopping, automatic teller machines, restaurants, etc. One important issue was to define the parameters used to identify centers of daily life and to differentiate usual activity places from others (Flamm and Kaufmann, 2006). Flamm and Kaufmann used the Moby drive dataset and determined anchor points on the basis of the following criteria: a) Centers of daily life were defined as places where an individual spent at least 67 hours over the entire six-week survey period. This value was chosen in order to include all places where individuals were employed for 20 percent of the time. Usual places were defined as activity places that were visited at least three times during the survey period, i.e. places with at least a bimonthly visit frequency. This criterion appeared to be most appropriate, since the survey period of six weeks did not allow detection of a monthly visit frequency (Flamm and Kaufmann, 2006).

From this theoretical and methodological basis, we began the compilation of our model or framework for the determination of meaningful places (same as personal anchor points or regularly visited places here) using passive mobile positioning data. As mentioned above, passive mobile positioning data have their peculiarities as there are a huge number of location points over a relatively long time period, and very little information about respondents' social features. Compatible methods have been developed for such data sources as the number of GIS-based analyses of mobile trajectories (Mountain, 2005; Zhou *et al.*, 2007) and GPS-based experiments in travel behavior studies (Wolf *et al.*, 2001). Computer scientists have also performed interesting studies to determine meaningful places from mobile positioning data and to develop algorithms for location-based services with different methods such as clustering movements and places (Kang *et al.*, 2004). It must be mentioned that the clustering method developed by Nurmi and Koolwaaij (2006) and Nurmi and Bhattacharya (2008) is another approach for determining important locations and movement types in cellular networks.

Data

Passive Positioning Data

In the current study, we use data from the databases of the largest Estonian mobile operator—EMT. EMT covers nearly 99.9 percent of the total land area of Estonia, which measures 46,000 km^2. Eurobarameter estimates that 87 percent of the Estonian population of 1.38 million have mobile phones (Eurobarometer 2007).

The database we use in the current study is passive positioning data—the locations of all "calls out" (calls initiated by the respondent, not received calls) with the precision of Cell IDs made in the Estonian largest network, EMT, during a period of one year. EMT uses network hardware and positioning middleware from Ericsson Ltd (Ericsson, 2008). The system uses Ericsson's GSM, GPRS/EDGE and WCDMA/3G network technologies and latest mobile positioning platform, MPS 9.0. The entries include the following parameters for every outgoing call of a phone registered in the network: a) the exact time of the call activity; b) the random ID number for the phone (not related to phone or SIM card number); c) the cell ID with the geographical coordinates of the antenna. Due to privacy issues, we do not possess any personal information about the respondents, but only a randomly assigned ID for every phone. The random ID generated by the operator enables us to identify the calls made by one person during the study period. There are hundreds of locations of calls for every ID, and we begin our analysis with this data. An example of the database is: *Time September 8. 2007. 22:03:11; ID 64353; Location E27-44-39.00 N59-25-49.00.*

Data are gathered by the company Positium LBS, which has a contract with the two major mobile operators in Estonia regarding the use of LBS data (Positium 2008). The current model was developed through a cooperative arrangement of the Positium LBS (authors E. Saluveer and O. Järv–) and the chair of human geography at the University of Tartu. The modeled data we present herein were initially used to analyze traffic on the Tallinn-Tartu Highway (Järv *et al.*, 2007) and to analyze urban sprawl (Silm *et al.*, 2007; Silm and Järv, 2007; Ahas *et al.*, 2007b). The Positium LBS database is powered by the open-source PostgreSQL database manager. As a monthly average, the database contains the calls of 592,553 people for every 12 months from November 1, 2006 to October 31, 2007. The average number of calls per month was 65.3 million. There are a total of

783 million location points in the database for the researched period. The distribution according to the ID and the number of calls is provided in Figure 1. There is an average of 100-120 calls per ID per month.

Privacy Issues

The use of positioning data brings up the issue of privacy and surveillance. This is the major concern of phone holders, operators, researchers, and the general public. Therefore, the research team together with the mobile operator and the Estonian Data Protection Inspectorate checked carefully the accordance of data use with Estonian legislation and EU directives (European Parliament, 2002). Our study and discussions concluded that the personal privacy of respondents is protected. There is no personal information connected with movement vectors, and the generalization level of the analysis does not allow the identification of persons on geographical or temporal grounds. It is not possible to extract individual movements from the data. Nevertheless, there is concern about issues of privacy and ethics, as any use of mobile positioning data is very sensitive in this respect. Our approach in Estonia was to make all research and results transparent, and we published our plans and results step by step in newspapers, on TV, and also on a website during the eight years of our studies in this field. Thus, everyone could see how "their" data were used and what the benefit of the research was. Nevertheless, we received many bad comments as feedback, but the mobile operator was happy with the transparency and the public comments, and we got the "green light" for the study.

In the current research, the data of the Population Register from January 1, 2007 at the municipality level were used as comparative material alongside the passive mobile positioning data.

Geographical Distribution of the Mobile Network

The mobile network is usually distributed unevenly over a territory, following population density patterns and transportation infrastructure. As a result, the outcome in a geographical sense is fairly objective, since positioning accuracy is

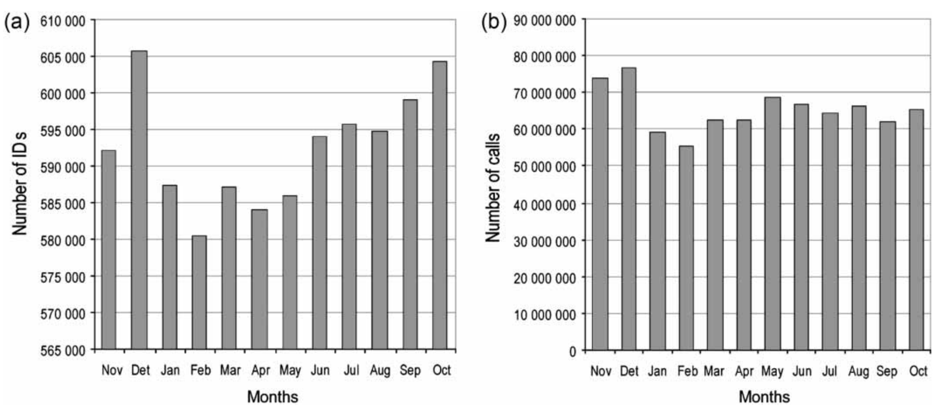

Figure 1. The distribution of the number of IDs and Calls in the passive positioning database used per month

smaller in less populated areas than in densely populated areas, and the overall error is smaller. Figure 2 represents the EMT mobile network in Estonia as of January 1, 2007. The mobile network is in a constant state of flux, with new base stations and antennae developed by all progressive operators.

In the present survey, local municipalities have been used as the subjects of research, and the data from mobile antennae have been generalized to match the level of local governmental units. In reality, mobile networks do not follow the borders of local governmental units. As a result, the data from mobile antennas are restricted to the local municipality on whose territory the mobile antenna are located. As a result, there are 12 local municipalities in Estonia (mostly in rural areas) about which there are no data, because there are no mobile antennas situated in those areas. For these units, a mean value has been calculated, based on the average data from other neighboring governmental units.

Description of Model

General Points

The passive positioning database is composed of the location and time of outgoing calls and of the random ID assigned to the respondent, which is not connected to the phone number but remains constant for every phone in the network. Based on random ID, the person's calls can be linked to each other through the study period. Personal anchor points, presumably home, work and other places, can be found as the locations the respondent visits regularly (hereafter, anchor points). Anchor points are important variables in describing humans' behavior

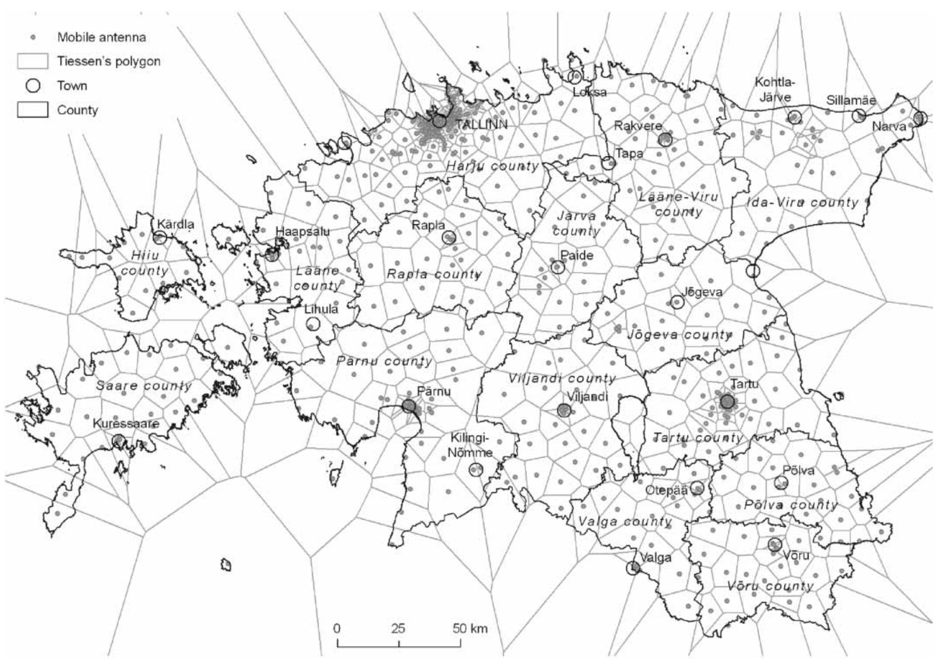

Figure 2. The EMT mobile network (mobile antennas and tiessen's polygons), counties, and major towns

in time and space. Anchor point modeling is one of the possibilities for making useable the anonymous data of passive mobile positioning, GPS tracking, etc.

In accordance with the frequency and regularity of the respondents' calls, we have chosen a calendar month to be the timestep for determining anchor points in the current paper. Thus we analyzed 12 months separately for every person, and the results are presented here as an average of 12 analyzed months for every person. Presumably, periods of one week would also produce results, but we believe that in Estonian conditions, the outgoing calls database yields better results for one-month periods.

The accuracy of pinpointing anchor points is at the level of the service area of a mobile antenna (a network cell) (See Figure 3). In different network systems, it is actually possible to divide the network cell into smaller cells, but this depends on the methods of data collection and storing. In this survey, we have used the Ericsson system, which dispenses sites of mobile antennas as a network cell because of the optimal price and quality of the data.

Terms Used

Respondent—a person (sometimes the term *phone* is also used), who owns a cell phone connected to the EMT network, who has performed calls on it, and who has been added into the passive positioning database by their random ID.
Regular Cells—network cell regularly visited by one respondent and from which the respondent has made calls on at least two different days a month.
Random Cell—a network cell from which the one respondent has made calls on only one day a month.
Meaningful Place—a personal anchor point or regularly visited place which has a significant meaning in the everyday life of the respondent.
Everyday Anchor Point—anchor point in which the respondent has spent time on most days, and which have thus been assigned as home or work places.
Secondary Anchor Point—anchor points that have lower visiting regularity than everyday anchor points.
Home Anchor Point—An everyday anchor point, at which the probable location of the respondent's home is determined, based on the model.
Work-Time Anchor Point—an everyday anchor point, at which the probable work-time location of respondents is determined, based on the model. The anchor is called a work-time location because it is not possible to differentiate between work, school, and other activities in the place where a person regularly and most often spends time in business hours during a month.
Multifunctional Anchor Point—an everyday anchor point in which the home and work-time anchor points are located in the same network cell and cannot be separately identified.

Description of Model

A model for detecting meaningful places or personal anchors of persons consists of eight steps (See Figure 3), of which each stage actually contains a complex PostgreSQL database manager query or framework. The important stages of model calculation have been brought out in the following list.

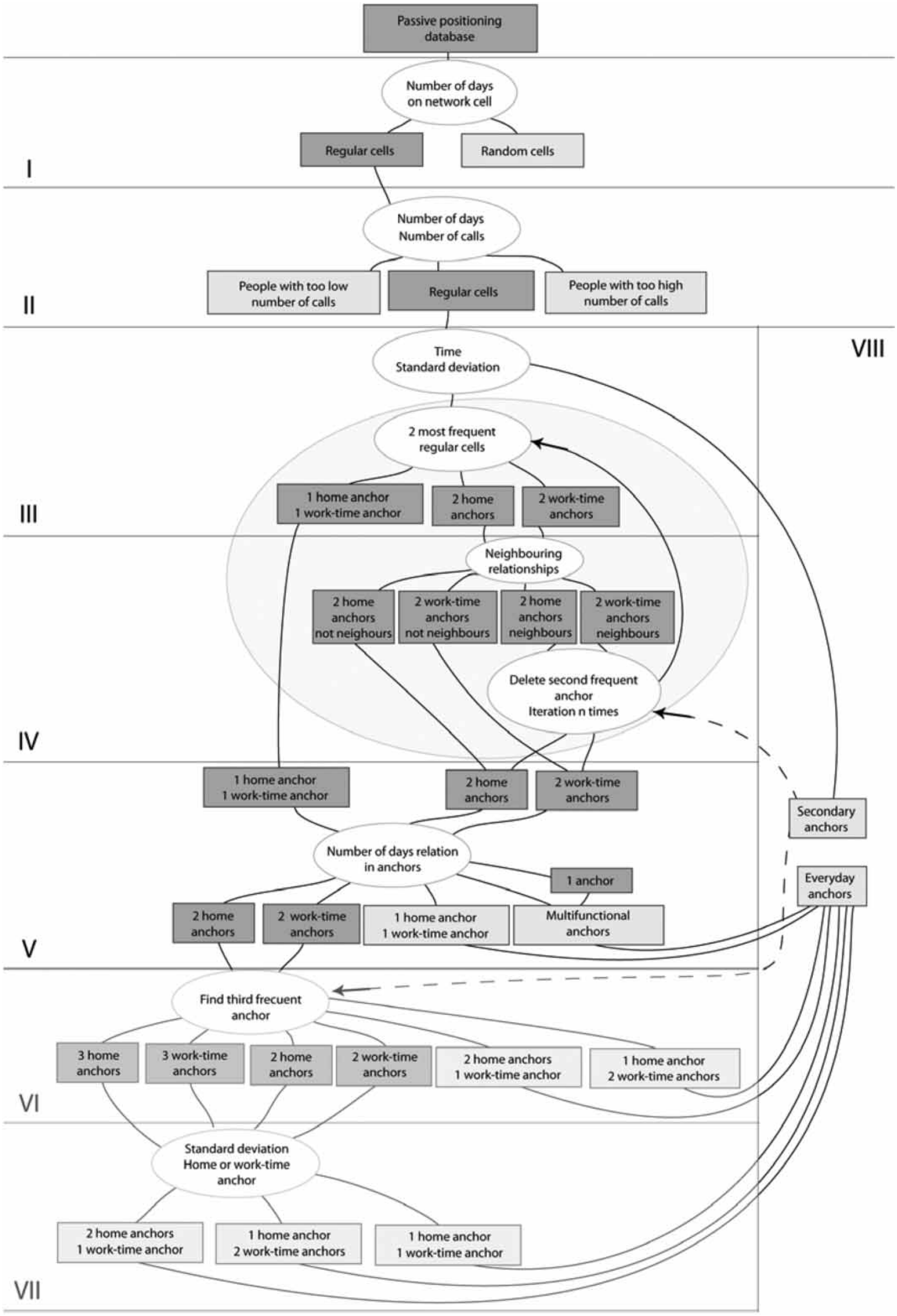

Figure 3. Anchor point determining model

In order to develop the model, we studied in detail 14 selected friends or relatives (five urban, five rural, and four suburban residents) whose logs of calls out were recorded during 12 months. We asked personally about all locations of calls from all respondents during interviews. The model was developed step by step, testing entries in log file and meaning of locations of those calls during 12 months. We compared the modeled home and work-time anchor points with each individual's actual places of residence and work. Figure 4 shows the location of one of our colleague's anchor points in the city of Tartu in November 2007.

Determining Points of Regular Cells and Separating Them from Random Cells. For every respondent (i.e., ID), the number of calls made is determined for each separate network cell. Based on this information, regular cells (network cells) where the respondent has performed calls on at least two separate days a month will be separated. Random cells from which the ID has made calls on only one day of the month will be left out of further analysis.

Removal of IDs with Too High or Too Low a Number of Calls Made. The next stage of the model removes respondents with too high or too low a number of calls from the database. For that purpose, call frequencies for every ID in every network cell are calculated. Regular cells are sorted by the number of days with calls, and if days overlap, by the number of calls. Then the respondents (IDs) who have made calls from their most visited network cell on fewer than seven days a month will be removed from the database. If the number of calls made is too low, it is not possible to calculate the anchor points.

On the basis of the same sorted list of regular cells, respondents with too many calls are also excluded. This exclusion is made on the presumption that the reason for there being too many calls is an organized call procedure (service center, etc.) or a technical device using a GSM network. Respondents are

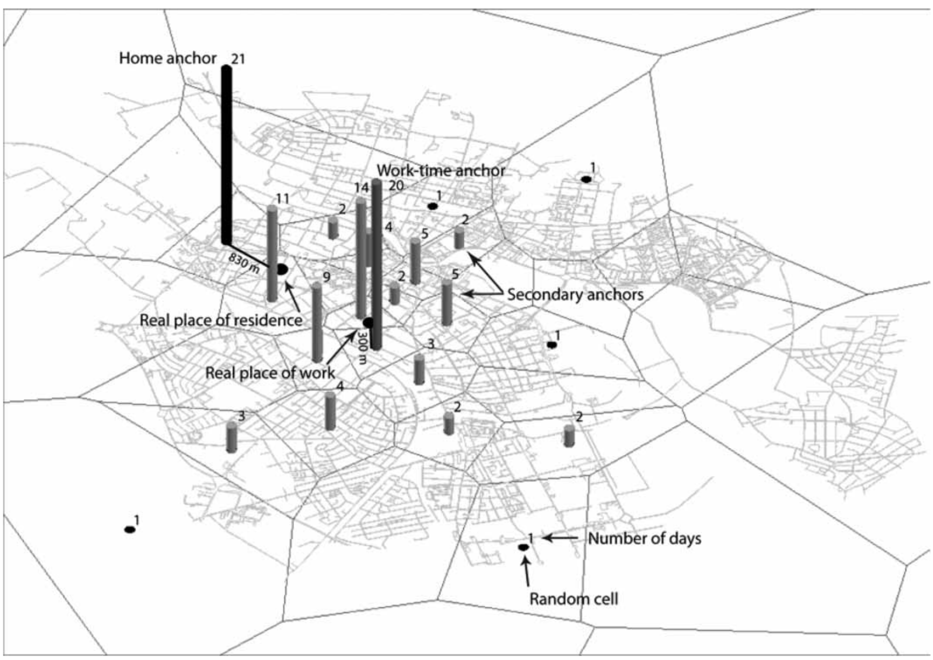

Figure 4. One of the respondent's anchor-point locations in the city of Tartu in November 2007
Note: Determined by means of passive positioning. The respondent's (Prof. Ahas) activities took place only in the city of Tartu.

considered to have too many calls if they have made more than 500 calls a month in their two most frequently visited network cells. In comparison, the average number of calls made per month in the two most often visited network cells by one person is 81.

Determining Home and Work-Time Anchor Points. The two regular cells that had the highest number of days with calls are selected for the calculation of home and work-time anchor points, and the rest are moved directly to the last stages of the model, where they are called secondary anchor points. For every regular cell, the average start time of calls (average of all calls made during a day) and the standard deviation of call beginning times are calculated. Home and work-time anchor points are determined from regular cells based on those two variables.

This model calculates the home anchor point from the regular cells where the average starting time of daily calls is after 17:00. If the value of the standard deviation of the beginning times of calls performed by a respondent is greater than 0.175, the home location is also determined for respondents whose average beginning time is before 17:00. If the average start time of calls is before 17:00 and the value of the standard deviation is less than or equal to 0.175, then these regular cells are assigned to be work-time anchor points.

To distinguish home anchor points from work-time anchor points, the standard deviation of the start times of calls is applied, since people's work periods have different durations, and thus the time is not sufficient for the determination of anchors. Our analyses (Järv *et al.*, 2007) have shown that the chronological variability (and, therefore, also the standard deviation) of calls performed at work is lower than that of calls made from residential locations (See Figure 5). This is probably explained by the fact that people are at home and also use their telephone at different times—at night, at lunch time, on the weekend, and during holidays, etc. The time period spent at work is more compact, and outside working hours people rarely spend time at the office. Of course, this model has weaknesses, for example if the respondent has more than one workplace or a very mobile job, or if they work at home.

In differentiating home and work-time anchor points, it is important to note that in addition to the common situation of having one home and one work-time

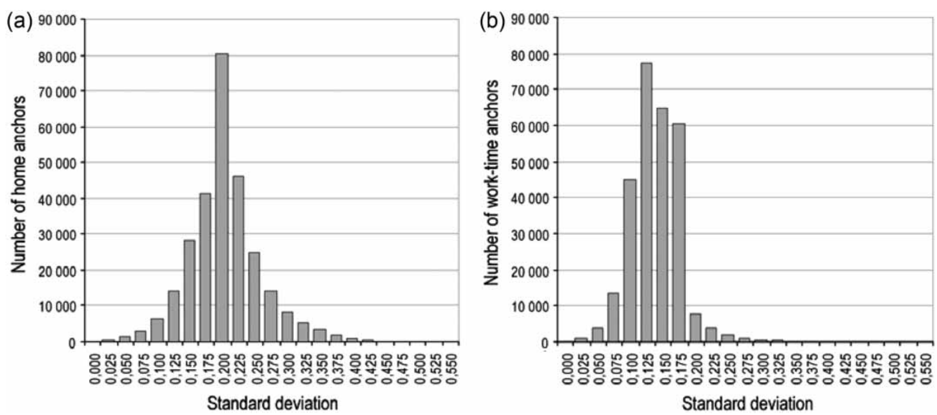

Figure 5. Distribution of home (a) and work-time (b) anchor points according to standard deviation of calls

anchor point, it is also possible that the respondent has two home or two work-time anchor points among the two most visited regular cells. At this point the model becomes branched. If the respondent has two home or two work-time anchor points, they move on to the next stage (Stage 4). Respondents with one home and one work-time anchor point that have been properly determined will skip the next stage and will be analyzed further in Stage 5.

Consideration of Neighboring Relationships in the Case of Two Home or Two Work-Time Anchor Points. If one respondent has two home or two work-time anchor points, this is probably due to the very common problem of "switching" or "tossing" connections between two neighboring antenna even if the person stays in the same room. It is typical that if the phone is situated in the border area of the network cell, the phone will be technically handed to the neighboring base station that has fewer users or better visibility. First the model determines whether those parallel anchor points are situated in neighboring cells or not. If the anchor points are situated in neighboring cells, the analysis is continued at the same stage. However, if they are not situated in neighboring cells, the two home or two work-time anchor points move on to the next stage (Stage 5) of the model.

In case they were located in neighboring cells, the model erases the second most frequently visited anchor point, first selecting the one with the higher number of days, and when the number of days is the same, on the basis of the number of calls. The most visited anchor point will also remain. After that, the model will use the earlier list of regular cells of the same respondent again, and the next most frequent, i.e. the third regular cell will be employed. The model will re-analyze it in Stage 3 and try to determine the missing home or work-time anchor point in the previously described manner. If the third most frequently visited regular cell differs from the previous two in terms of parameters, it is classified as the respective anchor point, and the respondent will then have one home and one work-time anchor point, which will be analyzed in Stage 5.

Similarly, anchor points will move on to the next stage (Stage 5) if the new determined anchor point is of the same type as the previous one, but is located in a non-neighboring cell. In that case, the respondent will actually have two locations that can be characterized as home or work-time anchors. Of course, the real meaning of those points is not recognizable by the data we possess. If the third most frequent anchor point does not qualify as the required missing anchor point, according to time, standard deviation, and location metrics, the model will search for required parameters from the data of the fourth regular cell. The iteration can be repeated n times. In our model n equals 3, meaning the same procedure will be executed, if necessary, until the fifth regular cell. However, if after n iterations it is still impossible to determine the missing anchor point, the person with two home or two work-time anchor points will move on to the next stage (Stage 5).

Assessment of the Proportion of Days Spent at an Anchor Point. If the most frequently visited anchor point covers more than 75 percent of the days a respondent stayed at the two most frequently visited anchor points, it is classified as the multifunctional (home + work-time) anchor point. In this case, the second most frequent anchor point will not be analyzed further, and it moves on to the eighth stage of the model, where it will be classified as a secondary anchor point.

Likewise, multifunctional anchor points are also formed when the respondent is left with only one anchor point, either home or work-time.

By the end of the fifth stage there are four different possible cases of respondents' home and work-time anchor points: 1) one home and one work-time anchor point, 2) one multifunctional anchor point, 3) two home anchor points or 4) two work-time anchor points. The first two cases are used to determine anchor points and then move on to the eighth stage, where they will be classified as everyday anchor points. The respondents with two home and two work-time anchor points will be analyzed further in Stage 6 of the model.

Determining the Missing Home or Work-Time Anchor Point by the Addition of a Third Point. In the case of persons with only two home and two work-time anchor points, the model will try to determine the missing anchor point by using the list of regular cells. The next most frequently visited regular cell will be selected from the list and assessed according to the average time and standard deviation of calls as in Stage 3. If the next most frequent regular cell differs from the two previous anchor points, the respondent will have either two home and one work-time anchor point or one home and two work-time anchor points. These anchor points are complete and they will move on to the eighth stage, where they are classified as everyday anchor points.

If the next most frequent regular cell employed is similar to the previous ones, the respondent will have either three home or three work-time anchor points. If the respondent does not have more regular cells to analyze in the database, only two home and two work-time anchor points will remain, these four possible variants will be analyzed further in Stage 7 of the model.

Classifying an Anchor Point as the Missing Home or Work-Time Anchor Point. In this stage the model again attempts to create both a home and a work-time anchor point for the respondent with two similar types of anchors. For this purpose, the standard deviation of existing anchor points is assessed once again. In the case of two home or two work-time anchor points, the anchor point with a greater standard deviation is classified as the home and the other as the work-time anchor point. In the case of three home or work-time anchor points, the standard deviations of the first two anchor points are similarly compared to the previous case, and the anchor point with the larger standard deviation is classified as the home and the other as the work-time anchor point.

By the end of this stage the respondent has either two home and one work-time anchor point, one home and two work-time anchor points or one home and one work-time anchor point. All of these anchor points move on to the last stage, where they are classified as everyday anchor points.

Formation of Everyday and Secondary Anchor Points. Everyday anchor points are most frequently visited, and the possible combinations for one person are: 1) one home and one work-time anchor point, 2) a multifunctional anchor point (home and work-time anchors in one cell), 3) two home and one work-time anchor point, 4) one home and two work-time anchor points. The rest of the regular cells, which are not everyday anchor points, will be classified as secondary anchor points.

Results

Results of the Model Calculation

The model calculations were carried out separately for each month during the research period that extended from November 1, 2006 to October 31, 2007. The results are given as an average of the 12 months. The following are the results of the model calculations by stages. Important statistics are given in Table 1.

Regular Cells. 556,681 respondents have a total of 3,589,351 points (network cells) from which they have made calls on at least two days, meaning an average of 6.4 regular cells per respondent. Calls placed by a respondent in a network cell on only one day a month were interpreted to be random cells and were removed from the calculations. Nearly 6.8 million calls or a 10 percent of the average number of calls were removed each month. The total number of respondents who did not place two calls in any of the network cells in one month was 35,872, and they were removed from further calculations, since in their case none of the network cells could be determined to be regular cells.

IDs with Too Low or Too High a Number of Calls. In the event of too few calls (most visited network cell less than seven days a month), an average of 102,688 respondents were left out every month. This makes 18 percent of all persons with regular cells (minimum in May: 92,883; maximum in February: 108,541). Likewise, respondents with too many calls were excluded, i.e. those who placed over 500 calls a month in the two most visited network cells, on average 4,200

Table 1. Modeling steps for determining anchor points and the number of entries

Stage		Anchor points	Number of IDs	Calls	
	Passive positioning database		592,553	65,279,606	
1	Random cells, *network cells in which a person made calls during only one day*		35,872	6,842,664	Removed
	Regular cells, *network cells in which a person made calls on at least two different days*	3,589,351	556,681	58,436,942	
2	IDs with too few calls *(no cells with two days of calls)*	240,709	102,668	1,476,666	Removed
	Number of regular cells after removing IDs with too few calls	3,348,642	454,012	56,960,276	
	IDs with too many calls *(more than 500 calls in two of the most visited network cells)*	71,583	4219	4,766,030	Removed
	Number of regular cells after removing IDs with too many calls	3,277,059	449,793	52,194,246	
3–7	Home anchor points	282,572			
	Work-time anchor points	284,859			
	Persons with both home and work-time anchor points	247,952			
	Multifunctional anchor points	178,458			
	Two home and one work-time anchor point	10,777			
	One home and two work-time anchor points	13,065			
8	Everyday anchor points, total	745,889	449,793		
	Secondary anchor points, total	2,531,170			

respondents, on average 1 percent of all respondents who have regular cells. In our assessment, the calls were made by automated devices, call centers, or unusually active persons. In comparison, the average number of calls for the remainder in locations of regular visits was 81 calls per month. The division of people by their most visited location based on the number of days spent there and the division of people with too low or too high a number of calls placed is presented in Figure 6.

Home and Work-Time Anchor Points. The analysis of the data found that respondents most often have two everyday anchor points: a home and a work-time anchor point. As a result of the model calculation, every month the model was used to locate an average of 282,572 persons' homes as places where people regularly spend their nights and also used their telephone. The model also located an average of 284,859 work-time anchor points a month where time is spent during the day (at work, school etc.). The average number of people who were simultaneously assigned a home and work-time anchor point was 247,952 over 12 months.

A relatively large number of respondents (178,458) had only one everyday anchor point, i.e. a unified multifunctional anchor point. The existence of one everyday anchor point is largely connected with the size of the network cells. There are rural areas where the density of the network is low and therefore there is quite a high probability that a person's home and work places are geographically in different locations but remain in the coverage area of the same antenna. These are therefore determined to be multifunctional anchor points. The distribution of multifunctional anchor points in Estonia shows their dominance in less populated rural areas and a lower number in cities and surrounding areas. The other presumable reason for the numerous occurrences of only one anchor point is people's high level of domesticity. As a result of the model calculation, a certain number of people were found to have two home and two work-time locations, but their numbers are relatively low: 10,777 with 2 home

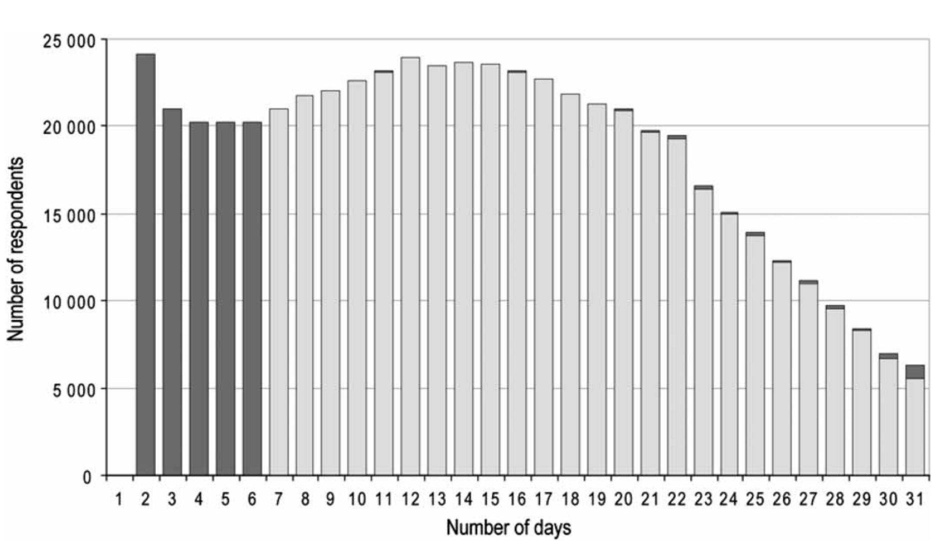

Figure 6. The division of people based on the number of days spent in the most visited location
Note: Anchor points highlighted in dark indicate persons with a number of calls that is too high or too low to be used in the analysis (removed from database).

and 1 work-time anchor points and 13,065 with 1 home and 2 work-time anchor points. (See Table 1.)

Everyday and Secondary Anchor Points. Home and work-time anchor points, which were calculated in stages from 5 to 7, were classified as everyday anchor points, and their monthly average was 745,889. In addition to everyday home and work-time anchor points, model calculations also determined the secondary anchor points (third most frequent regular cell etc), where people reside often, yet the frequency of visiting does not parallel the parameters related to home and workplaces. These might be locations where people shop, run errands, spend free time, visit family and friends, etc. During the 12 months there where 2,531,170 secondary anchor points. It was not the aim of the current research to study and find the meaning of secondary anchor points in great depth.

Geographical Distribution of Modeled Anchor Points

The geographical distribution of respondents' homes calculated using our model is presented in Figure 7. To determine the location of homes, home anchor points and multifunctional anchor points were summed. Places of residence are concentrated in larger cities and their surroundings, and more homes are also distributed in other more densely populated areas of Estonia, for instance the industrial North-East (Ida-Viru County) and the agricultural centers of Southern Estonia. The compared proportion of counties' populations in Estonia (See Table 2) shows that more than 38 percent of all home anchor points are focused in Harju County, followed by Tartu County with 12.7 percent, Ida-Viru County with 8.2,

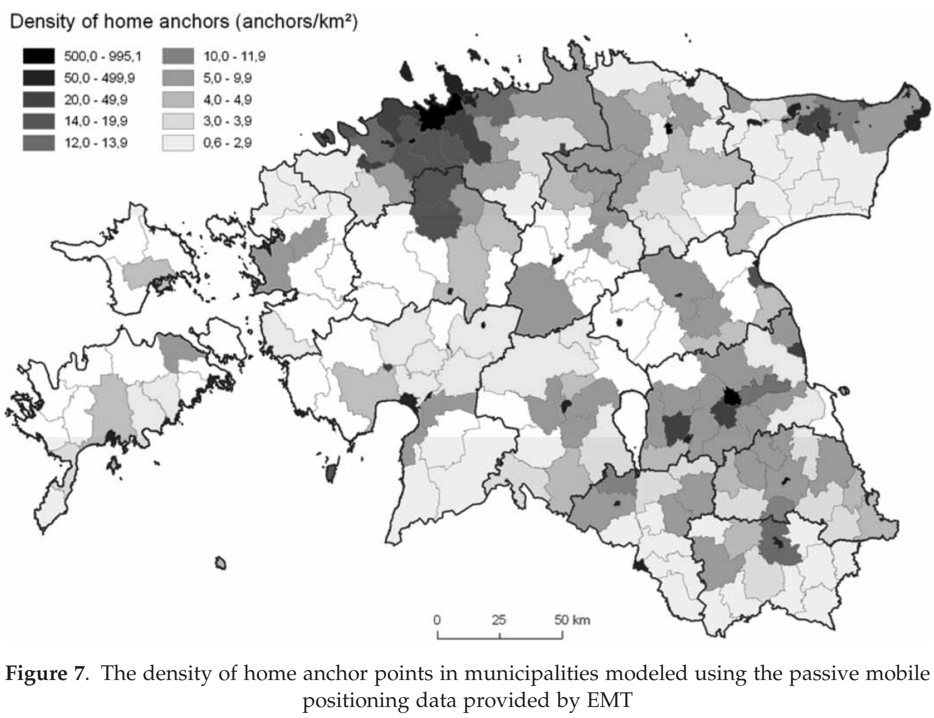

Figure 7. The density of home anchor points in municipalities modeled using the passive mobile positioning data provided by EMT

and Pärnu County with 6.5 percent. Tallinn, the capital of Estonia, hosts 68 percent of home anchor points in Harju County.

The geographical distribution of the modeled work-time anchor points (See Figure 8) also matches the major patterns of population densities. Work anchor points are composed by both work-time anchor points and multifunctional anchor points. The comparison of the geographical distribution of home anchor points and work-time anchor points showed that in larger cities there are more jobs than homes. For example, in our modeled data from Tallinn, there are 94,566 homes and 108,579 work-time locations, whereas in Tartu there are 18,190 homes and 32,348 jobs. Places of residence dominate over work-time locations in local governmental units in the surroundings of bigger cities. The proportion of work-time locations in Estonian counties is also quite similar to the above-mentioned locations of home anchor points.

Comparison of the Modeled Data with the Data from the Population Register

In order to verify outputs of our model calculations and unusual mobile positioning data, the distribution of homes calculated by the model was compared to the data from the Estonian Population Register. It was impossible to retrieve respective data of adequate quality and accuracy concerning the locations of work places.

There is a linear correlation ($r = 0.99$) between the number of modeled homes and the number of residents in the Population Register in Estonia's 227 municipalities. Local municipalities with larger populations also have more anchor points in our model. (See Figure 9.) Due to regional and social peculiarities, there are a number of deviations from the curve. Excluding the greater deviations in major cities Tallinn, Tartu, Narva, Pärnu, Kohtla-Järve, the correlation was $r < 0.86$.

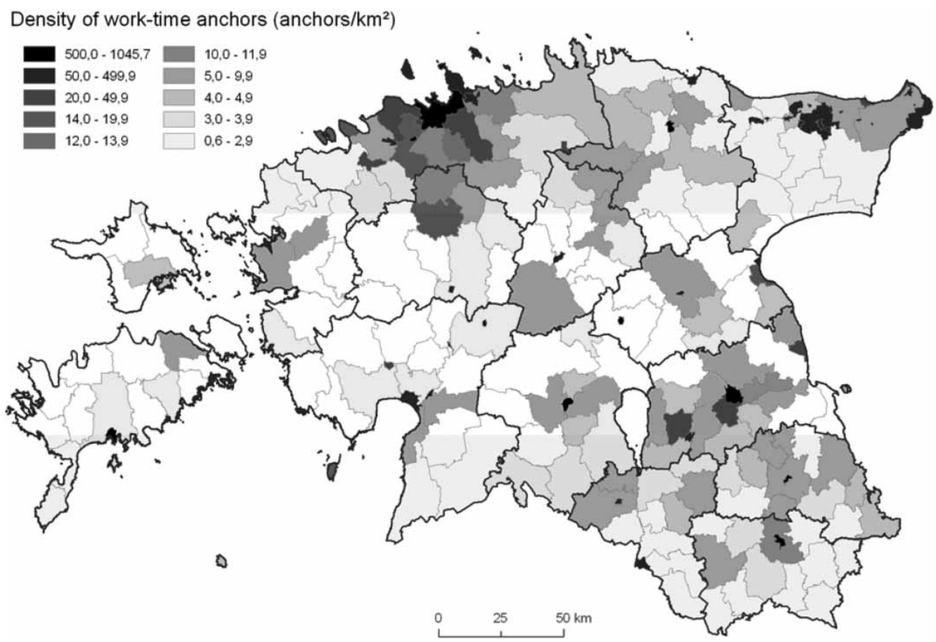

Figure 8. The density of work-time anchor points in municipalities modeled using passive mobile positioning data provided by EMT

Table 2. Share of the Estonian home anchor points and population register data by countries

	Place of residence				Work-time anchors	
	Number		Percentage		Number	Percentage
County	Mobile positioning	Population register 2007	Mobile positioning	Population register 2007	Mobile positioning	Mobile positioning
Harju county	181,969	536,152	38.4	39.4	187,308	39.5
Tartu county	59,997	146,749	12.7	10.8	61,017	12.9
Ida-Viru county	38,757	172,339	8.2	12.7	38,695	8.2
Pärnu county	30,564	90,994	6.5	6.7	30,583	6.4
Lääne-Viru county	24,107	68,090	5.1	5.0	23,734	5.0
Viljandi county	21,391	55,547	4.5	4.1	20,977	4.4
Rapla county	15,934	37,341	3.4	2.7	13,966	2.9
Põlva county	15,677	32,062	3.3	2.4	14,999	3.2
Võru county	15,487	39,058	3.3	2.9	15,362	3.2
Jõgeva county	14,511	36,698	3.1	2.7	13,963	2.9
Saare county	14,492	36,587	3.1	2.7	14,377	3.0
Valga county	14,269	35,007	3.0	2.6	13,872	2.9
Järva county	13,114	36,283	2.8	2.7	12,695	2.7
Lääne county	9809	28,249	2.1	2.1	9497	2.0
Hiiu county	3609	10,623	0.8	0.8	3549	0.7

In order to determine the deviations between two databases, the share of modeled homes and registered persons of every municipality in Estonia (total 100 percent) were calculated more precisely. These two percentage-based datasets were compared. There are more home anchor points based on the data of mobile positioning than those registered in the Population Register in the near vicinity of larger cities, especially Tallinn, e.g., municipalities such as Harku, Rae, and Loksa. This pattern also applies to some cities such as Tartu and Võru and their nearby surroundings. However, those deviations of homes were relatively small, with a maximum of 0.64 percentage points in Harku. The most probable reason for this is the intensive suburbanization that is taking place in Estonia (Tammaru *et al.*, 2009).

There are fewer home anchor points derived from mobile positioning data than those registered in the Population Register in most of cities of Estonia, for instance Tallinn, Narva, Sillamäe, Kohtla-Järve, Maardu, Pärnu, Elva, Valga, Jõgeva and in local governmental units farther from cities in central, western, and southern Estonia. The differences in the share of the total number of modeled homes from the Population Register reach 3.4 percentage points in Tallinn and 2.9 in Narva. In Kohtla-Järve, Sillamäe, and Maardu, the differences are about one percentage point.

If generalized to county level, the greatest differences between mobile positioning and Population Register data are found in North-East Estonia (Ida-Viru County), as according to mobile-positioning-based model output, there are 4.5 percent fewer people living and working there than shown by the Population Register. Poverty and related social problems may be one of the reasons why mobile data is underappreciated in the North-Eastern Estonian industrial cities, which results in the fact that many might not own a mobile phone or use cheaper calling cards. Data from the current study originates from EMT, which has both one of the best radio coverage and also above average price levels. In spring 2008 (February-April) we ordered a special survey of mobile telephone penetration and geographical share between operators from polling firm TNS

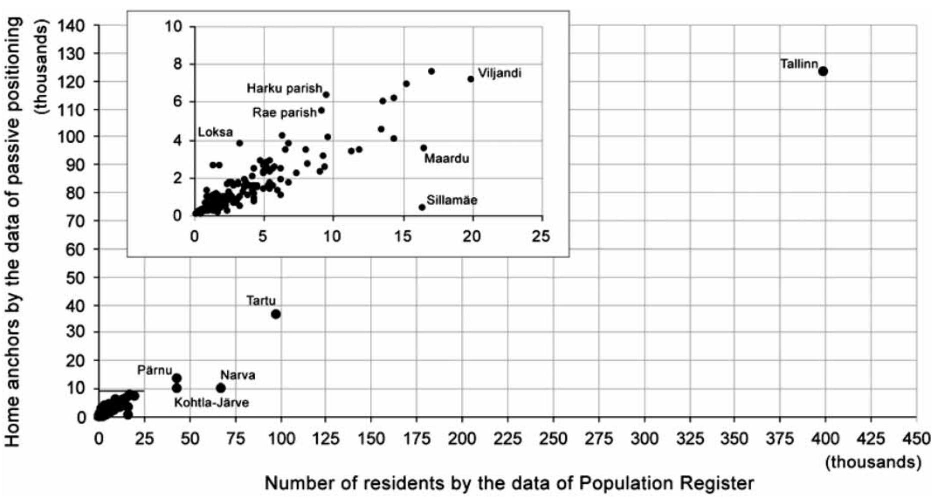

Figure 9. The correlation between the number of modeled home anchor points and the number of registered persons in local municipalities in Estonia

EMOR with 2000 respondents. The results showed that 95 percent of Estonians have a mobile phone, and that figure was 96 percent in cities. EMT had the greatest (70 percent) share of the market in South-Eastern Estonian rural Põlva county (74 percent), and 59 percent in Võru county. The smallest share of EMT was in Tallinn, with 39 percent, and northeastern Estonian counties Ida-Viru 37 percent and Lääne-Viru 39 percent.

The other area with greater differences is Tartu County, as mobile positioning data for that region shows a population that is 1.9 percentage points larger than shown by the Population Register.

Discussion

Model and Sampling

The relatively complex model of home and work-time anchor points that was developed during our practical traffic, commuting, and internal tourism research has been proven to be somewhat scientifically acceptable. Nevertheless, the whole system of inquiry and algorithm development need more exact analysis from the point of view of geographical accuracy, social characteristics, and the interpretation of an anchor point. All of the eight modeling stages need to be tested with databases of different levels. Also, depending on the use of telephones, the limits of standard deviations of our case by case set, temporal filtering times and calls may vary. The important thing about our model is its applicability in processing a large amount of data; our database had over 800 million data entries, and performing so many inquiries required a completely different approach. Our database and processing is based on PostgreSQL database manager inquiries, which can manage even larger amounts of data. Our discussions during seminar SPM2008 in Tartu showed that one of the greatest problems in using mobile positioning and GPS data in geographical studies is the great amount of data involved, which are not manageable with traditional programs and skills.

We believe that the model is accurate in finding routine places of personal activity spaces. The comparison with the Population Register was relevant. There is, however, still the possibility for systematic errors to occur, and these need to be controlled in future. For example, perhaps commuters' "most called" places are on their everyday route from home to work and vice versa, not home or work. Therefore the term of "meaningful place" is introduced from the beginning and used parallel with anchor point and regularly visited place. We think that because of specific passive positioning data, the term "meaningful place" fits better than anchor point which normally has fixed and known location. We also need to learn more about the meaning of quantitatively measured anchor points because such a quantitative model is good for monitoring larger areas or sampling, but is not good for analyzing personal activities, or places and their meaning.

Another important topic of discussion is the model's compatibility with different data sources. The model should be compatible with different sources, since mobile operators' infrastructures and information systems are different, and data outputs may exist on different levels in the mobile network. In this way, passive mobile positioning data may be obtained from terminals, positioning servers, billing memory, or indirect sensors of radio location. Different network standards differ in terms of data parameters for passive data outlets. For example, the size of network cells and protocol for hand-over differs in Nokia and Ericsson network systems (antennae, information system, positioning server, etc.) for GSMs used in Estonia's largest networks. A lot of work must be done to harmonize data outputs for CDMA, GSM, 3G, A-GPS and other types. The model must also be tested with data from GPS-based tracking experiments, which are very popular in studying space-time behavior.

A very important aspect to discuss is sampling. The data source is new, and there is a need for future analysis concerning technical standards, sampling, and methods (Ahas *et al.*, 2008a). The sample is influenced by technical aspects such as battery lifetime and the quality of radio coverage. For sampling, it is important to be aware of regional and social differences in the penetration of phones and operators, as well as peculiarities of phone use. For example, different ethnic groups may use phones differently in Estonia because of the different costs of calls. The activity of phone use is also determined by the age, profession, or travel behavior of the respondent. All of these aspects do influence the location of call activities and our geographical studies that depend on this data.

Geographical Accuracy

From the standpoint of the anchor point model, future discussions concerning geographical accuracy are necessary. Today's network cell accuracy makes it possible to collect comparative results, but there are still problems in sparsely populated rural areas where network cells and errors are greater. Confusion is also created when work-time and home anchor points are placed in the same cell in these big network cells; this was the case in 24 percent of the everyday anchor points in this survey. This makes certain research tasks significantly more difficult to fulfill. In terms of geographical distribution, the advantage of mobile networks is that they follow persons, their living spaces and roads, and because of that are geographically reasonably positioned (Ahas *et al.*, 2008a).

In general, it is possible to raise the geographical accuracy of the passive positioning data with the cooperation of operators, e.g., dividing cells into sectors or using better positioning techniques for data collection. However, more accurate data is always more expensive, and processing is more demanding. A technique to manage the "tossing between antennas" is also necessary. Tossing means that even if you are actually in one place, your phone may connect to different antennas and swap (toss) them automatically, depending on the antennas' workloads. In GSM networks this poses a great problem for the Cell ID data and our model. In the database, the location of the call is determined by the cell where it began. In a 3G network, the phone is simultaneously connected to three antennas and the calculations of location take all three into account. These technical problems need to be dealt with very thoroughly, and models must be able to unify data from different technical platforms.

Summary

The objective of the current study was to develop and to test the model for determining meaningful locations of mobile phone users as locations of homes and work-places. The database of Estonia's largest mobile operator, EMT, which covers the whole country, was used for modeling. In our opinion the results of the model calculations were quite good; this is supported by comparisons of our modeling results to those of the Population Register. Differences between the data of mobile positioning and population register were greater in socially unstable areas and regions affected by intensive urban sprawl. In a broader sense, the mobile positioning based data and methods may be the best or only available source for monitoring population processes in such unstable or rapidly developing regions.

We conclude that it is possible to monitor a population's geography and mobility by using the passive databases of mobile operators. However, the development of a perfect system demands a great deal of additional work, as the input data and model needs to be standardized for different sources and conditions.

In terms of geographical research, the current methodology is promising, and we sought to distinguish new features concerning different location points of the majority of population with relatively low research costs. This makes the database attractive for different research perspectives and applications. There is great potential for the development of real-time monitoring tools and geographical applications. There is growing interest in such geography of regularly visited places in geography for tourism development, traffic management, and urban planning applications. The information technology commercial sector can also use the model to develop and personalize mobile services; to develop personalized and location-aware advertisements; and to optimize radio coverage and other network services.

Acknowledgments

The authors wish to thank EMT Ltd., Ericsson Ltd., Positium LBS and all the people who participated in the experiment for their cooperation. The project was funded by Target Funding Project No. 0182143s02 of the Ministry of Education and Science and Grant of Estonian Science Foundation No. ETF7562 and Estonian Information Technology Foundation (EITSA).

Bibliography

R. Ahas, A. Aasa, A. Roose, Ü. Mark, and S. Silm, "Evaluating Passive Mobile Positioning Data for Tourism Surveys: An Estonian Case Study," *Tourism Management* 29 (2008a) 469–486.

R. Ahas, A. Aasa, S. Silm, R. Aunap, H. Kalle, and Ü. Mark, "Mobile Positioning in Space-Time Behavior Studies: Social Positioning Method Experiments in Estonia," *Cartography and Geographic Information Science* 34:4 (2007a) 259–273.

R. Ahas, J. Jauhiainen, S. Silm, M. Nuga, A. Tähepõld, E. Saluveer, and K. Kivi, *Cooperation Perspectives between Municipalities of the Urban Region of Tallinn (Tallinna ja ümbritsevate omavalitsuste koostöövõimalused ja perspektiivid valglinnastumise kontekstis)* (Tartu: University of Tartu, Institute of Geography, 2007b).

R. Ahas and Ü. Mark, "Location Based Services: New Challenges for Planning and Public Administration?" *Futures* 37 (2005) 547–561.

R. Ahas, E. Saluveer, M. Tiru, and S. Silm, "Mobile Positioning Based Tourism Monitoring System: Positium Barometer," in J. ÒConnor, W. Höpken, and U. Gretzel, eds., *Information and Communication Technologies in Tourism* (Wien, New York: Springer, 2008b), pp. 475–485.

Y. Asakura and E. Hato, "Tracking Survey for Individual Travel Behavior Using Mobile Communication Instruments," *Transportation Research Part C* 12 (2004) 273–291.

K.W. Axhausen, A. Zimmermann, S. Schönfelder, G. Rindsfüser, and T. Haupt, "Observing the Rhythms of Daily Life: A Six-Week Travel Diary," *Transportation* 29 (2002) 95–124.

X.L. Chen, H.M. Zhao, and P.X. Li, "Remote Sensing Image-Based Analysis of the Relationship between Urban Heat Island and Land Use/Cover Changes," *Remote Sensing of Environment* 104:2 (2006) 133–146.

M. Dijst, "Two-Earner Families and Their Action Spaces: A Case Study of Two Dutch Communities," *GeoJournal* 4:3 (1999) 195–206.

K.V. Dürckheim, "Untersuchungen zum Gelebten Raum: Erlebniswirklichkeit und ihr Verständnis Systematische Untersuchungen II," *Neue Psychologische Studien* 6 (1932) 383–480.

N. Eagle and A. Pentland, "Reality Mining: Sensing Complex Social Systems," *Personal and Ubiquitous Computing* 10:4 (2005) 255–268.

Eurobarometer, *E-Communications Household Survey, Fieldwork November – December 2006, Special Eurobarometer 274/Wave 66.3* (TNS Opinion & Social, 2007), <http://ec.europa.eu/public_opinion/archives/ebs/ebs_293_full_en.pdf>

Ericsson, *Home* (2008) <www.erricsson.com> Accessed April 2008.

European Parliament, Council of 12 July 2002, Directive 2002/58/EC (Directive on Privacy and Electronic Communications) <http://eur-lex.europa.eu/LexUriServ/site/en/oj/2002/1_201/1_201 2002073en00370047.pdf> Accessed April 2008

M. Flamm, *Comprendre le Choix Modal – Les Déterminants des Pratiques Modales et des Représentations Individuelles des Moyens de Transport* doctoral dissertation (Lausanne: Ecole Polytechnique Fédérale de Lausanne, 2004).

M. Flamm and V. Kaufmann, "The Concept of Personal Network of Usual Places as a Tool for Analyzing Human Activity Spaces: A Quantitative Exploration," paper presented at 6th Swiss Transport Research Conference, Ascona 2006).

G. Gartner, "Location-Based Mobile Pedestrian Navigation Services: The Role of Multimedia Cartography," paper presented at UPIMap2004 (Tokyo, Japan, 2004).

R.G. Golledge, "Cognition of Physical and Built Environments," in T. Garling and G. Evans, eds., *Environment, Cognition and Action: A Multidisciplinary Integrative Approach* (New York: Oxford University Press, 1990).

R.G. Golledge and R.J. Stimson, *Spatial Behavior: A Geographical Perspective* (New York: The Guilford Press, 1997).

R. Herman, "Monitoring Traffic in Real Time From the Comfort of Your Car," *Computerworld* 40:39 (2006) 31.

T. Hägerstrand, "What about People in Regional Science?" *Papers and Proceedings of the Regional Science Association* 24 (1970) 7–21.

O. Järv, E. Saluveer, and R. Ahas, *Analysis of Traffic Distribution and Frequencies in Kose-Võõbu ja Võõbu-Mäo Units of E263 Tallinn-Tartu-Luhamaa Highway (E263 Tallinna-Tartu-Luhamaa maantee Kose-Võõbu ja Võõbu-Mäo lõikude Liiklusuuring Mobiilpositsioneerimise Abil)* (Tartu: Positium LBS, 2007).

J.B. Kang, W. Welbourne, B. Stewart, and G. Borriello, "Extracting Places from Traces of Locations," in *Proceedings of the 2nd ACM International Workshop on Wireless Mobile Applications and Services on WLAN Hotspots* (Philadelphia, PA, USA: ACM Press, 2004).

T. Kopomaa, "Speaking Mobile: Intensified Everyday Life, Condensed City," in S. Graham, ed., *The Cybercities Reader* (London: Routledge, 2004).

E. Kutter, "Aktionsbereiche des Stadtbewohners: Untersuchungen zur Bedeutung der territorialen Komponente im Tagesablauf der Städtischen Bevölkerung," *Archiv für Kommunal- wissenschafte* 13 (1973) 69–85.

M.P. Kwan, "Interactive Geovisualization of Activity-Travel Patterns Using Three-Dimensional Geographical Information Systems: A Methodological Exploration with a Large Data Set," *Transportation Research Part C Emerging Technologies* 8:1–6 (2000) 185–203.

R.B. Mitchell and C. Rapkin, *Urban Traffic: A Function of Land Use* (New York: Columbia University Press, 1954).

D.M. Mountain, *Exploring Mobile Trajectories: An Investigation of Individual Spatial Behavior and Geographic Filters for Information Retrieval* doctoral dissertation (London: Department of Information Science, City University London, 2005) PhD Thesis.

D.M. Mountain and J. Raper, "Modeling Human Spatio-Temporal Behavior: A Challenge for Location-Based Services," in *Proc. of the 6th International Conference on GeoComputation*, (Brisbane, Australia: University of Queensland, 2001).

P. Nurmi and S. Bhattacharya, "Identifying Meaningful Places: The Nonparametric Way," J. Indulska *et al.*, eds, in *Proceedings of the 6th International Conference on Pervasive Computing* (*Pervasive'08*), (Sydney: Springer-Verlag, 2008), 111–127.

P. Nurmi and J. Koolwaaij, "Identifying Meaningful Locations," paper presented at 3rd Annual International Conference on Mobile and Ubiquitous Systems: Networks and Services (MobiQuitous), San Jose, CA, 2006).

R. Oldenburg, *The Great Good Place: Cafes, Coffee Shops, Community Centers, Beauty Parlors, General Stores, Bars, Hangouts, and How They Get You Through the Day* (New York: Paragon House, 1989).

Positium LBS (2008) <www.positium.ee> Accessed April 2008

C. Ratti, "Mobile Landscape: Graz in Real Time," in *Proceedings of 3rd Symposium on LBS & TeleCartography*, (Vienna Vienna University of Technology, 2005).

J.G. Raymer, G. Abel, and P.W.F. Smith, "Combining Census and Registration Data to Estimate Detailed Elderly Migration Flows in England and Wales," *Journal of the Royal Statistical Society Series A-Statistics In Society* 170:4 (2007) 891–908.

J. Reades, F. Calabrese, A. Sevtsuk, and C. Ratti, "Cellular Census: Explorations in Urban Data Collection," *IEEE Pervasive Computing* 6:3 (2007) 30–38.

T. Roos, P. Myllymäki, and H. Tirri, "A Statistical Modeling Approach to Location Estimation," *IEEE - Transactions on Mobile Computing* 1:1 (2002) 59–69.

R.C. Schank and R.P. Abelson, *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures* (Hillsdale: Lawrence Erlbaum, 1977).

B. Schwarze and S. Schönfelder, *ArcView-Extension VISAR: Visualisierung von Aktionsräumen, Version 1.6* (Zürich: Institut für Verkehrsplanung, Transporttechnik, Stassen- und Eisenbahnbau [IVT], 2001).

S. Schönfelder and K.W. Axhausen, *Structure and Innovation of Human Activity Spaces* (Arbeitsbericht: Verkehrs- und Raumplanung 258, 2004).

M. Sheller and J. Urry, "The New Mobilities Paradigm," *Environment and Planning A* 38:2 (2006) 207–226.

N. Shoval, "Sensing Human Society," *Environment and Planning B* 34 (2007) 191–195.

S. Silm, R. Ahas, and M. Nuga, "Gender Differences in Space-Time Mobility of Dwellers from New Residential Areas in the Suburbs of Tallinn," (Tartu: Study Report, Department of Geography, University of Tartu, 2008).

S. Silm and O. Järv, *Summerhouses, Users and Usage Frequencies in Saku County (Saku valla suvilate kasutuse uuring)* (Tartu: Positium LBS, 2007).

S. Silm, O. Järv, and R. Ahas, *Urban Sprawl in Kuressaare Urban Region (Valglinnastumine ja sellest lähtuvad mõjud Kuressaare linnaregioonis: esmane analüüs ja ettepanekud planeerijatele)* (Tartu: Positium LBS, 2007).

J.E. Spinney, "Mobile Positioning and LBS Applications," *Geography* 88 (2003) 256–265.

T. Tammaru, K. Leetmaa, S. Silm, and R. Ahas, "Temporal and Spatial Dynamics of the New Residential Areas around Tallinn," *European Planning Studies* 17:3 (2009) 423–439.

V. Vidakovic, "De Toedeling van Reistijd aan Activiteiten: Eerste Resultaten van Waarnemingen in Amsterdam," *Verkeerskunde* 32 (1981) 442–444.

J.L. Wolf, R. Guensler, and W.H. Bachmann, "Elimination of the Travel Diary: An Experiment to Derive Trip Purpose from GPS Travel Data," *Journal of the Transportation Research Record* 1768 (2001) 125–134.

Y. Zhao, "Standardization of Mobile Phone Positioning for 3G Systems," *IEEE - Communications Magazine* 40:7 (2002) 108–116.

C. Zhou, D. Frankowski, P. Ludford, S. Shekhar, and L. Terveen, "Discovering Personally Meaningful Places: An Interactive Clustering Approach," *ACM Transactions on Information Systems (TOIS)* 25:3 (2007) 1–31.