Estimating the capacity of the Location-Based Advertising channel

Győző Gidófalvi* and
Hans Ravnkjær Larsen

Geomatic ApS, Center for Geoinformatik,
Nybrogade 20, 1203 Copenhagen,
Denmark
E-mail: gyg@geomatic.dk
E-mail: hrl@geomatic.dk
*Corresponding author

Torben Bach Pedersen

Department of Computer Science,
Aalborg University,
Fredrik Bajers Vej 7,
9220 Aalborg, Denmark
E-mail: tbp@cs.aau.dk

Abstract: Delivering 'relevant' advertisements to consumers carrying mobile devices is regarded by many as one of the most promising mobile business opportunities. The relevance of a mobile ad depends on at least two factors: (1) the proximity of the mobile consumer to the product or service being advertised, and (2) the match between the product or service and the interest of the mobile consumer. The interest of the mobile consumer can be either explicit (expressed by the mobile consumer) or implicit (inferred from user characteristics). This paper tries to empirically estimate the capacity of the Mobile Advertising channel, i.e. the number of relevant ads that can be delivered to mobile consumers. The estimations are based on a simulated mobile consumer population and simulated mobile ads. Both of the simulated data sets are realistic and derived based on real-world data sources about population geo-demographics, businesses offering products or services, and related consumer surveys. The estimations take into consideration both the proximity and interest requirements of mobile ads, i.e. ads are delivered only to mobile consumers that are close-by and are interested, where interest is either explicit or implicit. Results show that the capacity of the Location-Based Advertising channel is rather large, which is evidence for a strong business case, but it also indicates the need for user-control of the received mobile ads.

Keywords: capacity; estimation; Location-Based Advertising; LBA; Mobile Advertising; MA; simulation.

Reference to this paper should be made as follows: Gidófalvi, G., Larsen, H.R. and Pedersen, T.B. (2008) 'Estimating the capacity of the Location-Based Advertising channel', *Int. J. Mobile Communications,* Vol. 6, No. 3, pp.357–375.

Biographical notes: Győző Gidófalvi is an Industrial PhD student at Geomatic ApS, Centre of Geoinformatik, working on Spatio-Temporal data mining for Location-Based Services. His professional interests also include machine learning and data mining, in particular, financial text mining. In 2000, he received a BS with honours in Computer Science from San Diego State University. In 2002, he received an MS in Computer Science and Engineering with a special focus on Artificial Intelligence from the University of California, San Diego. He is scheduled to finish his PhD at Aalborg University in the autumn of 2007.

Hans Ravnkjær Larsen is the Head of Marketing and Communication of Geomatic ApS, the Danish knowledge Centre for Geo-demographical Market Analysis. He received a MS in Geography and Mathematics from Roskilde University. As a Product Developer at the National Survey and Cadastre, he was one of the founders of the National Square Grid Denmark (a geo-reference system). He is also part of the editorial board of Geoforum Perspektiv – *the Journal of the National Geodata Society*. At Geomatic, he is responsible for marketing and maintaining both internal and external relations.

Torben Bach Pedersen has received his PhD in Computer Science from Aalborg University, and his MS in Computer Science from Aarhus University. He is currently an Associate Professor of Computer Science at Aalborg University. His research interests include OLAP, multi-dimensional databases, data integration, Location-Based Services, analysis of web-related data, privacy, data mining and business intelligence applications. He has published a large number of papers in peer-reviewed journals and conferences.

1 Introduction

Mobile or Location-Based Advertising (MA or LBA), i.e. sending electronic advertisements to consumers carrying mobile devices, is considered by many as one of the most promising business opportunities amongst Location-Based Services (LBS) (Kölmel, 2002). A recent mobile marketing survey suggests that about 7% of the mobile consumers would be willing to receive promotional text messages ‘if they were relevant’ (Gartenberg, Matiesanu and Scevak, 2006). According to other surveys, an even larger percentage of the mobile consumers are interested if they are rewarded in some way (Leppaniemi and Karjaluoto, 2005). In this paper, mobile ads are regarded as a means of presenting relevant information to a recipient, be it a commercial offer on an item on sale, traffic information, or a piece of public information.

To many people, the world seems to be more and more difficult to guide oneself through, thus the art of targeting information and services will prove to be of immense value. Only, efficient business cases have so far been very few, in spite of the market’s expectations. A broad range of aspects, or variables, determine the relevance and context of a mobile ad: distance to the mobile user, explicit or implicit interest of the mobile user, uniqueness (do not send ad twice within some interval), time and place of delivery, etc. To this extent, this paper describes an LBA framework and a LBA database that can be used for the management of mobile advertisements.

In lack of comprehensive, real data on the movements and behaviour of the population, estimation or simulation is extremely useful, bringing the models to life with real and well-documented consumption patterns. Using a simulated but realistic mobile consumer population and a set of mobile ads, the LBA database is used to estimate the capacity of the MA channel, i.e. the number of relevant ads that can be delivered to mobile consumers. Apart from this use, the LBA database and the estimates derived from it can also be used in mobile catchment area analysis to estimate business exposure. Results show that the capacity of the LBA channel is rather large (approximately 100 mobile ads per user within a single day), giving strong support for a business case. The same results can also be viewed as alarming, and indicate the need to incorporate user-control of the received mobile ads in the LBA framework, as suggested by the Mobile Marketing Association (2007).

The remaining of this paper is organised as follows. Section 2 reviews the related work. Section 3 defines the estimation problem both in the case of explicit and implicit interest. Section 4 describes the simulated data sets and their derivations from real-world data sources. Section 5 describes the method and technical foundations for delivering mobile ads while taking into account both the advertisers' and mobile users' interests. Section 6 proposes a revenue model for LBA. Section 7 describes the experiments and discusses the estimates resulting from them. Finally, Section 8 concludes and points to future research directions.

2 Related work

The estimations in this work are based on simulated movements of mobile users. Movements of users are influenced by physical, social and geo-demographical aspects of mobility. To aid the development in mobile data management, a number of moving object simulators have been proposed in the literature that model primarily the physical aspects of mobility to various extents. Since most objects use a network to get from one location to the other, a state-of-the-art framework for network-based moving object simulation is presented in Brinkhoff (2002). The behaviour of a moving object in this framework is influenced by

1 the attributes of the object having a particular object class

2 the combined effects of the locations of other objects and the network capacity

3 the location of external objects that are independent of the network.

For a review of other moving object simulators the reader is referred to Gidófalvi and Pedersen (2006).

Moving object simulators generally neglect the social and geo-demographical aspects of mobility. These social and geo-demographical aspects of mobility introduce patterns in the movement of users and give rise to a unique Spatio-Temporal (ST) distribution of users. ST-ACTS is a ST ACTivity Simulator that using real-world data sources models, some of these neglected aspects of mobility (Gidófalvi and Pedersen, 2006). To make the estimations in this work as realistic as possible, movements of mobile users are obtained from ST-ACTS, which is further described in Section 4.4.

Database indices allow the effective management and retrieval of data in large databases. Spatial and geographical databases manage information about spatial objects, i.e. objects that have physical properties such as location and extent. An R-tree is a widely used index structure that allows the effective management and retrieval of spatial objects (Guttman, 1984). An R-tree splits space with hierarchically nested and possibly overlapping Minimum Bounding Rectangles (MBRs). Search algorithms that test spatial relationships (e.g. intersection, containment, nearest) between spatial objects can effectively use MBRs to decide whether objects in two MBRs satisfy a specific spatial search criterion.

The location of a moving object changes over time. Thus, the path of a moving object is commonly described as a sequence of coordinate and timestamp pairs and is referred to as the trajectory of the moving object. Moving objects databases are databases that represent and manage changes related to the movement of objects. ST indices such as the STR-tree and Trajectory-Bundle tree (TB-tree) allow the effective management and retrieval of information about moving objects (Jensen, Pfoser and Theodoridis, 2000). An STR-tree organises line segments of a trajectory according to both their spatial properties and the trajectories they belong to, while a TB-tree preserves only the trajectories. While these ST indices are designed to effectively manage trajectories, they are not available in commercially available Relational Database Management Systems (RDBMSs). Hence, the herein presented method uses the widely available R-trees.

Time geography (Hägerstrand, 1975) is a conceptual basis/paradigm for human space–time behaviour which considers

1. the indivisibility or corporeality of the human condition
2. that humans typically operate over finite intervals of space and time
3. the natural laws and social conventions that partially constrain space-time behaviour
4. that humans are purposive.

The movements of mobile users used in the estimations are derived from ST-ACTS (Gidófalvi and Pedersen, 2006), which models some aspects of this paradigm.

Research has shown that LBS have not yet been as widely used as expected (Kaplan, 2006). In opposition to earlier forecasts and market expectations, technology has not been ready until now. Furthermore, and no less important, is that the user and the user needs have not been fully understood. Recent research shows that in order to succeed with location-based ads, content is imperative (Komulainen, Ristola and Still, 2006) as well as taking the consumers' permission, acceptance and responsiveness into account is crucial (Barnes and Scornavacca, 2004; Heinonen and Strandvik, 2007). Being aware of this, the concept proposed in the present paper takes off from a point where any user will benefit directly from the use of the system, thus leading to an increased opt-in readiness.

Ensuring full user responsiveness is by no means trivial; research indicates that behavioural intention to use mobile commerce can be greater for mobile commerce-non-users than for users. Those most used to mobile commerce are not as prone to take action on it. However, although this documents that the triggers of mobile commerce acceptance is not yet fully understood (Lee and Jun, 2007), mobile users are found to be eager to make use of their phones in new ways, and methodologies are proposed to model user willingness (Leppaniemi and Karjaluoto, 2005; Pousttchi and Wiedemann, 2006; Fox et al., 2006). In the present paper, by targeting content to the right recipients, our setup is realistic and should be attractive for the typical future user.

Market research document a significant increase in sales to customers who were exposed to MA compared to those who were not exposed (Merisavo et al., 2006). In other words, it works, if only the above-mentioned issues are treated seriously. It is the intention with the present paper to contribute in furthering the spread of LBS.

3 Problem statement

Let $A = \{a_1,\ldots,a_n\}$ be the set of ads. Each ad a has a location *adloc*(a) and is for a certain product *prod*(a). Let $U = \{u_1,\ldots,u_n\}$ be the set of (moving) users. Each user u has a location *uloc*(u,t) depending on the time t, an explicit interest profile *expint*(u) containing a set of products, and an implicit interest profile *impint*(u), containing a set of demographic variable values. We also assume a scoring function *score*(u,a) that given a user u and an ad a returns a value between 0 (no match) and one (perfect match) that predicts how interested the user u is in product *prod*(a) based on the values in *impint*(u). Given a maximum distance maxdist between user locations and ad locations, and a timespan $T = [t_{\text{start}} : t_{\text{end}}]$, the explicit location-based ad delivery estimation problem is to estimate how many times a user u has a location *uloc*(u,t) within maxdist from *adloc*(a) for an ad a in A and a time t in T where *prod*(a) $\in$ *expint*(u). Given a maxdist between user locations and ad locations, a minscore, and a timespan $T = [t_{\text{start}} : t_{\text{end}}]$, the implicit location-based ad delivery estimation problem is to estimate how many times a user u has a location *uloc*(u,t) within maxdist from *adloc*(a) for an ad a in A and a time t in T where *score*(u,a) $\geq$ minscore.

4 Data

The estimations stated in Section 3 are based on a number of real-world data sources. The use of real-world data sources is important to derive realistic estimates. While the data sources refer to the Danish market and population, similar data sources are available for other major markets (ESRI, 2007; Experian, 2007). The following subsections describe in detail the data sources used to derive the estimates.

4.1 conzoom® demographic data

conzoom® is a commercial database product that contains fine-grained, geo-demographic information about Denmark's population (Geomatic aps – Centre of Geoinformatik, 2007). The variables that describe the statistical characteristics of the population can be divided into three groups: person, housing unit, and household variables. Table 1 shows these variables and the number of categories for each.

In Table 1, variables that have 'type' in their names are categorical variables; variables that have 'count' in their name are counts of the corresponding entities within a 100-m grid cell; and finally, the rest of the variables are continuous variables that have been categorised into categories that are meaningful for market segmentation. Since, for example in the countryside, the number of persons, households or units could be very low in a 100-m grid cell, grid cells are grouped together into meaningful, large enough clusters to comply with social and ethical norms and preserve the privacy of individuals. The basis for clustering is twofold: geography and the publicly available one-to-one housing information. The intuition behind the basis is also twofold. First, people living in a given geographical region (be that a state, a county, or a postal district) are similar in some sense; e.g. they might be more likely to have a certain political orientation than people living in another geographical region. Secondly, people living in similar houses are likely to be similar in other demographic variables; e.g. an established family with a stable source of income is more likely to be able to buy a larger, more expensive house than a person who just started his/her career.

Table 1 Variables in conzoom®

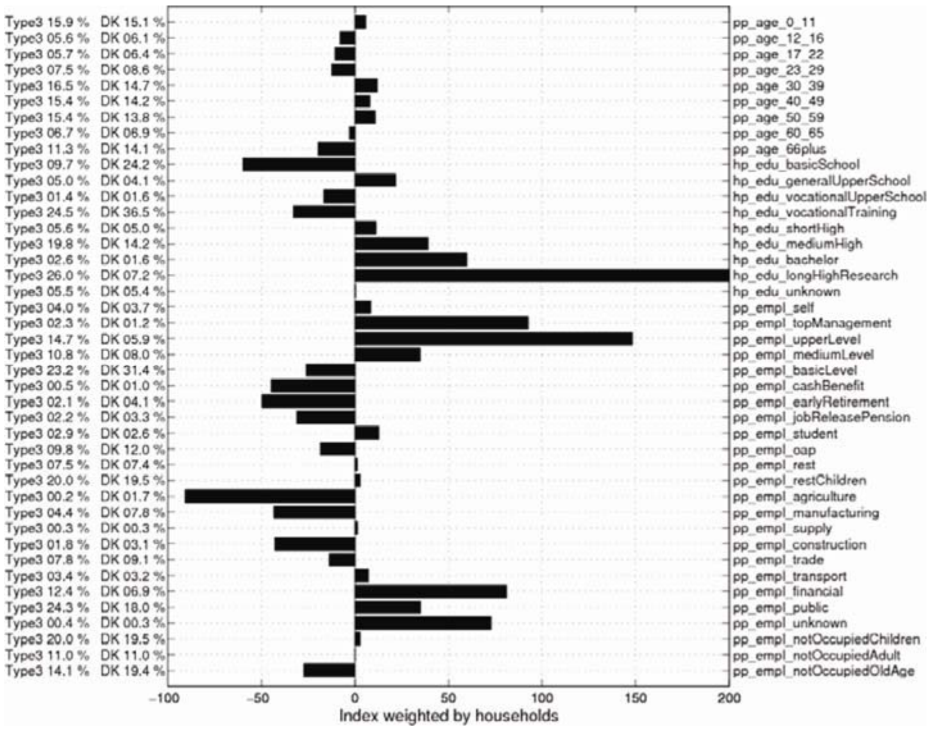
Referred entity	*conzoom® variable*	*Categories*
Person count	Person count	1
	Age	9
	Education type	9
	Employment status type	12
	Employment branch type	12
Housing unit	Unit count	1
	House type	6
	House ownership type	4
	House area	5
Household	Household count	1
	Family type	5
	Fortune	6
	Personal income	5

As mentioned earlier, to preserve the privacy of individuals, the clusters are constrained to contain at least some fixed number of households. Statistics for the variables, depending on the sensitivity of the information contained in them, are obtained from Statistics Denmark (2007) for clusters constructed at an appropriate level of cluster size constraint, e.g. 20, 50, 100 and 150 households per cluster. In case of a continuous variable, e.g. age, counts of the corresponding entities (in this case persons in the cluster) are obtained for the categories of the given variable. Due to this constrained geo-clustering method, the conzoom® clusters obtained comply with the social and ethical norms and preserve the privacy of the individual, yet the statistics obtained are accurate enough for effective market segmentation. This segmentation results in a grouping of the Danish population into 29 conzoom® types, one of which is defined for each 100-m grid cell. Cosmopolitan (type 3) is one example of the 29 conzoom® types. Comparing the demographics of type 3 to the demographics of the rest of Denmark's population gives the demographic profile of the type. This profile is partially shown in Figure 1. It roughly describes individuals that are more likely:

- to be middle aged (30–59 years old)
- to live in larger cities in larger, multi-family houses that are either owned by them or private rentals
- to be mostly couples with children
- to have a medium to long higher education

- to hold higher level or top management positions in the financial or public sector
- to have a better household economy (in terms of wealth and income) than the average Dane.

Figure 1 Partial profile of conzoom® type 3



4.2 *GallupPC® consumer survey data*

GallupPC® is a commercial database product and as the name suggests, it contains detailed survey responses of consumers about their demographics; interests such as culture, hobbies, and sports household consumptions, purchasing habits; transportation habits; views on various subjects; attitudes and exposure to various advertisement media (The Gallup Organization, 2007). The questions in the surveys are yes/no questions. To measure the magnitude of the consumer's interest in a specific area, the original yes/no question is re-phrased as categorical questions. For example, the original yes/no question 'Are you interested in fashion?' Is re-phrased to five yes/no questions using the following answer possibilities: very, rather, somewhat, not very, or not interested.

4.3 *bizmark™ products and services*

bizmark™ is a commercial database product that contains detailed information about Danish businesses both in the public and private sector (Geomatic aps – Centre of Geoinformatik, 2007). Some of the one-to-one information that is available about businesses is their location, the number of employees working in them, the physical size of the business facility, and the international branch codes the businesses fall under. Using the hierarchy of international branch codes 40 product and service categories were identified for which related consumer surveys were also available. The product and service categories are as follows: classical concert; pop/rock concert; discothèque; art exhibition; museum; cinema; theatre; pharmacy; bicycle/moped; car, stereo/HI-FI; CDs/DVDs; computer/internet; new technologies/telecommunication; do-it-yourself; fashion; cosmetics/skincare; glasses/contacts; hairdresser; jeweller/watches; interior design; travel; pets, fast-food; and 14 brand specific supermarkets. Based on the international branch codes a one-to-many relationship has been established between a subset of the businesses in bizmark™ and the 40 product/service categories.

4.4 Simulating mobile users with ST-ACTS

ST-ACTS is a probabilistic, parameterisable, realistic ST ACTivity Simulator (Gidófalvi and Pedersen, 2006). ST-ACTS is realistic in the sense that it is based on a number of real-world data sources (among others, the data sources described above) and a set of principals that try to model the social and some of the physical aspects of mobility. The modelled principles that govern the social aspects of mobility are:

1 People move from a given location to another location with an objective of performing some activity at the latter location.

2 Not all are equally likely to perform a given activity. The likelihood of performing an activity depends on the interest of a given person, which in turn depends on a number of demographic variables.

3 The activities performed by a given person are highly context dependent. Some important parts of context are: the current person location, the set of locations where a given activity can be performed, the current time, and the recent history of activities of the person.

4 The locations of facilities, where a given activity can be performed, are not randomly distributed, but are influenced by the locations of other facilities and the locations of the users of those facilities.

The output of ST-ACTS is a population of simulated persons, each described by a set of demographic variables and associated with a trajectory. The trajectories are sequences of time-stamped activities performed at particular physical locations, i.e. coordinates. In addition to the four principles above, the simulated activities also obey the following constraints. First, the temporal activity constraint, which states that certain activities are more likely to be performed during some periods than others. Secondly, the activity duration constraint, which states that not all activities take the same amount of time.

Thirdly, the maximum distance constraint, which states that for most activities there is a maximum distance a person is willing to travel. Finally, the trajectories assume linear movement between two consecutive activities, i.e. locations, but obey some physical mobility constraints, namely, that it takes time to move from one location to another. The time it takes to move from one location to another is calculated based on the distance between the two locations and a realistic speed model that assigns lower speeds to shorter, and higher speeds (with larger variance) to longer distances.

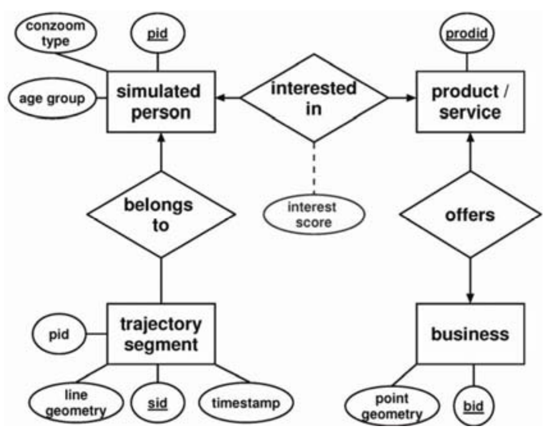
5 Method

The method presented here uses the Oracle RDBMS, and one of its extensions, Oracle Spatial, which provides advanced spatial features to support high-end Geographic Information Systems (GIS) and LBS solutions (Oracle Spatial and Locator, 2007).

5.1 LBA relational database

The objects or entities in the database are: simulated persons (or equivalently referred to as mobile users), trajectory segments, businesses, products and services. A simplified extended Entity-Relation (ER) diagram of the database is shown in Figure 2. In the extended ER diagram, square boxes represent entities, oval represent properties of entities and diamonds represent relationships between entities. Underlined properties represent primary constraints. The arrows between entities encode connectivity of relationships, i.e. arrow represents ‘one’ and no arrow represents ‘many’. For example, the ‘belong to’ relationship is a many-to-one relationship between trajectory segments and simulated persons, i.e. one trajectory segment belongs to exactly one simulated person, but many trajectory segments can belong to one simulated person. Mobile ads are indirectly modelled by in the relational database the many-to-many ‘offers’ relationship between businesses and products or services. Through this relationship, a mobile ad can be thought of as an entity having a unique identifier composed of a unique combination of `bid` and `prodid`, and having a location specified by the point geometry of the business offering the advertisement.

Figure 2 Simplified, extended ER diagram of the Mobile Advertising database



As it was introduced earlier, in the implicit case, the interest of a mobile user u and a mobile ad a about a product or service $prod(a)$ is not a Boolean function or binary relation. Rather, it is a continuous function that given the demographic characteristic $impint(u)$ of u, assigns a real valued interest score $score(u,a)$, usually from zero (not interested) to one (very interested), for $prod(a)$. In direct marketing, this function is termed as a scoring function, which encodes a particular scoring model. This real valued scoring function is untraditionally represented as a property of the ‘interested in’ relationship in the ER diagram.

5.2 Proximity requirements on mobile ads

A mobile ad *a* is likely to be considered relevant to a mobile user *u* only if at the time of delivery *t*, *u* is (or at some foreseeable future time point will be) within a maximum distance, maxdist, to the origin of the mobile ad *adloc*(*a*), i.e. the location of the business. Using the spatial features of Oracle Spatial, this proximity criterion between mobile ads and mobile users is tested as follows. The geometries of the businesses, equivalently mobile ads, are buffered to a maximum distance, and tested for any spatial interaction with the geometries of the trajectory segments, by performing a spatial join operation in the database. To make the join operation as fast as possible, geometries are indexed using R-trees.

5.3 Interests based on demography

The relevance of a mobile ad for a particular product or service is naturally influenced by the interest of the user for the given product or service. As described above, a subset of the GallupPC® consumer survey questions are related to products or services that can be directly linked to businesses in bizmark™, and measure the interests of the consumer in the products or services.

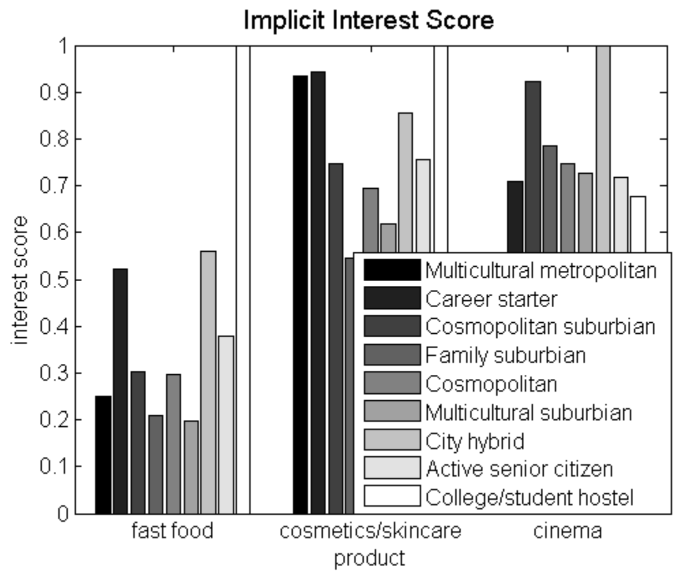
Using the geo-demographic parts of the surveys, each survey subject is assigned to one of the 29 conzoom® types. To derive a single indicator, an interest score, for how interested a given conzoom® type is in a given product or service, the answers to the questions processed as follows.

First, the five possible answer choices are associated with the following interest scores: very interested (1), rather interested (0.75), somewhat interested (0.5), not very interested (0.25), and not interested (0). Secondly, for a given conzoom® type and product or service the interest scores assigned to individual answers are averaged. Finally, the mean interest scores for a given product or service are scaled to the [0, 1]-interval amongst the 29 conzoom® types. Figure 3 shows a sample of these interest scores for a subset of the conzoom® types. In Figure 3, it can be seen that college students are most interested in fast food and cosmetics/skincare products, and among the conzoom® types listed, suburban families are least likely to be interested in the same.

5.4 LBA – implicit interest case

The interest score of a mobile user *u* in a particular product or service, which is advertised by mobile ad *a*, is implicitly encoded in the demographic characteristic, *impint*(*u,a*), or historical behaviour (reaction to previously received mobile ads) of *u*. The latter encoding is commonly referred to as relevance feedback in the scoring task in direct marketing, and while not considered in the current MA database, it can be naturally incorporated. In direct marketing, the model for this interest score is usually derived for one or many product(s) or service(s) of a particular company through the process of data mining or machine learning. This model can be represented as a company-specific interest score function in the MA database that for given user-demographics and historical user-behaviour, see Section 5.8, assigns an interest score to the user. In the estimations, however, these interest score functions are not company – but rather only product – and service-specific. Furthermore, due to their simplicity, they are implemented as table with the following schema: `interest_score=⟨conzoom_type,prodid,score⟩`.

Figure 3 Some interest scores for products or services for different conzoom® types



5.5 LBA – explicit interest case

Mobile users can also explicitly state their interest in certain products and services. In this case, the ‘interested in’ relationship is a binary relationship in the MA database. To provide a realistic estimates, the explicit interests of users are probabilistically simulated by randomly drawing a fixed number of products for every mobile user according to the distribution of interest scores given the conzoom® type of the user.

5.6 Uniqueness and user-defined quantitative constraints on mobile ads

Receiving the same ad multiple times naturally decreases the relevance of the ad as the therein presented information is not new. Primary key constrains in RDBMSs are an effective mechanism for guaranteeing that only unique combination of mobile users and mobile ads are considered for delivery. In the mobile marketing database, the delivered ads are stored in a table `mobile_ad_delivery=⟨pid,bid,prodid,delivery_time⟩`. Placing a primary key constrain on the first three columns guarantees that a mobile ad is delivered at most once to a mobile user. Recording the delivery time allows the control of the re-delivery of mobile ads after a certain period of time has passed. For clarity, the `mobile_ad_delivery` table is omitted from the ER diagram in Figure 2.

As the number of mobile ads increases, or the other constraints on the delivery of mobile ads weaken, the number of mobile ads delivered to a mobile user will naturally increase. After a certain number of ads have been delivered to the user, any additionally delivered ad, while maybe relevant, will likely be perceived as annoying. Hence, the mobile user's ability to limit the number of delivered ads is important. This user-control can be effectively facilitated by the top-k query mechanism which is provided in most RDBMSs.

5.7 User-defined ST constraints on mobile ads

Time and location are important aspects of the context of mobile ads. Most users would consider receiving a mobile ad as intrusive or disturbing when receiving it during work hours or after a certain time in the evening in their homes. Hence, the mobile user's ability to prevent the delivery of mobile ads in certain regions of space and time are important. While the user-control of ST constraints on mobile ads is not present in the MA database, the database can be easily extended to accommodate for this feature as follows. Users can specify mobile ad profiles by restricting certain spatial and/or temporal regions for mobile ad delivery. Then, ST joins between the mobile ad profiles and mobile ads can be performed to further control the delivery of mobile ads.

5.8 Inferring personal interests and relevance based on user-behaviour

Geo-demographic variables can be used to predict the general interests of an individual user, as explained in Section 5.3. However, since an individual user cannot be perfectly characterised by a few geo-demographic variables, it is likely that the predicted general interests differ slightly from the true, personal interests of the individual user. In the following, two methods are proposed to infer the personal interests of individual users.

The first method uses the locations that an individual user visits to infer the user's personal interests. Some of the locations a user visits are commercial in nature, i.e. businesses that offer products or services. During the lifetime of a user, the frequencies of how often the user visits particular businesses or types of businesses that offer particular products or services, can be recorded. Furthermore, periodical patterns can also be easily detected in the sequence of visits. An example of such a simple periodical pattern would be that a user visits a hairdresser approximately every second month. Basing the delivery of mobile ads on the personal frequencies of visited locations and periodic patterns ensures a closer match between the mobile ad and the true personal interest of the mobile user. The storage, maintenance, and derivation of the frequencies of visited locations and periodic patterns in those visits can either be done on the server or the client side. For the server side management, information about the frequencies of visits to particular types of businesses, offering specific products or services, are stored in a table with the following schema: `visit_frequency = ⟨pid,prodid,num_visits⟩`. Simple periodic patterns are stored in a table `pattern_period= ⟨pid,prodid,last_visit_time,period⟩`. To preserve clarity, these two tables are omitted from the ER diagram in Figure 2. The same information about personal interests can also be managed on the client side by a client application. Such a client side application, based on the current location of the mobile user, would have to be able to infer the type of business (product or service) that the user is currently visiting. This inference can either be aided by the server, or smart transmitters located at the businesses could communicate the required information to the client application.

The second method uses the information about how an individual user has reacted to mobile ads received in the past. More specifically, assume that a mobile user *u* received mobile ad a at time t. If at any time t' within a time period δt after time t the mobile user u visits the business that offered mobile ad a, i.e. $uloc(u,t') = adloc(a)$ and $t' - t \leq \Delta t$, it is considered as a strong indication of interest of the user towards the mobile ad. Similarly, if the user does not visit the offering business within the Δt period, it is considered as a weak indication of the user's indifference towards the mobile ad. The two events mentioned above can thus be considered as positive and negative feedback in an active control loop, respectively. The positive and negative feedback values can be quantified and summed over the lifetime of the user for specific businesses or types of businesses, i.e. products or services. After the control loop sum for a particular business or a product or service goes below a certain (user-defined) threshold, the business or product or service is 'black-listed' for the user, i.e. no more ads are delivered from the business or for the product and service to the user. The information needed to manage the control loops is already stored in the LBA database. In particular, the locations of users at time t are stored in the `trajectory_segment` table, and the received ads are stored in the `mobile_ad_delivery` table. Positive feedback conditions in the control loops can be checked by joining the recently changed location of mobile user to the locations of the locations of mobile ads that have been delivered within the last Δt period. Negative feedback conditions in the control loops are indicated by the ageing of delivered mobile ads, i.e. a tuple ⟨`pid,bid,prodid,delivery_time`⟩ in the `mobile_ad_delivery` table indicates a negative feedback condition for mobile user with ID `pid`, for a particular business with ID `bid`, or product or service with ID `prodid` if `delivery_time` + Δt is less than or equal to the current time. The same information can also be stored and managed in a similar fashion on the client side.

For both methods, both the server and client side approaches have advantages and disadvantages. The server side approach requires that personal behavioural data is stored on the server. This raises questions about scalability and privacy-related issues. In comparison, the client side approach requires a client application that manages the personal behavioural data on the client device. The client side approach seems to be more scalable and privacy-protecting, but in the case of loss or theft of the device, issues regarding the misuse of the sensitive personal information can arise.

5.9 An operational LBA database

The so far presented LBA database was developed for simulation and estimation purposes. As it is presented, it can be commercially used as a tool to forecast LBA exposure and penetration. However, the LBA database can easily be altered to support the online management of LBA. In an online operational setting it is assumed that the mobile units of the mobile users periodically, but not necessarily at regular time intervals, communicate their position to the server. In such an operational setting, the only necessary alteration to the LBA database is that instead of storing the historical trajectories of mobile users, the current locations of mobile users are stored. These locations can be represented as point geometries in the LBA database. Spatial queries to determine proximity between locations of mobile user and mobile ads can be implemented much in the same way using spatial joins. To manage mobile ads, periodically, relevant mobile ads are selected and delivered to mobile users who recently changed their location.

6 Proposal for a revenue model for LBA

A viable revenue model is a necessary prerequisite for successful, commercial LBA. There are essentially three parties involved in LBA:

1 the advertiser

2 the consumer (mobile user)

3 the LBA service provider or operator.

As in most other advertising media, and also in LBA, the advertiser pays for the majority of cost of advertising. These costs are for paying the other two parties for the participation in (consumer) and the facilitation of (operator) LBA. The incentive in doing so is clear: to increase the revenue of the business doing the advertising.

Most people do not like advertising. Some advertisements, like advertisements on billboards, they cannot escape. Some they are willing to endure in return of other services, newspapers and commercial TV. Finally, some, such as direct mail or commercial fliers, they may choose to opt-out from. Since, according to EU law, conducting LBA requires the informed consent of the consumer[1] (Cleff and Gidófalvi, 2007), the need for a clear consumer incentive is eminent. One way to motivate the consumer is to provide her/him with value-added services. One example of such a value added service could be a recently proposed Location-Aware Mobile Messenger that facilitates user-friendly communication and coordination between users (Cleff and Gidófalvi, 2007). Another possible value-added service can be the free delivery of non-commercial information, such as, e.g. information about traffic or weather. The consumer can also be financially motivated through electronic coupons or reward programmes.

The operator charges the advertiser for the services provided. The cost for these services can be determined based on

1 a flat rate per LBA campaign

2 the number of delivered mobile ads

3 the weighted number of delivered mobile ads taking into account the interests scores, or

4 the weighted number of delivered mobile ads taking into account the reactions of the mobile users to the received mobile ads, i.e. strong interest or indifference.

An accounting module for either one of the service cost schemes can easily be facilitated by the so far presented LBA database.

7 Experiments and results

Two sets of experiments (implicit and explicit interest case) were performed to measure the capacity of the MA channel under various maxdist and minscore settings. The estimation are based on

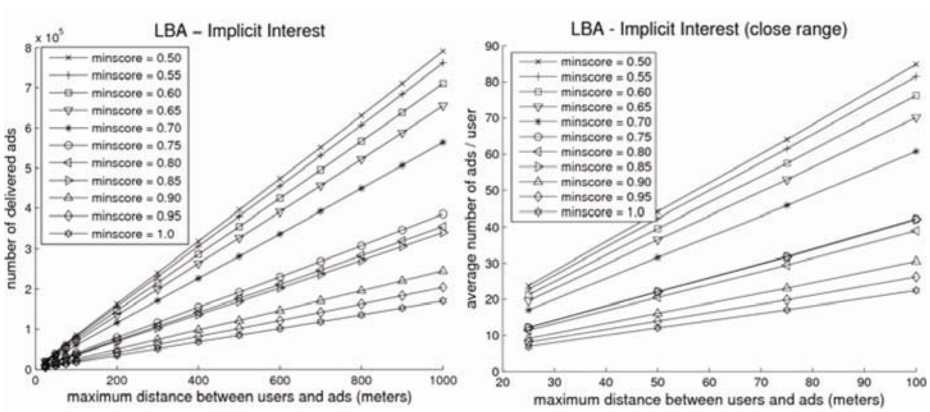
1 4,314 businesses in Copenhagen, Denmark offering one or many of the 40 hand-selected products or services

2 The simulated movements of 1,000 randomly selected simulated mobile users during the course of seven days (on average 3,800 trajectory segments per day).

Scores for implicit interests were modelled as described above. To simulate explicit interests, one product or service of interest was assigned to every simulated mobile user, as described above.

Figure 4(a) shows the number of delivered ads during the course of the first day in the implicit case. As expected, the number of delivered mobile ads increases as the minscore is decreased or the maxdist is increased. The rather surprising, close to linear relationship between the number of deliverable ads and the maximum distance criteria is due to the following facts. Simulated mobile users move from one location to another with the objective to perform an activity. These activities are tied to a subset of the businesses that advertise. Hence, the businesses that advertise the products or services often lie on the actual streets that the trajectories follow. If businesses are assumed to be uniformly distributed on those streets, then the relationship is indeed expected to be linear, as the number of businesses ‘reachable’ along a street grows linearly with maxdist. Another, rather interesting result is the sheer number of mobile ads that can be delivered to a small set of 1,000 users within a course of a day. Even for maxdist = 500 m (arguably a worthwhile detour for the mobile user) and minscore = 0.9 (quite high match in direct marketing) the average number of delivered ads to a user is about 100. This represents a huge marketing potential.

Figure 4 Number of delivered ads (implicit interest) to a population of 1,000 mobile users for various minscore and maxdist



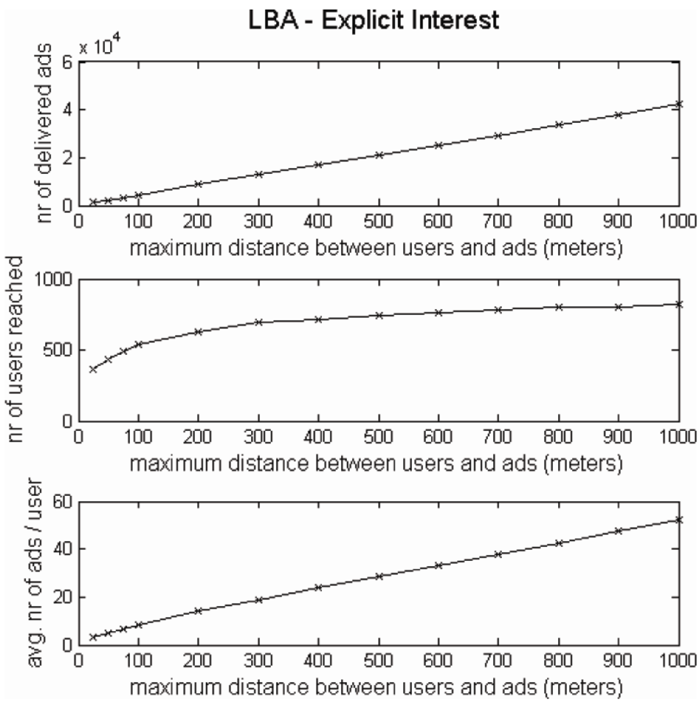
(a) Larger values for *maxdist*.

(b) Smaller values for *maxdist*.

The same numbers are likely to be viewed as alarming by many mobile users. As it is shown in Figure 4(b), even for rather high minimum interest scores for very low maxdist ranges the average number of ads delivered to a mobile user during the course of the first day is in the range of 6–40. This is a rather large number of ads to be received on a small, by many considered as extremely personal, mobile device. Hence, to avoid bad reputation, businesses interested in employing or facilitating MA should make great efforts to provide simple yet effective user-controls on the number of received mobile ads as suggested in Sections 5.6 and 5.7.

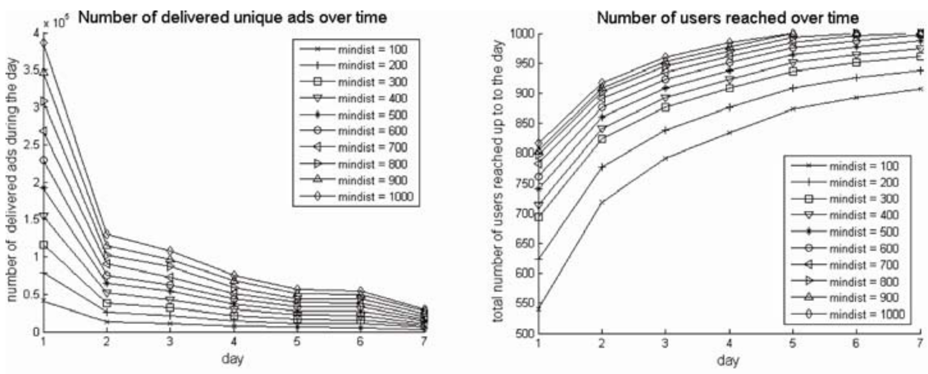
Figure 5 shows some statistics about the number of delivered mobile ads in the explicit interest case. Similar observations can be made about the relationship between the maxdist and minscore parameters and the number of delivered mobile ads as in the implicit case. However, it is surprising that even though every mobile user is only interested in exactly one of the 40 products or services, due to the presence of a large number of businesses offering those products and services, the number of deliverable ads is rather high even for small values of maxdist. It is also important to note, that for maxdist < 100, over half of the mobile users do not receive any mobile ads (middle graph). Hence mobile users, who are interested in getting good deals on products or services of their interests, have to set their maxdist values appropriately high.

Figure 5 Statistics about the delivered ads (explicit interest) for various maxdist



Since most mobile users need to perform mandatory activities at specific locations, such as going to work or coming home daily, and since they tend to perform their other activities either around, or along the way in-between, these specific locations, they tend to move around in approximately the same space from day to day. Hence, because of the uniqueness constraints on mobile ads, for a fixed set of static mobile ads, the number of deliverable mobile ads per day is expected to decrease over time. This decrease is shown in Figure 6(a) for the implicit case for an interest score of 0.75 for various mindist for a period of seven days. As it can be seen in Figure 6(a), the rate of decrease is also decreasing with time. This is because the number of infrequent and less regular (non-daily) destinations of an individual mobile user is limited, and over time these locations are eventually visited by the mobile user, at which time all the relevant mobile ads are delivered. Since the number of such irregular destination are limited, and the number of unvisited ones are decreasing as time progresses, the chance of a mobile user visiting an unvisited irregular destination is also decreasing with time. Consequently, the number of deliverable mobile ads per day also decreases with time. At the same time, because of the slight variation in day-to-day movements of mobile users, the penetration rates of LBA increase. More specifically, in the explicit interest case, Figure 6(b) shows that the number of reached users increases over time.

Figure 6 Effects of the uniqueness constraint on the number of delivered ads to a population of 1,000 mobile users over a period of seven days for various maxdist



(a) Number of deliverable ads by the day. (b) Number of reached mobile users over time.

User-defined ST constraints on mobile ads are used to disallow the delivery of mobile ads at specific locations and/or times. To test the reduction in deliverable mobile ads, in a set of experiments the delivery of mobile ads was not permitted when mobile users were either at home or at work[2] Somewhat counter intuitively, no reduction in deliverable mobile ads was observed. This result can be explained by the following two facts. First, in the LBA database movements of mobile users are represented as continuous trajectories. Secondly, the set of mobile ads used in the estimation were constant during the estimation period, i.e. all mobile ads were effective during the course of the whole simulation. Hence, after mobile users left their home for the first time, and travelled a short distance away from home, they received all the relevant mobile ads, and similarly they received all the relevant mobile ads as they for the first time approached their work place.

The effects of user-defined ST constraints on mobile ads would be more observable in LBA environments where mobile ads are dynamic and have short lifetimes. For example a cinema, after realising that over 90% of the seats are empty 30 min prior to the movie, might want to run a '50% off' LBA campaign for 30 min only. Such dynamic mobile ads with short lifetimes will be filtered out by user-defined ST constraints, if applicable. Extensions to the LBA database to handle such dynamic LBA conditions are trivial and are left for future work.

In summary, the experiments show that the capacity of the LBA channel is very high indeed, even for relatively small settings for minimum distance, and relatively specific interest settings. This is good news for LBA advertisers, as they can expect to reach a large set of potential customers.

8 Conclusions and future work

The aim of this paper was to investigate the capacity of the LBA channel. The paper proposed two types of LBA models (implicit vs. explicit interest) and described a relational database for the effective management of both types of LBA. Using a number of real-world data sources and simulated but realistic movement data of mobile users, the paper gave estimates on the number of deliverable mobile ads in both the implicit and explicit interest cases. Experimental results show that the capacity of the LBA channel is rather large implying a huge marketing potential. At the same time, the potentially large number of mobile ads could be alarming to mobile users, hence the paper warns businesses interested in LBA to provide the mobile users with adequate means to control the number of delivered ads and the time and place of delivery.

Future work is planned along two paths. First, while the presented LBA framework considers LBA both from the mobile users' and advertisers' perspective, the provided estimates are valid only if mobile users are willing to accept unlimited mobile ads at all times and places. Incorporating user-defined constraints on mobile ads, as described in Sections 5.6 and 5.7, will provide better estimates on the true audience size of LBA. Secondly, in the implicit interest case, the relevance of a mobile ad is estimated using a simple scoring model which is based on a consumer segmentation that divides users into 29 different consumer groups. However, as it is pointed out in Section 5.8, in real life, no two users' interests are exactly the same, hence a given mobile ad does not have the same relevance to them. Hence, altering the scoring model to include the personal interests of the individual mobile user, which are derived from historical behaviour of the mobile user – such as the type businesses the user has previously visited or the user's reactions to previously received mobile ads – will allow targeting the individual mobile user with even more relevant and personalised mobile ads. Since in the current simulation of mobile user movements, the possible influence of mobile ads on the future movements of mobile users is not accounted for, the evaluation of the effects of personal interest scoring are left for future research.

Acknowledgements

This work was supported in part by the Danish Ministry of Science, Technology, and Innovation under grant number 61480. The help from co-workers, Esben Taudorf, Kasper Rügge, and Lau Kingo Marcussen is also gratefully acknowledged.

References

Barnes, S.J. and Scornavacca, E. (2004) 'Mobile marketing: the role of permission and acceptance', *Int. J. Mobile Communications*, Vol. 2, pp.128–139.

Brinkhoff, T. (2002) 'A framework for generating network-based moving objects', *Geoinformatica*, Vol. 6, pp.153–180.

Cleff, E.B. and Gidófalvi, G. (2007) 'Legal aspects of a location-based mobile advertising platform', Accepted for publication in the Special Issue on "Identity, Privacy and New Technologies" of the International Journal of Intellectual Property Management (IJIPM) in 2008.

ESRI (2007) Business analyst. Available at: www.esri.com/businessanalyst.

Experian (2007) Available at: www.experiangroup.com.

Fox, P., Rezania, D., Wareham, J. and Christiaanse, E. (2006) 'Will mobiles dream of electric sheep? Expectations of the new generation of mobile users: misfits with practice and research', Paper presented in the Proceedings of the *ICMB*, p.44.

Gartenberg, J.M., Matiesanu, C. and Scevak, N. (2006) 'US mobile marketing forecast, 2006–2011', Paper presented in the Proceedings of the *Vision Report, Jupiter Research*, USA, October.

Geomatic aps – Centre of Geoinformatik (2007) Available at: www.geomatic.dk.

Gidófalvi, G. and Pedersen, T.B. (2006) 'ST-ACTS: a spatio-temporal activity simulator, Paper presented in the Proceedings of the *ACM–GIS*, pp.155–162.

Guttman, A. (1984) 'R-trees: a dynamic index structure for spatial searching', Paper presented in the Proceedings of the *SIGMOD*, pp.47–57.

Hägerstrand, T. (1975) 'Space, time and human conditions', in A. Karlqvist, Lundqvist, L. and Snickers, F. (Eds), *Dynamic Allocation of Urban Space*. Lexington: Saxon House Lexington Book, ISBN: 978–0347010528.

Heinonen, K. and Strandvik, T. (2007) 'Consumer responsiveness to mobile marketing', *Int. J. Mobile Communications*, Vol. 5, pp.603–617.

Jensen, C.S., Pfoser, D. and Theodoridis, Y. (2000) 'Novel approaches to the indexing of moving object trajectories', Paper presented in the Proceedings of the *VLDB*, pp.395–406.

Kaplan, M. (2006) 'Mobile advertising and marketing: market analysis and forecasts 2006–2011', Vision report, vision gain™, March.

Kölmel, B. (2002) 'Location based advertising', Paper presented in the Proceedings of the *ICMB*.

Komulainen, H., Ristola, A. and Still, J. (2006) 'Mobile advertising in the eyes of retailers and consumers – empirical evidence from a real-life experiment', Paper presented in the Proceedings of the *ICMB*, p.37.

Lee, T. and Jun, J. (2007) 'The role of contextual marketing offer in mobile commerce acceptance: comparison between mobile commerce users and nonusers', *Int. J. of Mobile Communications*, Vol. 5, pp.339–356.

Leppaniemi, M. and Karjaluoto, H. (2005) 'Factors influencing consumers' willingness to accept mobile advertising: a conceptual model', *Int. J. of Mobile Communications*, Vol. 3, pp.197–213.

Merisavo, M., Vesanen, J., Arponen, A., Kajalo, S. and Raulas, M. (2006) 'The effectiveness of targeted mobile advertising in selling mobile services: an empirical study', *Int. J. of Mobile Communications*, Vol. 4, pp.119–127.

Mobile Marketing Association (2007) Available at: www.mmaglobal.com.

Oracle Spatial and Oracle Locator: location features for oracle database 10g (2007) Available at: www.oracle.com/technology/products/spatial/.

Pousttchi, K. and Wiedemann, D.G. (2006) 'A contribution to theory building for mobile marketing: categorizing mobile marketing campaigns through case study research', Paper presented in the Proceedings of the *ICMB*, p.1.

Statistics Denmark (2007) Available at: www.dst.dk.

The Gallup Organization (2007) Available at: www.gallup.com.

Notes

[1]Directive on privacy and electronic communications (2002/58/EC, article 13(1)) involves asking the users' permission to send unsolicited advertising messages via all electronic communications for marketing purposes. Most countries outside of the EU also enforce similar legislative regulations.

[2]Home and work places for mobile users have been identified by slightly altering the output of ST-ACTS.