A Study on The Value of Location Privacy

Dan Cvrcek
Faculty of Informatics, Masaryk University
Botanicka 68a
602 00 Brno, Czech Republic
Cvrcek@fi.muni.cz

Marek Kumpost
Faculty of Informatics, Masaryk University
Botanicka 68a
602 00 Brno, Czech Republic
xkumpost@fi.muni.cz

Vashek Matyas
Faculty of Informatics, Masaryk University
Botanicka 68a
602 00 Brno, Czech Republic
Matyas@fi.muni.cz

George Danezis
K.U. Leuven, ESAT/COSIC
Kasteelpark Arenberg 10
B-3001 Leuven-Heverlee, Belgium
George.Danezis@esat.kuleuven.be

ABSTRACT

This paper introduces results of a study into the value of location privacy for individuals using mobile devices. We questioned a sample of over 1200 people from five EU countries, and used tools from experimental psychology and economics to extract from them the value they attach to their location data. We compare this value across national groups, gender and technical awareness, but also the perceived difference between academic use and commercial exploitation. We provide some analysis of the self-selection bias of such a study, and look further at the valuation of location data over time using data from another experiment.

Categories and Subject Descriptors

K.6.5 [**Management of Computing and Information Systems**]: Security and Protection; K.4.1 [**Computers and Society**]: Public Policy Issues—*Privacy*

General Terms

Security, experimentation.

Keywords

Economics, experiment, location privacy, mobile phone, mobile service.

1. INTRODUCTION

As technology increasingly permeates society we hear more and more about the need for privacy. Privacy can be partly modelled as confidentiality, or controlled disclosure, of some personal information. Yet people seem to be prepared to disclose some of their personal data for a very modest reward. A typical example is the popularity of loyalty cards,

allowing superstore chains to link data of particular customers regardless of the store used for shopping. People often do not realise what the data collected through royalty cards can be used for but their perceptions can change very quickly following an advertised privacy breach as we show later in this paper.

Privacy requirements are very hard to satisfy in information systems. One can implement access control mechanisms in trusted systems, but once the data leaks it is very hard to track their usage. There are two very different approaches to privacy problems. The first one uses the legal system to punish unlawful breaches of privacy. The second approach is to implement technical measures to guard the privacy of data stored and processed in information systems. The latter takes into account that law enforcement might be too slow or ineffective (due to the complexity of information technology) to enforce privacy requirements.

However, the technology needed to preserve privacy is very expensive. This is reflected in computational complexity, communication overheads and delays. Despite declaring some sensitivity to their personal information, users are often not prepared to accept the overhead or cost of privacy enabled technologies. The market failures of flagship products like the Freedom network [1, 2], an anonymous web browsing solution, illustrate this.

Yet more and more privacy intrusive technologies are deployed and become widely adopted. The GSM mobile network, to choose one that is considered indispensable in everyday life, allows for the tracking of powered-on handsets through the operator BTSs (Base Transceiver Stations). This allows real-time tracking of mobile phone devices with a granularity of a few hundred metres within cities, to a few kilometres in less populated areas. One may argue that the GSM network "requires tracking" so that voice communications can be efficiently routed. However triangulation based systems exist and are deployed to improve location accuracy and to determine the position of GSM handsets as precisely as possible. Such information can be recorded by mobile operators, and used for network management and law enforcement operations. Such tracking can also be offered by mobile operators as a charged service – parents may use it to locate their children mobile phones or employers their staff. Third generation GSM will provide location information at an even finer granularity.

It seems clear that the tools of computer security alone cannot give us the full picture of people's attitudes to privacy, and for this reason we turn to experimental psychology and economics to establish the "value" that individuals attach to their privacy. In particular, we shall generalise the study by Danezis, Lewis, Anderson [3] on the value of location information - and perform it on a wider and more varied population and across multiple E.U. countries. Location is quite a relevant aspect as mobile phones are ubiquitous and can be used to eavesdrop on users' movements. The original location privacy study [3] was done in a relatively small scale at Cambridge University. By measuring the same aspects of privacy, we can compare our results, and establish whether people's attitudes to privacy are uniform across the EU.

2. DESIGNING THE STUDY

Out of all possible aspects of personal data privacy we chose to focus our study on location privacy. Location is an important aspect as mobile phones can be used to eavesdrop on users' movements. A previous study by George Danezis, Stephen Lewis, and Ross Anderson also looked at location privacy through a relatively small study at the University of Cambridge [3].

Some studies about people's attitudes to privacy have already been published [3, 4], and experimentally pinpointing the value people attach to privacy is a difficult problem. We chose, in the tradition of the previous study [3], to use an auction, where people are required in effect to sell their private information. With an auction, participants have incentives to communicate to us the true value they attach to their privacy: bidding too high may exclude them from the study and its rewards, while selling at a lower price than the minimum they are expecting would provide too much discomfort for them to participate.

The key ideas behind the auction methodology is to create conditions that would push upper and lower bounds of bids (sums for which participants agree to be tracked) as close as possible. The lower bound is pushed up by the participants' endeavour to get as much money as possible while the upper bound is squeezed by a possibility to be left out in the auction. The sum the selected bidders obtain is equal to the amount of the lowest not accepted bid – this may bring in some bias but this is unavoidable.

It has also been shown that individuals tend to overemphasise their privacy sensitiveness if asked directly e.g. in a sociological study [5]. As a result we may obtain data about their behaviour that is generally not matched by their actions. For this reason we *deceive* the participants of our study, and make them believe that they are applying to take part in a fictitious study of mobile phone usage. The fictitious study was to take about a month and the participants had to be tracked during the period around the clock in five minute intervals. We further told the participants that as there was limited amount of money for the study we would use an auction to select participants. The lure of real financial returns, and the action structure of the study, gives all participants incentives to state their real privacy valuation on the disclosed location data.

Our study was conducted in the context of, and with the help of, the FIDIS[1] network. Beside the fact that partners in different European countries have been instrumental in allowing us to gather data across Europe we created an image of a potentially large study. This implied very hard assessment (by participants) of the total amount of money available for the study and thus very good initial conditions of this "one-run game".

3. DECEPTION AND EXTENSIONS

As argued above, our study required a small amount of deception. We were interested in the monetary value people would attach to a month of location data, yet we pretended that we were looking for volunteers for a study of mobile phone usage. This approach was deemed necessary to reduce the bias of the resulting data. Despite deception being a common tactic in experimental psychology, it is not in computer science or security research, and our actions had to be cleared by ethics or management boards in all participating institutions. All deceived participants were also notified of the real purpose of the experiments and of the preliminary results after the end of the data collection phase.

The study was implemented using web forms containing a questionnaire. We advertised it using emails and posters addressed to university audience to involve as many people as possible. The text of emails also leaked to readers of a large mobile phone webserver in the Czech Republic thanks to Czech students contributing to its content. The information text we used can be found in Appendix A. A similar introductory text was sent out to university students in five countries (translated into the local languages).

The original study by Anderson et al. [3] that pioneered this approach was limited by several factors. Firstly, the population on which it was carried out had very specific characteristics: computer science undergraduates at the University of Cambridge. These were mostly male, highly educated and technically aware. Although some conclusions were drawn about the correlation between privacy sensitivity and general use of the mobile phone, as well as patterns of travel, it was very difficult to assess correlations with gender or nationality. Secondly, no serious analysis was performed on the "self-selection" bias of the participants: those with high privacy sensitivity might have selected themselves out of the experiment, and not participated exactly for this reason. This could have lead to a constant bias in the study, underestimating the value people attach to their private location data.

Our study attempts to remedy some of these shortcomings. We performed the same study *in five European countries, in which FIDIS institutions are based.* It allowed us to compare the results across different nationalities. Due to relatively large number of respondents, not only the median value attached to the data can be compared, insights into the privacy sensitivity of local populations are also provided by the overall distributions of bids (privacy valuations).

A second considerable improvement over the original experiment was the use of a more gender balanced population. Replies from such a population allow us to draw conclusions about a possible gender bias between men and women when it comes to the sensitivity of location information. Although the majority of respondents were males, we managed

---

to attract a sufficient number of female participants to compare opinions of sexes against each other. The same holds for less technically savvy population, and so we are able to draw better conclusions about the sensitivity of an average individual.

4. IMPLEMENTING THE STUDY

This section introduces the sets of questions used for the study questionnaire. The questionnaires were implemented as a web application allowing for an easy access of participants. Access to the forms of the questionnaire themselves was authenticated. The questionnaire was structured into four logical parts. The first part contained a longer version of the introductory email (very much similar to subsection 3.3 in [3]). We put a question about the respondent's interest in the study at the end of the text. The question had three options for the answer a potential participant would choose from:

1. I do not have a mobile phone.
2. I do have a mobile but I am not interested to participate.
3. I do have a mobile phone and I am interested to participate.

Those selecting option 3 for the answer were presented with a request for their e-mail address. An email was then sent to them with a login name and a password for further communication through the web interface. After a successful enrollment and subsequent authentication, the following questions (with predefined options for answers 1-6 as pull-down menus) would be presented:

- Is your background (area of study)?
  a) Computer Science, Comp./El./Comm. Engineering, Informatics, b) Law, c) Other areas.
- What is your gender?
- What network do you use for your main mobile phone? <list of local operators>.
- Do you carry a mobile phone with you most of the time?
  a) yes, b) no.
- How often do you make irregular movements (such as shopping, going out with a friend, pub, visiting friends)?
  a) Several times a day, b) every day, c) every week, d) every month.
- With whom do you communicate using your phone?
  a) Friends, b) family, c) partner, d) business.
- How much compensation would you require to participate in our study for 30 days (in whole amounts of local currency)?
- Text array for free comments.

This was the main part of the questionnaire providing the most important information we needed. Note that we included questions to support the cover story but kept their number small not to distract the participants.

One of our objectives was to extract the "valuations" for different kinds of data usage. For this reason we used two additional forms. The second form (a very short one) was confronting the subject with a possible change in the use of the collected data: from academic to commercial exploitation. The text stated that "there is now some possibility of commercial interest (from partners of your mobile phone operator) in the data collected during our study. We would be grateful if you could let us know whether you would"

- not be willing to participate in the study if the data might be used by a regular business partner of your mobile phone operator;
- allow your data to be used by a regular business partner of your mobile phone operator for the same amount of compensation you originally bid;
- allow your data to be used by a regular business partner of your mobile phone operator only if you were allowed to revise your bid for compensation.

The second form may have benefited from being presented to subjects a few days after the first one. This would allow the subject to "forget" their valuation, rather than facing a sharp contrast between "academic" use, and "commercial" use. This could have been done using the collected email addresses. We conjecture that the results would not have been significantly different, yet confirming this experimentally should be the subject of a further study. We believe, however, that this kept more participants in the following rounds of the auction.

The final form further modified the parameters of the study. It stated that "a business partner of your mobile phone operator inquired about the option to extend the experiment period to 12 months in total. Would you"

- decline your participation in the experiment;
- participate for the following amount of money <enter the sum>.

In fact, we exposed participants to, first, a qualitative change in the use of the data, from academic to commercial, and then to a quantitative change from 1 month to 12 months. As we shall see the latter provided very interesting insights on how people value data collected over different time intervals.

5. SOURCE OF INTEREST IN THE STUDY

When we started analysing the obtained data immediately after closing the web questionnaire, we realised from the preliminary results that there is a real danger of the results deterioration due to participants' enthusiasm for the publicly announced goals of the study. The feedback we received from the participants was very often in the line of actual interest in the goals and willingness to take part in the project just to improve quality of the mobile network services.

We therefore promptly prepared a webpage with first results of the study and put the link into the emails sent out with a description of the real purpose of the study. We wanted to reduce expected negative reactions (few appeared anyway) and to prepare ground for the last question we prepared – "Why did you take part in the experiment". There were three possible answers that respondents could assign a relative weight. The options were: fun, money, and results.

We received over three hundred responses, which makes about 30 % of the participants. The analysis of this data

country	total	women
Belgium	37	3
Czech Republic	744	131
Germany	251	33
Greece	30	6
Slovak Republic	152	46

Table 1: Numbers of participants per country.

calling	friend	family	partner	business
Number	1076	975	598	358

Table 2: Who is being called by mobile phones.

Language	BE	CZ	DE	GR	SK
Not interested	12 %	6 %	12 %	25 %	12 %

Table 3: People not interested in the study.

shows that no substantial distortion of the results appeared. See section 6.5 for details.

6.   RESULTS OF THE STUDY

The results we obtained can be split into several parts. The first part (6.1, 6.2) contains data related to the organisation of the study. Although it may be seen as truly introductory, it contains some interesting insights into participants' behaviour. The second part (6.3, 6.4) encompasses results we were primarily interested in. The third, last part, (6.5) contains a further analysis of data correlating results to reasons why people took part in the study.

6.1   Demographics

This part of the results describes basic demographic data, as well as several details about the data collection. Our web questionnaire was up for about a month. We received the majority of responses during the 48 hours after emails were sent out, except for the Slovak Republic where the information was spread with paper notices. As a result, responses there were uniformly spread throughout the month of the experiment.

About 1200 participants answered the first set of questions. These were from five countries: Belgium, Czech Republic, Germany, Greece, and Slovak Republic – mainly representing people from the Central Europe but Southern and Western Europe was also represented. The split of the participant numbers according to their country of origin and gender is in Table 1. The sets of the Czech Republic, Germany, and the Slovak Republic are large enough to allow for very detailed structuring of results. Especially the size of the Czech sample is large enough to withstand the requirements used by many sociological surveys. The smaller sets of Belgium and Greece are used more just like "control sets" to verify general results.

The set of the participants can be also split according to their background. We obtained answers from 800 people with a primary background related to computers, 32 "lawyers", and 381 with another (unspecified) background. The set of respondents with the background in law is very small so we can use it again just as a control set confirming other results. We believe that participants with computer related background are most likely to introduce bias into data and that is why we wanted to separate them. The category "law" was introduced at a suggestion of a law department that took part in organising the study. All but ten participants carry their mobile phone all the time.

The second interesting demographic aspect concerns the frequency of irregular movements. The study of Danezis, Lewis, and Anderson used a bit more vague question in terms of frequency and destination, they were basically interested only whether the participants travelled outside Cambridge. In our case, we tried to split people according to

frequency of any irregular movements like shopping, going to a pub, visiting friends, meeting boyfriend/girlfriend, and so on. We also wanted to use this way to make people realise possible implications of being tracked for a certain period of time. The obtained results show that 483 people do such movements several times a day, 520 about once a day, 194 in a week intervals and only 15 monthly. As you can see, we have obtained three very large sets suitable for comparisons and analysis.

There was another "cover up" question in the forms – whom do you call most often. The participants could select one or more options from the set of friend, family, partner, business. As you can see in table 2, mobile phones are primarily used to communicate with friends and family. Much lower numbers are for partner and business options.

6.2   Cautiousness

The first question to be answered by potential participants was whether the respondents were interested in the study at all. From the three choices offered, we may ignore the possibility that the given person does not posses a mobile phone as only a few have selected this. We call those not interested in the study as "early drop-outs". We are now working with a sample of 2582 people with 239 early drop-outs. There were 11 Belgians, 85 Czechs, 65 Germans, 32 Greeks, and 46 Slovaks in absolute numbers. However, much more interesting are the relative results. The shares of early drop-outs on the sample sets are in table 3.

It is obvious there are two irregularities. Assuming that the sets of Czechs, Germans, and Slovaks are satisfactory (and statistical errors should be therefore well under 5 %) the enthusiasm of Czechs is surprising. We can only guess what reasons lies behind this number. We know that there was an extensive discussion about the study on student mailing lists praising the pretend goal of the study. We also heard rumours that students were discussing the issue among themselves – whether to take part or not. One of the possible reasons might therefore be that those opening the webpage have already made up their minds.

The second irregularity comes from a small sample of Greek respondents. More than half of them read the information and decided not to take part! These results are somewhat confirmed by the "standard drop-outs". Those drop-outs come from the fact that the potential participants had to enter an email address we sent an authentication data to. They were supposed to use the data in order to eventually get access to the web forms. The results are in Table 4. These results may come from different sources such as national characteristics. Belgians and Slovaks are very similar. Czechs differ from Germans very slightly – the difference is close to possible statistical error. The Greeks are again "the most privacy cautious" if we wish to use this hyperbole.
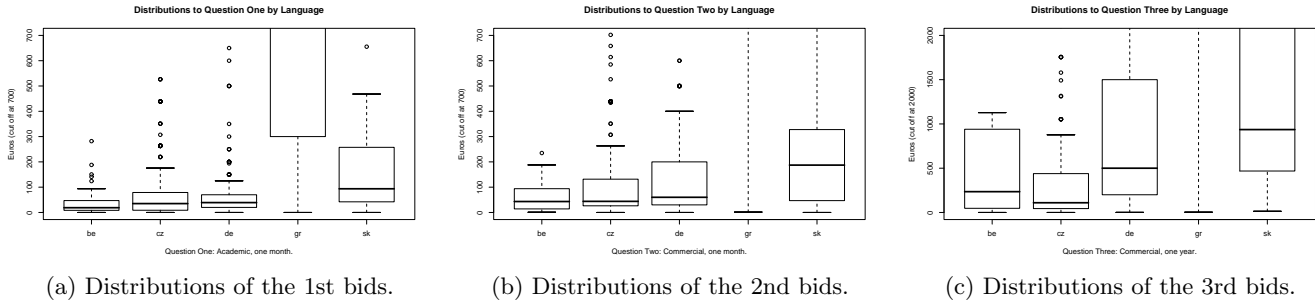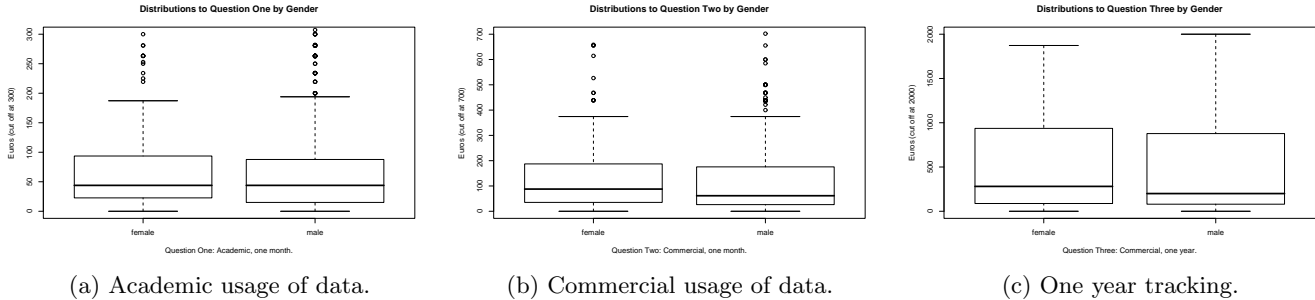
(a) Distributions of the 1st bids.   (b) Distributions of the 2nd bids.   (c) Distributions of the 3rd bids.

Figure 1: Distributions of bid values per country.

(a) Academic usage of data.   (b) Commercial usage of data.   (c) One year tracking.

Figure 2: Distributions of bid values per gender.

Language	BE	CZ	DE	GR	SK
Interested	44 %	56 %	52 %	32 %	42 %

Table 4: Fractions of respondents interested and eventually enrolling.

6.3 The Main Results

We are finally getting to the main results – the price of privacy. The countries we gathered data from have different currencies. There are Euros in Germany, Belgium, and Greece, and local currencies in the Czech Republic (Czech Crowns) and the Slovak Republic (Slovak Crowns). The plots in this section depict all values in Euros. We have also re-calculated the values of bids in different countries by using a "value of money" coefficient computed as a ratio of average salaries and price levels in particular countries – these data were taken from Eurostat statistics.

We have gathered bids for three auctions or scenarios. The first and second bids were for one-month tracking. The former data were to be used for academic purposes only, and the latter for commercial purposes. The third bids were for the scenario where participants agreed with a year long tracking and data free for commercial exploitation. Let us start with the first bids.

6.3.1 Differences among Countries

The distributions of the first bids are on the plot at fig. 1(a). You can see that when we mentioned Greeks as the most cautiousness nation, it was not for the last time. Although there are differences between all nations, the Greek bids are beyond our expectations. Figure 1(c) shows the same division of participants but the values of bids come from the third auction. The Czechs bid again lowest, while the me-

dian of the Greek bids went through the roof. The problem with Greek results, however, is the low number of respondents. We believe there is still some significance in the results, because of the immense difference from bids in other countries (the probability of this happening incidentally is low even for such a small sample set), but this needs to be confirmed with more representative data!

6.3.2 Men and Women

Our forms had few questions but they still allow for quite interesting analysis – especially of the first scenario bids. One of the questions we were particularly interested were possible differences between sexes. Figure 2(a) indicates that it is not the case, but fig. 2(c) tells a different story. When the participants were offered the possibility to increase the bid for a year-long study, males behaved differently from women. Although it is hard to see in the plots the medians of the first bids are equal, medians of the second bids are in ratio 1.4:1, and medians of the third bids are 1.8:1, with women bidding higher (numbers of women and men stayed at a roughly constant ratio). These results indicate that women are possibly more sensitive to what the collected data may be used for.

6.3.3 Mobility

While we did not expect substantially different behaviour between men and women, we definitely expected it with respect to the frequency of movements. Our expectations encompassed a correlation between the bid values and frequency of movements. What really happens when we split the bids according to the frequency of movements is shown in fig. 3. If we were talking only in terms of quartiles, there would be no difference that would be above a possible statistical error. What is visually different are particular bids
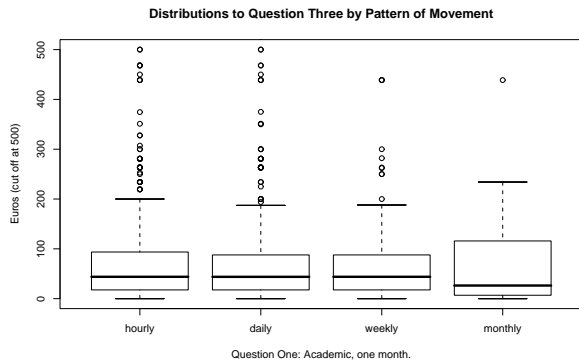
Distributions to Question Three by Pattern of Movement

Figure 3: **Correlation between distributions of bid values and frequency of movements.**

First Question Bid Distributions per Population
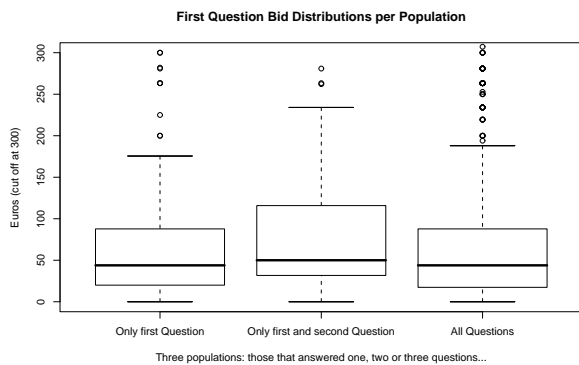
Figure 4: **Bid distributions for the first scenario per people cautiousness.**

Second Question Bid Distributions per Population

Figure 5: **Distributions of the 2nd bids of those who bid in the first two, and all three scenarios.**

above the third quartile. However, this is more likely to come from the differences in the sample sets size (daily 520, hourly 485, weekly 195, monthly 15). Despite general expectations, no correlation between the bids and frequency of movements was proved. Even though there might have been some misunderstanding of this question, it seems to us that the participants did not perceive their unusual movements as more sensitive than their every-day behaviour.

6.3.4 Impact of Scenarios

Another interesting comparison of the first bids is when reflecting behaviour of bidders in the second and the third auctions. Figure 4 shows differences of the first bids when the sample set was divided according to whether the participants kept bidding in the second and the third rounds. We believe that the results come from a combination of initial curiosity and privacy cautiousness. Although the differences are not conclusive, they are similar for all three quartiles. We may therefore entertain a bit of speculation here. The bidders who were really considering the difference between the scenarios bid higher and did not accept the conditions of the year long study – raw data confirm this as the median of those more "privacy aware" is about 20 % higher. There may be also a reason for the 1st bids of those who
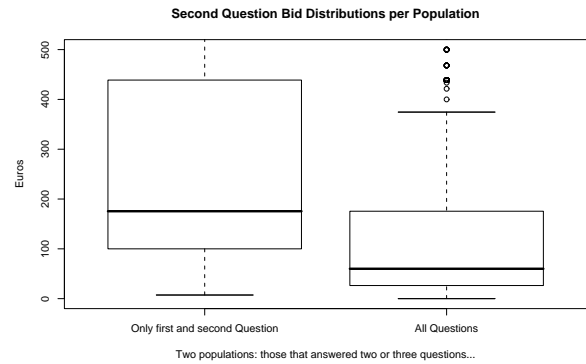
kept bidding till the last auction being lower than those of the former group. The people who put only the first bid did it only for curiosity and just did not want their data to be used for commercial purposes (the median of participants' bids does not change with the decision in the second auction round).

An important part of our study was an assessment of people's sensitivity to commercial exploitation of personal data. The second form was requesting respondents to submit a bid compensating their participation in the experiment involving a commercial partner, and were also given a chance to opt-out. The following table (tab. 5) shows divisions of participants according to their answer for the 2nd bid.

Table 5 indicates that participants from the two Central European countries (CZ and SK) were declining participation less often than respondents from other countries, whatever the reasons may have caused this. The figures 6(a) and 6(b) show the comparisons of the first bids according to how the participants reacted to the offers of the second and third round of the experiment.

These two graphs represent a bit contradicting evidence. The plot on fig. 6(b) quite clearly shows a correlation between answers in the last auction round and value of the first bids, and this is well beyond statistical inaccuracy. On the contrary, the left-hand side figure 6(a) does not demonstrate any such dependence. We think that the explanation of this phenomenon is not trivial and results of fig. 6(b) are quite surprising.

We have already presented one plot related to the third bid. It is however very interesting how the bids changed between the auctions. The two graphs of fig. 7(a) and fig. 7(b) show bid distributions of people who entered values for all three possible uses of data – they depict the same data in two different forms. On the left hand graph, the x-axis shows value of bids in EUR and the y-axis shows the fraction of bidders who entered bids of value lower or equal to a given amount.

One can see that the median of the bids increased about twofold when compared to the bids when the data were to be used not only for academic purposes but also for commercial purposes. The extension of the study period from one month to a year yielded another twofold increase in the median bids. This is a clear indication that participants were

Language	BE	CZ		DE		GR	SK	
Sex	M	M	F	M	F	M %	M %	F %
Declined	16 %	10 %	13 %	21 %	23 %	25 %	9 %	3 %
Same bid	39 %	47 %	42 %	51 %	45 %	50 %	49 %	40 %
Revised bid	45 %	42 %	44 %	28 %	32 %	25 %	42 %	57 %

Table 5: Distribution of types of answers to the change of data usage.
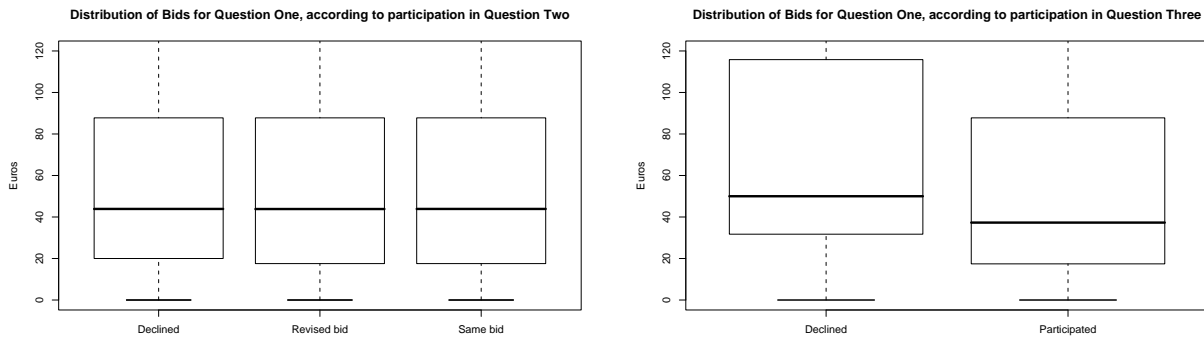
(a) Answers in the second scenario.

(b) Answers in the third scenario.

Figure 6: Distributions of the 1st bid values per behaviour in later rounds.
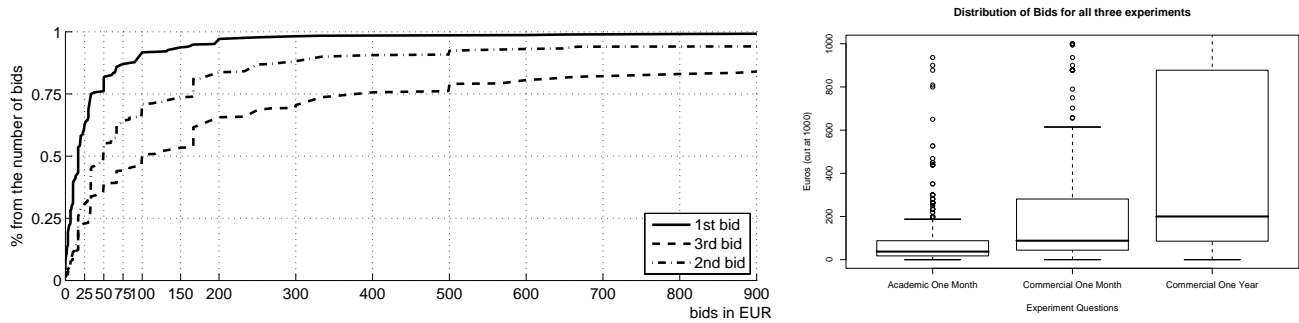

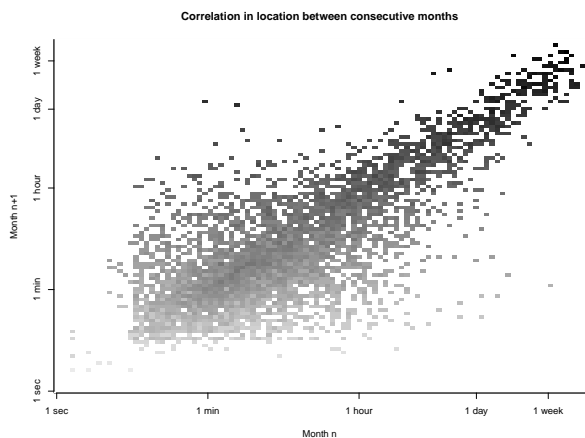Figure 7: Distribution of bids for all three study scenarios.

Figure 8: Correlation between time spent in the same cell for two consecutive months.

more sensitive to the *purpose* of the data collection, than the *duration and quantity* of data collected (the period has been increased from one to twelve months). Interestingly, there are huge differences among countries in the sensitivity to the time extension. Medians of the bids increased by 20 % with Czechs, 250 % with Belgians, and fivefold with Germans and Slovaks.

6.4 Exploring the Non-linearity in Time

As explained in the previous section, participants were presented with an option to take part in a commercial study. They were asked to provide valuations for studies of one month and twelve months durations. Perhaps surprisingly, the valuations were not linear in time – it is not the case that the amount sought in the second case was twelve time the amount in the first case. Instead we only observed a modest increase of a factor of two on average (five in two countries).

We see two possible explanations for this phenomenon. Firstly, we could explain it through hyperbolic discounting – people tend to value most things that are close in time, and often irrationally undervalue options that are farther away into the future. Studies of this phenomenon and privacy are presented in [6]. Yet the non-linearity is substantial, and only an extreme form of hyperbolic discounting behaviour could explain it.

A second hypothesis is that the value of the data extracted in months after the first one, is indeed of less value to the participant. We chose to test this hypothesis by using data from a real mobile phone usage study performed at MIT in the context of the Reality Mining project [7]. In this study, about a hundred participants were provided with mobile phones recording all their interactions with their phone, and the mobile cell they were in. This is exactly the type of data that one would expect to extract from our fictitious cover study.

Our thesis is that data after the first month is of less value to the participant, and also to third parties collecting it. The only reason we can see for this is that additional data provides very little new information to an observer about a participant's location, and therefore little additional privacy

infringement for the observed party. We analysed the reality mining data from two arbitrary months (month 9 and 10) from the reality mining project to test this. We extract and plot in fig. 8 the amount of time spent by each person in a particular cell for both months. The cells are aggregated by time spent in them, and the intensity of the graph indicated time spent in the cell in the second month.

We found that the correlation between the location of users in one month and the next one is striking. When a participant spends more than an hour in a cell at month n they are very likely to also spend a similar amount of time in the same cell at month $n + 1$. In case one spends more than one day in a cell, she is almost certainly going to spend a similar amount of time in the same cell the month after. These locations are likely to correspond to the accommodation used on a daily basis, and the place of work or study of the participants. These are indeed not likely to change for most people from one month to another. On the other hand, there is only a very weak correlation for cells in which the user spends less than 1 hour per month.

The analysis of this data seems to provide evidence to support the second hypothesis: participants are not irrational to value less subsequent months of surveillance, after the first one. An observer gets a lot of information at the start of the observation period, such as their usual movement pattern. Subsequent months add very little information, and can therefore be seen as less valuable both from the point of view of the observer, and the person observed.

6.5 Why People Took Part in The Study?

As the participants were mostly students (with possible exception in the Slovak Republic), we worried about possible biasing of the data from bids of students who took part only for fun or curiosity. For this reason we asked the participants directly, what their interest in the study was, after the data collection phase was over.

We received about 300 responses (quarter of the study participants) giving us importance of three factors that could potentially attract them to take part in the study. These were a) just for fun, b) curiosity for the results, c) financial benefits.

A simple comparison of reasons' weights for the decision whether to take part or not shows that all three reasons were in a rough balance: 38 % for money, 32 % for results and 29.7 % for fun. Can anything be deduced from this result? We think that the monetary value of bids we are talking about is only part of the reward people expected from the study. The open question may be how much may a good public-relation campaign reduce the financial rewards expected by people if an organisation was trying to collect some sort of personal data.

The interesting aspect is whether the types of interest correlate with values of bids. The graphs 9(a), 9(b) depict this aspect. One can see that the medians do not differ significantly. We can also look into the differences of "correlated interests" with respect to answers to the second auction – see fig. 10.

7. CONCLUSIONS

The paper summarises results of the study we have conducted primarily among university students. We used a cover story to promote the study and hide the real goals behind "a research into topology of mobile networks consid-
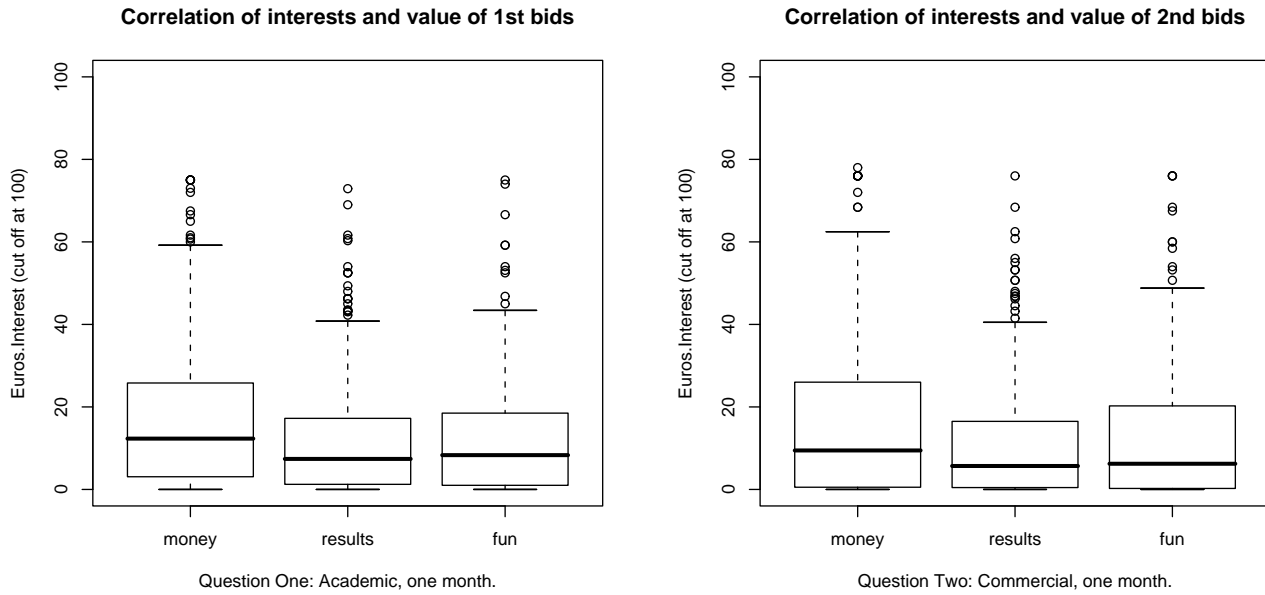
Correlation of interests and value of 1st bids

Euros.Interest (cut off at 100)

money   results   fun

Question One: Academic, one month.

Correlation of interests and value of 2nd bids

Euros.Interest (cut off at 100)

money   results   fun

Question Two: Commercial, one month.

Figure 9: Distribution of bids for all three study scenarios.

Correlation of interests – same bid

Euros.Interest (cut off at 100)

money   results   fun

Question One: Academic, one month.

Correlation of interests – revised bid

Euros.Interest (cut off at 100)

money   results   fun

Question One: Academic, one month.

Correlation of interests – declined

Euros.Interest (cut off at 100)

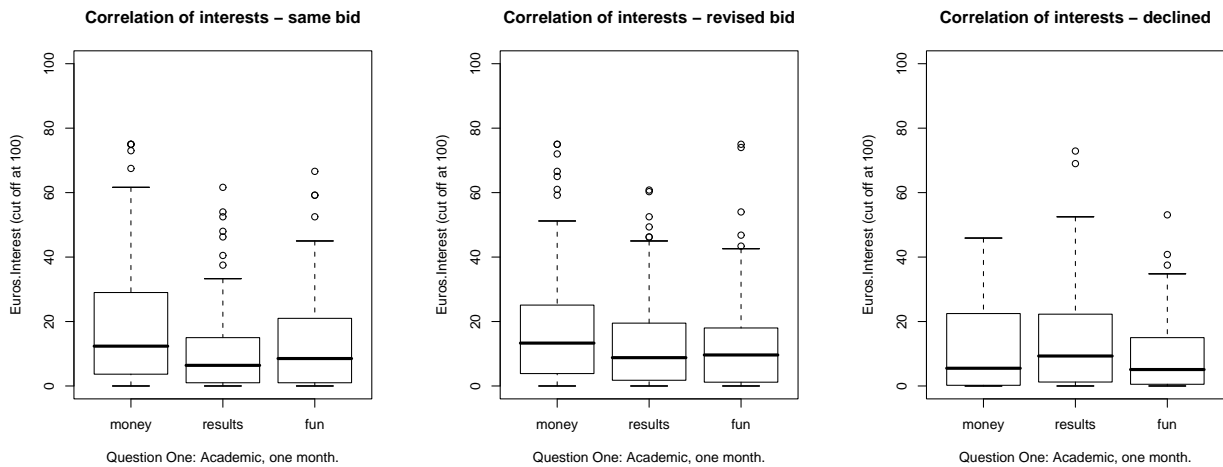money   results   fun

Question One: Academic, one month.

Figure 10: Correlation of interests and value of bids with respect do 2nd auction.

ering mobility of customers". We want to stress that the cover story did not substantially bias the results due to attractiveness of the "fake" study goals. There are differences between bids with respect to the reasons why participants took part in the study, but these differences are not fundamental.

There were about ten percents of participants bidding below 1 EUR. We believe that this fact comes from curiosity and enthusiasm of participants. We received a lot of feedback expressing interest in the results when the deception text (stating the improvement of the network quality as the main goal of the study) was sent out.

The highlight of the study could have been the evidence of Greek sensitivity to possible privacy breaches if a larger sample set was collected. Still, the findings (drop-out rates, values of bids) are worth mentioning and they deserve a follow-up study to be confirmed. The reason why we see so much importance in these results is a possible impact of an eavesdropping scandal [8]. Top Greek politicians were being wiretapped for a period of eleven months during and after 2004 Olympic Games as was confirmed at the beginning of February 2006 by the Greek government – just two months before this study actually took place.

Another highlight may be seen in the nonlinearity of bids in regard of the study length. A possible explanation was introduced in section 6.4, yet this may be an interesting problem for a further study.

Surprisingly, we have not found any correlation between valuations and the way the respondents commute or move in geographic terms. The results do not show any such relation. This result contradicts the Cambridge study [3].

A slightly different situation is in respect of gender. Bids differ notably only in medians and this difference increases with the number of auction rounds. Medians of women bids are higher with maximum 80 % difference in the third round.

Basic results confirm results of the Cambridge study in the overall value of bids – e.g. medians of bids are 20 GBP and 43 EUR (i.e. about 28 GBP at the August 2006 exchange rates) for non-commercial use of data, respectively.

Acknowledgments

8. REFERENCES

[1] Boucher, P., Shostack, A., Goldberg, I.: Freedom system 2.0 architecture. whitepaper, Zero-Knowledge Systems, Inc. (2000)

[2] Law, G.: Anonymity declines as zero-knowledge ends web service. PC World (2001)

[3] Danezis, G., Lewis, S., Anderson, R.: How much is location privacy worth? In: Fourth Workshop on the Economics of Information Security. (2005)

[4] Hann, I.H., Hui, K.L., Lee, T.S., Png, I.: The value of online information privacy: Evidence from the usa and singapore. In: International Conference on Information Systems. (2002)

[5] Acquisti, A., Grossklags, J.: Privacy and rationality in individual decision making. IEEE Security & Privacy **3**(1) (2005) 26–33

[6] Acquisti, A.: Privacy in electronic commerce and the economics of immediate gratification. In Breese, J.S., Feigenbaum, J., Seltzer, M.I., eds.: ACM Conference on Electronic Commerce, ACM (2004) 21–29

[7] Eagle, N.: Machine Perception and Learning of Complex Social Systems. PhD thesis, Massachusetts Institute of Technology (2005)

[8] Danezis, G.: Government communication illegally wiretapped in Greece. EDRI-gram `http://www.edri.org/edrigram/` (2006)

APPENDIX

A. INTRODUCTORY LETTER

Dear reader,

<Institution> participates in a European-wide study organised within the FIDIS (Future of Identity in the Information Society – *www.fidis.net*). This study involves gathering location data for a number of volunteers over a period of 30 days.

We are looking for people who will be monitored for the purpose of a sociological study into mobility of people and also with respect to the appropriateness of mobile phone network structures in regard to the requirements of mobile phone users. Please not that you should not switch off your mobile phone during this experiment.

The location data will be retained, and may be used again for future academic research. The location of mobile phones will be queried every 5 minutes 24x7 in cooperation with your mobile phone operator for the whole period of the study. The resolution of the position is within the phone network's "current cell" – about 800 metres in countryside, 100-200 metres in built-up areas. This querying will not be affecting any functions of mobile phones.

Each participant in the study will receive monetary compensation, and we are running an auction to select those who will take part. We invite you to submit a bid for the amount of money you require to take part in such a study. As our budget is fixed and limited, successful bidders will be those who bid the lowest amounts, and each will be paid the amount of compensation demanded by the lowest unsuccessful bidder.

Please visit the link www.buslab.org/FIDIS_experiment regardless of your intent to take (or not) part in this study.

Best regards,