

Information Bazaar: a Contextual Evaluation

Bernadette Kamleitner, Stephan Dickert
Vienna University of Economics and Business
Vienna, Austria

Marjan Falahrastegar, Hamed Haddadi
Queen Mary University of London
London, UK

ABSTRACT

The rise in the number of smart devices has created a large ecosystem centred on users' personal information and online activities. Numerous smartphone applications and social networking sites harvest and catalogue users' personal information, enabling brokers such as Google and Facebook to provide a platform for advertisers to use this information for targeted advertising. Despite the fact that the users of these services are at the heart of this ecosystem, there has been little effort in understanding individuals' perception of the value of their personal data in different contexts and situations.

In this work, we present the results of our large-scale, contextual study over ten days that used smartphones to collect data on user activities, location, and companionship, as well as the amount of money that individuals attach to such information. Our results indicate that people can be remarkably sensitive to situational cues and also be prone to valuation biases. This study represents a first step towards providing insights into the usefulness of a marketplace for information, where users, or their agents, can freely decide to auction off various pieces of their information within established contexts.

Categories and Subject Descriptors

J.4 [SOCIAL AND BEHAVIORAL SCIENCES]: Economics

General Terms

Measurement, Economics, Human Factors,

Keywords

Privacy, Activity, Location, Price.

1. INTRODUCTION

Harvesting user information through smartphones is economically attractive. This attraction is boosted by the large number of sensors on these phones, most importantly location and the bounty of applications offered by third party developers. Although these facilities are very useful for consumers, there is a high probability that consumers' private information is accessed without benefiting them. Smartphone applications are largely

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

HotPlanet'13, August 16 2013, Hong Kong, China
Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM 978-1-4503-2177-8/13/08...\$15.00.
<http://dx.doi.org/10.1145/2491159.2491161>

funded by support of integrated advertising modules within them [1]. Upon installation, a number of permissions are requested such as access to contact lists, location and call logs. Once active, apps pass this information on to advertising modules, independently from app operation. There have recently been a number of outcries over privacy issues such as access and storage of contact lists without asking for explicit permission from the user. This outcry meets with an increasing industry push towards profiling mobile user's behaviour. This information is of significant value for different sectors, in particular targeted advertising [2,3] or large-scale market research [4]. Yet from the users' point of view, such information is considered private and sensitive. Various studies across different fields such as economics, human behaviour, law, and computer science have been recently conducted to elaborate individuals' perceptions about privacy and the value assigned to personal information. However, important gaps in knowledge remain and the specific case of information garnered via mobile apps has barely been the subject of empirical scrutiny. More research aiming at understanding the value of personal information when traded in a market place is urgently needed. Currently, there is no encompassing and regulated trade space for users' personal information in which users partake. Profiling and personally identifying information is traded between organisations, advertisers and, in many cases, spammers in an unregulated manner. The first step to creating such a marketplace (see [5] and [6]), is in understanding the users' perception of the value of their personal information.

In our recent study¹ we aimed to understand the type and quality of information that people consider particularly valuable and whether individuals are capable of pricing their information. We used a mobile app to obtain real time evaluations of actual self-reported user activity patterns; thus assessing the value of information at different times of day, different places and in different situations. In order to explore whether people are in fact capable of coherently pricing this information we assessed whether valuations were sensitive to biases caused by different elicitation methods. Our app consisted of four different *experimental* versions, eliciting valuations of information in different ways (value to protect information, value to sell information using open prices and scales). The use of different value elicitation methods allowed us comparing the stability of value. We are hence able to determine whether people, in general, are able to consistently *price* their personal information. In this paper, we present a first look at the outcomes of this study and present the high level results obtained so far.

2. RELATED WORK

There have been a number of recent works on quantifying the value of personal information [16, 17]. In this paper we will focus on evaluating online and mobile phone data and location sharing preference while trying to infer the contextual monetary

¹ www.privacyvalue.org

evaluation ability of individuals. Our work is the first attempt at trying to capture the contextual evaluation of users, using a smartphone app study over duration of time, rather than relying on surveys. Our work in this sense is similar to efforts in [15]. Due to space limitations we do not represent all the related work on paper surveys or web surveys. Empirical studies suggest that value of personal information is not a constant and coherent figure and can vary based on different factors such as context of information or personal attitude toward a piece of information. Adams & Sasse [7] address a model of user perceptions of privacy, which consists of three factors: (1) data sensitivity (2) data receiver (3) data usage. They claim that individuals' decision about protecting their information depends on whether they know who can access it and how it will be used; in addition to how it is going to be interpreted by the receiver. The value of information also depends on the person itself. For example, Huberman et al. [8] show a certain degree of inertia in individuals' choice to sell information about their undesirable traits. Overweight subjects demanded higher prices than underweight or normal weight groups. Deciding how to price personal information as any other decision-making under uncertainty can be impacted by non-normative factors (such as hyperbolic discounting). For example, Acquisti & Loewenstein [9] pointed to the influence of the order of offers on the value that individuals assign to their personal information and showed that people are more sensitive to the purpose of the data collection, rather than to the duration and quantity of data collected. Riederer et al. [6] used a browser plugin to understand the way users evaluate different types of personal information and their perceptions with regard to exploitation of this information by online service providers. Their results indicate that, in exchange for personalized services, many users settle for low prices for their personal information.

3. USER STUDY

One of the main contributions of the current study is its immediacy. Existing research nearly exclusively asked for the value of information that has either already been available for some time (e.g. Facebook) or that will be available in the future (e.g., all in one selling of location data). This temporal asymmetry between generation of information and its sale may have substantial effects on the value assigned and distort the actual perceived value of information. Documented biases that are in line with a mismatch between real time and delayed estimation include hyperbolic discounting phenomena and forecasting errors.

Another important distinction between the current valuation and other prior studies (in particular comprehensive studies on social media users) concerns the usefulness of information for the potential information provider. Studies on Facebook assessed the value of Facebook profiles including all information that is stored on the profile. This mainly includes user-generated content. User generated content clearly goes beyond information about the user. Users can and do "use" their content for their own purposes whereas personal information that is simply traced via a device is at best occasionally of use to the user, for example, activity data used to monitor fitness or health.

We asked the participants to assign a monetary value to their Personal Information (PI) using a mobile phone app. Previous literature shows that such self-generated prices may be subject to biases. For example, people may construct values by anchoring on other, possibly irrelevant cues (anchoring effect [10]). In this sense any numbers presented when asking for a value may influence the value estimate given. Another potential bias for

value estimates stems from different evaluation perspectives. Ambiguity may arise as to whether it is a minimum value to keep information safe or a maximum value to sell this information. Rational actors should value information equally regardless of how they are asked for it. However, in line with psychological insights formalised in *Prospect Theory* [11], valuations are affected by their context. Changes that are framed as losses are weighted more heavily than changes framed as gains. Inspired from this theory, Thaler hypothesised that a person requires higher compensation to give up a product (willingness to accept, WTA) she would be willing to pay (WTP) to gain the same good. This is called the *endowment effect* [11] and it complicates the assessment of value. We dealt with this challenge by using different ways of assessing value.

One of the main questions in the mobile app survey asked participants to provide a monetary value for their current information, including their location, company and activity. If participants had no clear notion of how to derive such a value, estimates should be extraordinarily prone to biases such as anchoring and endowment effects. We hence observed the extent of these biases to get an idea of whether people are capable of consistently pricing their real-time in situ information. This was achieved by experimentally varying the way the valuation question was posed. Experimental versions can be logically divided into 2 main conditions: (1) WTP (What is the minimum amount that you would be willing to pay to restrict others from accessing this information?) and (2) WTA (What is the minimum amount that somebody would have to pay you to get access to this information?). Changing the way the answers were provided systematically varied WTA assessments. In the open Price condition, participants could write down any price in an open text field. In the £5 price condition participants could move a slider on a scale between £0 to £5, to indicate their desired price for that piece of information. Underneath this scale, there was an open text field in which participants could indicate any other price. In the £10 condition the procedure was analogous to the £5 condition except that the scale went from £0 to £10. Overall this yielded four experimental conditions to which participants were randomly assigned.

Our mobile application was designed to automatically show a pop-up message on the mobile screen and invite mobile holders to answer the questionnaire. The message would pop up twice per day and participants were supposed to answer it for ten days. This data was then collected and stored anonymously according to QMUL ethics approval.

One hundred and one participants provided a price for their information at least once. Participants' age ranged between 18 and 64 with an average age of 31 years. Approximately 80% of participants were under or equal to 35 years old. Only 18% of participants were female. Of the participants, 67% had an income between 20k-30k or less. The majority of participants were full-time workers or students, respectively, 56% and 32%.

With regard to the amount of apps installed, self-reports indicate that the majority of participants (53%) have between 10-40 applications on their phone. Over 50% of participants consider themselves as mid to highly privacy-concerned. The developed application was compatible with all Android devices but no other mobile operating system, which might have biased the sample somewhat. 95% of participants were using smartphones for the study.

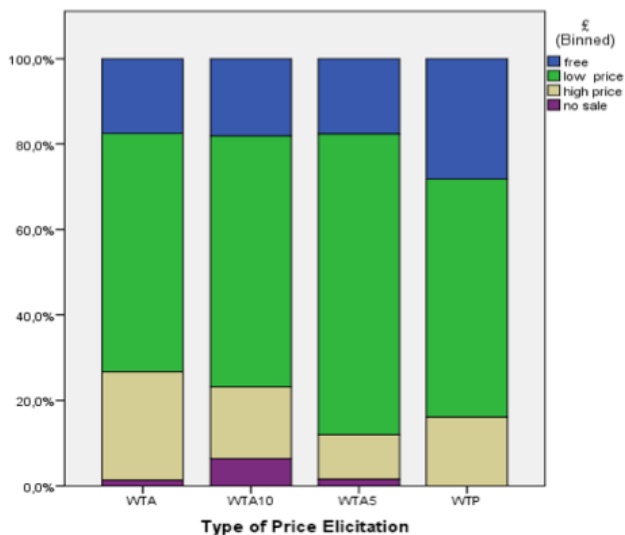


Figure 1: Price categories by condition

4. PERCEIVED VALUE OF INFORMATION

Overall, 1,425 valuations were obtained (an average 7.8 per participant). Here we present a few selected results from the experiment. In 79.4% of cases the price demanded did not exceed £10. In 20.4% of cases a value of zero was indicated.

To get a better idea of how to interpret the established differences we established different categories of valuations, similar to Bauer, Korunovska, and Spiekerman [12]: (a) instances in which information was considered free (20.42%), (b) instances in which a “low” price of a maximum of 10 GBP was demanded (59.02%); this border was chosen because it functioned as the highest possible price in the sliding scale conditions and also acts as an important psychological price marker, (c) instances in which a higher price of up to 500 GBP was demanded; we consider information in these cases as particularly valuable (17.33%) and (d) instances in which even higher prices (>500 GBP) were interpreted as indicating a general unwillingness to part with the information (3.23%).

We analysed whether the four experimental conditions yielded differences in terms of the prevalence of these valuation categories. A chi square test indicates that different ways of eliciting values indeed led to different responses, $\chi^2 (df=9) = 136.40, p < .001$, but an uncertainty coefficient of .04 indicates that the influence of the manipulation on the valuation categories is rather weak. Figure 1 shows that the following differences seem to emerge. First, participants who were asked how much they would be willing to pay in order to keep their information (WTP) were significantly more likely to treat information as free (around 28% of cases) than those asked how much money they demanded (WTA; free in around 17% of cases). This is in support of an endowment effect.

A comparison of the WTP conditions in which participants had different sliding scales gives first insights into whether participant valuations were subject to anchor effects. In line with anchor effects, participants who were provided with a sliding scale of up to 5 GBP were more likely to demand prices of up to 10 GBP than all other groups. Participants who were provided with a sliding

scale of up to 10 GBP were most likely to indicate unwillingness to sell by demanding prices in excess of 500 GBP.

Overall, even though there are variations in answer patterns that conform to endowment and anchor effects, it is remarkable how similar these patterns are in many other respects. Each experimental condition deviated from the other conditions with respect to one answer category only. Regardless of answer format participants across all 3 willingness to accept conditions were equally likely to give information for free and unless a comparably low reference price was provided a constant 55% of participants demanded “low or reasonable” prices.

In a next step, we analyzed whether valuations differed across demographic characteristics. There was a difference across gender with women being more likely to either provide information for free or to not sell it (see Figure 2)

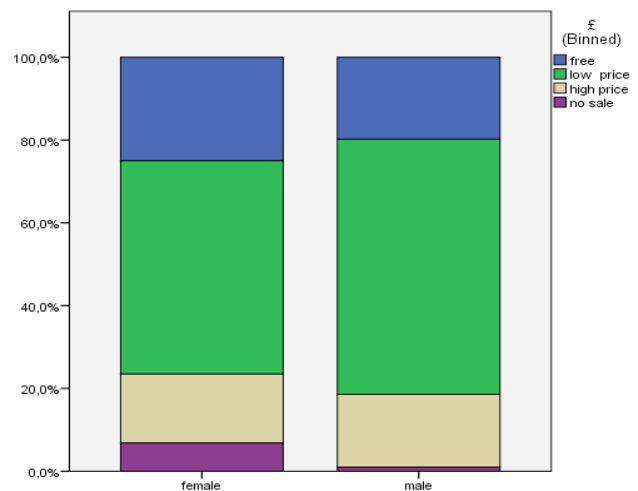


Figure 2: Price across genders

We also identified significant age differences with older people being less likely to provide information for free (see Figure 3).

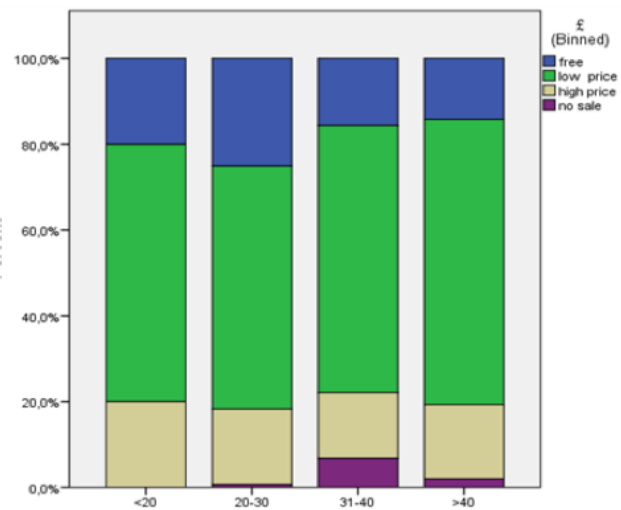


Figure 3: Price categories by age

In a next step, we aimed to understand how the concrete situation a person is in influences the value of situational information. First

we looked at the value of the location a person is in. We used χ^2 test to compare the prevalence of different valuation categories across 3 generic locations. The test proved marginally significant revealing that people who are outdoors are most likely to provide information for free.

We conducted the same analyses for activity categories but surprisingly found no significant difference in price distributions across broad activity categories, suggesting that people demand the same type of price regardless of what activity they are engaged in. The same holds for the type of company people are in. This is surprising but may be due to the somewhat rough categorizations used for these analyses and it may hide more complex situational patterns.

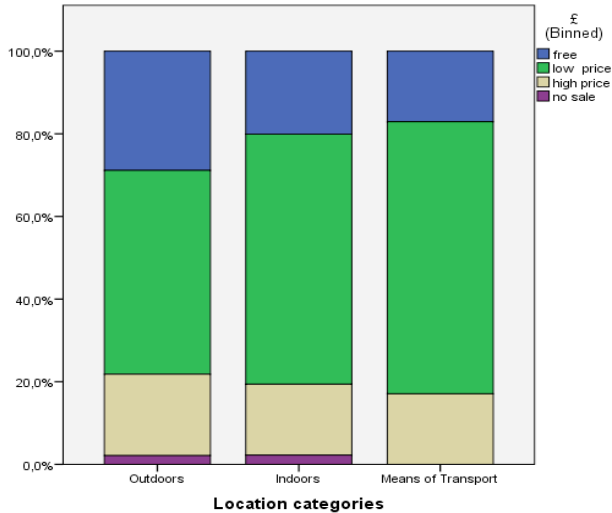


Figure 4: Price categories by location

Each data entry consists of the description of a complete situation: current detailed activity category ($n = 36$), location category ($n=12$) and company ($n = 6$). Naturally these pieces of information are often related, and form activity patterns, such as attending meetings at the work place with colleagues. To capture this holistic nature of a situation and to become capable of dealing with a dataset in which many individual situation cells contained zero observations we resorted to establishing distinct situation profiles. A two-step cluster analysis was used to condense information about detailed activities, locations and companions into 5 homogenous situation profiles with interpretable clusters (n between 170 and 311). Table 1 describes the respective clusters by documenting cluster name and most frequent locations, activities and companions per cluster.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Name	Family Matters	Me Time	Being Out and About	Being Social	Being at Work
Location	home	home	public space	no clear place	work place
Activity	eat/drink	watch TV/read	get somewhere	visting, study, eat	work
Company	family	alone	alone	friends	alone, colleagues
Size	N = 233	N = 311	N = 170	N = 266	N = 303

Table 1. Clusters of situations

In order to understand whether and how the situation and the method of eliciting valuations influences the price of information, we conducted an analysis of variance on the price demanded. Note that we first reduced skewness by capping (at 100 GBP; less than 5% of the data) and log-transforming participants' valuations. By simultaneously using situation and elicitation method as predictors we are able to determine whether there are situations in which participants are particularly sensitive to anchor effects and endowment effects respectively, in other words whether there are situations in which participants are less sure about the actual price and hence more sensitive to the elicitation method. To control for potential income differences and possible learning effects we included income and the number of prior valuations given as covariates. The resulting 4 (elicitation method) x 5 (situation) ANOVA yielded a main effect of the type of elicitation method used qualified by a significant interaction between situation and elicitation method. Surprisingly, the situation itself had low significant effect on price estimates received. In addition, income emerged as a significant covariate whereas the number of prior evaluations given did not.

Figure 5 illustrates the interaction in detail. To simplify visual interpretation it shows median prices demanded rather than mean log prices. Follow up ANOVAs per situational contrast in combination with simple contrasts provide statistical insights into the interpretation of the depicted interactions. We discuss variations across elicitation methods for each situation in turn.

When engaged in “family matters (Cluster 1)” participants seem to be somewhat receptive to anchoring effects. Participants in the WTA10 condition asked for significantly more money than participants in the WTA open price condition, however, no anchoring effect was observed in the WTA5 condition in which valuations were no different from the WTA open price condition. The difference between WTA open price and WTP open price approached significance indicating a small endowment effect.

When being “home and alone (Cluster2)” participants also seem to be prone to anchoring effects. Participants in the WTA10 condition asked for significantly more money than participants in the WTA open price condition. A downward anchor was also observed with participants in the WTA5 condition asking for less money than participants in the WTA open price condition. In addition there was a significant difference between WTA open price and WTP open price indicating the existence of an endowment effect.

When being “out and about (Cluster 3)” participants only seem to be prone to a downward anchoring effect. While there is no difference between participants in the WTA10 and WTA open price conditions, participants in the WTA5 condition asked for less money than participants in the WTA open price condition. In addition there was a significant difference between WTA open price and WTP open price indicating the existence of an endowment effect.

When engaged in social activities with friends (Cluster 4), participants do not seem prone to anchoring effects at all. There was neither a difference between participants in the WTA10 and WTA open price conditions, nor between participants in the WTA5 condition and the WTA open price condition. However, the difference between WTA open price and WTP open price approached significance indicating the existence of a small endowment effect.

When being “at work (Cluster 5)” participants only seem to be prone to a downward anchoring effect. While there is no difference between participants in the WTA10 and WTA open

price conditions, participants in the WTA5 condition asked for less money than participants in the WTA open price condition. In addition there was a significant difference between WTA open price and WTP open price, indicating the existence of an endowment effect.

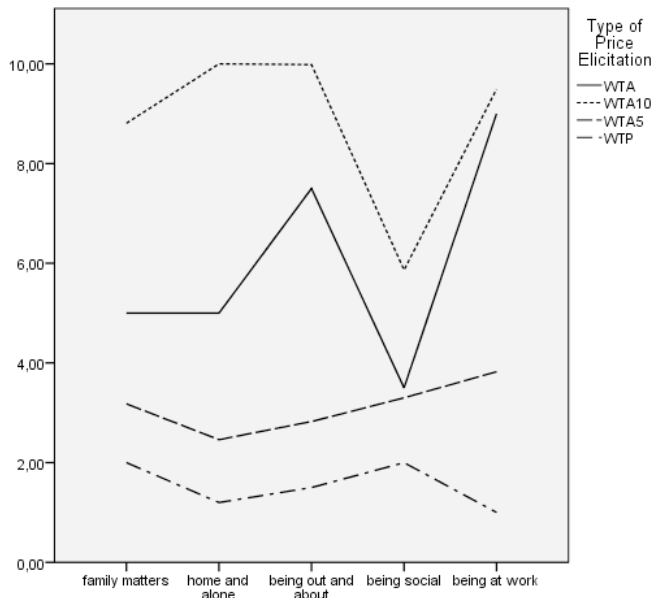


Figure 5: Median user pricing in £ in different situations

To sum up results indicate that different situations do not seem to come at different prices. However, how consistently the same type of situation is priced when there are bias-inducing cues present varies across situations. In general value estimates for in situ information are no more (if at all less) prone to anchoring and endowment biases than other objects. In fact in some situations people do not seem to exhibit commonly observed biases at all. When being engaged with others in a social situation, people’s valuations overcome otherwise pervasive anchoring effects and even the usually robust endowment effect only approaches significance. It could be that this reflects that a truly shared situation cannot be at the extreme ends of the valuation spectrum but further qualitative or experimental analyses would be necessary for being able to actually interpret these findings clearly. Nonetheless, coming back to our main point of interest: Taken together people seem able to attach a price to their in situ information, although we can not treat this as willingness to sell their information.

5. CONCLUSIONS

In this paper we have presented a first look at selected aggregate results of a large, multi-staged experiment in order to determine the contextual value of personal information from users’ perspective. In contrast with findings in [13], our survey results show that users do attach value to their information and many of them are prepared to sell them, with consistent awareness of the range of prices that this information could be realistically traded for. This is the first stage in enabling the creation of a market place for personal data, demanding user ownership over data, and clarification of the value of *free* services. It should also be noted, however, that although the price range was reasonable for most observations, we also found evidence for some of the more typical biasing effects in valuations (including anchoring and the

endowment effect). Interestingly, some situations seem to be more prone to these biases than others. Our analyses enable privacy-advocates, service providers and policy makers to better understand individuals’ perception of the personal data ecosystem and reach mutually benefiting agreements accordingly. Finally, our paper ties in neatly with current endeavors to communicate the importance of consumer protection [14].

6. ACKNOWLEDGMENTS

This work was funded by the EPSRC Pump-Priming grant number EP/J501360/1 at Queen Mary University of London. The authors wish to thank participants in the app trials and test studies. We also acknowledge constructive comments from the reviewers..

7. REFERENCES

- [1] Ilias Leontiadis, Christos Efstratiou, Marco Picone, Cecilia Mascolo, "Don't kill my ads! Balancing Privacy in an Ad-Supported Mobile Application Market " , In Proceedings of 13th ACM Sigmobility Workshop on Mobile Computing Systems and Applications (HotMobile 2012). San Diego, CA, USA. February 2012.
- [2] Hamed Haddadi, Pan Hui, Ian Brown "MobiAd: Private and Scalable Mobile Advertising", in The 5th ACM International Workshop on Mobility in the Evolving Internet Architecture (MobiArch2010), Chicago, Illinois, USA, September 2010.
- [3] Narseo Vallina-Rodriguez, Jay Shah, Alessandro Finamore, Yan Grunenberger, Konstantina Papagiannaki, Hamed Haddadi, Jon Crowcroft, "Breaking for Commercials: Characterizing Mobile Advertising", in ACM SIGCOMM Internet Measurement Conference, November 2012.
- [4] Hamed Haddadi, Richard Mortier, Steven Hand, Ian Brown, Eiko Yoneki, Derek McAuley and Jon Crowcroft: "Privacy Analytics". ACM SIGCOMM Computer Communication Review, April 2012.
- [5] Gkatzelis, Vasilis, Aperjis, Christina and Huberman, Bernardo A., Pricing Private Data (September 14, 2012).
- [6] C. Riederer, V. Erramilli, A. Chaintreau, B. Krishnamurthy, P. Rodriguez, For Sale: Your Data, By: You, ACM Hotnets 2011.
- [7] A. Adams and M. A. Sasse. Privacy in Multimedia Communications: Protecting Users, Not Just Data. 2001.
- [8] B. Huberman, E. Adar, and L. Fine. Valuating privacy. Security Privacy, IEEE, 3(5), Sept.-Oct. 2005.
- [9] A. Acquisti, L. John, and G. Loewenstein. What is Privacy worth? In Proceedings of the Twenty First Workshop on Information Systems and Economics (WISE), Dec. 2009.
- [10] A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases. In Science, volume 185, pages 1124-1131. 1974.
- [11] D. Kahneman and A. Tversky. Prospect Theory: An Analysis of Decision under Risk. In Econometrica., volume 47, pages 263-291. 1979.
- [12] C. Bauer, J. Korunovska, and S. Spiekermann. On the value of information; What Facebook users are willing to pay. In ECIS 2012 Proceedings. Paper 197.
- [13] J. Urban, C. Hoofnagle, S. Li, Mobile phones and privacy. In BCLT Research Paper Series; UC Berkeley Public Law Research Paper No. 2103405, 2012.

- [14] Federal Trade Commissions (2012). Protecting consumer privacy in an era of rapid change: Recommendations for businesses and policy makers. Retrieved from <http://www.ftc.gov/os/2012/03/120326privacyreport.pdf>.
- [15] Fehmi Ben Abdesslem and Tristan Henderson. Understanding mobile social behaviour using smartphones. Proceedings of the First Workshop on Observing the Mobile User Experience, Reykjavik, Iceland, October 2010.
- [16] Sasank Reddy, Deborah Estrin, Mark Hansen, and Mani Srivastava. Examining micro-payments for participatory sensing data collections. UbiComp '10, 2010.
- [17] George Danezis, Stephen Lewis, Ross Anderson, How Much is Location Privacy Worth? WEIS 2005