Comparing User Incentives to Participate in Crowd Sourcing Automobile Sensor Data

MARTIN DÜREN[1]

On Board Diagnostics 2 (OBD 2) Ports sind standardisierte Schnittstellen in Fahrzeugen, die es möglich machen vielfältige Sensorinformationen im Zusammenhang mit dem Abgassystem eines Autos auszulesen. Diese OBD Schnittstellen werden mittlerweile nicht mehr nur von Professionellen Nutzern genutzt, um Sensorinformationen auszulesen, sondern können auch von Amateuren genutzt werden. Möglich macht dies die wachsende Popularität und der technische Fortschritt von mobilen Endgeräten, die immense Rechenpower für jedermann verfügbar machen. An der WWU Münster wurde von Studenten eine Community Plattform mit dem Namen "enviroCar" entwickelt. Auf enviroCar können sogenannte Citizen Scientists (Bürger-Wissenschaftler) Daten ihrer Fahrzeuge sammeln und veröffentlichen. Dabei steuern sie zu einem großen Datensatz bei, der Informationen von vielen verschiedenen Automobilen umfasst und auch international gesammelte Sensordaten öffentlich bereit stellt. Dieser Ansatz wird Crowd Sourcing genannt, das heißt eine umfangreiche Aufgabe wird auf eine Menschenmenge aufgeteilt, bei der jeder einen kleinen Teil der Daten beisteuert. Oft ist eines der Hauptprobleme bei solchen Crowd Sourcing Anwendungen der Mangel an einer ausreichenden Anzahl von Teilnehmern. In diesem Paper wird genau diese Problematik aufgegriffen und versucht mit folgenden Forschungsfragen eine Lösung zu finden: "Wie können Crowd Sourcing Teilnehmer angeworben werden und gleichzeitig zum Beisteuern von Automobilen Sensordaten motiviert werden?" dabei liegt der Fokus speziell auf der Frage wie Web-Technologie genutzt werden kann, um dieses Ziel zu erreichen. Im Rahmen dieser Arbeit wurde eine Fallstudie erstellt, in der Test-Nutzer eine Plattform ähnlich wie enviroCar nutzen konnten. Diese Prototypische Plattform setzte dabei auf verschiedene Methoden, um Anreize zur Nutzung zu schaffen und die Nutzer zum Beisteuern von Sensor Daten zu motivieren. Zu den genutzten Methoden gehören Gamification Elemente, Quantified Self Features und Prinzipien, Community Aspekte sowie Umweltinformationen, die aus den Daten gewonnen werden können. Es wurden also vier verschiedene Möglichkeiten Nutzer zu motivieren getestet. Jeder Testnutzer konnte jeweils nur eine dieser Methoden zur gleichen Zeit testen. So wurde von jedem Tester jede Methode eine Woche lang getestet. Zum Beispiel ein Nutzer, der nur Umweltinformationen als „Belohnung" für seine durch Crowd Sourcing beigesteuerten Daten zu sehen bekam, konnte nach dem Aufzeichnen einer Autofahrt sehen wie viel CO_2 *ausgestoßen wurde. Dies wurde mithilfe einer Heatmap auf einer Karte visualisiert; um diese Zahlen zusätzlich zu verdeutlichen konnte der Nutzer verschiedene Umweltfakten im Zusammenhang mit seinen Daten sehen (z.B. wie lange kann eine Boeing 737 fliegen bis sie die selbe Menge an* CO_2 *ausgestoßen hat). Die Tester wurden zusätzlich in unterschiedliche Gruppen eingeteilt; jeder dieser Gruppen wurden die unterschiedlichen Methoden zur Schaffung von Nutzungs-Anreizen in einer anderen Reihenfolge zur Verfügung gestellt, um Folgeeffekte wie z.B. Ermüdung auszuschließen.*

Die Resultate der Fallstudie wurden zudem mit zwei zusätzlichen Umfragen abgeglichen. Die erste Umfrage richtete sich an die generelle Öffentlichkeit und die zweite Umfrage wurde den Test-Nutzern im Anschluss an die Fallstudie ausgehändigt. Die Fallstudie sowie die Umfragen lieferten ähnliche Ergebnisse. Sie zeigten beide, dass Quantified Self Aspekte

[1] Martin Düren, Institut für Geoinformatik der Westfälischen Wilhelms Universität Münster

am besten genutzt werden können, um Nutzer einer Crowd Sourcing Plattform zu motivieren. Das indiziert, dass hauptsächlich Egoistische Motive zur Motivation von Nutzern beitragen und somit auch am effizientesten genutzt werden können diese Motivation weiter zu fördern. Konkrete Features im getesteten Prototypen waren genaue Statistiken zum Fahrverhalten sowie zum Benzinverbrauch. Dem Nutzer wurde der Preis des verfahrenen Benzins angezeigt, eine Kartenansicht mit verschiedenen Datenebenen, die zusätzlich in einem Graphen veranschaulicht wurden. Dazu gehörten die Werte zur gefahrenen Geschwindigkeit, Umdrehungen des Motors und Verbrauch. So konnten Nutzer deutlich mithilfe der Karte und des Graphen ihren Fahrstil analysieren. Andere Features des Prototypen wie z.B. Gamification Elemente (Auszeichnungen, Ranglisten) oder die Community Elemente (Karte mit gesammelten Informationen aller Fahrer) stellten sich im Gegensatz zu den Quantified Self Features als unbeliebt heraus.

1 Introduction

In 2010 the number of cars on the world's roads surpassed 1 billion cars (Sousanis, 2012) and still 60.000.000 new cars have been produced in that year (Worldometers, 2012). An average German citizen produces 11 tons of CO2 annually; 2.5 tons of these emissions come solely from transportation (Katharina Schächtele, 2007). According to the Intergovernmental Panel on Climate Change (2007) nearly 14% of the world's carbon dioxide emissions are caused by burning fossil fuel. Modern automobiles have become more efficient and Governments and international councils set out to establish new directives and guidelines to control CO2 emissions (Bundesministerium für Umwelt, 2008). Also the development of computer technology provides more sophisticated ways to monitor and control our emissions. Especially the development of mobile computing, sensor technology and sensor interfaces makes collecting detailed statistical data possible. The broad availability of smart phones, which recently in 2012 topped one billion units worldwide (Lunden, 2012) puts mobile computing in the consumers hands and gives developers new options and opportunities.

Since 1995 every produced car provides an On Board Diagnostics 2 (OBD 2) port that makes it possible to exchange real-time data between any connected computing device and the car (Hilpert et al., 2011; Burelle, 2004; Cross, 1998). Using this interface a number of mobile applications already provide various monitoring possibilities. The best known application is called "Torque" (Torque, 2012), which is targeted at car enthusiasts, displays real time data on the smart phone such as a dynamo-meter and horsepower or car health statistics as well as a range of other features. Technology like this can be used for manifold purposes as there are more than 150 individual types of sensor information available through OBD 2 in most cars (OBD-Codes, 2012). Information gathered from these sensors could be used to create emission profiles for streets, monitor traffic data or help citizen scientists who are in need of such data. Concepts often used to collect information from sensors are crowd sourcing and participatory sensing. Both are methods where members of a community contribute to a sensor network by providing sensor data. Burke et al. (2006a) describe participatory sensing as sensor networks that give public and professional users the ability to gather, analyze and share data. This method also offers the opportunity to enhance the quality and credibility of large data sets but requires participation at a personal level. The ability to monitor the physical state of the world, such as pollution levels, can at the same time help to increase environmental awareness (Stevens and D'Hondt, 2010). By using smart phones as mobile sensor stations their sensors become "smart items" that have full Internet connectivity, memory and advanced elaboration capabilities (Atzori et al., 2010) and at the same time provide a tool for participatory sensing, which is usable by nonprofessionals.

At the University of Münster a group of Students implemented a novel Internet platform called *enviroCar*, where a community of car drivers is able to collect their cars' sensor data and share this data on the community platform. For enviroCar the students implemented a server component with a publicly accessible API, an Android application and the website front-end. The whole project aims at publishing environmentally relevant sensor data in order to provide this data to citizen scientists.

Figure 1: Architecture of the used prototype.

2 Problem

The enviroCar platform makes use of the participatory sensing approach to collect sensor data, but the ability to collect raw usage statistics of automobiles does not offer motivation to participate on a personal level, nor does it guarantee privacy or securities, which are some of the central issues in connection with this method. An enviroCar user who is not a citizen scientist and does not need a large community data set of automobile sensing data, will ask himself why he should participate in collecting this data and why he should spend his private time

participating in the community by recording his or her car rides. Also by submitting data to a platform such as enviroCar the users expose their data to the public.
Anhai et al. (2011) define four key challenges in crowd sourcing systems: „How to recruit contributors, what they can do, how to combine their contributions and how to manage abuse". Other challenges are scalability, raising incentives, cost sharing, security, privacy and trust (Dua et al. (2009), Kansal et al. (2007)). On the technical side central issues are content integrity and providing full interoperability of interconnected devices; especially connected sensor systems need inter-operable interfaces (Atzori et al. (2010), Gorman et al. (2009), Dua et al. (2009)). Popular examples of crowd sourcing successes are Linux and Wikipedia. While Wikipedia works using a credit system and the mutual information sharing interests to raise user incentives, Linux works mainly through Linux enthusiasts as the driving force behind single Linux distributions. Another more related example is the Intelligent Transportation Systems Laboratory (ITSL) at Portland State University, who analyze data sampled from a small number of cars to optimize commute paths and bypass congestion, but the problem is that they do not have enough data (Dua et al., 2009). These applications and all similar projects need to find a good balance between openness and quality (Anhai et al., 2011) and need effective ways to recruit and motivate contributors. Additionally the available hardware of OBD adapters and smart phone connectivity has not yet arrived at an end-consumer level and thus does not provide easy access and easy to use features. The current state is that OBD interfaces are solely used by car-enthusiasts and professionals.

3 Research Question

As described above current research goes into maintaining privacy and raising user incentives in participatory sensing. In order to successfully build a participatory sensing platform where a crowd of participants gathers sensor data it is necessary to provide incentives for participation on a personal level. So on the one hand is the opportunity of collecting automobile sensor data that can be used for environmental purposes or even be used on an end-consumer level and on the other hand there is a lack of participatory sensing contributors due to missing user incentives. In this context the main research goal was *"assessing incentives raised in participatory sensing automobile sensor data by comparing features of a web portal that provides personal benefits for users versus benefits resulting from collecting data under the aspect of an online community platform"* with the focus on the question: *"Do personal benefits outweigh community benefits in crowd sensing automobile sensor data and how can different visualizations help in raising user incentives"*. The hoped for aim of this research was to provide a model for effective incentive mechanisms that can be used in future projects to raise user incentives in participatory sensing.

4 Experiment Design

To answer these questions different methods used for raising user incentives were compared. Initial research dealt with the comparison and combination of different incentive mechanisms which could be used in applications to raise user incentives as extracted from literature and real world examples. Additionally different incentive mechanisms were evaluated in group talks and brain storming sessions with potential users and developers. A questionnaire was also conducted to gain additional insight on the possible effectiveness of the previously researched incentive mechanisms. A prototypical website and mobile application similar to the enviroCar platform were used to evaluate incentive mechanisms.

From the initial research and the questionnaire the most effective and feasible incentive mechanisms were implemented in a prototype as web portal features. Within the initial research four features were chosen to be used for motivating users to participate on a crowd sensing platform. The chosen features to raise user incentives were:

1. Detailed statistics such as motor data, fuel consumption and the associated expenses visualized in a map and a graph (Quantified Self inspired features).
2. Aggregated community data shown on a map that provides different information products such as average speed, consumption and emissions on the streets, average waiting time and maximum speed on street segments (Community features).
3. Badges and driver ranking points rewarded for an ecological and generally good driving style (Gamification features).
4. Environmental data in the form of CO_2 emissions to raise the drivers' environmental awareness (Environmental features).

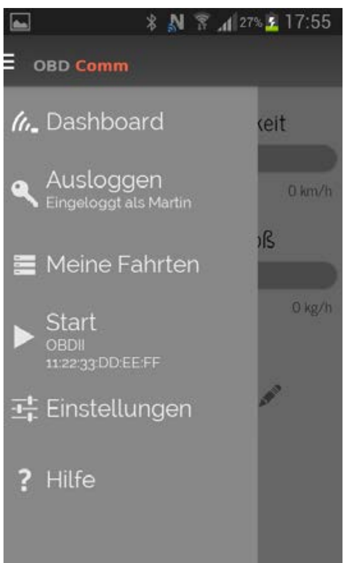

Figure 2: Screenshot of the prototype's Android application interface.

In order to evaluate the effectiveness in raising user incentives during a case study the prototype was handed to eight test users for the duration of one month. However, it was important that each feature could be evaluated independently. Each user was given every feature to test in an isolated fashion. Only one feature was tested at a time by each tester. The sequence in which features were given to users differed between every test user. This experiment approach is called counter balanced measures design, which is a commonly used method for experiments in psychology where the sequence of the tests is based on a balanced Latin square (Reese, 1997). This approach was used to counter possible interference or learning effects such as users losing interest in between the features to be tested.

It was necessary to observe the test users' behavior during the study. Thus their activity within the web application was logged during the experiment. Every click and the time spent in the UI were being saved to a database. This way it was possible to draw conclusions towards the raised incentives by comparing click and page view rates. In addition to this evaluation study the test users were handed a questionnaire after participating in the case study. The results were used to answer the posed research question and can be used to identify possible usability issues or other limiting factors. The hoped-for conclusions to be drawn from this approach were a clear distinction between the effectiveness in raising incentives of different sets of features as provided by the prototype. The

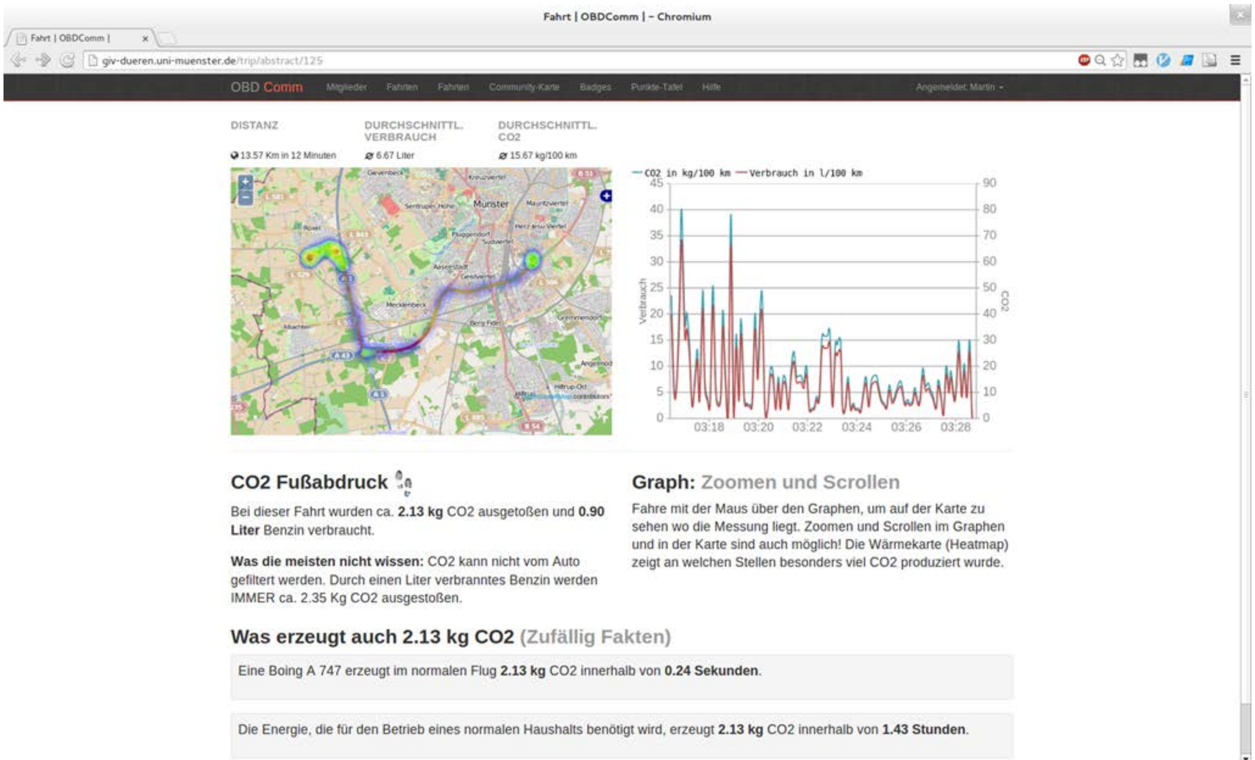

Figure 3: Screenshot of the prototype's web portal interface showing a view on the environmental features.

aim was to find out, if after the user testing it will be clear, if one set of features clearly outweighs the other ones and could be recommended to be focused in future projects.

5 Experiment Evaluation

The results show that the usage between the individual features varies significantly for almost all users which gives room for conclusions on the users' individual preference of the utilized features; thus allowing to draw conclusions on which method was the most effective one to raise user incentives. However the preferences of the users are very dissimilar and no clear outcome can be derived from simply looking at the page views of each user. Although the trend can be observed that the Quantified Self and environmental features are more favored by the users than the gamification and community features even if that was not the case for all users.

	Features (Views)	Other Pages (Views)	Interactions
Quantified Self	76	60	374
Community	43	99	118
Gamification	61	62	n/a
Environmental	67	64	95

Table 1: Number of page visits per feature for all four sets of features. (Note: Gamification features did not include any interactive elements)

The numbers alone suggest that the community features are the least popular ones. During the case study the users had only access to one feature at a time, but not only to the pages of the activated feature but also to some common pages such as profile pages. A clear trend is observable in regards to an increase of common page views during the activation of community features. When the users had only access to those features, they viewed the common pages more than the actual feature's pages.

In the follow-up questionnaire the testers were asked to state their preference for all features directly and asked to fill out a question used to extract calculated preference weights using conjoint analysis later on. The results show that their stated preference differed from the observed preference, but also differed from the calculated preference obtained through conjoint analysis. However all methods show the same trend, that Quantified Self seems to be the most effective way to motivate users and make the software personally beneficial for them. This result confirms the result from the first questionnaire, which was conducted before the case study took place. The second most favored features by stated preference are the environmental features followed by gamification and then community data. To eliminate usability failure as a factor in the statistics the testers were also asked to give their opinion on how much potential they saw in the presented features if those were better developed. The result did not differ significantly from their stated preference.

6 Conclusions

The question if personal benefits outweigh community benefits in crowd sensing automobile sensor data was investigated. Crowdsourcing and in this context participatory sensing are a popular approach to gather extensive data sets in research as well as in commercial projects. The research question picks up the problem that participatory sensing needs to rely on a broad user base to account for credible data. These users need to be recruited and stay motivated to ensure them contributing data to the community. A combined approach of conducting questionnaires with prospective and actual users and a case study was used to find out how users can be motivated best in order to contribute automobile sensor data on an online community platform. For this purpose four sets of suitable features for raising user incentives were extracted from existing literature and surveys and tested within the case study.

The questionnaires and the case study allow to draw the same conclusions: The gathered data indicates that Quantified Self inspired features are a good way to trigger user incentives and make a participatory sensing product or software attractive in the eyes of a user. While participatory sensing itself is generally not used for the personal benefits of the contributors, it is necessary to motivate the participants on a personal level in order to account for good data. The findings of the research suggest that while Quantified Self features are a good way to raise user incentives, the motivation provided through the available community data provides less user motivation. Individual interviews and questionnaires with the testers have shown that all of them liked the tested software without exception, and they would want to continue using the provided system. Some of them even mentioned after the case study that testing the software changed their driving behavior and views on ecological driving for the better.

Future research and work on this topic needs to address the problem stated in the introduction of this thesis: Project planners need to think about how to motivate their prospective users. While the main findings of this thesis show that personal benefits are of importance for participatory sensing contributors, literature suggests other problems that need to be addressed in this context and especially the importance of community incentives. Forte and Bruckman (2005) state that in order to sustain user involvement it is necessary to "meaningfully structure participants' contributions", the authors also state that users can achieve higher levels of efficacy and responsibility within the community if their contributions are presentable and are of importance for the project. Woolgar and Latour (1986) found as early as 1986 that within the scientific community the cycle of credit is the most important incentive system. This also needs to be focused in future projects involving citizen scientists. The presented research was conducted with a low number of testers and participants. The results can show up a trend but future research can also go into consolidating these findings by addressing the community issues more and testing the proposed feature and incentive mechanisms in a real community environment for a longer duration.

7 Literature

Doan Anhai, Ramakrishnan Raghu, and Halevy Alon Y. Crowdsourcing systems on the World-Wide Web. Commun. ACM, 54(4):86–96, 2011.

Luigi Atzori, Antonio Iera, and Giacomo Morabito. The Internet of things: A survey. Computer Networks, 2010.

Bundesministerium für Umwelt. Die EU-Verordnung zur Verminderung der CO2 Emissionen von Personenkraftwagen, 2008. URL http://www.bmu.de/files/pdfs/allgemein/application/pdf/eu_verordnung_co2_emissionen_pkw.pdf.

Carol Burelle. On-Board Diagnostics II (OBDII) and Light-Duty Vehicle Emission Related Inspection and Maintenance (I/M) Programs. D. Cope Enterprises, 2004. URL http://www.ccme.ca/assets/pdf/jia_trnsprt_obd_e.pdf.

Jeffrey A Burke, Deborah Estrin, Mark Hansen, Andrew Parker, Nithya Ramanathan, Sasank Reddy, and Mani B Srivastava. Participatory sensing. 2006b.

Robert Cross. Message to All Car Manufacturers. On Board Diagnostic II Compliance Guidelines. Air Resources Board California Environmental Protection Agency, 1998. URL http://www.arb.ca.gov/msprog/obdprog/mailout9801.pdf.

Akshay Dua, Nirupama Bulusu, Wu-Chang Feng, and Wen Hu. Towards trustworthy participatory sensing. In HotSec'09: Proceedings of the Usenix Workshop on Hot Topics in Security, 2009.

Andrea Forte and Amy Bruckman. Why do people write for Wikipedia? Incentives to contribute to open–content publishing. Proc. of GROUP, 5:6–9, 2005.

B. L. Gorman, D. R. Resseguie, and c. Tomkins-Tinch. Sensorpedia: Information Sharing Across Incompatible Sensor Systems. In International Symposium on Collaborative Technologies and Systems, 2009.

Hendrik Hilpert, Lars Thoroe, and Matthias Schumann. Real-Time Data Collection for Product Carbon Footprints in Transportation Processes Based on OBD2 and Smartphones. In Proceedings of the 44th hawai international Conference on System Sciences, 2011.

Intergovernmental Panel on Climate Change. Climate change 2007: Synthesis report, 2007. URL http://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr.pdf.

Aman Kansal, Suman Nath, Jie Liu, and Feng Zhao. Senseweb: An infrastructure for shared sensing. IEEE MultiMedia, 14(4):8–13, October 2007.

Ingrid Lunden. Mobile milestone: The number of smartphones in use passed 1 billion in q3, 2012. URL http://techcrunch.com/2012/10/16/mobile-milestone-the-number-of-smartphones-in-use-passed-1-billion-in-q3-says-strategy-analytics/. Accessed 6 December 2012.

OBD-Codes. Obd-ii trouble codes, 2012. URL http://www.obd-codes.com/.

Hayne W Reese. Counterbalancing and other uses of repeated-measures latin-square designs: Analyses and interpretations. Journal of Experimental Child Psychology, 64(1):137–158, 1997.

Katharina Schächtele. Die CO2 Bilanz des Bürgers. Umwelbundesamt, 2007. URL http://www.umweltdaten.de/publikationen/fpdf-l/3327.pdf.

John Sousanis. World vehicle population tops 1 billion units, 2012. URL http: //wardsauto.com/ar/world_vehicle_population_110815. Accessed 6 December 2012.

M. Stevens and E. D’Hondt. Crowdsourcing of Pollution Data using Smartphones. 1st Ubiquitous Crowdsourcing Workshop at UbiComp, 2010.

Torque. Engine performance and diagnostic tool for automotive professionals and enthusiasts, 2012. URL http://torque-bhp.com/.

Steve Woolgar and Bruno Latour. Laboratory of Life: The Construction of Scientific Facts. Princenton University, 1986.

Worldometers. Cars produced in the world, 2012. URL http://www.worldometers. info/cars/. Accessed 7 December 2012.