# Personalized Vehicle Insurance Rates

## A Case for Client-Side Personalization in Ubiquitous Computing

**Vlad Coroama, Marc Langheinrich**
Institute for Pervasive Computing
ETH Zurich
{coroama, langhein}@inf.ethz.ch

### ABSTRACT

In the foreseeable future, miniature sensor and networking technology will allow insurers to offer personalized insurance rates based not only on demographic data, but on actual customer behavior. Behavior-based personalized rates can significantly lessen the traditional information asymmetry between insurers and insured, and thus prompt significant cost savings both on a personal and societal level. However, initial prototypes have so far required consumers to disclose their entire data set in order to enjoy personalized cost savings. We have opted to follow a more privacy-friendly approach, by performing the entire personalization on a trusted platform on the client-side and only reporting gross rates to the insurer. This article describes our system and a set of requirements we derived for client-side personalized insurance rates.

### Author Keywords

Ubiquitous computing, personalized insurance rates, client-side personalization.

### INTRODUCTION

An old and well-known management adage points out that "you can't manage what you can't measure." With today's rapid spreading of ubiquitous computing technologies, more and more real-world actions will become measurable, thus manageable, and eventually billable. Whole new economic applications are hence enabled by ubiquitous computing [9, 13]. An area that might be particularly affected by better measurements is the insurance market in general, and vehicle insurances in particular [7, 8, 12]. Insurers often suffer from a lack of knowledge on how their clients treat the insured goods, a phenomenon known as information asymmetry [2]. If the insurers could know precisely how the insured goods are treated, they could use this data to better compute the risk of an accident, and thus an insurance rate. Such highly *personalized insurance rates* would have several advantages over today's demographics-based systems: Firstly, personalized insurances would be fairer, since they would reflect more exactly the individual risk, rather than having less "risky" customer cross-finance the higher risk of the more daring ones. Secondly, measurable insurances would at the same time give clients an incentive to be more careful in handling the insured goods, since more careful behavior will directly lower their costs. Last not least, deploying individualized insurances might prompt other beneficial societal effects, such as increased safety or ecological gains [6, 10]. Obviously, the price to pay for all these improvements would be the individual's privacy: computing a client's risk and properly rewarding safe customers typically requires insurers to observe the individual's behavior in great detail, potentially 24 hours a day, seven days a week. A personalized car insurance, for example, would record all trips taken with the family car, possibly taking into account not only where, when, and how the car was driven, but also by which member of the family. Such data would not only be useful for the insurers themselves, but also to marketers, neighbors, political enemies, or law enforcement.

In this paper, we try to devise an alternative personalization method that retains all benefits of individual insurance rates while preserving personal privacy as much as possible. Such insurance schemes based on *client-side personalization* [4, 11] restrict the amount of information available to insurers by computing insurance rates locally, and only uploading monthly or yearly rate totals to the insurer. The remainder of this paper will describe a first prototype based on such a scheme that we built in the domain of vehicle insurances, the *Smart Tachograph*, and will point out challenges and potential solutions implied by the use of client-side personalization.

### PERSONALIZED VEHICLE INSURANCE RATES

Various authors argue that today's classification of drivers into a few classes based, e.g., on their driving experience, accident history, and type of driven car is not optimal [10]. [12] points out that within such a class (of presumably similarly skilled drivers), there is still a large spread of risks, depending on such factors as the annually driven mileage, the time of day and season typically driving, the predominant weather conditions when driving, the type of route driven, or the neighborhood where the car is usually parked.

Thus, several insurers have investigated over the past years the possibility of personalizing insurance rates to a higher degree than they are today – through studies, pilot projects, and more recently also market products. Progressive, a US-insurer, deployed one of the first pilot projects between 1998 and 2000. A black box containing a GPS-sensor measured the distances driven and reported them to the in-

surer's server, allowing Progressive to offer a mileage-dependent insurance rate.[1] More recently, Norwich Union, a UK-based insurer, offers a similar unit for what they call "pay-as-you-drive" insurance.[2] While they disclose that they take the time of driving into account, no further information is available. According to Norwich Union, customer privacy is not an issue: "The black box device measures vehicle usage and sends data directly to Norwich Union using similar technology to that used by mobile phones." Progressive also started last year to offer a more sophisticated insurance product, called TripSense.[3] TripSense is based on a black box that has to be installed in the car as well, but their webpage is more detailed about what it will record: It "measures your actual driving habits and allows you to earn discounts on your insurance by showing us how much, how fast and what times of day you drive." The driver may analyze the data recorded over several months at her PC at home and decide for herself whether she wants to send the data to the insurance company or not. If the data is not sent, a no-punishment policy is advertised.

While seemingly a win-win situation, Progressive's voluntary trip disclosure still carries strong privacy implications. Although Progressive does not force the customer to send her data, it needs to be fully disclosed in order to profit from the offered savings. Given the popularity of supermarket loyalty cards all around the world, already a miniscule price reduction might prompt a majority of customers to hand out their data.[4]

The more important point, however, is that the responsibility lies with the customer. How can she properly assess the longer-term consequences of such an action? Not only is it hard to imagine the countless situations in which such a comprehensive record might be undesirable, it is also unclear how the disclosure or non-disclosure affects the customers relationship to the insurer itself: What will happen when the contract is up for renewal and the driver has not sent "enough" data to the insurer? Or what if the data was sent, but the sophisticated data mining algorithms of the insurer detected a sublime risk pattern in the otherwise safe driving style that will lead to significantly increased rates? And of course: Who else will gain access to the data, and will "my" data be used against me? Progressive states that: "We may retain the information that you send to us indefinitely" and further "If you are in an accident, you may have a legal obligation to preserve the information on the TripSensor. This information may be sought by opposing parties in a civil lawsuit or by police when investigating the cause of an accident. We may be legally obligated to provide such information in response to a subpoena or as otherwise required by law."

## THE SMART TACHOGRAPH

In the "Smart Tachograph" project [7] we tried to provide evidence that highly personalized insurance rates are also feasible without such a significant loss of privacy and control. The Smart Tachograph is a generic platform that allows the precise and individual measurement of traffic-related costs (including the costs of individual insurance risks) and charging these costs to their originators. The system, like those previously described, consists of a sensor box that is installed in the car. In contrast to the other presented projects though, the acquired telemetric data is not sent to the insurer directly. Instead, the box locally computes the current insurance rate (which in our prototype depends on the distance driven, the current weather and traffic conditions, as well as the individual driving style) and periodically sends a gross total (e.g., once a day, or once a month) to the insurer.

The system provides a generic platform to gather and analyze sensory input (or more specifically: telemetry data). Upon system start, the vehicle owner may choose between different insurance offers that the Smart Tachograph automatically acquired through a publish/subscribe mechanism from various insurers. After the driver has chosen a specific insurance contract, a piece of software (a Java class) is downloaded from the insurer to the vehicle's computer that implements the chosen contract. Contract software must adhere to a specific API in order to pick up data from the vehicle's sensors. It can then calculate the corresponding insurance rate for a given period of time and send the total amount at the end of that period to the insurer.

By sending insurers only a gross total at the end of a billing period, no inferences on the individual driving style of the insured drivers can be drawn. Even if a driver ends up with a high rate at the end of a period, this sum may be result of a risky driving style, but could also be caused by a very cautious, but heavy-mileage driver. Moreover, since the data is cached only client-side, the driver retains control over her data, which reduces the possibility that data gained by the sensors will be used against her.[5]

## SYSTEM CHALLENGES

A billing system based on client-side personalization poses a range of new challenges. They all ultimately require the system to be designed in such a way that neither party may tamper with the system (or at least that such an attempt will be noticed). Customers need to be assured that the

[1] See www.epa.gov/projectxl/progressive/index.htm.

[2] See www.norwichunion.com/pay-as-you-drive/.

[3] See https://tripsense.progressive.com/home.aspx.

[4] Loyalty cards typically offer less than 0.5% in savings.

[5] Of course, this does not preclude court-orders that might force her to turn such data over to law enforcement – in such cases, only an officially sanctioned deletion strategy, e.g., after several months, might limit the amount of data that could be disclosed.

downloadable code bills them according to the agreed-upon contract, while insurers must be satisfied that the system will not allow customers to submit false (i.e., lower) rate totals.

### Insurer Safety

Drivers have several options for attacking the system. An obvious one would be to modify the software class *S* received from the insurer into a new software piece *S'*, which then sends an insurance rate *R'* lower than the original rate *R* that would have been sent by *S*. The system thus needs a mechanism for *tamper-proof software distribution*, which in turn is one of the core functionalities of the Trusted Computing Platform (TCP).[6] Originating from early work on secure bootstrapping [3], TCP-based systems will allow software distributors to securely deploy software programs by using special hardware chips to secure a machine's initial state, memory, and computation, as well as having specific software modules in both the operating system and the application software. While the deployment of trusted computing is highly controversial in the area of consumer devices (e.g., personal computers, media players) [14], such an infrastructure can play an important part in safety-critical systems, where there is less of a threat for curtailing individual rights such as fair use. For our purposes, the implementation of a TCP inside the black box will effectively ensure the correct deployment of the insurer's billing code.

Given a correctly functioning software *S*, however, the vehicle's owner may alternatively try a man-in-the-middle attack by intercepting the message *C* from *S* to the insurer and replacing it with another message *C'*. To counter this, we require a black-box-specific secret key $s_c$, which then allows us to digitally sign all messages sent by *S*. Such a key might either be part of the TCP on the black box, or a customer-specific key given out by the insurer (e.g., as part of the software distribution step). Given the sealed storage property of TCP, this key is not known accessible from outside our black box, thus making it impossible for the vehicle owner to replace the message *C*. By including a timestamp or serial number with each message *C* we can also prevent replay attacks (i.e., replacing a message $C_h$ reporting a high rate with an older message $C_l$ that reports a lower rate).

Last not least, a driver could try to tamper directly with the vehicle's sensors, e.g., inserting a device that would cap the reported speeds to always stay below a certain (safe) limit, or simply covering up the rain sensor, so that it never reports rainfall (as driving under rainy conditions would make the insurance more expensive). Note that this is not an attack specific to the client-side personalization paradigm; it would be equally effective in systems such as Progressive's TripSense. However, by employing data mining algorithms, similar to those used by credit card companies to empirically discover a high probability fraud, insurers might be able to discover such tampering based on extensive analysis of valid driving records. For example, a security system on the insurer's side could issue an alarm if a certain change speed would be accompanied by certain non-standard acceleration changes, or if the reported telemetric data does not follow the path of the car's current position.

To counter a sensor tampering attack, we see a number of possibilities. The most obvious and probably least realistic would be to include such detection heuristics into either the operating system of the black box, or alternatively into each insurer's billing software. Unless such detection methods could be boiled down to some simple algorithms, the limited computing and storage facilities of the box might render this infeasible. A second, less powerful alternative would be to have the box periodically send position-neutral telemetric data to the insurer, in order to at least cross-verify the various sensors for acceleration, speed, temperature, and road type. The rogue customer taping over the car's rain sensor might thus get caught as the reported dryness would not match the also-measured humidity levels or the traction feedback from the wheels. More promising might be the approach of using data perturbation techniques from the field of statistical databases [1] to randomize the driving record before sending it off to the insurer, thus allowing for a statistical validation of the telemetric data without disclosing individual trips. One might further envision building equally tamper-proof sensors that would communicate with the black box using encrypted and authenticated channels, though it might still be several years before this would be cheap enough. An even simpler solution may be to include as many sensors as possible into the tamper-proof black box hardware, although this approach may not work for some of the sensors (e.g., the rain sensor).

### Customer Safety

Equally imperative for the acceptance of such a system will be the customer's ability to verify the correct functioning of the billing mechanism, i.e., that the daily or monthly charges correspond to the agreed-upon insurance contract.

Rogue insurers might be tempted to simply charge a sum *C'* different (i.e., higher) from the 'true' sum *C* reported by the car's box. To this extend, customers might want to verify not only the values that the unit is reporting back to the insurer, but also keep a local log of these sums in order to verify the totals billed. In order to allow verification of the encrypted total *C*, our previously stipulated secret key $s_c$ must be part of a *public key cryptography* system. This enables a customer to decrypt this value using the corresponding public key $p_c$ (which would be made available together with the installed black box or the downloaded billing module). Equally important would be to have the billing software also available in unencrypted form in order to allow customers to locally verify the computed values.

Note that while market pressures might prompt most insurers not to cheat in such a blunt way, the ability to quickly

[6] See https://www.trustedcomputinggroup.org/home

switch insurers with such an infrastructure (potentially on a trip-by-trip basis [6]) might prompt developments similar to today's telephony market, where many consumers simply choose the cheapest provider on a call-by-call basis, thus risking to fall prey to scrupulous fraudsters.

The second customer-safety requirement concerns the correspondence of the downloaded software to the advertised rates. One option would be to standardize the computation in such a way that insurers would only publish and download the specific parameters. However, this might not only restrict the insurers' ability to differentiate their products sufficiently, but also turn out to be too simple to properly compute the correct risk value of a specific driving style. Alternatively, consumer interest groups or government bodies might offer a testbed-environment in which published billing algorithms could be tested under a variety of simulated conditions (potentially using tools from formal software verification to fully automate this task) and which would award trust-seals to verified algorithms. Note that such open-source billing algorithms would not necessarily infringe on an insurer's trade-secrets, as these typically lie with the way these algorithms or parameters are chosen, rather than with the published rates themselves. Also, our research has shown that many insurers would prefer rather simple algorithms based on few parameters only, in order to better communicate (i.e., sell) these to their customers [5].

## CONCLUSION

Ubiquitous computing offers the possibility to measure a range of real-world features more precisely and continuously than before, thus opening up many new economic options. Personalized insurances that take into account an *individual* rather than an *average* risk-level are one example of such a development. In order to provide such a product in a non-intrusive way, *client-side personalization* is of utmost importance. However, in contrast to many other personalization areas, personalized insurances need to take malicious users into account, which brings up a range of new challenges to the field: How can one ensure that users trust the personalization method offered by the insurer? How can such a personalization method be understandable to the user, while at the same time be expressive enough for modeling complex patterns? And how can an insurer trust the modeling parameters sent back from a customer to correctly reflect the customer's behavior?

With the Smart Tachograph, we have implemented a first prototype of such a personalized insurance scheme that directly tries to address such questions. While the current system only reflects the basic data flow model (i.e., only rate totals are being sent back to the insurer), it serves as a testbed application for devising and implementing the above-mentioned trust requirements. It also serves as a demonstrator to a range of stakeholders by illustrating the possibilities and implications of ubiquitous measurement systems, not just for vehicle insurances, and initiating a much-needed debate on fairness and efficiency in society.

## REFERENCES


1. Adam, N. R. and Worthmann, J. C. Security-control methods for statistical databases: a comparative study. *ACM Computing Surv*eys. Volume 21, Number 4, ACM Press (1989), 515–556.
2. Akerlof, G. The Market for Lemons: Qualitative Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics* (1970) 84(3):488–500.
3. Arbaugh, W., Farber, D., and Smith, J. A Secure and Reliable Bootstrap Architecture. In *Proc. of IEEE Symp. on Security and Privacy*, (1997), 65–71.
4. Cassel, L., Wolz, U. Client Side Personalization. In *Proc. of DELOS Workshop* (2001).
5. Coroama, V. The Smart Tachograph – Individual Accounting of Traffic Costs and its Implications. To appear in *Proceedings of Pervasive 2006*, Springer-Verlag (2006).
6. Coroama, V., Bohn, J., Mattern, F. Living in a Smart Environment – Implications for the Coming Ubiquitous Information Society. In *Proc. of the IEEE Conf. on Systems, Man and Cybernetics*, (2004), 5633–5638.
7. Coroama, V., Langheinrich, M. The Smart Tachograph. Video submission abstract. *Adjunct Proceedings of UbiComp 2005*.
8. Filipova, L., Welzel, P. Reducing asymmetric information in insurance markets: Cars with black boxes. In *Proc. of the 32nd Conference of the European Association for Research in Industrial Economics (EARIE).*
9. Fleisch, E., Dierkes, M. Betriebswirtschaftliche Anwendungen des Ubiquitous Computing. In F. Mattern (Hrsg.) *Total vernetzt*, Springer-Verlag (2003).
10. Litman, T. Distance-based Vehicle Insurance. *Victoria Transport Policy Institute* (2003).
11. Mulligan, D., Schwartz, A. Your Place or Mine?: Privacy Concerns and Solutions for Server and Client-Side Storage of Personal Information. Computers, Freedom, and Privacy Conference (1999), 81– 84.
12. Oberholzer, M. Strategische Implikationen des Ubiquitous Computing für das Nichtleben-Geschäft im Privatkundensegment der Assekuranz. *PhD thesis*, Basel University, Switzerland (2003).
13. Siegele, L. How about now? A survey of the real-time economy, *The Economist*, 362(8257):3–18, January 2002.
14. Stallman, R.M. Can You Trust Your Computer? *Free Software, Free Society: Selected Essays of Richard M. Stallman*. GNU Press (2002), 115–118. Available from: http://www.gnu.org/philosophy/fsfs/rms-essays.pdf.