PriPAYD: Privacy Friendly Pay-As-You-Drive Insurance

Carmela Troncoso[1] George Danezis[1] Eleni Kosta[2] Bart Preneel[1]

[1]K.U.Leuven, ESAT/COSIC,
Kasteelpark Arenberg 10,
B-3001 Leuven-Heverlee, Belgium.
firstname.secondname@esat.kuleuven.be

[2]K.U.Leuven, ICRI,
Sint-Michielsstraat 6,
B-3000 Leuven, Belgium.
eleni.kosta@law.kuleuven.be

ABSTRACT

Pay-As-You-Drive insurance systems are establishing themselves as the future of car insurance. However, their current implementations entail a serious privacy invasion. We present PriPAYD where the premium calculations are performed locally in the vehicle, and only aggregate data arrives to the insurance company, without leaking location information. Our system is built on top of well understood security techniques that ensure its correct functioning. We discuss the viability of PriPAYD in terms of cost, security and ease of certification.

Categories and Subject Descriptors

C.3 [**Computer Systems Organization**]: Special-purpose and application-based systems—*Real-time and embedded systems*; J.7 [**Computer Applications**]: Computers in other systems—*Consumer products*

General Terms

Design, Security, Legal Aspects

Keywords

Privacy, Car Insurance, Pay-as-you-drive

1. INTRODUCTION

Insurance is a large percentage of the cost of owning a car. In order to decrease costs for both owners and insurers, insurance companies have developed Pay-As-You-Drive (PAYD) (or Pay-Per-Mile) models. In contrast to the current pay-by-the-year policy, customers are charged depending on where and when they drive, instead of a fixed amount per year. For each kilometer that a car is driven the statistical risk of accident, depending on the road and the time of the day, is calculated and translated to personalized insurance fees. To make possible the estimation of the monthly premium by the client, the fares applied for the billing are made public beforehand. Depending on the policy followed by the insurer these fares can be: price per type of road, price per time of driving,. . .

Pay-As-You-Drive insurance models are hailed as the future of car insurance due to their advantages for users and companies [22, 42]. First, the insurance fees applied to each user are fairer than the ones in the pay-by-the-year scheme, as customers are only charged for the actual kilometers they travel. Customers can also reduce their monthly bill by choosing cheap itineraries or by just not using their car. This makes vehicle insurance affordable for lower-income car users (e.g. young people) or for people who wish to have a second vehicle. Second, PAYD policies are socially beneficial, as they encourage responsible driving, decreasing the risk of accidents, which in turn saves money for users and insurers (aside from saving lives). Finally, PAYD has an environmental benefit, as it discourages driving, hence reduces energy consumption and pollution emissions. Due to all these advantages, PAYD insurance policies are supported by motorist associations like the National Motorist Association [9] and the American Automobile Association [8]; and they are being widely developed by insurance companies all over the world like Norwich Union [38] (UK), Aioi [1], Toyota [12] (Japan), Hollard Insurance [19] (South Africa), etc.

Although PAYD insurance seems to have many advantages, its current implementations involve an inherent threat to user's privacy. The full information used for billing (the time and position where the car was) is gathered by a black box in the car, and transferred to the insurance company (and, in some of the cases, to a third company providing the location infrastructure). In this model, the insurance company has the ability to track any of its users with ease and precision.

We propose PriPAYD a privacy friendly scheme, where the premium computation is done in the car's black box, and only the minimum information necessary to bill the client is received by the insurance company. We provide an overview of our architecture, where well understood techniques are combined to give assurance to the user that the insurance company does not get more information than necessary, while granting him (or a judge in case of dispute) access to all the data. Our techniques also permit easy management and enforcement of the policies by the insurer. A similar case has been pointed by LeMay et al. in [18] in the field of electric metering, where a remote billing system that preserves user privacy against the electric company and eavesdroppers is proposed.

The rest of the paper is organised as follows: Sect. 2 presents a survey on the current implementations of PAYD

WPES'07, October 29, 2007, Alexandria, Virginia, USA.
Copyright 2007 ACM 978-1-59593-883-1/07/0010 ...$5.00.

policies. In Sect. 3 we give a detailed description of our privacy friendly scheme. We discuss the feasibility of our scheme and compare it with the previous work in Sect. 4. Finally, we conclude in Sect. 5.

2. A SURVEY OF CURRENT IMPLEMENTATIONS

Pay-As-You-Drive plans are offered by many insurance companies around the world, gathering the data in a variety of ways. Depending on how privacy invasive they are, we can distinguish three type of policies. Some of them do not imply any breach of privacy since the data about the amount of kilometers traveled (no location information) needed to compute the premium, is provided only once a year from a fixed location. The second type, despite not recording location information, collects data in geographically distributed points, which allow the insurance to estimate the movements of the user. Finally, the last model collects GPS data to track all the car's movements. In the rest of the section we present real-world systems that fit in these three categories (Table 1 offers a summary.)

The first type of systems, that are the least privacy invasive, are also the least numerous. Examples of this model are Corona Direct (Belgium) [13] and Polis Direct [14] (Netherlands). They only use the data from the car odometer, obtained in annual vehicle inspections, and per-kilometer premiums are calculated by dividing current premiums by the current policy maximum annual kilometers. This is hardly privacy sensitive, since it does not reveal where the car has been over time.

WGV [40], a German insurance company, offers a different scheme that does not break into the user's privacy either. They collect the car speed and they use GPS to locate the road where the car is driving, but with the sole purpose of checking that speed limits are being observed and without saving the location data. When the speed limit for a given route is exceeded, the user collects "negative" points that will have repercussions on his final premium.

In the second group of PAYD policies we can find models such as the one from Aryeh [6], an Israeli company, that calculates premiums monthly using mileage data. This data is collected by receivers at fuel pumps (offered by the country's largest petroleum company) through small wireless transmitters in vehicles. Another example is Nedbank [28], from South Africa, that offers "Pay-Per-K", where every time that the vehicle is refueled using a Nedbank card, an odometer reading will be recorded on the transaction document. A similar scheme is deployed by Aioi [1], a Japanese insurance company. They install a device in the car that records the odometer value, the car condition and the time. This information is collected by receivers placed by the road, thus allowing to approximate the car trips. This data is sent to the insurance central database for billing purposes and, also, to the database of the company that provides the data collectors.

Two companies, Progressive Casualty Insurance (US) and AVIVA (Canada), supply devices (TripSense® [32] in the first case and Autograph® [10] in the second) that can be easily connected by the user to the OBDII (On Board Diagnostics II) port of the car. This device collects: trip start and end time, miles driven, duration of trip, number of sudden starts and stops, and time and date of each connection/disconnection to the OBDII port. This data can be seen by the client in a personal computer and can be exchanged for discounts if sent to the insurer. In Germany, a similar device is used by Swiss Re [33], and a variant is adopted by DVB Winterthur [41] giving the user the opportunity of exchanging data for discounts. They collect, through the use of GPS, the route information of the vehicle, from which they infer the kilometers traveled, the speed and the behaviour of the user.

In the US, Pay&Go [29] is a third party for "privacy friendly" gathering of PAYD data. Their claim their device only records the time of driving (neither the location nor the speed), but this information is collected by intermediate stations which give strong hints about where a vehicle has been over time.

Many patents propose models that also fit in the second group. For instance, the one registered as ES2108613 [30] suggests a model where the car is fitted with speed sensors and accelerometers, and also collect data from special devices on the roadside. The gathered data is sent to the insurer via "data collectors" present in garages and petrol stations.

Finally, we can find models that base their premium calculations on continuous collection of data, that leads to the gravest invasion of customers' privacy. Many insurance companies have chosen to follow this model, for example: Hollard Insurance [19] performs a PAYD insurance based on Skytrax GPS service (supplied by Mobile Data [36]) in South Africa. This GPS module is installed in the car, records all the data (position, time, speed,...) and stores it in a server, where the client can access it from the Internet. This is privacy-wise the worse model, as not only the insurance company gets the client's data, but also a third party has access to it.

Progressive Insurance Corp. (US) [31], registered the US Pat. US5797134 [21], in which they propose to gather the necessary data for billing (where, when and how much the car has been driven) using GPS. At the end of each month, a GSM phone fitted in the car (which is part of the policy) reports to Progressive all driving patterns. They go even further, proposing the collection of data that would give an idea of the safe operation of the vehicle by the driver such as speed, safety equipment used (seat belt, turn signals,...), rate of acceleration, rate of braking, or observation of traffic signs. This scheme is closely followed by Norwich Union [38] in the UK, owner of European patent (EP) number 0700009 [30]. They base their policy on less data as they only consider the time of the day, the type of road (more or less dangerous) used and the number of kilometers driven. Nevertheless, Norwich Union keeps all the location and timing data collected from the GPS signal, that is transferred to their central database via GSM. We find a very similar scheme in Austria, where Uniqa Group [39] offers an insurance that uses a GPS device in the car to collect location data and transmits it once a day, via GSM, to the base station of the company. The data is then used to calculate the monthly premium of the client.

"SaraFreeKm" is offered by the Italian insurance company SARA [7], in which customers install a GPS device (supplied by Movitrack [25]). The company calculates from the satellite data the client premium based on the actual kilometers driven. Also in Spain insurers provide PAYD policies. MAPFRE [23] offers the installation of a black box in the car

Table 1: Current PAYD implementations.

Company	Country	Method to gather data	Method to transmit data	Known Patent	Third Party	Privacy invasive
Polis Direct [14]	Netherlands	Odometer read yearly	Read by the mechanic	-	No	No
WGV [40]	Germany	GPS	User gives the info	-	No	No
Aioi [1]	Japan	Device in car	Radio or GSM	WO2005/08365 [2]	Yes	Medium
Aryeh [6]	Israel	Odometer sent at refueling	Unclear	-	Fuel Company	Medium
NedBank [28]	South Africa	Odometer sent at paying with bank card	Included in bank transaction	-	No	Medium
Progressive Casualty TRIPSENSE [32]	US	Device in car and software	User send info through internet	-	No	Medium
Toyota [12]	Japan	Device in car	Radio or wired communications	JP002259708 [27]	Garage	Medium
Hollard Insurance [19]	South Africa	GPS	GSM network	-	Skytrax [36] (Mobile Data)	Yes
iPAID$^{\text{TM}}$ [20]	Canada	Device in car	USB key, Bluetooth, GPS	-	Themselves	Yes
MAPFRE [23]	Spain	Full GPS data	GSM network	-	Yes	Yes
Norwich Union [38]	UK	Full GPS data	GSM network	EP0700009 [30]	No	Yes
Pay&Go [29]	Israel US	Device in car	Use of intermediate stations	-	Themselves	Yes
Progressive Auto Insurance [31]	US	Full GPS data	GSM network	US5797134 [21]	No	Yes
Sara [7]	Italy	GPS	GSM network	-	Movitrack [25]	Yes
STOK [37]	Nederlands	GPS	GSM network	-	Themselves	Yes
Uniqua [39]	Austria	Full GPS data	GSM network	-	No	Yes

that records: kilometers driven, type of roads used, average length of the trips, time of the day, regions in Spain where the car has been driven, average speed, and percentage of night hours. In order to obtain these data, the company counts with a third party that receives the raw data and performs the processing.

There are also third parties that offer insurers the necessary technology in order to implement GPS PAYD policies like STOK [37] (Nederlands). This company offers a system to be installed in cars, as well as the means to transfer the information collected to the insurance company and present it to the client (while having the data themselves). A more developed technology is the one introduced by iPAID$^{\text{TM}}$[20] (Canada); they present a GPS tracking solution for driving data collection. It records when, where, how far, how fast and how aggressively a vehicle was driven on the in-vehicle iPAID$^{\text{TM}}$unit. These data can be transmitted to the central server in a passive way (via a USB key, Bluetooth or wirelessly) or an active way (using the GSM network), which compiles it in statistics and trip logs, which the user can look up through the web. These statistics are also given to the insurance company in order to calculate the premium.

2.1 The abstract ‘Current Model’

We chose to model one of the most privacy invasive, data hungry PAYD model that is available today. We call ‘Current Model’ any system in which the data is collected by GPS, using a black box installed in the car, and then sent to the insurance company (directly or through an intermediary.) This model is a generalization of all the other models, meaning that less privacy invasive policies (such as those that only take into account odometer readings) can also be implemented using it.

The GPS-based pay-as-you-drive insurance is illustrated in Fig. 1(a). It works as follows: as the car is being driven, GPS data is collected by the insurance *black box.* The full data gathered is sent to the insurance company, who will do the accounting to obtain the client’s premium and send the bill by traditional post, together with a user friendly (reduced) version of the full GPS data. (This is very close to the Norwich Union [38] operating procedures.)

It is important to note that the correctness of the billing depends on the black box. For this reason both the customer as well as the insurer have stakes in its correct functioning, and incentives to game it to their advantage. To prevent malicious behavior in practice, the boxes are provided by the insurance company and should be protected using tamper evidence and tamper resistance techniques [3] making it hard for the car user to modify their behavior. Moreover, the car user receives a detailed bill that allows him to audit the trips contributing to his premium and legally challenge the premium if they do not correspond with actual car movements.

The 'Current Model' has the key advantage that is very flexible: the premium rates can be changed at the central database where they are calculated. However, such changes are restricted by the fact that the insurance policy should be predictable and easy to understand by the customer (which is required for a contact to be fair.) A second advantage of this model is that computation can be offloaded onto powerful servers having access to on-line up-to-date data sources. In the case of the Norwich Union policy [38] GPS data points (coordinates and time) have to be mapped to different road categories (more or less dangerous). Such classification from raw GPS data only requires access to an electronic road map and some computations to match the coordinates to a road type.

This model puts service providers (either insurers or third parties) in a business advantage position. With all the data collected, new services (traffic information, pollution information, . . .) can be offered to customers. It also allows providers to perform data mining to detect potential fraud.

However, the obvious disadvantage of the 'Current Model' is that it is privacy invasive, as the data collected by the insurance company is sufficient to track almost every movement of a car over time. The data is transmitted sometimes over third parties, such as the GSM[1] network or a third party location data provider. Once the location data has been transmitted the data subject has little control over it. This data could be stored or retained for long periods as well as used for other purposes than the ones it has been collected and, although Data Protection legislation may impose limits on what can be done with it, the penalties for breaching them are often very light.

3. PriPAYD: PRIVACY FRIENDLY PAYD INSURANCE

We present the PriPAYD architecture (see Fig. 1(b)) that follows closely the 'Current Model' with the exception that the raw and detailed GPS data is never provided to third parties. The main advantage of PriPAYD, is that the insurance receives only the billing data instead of the exact vehicle locations (thus cannot invade the user's privacy) while being sure he is receiving the correct data. The client can check that only the allowed data is getting in the insurance company database and the raw data is available for the client to check the correctness of the bill in case of dispute between user and insurer.

Before diving into the details of the scheme it is important to delineate our threat model. There is little point for our system to try to protect user's privacy beyond what road users already expect today. We assume that any third party adversary that has extensive physical control of the car will be able to track it (by simply installing their own tracking system.)

The objective of PriPAYD is to limit casual surveillance by the insurance company or any third parties (with limited physical access to the car), as well as preventing the aggregation of masses of location information in centralized databases. Fine grained location/timing information should be hard to obtain for any third party except the policy holder, who has the right to audit the bill and ensure its fairness. This protection still allows for surveillance of the drivers (in case they differ from the policy holders), but we are satisfied that no systemic surveillance risk is introduced beyond what is already possible today.

3.1 Description of PriPAYD

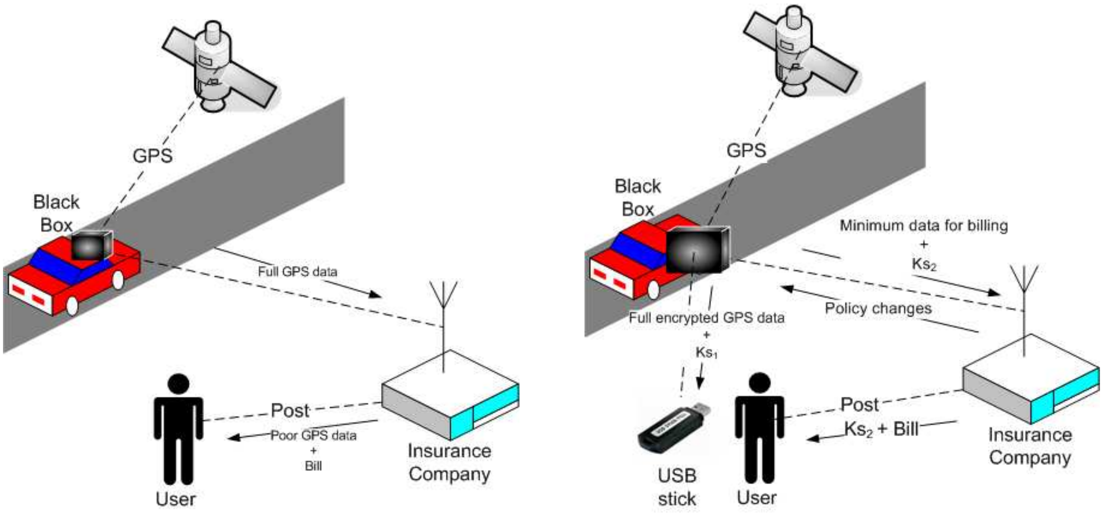
The key difference between PriPAYD and the 'Current Model' is that all computations transforming the GPS data into billing data are performed in the vehicle black box. The data involved in the calculation of the final premium are the number of kilometers traveled, the hour of the day, the road the user has chosen, and the rate per kilometer (hour and road type) given by the insurer (following the NU model [38]). To perform the conversion, maps have to be available to the device, and calculations have to be performed to match the coordinates with road types. These are no more complex than the operations already supported by any off-the-shelf commercial GPS navigation system.

The rates imposed by the insurer or other policy parameters can be initialized in the black box when installing it, and can be updated later in a trustworthy manner through signed updates. For the purposes of this work we consider that policies have a unique ID_{policy} that uniquely identifies the rates interpreted by the black box. A similar mechanism can be used to perform software upgrades (uploading new firmware to the black box) with identifier ID_{code}.

Once the premium for a period of time is calculated, the amount to be payed is sent in a secure way to the insurance company via GPRS, or even the cheaper SMS services (along with the current policy, ID_{policy}, and code version, ID_{code}). The data is signed using the black box key, and encrypted under the public key of the insurance company, in a special way (see Sect 3.2) that allows the policy holder to check that only the minimum billing information is transmitted.

To ensure that the black box is not acting maliciously in favor of the insurance company, we need to allow a car user or owner to audit the billing mechanism. For this purpose, we propose the use of an off-the-shelf USB memory stick. The data is recorded in an encrypted way on this token so that only the policy holder can access it, and it is signed by the black box to be usable as evidence. The symmetric encryption key is generated by the black box and provided to the policy holder in two shares: one written on the USB stick and the other relayed through the insurance company and delivered by post with the bill. A simple mechanism, such as pushing a button on the box for some time, allows the encryption key to be reset.

[1] GSM lawful interception interfaces could be used by the authorities to get access to the location data without the knowledge of the users or even insurers [15]. Given the legal void, we fully expect to see such attempts. Moreover, after 20 years the security of the original A5/1 and A5/2 GSM encryption algorithms has been degraded to an extent that production cryptanalysis on massive GSM traffic is within reach of many organizations.



(a) Current PAYD model

(b) Privacy-friendly PAYD model

3.2 The Security of PriPAYD

At the heart of PriPAYD we have a two level Bell-La Padula policy [11]: the confidential (high) level contains the sensors and records of the vehicle position and at the restricted (low) level we have the billing information. The only party that is authorized to access the confidential information is the policy holder, while the insurance company is only authorized to access the billing information. (Note that there is no restriction in the insurance company sending information up to confidential, i.e. policy or software updates.) In this context transferring billing information to the insurance company is an act of *declassification*, since the data at high is sanitized (only the amount of the final premium is sent) to not leak any information, and sent to low. The policy holder providing the detailed location records, as part of a dispute, is an even more radical act of declassification.

Three key security properties are required from the channel that transfers the billing data from the vehicle to the insurance company:

Authenticity. Only the black box can produce billing data that is accepted as genuine by the insurer or any other third party.

Confidentiality. Only the insurer and the car owner should be able to read the billing data transmitted.

Privacy. The policy holder should be able to verify that *only* the billing data is sent to the insurer.

A public key signature scheme [24] can be used to certify that the data has been generated and sent by the black box. The signature key in the black box is difficult to extract due to a custom tamper resistant solution [4] or established smart-card [26] technology. Public key encryption [24] can be used to encrypt the billing information under the public key of the insurer. There is no key distribution problem since the fingerprints of all public keys can be exchanged when the device is fitted.

We denote a message sent by the black box to the insurance company,

$$\mathrm{Enc}_{\mathrm{Insurer\ Key}}(D = (\mathrm{Data}, ID_{\mathrm{policy}}, ID_{\mathrm{code}}), \mathrm{Sig}_{\mathrm{Box\ Key}}(D)),$$

and note that the Privacy property, that allows the user to verify that only billing data is transferred, can also be enforced. Any signature scheme ($\mathrm{Sig}_{\mathrm{Box\ Key}}(\cdot)$) as well as public key encryption scheme ($\mathrm{Enc}_{\mathrm{Insurer\ Key}}(\cdot)$) are verifiable: the policy holder can be convinced that the encryption is correct by being given the randomness used to perform the encryption operation (in the detailed audit log.) The signature can then be verified to ensure it is correctly computed on D. Verifying these only requires the public key of the insurance and the verification key of the black box, that are public.

The task of verifying that no other information is contained in the messages is made difficult by the existence of *subliminal channels* [5, 35] (or covert channels) in signature schemes with the potential to leak information from a maliciously programmed black box back to the insurance company. Subliminal channels, as well as techniques to limit their capacity, have been extensively studied in the multi-level secure systems literature. PriPAYD implementations should either use signature and encryption schemes that are free from such channels, or estimate their capacity and keep it under a certain threshold [16].

A detailed log of all the vehicle's movements, and other audit information, should only be accessible to the policy holder. This is not a trivial requirement to fulfill since the black box and the policy holder need to share a symmetric key, unknown to any third party. We solve the key exchange problem by having the black box generate the symmetric key and sending two shares of it (using a secure secret sharing scheme [34].) The first part (Ks_1) is written to a USB stick along with the full encrypted log, and the second part (Ks_2) is relayed through the insurance company and received by the user as part of their bill in a sealed envelope. Both key parts are necessary to decrypt the detailed log of location data, and check its correctness. (Special software can be provided by any third party to re-assemble the key parts, decrypt, read and present the detailed location logs on any commodity computer.) Through this mechanism, we ensure that only the policy holder can access this data, as neither the insurer nor any person with direct access to the car (e.g. garage mechanics) will have access to the whole key.

Although the black box is trusted for correctness, if both shares of the key are stolen, the privacy of the client may be compromised. For that matter, the user can change the symmetric encryption key that he shares with the black box. He can do it directly on the black box (e.g. by pushing a button more than five seconds) which will then record the new Ks_1 on the USB stick, and send the new Ks_2 to the insurance company. When this happens, the key to decrypt any previous data is lost, and so is the past audit data. In effect the user can guarantee his privacy by deleting the detailed record at the cost of not being able to contest the correctness of past premiums (unless he had previously recorded a copy of these data).

4. DISCUSSION

4.1 Legal considerations

PAYD insurance based on the 'Current Model' transmits the full GPS data of the car's location to the insurance company for the calculation of the premium. As we have discussed, not all this data is necessary for the accomplishment of the wished purpose, namely offering PAYD insurance.

PriPAYDis based on the processing of only the relevant and absolutely necessary data, in compliance with the basic privacy principles of proportionality and data minimisation. For the calculation of the bill, the insurance company does not need to know the detailed route a car has followed, but only the number of kilometers it has covered and the type of road used. Similarly the exact time of a journey is not relevant, but in some types of PAYD insurance the approximate time of day is used (for instance between 2a.m. and 6a.m). The PriPAYDmodel allows the insurance company to calculate the bill using the aforementioned necessary information in the black box, while it avoids the collection of any other tracking data.

One implementation of PAYD has been criticised by the French Data Protection Authority (CNIL)[2]. The insurance company MAAF Assurances S.A. wanted to launch a new insurance policy for young drivers, according to which the latter would agree not to drive during the weekend at night or longer than two hours as well as not to exceed the speed limit. To check compliance with the policy the insurance company would collect data related to the car's location, speed, type of road, hours and driving duration and transmit them every two minutes. The CNIL refused its authorisation for the processing of the data. It argued that via the proposed system the insurance company would get information about violations of the speed limit. Such processing would involve sensitive data and would infringe Article 9 of the French Data Protection Act, according to which private entities are not authorised to process data relating to criminal offences. Although the data relating to offences is not considered as sensitive data in all Member States, the European Data Protection Directive allows the Member States to foresee derogations regarding them[3]. The CNIL further argued that the monitoring of data that reveals the movements of the car and consequently its driver with the exclusive aim the control of the respect of the contractual obligations of the driver infringes the principle of proportionality and the European principle of free movement of persons. PriPAYDon the other hand, does not allow the monitoring of the car and its driver and it limits the processing of the data to the absolutely necessary for the provision of the PAYD insurance, in full respect of the principle of proportionality.

The data collected in the car's black box is transmitted to the insurance company and to the other parties involved for the *provision of the service.* This defines the purpose for which the data is processed and any further processing must be compatible with it. In the 'Current Model' the abundance of collected data may tempt companies to further process them, even in ways that can not be considered compatible, to gain some business advantage. Full anonymisation of the data, would allow any processing, as it would not qualify as personal data, but it is very difficult, if not impossible. We note that as long as the data can be linked, with no excessive effort, to the original user, it cannot be processed for purposes that are incompatible with the original ones.

A crucial issue for the privacy offered by PriPAYDis the ownership of the recording of the black box; one of the principle goals of the model is that the company does not get data that are not necessary for the calculation of the insurance. The full content of the black box with the detailed information that can serve for the audit of the bill on behalf of the policy holder must only be accessible to them. It is therefore important that the contract between the insurance company and the policy holder clarifies that, even if the actual black box belongs to the company, the contained records belong exclusively to the user. This presupposes that it is feasible to erase completely any traces of the information that has been stored in the box (which might involve expensive certification of black boxes). In case this can not be guaranteed it is safer for the ownership of the box to be transferred to the user from the moment of its installation in the car.

The transmission of location data in the 'Current Model' can be achieved using the GSM network and as a result the mobile communications operator can record all the location data of the GSM device. In the 'Current Model' the mobile operator is offering a value added service, in the sense of a service that requires the processing of traffic or location data beyond what is necessary for the transmission of a communication or for the billing thereof[4], commonly known as a Location Based Service. The insurance company is the actual service provider and is obliged to inform the other contract party, prior to obtaining his consent, of the type of location data that will be processed, of the purposes and the duration of the processing and of the fact that the data will be transmitted to the mobile operator for the provision of PAYD insurance[5].

The insurance company has the obligation to delete any collected data after a period of time, when it is no longer necessary for the calculation of the insurance premium and after the end of the period during which the bill can be dis-

[2] CNIL, Délibération 2005-278 du 17 novembre 2005, portant refus de la mise en oeuvre par la MAAF Assurances SA d'un traitement automatisè de données à caractère personnel basé sur la géolocalisation des véhicules.

[3] Article 8 (5) data protection directive.

[4] Article 2 (g) ePrivacy directive.

[5] Such interpretation is favoured by the Article 29 Working Party, composed of the Data Protection Commissioners from the Member States together with a representative of the European Commission. See. Article 29 Data Protection Working Party, Opinion on the use of location data with a view to providing value-added services, adopted on 25 November 2005, 2130/05/EN (WP 115), p. 5.

puted[6]. A similar obligation exists for the mobile operator that has to delete the data after the provision of the service[7]. However the mobile operator falls under the scope of application of the data retention directive[8], according to which the operator has to retain specific categories of traffic and location data for a period between six months and two years and have them available for law enforcement purposes. It is to be pointed out that such an obligation does not exist for the insurance company, as the directive creates an obligation only for providers of publicly available electronic communications services or of public communications networks.

There are cases when the policy holder and the driver of the car do not coincide, as for rental cars or company vehicles. In such cases the driver should be informed about the presence of the black box and its functionality. Especially in the case of rental cars this should be explicitly mentioned in the car rental contract. When company cars are used by employees during their working hours, they should also be informed about the installation of the system. It is still an open question whether a company has the right to choose a PAYD insurance for a company owned car, which is used by an employee outside working hours. Current practices require the consent of both the employer and the employee.

4.2 Cost

PriPAYD does require more computations and mapping data in the black box than the 'Current Model'. Yet these are comparable to what current commercial GPS navigation systems do. Since the 'Current Model' already relies on tamper resistance for security no additional costs should be expected from this either.

Another source of costs is GSM communications and the PriPAYD model should be cheaper since less data is transmitted. Billing data can be aggregated to reduce those costs further. Updates containing new rates, maps and policies, can be pushed to the black box either through the GSM communications or during the servicing of the car.

The PriPAYD design keeps the trust infrastructure to a minimum, and particularly does not require a public key infrastructure. The identity infrastructure is based on the pre-existing relationship of the policy holder with the insurance company that is used as part of the key distribution mechanism. Hence there is no cost associated with either of these.

Finally, one has to take into account the cost of development and maintenance of the infrastructure. The technology and cryptography used is available off-the-shelf and developing PriPAYD should not be more expensive than the 'Current Model'. The additional engineering that is required for building a slightly more complex black box should be more than balanced by the reduced costs of the back-office systems, since they handle less, as well as less sensitive, data.

4.3 Strengthening Privacy

Some additional privacy concerns should be tackled as part of a real-world implementation.

One need to ensure that past location information can easily be deleted. We would advise implementers to never store encrypted GPS data from the audit record; and users to keep this, or key material, only on the USB stick to which they were written by the black box. This allows the user to easily destroy the data by destroying or deleting the USB stick. The downside of having a token that can be destroyed as an easy and intuitive operation (and a better paradigm for destroying the private data than electronic equivalents) is that, once audit records of the detailed locations have been deleted, it is difficult to challenge any bills that seem incorrect. This can be prevented by making a back-up of his records by simply copying the files to a computer or other storage.

A further concern is the use of GSM to transfer the data back to the insurance company. In our scheme the billing data does not contain any sensitive location information, but an active GSM device registered in the network does leak the cells the car is visiting. Hence it is prudent, to keep the GSM system powered down at all times except when transmitting. The transmission time and location must be chosen to minimize location leakage because of the GSM technology. Defining and using a known 'home' location should easily address this concern.

4.4 Certification and Independent Monitoring

The key objective of our design is to *not require* a trusted black box to guarantee user's privacy. This is an important requirement: the black box is commissioned by the insurance company and the user has only a limited capacity to discern its functioning. Furthermore, independent certification of even simple devices (such as the black box described) is expensive and appropriate certification criteria have hardly been established.

Our design choice is to allow policy holders to have a full view of the output of the black box and to ensure that only the minimum billing information is transmitted. One option is to allow a device (again a USB mass storage device would be sufficient) to record all data sent between the black box unit performing the calculations, and the GSM subsystem that relays all the information back to the insurer. This solution is not invulnerable to a maliciously programmed black box that only reveals part of the conversation. On the other hand it makes certification easier, since only a trivial property needs to hold: that all data transmitted using GSM is also recorded on the auditing device. A second approach, that offers stronger guarantees, is to physically separate (and shield) the black box from the GSM transmitter, and link them with a recording device controlled by the user. This device would record all traffic, and allow the users to verify that the data transmitted only contains the billing information. Recording cables could be sold by multiple manufacturers, or provided by privacy advocacy groups or data protection authorities.

We note that trusting the black box is still necessary for correct billing, and that the encrypted audit trail can be used to check bills or dispute them. Without third-party certification it is impossible to ensure that the black box is not recording precise location data with the intent to provide them to a third party. Since such a box has no way of

[6] Article 6 (e) data protection directive.

[7] Article 9 ePrivacy directive.

[8] Directive 2006/24/EC of the European Parliament and of the Council on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC, Official Journal L105, pp. 54-63 (March 15, 2006).

transmitting the recorded data over the air, physical access would be required to extract the data, making it difficult to turn this weakness into a mass surveillance tool. This is a known open problem [17].

5. CONCLUSIONS

Pay-As-You-Drive insurance policies, due to their advantages, are bound to gain popularity or even dominate the car insurance market. However their most advanced current implementations show a fundamental disregard for the privacy of car owners, which might even slow or limit their deployment. Our survey of existing systems and practices sadly documents a move towards more, not less, privacy invasive systems.

PriPAYD is a system that can support the deployment of very fine granularity PAYD policies while also providing strong privacy guarantees. Its core security architecture is based on simple and well understood multi-level security components, that have been the subject of extensive study in the field of computer security since the nineteen-seventies. The PriPAYD architecture relies (as previous systems) on secure hardware for correct accounting, but privacy properties can be checked without relying on its correctness. This separates correct accounting from privacy concerns, allowing black boxes to remain fully under the control of insurance companies, while users can be sure that none of their location data is leaking. Our approach follows the paradigm of many security metering systems used for electricity or gas distribution that only record aggregate use.

There is no component or infrastructure required by PriPAYD that would make it much more expensive than current systems. One could in fact argue that in the long run running PriPAYD as any other privacy enhanced technology, is cheaper than privacy invasive systems. The costs of protecting private data stores is often overlooked in the accounting of costs, as is the risk of a single security breach leaking the location data of millions of policy holders. In addition, PriPAYD keeps sensitive data locally in each car, in a simple to engineer and verify system. Requiring off-the-shelf backend system to provide the same level of privacy protection to masses of data would make them, not only prohibitively expensive, but simply unimplementable.

Acknowledgements

This research work was partially supported by the IWT SBO ADAPID project (Advanced Applications for e-ID cards in Flanders), GOA Ambiorics and IUAP p6/26 BCRYPT. G. Danezis is funded by a research grant of the Katholieke Universiteit Leuven. C. Troncoso is funded by a grant of the Foundation Barrie De la Maza from Spain.

6. REFERENCES

[1] Aioi. `http://www.ioi-sonpo.co.jp/`.

[2] Aioi. Telematics insurance system.

[3] R. Anderson. *Security engineering.* Wiley New York, 2001.

[4] R. Anderson, M. Bond, J. Clulow, and S. Skorobogatov. Cryptographic Processors-a survey. *Proceedings of the IEEE*, 94(2):357–369, 2006.

[5] Ross J. Anderson, Serge Vaudenay, Bart Preneel, and Kaisa Nyberg. The Newton channel. In Ross J. Anderson, editor, *Information Hiding*, volume 1174 of *Lecture Notes in Computer Science*, pages 151–156. Springer, 1996.

[6] Aryeh. `http://www.aryeh.co.il/`.

[7] Sara Assicurazioni. `http://www.saraassicurazioni.it/`.

[8] American Automobile Association. `http://www.aaa.com/`.

[9] National Motorist Association. NMA's position on auto insurance. `http://www.motorists.org`, 1998.

[10] Autograph. `https://secure.avivacanada.com/autograph/product.php`.

[11] D.E. Bell and L.J. La Padula. *Secure Computer Systems: Mathematical Foundations and Model.* Mitre, 1974.

[12] Toyota Motor Corporation. `http://www.toyota.co.jp/`.

[13] Corona Direct. `http://www.kilometerverzekering.be/`.

[14] Polis Direct. `http://www.kilometerpolis.nl/`.

[15] Alberto Escudero-Pascual and Ian Hosein. Questioning lawful access to traffic data. *Commun. ACM*, 47(3):77–82, 2004.

[16] Virgil D. Gligor. *A Guide to Understanding Covert Channel Analysis of Trusted Systems.* National Computer Security Center, ncsc-tg-030 version-1 edition, 1993.

[17] Vanessa Gratzer and David Naccache. Alien *vs.* Quine, the vanishing circuit and other tales from the industry's crypt. In Serge Vaudenay, editor, *EUROCRYPT 2006*, volume 4004 of *Lecture Notes in Computer Science*, pages 48–58. Springer, 2006.

[18] Michael LeMay, George Gross and Carl A. Gunter and Sanjam Garg. Unified Architecture for Large-Scale Attested Metering. In HICSS '07: Proceedings of the *40th Annual Hawaii International Conference on System Sciences*, 2007.

[19] Hollard Insurance. `http://www.payasyoudrive.co.za/`.

[20] iPAID. `http://www.ipaid-insurance.com/`.

[21] McMillan Robert John, Craig Alexander Dean, and Heinen John Patrick. Motor vehicle monitoring system for determining a cost of insurance, August 1998.

[22] Todd Litman. Distance-based vehicle insurance feasibility, costs and benefits. Technical report, Victoria Transport Policy Institute, 2007.

[23] MAPFRE. `http://www.jovenesdesiguales.com/`.

[24] A.J. Menezes. *Handbook of Applied Cryptography.* CRC Press, 1997.

[25] Movitrack. `http://www.movitrack.it/`.

[26] D. Naccache and D. M'Raihi. Cryptographic smart cards. *IEEE Micro*, 16(3):14–24, 1996.

[27] Shigeru Nakagawa, Kenji Mori, Akira Shinada, Katsuhiko Nunokawa, Hiroaki Okajima, and Makoto Sasaki. Vehicle insurance premium calculation system, on-board apparatus, and server apparatus, March 2001.

[28] NedBank. `http://www.nedbank.co.za/`.

[29] Pay&Go. `http://paygo-system.com/ShamirWeb/`

PublicSite/default.html.

[30] Salvador Minguijon Perez. Individual evaluation system for motorcar risk, December 1997.

[31] Progressive. http://www.progressive.com/.

[32] TripSensor Progressive Casualty Insurance. https://tripsense.progressive.com/.

[33] Swiss Re. http://www.swissre.com/.

[34] Adi Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, 1979.

[35] G.J. Simmons. Subliminal communication is easy using the DSA. In T. Helleseth, editor, *EUROCRYPT 1993*, volume 765 of *Lecture Notes in Computer Science*, pages 218–232. Springer, 1993.

[36] Skytrax. https://www.skytrax.co.za/index.asp.

[37] STOK. http://www.stok-nederland.nl/.

[38] Norwich Union. http://www.norwichunion.com/pay-as-you-drive/.

[39] Uniqa. http://www.uniqa.at/uniqa_at/.

[40] WGV. http://www.wgv-online.de/index.htm.

[41] DBV Winterthur. http://entry.dbv-winterthur.de/.

[42] Fayyaz Zahid and Craig Barton. Pay per mile insurance. Technical report, Davenport University, 2004.