PRIVACY AND DISTANCE BASED CHARGING FOR ALL VEHICLES ON ALL ROADS

Stefan Eisses[1], Wiebren de Jonge[2] and Vincent Habers[3]
1. Rapp Trans NL – Get ID B.V., P.O Box 2976, 1000 CZ Amsterdam, The Netherlands, Phone +31 6 45696864, stefan@getid.nl
2. Vrije Universiteit and Tip Systems B.V., The Netherlands, wiebren@cs.vu.nl
3. Rapp Trans NL – Get ID B.V., The Netherlands, vincent@getid.nl

ABSTRACT
Privacy concerns for a distance-based charge for all motor vehicles on all roads are of a more serious nature than for existing implementations of electronic tolling or congestion charging systems. Privacy legislation favours On-Board Aggregation over Central Aggregation solutions. As the impact of either concept on costs, risks and other aspects is not sufficiently clear yet, both types of options still have to be kept open and investigated further.

KEYWORDS
Privacy, distance-based road user charging, centralized vs. distributed processing, thin-client vs. thick-client, on-board vs. central aggregation, electronic fee collection, GNSS/GSM-based tolling.

1 INTRODUCTION
The European Commission regards road user charging as an appropriate means to fairly allocate the costs of road usage, including the so-called external costs. Whether considered fair or not, road user charging on a large scale is likely to lead to a reduction of fuel consumption, a smaller impact on the environment and less traffic congestion.
The magnitude of these effects depends on various factors – not in the least the height of the tariffs. An ultimate form of road pricing is distance-based charging for all motor-vehicles and every distance traveled, with tariffs that may depend on time, location[1] and vehicle class. Throughout this document we will use the acronym DBCAVAR (Distance-Based Charging for All Vehicles on All Roads) to refer to this form of road pricing. DBCAVAR may substitute existing fixed taxes on vehicle ownership and purchase.

Due to the enormous number of road segments to cover, solutions requiring substantial roadside infrastructure to determine the amount of road usage – as in existing DSRC-based systems – are less suited in case all roads are to be charged. Systems based on 'autonomous' On-Board Equipment (OBE) – using e.g. satellite-based positioning and some form of wireless communication to transfer usage details – are a better answer to the problem.

It should be noted that no country[2] has yet introduced DBCAVAR. The UK and Dutch governments have expressed more or less concrete intentions to introduce DBCAVAR, but

[1] 'Location' may include 'type of road'.
[2] The German LKW-Maut and Swiss LSVA system are distance-based and make use of autonomous OBE, but the charges apply only to heavy goods vehicles.

certainly not earlier than 2012. Apart from the political difficulties, there are technical challenges to overcome and a considerable investment to be made. If costs or risks seem to be too high, a positive decision cannot be expected on a short term. An important issue is the protection of privacy of the road users.

The authors have been involved in road pricing projects for the EU, in the Netherlands, the UK, Sweden and Slovenia. For the Dutch distance-based charging project 'Kilometerheffing' (2001) they were involved in the privacy concept definition and security architecture. The first author currently assists the Dutch Ministry of Transport in the market consultation for DBCAVAR in the Netherlands.

2 WHAT IS DIFFERENT ?

Privacy protection is to be taken care of in any EFC system, yet for most operational systems one or more of the following circumstances apply:

- The alternative of manual payment exists. The user is free to choose for passing anonymously by stopping and paying in cash, or to choose that some data relating to his journey are logged in exchange for the convenience of not having to stop.
- A non-charged road exists that is a reasonable (though possibly more busy) alternative to the toll road.
- The charge only applies to commercial or heavy goods vehicles. Driver privacy then is less a concern. First, because most trips – trips with personal cars – are not involved. Second, because trips with commercial or heavy goods vehicles usually are not made for personal motives but on behalf of a company, which often already gathers detailed information on the whereabouts of its vehicles/drivers anyhow.

Clearly, for the case of DBCAVAR considered in this paper these 'alleviating circumstances' do not apply and the impact on privacy could be enormous as complete travel patterns of all vehicles are involved. As a consequence, the issue of privacy protection is a key issue for DBCAVAR. It has to be addressed properly from the very beginning, i.e. already when defining the system concept.

3 PRIVACY LEGISLATION AND IMPACT ON ROAD USER CHARGING

The EU data protection directive 95/46/EC, [1], provides a legal framework for the treatment of privacy. EU countries must implement the directive in their national legislation and may have additional regulations on specific aspects identified in the directive. Western countries outside the EU mostly have similar privacy legislation largely grounded on similar principles. Referring to the EU directive, the most determining elements of privacy legislation for road user charging seem to be the following:

Article 7

Member States shall provide that personal data may be processed only if:

(a) the data subject has unambiguously given his consent; or
(b) processing is necessary for the performance of a contract to which the data subject is party [...]; or
(c) processing is necessary for compliance with a legal obligation to which the controller is subject; or
(d) processing is necessary to protect the vital interests of the data subject; or
(e) **processing is necessary for the performance of a task carried out in the public interest** or in the exercise of official authority vested in the controller or in a third party to whom the data are disclosed; or

(f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed, except where such interests are overridden by the interests for fundamental rights and freedoms of the data subject which require protection under Article 1.

Article 6

1. Member States shall provide that personal data must be:
   (a) processed fairly and lawfully;
   (b) collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes. [...]
   (c) adequate, relevant and **not excessive in relation to the purposes for which they are collected** and/or further processed;
   (d) accurate and, where necessary, kept up to date. [..]
   (e) kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed. [..]
2. […]

Article 2 Definitions

For the purposes of this Directive:
   (a) 'personal data' shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.
   (b) 'processing of personal data' ('processing') shall mean any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organization, storage, […], use, disclosure by transmission, dissemination […];
   […]

Thus, data about a certain vehicle's trips are personal data if they can be linked to the vehicle keeper or to any user of that vehicle. Such is, for example, the case if the vehicle's license plate number is available. This also applies if the identity of the vehicle keeper is not known by the processor. And to illustrate the concept of indirect identification: trip data not linked to any ID may even be judged personal data in case the data itself enable determination of the user, e.g. from the location where the car is parked regularly (e.g. at night).

It seems reasonable to assume that, if there is sufficient political support for DBCAVAR, clause (e) of Article 7 would be satisfied. Article 6 clause (c) is crucial and implies two principles[3] that determine the legitimacy of the processing of personal data:

- **proportionality**: the purpose must justify the impact of the implied processing of personal data;
- **subsidiarity**: intended 'processing' is not allowed if the purpose can also be achieved with less or no processing of personal data.

[3] These two principles stand out in e.g. clause (e) of Article 7 and clause (c) of Article 6. Note that proportionality and subsidiarity here must be understood in the context of law in general, and of privacy law in particular. (And thus **not** in the context of delimiting EU transnational influence.)

Subsidiarity requires a *comparison* of different solutions in order to select the solution with minimum processing of personal data.

However, a solution may be optimum from a privacy point of view, but at the same time have such a bad score on other relevant criteria – e.g. costs, fraud resistance, flexibility, risks and user convenience – that it should be considered 'impossible'. So, an important underlying problem is whether sufficiently objective weights can be assigned to these criteria in order to support a scrupulous decision on whether a solution is 'possible' or not[4].

Finally, the proportionality principle requires a judgement whether the objective justifies the impact of the selected solution. This will be the domain of politics rather than legislation.

4 TWO BASIC CONCEPTS FOR DISTANCE-BASED CHARGING

Before illustrating and discussing possible implications of privacy legislation (in Section 5), we describe two different approaches for implementing DBCAVAR: one based on back-office processing of position data, and one on performing more data processing on-board. Many different variations on these basic concepts are possible.

4.1 Basic concept 1: Central Aggregation

On each trip of a vehicle, its OBE keeps track of time and position. Periodically –once per certain time period or distance travelled – the OBE sends to a back-office via wireless communication a declaration – i.e., a signed message – containing a unique ID of the OBE[5] and all[6] positions with timestamps since the previous declaration. In back-office the charge[7] is calculated after reconstruction of the corresponding vehicle's declared trips by matching the declared positions to the road network. Payment may occur by deduction of the amount from a pre-paid or post-paid account.

To discourage abuse (turning off or sabotage of the OBE, shielding the antenna, etc.) some form of spot-checking is necessary. Spot-checking can be rather simple: register vehicle license plate numbers at many – including random – times and locations. In the enforcement back-office the spot-checking data – i.e., each triplet with license plate number, location and time – can be compared to declarations received from the vehicle in question. If a triplet is not covered by a reconstructed trip, the keeper of the vehicle will be called to account.

The concept described is an example of what is often referred to as a 'thin client' approach. Since detailed movement data are collected that can be linked to an individual, this concept clearly requires strong provisions for preventing abuse, i.e. for privacy protection.

4.2 Basic concept 2: On-Board Aggregation

On each trip of a vehicle, its OBE continually keeps track of time, tariff currently applicable, position and/or distance travelled. In the variation we will describe here[8], the OBE frequently

[4] Perhaps one could say that a concept is 'impossible' if the democratic decision to implement the measure with that concept – judging benefits versus costs and other consequences – is negative.

[5] In principle, one can use any ID that can be uniquely linked to the OBE (or a Trusted Element inside) or the vehicle in question. For privacy protection, an OBE identification number is preferable to the vehicle's license plate number (cf. Section 5.2).

[6] In a practical implementation some data reduction would be applied. This does not affect the discussion above.

[7] Time of travel, location, and vehicle characteristics clearly all can be parameters.

[8] For example, we do not treat another interesting option, in which each entry is immediately broadcast via a transmitter with a short range (say, 100 meters) to the outside world in the direct environment of the vehicle.

– i.e., once per some relatively small unit of time or distance, or at each designated road segment – computes the usage (in terms of distance per tariff or costs) and performs an 'irreversible' registration of this usage. This can be seen as an internal declaration ('micro-declaration') to a secured element of the OBE of which the data storage can only be influenced through a limited set of defined operations. For example, this element may keep a distance counter for each tariff that can only be incremented - never decremented or reset. Periodically – e.g., once per month – the OBE sends to the back office an electronic declaration containing the OBE ID (cf. footnote 5) and the current readings of the distance counters. Then, the back-office can compute the charge due since the previous declaration and can take care of billing and payment. The OBE may also keep a detailed travel log accessible only by the vehicle keeper. To be useful as evidence in case of a dispute (e.g. about malfunctioning equipment) with the operator of the scheme, the entries in this log may be signed by the secured element. For enhancing fraud detection and prevention, the OBE may also keep a secured log of anomalies detected by one or more of its elements.

Spot checking may occur by interrogation of OBE from the roadside (e.g. via DSRC) at many – including random – times and locations. It can be checked whether the OBE is/was operational and functioning correctly during the spot check and – optionally – some short period before. Registration of the license plate number and/or the OBE ID is only necessary if the spot-check reveals a defect or possible case of fraud.

This concept is sometimes referred to as a 'thick client', or 'intelligent OBE' as proponents prefer to denote it. Clearly, this concept requires less processing of personal data than the Central Aggregation approach.

5 DISCUSSION

To arrive at a solution that satisfies all privacy requirements, a stepwise approach is sensible.

First, it should be determined if the objective can be realized without processing personal data (see Par 5.1). Only if this proves not to be feasible, personal data can be processed. In that case a solution shall be derived that allows the processing to be the 'minimum' to realize the objective (see Par. 5.2). As suggested in Section 3, the solution has to be a *realistic* as well. This inevitably introduces other criteria of assessment, and may eventually lead to a 'second best' solution in terms of privacy (see Par. 5.3).

Having defined the appropriate minimum, the requirement of proportionality has to be satisfied. This aspect is not elaborated any further in the discussion below.

Assuming that the selected concept and architecture is judged to be in accordance with the above, still various measures have to be taken in the implementation of system and organization that guarantee proper handling of personal data during operations. This is briefly discussed in Par. 5.4.

5.1 Can anonymity be preserved ?

The first question to be answered is: "Is it 'reasonably' possible to achieve the objective (i.e., to implement DBCAVAR) while preserving anonymity?" If the answer would be a clear 'yes', the subsidiarity principle would require following such a route. In the following we concentrate on giving users the possibility to participate anonymously, at least as long as they

correctly pay for their usage and do not commit fraud[9]. Clearly, in case of DBCAVAR for every motor-vehicle involved the amount due for its use during each and every period must be registered and paid somehow. Thus, one challenge is to keep the payer anonymous.

This may be realized if pre-paid accounts[10] are used: the operator/payment collector has received the payment in advance and has no credit risk. Even if prepaid accounts are used, it should be noted that still an account ID has to be linked to an OBE ID that identifies the usage data from a particular vehicle. In the following, we concentrate on the fixed relationship between an account ID and a payer (usually the vehicle keeper and/or user) that in case of anonymous payment still would exist logically; however, it would simply be unknown by the collecting party or any other party except the payer/user.

The anonymity depends on keeping this relationship secret. Topping up the account balance by bank transfer, direct debit or credit card transaction would normally disclose the identity of the payer[11]. To keep the secret, one would have to implement such options as reloading via scratch cards (cf. prepaid GSM, using activation via phone or internet) that can be purchased for cash or reloading via machines that accept cash payments. Such options tend to be inconvenient for the user and very costly to operate. But in a situation where the volume is limited and the majority of users would voluntarily choose more efficient and convenient types of payment, the effects on overall costs could be acceptable. Note that data protection agencies and civil right watchers would argue that the user shall not have to pay for privacy – not in financial terms, nor in additional effort. But some compromise may be found here.

Still, the provided anonymity for the prepaid account is vulnerable as the relation between account ID and payer or number plate (indirectly identifying a vehicle keeper) may be disclosed by:

- A roadside enforcement spot-check in case a picture of the license plate of the vehicle is taken. Depending on the concept, see Sections 4.1 and 4.2, all passing vehicles may be subject to license plate number registration or only suspect ones. Either way, once the license number is linked to the OBE's ID and the account ID, anonymity may be 'lost forever'[12] or restored only by installation of a new OBU or Trusted Element.
- In case the user wishes to make specific inquiries or raise complaints, disclosure of the link between user and account ID and/or OBE's ID may be hard to avoid.
- Detailed charging data may identify a user indirectly (cf. Section 3). E.g. the location where a vehicle regularly is parked during nights may often reveal (the home address of) the vehicle user, thus breaking the link's confidentiality.

To conclude, anonymous participation in DBCAVAR is costly and inconvenient, but may be offered as an option. The anonymity offered will always be vulnerable and will not eliminate privacy concerns completely.

---

[9] Clearly, perpetrators must always be identifiable.

[10] In principle, a pre-paid account can be implemented 'on-board' or 'centrally-held' as for prepaid GSM. The centrally held prepaid account is the most realistic option in the given context.

[11] One may argue that the link is not disclosed , since the payer is not necessarily the vehicle keeper and/or user. However, in practice the majority of payers will be (identical to) the vehicle keeper and/or user.

[12] Obviously various technical and procedural measures can be taken to limit the risk that such data are disclosed to other persons or other processes. This discussion is merely intended to illustrate the difficulty of realizing a system that preserves anonymity of users.

5.2 Finding a minimum of processing of personal data

The issue is to find a 'realistic minimum' of processing of personal data. First, we will assess the On-Board Aggregation concept. In the extreme form described in 4.2, only the amounts due for each period (which are unavoidable; see Section 5.1) are reported[13]. As the privacy impact of the declarations decreases with decreasing reporting frequency[14], minimizing 'processing' for this part/aspect comes down to maximizing the size of the periods as much as possible within the limits imposed by the required level of fraud resistance. A minimal set of personal data will further include a realistic minimum of spot-checking data.

With Central Aggregation it is difficult to defend that the data collected are 'minimum', as the data reported reveal all trips in detail. As suggested before, one should separate the trip reconstruction domain as strictly as possible from the domain where the mapping between OBE identification numbers and license plate numbers is kept, and from the billing and payment domain, where the link between OBE ID and account ID (and possibly account holder) is known[15]. Furthermore, it must be prevented that any of the parties or their employees are linking travel patterns to individuals. Processing the raw data (almost) only in a fully automated way may help to limit undesired access to these data considerably.

5.3 Determining a *realistic* concept with minimum processing of personal data

The obvious conclusion is: if privacy protection would be the only criterion, then On-Board Aggregation would clearly be the concept that must be chosen. The balance may however still tip in favor of a Central Aggregation solution, if On-Board Aggregation solutions would appear to have an unacceptable score on other criteria deemed vital to meet the objectives for a DBCAVAR implementation. This may lead to a judgement that the On-Board Aggregation concept is 'infeasible' to realise.

Let us look at what proponents of Central Aggregation bring forward in favour of Central Aggregation. They argue that the following serious disadvantages are inherent in On-Board Aggregation solutions:

- Higher costs of investment due to additional functionality required in the OBE
- Complex and inflexible enforcement
- High risk of critical bugs due to the complexity of distributed software
- Difficulties in software lifecycle management, since clients are not always online and since wireless data bandwidths and latencies may be insufficient to guarantee successful updates within a given time frame
- Difficulties in geographic and tariff data management if map matching is to be performed in the OBE, for similar reasons as above.

It should be noted that this opinion is still only qualitative and not shared by the entire road pricing community, which also includes proponents of On-Board Aggregation[16]. It should also be noted that emerging technologies may alleviate the claimed problems over time. Given the fact the DBCAVAR is not deployed anywhere yet, it seems too early to prove either side wrong.

---

[13] Slightly more detail in the declarations, e.g. reporting counters per tariff category, hardly changes the overall impact on privacy.

[14] E.g., a usage total over a whole year provides far less information than a total for each single day.

[15] Note that full separation does not offer full protection. See the example at the end of Section 5.1.

[16] On-Board Aggregation has not only disadvantages, but also advantages compared to Central Aggregation.

5.4 The story does not end with a good choice of concept

It seems worth mentioning that the impact of privacy legislation does not end with choosing the concept enabling a 'realistic minimum' of personal data processing. The information processing shall be designed and operated in such a way that data are only distributed on a need-to-know basis, shall be protected against unauthorised disclosure of data and shall incorporate measures to minimize possible consequences of unauthorized disclosure.

Furthermore, measures are to be taken to maintain the quality of data stored on individuals, to inform persons involved on the purpose of the registration, to enable them to inspect data concerning them and adequately handle appeals that data are inaccurate, etc.

6 CONCLUSIONS

Privacy concerns for a distance-based charge for all vehicles on all roads (DBCAVAR) are of a higher order than for existing implementations of electronic tolling or congestion charging. An attempt has been made to apply the European privacy directive to the context of DBCAVAR.

Privacy should be taken into account starting from the basic concept of the system. The best solution for privacy – no processing of personal data at all, and thus full anonymity - cannot be realized in practice for DBCAVAR. The directive further requires that processing of personal data must be reduced to the minimum required for meeting the objective. However, if the second best solution – a DBCAVAR system with minimal processing of personal data – would appear to be too costly and risky, any government will refrain from implementing it. In that case the concept may be regarded an *unrealistic* solution, and 'next best' concepts in terms of privacy can be considered.

Two basic system concepts have been discussed: 'Central Aggregation' and 'On-Board Aggregation'. On-Board Aggregation allows a minimal processing of personal data and is therefore preferable from a pure privacy point of view. However, the score on other important criteria, such as costs and risks, is still uncertain as no real-life implementation of ABCAVAR has been realised yet.

Until more evidence has come available, both approaches should be kept open and further investigated.

7 REFERENCES

[1] **Directive 95/46/EC of the European Parliament and of the Council** of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, *Official Journal of the European Communities*, No L. 281 (23 November 1995), pp. 31-50.