Data Protection

# SAFETY IN NUMBERS

## ROAD PRICING BEYOND 'THIN' AND 'FAT'

**BART JACOBS and WIEBREN DE JONGE may have found a method for failsafe fraud resistance and optimal privacy protection in road pricing**

**A new approach to time, distance and place-dependent road pricing can offer, both with 'thin' and 'fat' on-board equipment (OBE), far greater privacy protection and fraud-resistance. It allows the hiding of the amount of usage between two spot-checks and an enforcement approach that does not rely on physical protection of OBE against traffic data manipulation.**

Most countries currently have a road tax for cars, involving a fixed charge that does not depend on the actual road use, but that may depend, for example, on the car's weight and engine category. Toll charges for roads, bridges or tunnels introduce a certain level of usage dependence, but only at a limited number of specific locations. Several countries are considering to go one step further and to replace these (road tax and toll) charges with a new wide area traffic pricing system in which road charges depend on the actual distances driven and possibly also on the time of day and on location. Such time-distance-place road pricing is not only considered to be fairer than flat charges, but also allows targeted congestion and pollution reduction by applying a higher tariff per kilometre for busy areas or road segments and for environmentally un-friendly cars. Furthermore, it can solve the (fuel) taxing problems associated with the transition to plug-in hybrid and fully electric vehicles.

Global navigation satellite systems, like GPS or Galileo, usually form the basis for such road pricing. The idea is that cars will be equipped with OBE for registering their successive locations and transferring relevant information to the traffic fee service provider (TSP) or traffic fee charger (TC) for billing and checking purposes.

### Fraud-resistance and privacy

It is obvious that fraud-resistance is important and also that the OBE resides in a possibly hostile environment. In general, it is rather unwise to fully rely on physical (ie, hardware) protection measures, since these are never perfect and are always subject to an arms race. Furthermore, attacks may include the switching off of power, screening the OBE from the satellite signals or feeding it false signals, for example. Therefore, spot-checks are always needed to verify whether the OBE in practice really functions appropriately. To be really effective, these spot-checks should be performed by surprise at random locations and times, and also passively (without two-way communication), because otherwise future networked cars may warn each other instantaneously and fully automatically.

Privacy protection is also a real concern in road pricing. Any system making it feasible to get detailed whereabouts of every individual car raises serious concerns about surveillance and control (Big Brother), possible abuse and personal security. Broad acceptance simply requires that privacy protection is dealt with adequately and convincingly, in a transparent manner that people can trust. Proper privacy protection must be built into the architecture of the system and cannot consist of just procedural measures, since such measures can be ignored or changed easily.

> "Broad acceptance simply requires that privacy protection is dealt with adequately and convincingly, in a transparent manner that people can trust"

### Thin and fat on-board equipment

There are two main approaches to satellite-based road pricing: 'thin' and 'fat'. Here, OBE is called thin if the usage calculation occurs outside the vehicle and fat if it occurs inside the vehicle.

In a ('conventional') thin approach the OBE registers the car's location and passes this information on to the TSP's back-office, say every X minutes or hours. Based on this information, the TSP calculates the usage and can send bills to individual car owners, say every three months. Spot-checking can simply be based on observations, such as videos or photographs, that are made by surprise - or even secretly - and that prove a vehicle's presence at a certain time and location. These proofs can be compared (later) to the location data transferred by the OBE. In case a discrepancy is observed, a fine may be imposed. This approach is simple and robust, but too weak on privacy protection.

In a ('conventional') fat approach the OBE must store its own location data safely and tamper-free, possess an up-to-date tariff map (or tariff category map), be able to perform the usage calculations and transmit regularly the results to the (TSP's or TC's) back-office.

Furthermore, its tariff map and extensive software need to be updated regularly (think of changes, bugs and vulnerabilities). Hence, fat OBE is much more complex than thin OBE. Spot-checks are also much more complex, since they (a) are based on real-time two-way (request-response) communication; (b) require swift mutual authentication of OBE and inspection device; and (c) must somehow verify whether usage is registered - and will later be reported - correctly. In view of this, 'conventional' fat OBE requires more and better physical protection. On the positive side, fat OBE allows proper privacy protection, with decentralised processing and storage of the privacy-sensitive location data.

### The best of both worlds

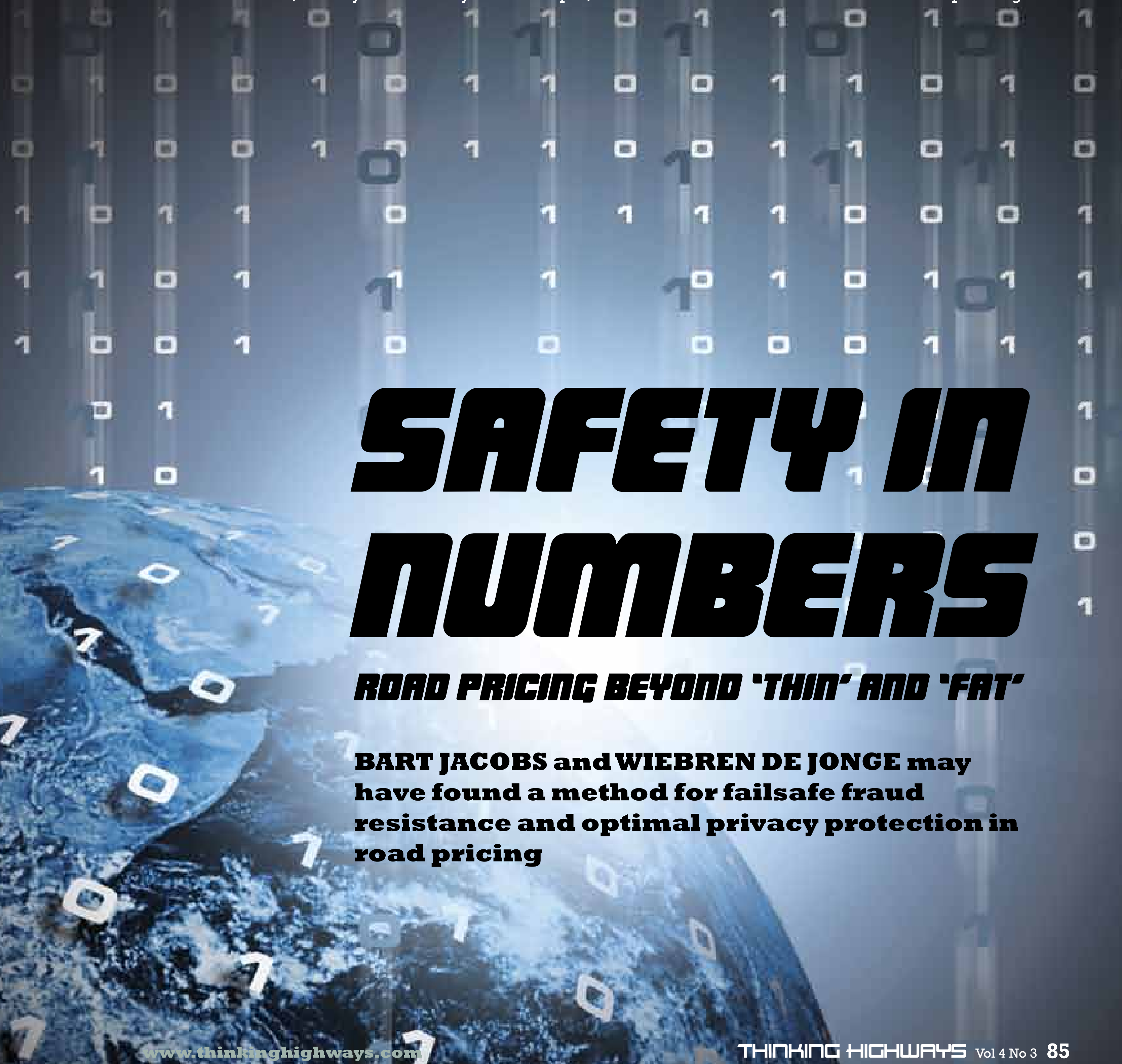
Except for its serious privacy problems, the 'conventional' thin approach is actually very good. In contrast, the 'conventional' fat approach can score much better on privacy protection, but is considerably weaker on properties related to fraud-resistance (see a later section). Our novel approach addresses both issues while allowing simple spot-checks and both fat and thin OBE. An elaborate explanation is given in an earlier paper[1]. Below we only sketch some main lines.

### Keeping things private - even with thin OBE

In order to make (both the thin and fat varieties of) our new approach privacy-friendly as well as fraud-resistant, we require that the OBE regularly - say, every X minutes or hours - 'commits' itself to the car's location data, without revealing the content. This can be done, for example, by simply transferring a hash value (see Fig. 1 and 2). The location data themselves can remain under full control of the car owner.

Only in case of a spot-check, some of the original (ie, 'pre-image') location data must be produced (see Fig. 2). Thus, only a very small subset of all location data needs to be revealed. Furthermore, we just mention that it can be made impossible to deduce (from 'commits' received) when the vehicle is driving and when not. Usage calculation based on pieces of traffic data can be anonymous (since it does not require a vehicle-ID) and under user control, in many different ways. For example, usage can be calculated inside the OBE or - after transferring the data from the OBE - by the car owner's PC or by third parties, which do not have to be trusted by the traffic fee charger. Again, 'non-revealing commits' are used, this time to commit to the results of usage calculation[1]. Again, accuracy can be verified with spot-checks[1].

Our approach is in many respects (among which, spot-checking) quite similar to the 'conventional' thin approach. However, there are a few prominent differences. In case of a 'conventional' thin approach the OBE 'commits' to location data by transferring all location data, together with absolute timestamps and an identification (of the vehicle or OBE), to one TSP that the user must choose from a list of TSP's cooperating with



- and approved or trusted by - the traffic fee charger. Thus, this TSP receives all details of all trajectory parts travelled, including vehicle identity.

In our approach - and particularly also in its thin variety - this is different. Our approach uses 'non-revealing commits' and gives the user full freedom in choosing any party or parties for the usage calculations. Furthermore, these calculations do not require supply of absolute timestamps, nor of any identification. Therefore, the user may choose to 'anonymise' the location data before anonymously distributing them - and thus the calculations - bit by bit over many different calculation services. Alternatively, if the user fully trusts one particular calculation service (eg, on his/her own PC) to protect his/her privacy, then (s)he can use that one for all trajectory parts. In either case, a party that is not fully trusted by the user, will receive no privacy-sensitive data on whereabouts, or very little (which is true for enforcement units).

> **"Users can freely choose whether - and, if so, to which parties - they are willing to reveal more information"**

### Parallel protection - even with fat OBE

In 'conventional' fat approaches, the enforcement by spot-checking crucially depends on proper, very long-lasting physical protection of the OBE. If this protection is breached, then the spot-checking will fail and the road pricing system may collapse, since changing or adapting the physical protection of the OBE in all cars may take months. In short, the physical protection of 'conventional' fat OBE is crucial and strictly necessary.

Our observation is that this is not true for our new approach, nor for a thin approach. Their simple observation-based spot-checking can be made effective even if the OBE has no physical protection at all against traffic data manipulation by a hostile driver, or if such protection is breached. Hereto, one only has to prevent the drivers from determining the time and location of the majority of relevant observations within the limited time available before the corresponding traffic data must be 'committed' to. For example, one can choose the parameter X, which determines the frequency of 'committing', and organise the observation process both in such a way that a sufficient percentage of the observations will remain undetected for a period of sufficient length (which will never exceed X minutes or hours).

If physical OBE protection against traffic data manipulation is not strictly necessary but present, it works in parallel with (ie, additional to) the 'logical' protection offered by spot-checking and penalties. This key difference between fraud-resistance in 'conventional' fat on the one hand and in 'thin' or 'new' (including also 'new fat') on the other hand, is illustrated in Fig. 3a and 3b respectively, where an analogy is suggested between keeping up a certain protection level with logical and physical protection measures and keeping up a weight with two ropes. Note that the strength of the logical protection is variable, since spot-checking intensity can be adapted easily in relatively short time. The strength of physical protection is more or less fixed, since adapting this strength is much more difficult and time-consuming. Actually, the strength of physical protection degrades monotonically due to advances in technology.

### A sure success

Parallel protection offers much better operational continuity than serial protection: in case parallel physical protection is breached, the spot-checking still works and can relatively easily be intensified temporarily until the physical protection problems are solved. Further, parallel protection offers at least twice as much fraud-resistance as serial protection (or the same level of fraud-resistance at lower costs). It also offers more flexibility in designing and implementing fraud-resistance, since one can choose any combination (eg, a cost-optimal one) between the extremes of 'spot-checking only' and of 'physical protection only'.

In addition, it offers more operational flexibility, since intensifying spot-checking always raises the aggregate protection level (no limit is imposed by physical protection). Finally, only in the case of parallel protection can remote spot-checks be used to monitor the aggregate fraud-resistance level that is actually achieved, in terms of percentages of violations. For example, a traffic fee charger can monitor easily whether (and to what extent) various traffic fee service providers succeed in keeping fraud below a level agreed upon.

### Free competition and easier interoperability

Our approach allows for a minimal uniform infrastructure, dealing with the collection of hash values and requests for original data needed for spot-checks, on top of which many different (thin and fat) implementations may be offered by various commercial and non-commercial parties. Users can freely choose whether - and, if so, to which parties - they are willing to reveal more information, such as (part of) their location data or just trajectory travel durations realised.

Finally, we mention that our approach eases scaling up to an international setting due to its monitoring capability, its easy and robust observation-based enforcement and its hash values, which can be exchanged between different countries - or collected centrally - without endangering privacy. For example, for enforcement purposes one could set up one European hash value collector that after each legitimate request supplies a specific hash value to an enforcement unit. TH

*For more information and details, please contact Wiebren de Jonge (wiebren@cs.vu.nl).*

*Bart Jacobs is Full Professor of Software Security and Correctness at the Radboud University Nijmegen and Wiebren de Jonge is Associate Professor of Computer Science at the VU University Amsterdam.*

[1] W. de Jonge & B. Jacobs (2009), "Privacy-friendly Electronic Traffic Pricing via Commits" in P. Degano, J. Guttman & F. Martinelli (Eds.): *FAST 2008, Springer Lecture Notes in Computer Science 5491*, pp143-161 (available at www.tipsystems.nl/files/ETPprivacy.pdf).



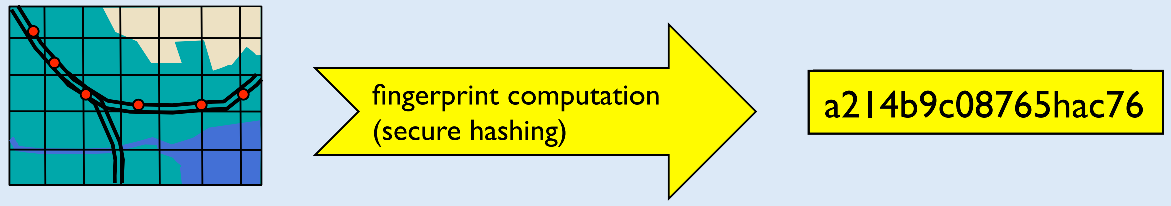


## Figure 1. Hash values used as 'anonymous' fingerprints

Secure hash functions are basic operations in cryptography that transform an input text to completely garbled output. For instance the (MD5) hash value of the sentence "traffic pricing is hot" is "4f5688b38731ca84e4fc4b13de692f7f". By changing only one character in the input, the output is totally different: the hash value of "traffic pricing is hit" is "a7f73aa6a9d4007ac3e638c6b8d76852". The important thing is that given a certain hash value (like "4f5..." as above): a) there is no feasible way to reconstruct its original input (also known as the 'pre-image') or any part of its input, while b) one can easily check whether some given sentence (eg, "traffic pricing is hot" or "… hit") must have been the original input, simply by checking whether the hash value of this given sentence is the same (ie, "4f5…"). In short, hash values can work as 'anonymous' digital fingerprints of their input.

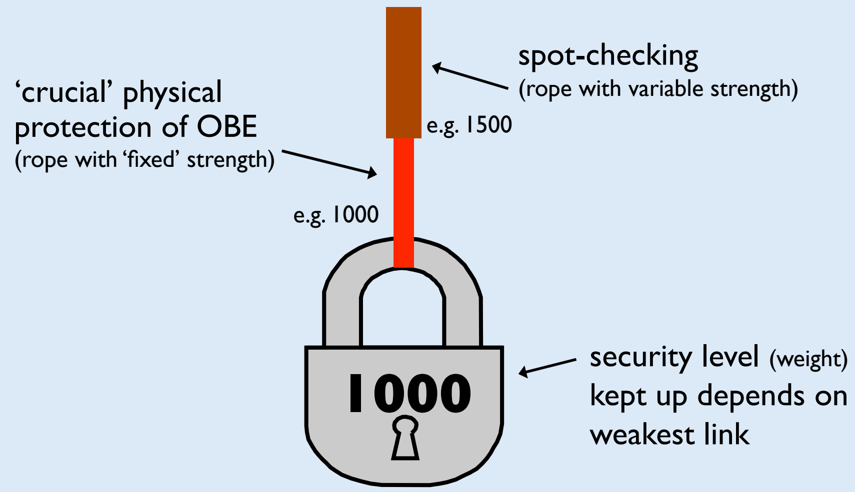
## Figure 2. Observation-based spot-checks and non-revealing commits based on hash values

If OBE regularly sends hash values of its location data to the traffic fee charger, it commits itself to these location data without revealing any content (see Figure 1).




Observation-based spot-checks then can be performed as follows. After a road-side observation and after receipt of the commit (ie, hash value of the traffic data that correspond to the time of observation), one demands the 'pre-image', that is, the original traffic data committed to. Then one verifies whether the data returned indeed produce the correct hash value and also cover the location of the observation correctly.

## Figure 3a. Serial protection adhering to 'conventional' fat approach




## Figure 3b. Parallel protection as possible with (thin and fat varieties of) our new approach