Pay-As-You-Drive applications

PRIVACY IMPLICATIONS AND POSSIBLE SOLUTIONS

Carmela Troncoso (KU Leuven-Cosic)

TSS Seminar – University Illinois at Urbana-Champaign – 3 Nov 2009

Acknowledgements: Josep Balasch (for some of the slides) and Fonds Wetenschappelijk Onderzoek Flanders (for funding my trip)

Outline

- Pay-as-you-drive: the concept
- Current implementations
  - Insurance
  - Road tolling
- Legal implications in the EU
- Possible solutions
- Conclusions

Pay-As-You-Drive: the concept

- Flat fees are not fair for everyone

- Users should pay depending on their use of the car and roads:
  - Long drives, high density roads, rush hours: higher fee
  - Sporadic use, second vehicle for weekends, young drivers with small salary: smaller fee

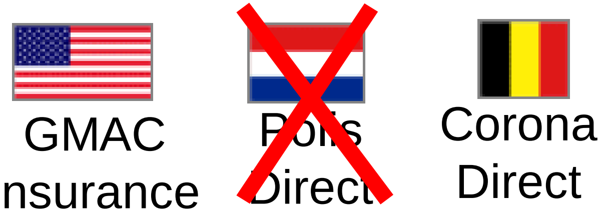
- Applicability:
  - Vehicle insurance
  - Road Charging (taxes)

Pay-As-You-Drive: pros

- Fair fees
  - For customer and companies
- Customer can “choose” his premium
  - Young drivers, second cars
- Social benefit
  - Less use of cars, responsible driving, less accidents, improve road mobility…
- Environmental benefit
- Business advantage position
  - Data mining
  - Additional services (LBS, targeted advertising,...)

Insurance: current implementations (I)

- **First Group** (Not privacy invasive):
  - data from odometer, recorded once/twice a year.



GMAC Insurance

~~Polis Direct~~

Corona Direct

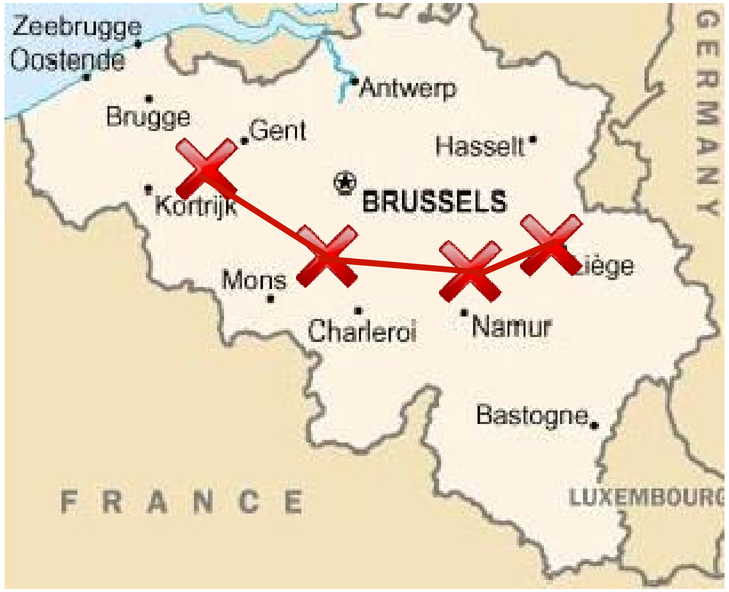
- Not viable
  - Costs of reading the car odometer high
  - Low benefits for client and companies

Insurance: current implementations (II)

- **Second Group** (medium privacy invasive):
  - data from geographically distributed points (gas stations, credit card payments,…)
  - change data for discounts
  - more information

Aryeh Nedbank Aioi AVIVA

Pay&Go (3rd Party) DVB Winterthur Progressive Casualty



Insurance: current implementations (III)

- **Third Group** (very invasive):
  - continuous collection of data
  - use GPS for location
  - use GSM for transmission (continuously or not)
  - more information
  - third parties

Hollard (Mobile Data)

STOK (3rd party)

iPAYD (3rd party)

Octo telematics (3rd party)

WGV

Progressive Insurance

~~Norwich Union~~

Uniqa Group

AVIVA

MAPFRE

Zeebrugge
Oostende
Antwerp
Brugge
Gent
Hasselt
BRUSSELS
Kortrijk
Mons
Liège
Charleroi
Namur
Bastogne
FRANCE
LUXEMBOURG
GERMANY

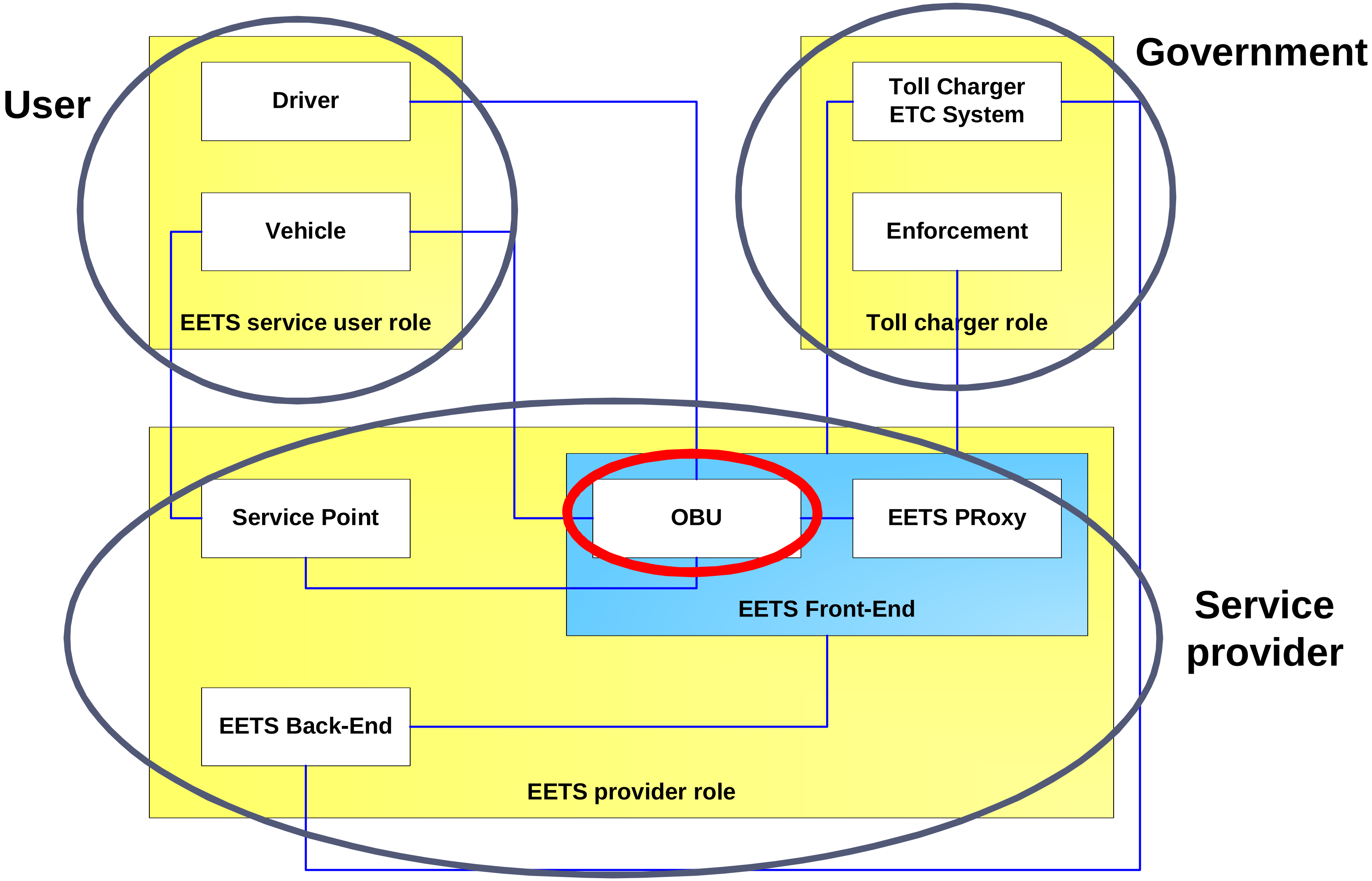
Road Tolling: EU EETS Decision

- European Electronic Toll Service
  - 6 Oct 2009
  - Coordinates exchange of information between Member States, to ensure the correct declaration of tolls
  - Defines the actors involved: EETS architecture
  - Defines the interfaces and capabilities
    - GNSS: Global Navigation Satellite System
    - DSRC
    - GPRS/GSM network

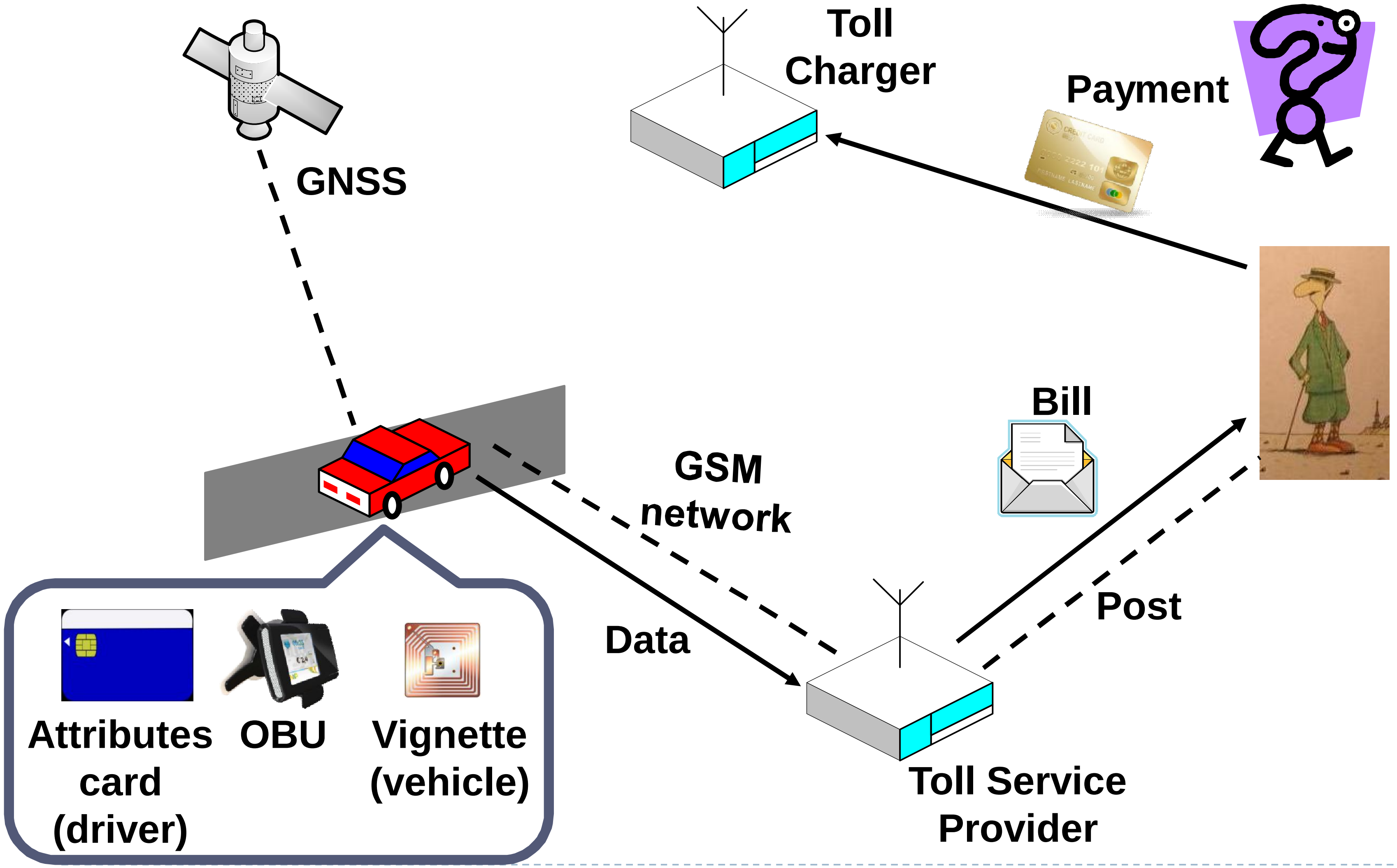
- Within **three** years for vehicles above 3.5 tons, all other vehicles within **five** years.

http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:268:0011:0029:EN:PDF

EETS architecture



Basic idea for the implementation



EETS Decision: Security and Privacy

- Protection against fraud/abuse for
  - Toll chargers
  - EETS providers
  - Users

- Protection of data under Directive 95/46/EC (Data Protection Directive)
  - Storage
  - Processing
  - Transfer

Data Protection Directive 95/46/EC

- Protection with respect to the processing or movement of **personal data**
- Two main actors:
  - **Data subject:** individual to whom the personal data refers
    - Right to access, rectification and deletion of all data processed about him.
  - **Data controller:** determines purpose and means
- Three principles
  - **Transparency**: data subject has the right to be informed when his personal data are being processed
    - Consent, or contract, or legal obligation, public interests, safeguard subject interest, safeguard controller's interest
  - **Legitimate purpose**: purpose must be specified and data may not be processed further
  - **Proportionality and minimization**: collect and process only adequate for the purpose for which they are collected

PAYD involves personal data?

- **Personal data:** any information relating to an identified or identifiable natural person ("data subject")
- Work/home is enough for re-identification [Golle and Partridge 09]
  - Given home and workplace (can be deduced from a location trace [Krumm 07]), then median size of the individual's anonymity set in the U.S. working population is 1
  - Inferences about driver [Iqbal 07]: personal, government, businesses
- Anonymization **very** difficult
  - What is anonymity?
    - property of an individual of not being identifiable within an anonymity set
    - probabilistic concept
    - cryptographic protocols (identity management) anonymity achievable but…
  - Traffic analysis -> anonymity extremely hard
    - Tracking techniques [Gruteser and Hoh 05][Haas et al 09]
      - Exploit spatio-temporal relations

Data protection does not "protect"

- Data security is hard to achieve:
  - Even if a system it Data Protection compliant…
    - Accidental leaks (Toyota, Norwich Union)
    - Insider attacks (Greek Mobile Phone Scandal)
    - Outsider attacks (10, 000 Hotmail passwords released by hacker – 6th Oct)
    - Today: medical data from 173 people found in Barcelona besides a container
  - … and once data is leaked, there is no control over it
    - Harvard Student database on BitTorrent 2008 (name, Social Security number, date of birth, address, e-mail address, phone numbers, …)
  - How long should data be kept?
    - Data retention
  - Liability
    - What if data is lost/tampered?
    - Need for certification

Mapping Data Protection to PAYD

- Data subject:
  - Car vs driver
  - Children vs parents
  - Employer vs employee
  - Insurance/Provider (box) vs user
- Data Controller:
  - Box vs Insurance company
  - Telecom provider
- Data minimization and proportionality:
  - GPS data reveal far too much information (e.g., speed, inferences)
- Secondary use of data (collides with legitimate purpose of the service)
  - Back to anonymization problem …

Third parties, covered by Data Protection?

- False sense of privacy
  - AVIVA in France, MAPFRE in Spain, ...
- Aggregation of data
  - Larger databases (Octo Telematics: 30 insurance companies / 858.775 users)
- Data security
  - More entities involved make securing data even more difficult
  - Data controller?

… and then Data Retention

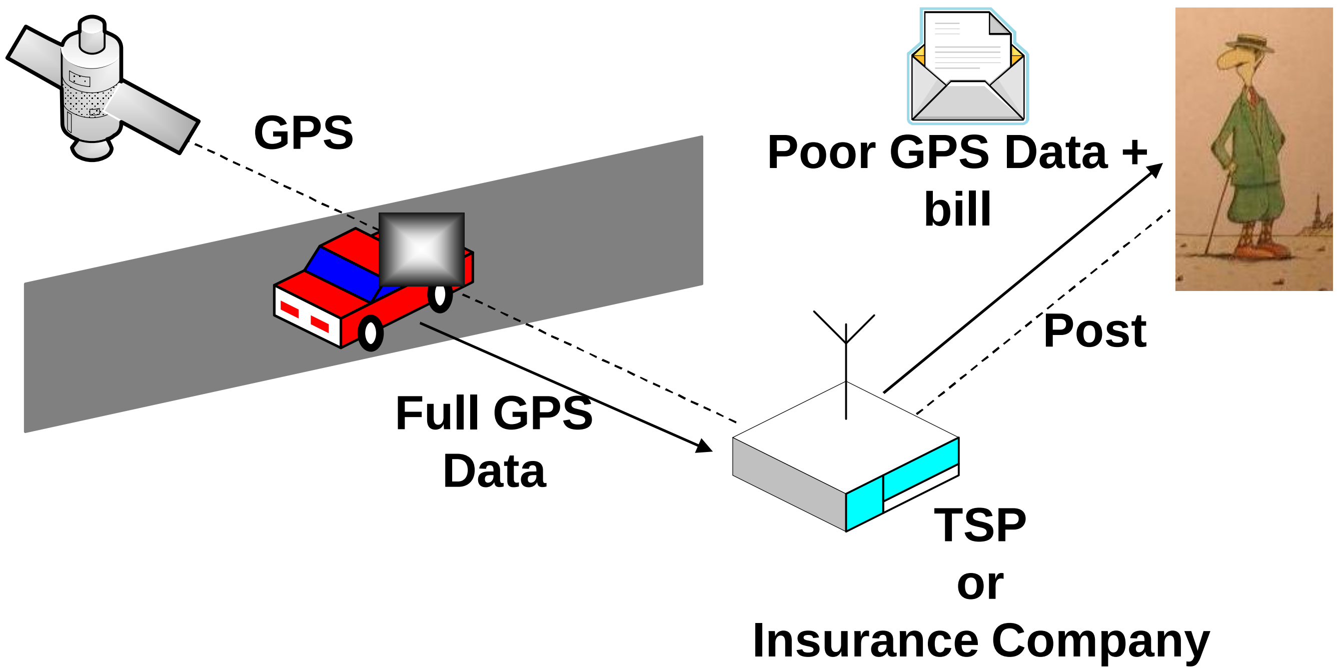
- Directive 2006/24/EC: retention of data generated or processed in electronic communications services or of public communications networks for enforcement
  - for a period of between 6 months and 2 years, necessary data:
    - **source** of a communication; **destination** of a communication; to identify the **date**, **time** and **duration** of a communication; to identify the **type** of communication; to identify the communication **device**; to identify the **location** of mobile communication equipment.
- GSM operator falls under Data Retention
  - And the insurance company or the Toll Service Provider?

Other legal issues

- Who is in charge of enforcement?
  - Toll Service Provider vs Toll Charger
    - Constraints on the collected data
- How will the tariffs be? Are dynamic fees legal?
  - Constraints on the implementation
- Is traffic congestion further processing of the data?
  - The data is collected for tolling…
- Other applications in the OBU?
  - eCall

Straightforward implementation

- OBU + GPS + (third party) + transmit



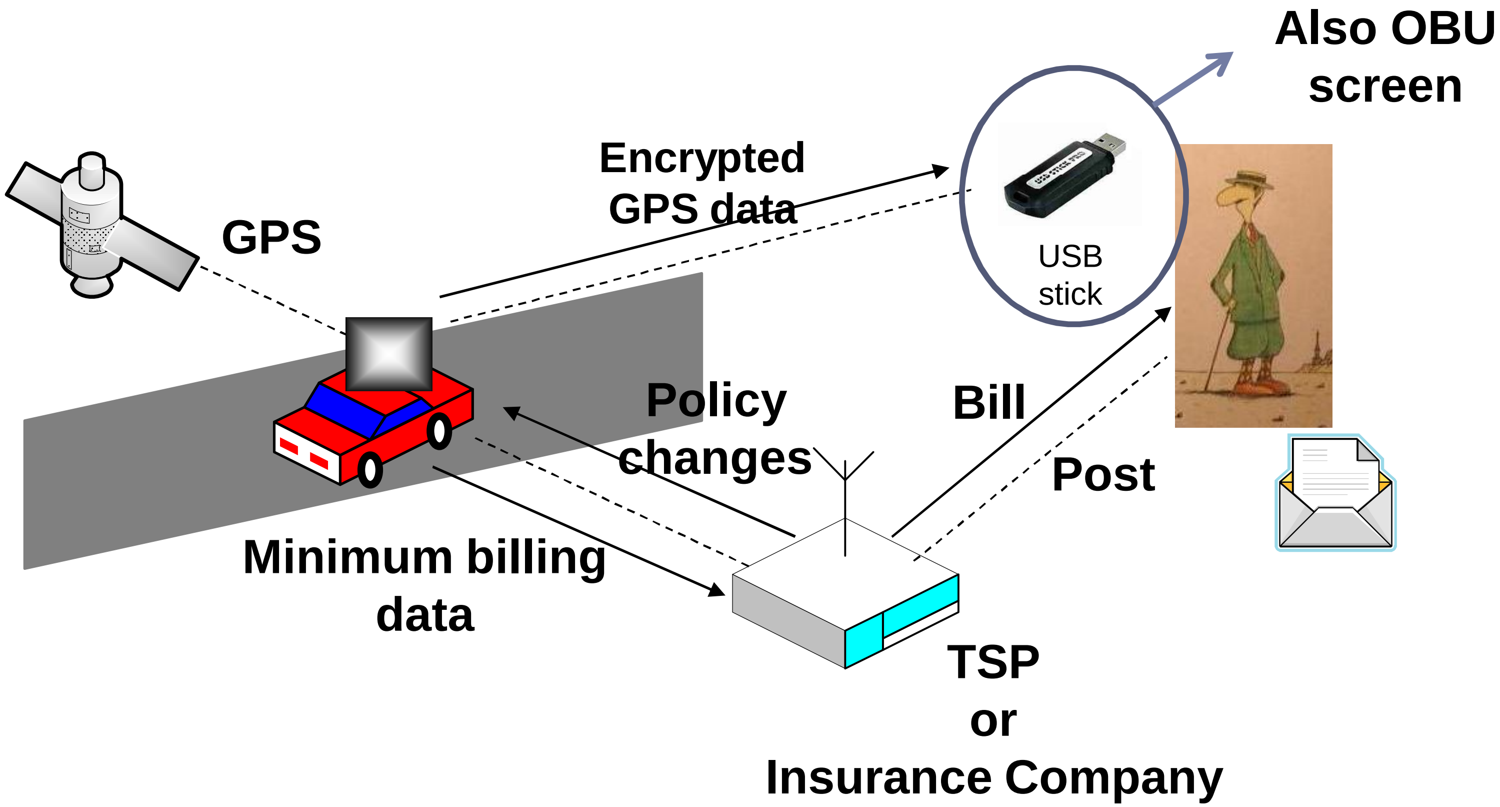
Straightforward implementation

- Flexible: any fee is possible
- Easy computation
- Easy updates
- Enforcement: use data mining
- Business advantage: data mining and new services

- **Privacy invasive: tracking**
- Upstream transmission of data
- Third parties (legal implications)

PriPAYD model [Troncoso et al 07]

- GPS + OBU (computation) + transmit billing



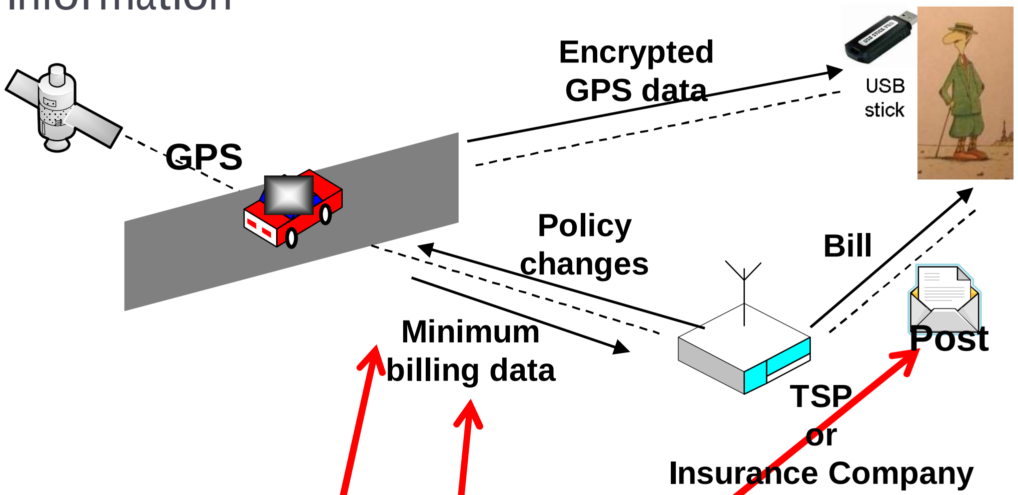
PriPAYD

- 🟩 Privacy friendly
- 🟩 Easy computation
- 🟩 Small upstream transmission
- 🟩 Third parties do not carry personal data

- 🟥 Difficult to update
  - 🟥 Large amount of vehicles
  - 🟥 Driving into another country (in Europe is easy...)
  - 🟥 Digital maps cannot be partially updated
- 🟥 Less flexible
- 🟥 Downstream transmission of data
- 🟥 Difficult enforcement

The security of PriPAYD

- Two-level Bell-LaPadula
  - high: complete position (and others) records
  - low: billing information



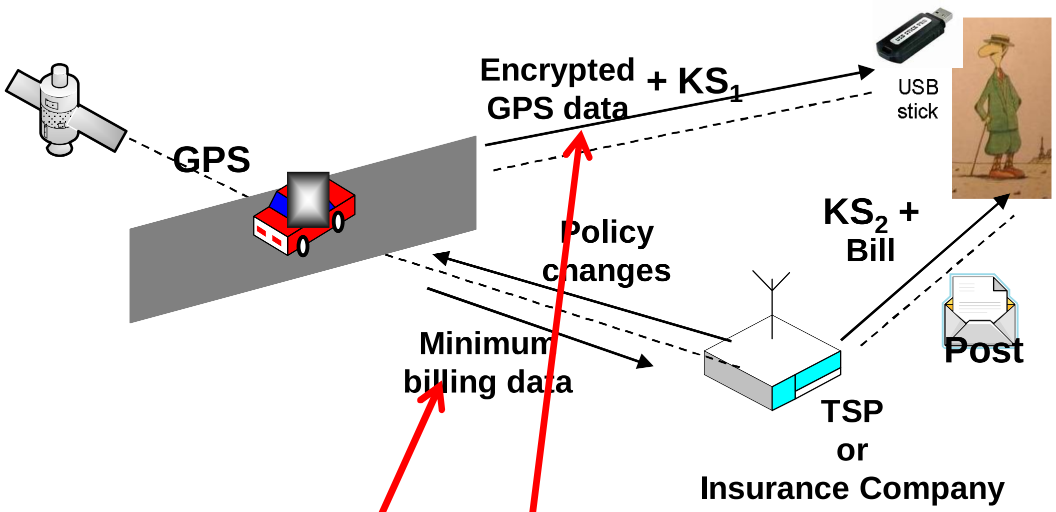
- **Authenticity**: data comes from black box

  Signature scheme (box should be tamper resistant)
- **Confidentiality**: only insurer and customer read billing data

  Public Key Encryption

  $\mathtt{Enc}_{\mathtt{InsKey}}\ (\mathtt{D=(TS,\ Data,\ ID}_{\mathtt{policy}},\ \mathtt{ID}_{\mathtt{code}}),\ \mathtt{Sig}_{\mathtt{BoxKey}}\mathtt{(D))}$

The security of PriPAYD



- **Privacy**:
  - only billing data transferred, avoid *covert channels*

    Signature schemes free or limited
  - logs only accessible to customer

    Symmetric key between box and customer:

    KS_1 and data from black box through USB stick

    KS_2 relied through insurer

    Possible change but loose contest ability

Cost: invasive vs friendly

- More computation in the black box:
  - commercial GPS,
  - tamper resistance is already in the straightforward implementation

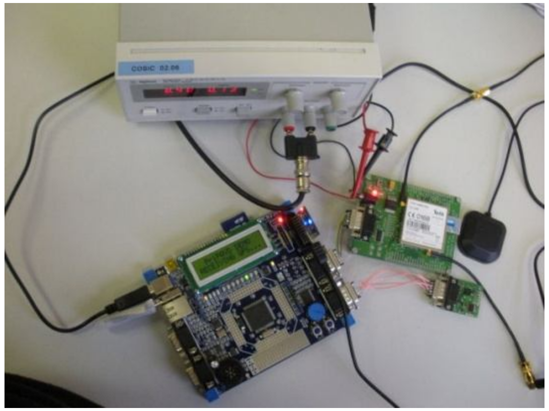
- Cheaper communications:
  - aggregate billing data (even SMS)

- Minimum trust architecture:
  - no PKI (relationship user – insurer/government)

- Same development cost:
  - off-the-shelf
  - more engineering
  - But… back-office simpler (no personal data)

Our prototype [Balasch and Verbauwhede08]

- Components
  - NXP LPC2388 processor (ARM7TDMI architecture)
    - Not the most powerful in the market
  - Telit GM862-GPS
  - External memory (SD Card) for the insurer‘s policy, digital road maps (OpenStreetMap), and encrypted GPS data
- Achieves real time computation
- Tested in 1h trip around Leuven
- Cost: ~500€
  - Production cost: ~50€
  - Less features needed
- Lots to do…



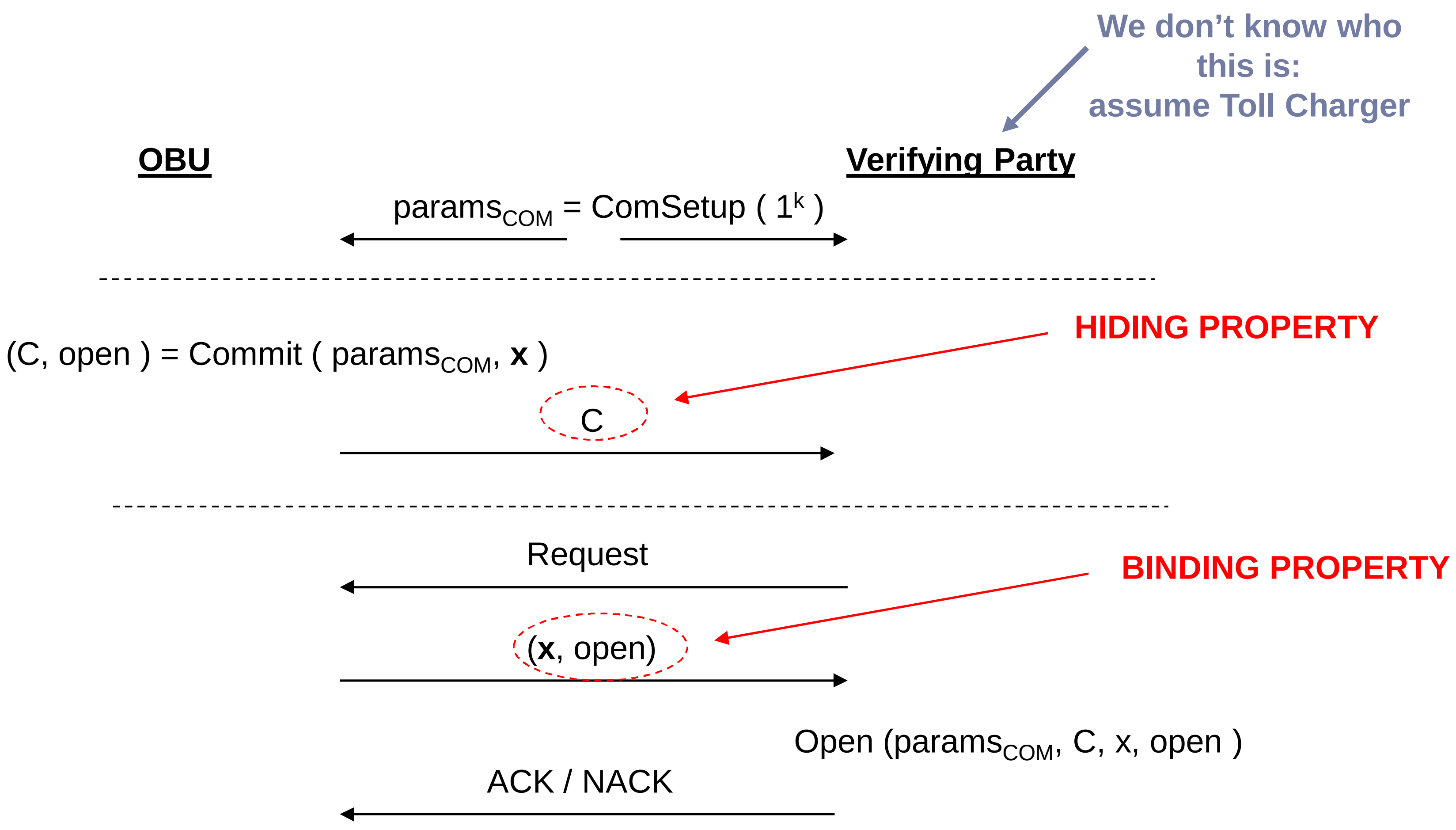
Enforcement

- Control mechanisms applied by the Toll Charger to detect misuse of the system
  - Law-enforcement

This can only be done by visual inspection or DSRC

- Includes...
  - 1) Detect vehicles with inactive OBUs
  - 2) Detect vehicles reporting false location data
  - 3) Detect vehicles using incorrect road prices
  - 4) Detect vehicles reporting false final fees
- ... in a privacy-friendly way
  - Minimize disclosure of location data

Non-Interactive Commitment Schemes



We don't know who this is: assume Toll Charger

OBU

Verifying Party

params_{COM} = ComSetup (1^k)

(C, open) = Commit (params_{COM}, **x**)

C

HIDING PROPERTY

Request

(**x**, open)

BINDING PROPERTY

Open (params_{COM}, C, x, open)

ACK / NACK

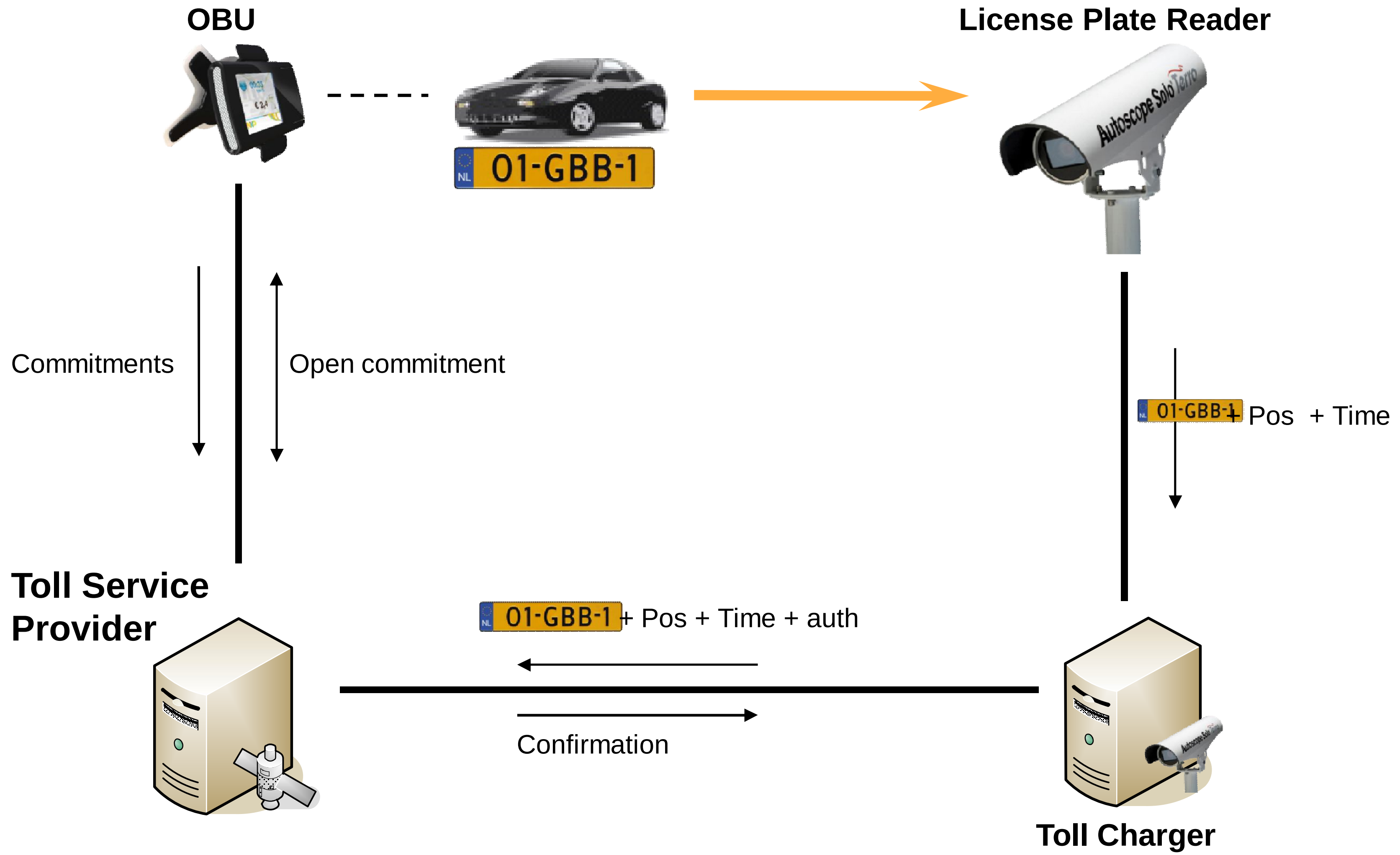
Mode of Operation

- Assumptions
  - Roads have assigned a price per Kilometer (or similar)
    - e.g. *Road price* = *f* (*type road*, *time day*)

	00u00 – 07u00	**. . . . .**	**22u00 – 00u00**
Highway	p_1	. . . .	p_2
Primary	p_3	. . . .	p_4
. . . .	. . . .	. . . .	. . . .
Residential	p_{n-1}	. . . .	p_n

  - OBU sends commitments based on distance
    - e.g. a commitment per Km (or similar)

How does it work?

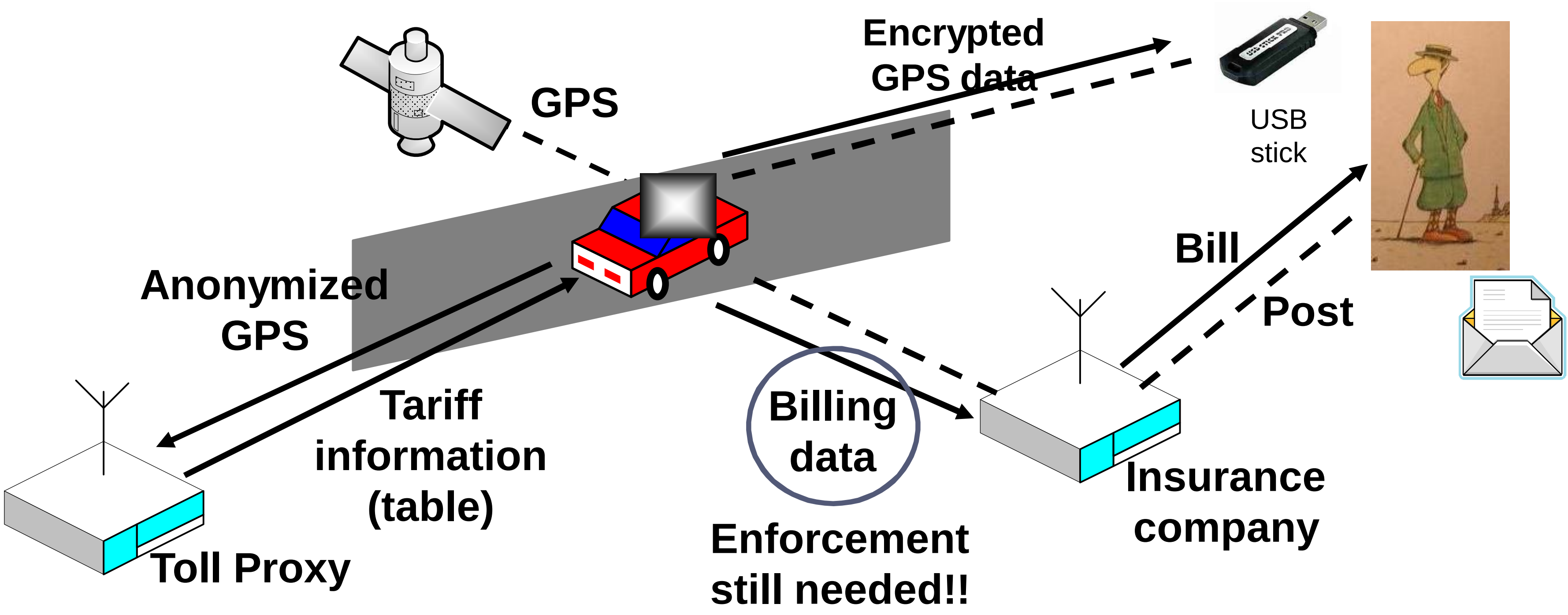


What can we prove?

- OBU used correct prices
  - Prices in the table signed by Toll Service Provider
- OBU was at reported location
  - Compare photo location with commited location
- OBU made correct operations
  - Homomorphic commitments
- Ongoing work: theory and implementation
  - Similar to [Popa et al 09], more flexible

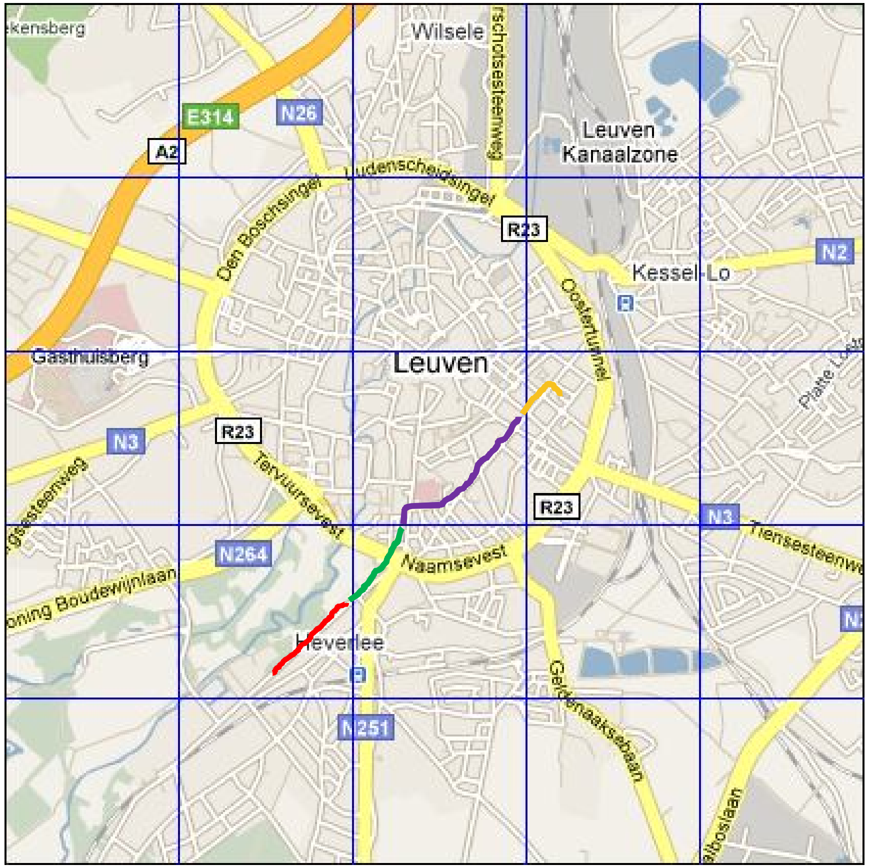
Meet-in-the-middle solution

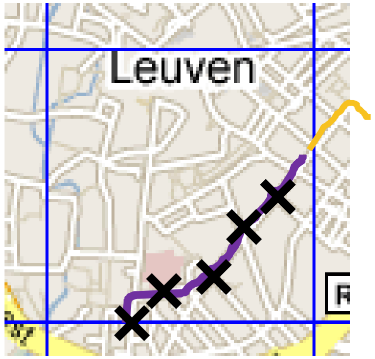
- Use a proxy to compute fees
  - Flexible policies
  - Easy updates



Anonymization

- Divide trajectories in segments: convert map in grid
  - **Remove time information**
  - Send segments “mixed”
    - Space wise
    - Time wise
  - Synchronize vehicles
  - Remove (or change speed)





Anonymization

- Use GSM operator as anonymizer proxy
  - GSM NAT hides IP addresses
  - Encrypted data for the Toll Proxy

- Can trajectories be linked back?
  - What about “disclosure attack”?

- Optimal grid size?
  - Overhead
  - Privacy

Conclusions

- PAYD has many advantages but its implementation may have catastrophic privacy consequences
  - Issues
    - Sensitivity of location data (Difficult to anonymize, allows inferences)
    - Data security (Leakage can always happen)
    - Legal issues (actors difficult to distinguish)
    - Third parties (false sense of privacy)
    - Law-enforcement
    - …
- **It is coming whether we like it or not….**
- Privacy-friendly solutions
  - Computation in the box (PriPAYD [Troncoso et al 07])
  - Half-way solutions (working on it…)

Thanks for your attention!

QUESTIONS?

Carmela.Troncoso@esat.kuleuven.be
http://homes.esat.kuleuven.be/~ctroncos/

- Further reading:
  - C. Troncoso, G. Danezis, E. Kosta, and B. Preneel, "PriPAYD: Privacy Friendly Pay-As-You-Drive Insurance," In Proceedings of the 6th ACM workshop on Privacy in the electronic society (WPES 2007), T. Yu (ed.), ACM, pp. 99-107, 2007
    - Extended version under submission
  - J. Balasch and I. Verbauwhede, "An Embedded Platform for Privacy-Friendly Road Charging Applications." Under Sumbission to Design, Automation and Test in Europe (DATE 2010), 2009.
    - Demo needs to be improved
  - Soon more ☺