A Privacy-preserving Architecture for Ubiquitous Public Transport Systems based on E-ticketing

Ivan Gudymenko*

Chair of Privacy and Data Security, Faculty of Computer Science, TU Dresden
01062 Dresden, Germany
ivan.gudymenko@mailbox.tu-dresden.de
http://dud.inf.tu-dresden.de/

Abstract. In this paper, we present a privacy-preserving architecture for a public transport system. The Transport Authority (TA) is prevented from learning e-ticket IDs and operates only on pseudonyms created by a trusted third party (TTP). Furthermore, the widely distributed terminals are prevented from tracking valid e-tickets during validation. Mutual authentication between terminals and e-tickets is performed at the beginning of each validation session to prevent man-in-the-middle attacks. Our approach allows for regular billing and by this enables to deploy flexible fare policies together with different loyalty programs attractive to customers. User identification together with end billing is performed by the TTP.

Keywords: Privacy protection, E-ticketing, Public Transport

1 Introduction

In the last decade, the ubiquitous computing concept has affected many areas of public life. The public transport sector is no exception [1].

The introduction of the so-called electronic ticketing (e-ticketing) has revolutionized the process of automatic fare collection (AFC) allowing for deployment of flexible fare policies attractive to customers and profitable for public transport companies. As a result, such systems have already been deployed in many countries around the world, e.g., Dutch OV-Chipkaart [2], London Oyster Card [3], EZ-Link in Singapore [4], Hong-Kong Octopus Card [5], etc. Despite introducing noticeable benefits, these systems raise serious concerns over the user privacy. In [6], we addressed this issue and presented a classification of privacy threats endemic to such systems. Moreover, a holistic framework for the development of countermeasures was outlined as well. In this paper, we continue our work and propose a privacy-preserving architecture that is aimed at protecting customer privacy in ubiquitous e-ticketing systems for public transport. Section 2 presents our concept. A short discussion is given in Section 3 with related work following in Section 4. Section 5 concludes the paper.

*This work is supported by the Free State of Saxony and the European Social Fund (ESF). The author would also like to thank his colleagues from Chair of Privacy and Data Security for fruitful discussions and feedback.

2 A Privacy-preserving System Architecture

In this section, our privacy-preserving architecture for the target system is presented together with the attacker model and the core requirements which must be satisfied. The public transport system under concern consists of front-end (FE), back-end (BE), and a bridging element (terminals). E-tickets interacting with terminals compose the system FE. In this scenario, an e-ticket is an electronic medium (e.g., a smart card or an NFC-enabled smart phone) holding the digitalized version of rights to claim the public transport service. BE incorporates powerful interconnected computing centers which maintain system functionality. To enhance user privacy, we additionally consider an external trusted third party (TTP) which acts as a trusted mediator between users and the Transport Authority (TA), see Section 2.3 for details.

2.1 Requirements

Based on the generic description of an e-ticketing system for public transport provided in [6], the following requirements for its privacy-respecting architecture were determined:[1]

1. *E-ticket privacy.*
   (a) *Privacy against terminals.* Terminals must not be able to identify and track valid e-tickets.
   (b) *Privacy against back-end.* Back-end is allowed to correlate travel records related to a single e-ticket while being prohibited from identifying valid e-tickets.
   (c) *Privacy against external observers.* An external observer must be prevented from deriving any identifying information from interactions between e-tickets and terminals.
2. *Billing*
   (a) *Regular Billing.* The target system must support the (fine granular) regular billing[2] (e.g. monthly billing)
   (b) *Billing Correctness.* The billing procedure must be performed correctly (i.e. with strict accordance to the deployed fare policy).
3. *Efficiency.* Check-in/out events handling must comply with the timing requirements.[3]

Requirements *1* and *2a* are particularly conflicting, since fine granular billing implies the ability to track the users' movements in some way to be able to compute the bill. For privacy reasons, on the other hand, tracking should not be possible. A straightforward way to solve this conflict would be to abandon Requirement *2a* and construct a fully anonymous system (as described in [8, 9], for example). We argue, however, that the ability to offer highly flexible

[1] The focus was explicitly made on privacy issues. Detailed functional requirements or any interdisciplinary ones are not considered in this paper.

[2] The reason for this is to provide the ability to support highly flexible tariffs and better loyalty service.

[3] In practice, a maximum tolerated value ranges from 0.2 sec (London Oyster Card) to 2 sec (Singapore EZ-Link) [7].

fare polices supporting different loyalty programs (which, in contrast to the pay upfront approaches, require fine granular billing) is essential in the current state of the market for PT. In [1], for instance, it was stated that a considerable number of customers choose personalized cards since they provide more services. Currently, several systems for PT have already reacted to this trend offering a regular billing approach [10, 11]. We aim at providing the solution independent of any concrete fare policy in use *fully decoupling* the issues of system architecture (specifically privacy protection) from the development of tariff schemes (unlike it was done, e.g., in [12]). In our system, therefore, a reasonable trade-off is allowed, namely while different sessions between e-tickets and terminals are completely unlinkable in FE, the system BE is allowed to correlate different travel records pertaining to a customer *pseudonym* for billing purposes (see Section 2.3 for details) at the same time being prohibited from learning the underlying user identity (Requirement *1b*).

2.2 Attacker Model

Taking into account the arguments presented in previous section, the following attacker model is considered:

1. (Outsider) An observing attacker (outsider) must be prevented from deriving any identifiable information from interaction between terminals and e-tickets.
2. (Insider) Terminals must not be able to identify and track valid e-tickets[1].
3. (Insider) Back-end must be prevented from learning the identifiers of e-tickets.

2.3 Architecture Outline

Having considered the aforementioned requirements (Section 2.1) and the attacker model (Section 2.2), the following system architecture is proposed (see Figure 1). After acquiring an e-ticket, a user checks in at the terminal on entering a vehicle. The respective check-out operation is performed when the vehicle is left. Check-in/out processes consist of three major phases. Firstly, a terminal and an e-ticket establish a secure channel (see *SC Establishment*, Figure 1) using standard methods described, for example, in ISO 7816-4 [13] or in NFC-SEC-01 [14] depending on the carrier medium used to manage an e-ticket. All subsequent messages between the terminal and the e-ticket are, therefore, secured against eavesdropping attacks. During the second phase, a mutual authentication takes place (*Mutual Authentic.*, Figure 1), thus securing the communication entities against unauthorized interaction and man-in-the-middle attacks. It is important that the terminal does not gain any identifiable information about the e-ticket except that the latter is a valid one[2]. Lastly, the terminal checks if the e-ticket has not been blacklisted (*BL Check*). To perform this in a privacy-preserving way, a pseudonymization technique is used (see Section 2.4). If the e-ticket is a

[1] Terminals are widely distributed along the transport network and are, therefore, situated for the most part in an unsecured area (and may even be subject to compromisation).

[2] For instance, it belongs to a valid ticket group.

legitimate one, the terminal accepts it and creates a respective travel record (containing a timestamp, e-ticket pseudonym, location, etc.). Otherwise, the e-ticket is rejected. A set of travel records maintained by the terminal, *TR*, is regularly sent to the back-end system (BE) via the backbone network (*Send TR*, Figure 1) where they are processed for billing purposes (see Requirement *2*). Terminal-side blacklists are repeatedly updated[1] as well (see *Update BL*).

In order to preserve the privacy of e-tickets (Requirement *1*), the transport authority (TA) operates only on pseudonyms created by an external Trusted Third Party (TTP) for each e-ticket during the initialization phase (e.g. on e-ticket acquisition). The bills regularly computed by TA are sent together with the respective (static) pseudonym to TTP which in turn identifies the target user (mapping a pseudonym to the respective user ID) and sends her/him the bill. The respective bill payments are transferred back from users to TA through TTP in an aggregated form to prevent correlation. TTP, therefore, acts as a trusted mediator between the TA and end users. Namely, TA trusts TTP that users are correctly billed (together with payment enforcement) while the customers rely on TTP to protect their privacy and to forward payments to TA.

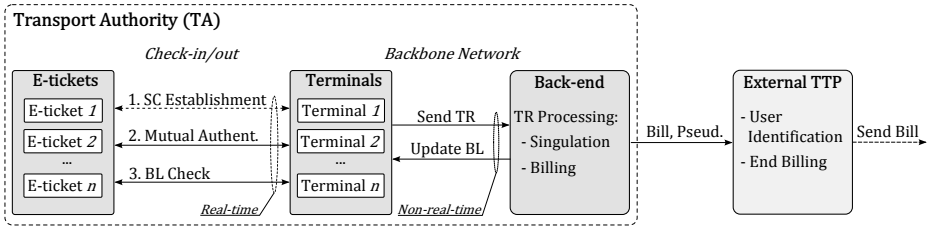

Fig. 1. An e-ticketing system under concern: a high-level architecture overview
TR – Travel Record; *BL* – Blacklist; *SC* – Secure Channel; *TTP* – Trusted Third Party.

In what follows, we elaborate on the privacy-preserving architecture outlined above and discuss its constituents in more detail.

2.4 Pseudonymization

To satisfy Requirement *1* (e-ticket privacy), the following pseudonymization technique is applied. During the initialization phase, a static pseudonym P_i^T is created by TTP for each e-ticket ID. The mapping[2] between P_i^T and the respective e-ticket ID is kept secret. P_i^T is then sent to TA to be included into the TA's pseudonym set P^T. By this, e-ticket privacy is secured against Attacker *3* (back-end, see Section 2.2), since TA is only operating on pseudonyms. TA further encrypts[3] each pseudonym received from TTP with its public key (k_{ta}^+) using some deterministic one-way (trapdoor) function: $P_i^A = E_{k_{ta}^+}\left(P_i^T\right)$.

[1] The frequency of such updates is mainly determined by the connection between terminals and BE (e.g., nightly updates as considered in [15] or more frequent updates if the connection allows).

[2] One of the ways to implement such mapping is to probabilistically encrypt the e-ticket ID (for semantic security) with the private key of TTP and to keep the latter secret.

[3] This is necessary for privacy-preserving blacklist checking, see Section 2.6.

In order to prevent terminals from tracking[1] e-tickets (Attacker *2*), a session pseudonym SP_j is created at the e-ticket side on each interaction with a terminal:

$$SP_j = E_{k_{ta}^+}\left(P_i^A \cdot r_j\right), \tag{1}$$

where r_j is nonce number generated by the e-ticket. Since a terminal is not in the possession of a TA's decryption key (k_{ta}^-), it is infeasible for it to tell if two pseudonyms obtained from different sessions pertain to the same e-ticket or not. Neither can the terminal gain any knowledge from interaction with an e-ticket about the static pseudonym (P_i^A) of the latter.

In order to enable billing in the back-end (BE), the pseudonym singulation step (see Figure 1) is required to correlate different session pseudonyms $\{SP_j\}$ with the respective static one P_i^A using k_{ta}^-. Afterwards the billing process is carried out on static pseudonyms $\left\{P_i^A\right\}$ which are finally decrypted to the initial TTP pseudonyms $\left\{P_i^T\right\}$. The result of the billing step is the set of pairs $\left(bill, P_i^T\right)$ which is regularly (e.g., monthly) send to TTP for end user billing.

2.5 Mutual Authentication Between E-ticket and Terminal

After the secure session is established between an e-ticket and a terminal, a mutual authentication is performed. We suggest that it is carried out using a certificate-based approach, i.e. a terminal provides its unique signature, which is in turn signed by the transport authority (TA). An e-ticket also possesses a signature certified by TA. Unlike the terminal's signature, the one of an e-ticket solely proves that the latter belongs to a valid *ticket group* (e.g., a monthly or a yearly ticket) and does not reveal any identifiable information on each particular e-ticket. This can be done, for example, by using the concept of group signatures, see [16, 17], for example. After successful authentication, e-ticket is checked against a blacklist which is discussed in the next section.

2.6 Terminal-side Blacklist Checking

The suggested mechanism for a terminal-side blacklist checking is based on the inherent homomorphism of an encryption scheme[2] in use. In this case, we exploit the following property:

$$E(x \cdot r) = E(x)^r, \tag{2}$$

where $(x \cdot r)$ represents the (TA) pseudonym of an e-ticket x masked by a nonce r as described in (1). The relevant notations used throughout the paper are summarized in Table 1 for clarity.

The terminal-side blacklist (BL) contains a set of e-ticket static pseudonyms[3], $\{y : y \in BL\}$, which are checked against during the e-ticket verification procedure. After mutual authentication, an e-ticket presents its Session Pseudonym $E_{k_{ta}^+}(x \cdot r)$ to a terminal along with the encrypted nonce value $E_{k_{ta}^+}(r)$ used for masking thus forming a so-called *Session Pseudonym Tuple (SP tuple)*:

[1] Tracking capabilities at the terminal side are not required to keep the system operating. Therefore, following the data minimization principle, terminals must be prevented from doing so (which supports Requirement *1a*).

[2] An example of a suitable encryption scheme is given in further in Section 2.7.

[3] That is, the pseudonyms created by the TA from the TTP ones.

Table 1. A summary of the notations used.

Notation	Meaning
x	an e-ticket static pseudonym (created by TA)
y	a blacklisted e-ticket TA pseudonym
$BL : \{y\}$	a blacklist
r	a random nonce
$E(x \cdot r)$	a session pseudonym, SP
$\big(E(x \cdot r), E(r)\big)$	a session pseudonym tuple, SPT

$SPT \leftarrow \big(E_{k_{ta}^+}(x \cdot r), E_{k_{ta}^+}(r)\big)$. Having obtained this tuple, a terminal can use the malleability property (2) in order to perform blacklist checking. For this, a terminal creates an auxiliary temporary check set C and computes its elements as follows:

$$\forall y \in BL, E_{k_{ta}^+}(r) \in SPT \ : \ c \leftarrow E_{k_{ta}^+}(r)^y.$$

Then a terminal pairwise compares the computed c elements with the delivered Session Pseudonym: $c \stackrel{?}{=} E_{k_{ta}^+}(x \cdot r) \ \ \forall c \in C$. If a match is found, the e-ticket is in the Blacklist set BL and must be rejected.

2.7 A choice of a Suitable Encryption Function

As an example of an encryption function possessing the homomorphic property (2), the scheme based on the intractability of the Discrete Logarithm Problem[1] (DLP) can be used. Thus, $\forall x \in \mathbb{G}_q : \mathbb{G}_q \subset \mathbb{Z}_p^*$ (with p, q being large primes, $q|p-1$) the encryption can be written as $E(x) = g^x \ (mod\ p)$. A session pseudonym (1), therefore, can be expressed as:

$$SP_j \leftarrow g^{x \cdot r_j} \qquad (mod\ \ p), \tag{3}$$

where x is an e-ticket pseudonym, r_j is a session nonce generated to mask x; $r_j, x \in \mathbb{G}_q$.

The homomorphic property (2) can then be expressed as follows:

$$\begin{aligned} E(x \cdot r) &= g^{(x \cdot r)} \\ &= \big(g^x\big)^r \qquad (mod\ \ p) \\ &= E(x)^r. \end{aligned}$$

In principle, any other inherently homomorphic (deterministic) encryption function can be used. The DL exponentiation function was chosen in this paper since it is well known and has been extensively studied.

3 A Short Discussion

The proposed solution satisfies the requirements presented in Section 2.1 within the assumed attacker model (Section 2.2). Namely, observing outsiders and terminals can not identify and track (valid) e-tickets (Requirements *1c*, *1a*). Backend can correlate different travel records pertaining to a certain static pseudonym but is not able to identify an e-ticket (Requirement *1b*). Our system inherently allows for regular billing by design (Requirement *2*). The billing procedure is distributed between BE (under control of TA) and external TTP to protect

[1] DLP follows from the hardness to extract x out of g^x in $\mathbb{Z}_p^*$ (see, for example, [18]).

customer privacy. Due to loose-coupling between BE and terminals, e-ticket validation can be performed locally at the terminal side thus supporting Requirement *3*. Moreover, to boost the performance of the black list checking procedure, an e-ticket can additionally deliver its k-anonymous identifier [19] to a terminal after mutual authentication substantially lowering the search time over the partitioned black list. Furthermore, sensitive pieces of information pertaining to the valid e-tickets are not stored at the terminal side which further enhances the privacy-friendliness of the system (especially in case one of the terminals gets compromised).

4 Related Work

In [8], the authors presented a privacy-preserving framework for public transport based on e-cash, anonymous credentials, and proxy re-encryption. The main drawback of this approach is, however, that it considers a fully online system for e-tickets validation (i.e. BE and terminals are tightly coupled) which is likely to introduce a serious bottleneck (does not scale well). Moreover, it does not allow for regular billing. The approach presented in [15] does not require BE to be always online but assumes it to be fully trusted[1]. Furthermore, in order to authenticate an e-ticket, a terminal must perform an exhaustive search (with additional computations) in the precomputed database[2] with the size of the the overall number of valid e-tickets circulating in the system. Regular billing support was not considered either. The authors of [20] introduced the so-called trusted anonymizers which can be used in an add-on fashion by a customer and are decoupled from the direct functionality of a system. That is, the e-ticket can still be validated in a non-privacy-preserving way if the respective customer's anonymizer is for some reason unavailable. The solution is based on secure key storage with physically unclonable functions (PUFs), symmetric key based authentication, and re-randomizable encryption. Similarly to [15], it allows for local validation by terminals but does not allow for regular billing.

The authors of [21] focused on the neighboring area of toll collection and considered a client-aided fare calculation approach[3] where the billing process is distributed between several non-colluding parties (including the user) and the calculated bill is regularly submitted to BE by a customer in a declarative way. An immediate application of this solution to public transport, however, is not apparent. Furthermore, as it was mentioned earlier in [22], a client-based fee calculation requires more complex user equipment and software (making the system more error prone) and up-to-date tariff/road map data. Moreover, it implies additional overhead for users (beyond simply using the transport service) and less control over the enforcement from the side of TA[4].

None of the reviewed approaches, therefore, fully satisfies the requirements set presented in Section 2.1.

---

[1] BE, therefore, knows all e-tickets identifiers.

[2] For e-tickets authentication, the authors of [15] considered a terminal-side check database (computed by BE and uploaded to terminals) with entries for every e-ticket circulating in the system.

[3] The solution is targeted at the specific fare policy, namely the travel area is divided into cells and the pricing function is assumed to be linear (with cell-based fares).

[4] Both issues are likely to have a negative impact on the overall system acceptance.

5 Conclusions and Future Work

In this paper, a privacy-preserving architecture for public transport systems based on e-ticketing was proposed. The focus was made on personalized e-tickets, for which a privacy-preserving blacklist checking mechanism was suggested. Check-in/out event handling is performed without the necessity to maintain constant connection to the back-end. Therefore, travel records processing performed in BE for billing is decoupled from the front-end and can be carried out periodically (e.g. once a month) in an offline fashion. Moreover, to prevent TA from learning e-tickets IDs, it operates only on pseudonyms created by TTP. The latter is responsible for user identification and end billing.

As future work, we would like to further elaborate on the suggested solution by conducting an in-depth evaluation of our architecture and extending it to efficiently handle anonymous and one-time e-tickets.

References

1. Andrea de Panizza et al. RFID: Prospects For Europe. Item-level Tagging And Public Transportation. Report, eur 24416 en, European Commission, JRC, 2010.
2. Trans Link Systems. OV-Chipkaart. `http://www.ov-chipkaart.nl/`, 2012. Accessed on 30.10.2012.
3. Transport for London. Oyster Online. `https://oyster.tfl.gov.uk/oyster/entry.do`, 2012. Accessed on 30.10.2012.
4. Land Transport Authority. EZ-Link. `http://www.ezlink.com.sg/index.php`, 2012. Accessed on 30.10.2012.
5. Octopus Cards Limited. Octopus. `http://www.octopus.com.hk/home/en/index.html`, 2012. Accessed on 30.10.2012.
6. Ivan Gudymenko. On Protection of the Users Privacy in Ubiquitous E-ticketing Systems Based on RFID and NFC Technologies. In *PECCS-2013*, February 2013.
7. Mohamed Mezghani. Study on electronic ticketing in public transport. Technical report, European Metropolian Transport Authorities (EMTA), May 2008.
8. Thomas Heydt-Benjamin et al. Privacy for Public Transportation. In George Danezis and Philippe Golle, editors, *PETs-2006*, volume 4258 of *LNCS*, pages 1–19. Springer Berlin Heidelberg, 2006.
9. Foteini Baldimtsi et al. Pay as you go. In *Workshop on hot topics in privacy enhancing technologies, HotPETSs 2012*, 2012.
10. Großraum-Verkehr Hannover. HANNOVERmobil. `http://www.gvh.de/hannovermobil.html?&L=1`, 2013. Accessed online on 19.04.2013.
11. Phoenix Valley Metro. Platinum Pass Program. `http://www.valleymetro.org/employer_programs/platinum_pass`, 2013. Accessed online on 19.04.2013.
12. Jaap-Henk Hoepman and George Huitema. Privacy Enhanced Fraud Resistant Road Pricing. In Jacques Berleur et al., editors, *HCC9/CIP-2010*, volume 328 of *IFIP AICT*, pages 202–213. Springer Berlin Heidelberg, 2010.
13. ISO. ISO/IEC 7816-4:2005. Identification cards – Integrated circuit cards – Part 4: Organization, security and commands for interchange, 2005.
14. ECMA International. NFC-SEC. NFCIP-1 Security Services and Protocol. Cryptography Standard using ECDH and AES, 2008. White paper.
15. Gildas Avoine, Cdric Lauradoux, and Tania Martin. When Compromised Readers Meet RFID. In Heung Youl Youm and Moti Yung, editors, *Information Security Applications*, volume 5932 of *LNCS*, pages 36–50. Springer Berlin Heidelberg, 2009.
16. Aggelos Kiayias, Yiannis Tsiounis, and Moti Yung. Traceable Signatures. In Christian Cachin and Jan Camenisch, editors, *Advances in Cryptology – EUROCRYPT 2004*, volume 3027 of *LNCS*, pages 571–589. Springer Berlin Heidelberg, 2004.

17. David Chaum and Eugne Heyst. Group Signatures. In Donald Davies, editor, *Advances in Cryptology – EUROCRYPT 91*, volume 547 of *LNCS*, pages 257–265. Springer Berlin Heidelberg, 1991.
18. Oded Goldreich. *Foundations of Cryptography: Volume I Basic Tools*. Cambridge University Press, 2004.
19. Latanya Sweeney. k-anonymity: A Model for Protecting Privacy. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, 10(5):557–570, October 2002.
20. Ahmad-Reza Sadeghi et al. User Privacy in Transport Systems Based on RFID E-Tickets. In Claudio Bettini et al., editors, *PiLBA-2008, Malaga, Spain*, 2008.
21. Jaap-Henk Hoepman et al. Privacy and Security Issues in e-Ticketing – Optimisation of Smart Card-based Attribute-proving. In V. Cortier et al., editors, *Workshop on Foundations of Security and Privacy, FCS-PrivMod 2010*, 2010.
22. Wiebren Jonge and Bart Jacobs. Privacy-Friendly Electronic Traffic Pricing via Commits. In Pierpaolo Degano et al., editors, *FAST 2009*, pages 143–161. Springer-Verlag, Berlin, Heidelberg, 2009.