Pseudonymous Context-Aware Transport Applications[*]

David Evans
David.Evans@cl.cam.ac.uk

Alastair R. Beresford
Alastair.Beresford@cl.cam.ac.uk

Computer Laboratory,
University of Cambridge,
Cambridge, United Kingdom. CB3 0FD

Abstract: *This position paper examines some of the privacy features required in context-aware transport applications. We begin by describing why pseudonymity is preferable to access control for protecting privacy. We then present three sample transport applications and describe the key pseudonymity features that could be provided by a transport middleware framework; these features enable applications to be privacy-aware, limiting exposure of the participants' identities and locations.*

Context-aware applications use sensors to measure their surroundings and use these data to construct a world model of the local environment. This model allows an application to perform operations for its users in an automatic or semi-automatic way. For example, a bus arrival alert application may notify an individual when to leave work for the bus stop in time to catch the bus. To make the alert message timely, the application's context may encompass location data collected from the bus combined with prevailing congestion information and the juxtaposition of bus stop and office location.

Relevant context information, by necessity, includes information about people. This means that as context-aware applications proliferate, computers will gather, collate, and distribute much more personal information than they do today and computing could have a major impact on the privacy of individuals. Our aim is to minimise this impact whilst still allowing context-aware applications to function.

Access control is often cited as a method of improving the privacy for individuals in a ubiquitous computing environment, but we believe that it is unsuitable in many cases. Access control requires configuration: users must enter their access control decisions into the system and this runs counter to ubiquitous computing's aim for applications to disappear into the background; access control requires technical competence: correct configuration necessitates in-depth technical understanding which many users will not have; and access control demands trust: submitting personal data to an application requires trust that access control constraints will be observed. In general, we believe that the actions required by the user in effecting access control do not afford natural user interfaces for context-aware applications.

Anonymisation offers a alternative—removing personally identifiable information from the context data processed by applications assuages privacy concerns. Anonymisation does not require configuration, technical competence, or trust in the applications which receive the data. However it is not a panacea. Anonymisation does not work universally since some applications, such as "Find Alice", require a concrete notion of identity and location.

Anonymisation constrains the distribution of concrete identities by replacing identifiable attributes, such as names, with pseudonyms. The use of pseudonyms allows controlled linkability for applications that require it, reducing the number of software components that receive personally identifiable information. This is particularly useful in the ubiquitous computing scenario, where software executing on many platforms, each under the control of a different party, makes the establishment of trust particularly difficult.

Simply replacing a direct identifier such as a name with a static pseudonym does not necessarily make it untraceable. For example, it is often possible to re-identify location data associated with a static pseudonym in an office environment [1]. More advanced statistical disclosure control techniques are required to ensure that pseudonyms issued to applications are unlinkable with real-world identities. Such techniques are domain specific because they require a model of the knowledge received by potential attackers [2].

In this position paper we investigate the domain of pseudonymous context-aware transport applications and suggest some services that should be included in a transport information middleware. We believe that a range of useful applica-

[*] This research is supported by EPSRC grant EP/C547632/1 and British Telecommunications plc.

tions can be constructed while requiring little or no explicit location information in their context data. This will bring us closer to realising our goal of minimising the exposure of participants' personal information—including identity and location—not only to other participants but to applications and the middleware itself.

Applications

We begin discussion by reviewing some example transport-related applications. Those we have selected are amenable to offering privacy protection using untraceable pseudonyms as they do not require a centralized notion of concrete identity. Furthermore, in the presence of pseudonyms, they require little explicit location information.

Friend Finder: Suppose that Alice and her friend Bob are travelling to a common destination using public transport. They live on the same transit route and a vehicle following this route will first pick up Alice and then, some time later, Bob. Bob would like to travel on the same vehicle as Alice. To do this, when Bob arrives at his transit stop, he needs to know whether Alice is on the vehicle that is approaching, whether she is on a vehicle yet to arrive, or whether she has passed his stop already.

Meeting Place Recommender: Suppose that people located at two or more places within a city would like to meet and want to find the best location and a route for each person to take to get there. "Best" is defined by a utility function of factors such as the amount of time until everyone in the group arrives and the total distance travelled.

Taxi Locator: Suppose that two or more taxi companies operate in a particular city and that these companies would like to use the transport information infrastructure to support their business. They want routing for dispatch (determining which car is closest to a customer and the best route to get there), to allow customers to determine which company can provide service first, to notify a customer of the current location of the car assigned to him or her, and to provide an estimate of the cost of a particular trip based on prevailing traffic conditions.

Middleware Services

In order to minimise the impact on privacy we need to reduce the breadth of application components and supporting middleware which is able link a concrete identity with a pseudonym and associate that pseudonym with a physical location. To this end we suggest the following three middleware services.

The middleware must provide, or at least specify, a mechanism for pseudonym creation and distribution. Therefore middleware or an application directly under user control should support an association between one or more pseudonyms and an individual, a group, a location, or a (geographical) route. Individuals can share pseudonyms out-of-band, such as by writing a number on a piece of paper, or sharing can be done in-band through the middleware.

Storage of pseudonym data for later distribution is also important. For example, to build the friend finder application we can augment vehicles with a wireless computer capable of storing passengers' pseudonyms. Travellers can upload their current pseudonym(s) to the vehicle when boarding and remove them when alighting. Pseudonym values may be collated and distributed by a transport middleware so that vehicle occupancy can be queried centrally. A vehicle's ability to store pseudonyms and respond to associated queries can also be used to construct the taxi locator.

The meeting place recommendation and taxi locator applications require suggested routes and locations based on various information provided by the participants and generated from sensor data. This is challenging when the participants' locations are concealed from the system. We envision a form of distributed Geographical Information System (GIS) that can operate on pseudonymous locations, individuals, and routes, providing functionality such as computing the distance between two points without knowing the location of the points.

Future Work

We believe that the three services described above—pseudonym creation and distribution, vehicular pseudonym storage, and "pseudonymous GIS"—are useful building blocks for constructing privacy-aware transport applications. We intend to test this theory by developing these services and using them to construct prototypes of our three applications.

References

[1] Alastair R. Beresford and Frank Stajano. Location privacy in pervasive computing. *IEEE Pervasive Computing*, 3(1):46–55, 2003.

[2] Leon Willenborg and Ton de Waal. *Statistical Disclosure Control.* Springer, 2001.