Privacy and Identity in a Mobile Pervasive Environment

Chris Mitchell
www.isg.rhul.ac.uk/~cjm

Agenda

- Privacy issues for mobile healthcare
- Identity, privacy and anonymity
- Technology support
- Major challenges
- Secure ad hoc communications
- Secure/trusted platforms
- Conclusions

Environment

- Ubiquitous computing environment …
- Assumed to mean an environment in which multiple devices, some personal, some mobile, combine to provide an all-pervasive computing and communications service to end-users.
- Requires automatic configuration of certain aspects of some devices, since it is assumed that there may be no global management infrastructure.

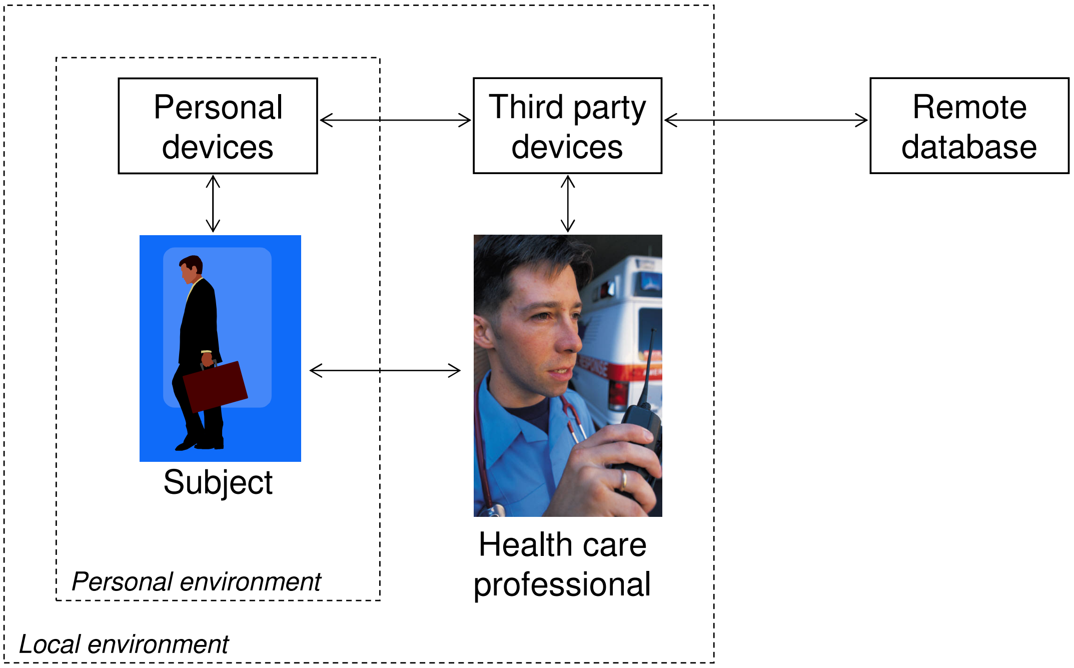
On privacy

- It is important to distinguish between security and privacy.
- Privacy is not just a special case of security – there are interesting interactions between security and privacy.
- It is important to appreciate that security and privacy are different notions – indeed the two sometimes conflict.
- Examples of conflicts:
  - accountability versus anonymity;
  - denial of service resistance versus anonymity.
- It is nevertheless true that supporting privacy requires the provision of security services, e.g. confidentiality for stored and transmitted data, and access control.

What is the problem?

- Before trying to decide how to provide and support privacy in a mobile health environment, we need to decide what privacy issues can arise.
- This requires identifying where Personally Identifying Information (PII) is at risk of disclosure.
- Note that disclosure can occur in a variety of ways (e.g. linking of sensitive information to a unique identifier, or collection of a ‘profile’ for an unidentified individual, which may eventually be linked to a particular individual).

Mobile health care data model



Health information handling

- Where in a mobile environment is sensitive information at risk?
  - at *point of capture* of information, e.g. by individuals (e.g. physicians, paramedics) or automatically (e.g. using sensors);
  - when *stored/used in personal devices* (e.g. smart card, sensor equipment, mobile phone);
  - when *stored/used in mobile third party devices* (e.g. belonging to a physician or paramedic);
  - when *stored in fixed databases* (e.g. in a hospital);
  - when *communicated between devices*.

Mobile/pervasive computing – new challenges

- Many threats arising in a mobile environment are the same as those arising in a more conventional environment.
- However, we also have new threats, e.g.:
  - mobile devices are more easily stolen;
  - mobile devices typically offer less physical security;
  - mobile devices may need to communicate with other devices where a single security infrastructure is not present (as might, for example, exist in a hospital);
  - mobile devices may capture personal/medical information in a way not requiring user consent or knowledge.

Agenda

- Privacy issues for mobile healthcare
- Identity, privacy and anonymity
- Technology support
- Major challenges
- Secure ad hoc communications
- Secure/trusted platforms
- Conclusions

Identities

- A user may have many identities (with associated identifiers) for use with different third parties.
- For example:
  - we all have a name (although this is typically not a unique identifier);
  - an employee may have an employee number for use with his/her employer;
  - a citizen has one or more numbers for interactions with government;
  - a health care user may have a government ID (e.g. the NHS number in the UK), and one or more health insurance identifiers.

Credentials

- To enable a provider of service to authenticate a user as a legitimate holder of an identity, the user may be required to provide one or more credentials.
- Possible credentials include:
  - a password;
  - a biometric sample;
  - a public key certificate;
  - a MAC computed using a shared secret key;
  - a signature on a challenge provided by the service provider.

Authorisation

- Once an entity has been authenticated, the provider needs to decide whether or not to grant the requested service.
- This is refereed to as *authorisation* (i.e. is the holder of this identity authorised to access this service?).
- This could, for example, be supported using server-held Access Control Lists (ACLs).

Anonymity

- A user may wish to be able to access a service in an *anonymous* way.
- Anonymity means that no party will learn any of the identities of the user.
- Providing anonymity for free services is relatively simple.
- If service requires using stored data, then some level of identification to the stored data provider is required.
- If payment is needed, then an anonymous payment system is needed, e.g. cash or e-cash.
- True ('absolute') anonymity is difficult to achieve, since even revealing an IP address or a MAC layer address to some extent compromises it.

Pseudonymity

- Pseudonymity is a lesser form of anonymity, in which the user reveals a special type of identity to the service provider known as a *pseudonym*.
- Typically, new pseudonyms will be generated regularly, i.e. pseudonyms are typically short-lived.

Unlinkability

- Unlinkability is a privacy property required to support the use of pseudonyms.
- Two pseudonyms are unlinkable if a third party cannot tell whether or not they belong to the same user.
- In practice, absolute unlinkability is often difficult to achieve, since the authorisation process may reveal information about the user.

Consent

- In many medical scenarios, the subject may be required to give implicit or explicit consent for
  - treatment to be given, and
  - information to be passed to a practitioner.
- In case a), some measure of non-repudiability may be required;
- In case b), the information source will need to authenticate the subject – potentially problematic since the information source may be remote and only communicating with a device belonging to the practitioner.

Auditing

- For a variety of reasons (e.g. to enable subsequent investigation on case of a dispute) all data handling actions may need to be auditable.
- That is, records may need to be kept of all significant actions involving sensitive data.
- These records may themselves be a threat to user privacy.

Agenda

- Privacy issues for mobile healthcare
- Identity, privacy and anonymity
- Technology support
- Major challenges
- Secure ad hoc communications
- Secure/trusted platforms
- Conclusions

Cryptography for security

- Many of the security issues which must be addressed to protect patient privacy can be solved using well-understood cryptographic techniques, e.g. encryption, signatures, etc.
- However, use of cryptography requires key management, typically supported by some kind of security infrastructure.
- Key management is also well-studied, e.g. in the form of PKIs for reliable public key distribution.

Privacy Enhancing Technologies

- A huge amount of effort has been devoted to a variety of technologies designed to enhance user privacy in a distributed environment.
- This includes a wide range of techniques, including:
  - anonymising networks (MIXes);
  - anonymising routing techniques (onion routers);
  - anonymous credential systems;
  - P3P (Platform for Privacy Preferences Project);
- There is a series of conferences (PET ’06, etc.) devoted to such technology.

Anonymous credentials

- Credentials here mean electronic documents designed to enable an entity to authenticate itself (and/or prove it has authorisation to achieve certain objectives).
- Anonymous credential systems (protocols) enable users to prove possession of properties (e.g. authorisations) without revealing their identity.

Trusted computing

- *Trusted Computing* is a rather different technology (examined in greater detail later in this talk).
- TC is implemented as a combination of hardware and software enhancements to a computing platform (e.g. a PC, PDA, server, or mobile phone).
- Hardware enhancements typically include a TPM (Trusted Platform Module) and modifications to the boot ROM.

Agenda

- Privacy issues for mobile healthcare
- Identity, privacy and anonymity
- Technology support
- Major challenges
- Secure ad hoc communications
- Secure/trusted platforms
- Conclusions

Lack of infrastructure

- There are a number of major new challenges arising in a mobile environment.
- However, as alluded to previously, at the root of most of these problems is the potential lack of a single pre-existing managed security infrastructure.
- Where such an infrastructure does exist, many of the problems become much less fundamental.

Trusted interactions

- Devices collecting, storing and/or using health data may need to share this data with other devices.
- E.g. a wireless heart monitor may need to pass data to a portable device used by a physician which integrates and displays the data to the physician.
- Clearly such data transfers should not take place without restriction for privacy/security reasons.
- What if the devices interacting do not all belong to the same individual?

Maximising privacy

- Stored data will typically need to be associated with a unique individual, and the capability to retrieve data for unique individuals is imperative.
- However, this does not mean that the health care worker necessarily needs access to unique identifying data.
- For example, a paramedic may need to know information regarding the allergies of a subject, but does not need to know the subject’s name or address.

Authorisation

- Assuming security and privacy measures are in place to prevent casual access to healthcare data, how is authorised access to be managed?
- Rapid access to critical data is required, perhaps without subject authorisation (e.g. if subject is incapacitated).
- A mobile health practitioner may only have access to mobile devices, and perhaps one or more devices (e.g. a smart card, mobile phone, …) of the subject.

Agenda

- Privacy issues for mobile healthcare
- Identity, privacy and anonymity
- Technology support
- Major challenges
- Secure ad hoc communications
- Secure/trusted platforms
- Conclusions

General problem

- An ad hoc network is a collection of communicating devices with no pre-existing relationships or infrastructure.
- A typical scenario for use of such a network is an emergency situation, e.g. a major transport accident.
- Many security issues arise in establishing working relationships in such a network, e.g.:
  - Initial trust setting;
  - Managing collaborative activities (e.g. routing);
  - Authentication, authorisation, …

Trust establishment

- One fundamental issue for two devices in an ad hoc network (with no pre-existing relationship) is deciding whether to trust one another (and by how much).
- What resources or services should one node make available to another?
- Can another node be trusted to provide a communications service without eavesdropping, manipulating messages, and/or selectively dropping packets?

Reputation schemes

- One solution is to try to dynamically ‘measure’ the trustworthiness of another node.
- Each node maintains an assessment (typically a numerical score) of the trustworthiness of its neighbour nodes.
- This would typically be derived by monitoring the behaviour of the node, possibly combined with assessments passed on by other nodes.
- Such schemes are widely used, e.g. on eBay.
- However, such schemes are also easily spoofed.
- Many schemes have been proposed, but the robustness of schemes against deliberate attack has rarely been assessed.

Currency schemes

- Another solution proposed in some scenarios, is to use a virtual ‘currency’ to reward nodes performing service in an ad hoc network.
- Currency needs to be unforgeable!
- Other problems arise – e.g.
  - shortages of currency can cause major inefficiencies in the network;
  - how to start the scheme started, i.e. how does the currency get allocated initially?

Infrastructure needed?

- These dynamic trust management schemes may be appropriate in some settings, perhaps where data are not of high value.
- However, such a solution seems inadequate in a healthcare setting, where users require guarantees that their data will not be abused.
- It seems that some kind of security infrastructure (to support entity authentication, etc.) will be needed.
- This is, of course, problematic, when the devices are not managed by a single entity.

Solutions

- In the remainder of this talk we look at one possible solution to the problem of providing a ubiquitous security infrastructure.
- We consider the possible use of trusted computing to provide this infrastructure.
- This is just one approach - however it appears that it can help solve two basic problems:
  - provision of a security infrastructure;
  - enabling one device to determine its level of trust in a another (‘foreign’) device.

Agenda

- Privacy issues for mobile healthcare
- Identity, privacy and anonymity
- Technology support
- Major challenges
- Secure ad hoc communications
- Secure/trusted platforms
- Conclusions

A trusted system

- A trusted system or component is one that behaves in the expected manner for a particular purpose.

  [Trusted Computing Group – www.trustedcomputinggroup.org]
- This is difficult to achieve for a PC – where typically there is no way of telling whether the 'real' (uncorrupted) Windows is running.
- As a result there is no way of getting any confidence in the correct running of applications. [Even if the operating system says that everything is OK, then this does not help because it cannot be believed].
- It is even more difficult to prove to a third party that the state of a PC is as claimed.

Fundamental requirements

- First and foremost we need to have a way of achieving assurance that the operating system has booted correctly.
- This requires assuming that the PC hardware has not been modified; this is made difficult, but not impossible, for the attacker by embedding key functions in a dedicated chip – the Trusted Platform Module (TPM).
- Need a way of checking the boot process.
- The component that checks the initial boot must be trusted – the ‘Core Root of Trust’ – this is hardware-based.
- If the loaded software has been checked (and hence is reliable), it can check the next software to be loaded, and again there is a solid basis for trust – this process is iterated.

Monitoring the checking

- As well as performing checks during the boot process, there needs to be a reliable way of recording the results of each of these checks.
- The trusted hardware incorporates hardware registers which store hash-codes of software that has been loaded – these registers provide a reliable record of all the software that has been executed on the trusted platform.
- Anyone wishing to check the state of the platform only needs to be given the contents of these registers (as long as they know what the values ‘ought to be’).

Building on the trusted base

- This base of trust can be used to support two fundamental trusted computing functions:
  - ***Attestation***, where a PC can reliably attest to its software state to a third party (by describing the contents of the registers which store hashes of software state);
  - ***Secure storage***, where a PC can store data in such a way that only if the PC is in a specific trusted state will the data be decrypted and available to an application (by linking the decryption keys to specific register contents).

Components of a trusted computing framework I

- Shielded locations and protected capabilities:
  - Protected capabilities are those capabilities whose correct operation is necessary for the platform to be trusted.
  - Shielded locations are areas in which data is protected against interference or snooping.
    - Only protected capabilities have access to shielded locations.
- Attestation:
  - Attestation by the TPM;
  - Attestation to a trusted platform (incorporating a TPM);
  - Attestation of a trusted platform;
  - Authentication of a trusted platform.
- Integrity measurement, storage and reporting.

[TCG specification Architecture Overview]

Components of a trusted computing framework II

Microsoft's additional components:

- Process isolation, whereby an integrated isolation kernel facilitates the execution of several compartments/domains in parallel on the same machine, and controls the access of applications/OSs running in these compartments to system resources.
- A secure path from the peripherals to trusted applications.

[Microsoft Security Model for NGSCB]

Components of a trusted computing framework III

- Confidentiality and integrity protection of application code and data during execution.
- Confidentiality and integrity protection of application code and data during storage.
- Integrity protection of the operating system and underlying hardware so that the above properties can be satisfied.
- Platform attestation.
- A trusted path to the user so that confidentiality of user input can be assured.
- Secure channels to devices and between applications to ensure the confidentiality, integrity, and authenticity of communicated data.
- Reliability assurance, necessitating size restrictions on trusted critical components.

[Sadeghi and Stüble: Bridging the Gap between TCPA/Palladium and Personal Security]

Components of a trusted computing framework IV

- Attestation – provides remote assurance of the state of the hardware and software stack running on a computer.
- Isolation – execution environments/domains/compartments.
- Secure storage:
  - Encryption;
  - Sealing (binding of data to specific machine state).
- Secure I/O.

Deployment of trusted computing

- The TCG specifications allow many possible implementation architectures on a variety of platform types.
- Most effort into PC implementations – many PCs now being shipped with TPMs, and Windows Vista will build on presence of TPM.
- However, range of working groups within TCG looking at other platforms.
- This include Mobile Phone Working Group (MPWG) – profiling TCG specifications for mobile phone use.

Applications – ubiquitous infrastructure

- Every TCG-compliant platform comes pre-equipped with an asymmetric key pair and a set of certificates/credentials.
- These provide a basis by which any platform can authenticate any other platform, and learn the platform type, and the level of physical security provided.
- This provides a simple basis for a key management infrastructure (a ubiquitous PKI).

Applications – stable identities

- What does authentication mean when one entity does not know in advance the name of the other entity with which it is communicating?
- We can authenticate against attributes rather than name.
- Trusted Computing enables devices to remain anonymous, yet their properties to be authenticated (using credentials issued, e.g. by the platform manufacturer and/or third party testing laboratories).
- It also enables degree of ‘linkability’ to be selected, e.g. enabling all interactions with one third party to be linked, but no linking between interactions with different third parties (hence notion of ‘stable’ unidirectional identities)

Applications – trusted interactions

- By exploiting the attestation function of TCG, one platform can determine the software state of another.
- That is, one device can be sure that the data passed to another device will be handled by trusted software before transferring the data.
- In a simple system (e.g. a simple mobile device) TCG attestation may be sufficient.
- In a complex system (e.g. a PC), the TCG attestation will need to be applied to a virtualisation layer, which will then guarantee the integrity of software running on top of it.

Applications – secure data storage

- TCG enables one device to ensure that data passed to another TCG device will only be available in specific circumstances.
- The device stores data encrypted such that it will only be decrypted if the device is in a specific configuration.
- Thus stored data can be made inaccessible to unauthorised software and even the owner of the device.
- This could be applied both to health care data itself and to audit trails created by health care applications.

Agenda

- Privacy issues for mobile healthcare
- Identity, privacy and anonymity
- Technology support
- Major challenges
- Secure ad hoc communications
- Secure/trusted platforms
- Conclusions

Fundamental problems

- We have identified fundamental problems that need to be solved to realise full potential of mobile ubiquitous computing for healthcare:
  - need for a ubiquitous security infrastructure to support secure communications between mobile devices;
  - need for one device to be able to verify the conditions under which data will be stored, handled, and retransmitted by another device.

Directions for future research

- Can trusted computing genuinely realise all the security infrastructure needs of future pervasive computing environment?
- Who will be the trusted third parties to support the trusted computing based security infrastructure?
- What if some mobile devices are trusted computing enabled and others are not?
- What other solutions are there?

More information

- TCG: `www.trustedcomputinggroup.org`
- MPWG: `www.trustedcomputinggroup.org/groups/mobile`
- OpenTC (an EU Integrated Project with the goal of providing open source software supporting TC): `www.opentc.net`
- Trusted Computing books:
  - S. Pearson (ed.), *Trusted Computing Platforms: TCPA Technology in Context*, Prentice-Hall, 2002.
  - C. J. Mitchell (ed.), *Trusted Computing*, IEE, 2005.
  - S. W. Smith, *Trusted Computing Platforms: Design and Applications*, Springer-Verlag, 2005.