Trusted Computing in Mobile Action

Horizontal integration of access technologies needs convergence of authentication technologies. The concept of establishing trust by trusted computing enables cross-domain authentication functionality, surmounting technical boundaries. The focus of target application scenarios lies in the realm of mobile networks and devices.

Nicolai Kuntze, Andreas U. Schmidt
Fraunhofer-Institute for Secure Information Technology SIT
Darmstadt, Germany

ISSA 2006 – Sandton, South Africa, 5 July 2006

Fraunhofer SIT Institut Sichere Informations-Technologie

issa information security south africa conference2006 from insight to foresight

The future role of Mobile Network Operators (MNOs)

From a recent study of KPMG Germany*:

„Throwing subsidised handsets on the market is not a sustainable strategy for success. It makes more sense to build a stable and loyal customer base with attractive and convergent services"

- **MNOs are privileged players as they already have a stable and huge customer base**
- **MNOs are advised to consider new business models from the Web 2.0 environment**
- **It is necessary to explain the new convergent services to the customer**
- **Opportunities:**
  Mobile entertainment; mobile network games are enabled by 3G and convergent technologies;
  Live TV and Video downloads are attractive to commit customers to a service provider
- **A majority of customers wishes a single service provider**
  for accessing the various services and coordinating charging and payment,
  making the various service providers transparent with respect to the user recognition
- Accounting and charging competence is going to be a key ability

And we may add: Trust is crucial

Fraunhofer SIT Institut Sichere Informations-Technologie

* KPMG: Consumers and Convergence - Challenges and opportunities in meeting next generation customer need. 2006.
http://www.kpmg.de/about/press_office/13611.htm

issa information security south africa conference2006 from insight to foresight

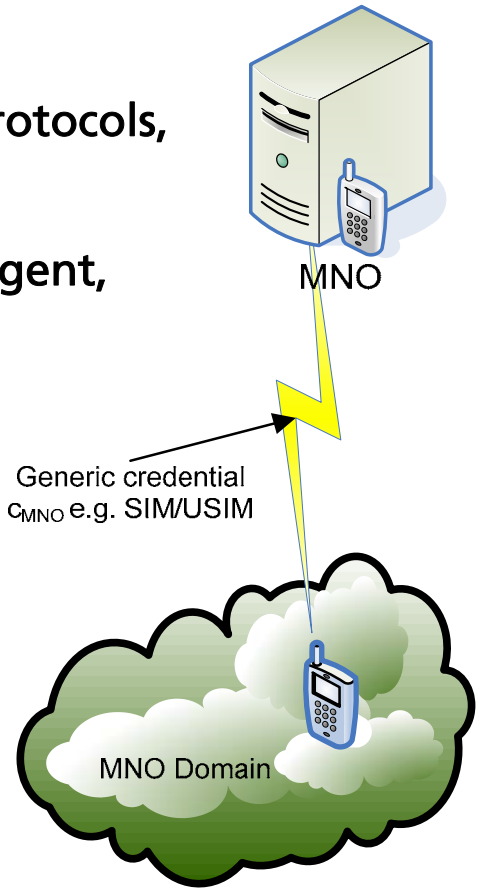
Trust in mobile service access

- **Mobile access to applications & content is becoming network-agnostic**
  - Customers attracted by attractive applications & content
  - Diversity of technologies (2G, 3G, WLAN, WiMAX, MobileIP)
  - Customers interested in optimizing price/performance ratio
- **Mobile devices are becoming very smart, multi-purpose devices**
  - More than voice comm., both consuming and providing applications, data and media
  - Network access is a commodity, customers expect additional features
  - Next step for MNOs (business models): providing customised/customisable services
- **Novel requirements for trust across domains – even technological boundaries**
- **Trusted computing (TC) can become the enabler for service provisioning**
  - Enables network- and device-agnostic trust relations on application-level
  - Uniform trusted platform for service provisioning
- **Credentials from various domains of trust, carried, managed and transmitted by TC-enabled devices can yield *transitive trust relationships***

My small world of trust

We take an authentication (minimalist) view on trust

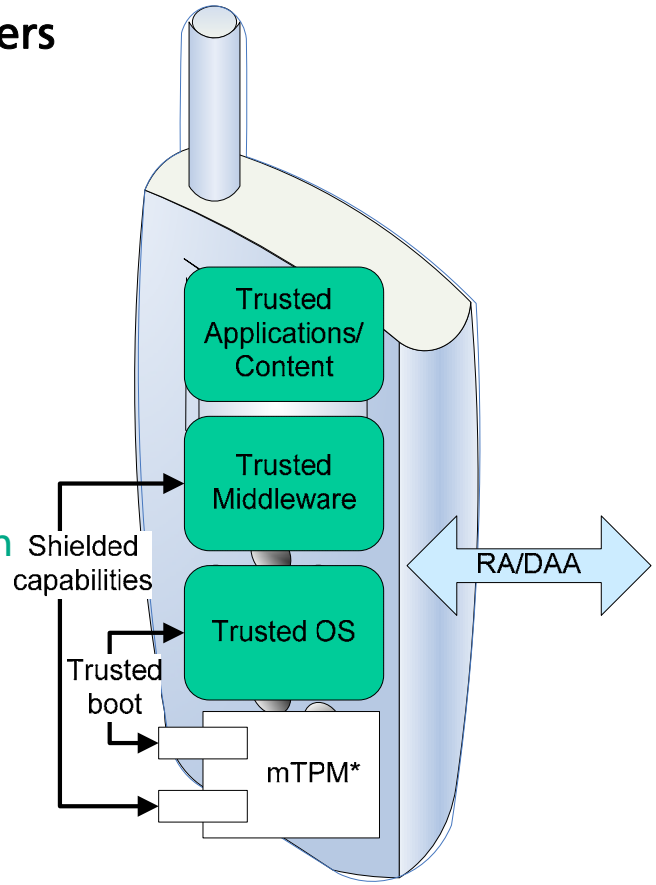
- By transmitting authentication data or exerting authentication protocols, agents enter the domain of trust of a principal
- The authentication process attests to the trustworthiness of the agent, i.e., makes implicit or explicit assertions to the principal about
  - the identity of an agent and/or
  - its being in a secure/trustworthy state
- the token/data for authentication and the embodied attestation is subsumed under the term credential
  - SIM/USIM, cryptographic certificates, shared secrets, PIN/TAN, (smart card-based) biometry, …
- Our aim: Complement generic domain credentials c by trust credentials t obtained somehow by use of TC, to enable referral trust and transitivity



Trusted mobile agents

A system equipped with a trusted platform module (TPM) as specified by TCG, is called a trusted platform (TP). It offers protected capabilities and shielded locations

- TPM provides (RSA) key management, i.e., methods for generation, storage, and usage of keys
- Trust measurements on the system environment exerted at boot- and run-time allow for trustworthy assertions about the current system state and a re-tracing of how it was reached
- The system state, and a measurement log (order matters) of how it was reached is securely stored in platform configuration registers (PCR) tamper-resistantly located inside the TPM.
- Memory curtaining, sealed storage, and secure I/O are enabled by pertinent TPM base functions.
- Trustworthy system and application software can build on this to establish authenticated communication with the exterior and transmit data protecting integrity and confidentiality.



*mobile TPM profile, currently under spec by the pertinent TCG WG

Remote attestation

- The TPM is a hardware (or virtualised) root of trust on which the two essential attestation protocols rest:
  - Remote attestation (RA), yieldig, conventional privacy using an ID provider, and
  - Direct anonymous attestation (DAA), employing zero-knowledge proofs

- RA enables an exterior entity to determine if the requesting agent is altered or not:
  - The TP presents the value of a certain PCR *and* the log how this value was created
  - This data block is signed using an Attestation identity key (AIK)

- Privacy: AIKs are generated by the TPM, and certified by a privacy CA (PCA)
  (the TP identifies itself towards the PCA using the endorsement key [EK])

- Based on this data two attributes can be attested to the external verifier:
  - **The data was produced by a genuine TP**
    by verifying the AIK-signature of the data with the corresponding PCA certificate.
  - **The integrity, i.e., unalteredness of the TP**
    by comparing the integrity values of the measurement log.
    A compromised system can tamper the log but cannot change the PCR values as these are protected by the TPM.

Practical caveat for attestation

- The verifier has to know a reference value for the reported values in the transmitted measurement log and PCR value

- This reference base is rather large and hard to maintain in the PC domain:
  There are many
  - hard- and software versions
  - admissible boot and runtime parameters
  - frequent updates

  A remedy might be virtualisation, implementing small trusted compartments managed by a hypervisor and virtual TPMs (currently under specification)

- The mobile domain is distinct
  - the number of hardware combinations is rather small
  - updates of the used software are also rare

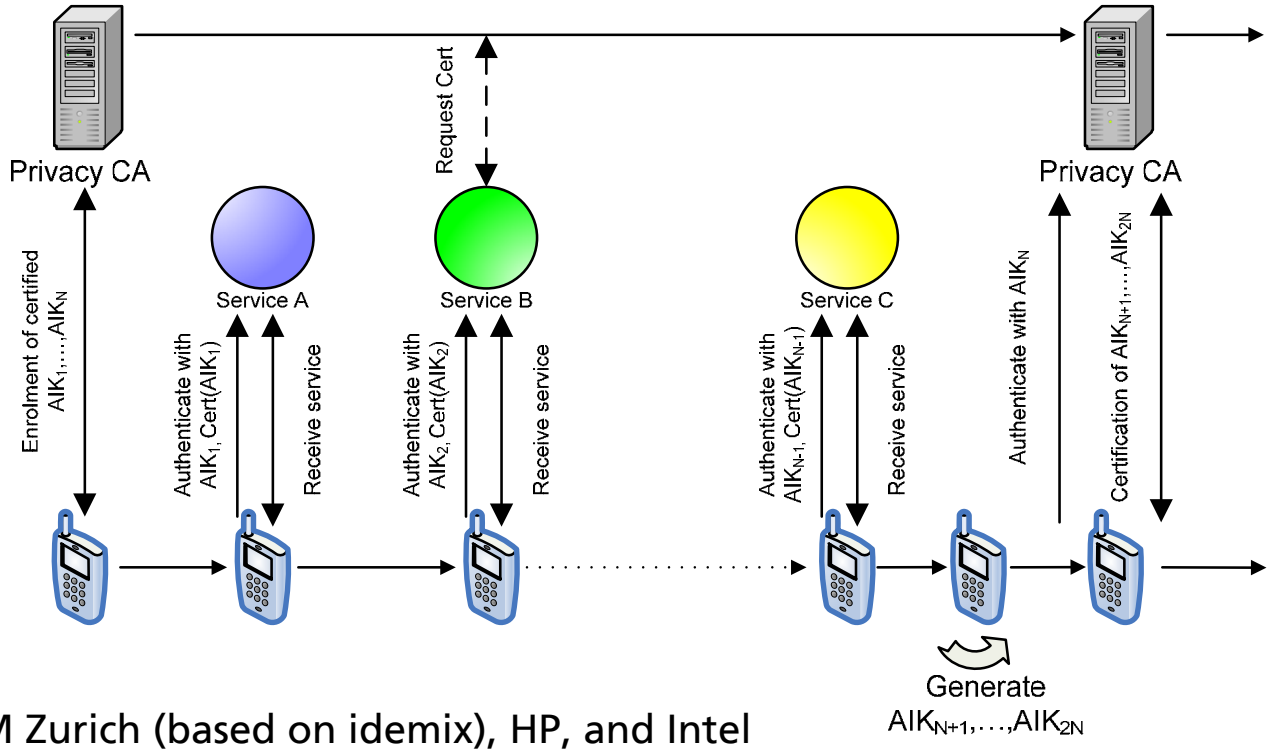
- RA therefore seems to have higher practical feasibility in the mobile domain

A bit of privacy by RA and AIKs

Assume AIKs are used for AA to service access

Then principals can annul the pseudonym annd identify users, by 1to1 association of genuine credentials to TPs (SIM)

Improvement: Use One-time AIKs (like one-time PIN/TANs) to prevent accumulation of profiles by principals and/or service providers



- Better: use DAA, created by IBM Zurich (based on idemix), HP, and Intel
- DAA enables to prove the same assertions as RA, without revealing the platform identity at all
- Needs initial enrolment with a trust domain and principal
- DAA is not used yet

Trust credentials

By RA (or DAA), a an agent a which is a TP can establish a trust credential t_a, embodying three fundamental assertions

1. The presence of a live and unaltered TPM.
2. The integrity of the system and its components.
3. That an existing credential $c_{a,A}$ is unaltered.
   (Established by trusted system software and components to access it)

3. builds on 1. and 2.

Applied categories of transitive trust

- **Restriction**
  An agent a in the domain of principal A is put in a subclass a‘,
  e.g. privy to special services or content

- **Subordination**
  An agent b (previously not in A‘s domain) is incorporated in it
  by referral through an agent a in that domain who vouches for him

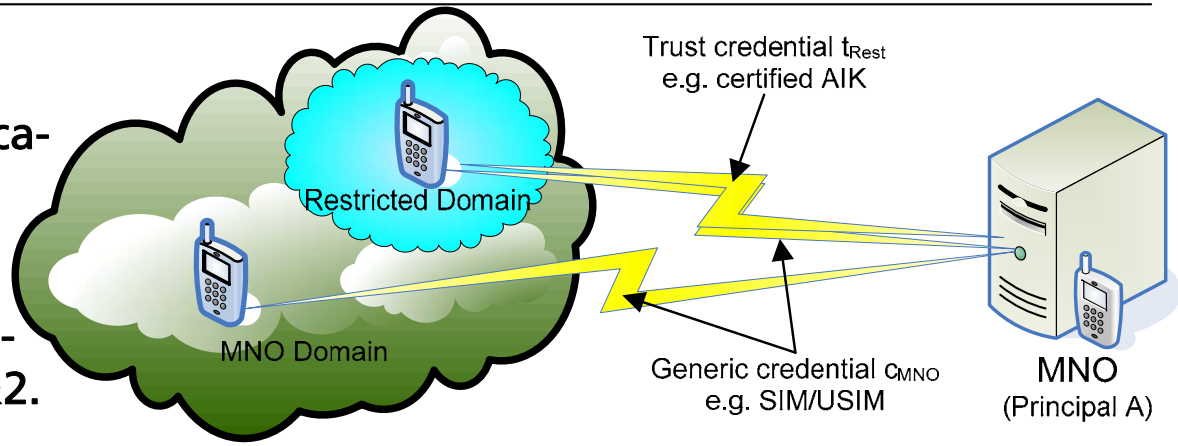
- **Transposition**
  authentication of agent b w.r.t. her own principal B is mediated through
  agent a in domain A

Categories centred on agents a in principal A‘s (the MNO‘s) domain
and involve increasing number (0,1,2) of other subjects

Mind: Applied categories, *not orthogonal* (e.g. transposition can sometimes be decomposed in twice subordination). Theoretical refinement seems possible

Restriction

- Restriction places agents a in a subgroup a'<a, by dual authentication, with generic credential $c_{a,A}$ and trust credential $t_{a'}$.
- $c_{a,A}$ and $t_{a'}$ are used independently, thus needs only assertions 1.&2.



- Restriction can be implemented in many ways:
  AIKs, ACLs, shared secrets or individual credentials residing in trusted storage space,…
- Security in restriction:
  - $t_{a'}$ may be stronger than $c_{a,A}$, but basic network access usually still requires $c_{a,A}$
  - Stronger authentication makes a'-agents privy to special services and/or content
  - Combination of credentials raises resilience against cloning (by checking consistency of creds)
  - Enrolment is key, highest security (against cloning) is only achieved if both $c_{a,A}$ and $t_{a'}$ are individualised and impressed under control of A – balance with privacy

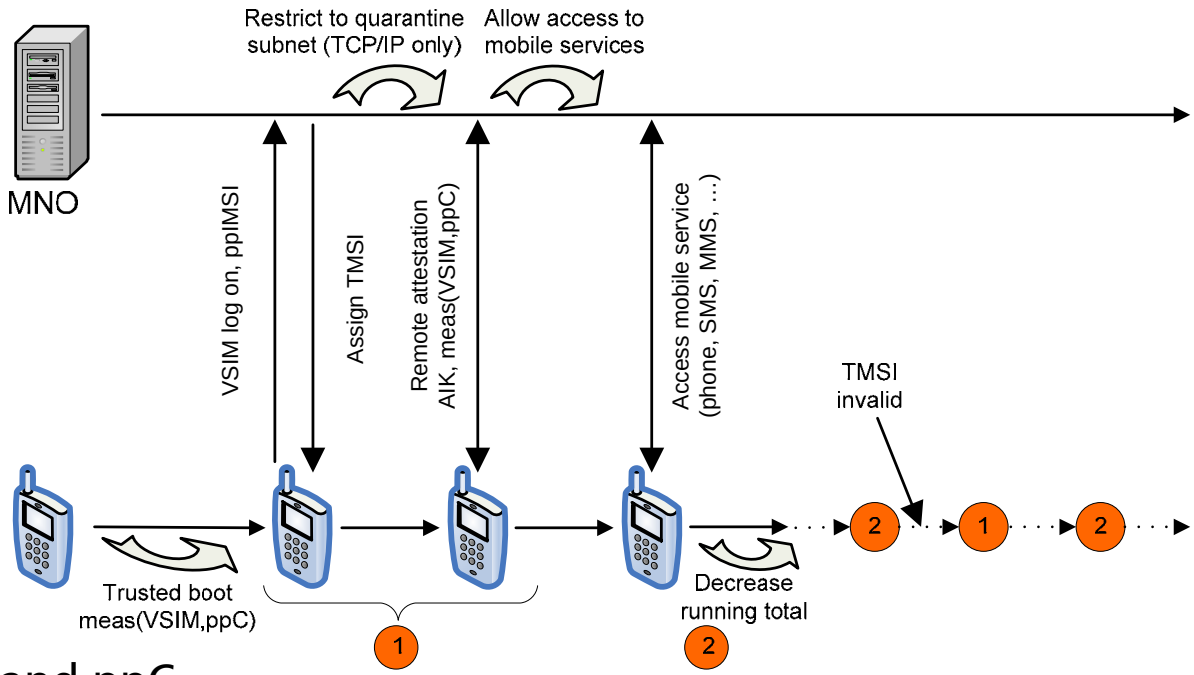
Restriction applications

Restriction is a general concept with manifold applications, a major instance of which is, from an MNO's viewpoint, and in accordance with statements from the industry

- Functional restriction
  - Finer-grained than SIM-lock
  - enables the production of single device with many appearances (cost-efficient)
  - Model appearance can be determined at roll-out or even at the POS (e.g. by user activation)
  - Dynamic, seamless up- and downgrading according to customer SLAs
  - High enforcement level. (This is as well the basis for DRM proper)
  - Location-based restriction, e.g. to counter industrial espionage
  - On-device management & transfer of sensitive user data (photos, messages,…)
  - …

Application of restriction: 'Anonymous' prepaid device

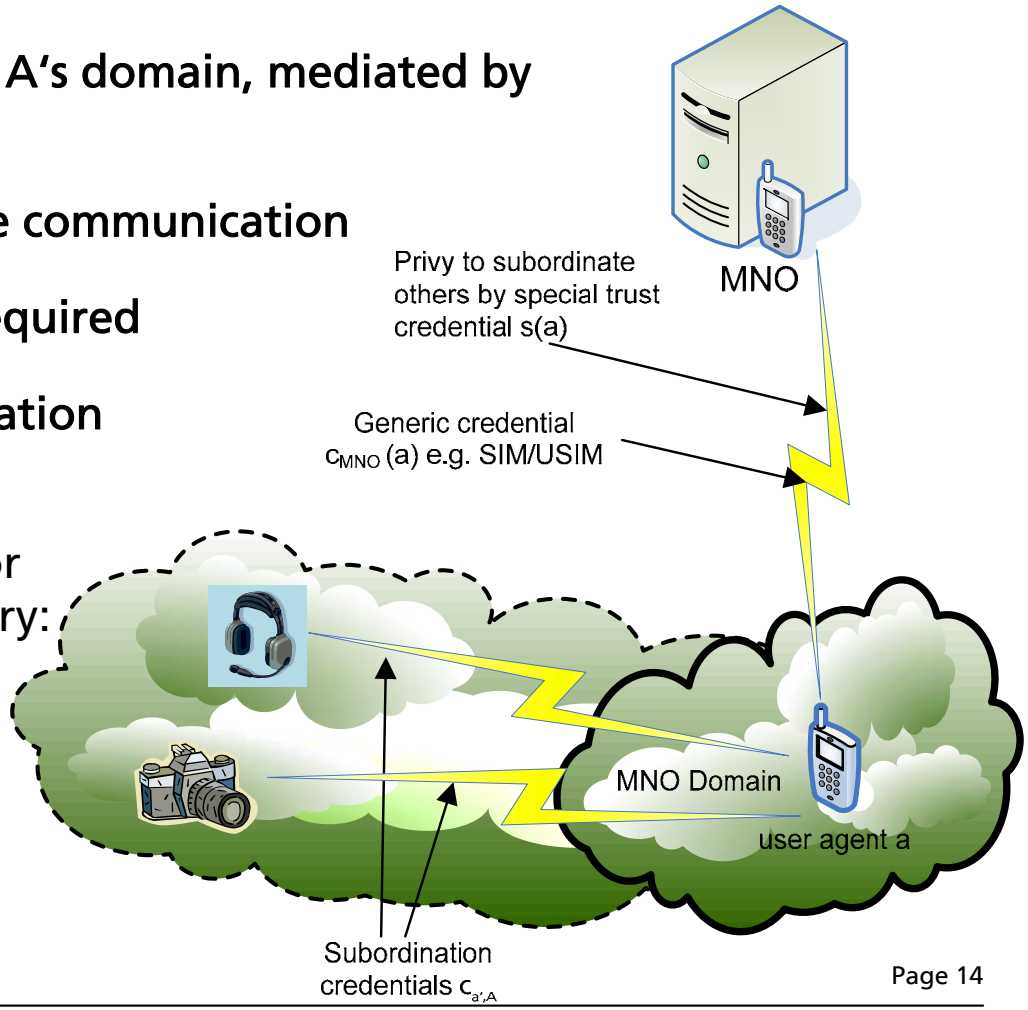
- A prepaid mobile device:
- Running total managed on device – no central accounting
- User can remain anonymous (not legal in the EU)
- Uses virtual SIM (VSIM) and a trusted prepaid client (ppC)



- Modified network log on 1. attests to the integrity of VSIM and ppC, after which access to network services is granted (2.), as usual using only a TMSI
- MNO can demand frequent re-attestation (e.g. by invalidating TMSI)
- Cheap one-way devices or recharging via third party SP

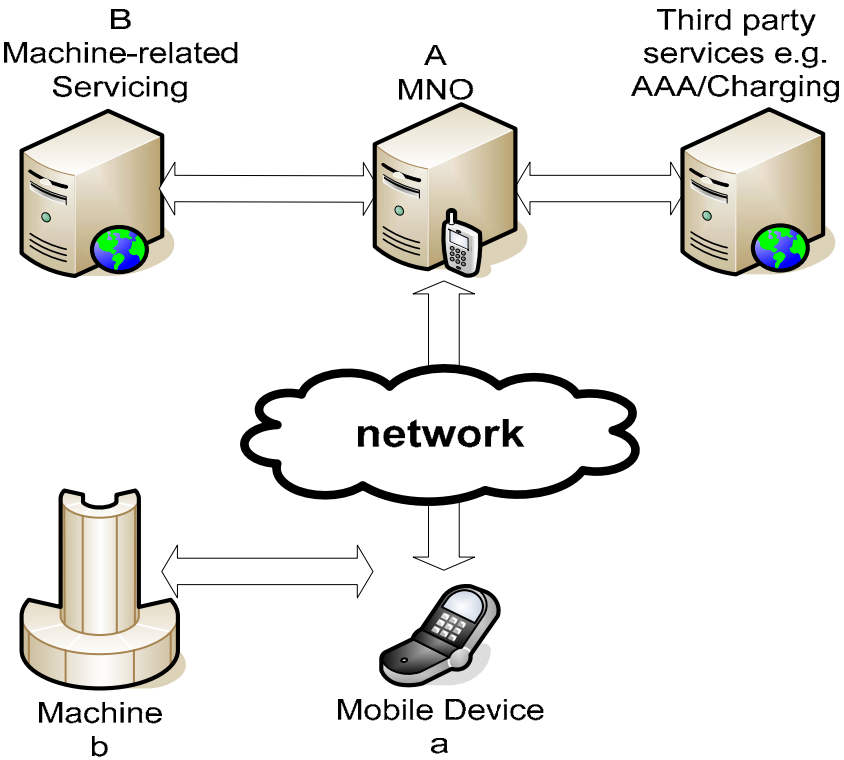
Subordination

- Subordination introduces new agents to A's domain, mediated by existing agents
- Subordinated agents can use short-range communication
- Direct communication to principal not required
- A can but need not partake in authentication
- Many possible variants
  - If a dedicated credential $c_{a',A}$ is used for sub-devs, trusted assertion 3. is necessary:
  - an existing credential $c_{a',A}$ is unaltered.
- Prime example: Bonding of accessories to mobile devices
- Extends range of SIM-lock – ‚customer retention'



Transposition

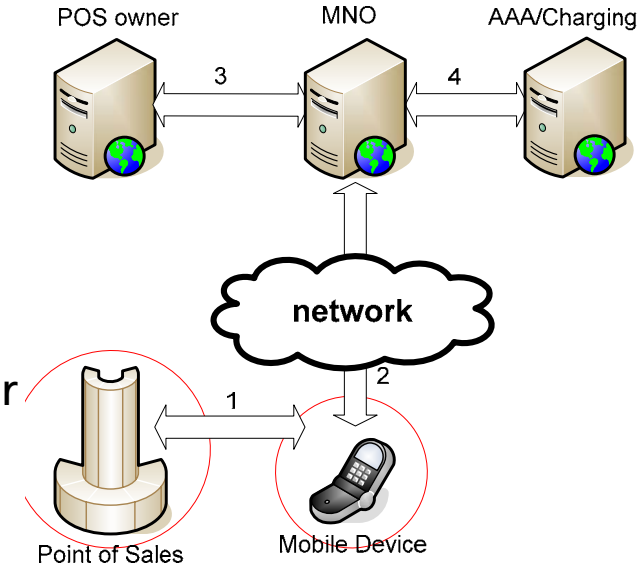
- Transposition can make sense if b cannot connect directly to principal B
- Mobile device a and machine b mutually authenticate using trust credentials t_a, t_b
- They thus establish a secure channel to convey b's generic credential $c_{b,B}$ to principal B
- Assertion 3. proves that $c_{b,B}$ is unaltered
- Variants of authentication of b toward B can involve A, depending on trust (e.g. contractual relationships)



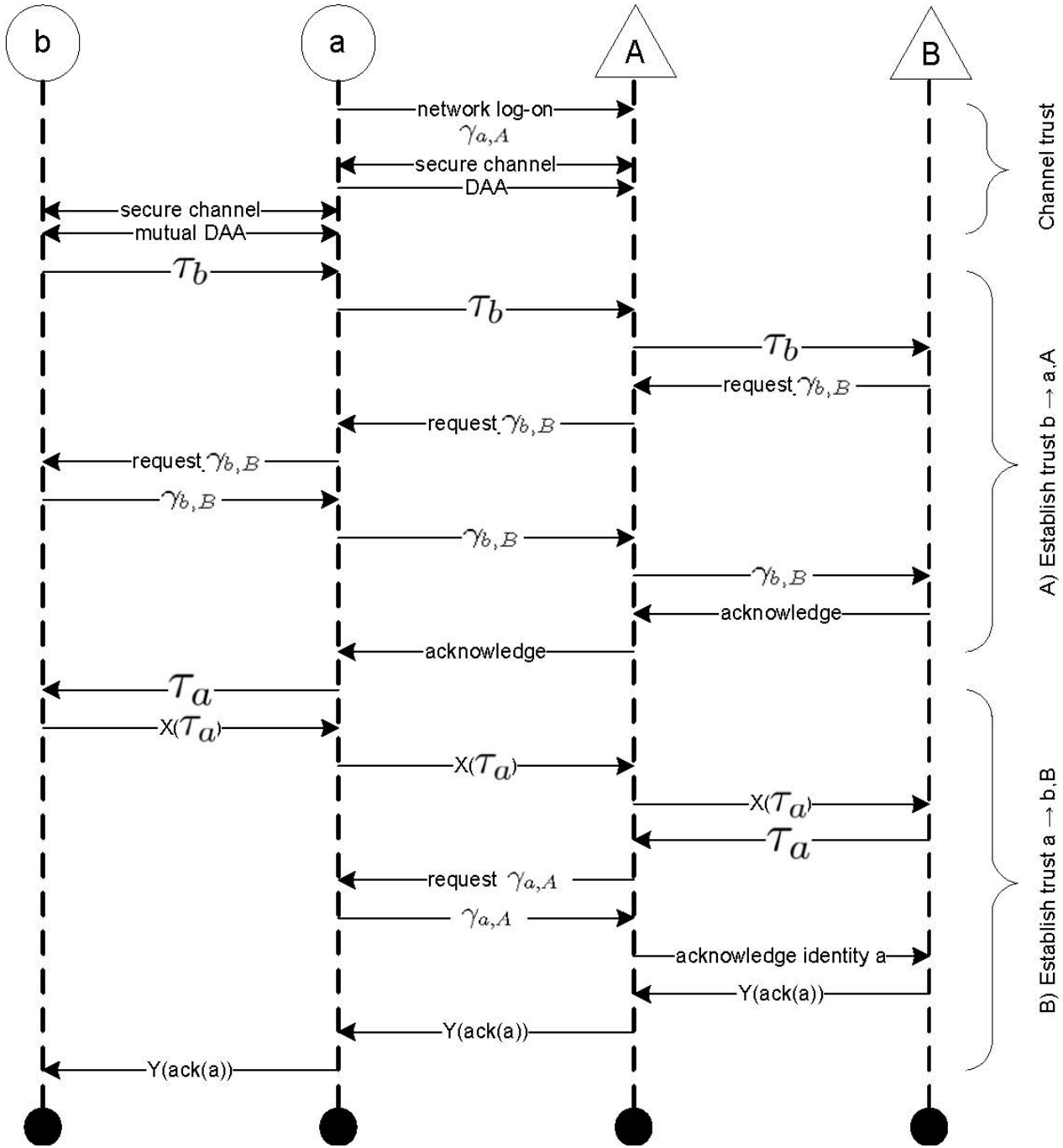
- AA can even be decentralised, i.e., left to agents a, acting as deputies
- Balance gains by outsourcing with secrecy
- Third party services, e.g., for accounting and charging can be included

Transposition application: Point of Sales

- (1) Mutual proof of integrity between POS and mobile device as trust base of the purchase operation
  Device and POS exchange price lists
  and payment modalities
- POS has to verify the device's authentication by connecting the POS owner infrastructure (via the mobile device). Alternatively POS connects the charging provider
- (2) Signed price and payment processing info is transferred to the MNO
- (3+4) The charging data is transferred to the POS owner where a special data package for the charging provider is generated
- After the confirmation of the charging the POS owner (or the MNO) acknowledges the purchase and the POS vending machine delivers the good.



Transposition realisation



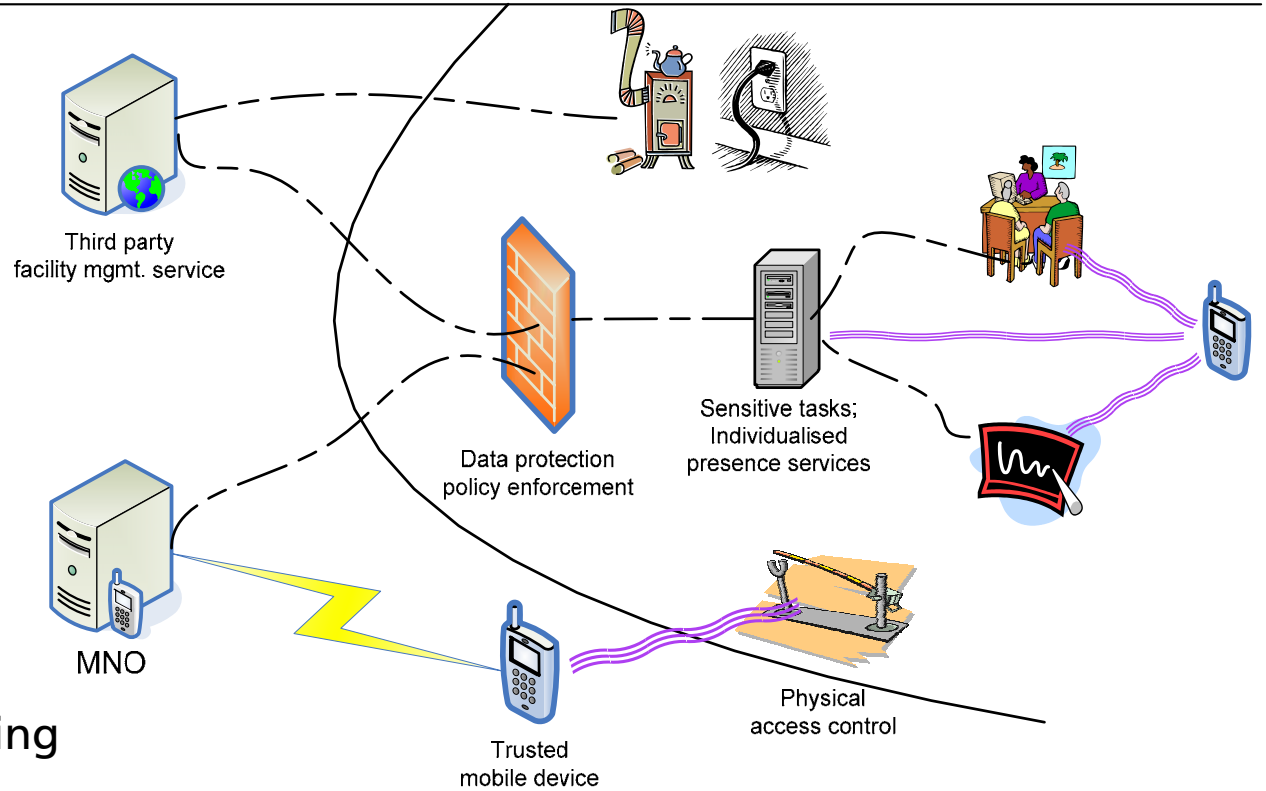
- The sequence shows a realisation variant
- It establishes maximal mutual trust: Both principals A and B can trust the involved agent of the other domain, resp. b, and a
- It is equivalent to two subordinations, with exchanged roles
- Many other variants are possible

Minimal need to know principle in transposition

- The POS owner wants to hide his business secrets from the MNO, e.g.
  - the location and number of its POS endpoints, sales volumes, and price structure
- The MNO likes to protect the privacy of customers w.r.t. the POS owner
  (and maybe even the charging provider)
- Individual identities of POS and device need not be revealed in the purchase process
  - Using the TC concept of a privacy CA and AIKs
  - The AIK can be used in combination with the PCA certificate as a pseudonym of the platform, e.g., one per purchase
  - POS and device can change their identity after a certain time
- Fine separation of duties (POS owner / MNO / charging provider) is helpful
  - e2e encryption protects individual communications
- Price lists need only be exchanged between POS and mobile device
  - Established trust assures that they won't leave the device
- Advanced scenarios may employ DAA

Integrated scenario: Facility management

- MNO offers services to fac. Managers
- No more specialised tokens to access a building – standard mobile devices can be used
- Authentication at gates is essentially transposition
- Functional restriction to, e.g., disable cameras and suppress MMS within building



- Tasks within the building can be fulfilled using mobile's short-range comm.
- Can save network infrastructure in the building

Trusted Computing research and application potential – traits of the technology and research perspectives

- The public's negative impression of TC is gradually changing,
  e.g., data-protection agencies note TC's potential for privacy-protection.
  TC has been functionally and organisationally separated from DRM

- Mobile devices and laptop PCs will soon provide a broad base of TC-equipped user agents

- The relationship between privacy, data protection, and TC should be further examined:
  privacy is not in opposition to TC, but rather privacy protection can benefit from TC
  (by, e.g, separation of duties, implementation of ‚minimal need to know' principles)

- TC can provide a de-centralised trust infrastructure,
  transgressing technical boundaries between, eg., authentication domains and methods –
  research on a fundamental and applied level is needed

- TC has great potential *in combination* with other technologies
  like RFID, mobile devices, PKI, idenitity management (IDM)

- TC has a potential to partially replace resp. complement or co-operate with PKI and IDM

Trusted Computing research and application potential – application and economic perspectives

- **TC supports two emerging and ongoing trends in ICT**
  - horizontal integration of access technologies
  - movement from closed to open systems in business environments

- **TC application should be explored in various (economic) sectors such as**
  - Mobile (broad user base, established AAA infrastructure)
  - E-Government (e-procurement, government rights management, controlled publication of data)
  - E-commerce (user-to-user transactions, commercial grade signatures)

- **TC can be an enabler for new business models and market mechanisms, e.g.**
  - de-centralisation of trust transactions, recommender systems, …),
  - integrated multi-VAS (provider) businesses
  - peer-to-peer, and superdistribution-based markets
  - Web 2.0+ business

We are looking for research partners!

- TC research in Europe is centred around the single outstanding project OpenTC (www.opentc.net, IST-027635)
  - integrated project with 23 partners'
  - approx. 17.1M€ total volume, funded within the last call of FP6,
  - running from Nov. 2005 – mid-2008.

- OpenTC aims at:
  - common infrastructure and OS basis (Linux based)
  - publicly available open source toolbox

- There will be room for applied TC research in FP7!
- The first FP7 call will be issued by December 2006.
- Anyone who would like to join us?

(participation of ZA institutions will be possible under the EU-ZA co-operation treaty and INCO policies)

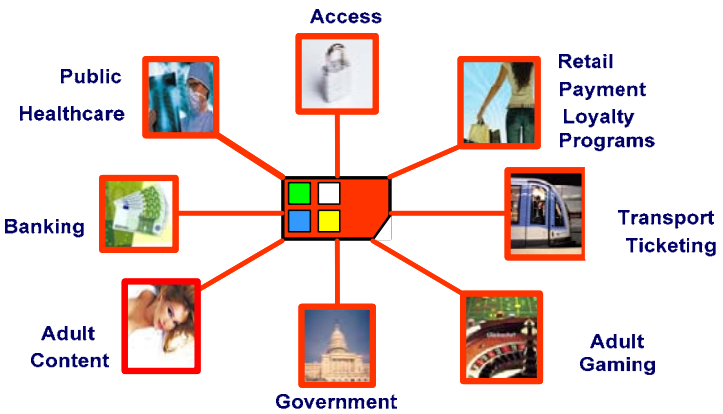
Contact: Dr. A. U. Schmidt, Fraunhofer SIT, Rheinstrasse 75, 64295 Darmstadt, Germany
e-mail: andreas.u.schmidt@sit.fraunhofer.de, phone: +49 (0) 6151 869 60227

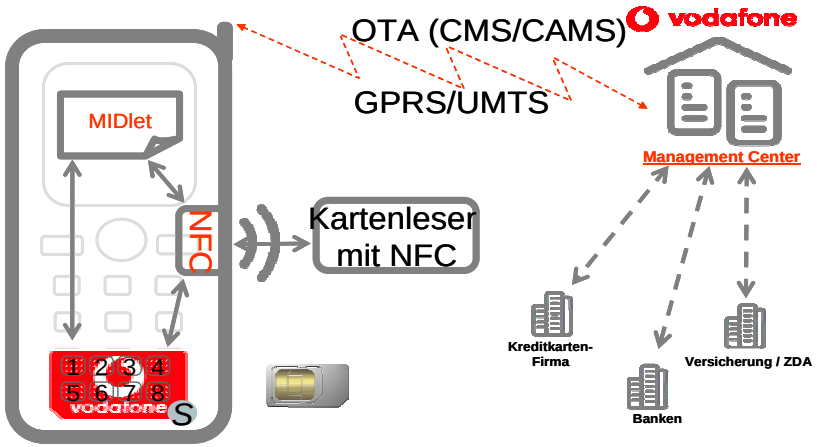
Fraunhofer SIT Institut Sichere Informations-Technologie

megaSIM vs. mTPM 1/3
The megaSIM, a powerful security device

- USIM functionality, security, and high-density flash
- fully functional SIM cards
- SIM card form factor
- offer secure personal storage in a two-chip solution
- Secure smart card controller with additional crypto engines, crypto libraries and MPU provide the highest level of secure operations
- A secure, high-performance controller, with embedded mask ROM, is based on a 32-bit ARM SC100 secure core. The controller empowers mSIM cards with multi-tasking capabilities.
- NAND flash memory and advanced crypto engines offer high-performance, low-power operation and secure personal storage.

megaSIM vs. mTPM 2/3
megaSIM potential application





A ‚Vision' of Vodafone Germany

- Goes so far as to propose megaSIM-enabled handsets as platforms for qualified electronic signatures
- Common handsets, Multi- application- SIM card as certified platform for service access, enables novel services
- Ongoing projects in Germany and Europe

  Ticketing, Payment, access control, Health service card
- Integration of VAS in the health sector

  Tele-medicine / Tele-monitoring

  Individualised care offerings, prevention, outpatient care,

  Fitness, Wellness, home care, using electronic patient files

  Assistance services based on LBS and GPS
- Will MNOs stand in for higher costs of megaSIMs?

megaSIM vs. mTPM 3/3
Assessment: Some arguments in favour of the mTPM

- **Better cost/efficiency ratio**
  - By the unique feature of trusted boot, mTPM can use the full resources of the device for security and services
  - While a megaSIM is a monolith – admittedly more powerful than saingle mTPM
  - For this sole reason, the mTPM is likely to even be cheaper than a megaSIM

- **Virtualisation is a powerful TC concept that can be used in conjunction with an mTPM to realise the concept of secure compartments**
  - Various vendor, MNO, and service provider specific trust and AAA can be realised with no physical limits
  - Even xSIM can be virtualised in TC-enabled devices
  - While a megaSIM offers only limited (security service) scalability

- **The MNO-centric business model to sell or rent megaSIM compartments to third party SPs, does as well apply to mTPM** – only better for the mentioned reasons

- **A little discussion at ETRICS 2006:**
  - A. U. Schmidt was asked after his talk: ‚Do You think that mTPM will soon be incorporated in a megaSIM?'.
  - Answer: ‚Such a device would perhaps not be removable anymore, since otherwise the trusted boot concept would be void. If it is not removable, why would anyone have two physical security anchors in one device (or are we building hybrid cars)?'
  - IMHO: Rather virtual xSIMs will run in an mTPM protected environment – as legacy applications, alongside with virtual smart cards, etc.