
Transitive trust in mobile scenarios

Horizontal integration of access technologies needs convergence of authentication technologies. The concept of establishing transitive 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**

ETRICS 2006 – Freiburg, 7. June 2006

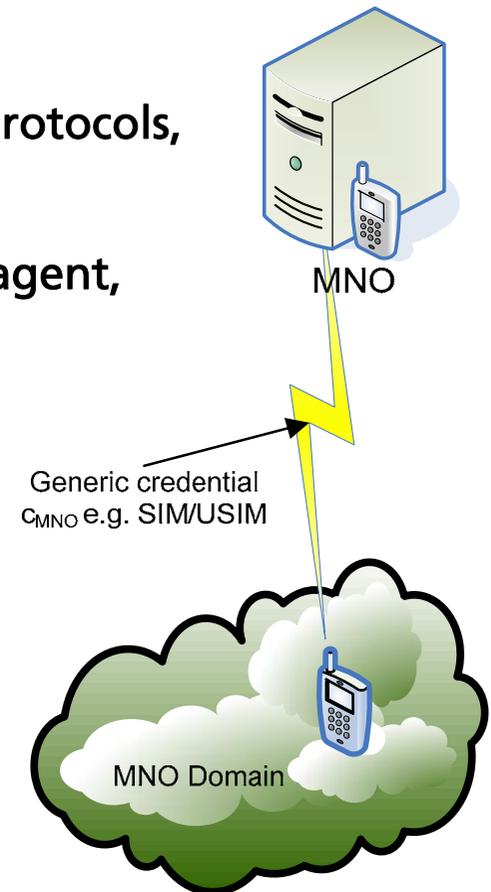
Trust in mobile service access

- 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**

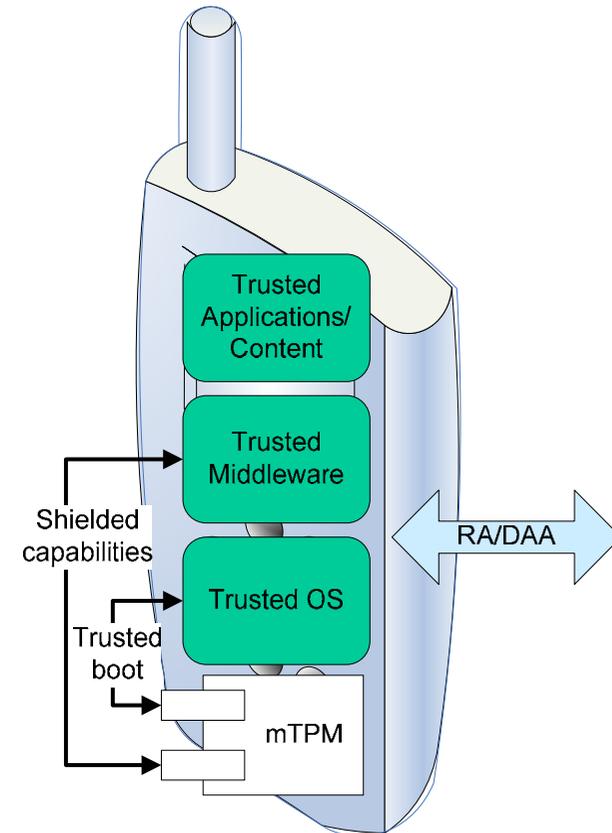


Trusted mobile agents

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

Memory curtaining, sealed storage, and secure I/O

- TPM offers (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 and sealed storage spaces 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.



Remote attestation

- The TPM is a hardware (or virtualised) **root of trust** on which the two essential **attestation protocols** rest:
 - **Remote attestation (RA)**, yielding, 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

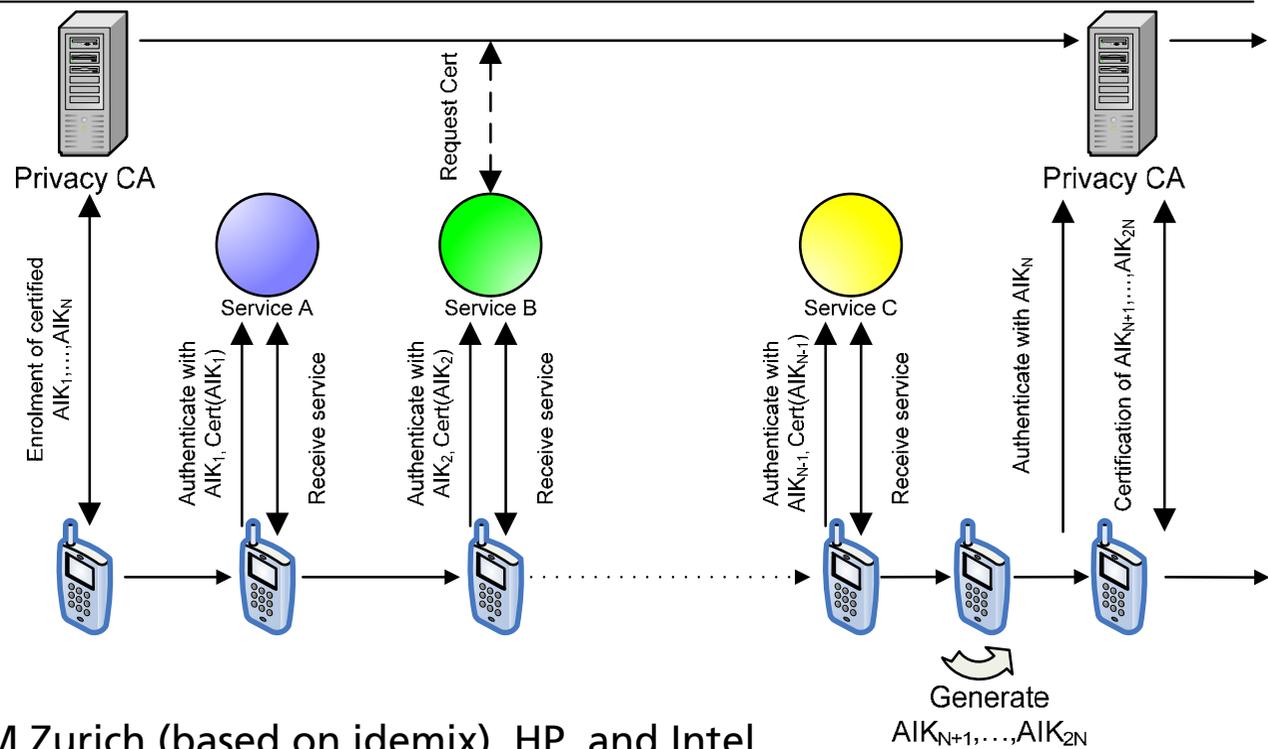
- The verifier has to know a reference value for the reported values in the transmitted measurement log and PCR value
- This reference base is rather large and hard to maintain in the PC domain:
There are many
 - hard- and software versions
 - admissible boot and runtime parameters
 - frequent updates
- 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 and 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 $a' \subset a$, by dual authentication, with generic credential $c_{a,A}$ and trust credential $t_{a'}$.

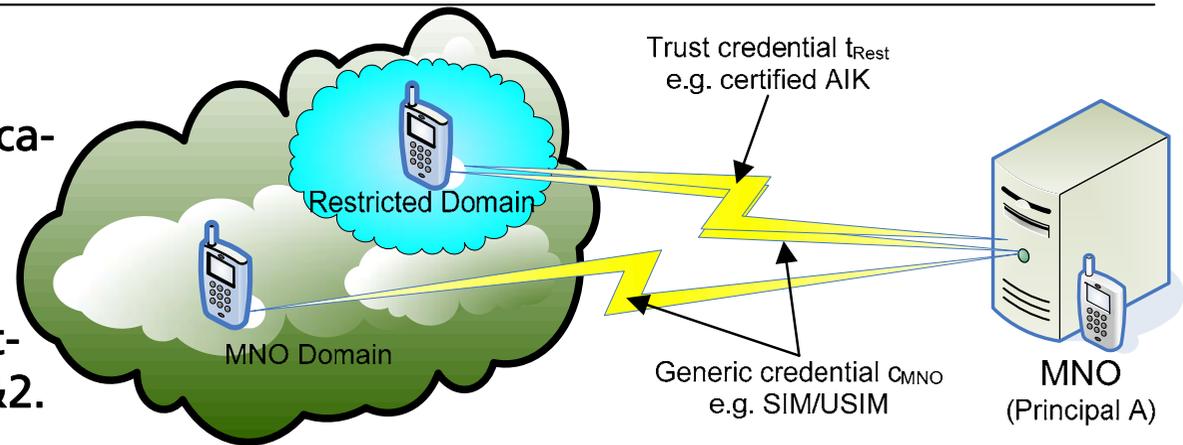
➤ $c_{a,A}$ and $t_{a'}$ are used independently, thus needs only assertions 1.&2.

➤ 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

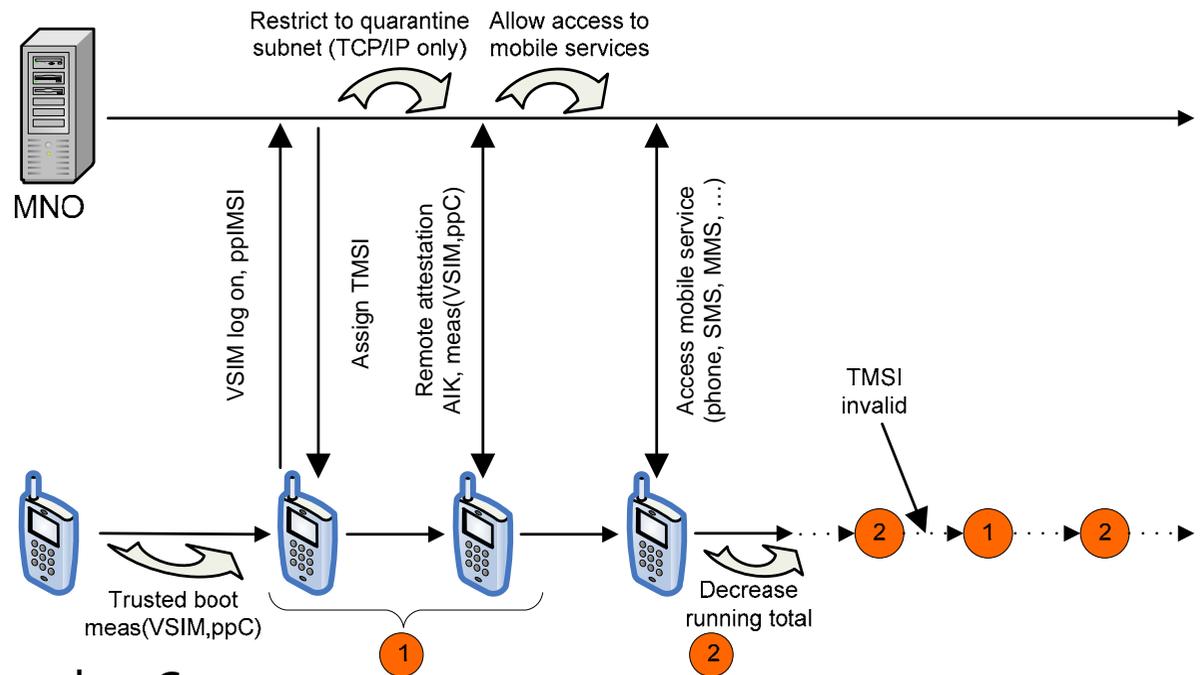
➤ User can remain anonymous (not legal in the EU)

➤ 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

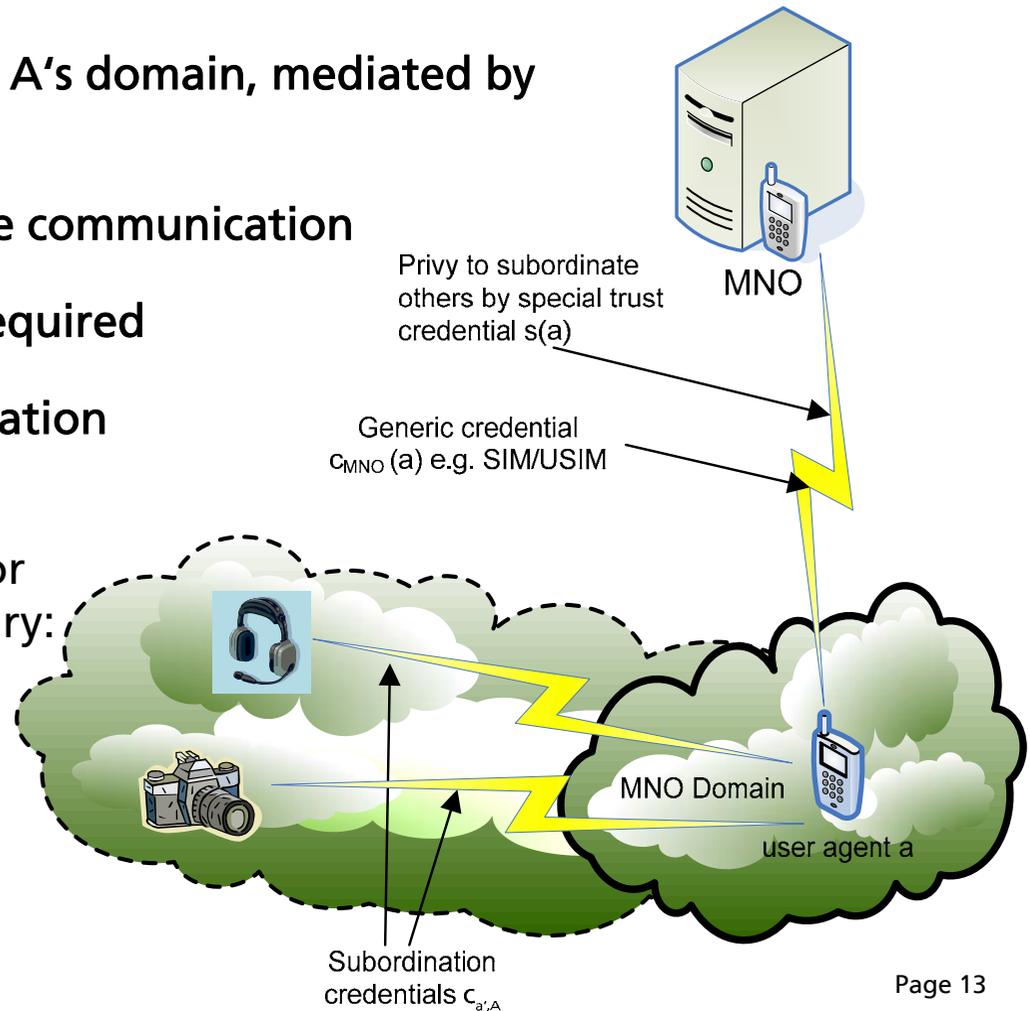
➤ MNO can demand frequent re-attestation (e.g. by invalidating TMSI)

➤ Cheap one-way devices or recharging via third party SP



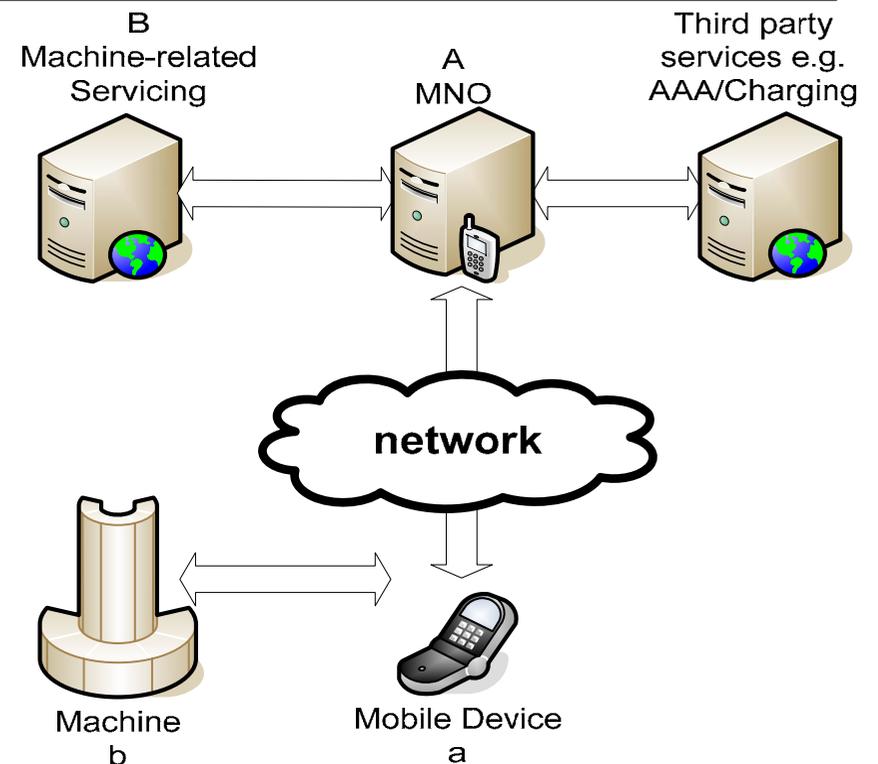
Subordination

- **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‘



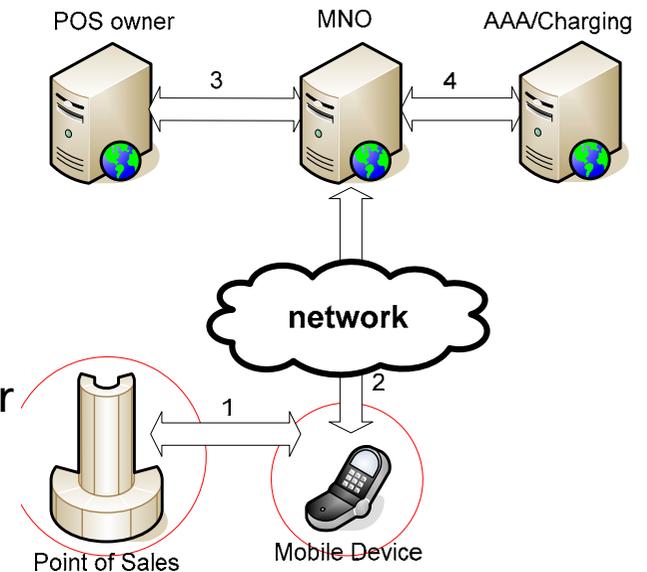
Transposition

- **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

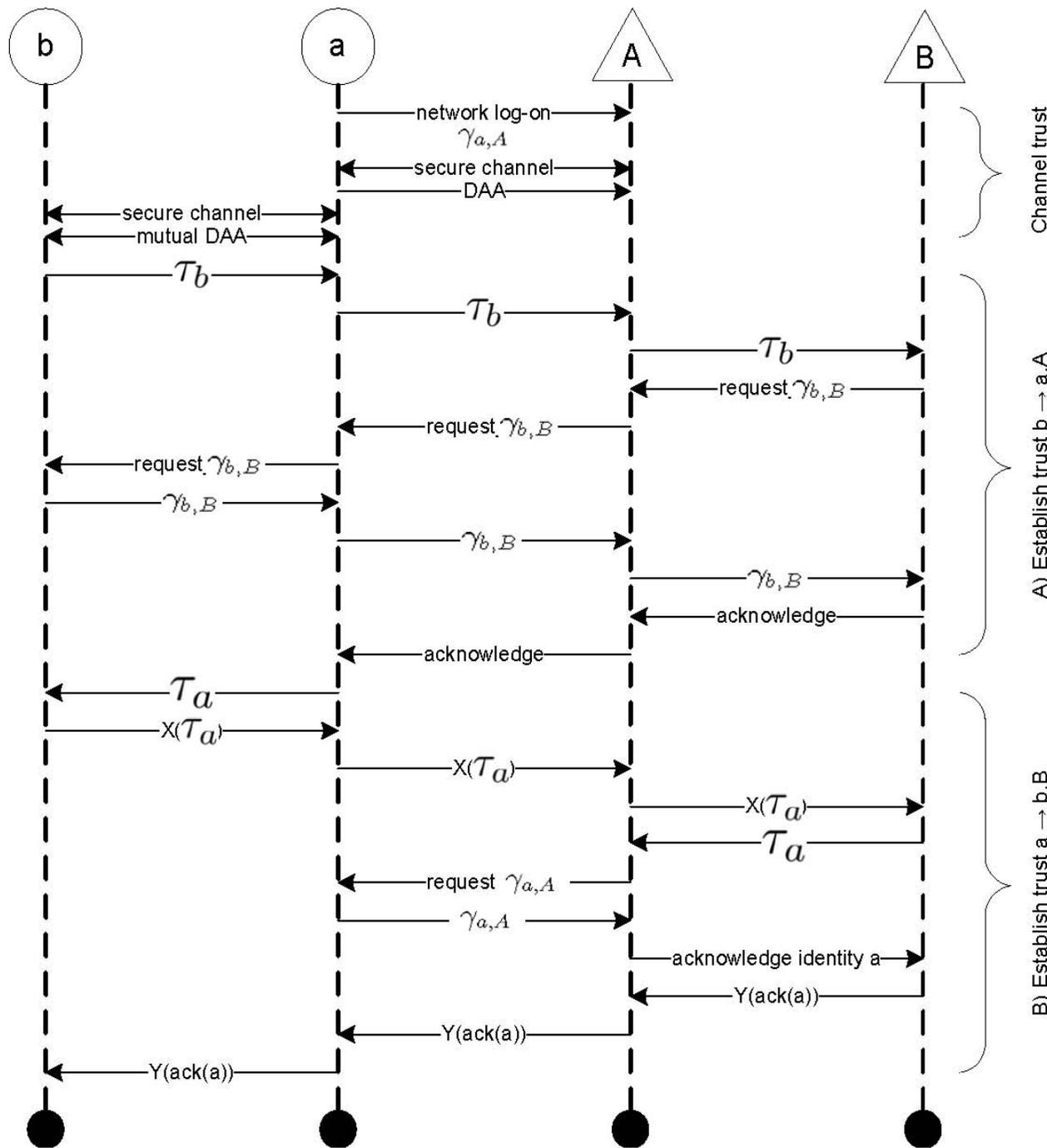


Transposition application: Point of Sales

- (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

