A Protocol’s Life After Attacks
let’s investigate beyond

Giampaolo Bella

Stefano Bistarelli

Fabio Massacci

@ Cambridge, Catania, Pescara, Pisa, Trento

Current verification setting

Yes! I found an attack!

Model Checker

I can assure there's no such!

Theorem Prover

Focus is in fact on "THE attack".
Is this all??

Glance at the physical world

We own a bakery, and one morning we find the window smashed. We can:

1. Suspect: no-one's around – it could have been any passer-by ☹
2. Detect: the burglars are still there, and no-one else's around – it was them! ☺
3. Retaliate: the burglars are caught and punished accordingly – by appropriate measures!! ☺☺

Idea: apply same concepts to security protocols

How and Why

- How? Must continue analysis after “THE attack”

  For example:

  - Model Checkers: If I find an attack, is there another one? (retaliation)
  - Theorem Provers: If I assume there is an attack, could anyone else mount the same attack? (detection)

- Why? Can get novel insights about protocols

  For example:

  - Is it really convenient to attempt attacks?
  - Do we need to redesign, or the bad guys are stopped by realistic threats?
  - What if the principals change their behaviours?

Example

Take Lowe’s middle-person attack on NS:
if *A* executes with *C* then *C* impersonates *A* with *B*

- Consequence (Lowe):
  if *B* is a bank, *C* can steal from A’s account
  C→B : {Na, Nb, “Transfer £1000 from A’s account to C’s}$_{Kb}$
- Extra consequence (last year’s workshop):
  if *A* is a bank, *B* can steal from C’s account
  B→A : {Na, Nb, “Transfer £1000 from C’s account to B’s}$_{Ka}$

Principals’ behaviours

Principals are divided according to their behaviours into three disjoint sets.

Good: $\mathcal{G}$ – conform to the protocol

Bad: $\mathcal{B}$ – attempt to break the protocol

Ugly: $\mathcal{U}$ – conform to the protocol but would collaborate with bad

Crucially: principals may decide to change behaviour!

Traces and attacks

Trace $\mathcal{T}$: conventional view of protocol history as log (of events or messages, or…)

Projection $\mathcal{T}/\mathcal{A}$: subtrace of $\mathcal{T}$ where some agent in $\mathcal{A}$ acted

Attack $\mathbf{A}$: some predicate $\mathbf{A}(\mathcal{T},\mathcal{G},\mathcal{B},\mathcal{U})$

Can make Spy, owner of the network, explicit.

Current verification setting
(more formal)

P vulnerable to $\mathbf{A}$ against $\mathcal{G}$ if $\exists\ \mathcal{T}\in P.\ \mathbf{A}(\mathcal{T},\mathcal{G},\mathcal{B},\mathcal{U})$

Model Checker

P immune to $\mathbf{A}$ against $\mathcal{G}$ if $\nexists\ \mathcal{T}\in P.\ \mathbf{A}(\mathcal{T},\mathcal{G},\mathcal{B},\mathcal{U})$

Theorem Prover

Retaliation

A protocol P allows *retaliation* of an attack $\mathbf{A}$ by $\mathcal{B}$ if

$\forall\ \mathcal{T} \in$ P, $\mathcal{G},\mathcal{B},\mathcal{U}$ s.t. $\mathbf{A}(\mathcal{T},\mathcal{G},\mathcal{B},\mathcal{U})$,

$\exists\ \mathcal{T}_r \in$ P extending $\mathcal{T}$,

$\exists\ \mathcal{G}',\mathcal{B}',\mathcal{U}'$ s.t. $\mathcal{B}' \subset \mathcal{G} \cup \mathcal{U}$ and

s.t. $\mathbf{A}(\mathcal{T}_r,\mathcal{G}',\mathcal{B}',\mathcal{U}')$

- if $\mathcal{B}' = \mathcal{G}$ *direct retaliation*
  - else, if $\mathcal{B}' \cap \mathcal{G} \neq \varnothing$ *combined retaliation*
    - else, if $\mathcal{B}' \subset \mathcal{U}$ *arbitrary retaliation*

Appears suitable for theorem proving…

Example (mvore formal)

Lowe's middle-person attack on NS:
if A executes with C then C impersonates A with B

- Consequence (Lowe):
  if B is a bank, C can steal from A's account
- Extra consequence (last year's workshop):
  if A is a bank, B can steal from C's account

Whenever

$$\mathbf{A}(\mathcal{T}, \mathcal{G}{:=}\{B\}, \mathcal{B}{:=}\{C\}, \mathcal{U}{:=}\{A\})$$

$\mathcal{T}$ can be extended as $\mathcal{T}_r$ s.t.

$$\mathbf{A}(\mathcal{T}, \mathcal{G}{:=}\{A\}, \mathcal{B}{:=}\{B\}, \mathcal{U}{:=}\{C\})$$

No Retaliation

A protocol P allows *no retaliation* of an attack $\mathbf{A}$ by $\mathcal{B}$ if

$\exists$ $\mathcal{T} \in$ P, $\mathcal{G},\mathcal{B},\mathcal{U}$ s.t. $\mathbf{A}(\mathcal{T},\mathcal{G},\mathcal{B},\mathcal{U})$,

$\forall$ $\mathcal{T}_r \in$ P extending $\mathcal{T}$,

$\forall$ $\mathcal{G}',\mathcal{B}',\mathcal{U}'$ s.t. $\mathcal{B}' \subset \mathcal{G} \cup \mathcal{U}$ and $\mathcal{B} \subset \mathcal{G}' \cup \mathcal{U}'$

holds $\neg\mathbf{A}(\mathcal{T}_r,\mathcal{G}',\mathcal{B}',\mathcal{U}')$

Appears suitable for model checking

Detection

A protocol P allows *detection* of an attack $\mathbf{A}$ by $\mathcal{B}$ if

$\forall$ $\mathcal{T}\in$P, $\mathcal{G},\mathcal{B},\mathcal{U}$ s.t. $\mathbf{A}(\mathcal{T},\mathcal{G},\mathcal{B},\mathcal{U})$,

$\forall$ $\mathcal{T}_r\in$P s.t. $\mathcal{T}/\mathcal{G} = \mathcal{T}_r/\mathcal{G}$ holds $\mathbf{A}(\mathcal{T}_r,\mathcal{G}',\mathcal{B}',\mathcal{U}')$

Appears suitable for theorem proving…

No Detection

A protocol P allows *no detection* of an attack $\mathbf{A}$ by $\mathcal{B}$ if

- $\exists\ \mathcal{T} \in$ P, $\mathcal{G}, \mathcal{B}, \mathcal{U}$ s.t. $\mathbf{A}(\mathcal{T}, \mathcal{G}, \mathcal{B}, \mathcal{U})$,
  - $\exists\ \mathcal{T}_r \in$ P s.t. $\mathcal{T}/\mathcal{G} = \mathcal{T}_r/\mathcal{G}$ and $\mathcal{T}_r \neq \mathcal{T}$ holds $\neg\mathbf{A}(\mathcal{T}_r, \mathcal{G}', \mathcal{B}', \mathcal{U}')$

Appears suitable for model checking…

Suspicion

A protocol P allows *suspicion* of an attack $\mathbf{A}$ if

$\forall\ \mathcal{T} \in P,\ \mathcal{G}, \mathcal{B}, \mathcal{U}$ s.t. $\mathbf{A}(\mathcal{T}, \mathcal{G}, \mathcal{B}, \mathcal{U})$,

$\quad \forall\ \mathcal{T}_r \in P$ s.t. $\mathcal{T} / \mathcal{G} = \mathcal{T}_r / \mathcal{G}$

$\quad\quad \exists\ \mathcal{B}', \mathcal{U}'$ s.t. $\mathcal{B}' \neq \mathcal{B}$ and $\mathcal{U}' \neq \mathcal{U}$

$\quad\quad\quad$ s.t. $\mathbf{A}(\mathcal{T}_r, \mathcal{G}, \mathcal{B}', \mathcal{U}')$

Appears suitable for theorem proving…

No Suspicion

A protocol P allows *no suspicion* of an attack $\mathbf{A}$ if

$\exists\ \mathcal{T} \in$ P, $\mathcal{G}, \mathcal{B}, \mathcal{U}$ s.t. $\mathbf{A}(\mathcal{T}, \mathcal{G}, \mathcal{B}, \mathcal{U})$,

$\exists\ \mathcal{T}_r \in$ P s.t. $\mathcal{T} / \mathcal{G} = \mathcal{T}_r / \mathcal{G}$

$\forall\ \mathcal{B}', \mathcal{U}'$ s.t. $\mathcal{B}' \neq \mathcal{B}$ and $\mathcal{U}' \neq \mathcal{U}$

holds $\neg \mathbf{A}(\mathcal{T}_r, \mathcal{G}, \mathcal{B}', \mathcal{U}')$

Appears suitable for model checking…

Conclusions

- There's life after attacks take place!
- Life that is worth investigating
- More complex properties of traces: at least two quantifiers (possibly alternated) where we used to have one only
- Theory now adapted. Can we adapt mechanised tool support?