Information Flow Analysis and Approximate Security

Alessandro Aldini
University of Urbino “Carlo Bo”, Italy
aldini@sti.uniurb.it
Outline

• Introduction: information flow analysis
• Formal framework: modeling the system and the adversary
• Approximating noninterference
• About the most powerful adversary
• Security properties
• Comparison and conclusion
Introduction

Multi-level secure computing

Information flow detection:  
Deterministic/  
Nondeterministic  
Noninterference

- exact analysis  
- 0/1 result

Nondeducibility (Sutherland), Causality (Roscoe)  
Nondeducibility on Strategies (Wittbold&Johnson)  
Generalized Noninterference, Restrictiveness  
(McCullough, McLean)  
Nondeducibility on Compositions (Focardi&Gorrieri)

Covert channel capacity:  
Probabilistic  
Noninterference

- approximate analysis  
Probabilistic noninterference  
(Gray, Dipierro et al., Aldini et al., Smith, ...)

- [0; 1] result  
  best-case/average/worst-case?
Modeling the system and the adversary: abstraction mechanisms

- **probabilistic** (internal) behavior affects probability distributions
- **nondeterministic** (external) behavior acts as a probabilistic scheduler that solves nondeterminism

---

**Adversary**

- Class of adversary strategies $\Rightarrow$
- **Noninterference property**

**Problem:** evaluating the interference of possibly infinite adversary strategies
Modeling the system: GRTS $S = (S, Act, T, s_0)$

$\forall s \in S. \sum\{|p| \exists a \in Act, t \in S : (s, a, p, t) \in T\} \in \{0, 1\}$

$\forall s \in S, \forall a_* \in Act. \sum\{|p| \exists t \in S : (s, a_*, p, t) \in T\} \in \{0, 1\}$
Modeling the interference

- Two security clearances: *high-level* and *low-level*.

- The high-level environment (*Adversary*) should not affect the view of the system observable by the low-level user (*Low*).

- Evaluating the interference of an adversary *A*:
  - analyze the semantics of the system in isolation (no adversary interference).
  - analyze the semantics of the system in the presence of *A*’s interactions.
  - comparing the two semantics views.
What Low observes whenever:

- the adversary does not interact through the high-level input (b).
- the adversary interacts through the high-level input (c).
[Synch. model of comm.] What Low observes whenever:

- the adversary does not interact through the high-level output (b).
- the adversary interacts through the high-level output (c).
Without adversary \([S \setminus A]\)

- The adversary \(A\) does not interact with the system in any way.
- All the transitions labeled with high-level actions are simply removed.

Input: GTRS \(S = (S, Act, T, s_0)\).
Output: GRTS \(S \setminus A = (S', Act, T', s_0)\).

\[
S' := \emptyset; T' := \emptyset; \\
No_{Adv}(s_0); \\
\text{where:} \\
No_{Adv}(s) : \\
S' := S' \cup \{s\}; \\
\quad \text{for each } (s, a, q, t) \in T \\
\quad \quad \text{if } a \notin Act_H \text{ then} \\
\quad \quad \quad T' := T' \cup \{(s, a, q/p(s), t)\}; \\
\quad \quad \text{for each } (s, \_\_ \_ t) \in T' \\
\quad \quad \quad \text{if } t \notin S' \text{ then } No_{Adv}(t); \\
\text{where} \\
p(s) = \sum\{ |p| \exists s' \in S, \exists a \in Act_L \cup \{\tau\}. (s, a, p, s') \in T \}.\]
Simple Adversary \( [S|A] \)

The adversary \( A \) is a probabilistic scheduler which determines \textit{a priori} the behavior of the high-level inputs/outputs.

Strategy on outputs: \( A_g \subseteq H \)
Strategy on inputs: \( A_r \subseteq H \times ]0, 1[ \)

Input: GTRS \( S = (S, Act, T, s_0) \).
Output: GRTS \( S|A = (S', Act, T', s_0) \).

\[
\begin{align*}
S' &:= \emptyset; T' := \emptyset; \\
S_{\text{Adv}}(s_0); \\
\text{where:} & \\
S_{\text{Adv}}(s) : & \quad S' = S' \cup \{s\}; \\
\text{Gen}(s, A_g) : & \quad \text{for each reactive bundle of type } h \in H \text{ enabled at } s \\
& \quad \text{if } (h, p_h) \in A_r \text{ then } \text{React}(s, h, p_h); \\
& \quad \text{for each } (s, \_, \_, t) \in T' \\
& \quad \text{if } t \not\in S' \text{ then } S_{\text{Adv}}(t); \\
\text{Gen}(s, I) : & \quad \text{for each } (s, a, q, t) \in T \\
& \quad \text{if } a \not\in \text{Act}_{\text{H}} \text{ then } T' := T' \cup \{(s, a, q/p(s, I), t)\}; \\
& \quad \text{if } a \in I \text{ then } T' := T' \cup \{(s, \tau, q/p(s, I), t)\}; \\
\text{where } p(s, I) & = \sum \{| p | \exists s' \in S, \exists a \in \text{Act}_{\text{L}} \cup I \cup \{\tau\}. (s, a, p, s') \in T \}; \\
\text{React}(s, h, p_h) : & \quad \text{if the gen. bundle is non-empty at } s \in S' \text{ then} \\
& \quad \text{for each } (s, a, q, t) \in T' \\
& \quad \quad q := q \cdot (1 - p_h); \\
& \quad \text{for each } (s, h^*, q, t) \in T \\
& \quad \quad T' := T' \cup \{(s, \tau, q \cdot p_h, t)\}; \\
& \quad \text{else} \\
& \quad \text{for each } (s, h^*, q, t) \in T \\
& \quad \quad T' := T' \cup \{(s, \tau, q, t)\}; 
\end{align*}
\]
Interactive Adversary $[S|A]$

The adversary $A$ is a state-dependent probabilistic scheduler which determines the behavior of the high-level inputs/outputs depending on the current high-level interface.

Strategy on outputs:
$$A_g : \mathcal{P}(H) \rightarrow \mathcal{P}(H)$$

Strategy on inputs:
$$A_r : \mathcal{P}(H) \rightarrow \mathcal{P}(H \times ]0, 1[)$$

Proposition $A_S \subset A_I$
History-dependent Adversary \([S|A]\)

The adversary \(A\) is a history-dependent probabilistic scheduler which determines the behavior of the high-level inputs/outputs depending on the previous history.

Strategy on outputs:
\[
A_g : Act^* \rightarrow \mathcal{P}(H)
\]
Strategy on inputs:
\[
A_r : Act^* \rightarrow \mathcal{P}(H \times [0,1[)
\]

Proposition \(A_S \subset A_{HD}\)

Input: GTRS \(S = (S, Act, T, s_0)\).
Output: GRTS \(S|A = (S', Act, T', (s_0, \epsilon))\).

\[
S' := \emptyset; T' := \emptyset; \quad H_{Adv}((s_0, \epsilon));
\]
where:
\[
H_{Adv}((s, Tr)) : \\
S' = S' \cup \{(s, Tr)\}; \\
\quad \text{for each } (s, a, q, t) \in T \\
\quad \quad \text{if } a \notin Act_H \text{ then } \\
\quad \quad \quad T' := T' \cup \{((s, Tr), a, q/p(s, A_g(Tr)), (t, Tr.a))\}; \\
\quad \quad \text{if } a \in A_g(Tr) \text{ then } \\
\quad \quad \quad T' := T' \cup \{((s, Tr), \tau, q/p(s, A_g(Tr)), (t, Tr.a))\}; \\
\quad \quad \text{for each reactive bundle of type } h \in H \text{ enabled at } s \\
\quad \quad \quad \text{if } (h, p_h) \in A_r(Tr) \text{ then } \\
\quad \quad \quad \quad \text{if the generative bundle is non-empty at } (s, Tr) \text{ then } \\
\quad \quad \quad \quad \quad \text{for each } ((s, Tr), a, q, \_ ) \in T' \\
\quad \quad \quad \quad \quad \quad q := q \cdot (1 - p_h); \\
\quad \quad \quad \quad \quad \quad \text{for each } (s, h_*, q, t) \in T \\
\quad \quad \quad \quad \quad \quad T' := T' \cup \{((s, Tr), \tau, q \cdot p_h, (t, Tr.h_*))\}; \\
\quad \quad \quad \quad \text{else } \\
\quad \quad \quad \quad \quad \quad \text{for each } (s, h_*, q, t) \in T \\
\quad \quad \quad \quad \quad \quad T' := T' \cup \{((s, Tr), \tau, q, (t, Tr.h_*))\}; \\
\quad \quad \quad \quad \text{for each } ((s, Tr), \_, \_, (t, Tr.\pi)) \in T' \\
\quad \quad \quad \quad \quad \text{if } (t, Tr.\pi) \notin S' \text{ then } H_{Adv}((t, Tr.\pi));
\]
Comparing different system views: $\approx_{PB}$

Definition. An equivalence relation $R \subseteq S \times S$ is a weak probabilistic bisimulation if and only if, whenever $(s, s') \in R$, then for all $C$ in the quotient set $S/R$:

1. $\text{Prob}(s, \tau^*a, C) = \text{Prob}(s', \tau^*a, C)$ $\forall a \in \text{Act}$
2. $\text{Prob}(s, a_*, C) = \text{Prob}(s', a_*, C)$ $\forall a_* \in \text{Act}$

Two states $s, s' \in S$ are weakly probabilistically bisimilar, denoted $s \approx_{PB} s'$, if there exists a weak probabilistic bisimulation $R$ including the pair $(s, s')$. 
Revealing the interference

$S \setminus A = (S, Act, T, s_0)$ and $S|A = (S', Act, T', s_0^A)$

Exact analysis: $S \setminus A \approx_{PB} S|A$
Approximating \( \approx_{PB} \)

A relation \( R \subseteq S \times S \) is a weak probabilistic bisimulation with \( \varepsilon \)-precision, where \( \varepsilon \in (0, 1) \), if and only if, whenever \( (s, s') \in R \), then for all \( C \in S/R \)

- \( | Prob(s, \tau^*a, C) - Prob(s', \tau^*a, C) | \leq \varepsilon \ \forall a \in Act \)
- \( | Prob(s, a_*^*, C) - Prob(s', a_*^*, C) | \leq \varepsilon \ \forall a_*^* \in Act \)

Two states \( s, s' \in S \) are weakly probabilistically bisimilar with \( \varepsilon \)-precision, denoted \( s \approx_{PB\varepsilon} s' \), if there exists a weak probabilistic bisimulation with \( \varepsilon \)-precision including the pair \( (s, s') \).
Measuring the interference (1)

$R$ approximates $\approx_{PB}$ if $(s_0, s_0^A) \in R$ and there exists at least a pair $(s, s') \in R$ such that $\text{Prob}(s, \tau^*a, C) \neq \text{Prob}(s', \tau^*a, C)$ for some $a \in \text{Act}$ and class $C$ in $S \cup S'/R$.

- $d^R_A(s, s', a, C) = | \text{Prob}(s, \tau^*a, C) - \text{Prob}(s', \tau^*a, C) |$
- $\delta^R_A = \max \{ |d^R_A(s, s', a, C)| : s \in S \setminus A, s' \in S \setminus A, (s, s') \in R, a \in \text{Act}, C \in S \cup S'/R \}$
- Let $\varepsilon_A = \inf_R \delta^R_A$. Then:
  
  1. $S \setminus A \approx_{PB\varepsilon_A} S \setminus A$
  2. There does not exist $\varepsilon < \varepsilon_A$ such that $S \setminus A \approx_{PB\varepsilon} S \setminus A$
Measuring the interference (2)

In order to discount the future, we consider the probability of being in $s$ and $s'$ when computing their distance:

$$
\bar{d}_A^R(s, s', a, C) = \text{Prob}(s_0, s) \cdot \text{Prob}(s_A^0, s') \cdot \left| \text{Prob}(s, \tau^* a, C) - \text{Prob}(s', \tau^* a, C) \right|
$$

$$
\delta_A^R = \max \{ | \bar{d}_A^R(s, s', a, C) | \mid s \in S \setminus A, s' \in S \mid A, (s, s') \in R, a \in \text{Act}, C \in S \cup S'/R \} \}
$$

**Reward-based DTMC analysis approach**

$$
| \pi(s) \cdot p - \pi(s') \cdot p' |
$$

where:

- $\pi(s)$: probability of being in $s$
- $p$: prob. of observing the event of interest at $s$

**Approx. bisimulation approach**

$$
\pi(s) \cdot \pi(s') \cdot | p - p' |
$$
Measuring the maximum interference

- $\mathcal{R}_S$: family of considered relations.
- $\mathcal{A}_C$: set of adversaries related to family $C \in \{S, I, HD\}$.
- $\varepsilon^C_R = \sup_{A \in \mathcal{A}_C} \delta^R_A$: maximum interference caused by the most powerful adversary $A \in \mathcal{A}_C$ whenever the considered relation is $R$.
- $\bar{\varepsilon} = \inf_{R \in \mathcal{R}_S} \varepsilon^C_R$: maximum interference caused by the most powerful adversary in $\mathcal{A}_C$ that maximizes the probability of revealing its presence to Low for the closest approximation $R \in \mathcal{R}_S$ of $\approx_{PB}$.
- $\bar{\varepsilon}$ can be interpreted as the distance induced by a specific operator norm in a linear operator framework.
Finding the mpa

Most powerful simple adversary:
(a) \( p_h = p_k = \frac{1}{2} \).
(b) \( p_h = \frac{2}{3} \).

Most powerful interactive adversary:
(a) \( \{h\} \rightarrow \emptyset, \{k\} \rightarrow \{(k, 1)\} \),
\( \{h, k\} \rightarrow \{(h, 1)\} \).
(b) \( p_h = \frac{2}{3} \).
Finding the mpa

Most powerful history-dependent adversary:
(a) $A_r(\epsilon) = \emptyset$, $A_r(l) = \{(k, 1)\}$, $A_r(l.k_*) = \{(h, 1)\}$.
(b) $A_r(\epsilon) = \emptyset$, $A_r(l) = \{(h, 1)\}$, $A_r(l.h_*) = \{(h, 1)\}$. 
Finding the mpa

Most powerful history-dependent adversary:

\[ A_g(\epsilon) = \{h\}, \ A_r(\epsilon) = \emptyset, \text{ and} \]
\[ A_g(h) = \emptyset, \ A_r(h) = \{(k, 1)\}. \]
Complexity of finding the mpa

Theorem

Let $S = (S, Act, T, s_0)$ be a GRTS such that $(-, \tau, -, -) \notin T$ and let $A \in A_{HD}$ be the m.p.a. for $S$ such that $S$ is safe w.r.t. $A$ and a security parameter $\nu$.

Then, $A$ is defined by a pair $(A_g, A_r)$ such that for each execution trace $Tr$ either $A_r(Tr) = \emptyset$, or $A_g(Tr) = \emptyset$ and $\forall h \in H$ such that $(h, p_h) \in A_r(Tr)$ it holds that $p_h$ is the limiting value 1.
Noninterference properties

- GRTS-based Probabilistic Process Algebra
- Probabilistic Noninterference Properties:

<table>
<thead>
<tr>
<th>BSPNI :</th>
<th>System model with restriction on the high-level actions</th>
<th>System model with hiding on the high-level actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBNDC :</td>
<td>System model with restriction on the high-level actions</td>
<td>System model in parallel with any high-level process</td>
</tr>
</tbody>
</table>
Adversaries vs. noninterference properties

\textit{BSPNI} vs. simple adversaries:
\begin{quote}
Result. $\mathcal{A}_{\text{BSPNI}} \subset \mathcal{A}_S$.
\end{quote}

\textit{PBNDC} vs. interactive adversaries:
\begin{quote}
Result. Same difficulties in evaluating the mpa.
\end{quote}

??? vs. history-dependent adversaries:
\begin{quote}
\textit{nondeducibility on strategies for nondeterministic state machines}
\end{quote}
History-dependent adversary via a strategy operator

\[(P, \gamma) \mid A\]

\[\uparrow \uparrow \uparrow \quad \text{process history strategy}\]

\[P \xrightarrow{a,p} P'\]

\[a \in \text{Act}_L \cup \{\tau\}\]

\[P \xrightarrow{\tau,p \cdot \nu(P,\gamma,A)} (P', \gamma a)\mid A\]

\[a \in \text{Act}_L\]

\[P \xrightarrow{h,p} P'\]

\[h \in A_g(\gamma)\]

\[P \xrightarrow{\tau,p \cdot \nu_h(P,\gamma,A)} (P', \gamma h)\mid A\]

\[h, p_h \in A_r(\gamma)\]

\[\nu(P, \gamma, A) = (1/g(P, A_g(\gamma))) \cdot r(P, A_r(\gamma))\]

\[g(P, A_g(\gamma)) = \sum \{ q \mid \exists Q \in \mathcal{G}, a \in \text{Act}_L \cup A_g(\gamma) \cup \{\tau\}, P \xrightarrow{a,q} Q \}\}\]

\[r(P, A_r(\gamma)) = \prod \{ (1 - p_i) \mid (h_i, p_i) \in A_r(\gamma) \land P \xrightarrow{h_i,*} \}\text{ such that}\]

\[r(P, A_r(\gamma)) = 1 \text{ if the multiset is empty}\]

\[\nu_h(P, \gamma, A) = q \cdot \prod \{ (1 - p_i) \mid (h_i, p_i) \in A_r(\gamma) \land h < h_i \land P \xrightarrow{h_i,*}\}\text{ such that}\]

\[\nu_h(P, \gamma, A) = q \text{ if the multiset is empty}\]

\[q = \begin{cases} p_h & \text{if } g(P, A_g(\gamma)) > 0 \vee \exists (h_i, p_i) \in A_r(\gamma). h_i < h \land P \xrightarrow{h_i,*} \\ 1 & \text{otherwise} \end{cases}\]
Probabilistic nondeducibility on strategies

Definition. $P \in PNDS$ if and only if

$$P \backslash H \approx_{PB} (P, \varepsilon) \mid A \quad \forall A \in A_{HD}.$$  

$P$ is $PNDS$-secure if its execution is invariant, from the viewpoint of Low, with respect to every possible strategy conducted by a polynomial-time history-dependent adversary.

Result. $A_{PNDS} = A_{HD}$.

Result. $PNDS \subset BSPNI$. 

An example

1. $A_r(\varepsilon) = \{(\text{receive request}, 1)\}$
2. $A_g(\text{receive request}) = (\text{send msg})$
3. $A_r(\text{receive request} \ast \text{send msg}) = \{(\text{receive ack}, 1)\}$
4. $A_g(\text{receive request} \ast \text{send msg} \ast \text{receive ack} \ast) = (\text{send msg})$
5. $A_r(\text{receive request} \ast \text{send msg} \ast \text{receive ack} \ast \text{send msg}) = \{(\text{receive stop}, 1)\}$
Comparisons [Desharnais et al.]

- **Model** labelled Markov processes
- **Semantics** logical characterization of bisimulation:
  \[ \mathcal{L} := \top \mid \phi_1 \land \phi_2 \mid \langle a \rangle_q \phi \]
- From the logic-base characterization to the functional expressions based characterization:
  \[ f := 1 \mid 1 - f \mid \langle a \rangle f \mid \min(f_1, f_2) \mid \sup_{i \in \mathbb{N}} f_i \mid f \ominus q \]
- \( s \) and \( s' \) are bisimilar iff they satisfy the same logical formulas iff they have the same values for each functional expression
- pseudometric: \( d^c(P, Q) = \sup_{f \in \mathscr{F}} |f_P(p_0) - f_Q(q_0)| \)
Comparisons [Desharnais et al.]

\[ d^c(P, Q) = \sup_{f \in F^c} |f_P(p_0) - f_Q(q_0)| \]

What about:

- diagnostic information
- complexity
- nondeterminism
Conclusions

- Characterization of the adversary for probabilistic non-interference properties.
- Estimation of the mpa: affordable for PNDS. What about complexity?
- To investigate: integration with the computational model of cryptography.
- Dealing with nondeterminism and time.