Architecture-based Uncertainty Impact Analysis to ensure Confidentiality
18th Symposium on Software Engineering for Adaptive and Self-Managing Systems, SEAMS’23

Sebastian Hahner, Robert Heinrich, Ralf Reussner
Motivation

Uncertainty has an impact on a software system’s confidentiality. Uncertainty sources exist in the system and its environment [1]. Design time analysis can find confidentiality violations [2, 3].

Challenges

- Uncertainty source and impact location in the system can differ
- Lack of comprehensive and precise modeling and analysis

### Modeling the Impact of Uncertainty on Confidentiality

- We distinguish between uncertainty sources and impact locations

<table>
<thead>
<tr>
<th>Uncertainty Source</th>
<th>Architecture Element</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Uncertainty Impact</th>
<th>Data Flow Element</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>1..*</td>
</tr>
</tbody>
</table>

- There are five uncertainty types with potential impact on confidentiality [4]
  - We extend data flow diagrams [5] to represent the impact of uncertainty

---


Uncertainty Impact Analysis on Confidentiality

Impact Analysis Algorithm
1) Annotate the uncertainty source
2) Calculate structural propagation based on change impact analysis \[6\]
3) Map all impacts to the data flow diagram \[2\]
4) Calculate the propagation along all affected data flows
5) Calculate the impact set by finding the maximum discontiguous data flows

\[
\text{max}_D \left( \{\{2, 3, DB, 4, 5\}, \{3, DB, 4, 5\} \} \right)
\]

Uncertainty Impact Analysis on Confidentiality

Formal Foundation of Impact Analysis

- Data flow diagrams can be represented as $DAG \ G = (V, E)$ with a strict partial order $u < v$
- We reuse the mapping $m(a)$ from the architecture $A$ to data flow nodes
- The impact analysis of an uncertainty source $S$ is a function $u : S \rightarrow X \subseteq V$
- The impact set is represented by an induced subgraph $G[X]$
- Uncertainty impacts follow the data flow: $\forall x \in X \subseteq V, \exists a \in A : m(a) \preceq x$

Online Shop

U1: Interface Uncertainty

m

Logistics

DAG G

1 2 3 4 5 6 7 8 9

Online Shop

Database

DB

Uncertainty

16 May 2023

S. Hahner, R. Heinrich, R. Reussner – Architecture-based Uncertainty Impact Analysis to ensure Confidentiality

KASTEL – Institute of Information Security and Dependability
DSiS – Dependability of Software-intensive Systems group
Case Study-based Evaluation

Goal Question Metric Plan

- **Accuracy**: How precise and complete are the calculated impact sets?
- **Effort reduction**: How many model elements must be considered in the analysis?

Case Study

- Corona Warn App, 19 components, 200 data flow diagram nodes
- 4 evaluation scenarios, comparing to confidentiality analysis [2]

Results

- High $F_1$ score of 0.94, analysis optimized for recall $R$ of 1.0 without false negatives
- Impact set ratio $r_i$ of 0.18 has slight overestimation of affected set ratio $r_a$ of 0.16

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision $P$</td>
<td>0.838</td>
<td>1.000</td>
<td>0.840</td>
<td>0.882</td>
<td>0.890</td>
</tr>
<tr>
<td>Recall $R$</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>$F_1$ score</td>
<td>0.912</td>
<td>1.000</td>
<td>0.913</td>
<td>0.938</td>
<td>0.942</td>
</tr>
<tr>
<td>Ratio $r_a$</td>
<td>0.155</td>
<td>0.080</td>
<td>0.105</td>
<td>0.300</td>
<td>0.160</td>
</tr>
<tr>
<td>Ratio $r_i$</td>
<td>0.185</td>
<td>0.080</td>
<td>0.125</td>
<td>0.340</td>
<td>0.183</td>
</tr>
</tbody>
</table>

Precision vs. Recall in Security Analysis

High precision is good… but not without high recall.
Related Work

Three Categories of Related Work:

- **Architecture-based uncertainty analyses** use fuzzy values [7, 8] or design space exploration [9] but do not focus on confidentiality and lack precision [3].

- **Uncertainty-aware confidentiality analysis** also use data flow-based methods [10, 11] but require expert knowledge and lack comprehensiveness.

- **Uncertainty propagation for self-adaption** acknowledges the analysis challenge [12], especially related to the uncertainty interaction problem [13].

---

Conclusion and Future Work

- **Problem:** Predicting the impact of uncertainty on confidentiality
- **Contribution:** Modeling and analysis of the uncertainty impact by combining architecture-based and data flow-based propagation
- **Benefit:** More precise and less labor-intensive prediction of the potential impact of uncertainty, both at design time and run time

**Future Work**

- Enhance the expressiveness of the uncertainty impact model, e.g., with variability modeling
- Combine uncertainty impact analysis with existing design time confidentiality analysis
- Extend evaluation with additional domains

https://abunai.dev
References