Incremental Fingerprinting in an Open World

Loes Kruger, Paul Kobialka, Andrea Pferscher,
Einar Broch Johnsen, Sebastian Junges, Jurriaan Rot

Radboud University Nijmegen
University of Oslo

7th of July, StANP 2026

 loes.kruger@ru.nl



UNIVERSITY OF OSLO

iCIS | Software Science
Radboud University



Bluetooth Devices

- Bluetooth devices are everywhere
- Often a security-critical role

iCIS | Software Science
Radboud University

Bluetooth Devices

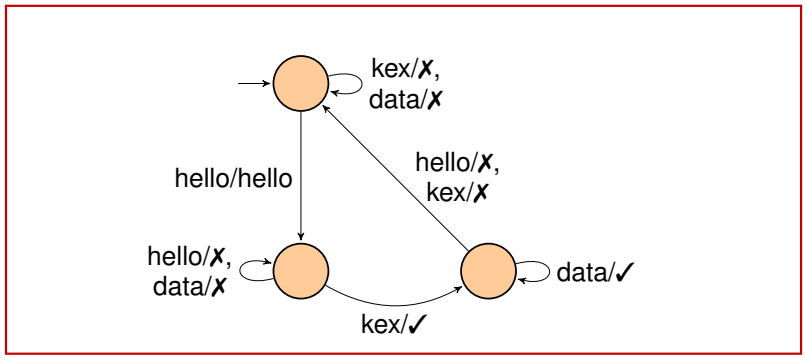
- Bluetooth devices are everywhere
- Often a security-critical role
- Recent vulnerabilities: Google Fast Pair, Airoha chips
- Vendors unaware of vulnerabilities
- How to determine whether your device is susceptible?

Fingerprinting!

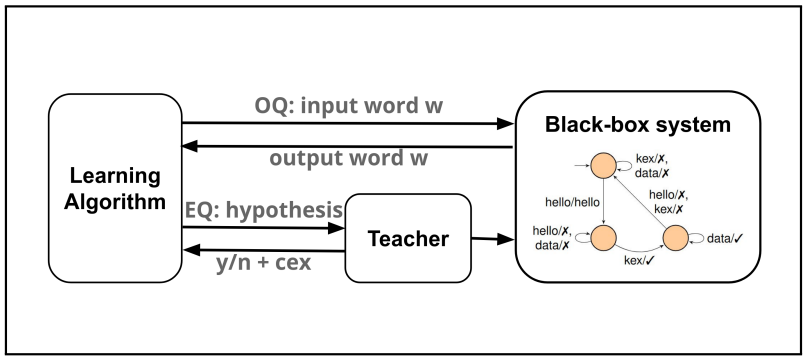
loes.kruger@ru.nl

iCIS | Software Science
Radboud University

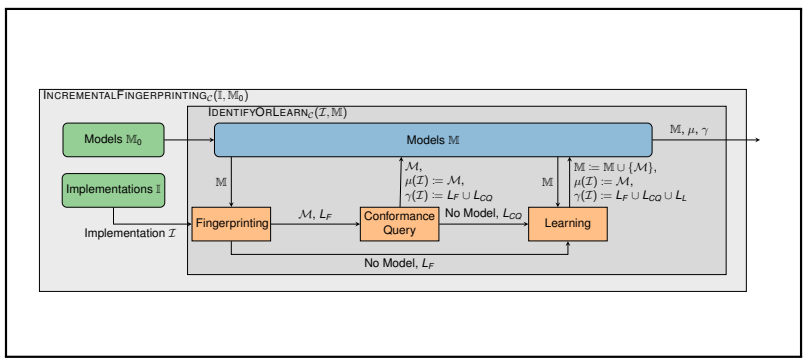
Overview



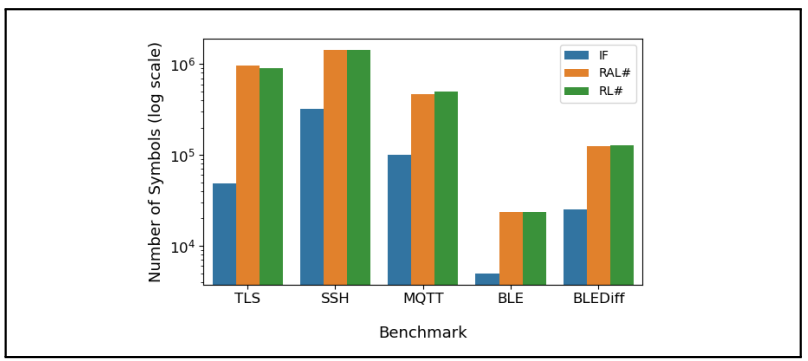
Closed-world Fingerprinting



Active Automata Learning



Incremental Fingerprinting



Experiments

loes.kruger@ru.nl

iCIS | Software Science
Radboud University

Fingerprinting

Network protocol fingerprinting: identify a **protocol implementation** by analyzing its **input-output behavior**

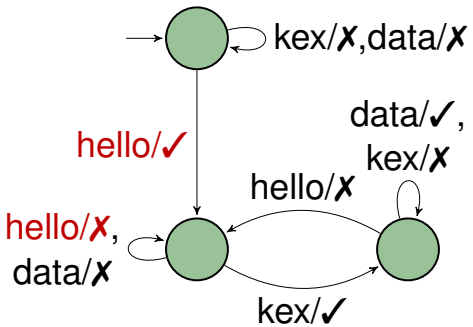
iCIS | Software Science
Radboud University



Fingerprinting

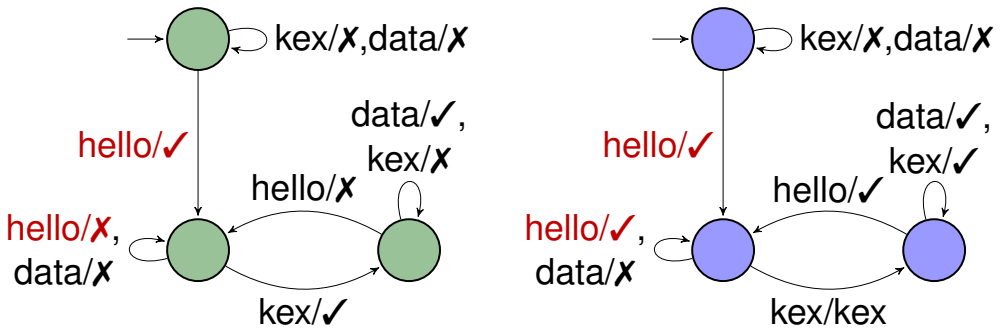
Network protocol fingerprinting: identify a **protocol implementation** by analyzing its **input-output behavior**

- The protocol implementation is a **black-box system**
- We want to identify several black-boxes (**group matching**)
- Protocols can be represented by **Finite State Machines** (FSMs)



iCIS | Software Science
Radboud University

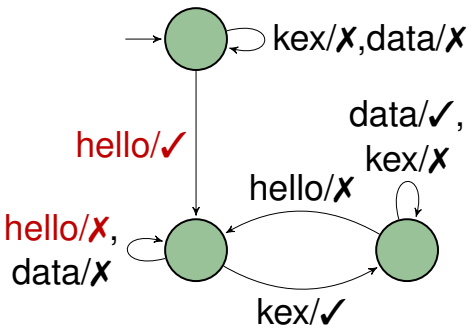
Fingerprinting in practice



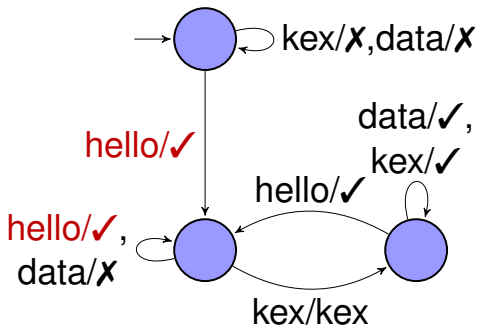
- Assume access to FSMs of protocols we want to identify
- Separating sequence: *hello* · *hello*

iCIS | Software Science
Radboud University

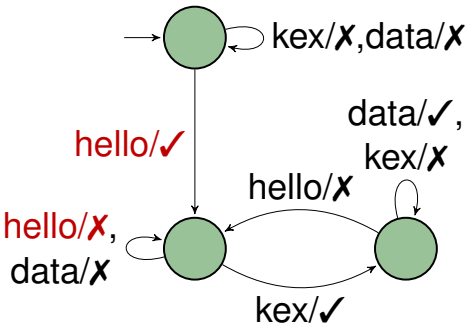
Fingerprinting in practice



Model 1



Model 2

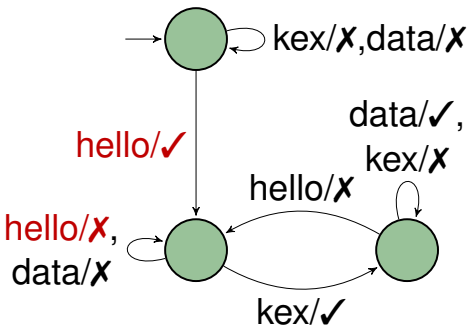


Black-box system

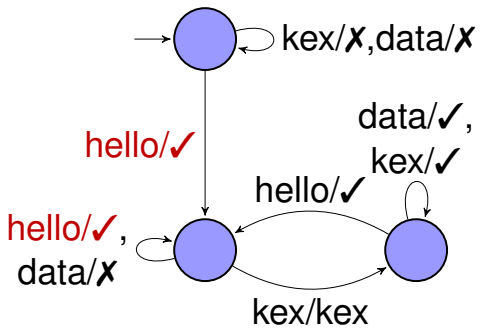
- Assume access to FSMs of protocols we want to identify
- Separating sequence: *hello* · *hello*

iCIS | Software Science
Radboud University

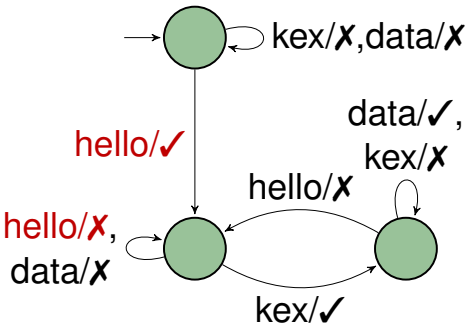
Fingerprinting in practice



Model 1



Model 2

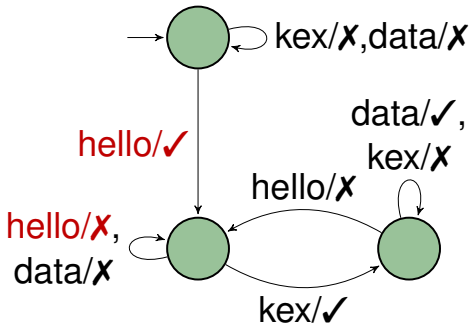


Black-box system

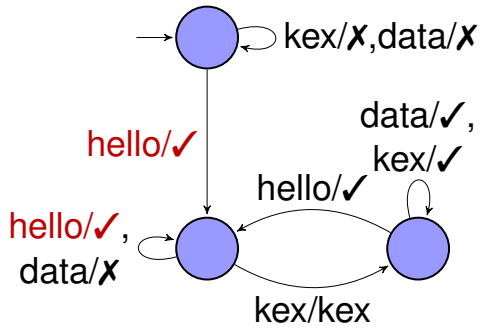
- Assume access to FSMs of protocols we want to identify
- Separating sequence: *hello* · *hello*
- Closed-world assumption: an exhaustive list of all possible protocol implementations exists

iCIS | Software Science
Radboud University

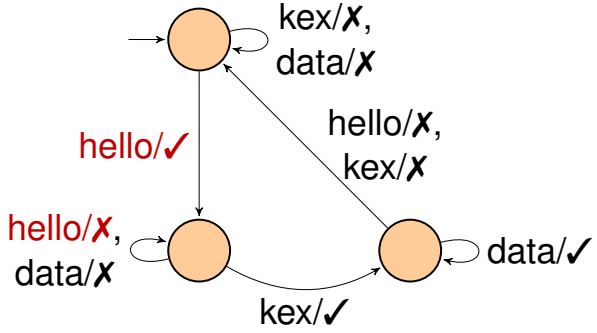
Fingerprinting in practice



Model 1



Model 2



Black-box system

- Assume access to FSMs of protocols we want to identify
- Separating sequence: *hello* · *hello*
- Closed-world assumption: an exhaustive list of all possible protocol implementations exists

iCIS | Software Science
Radboud University

Closed-world assumption does not hold

- An exhaustive list often does not exist
- Incorrectly assuming a closed-world assumption leads to
  1. Unnoticed misclassifications
  2. No matches
- If the model is missing, then 95% of the time it is misclassified

loes.kruger@ru.nl

iCIS | Software Science
Radboud University

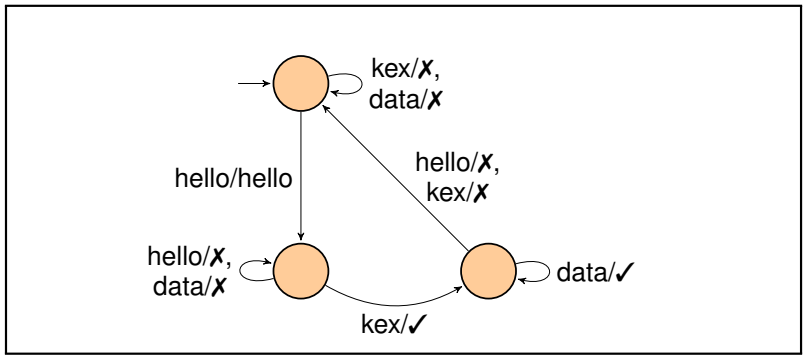
Open-world fingerprinting problem

1. Given a set of possibly incomplete models and
2. A list of protocol implementations,
3. Compute a list of models such that there exists an equivalent model for each protocol implementation.

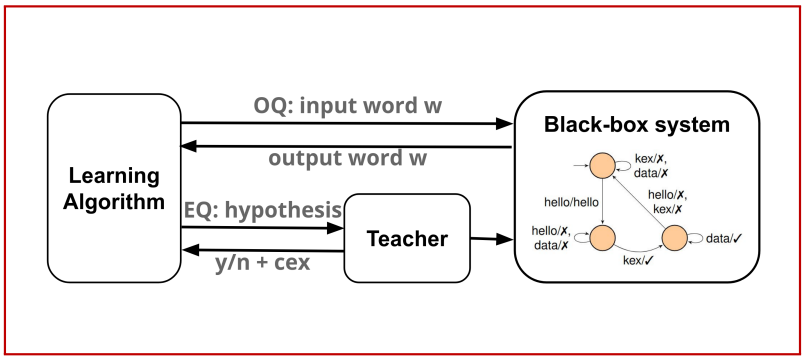
iCIS | Software Science
Radboud University



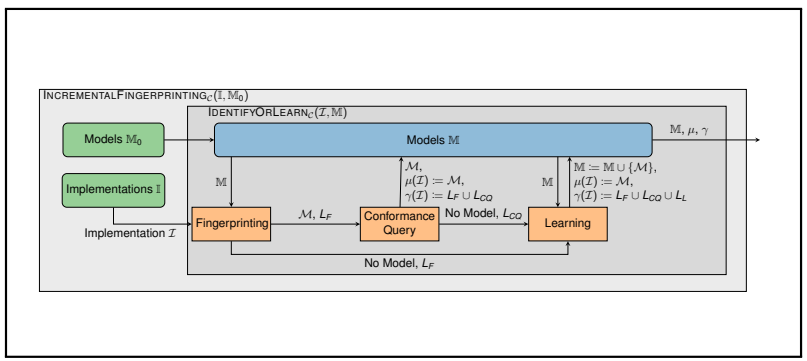
Overview



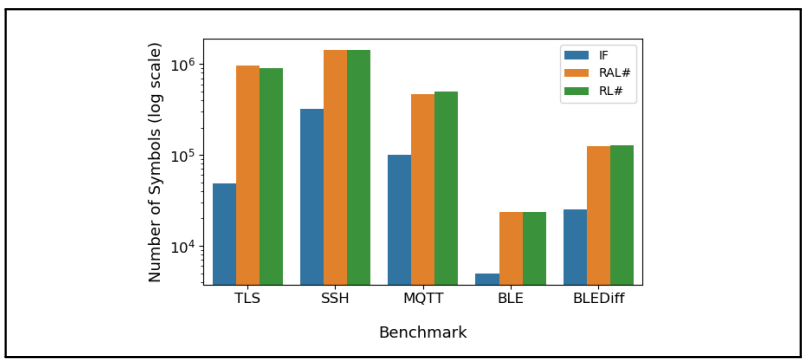
Closed-world Fingerprinting



Active Automata Learning



Incremental Fingerprinting



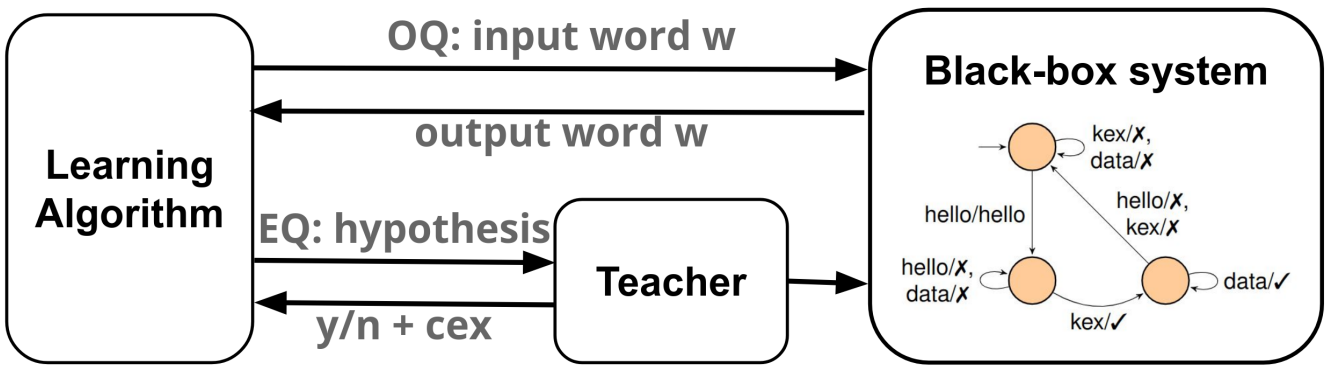
Experiments

iCIS | Software Science
Radboud University

Active Automata Learning

Active Automata Learning (AAL): learn an FSM of a black-box system

- AAL implements the MAT-framework by Angluin (1987)
- Output Query (OQ): what is the output of this input sequence?
- Equivalence Query (EQ): is this hypothesis equivalent to the black-box system?

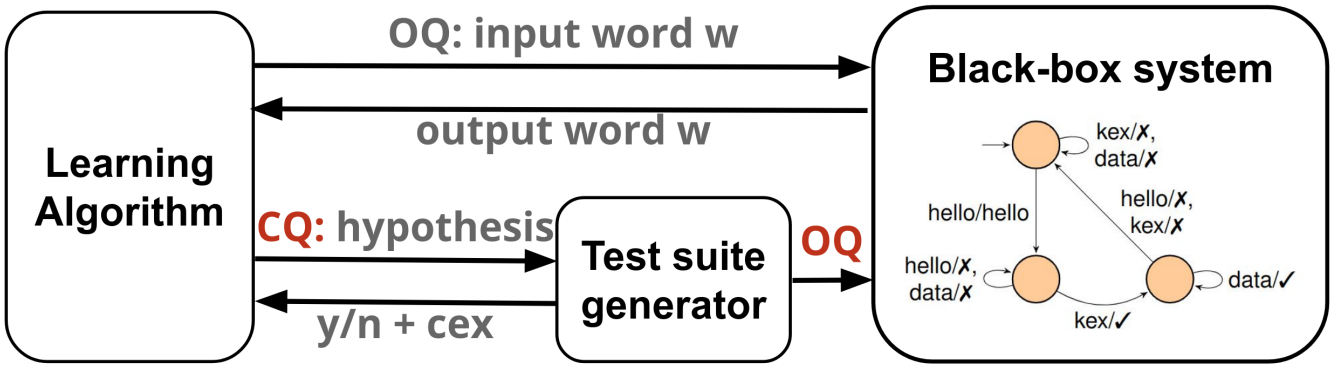


iCIS | Software Science
Radboud University

Active Automata Learning

Active Automata Learning (AAL): learn an FSM of a black-box system

- AAL implements the MAT-framework by Angluin (1987)
- Output Query (OQ): what is the output of this input sequence?
- Conformance Query (CQ): is this hypothesis equivalent to the black-box system?



iCIS | Software Science
Radboud University

Repeated AAL

- The goal is to learn several protocol implementations
- Apply AAL to each implementation separately
- Motivational experiment:
  - 596 TLS models, 22 unique models
  - Apply AAL to each of the 596 implementation separately

iCIS | Software Science
Radboud University

Repeated AAL

- The goal is to learn several protocol implementations
- Apply AAL to each implementation separately
- Motivational experiment:
  - 596 TLS models, 22 unique models
  - Apply AAL to each of the 596 implementation separately
    $\rightarrow$ 2.6 million total interactions, 75.9% incorrect models

iCIS | Software Science
Radboud University

Repeated AAL

- The goal is to learn several protocol implementations
- Apply AAL to each implementation separately
- Motivational experiment:
  - 596 TLS models, 22 unique models
  - Apply AAL to each of the 596 implementation separately
    $\rightarrow$ 2.6 million total interactions, 75.9% incorrect models
  - Improve conformance query
    $\rightarrow$ 14.3 million total interactions, 32.6% incorrect models

iCIS | Software Science
Radboud University

Pros and Cons

Repeated AAL

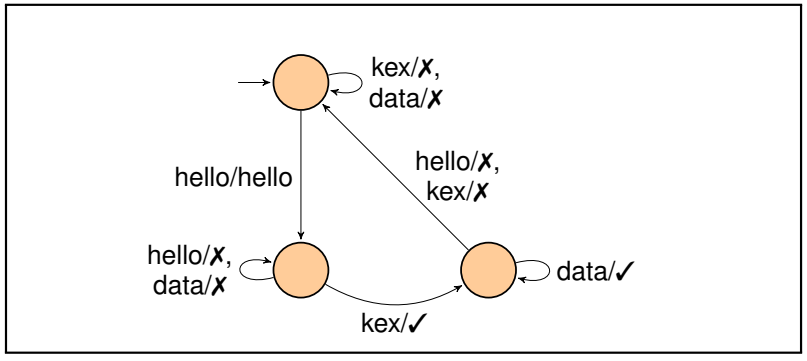
+ Few misclassifications given sufficient testing
- Difficult to judge when testing is sufficient
- Excessive number of interactions

Closed-world fingerprinting

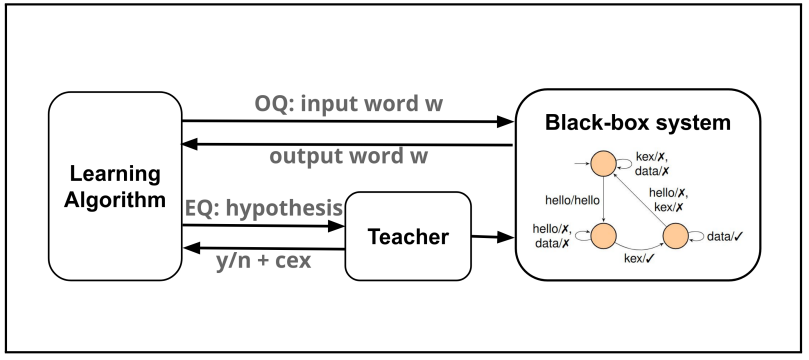
+ Requires only a few interactions
- Many unnoticed misclassifications

iCIS | Software Science
Radboud University

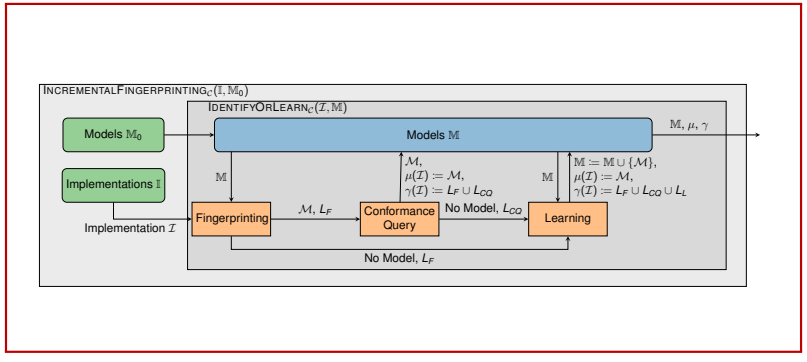
Overview



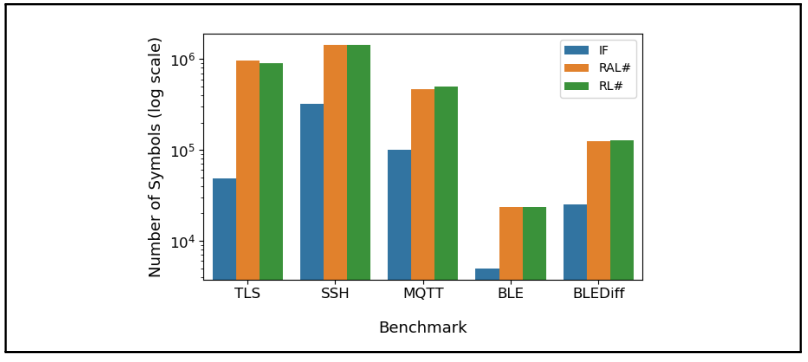
Closed-world Fingerprinting



Active Automata Learning



Incremental Fingerprinting



Experiments

iCIS | Software Science
Radboud University

Combining closed-world fingerprinting and AAL

Two key adaptations to repeated AAL:

1. Test whether the implementation matches a previous model to avoid unnecessary relearning
2. Use adaptive AAL to reuse previous models during learning

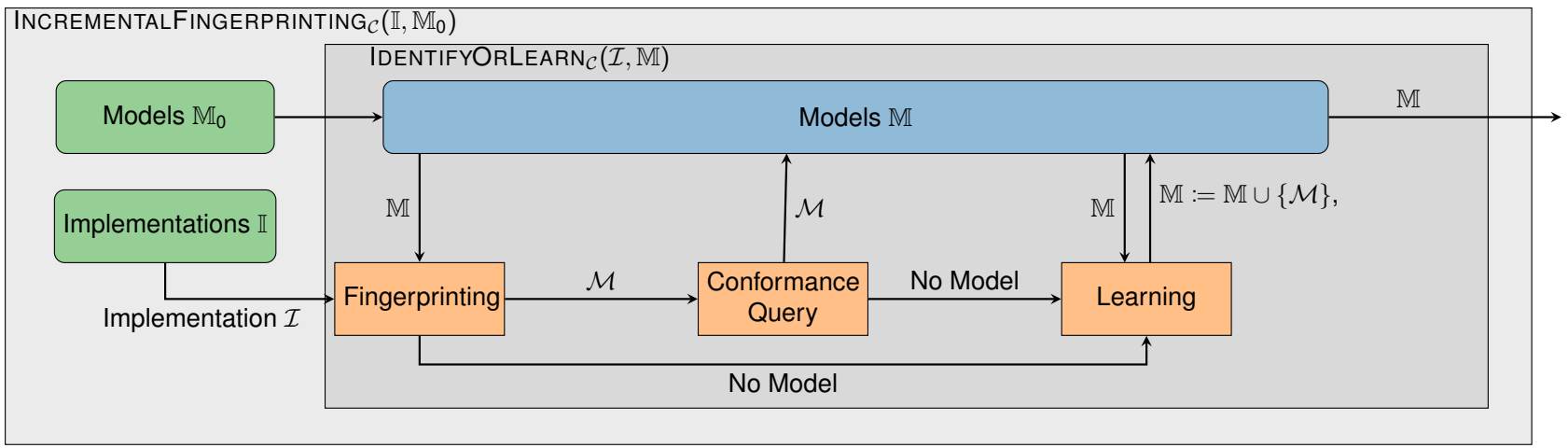
iCIS | Software Science
Radboud University



Combining closed-world fingerprinting and AAL

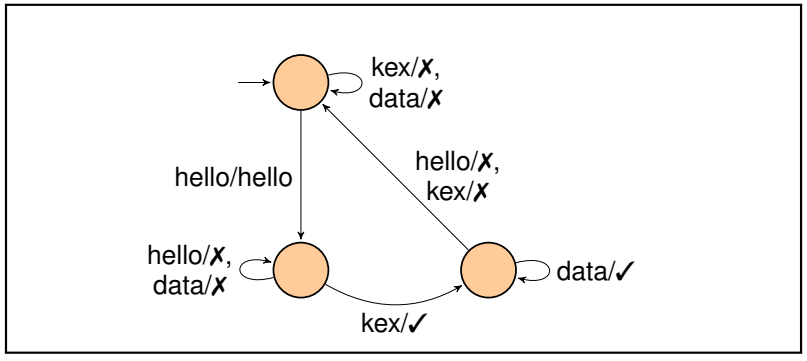
Two key adaptations to repeated AAL:

1. Test whether the implementation matches a previous model to avoid unnecessary relearning
2. Use adaptive AAL to reuse previous models during learning

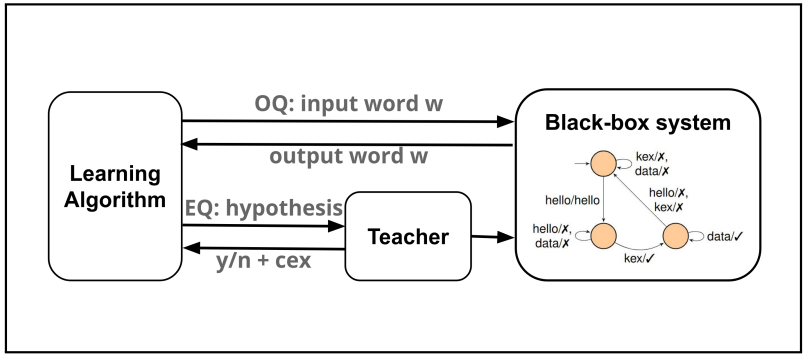


iCIS | Software Science
Radboud University

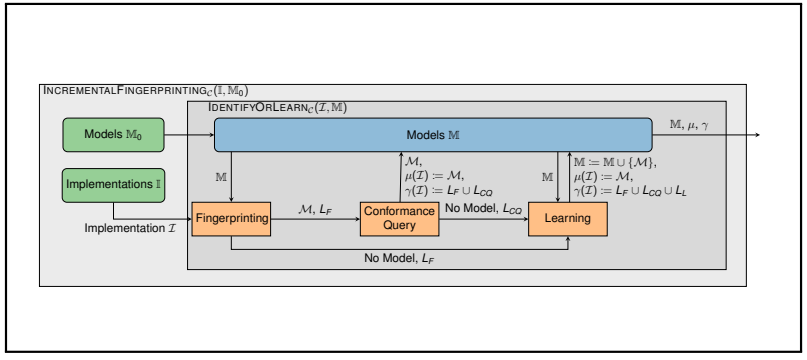
Overview



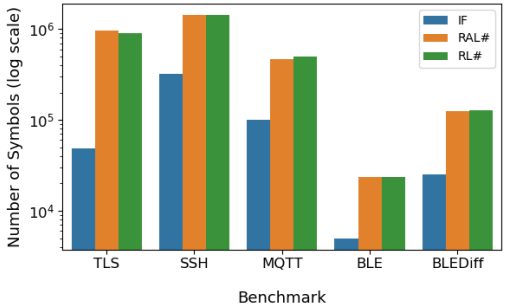
Closed-world Fingerprinting



Active Automata Learning



Incremental Fingerprinting



Experiments

iCIS | Software Science
Radboud University

Experimental Settings

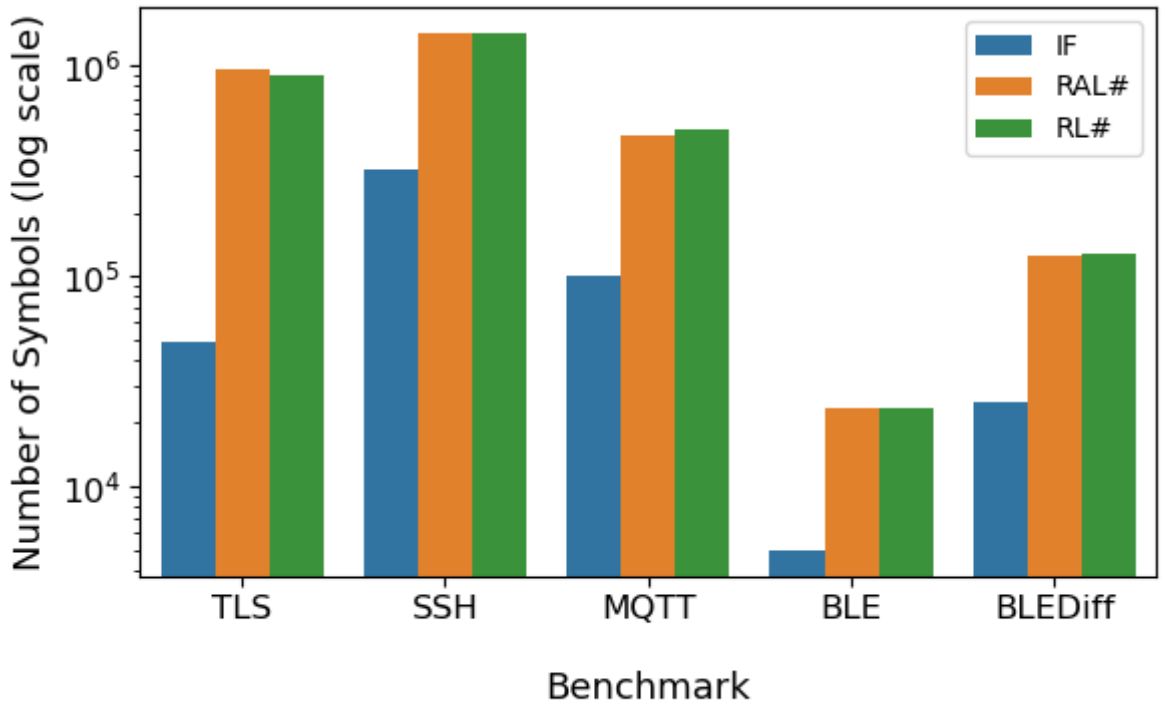
- INFERNAL: instantiation of incremental fingerprinting using state-of-the-art algorithms
- Baselines: repeated $L^{\#}$ and repeated adaptive $L^{\#}$
- Evaluation on various network protocol
- RQ: Better trade-off in misclassifications vs interactions?

iCIS | Software Science
Radboud University



Perfect Teacher

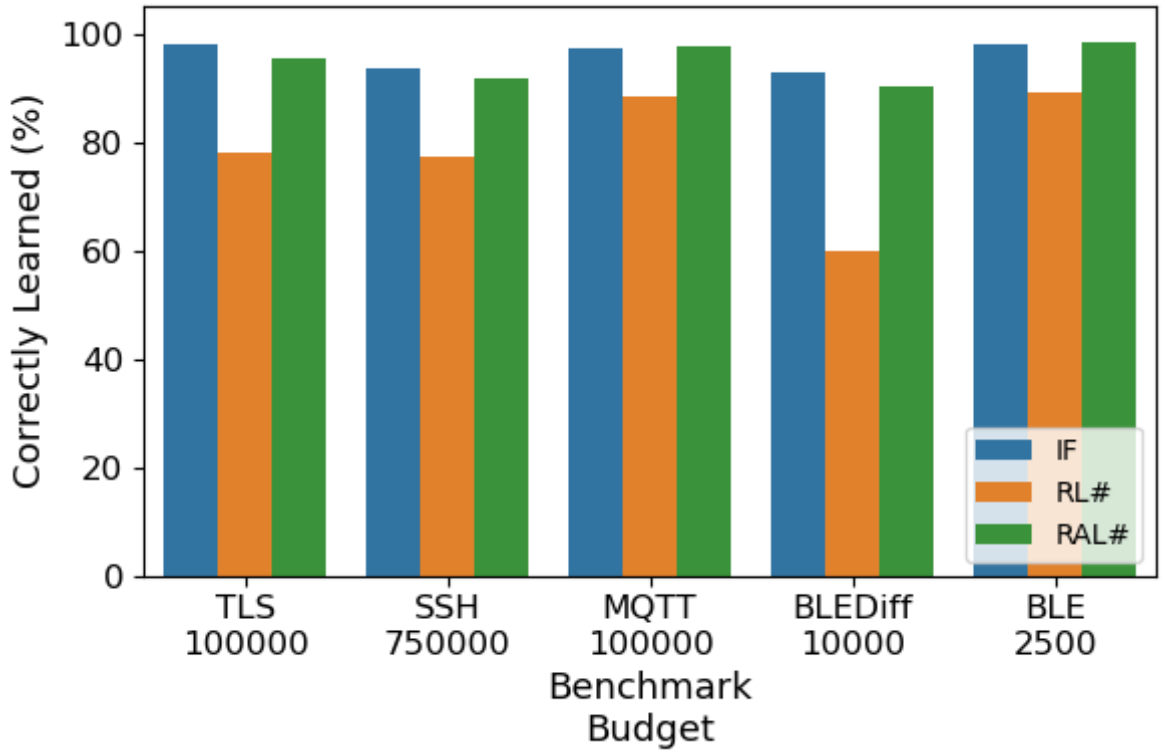
INFERNAL: fewer interactions



iCIS | Software Science
Radboud University

Interaction Budget

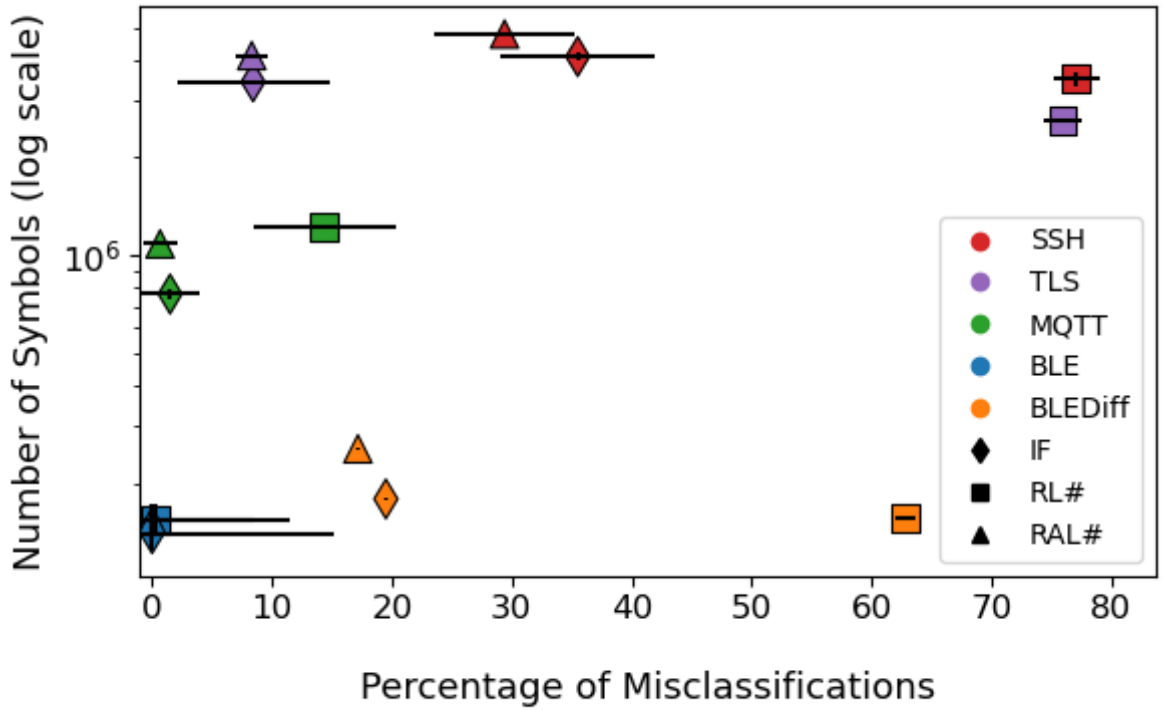
INFERNAL: fewer misclassifications within an interaction budget



iCIS | Software Science
Radboud University

Testing until all tests pass

INFERNAL: good trade-off misclassifications vs interactions



iCIS | Software Science
Radboud University

Applications of the framework

- Built set of initial models
- Fingerprinting new devices is relatively cheap

iCIS | Software Science
Radboud University



Applications of the framework

- Built set of initial models
- Fingerprinting new devices is relatively cheap
- Incorporate into IoT systems with automatic Bluetooth updates
  - Notify when new behaviour is different
  - Notify when new behaviour matches model with vulnerabilities
  - Add verification to the framework

loes.kruger@ru.nl

iCIS | Software Science
Radboud University

Contributions & Challenges

Contributions

- Formalized open-world fingerprinting problem
- INFERNAL with good trade-off misclassifications vs interactions
- Pipeline highly adjustable

Challenges

- Non-determinism due to environment
- Gray-box open-world fingerprinting
- Improve conformance query

CSF Paper

Code

loes.kruger@ru.nl

iCIS | Software Science
Radboud University