Master’s Thesis: Threat Analysis, Evaluation, and Mitigation for Smart Contracts Endorsed by TLS/SSL Certificates

Jan Felix Hoops, 12.04.2021, Final Presentation

Chair of Software Engineering for Business Information Systems (sebis)
Faculty of Informatics
Technische Universität München
wwwmatthes.in.tum.de

Outline

1. Background

2. Research Questions

3. Smart Contract Attacks

4. Typosquatting Detection

5. Evaluation

Background (1/2)

The Problem

Lack of Smart Contract Owner Authentication
There is no widely adopted, standardized way of authenticating the owner of an Ethereum smart contract. This is a security risk.

One important reason for this deficit is the **bootstrapping problem**.

TLS endorsed Smart Contracts (TeSC)
This proposal by Gallersdörfer envisions an authentication infrastructure leveraging SSL/TLS-certificates of the web.

Background (2/2)

TeSC

× N

Endorsement

- Part of every compliant smart contract
- Binds contract to domain

$$C = \{addr, cert_{domain}, exp, flags\}$$

$$S = \{sign(hash(C), cert_{privKey})\}$$

$$E = \{S, C, [cert_{fingerprint}]\}$$

× N

Verifier

- Off-chain software
- Verifies endorsements

Registry

- One smart contract
- Lists endorsed smart contracts by domain

Research Questions

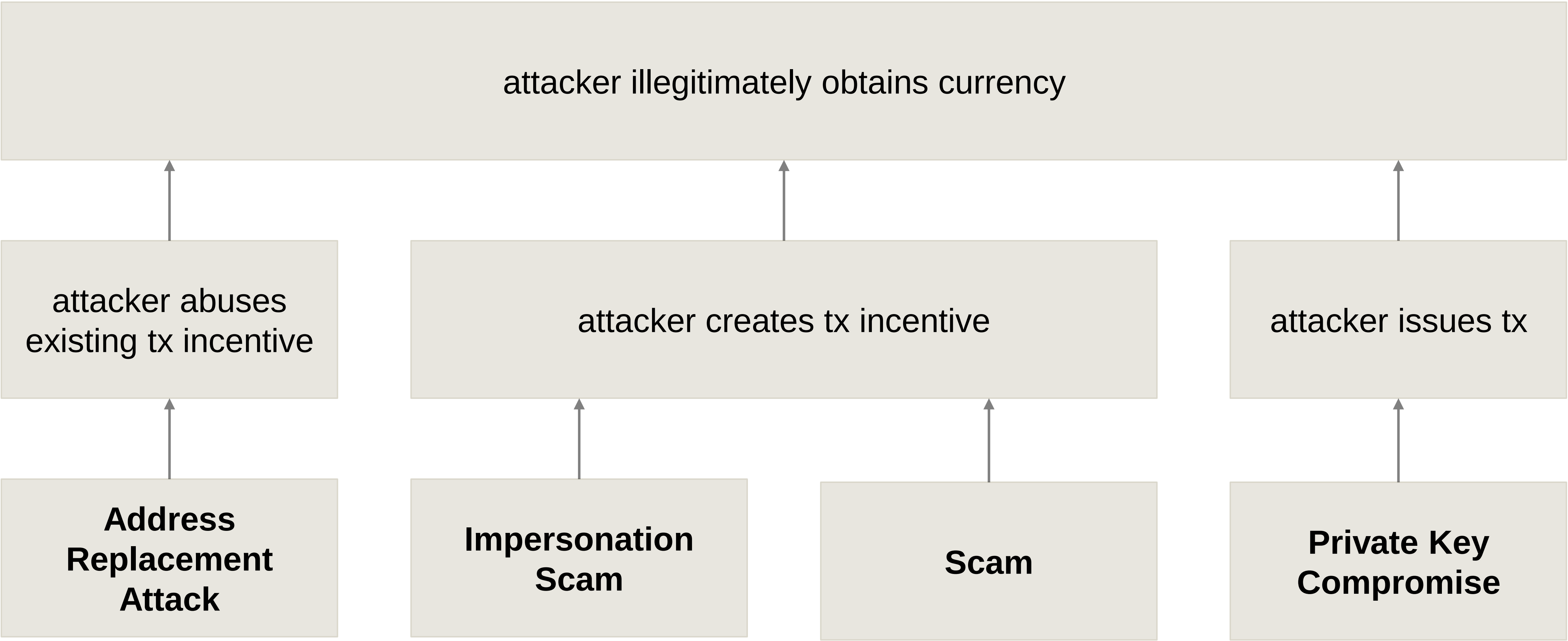
RQ1 What are actively used security mechanisms for the TLS/SSL certificate infrastructure on the web?

RQ2 What attacks could be performed against TeSC?

RQ3 How can TeSC be augmented to improve its security benefit?

Smart Contract Attacks (1/3)

Without TeSC



attacker illegitimately obtains currency

attacker abuses existing tx incentive

attacker creates tx incentive

attacker issues tx

Address Replacement Attack

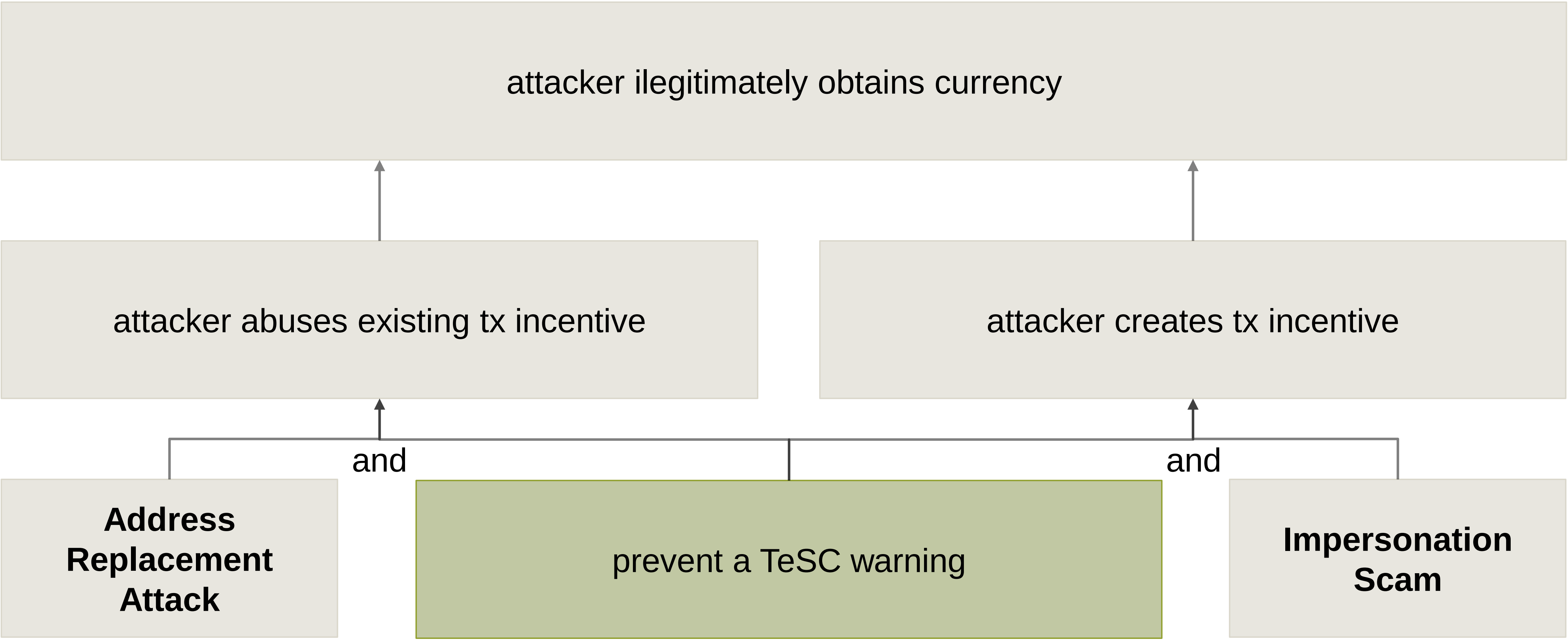
Impersonation Scam

Scam

Private Key Compromise

Smart Contract Attacks (2/3)

With TeSC



attacker ilegitimately obtains currency

attacker abuses existing tx incentive

attacker creates tx incentive

and

and

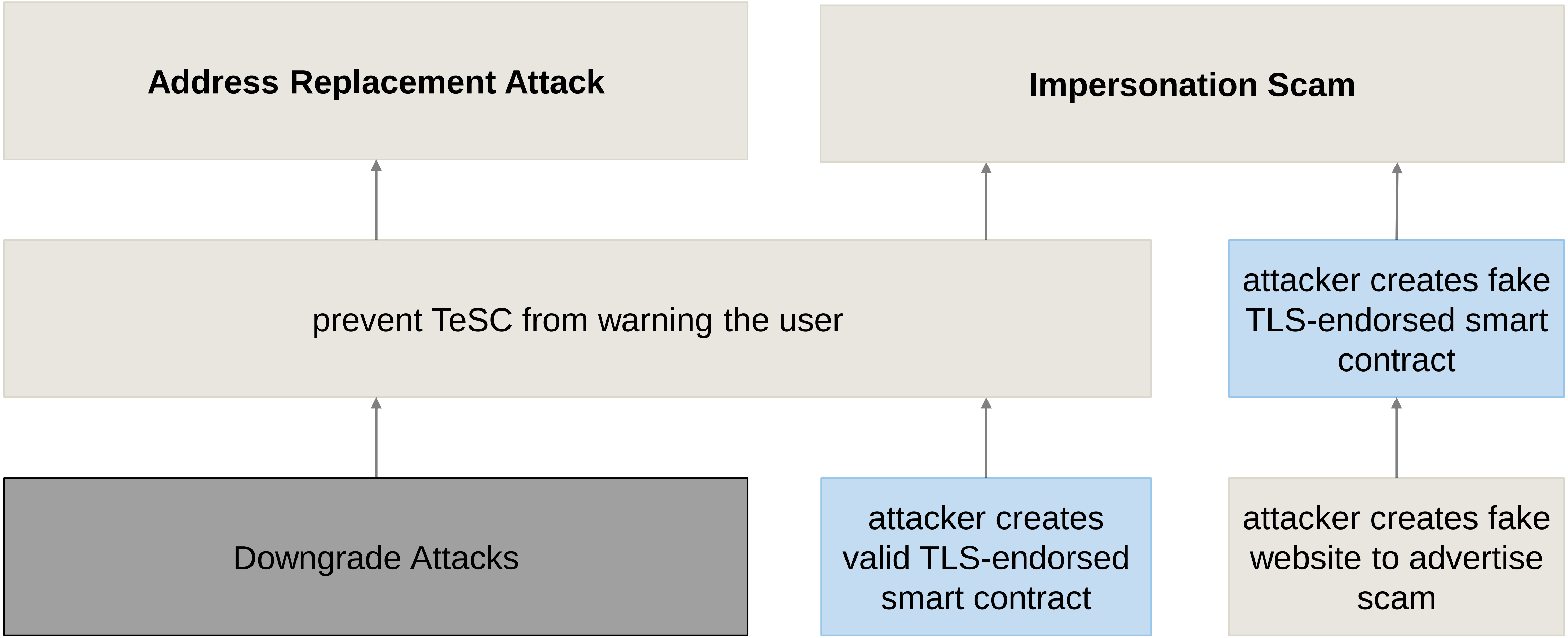
Address Replacement Attack

prevent a TeSC warning

Impersonation Scam

Smart Contract Attacks (3/3)

With TeSC



Address Replacement Attack

Impersonation Scam

prevent TeSC from warning the user

attacker creates fake TLS-endorsed smart contract

Downgrade Attacks

attacker creates valid TLS-endorsed smart contract

attacker creates fake website to advertise scam

Typosquatting Detection (1/7)

What is typosquatting?

Typosquatting is the practice of registering domains similar to well-established domains in bad faith.

e.g., turn.de

Simple Typo-Generation Models [Spaulding et al.]:

- Character-omission typo
- Character-permutation typo
- Character-duplication typo
- 1-mod-inplace

Further Typo-Generation Models:

- Homograph Attacks [Holgers et al.]
- Suffix Change

Typosquatting Detection (2/7)

Requirements

Language Agnostic
The algorithm must not depend on language.
Ensures universal applicability.

Device Agnostic
The algorithm must make no assumptions about user devices.
Ensures universal applicability.

Client Authority
All decisions must be made locally to ensure transparency of the decision process.

Typosquatting Detection (3/7)

Requirements

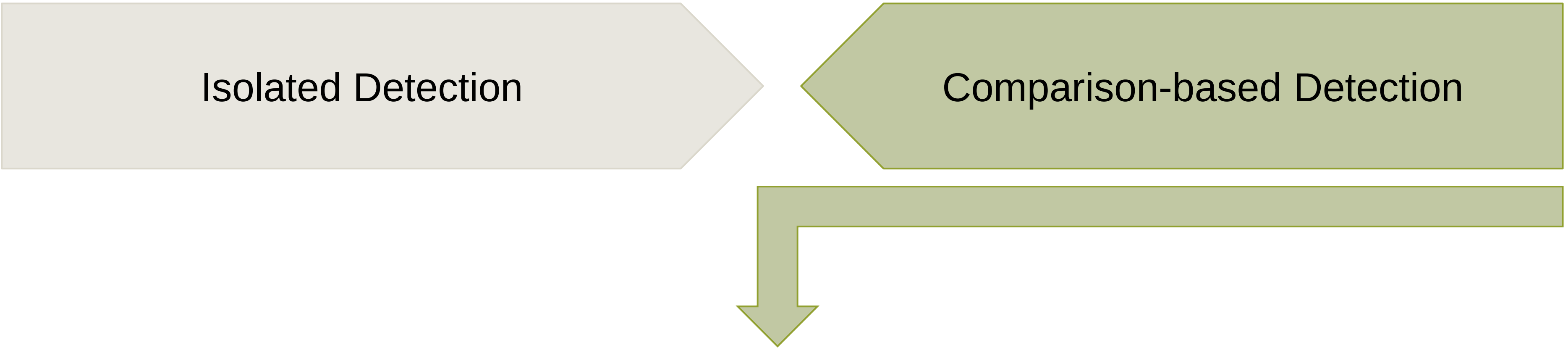
Independence
The algorithm must not depend on third parties or use unverifiable third-party metrics.

Minimal Knowledge Base
The static data required must be minimal to conserve memory on user devices and simplify maintenance.

Real-time Capable
The algorithm runs at least every time a user issues a transaction. Delays deter users.

Typosquatting Detection (4/7)

Design



Isolated Detection

Comparison-based Detection

1. **Candidate Detection**
   Identify pairs of suspicious domains.

2. **Candidate Evaluation**
   Consult additional information to possibly dismiss candidate.

Typosquatting Detection (5/7)

Candidate Detection

Can we rely on Damerau-Levenshtein distance (a.k.a. edit distance)?

tum.de vs. lmu.de

Can we rely on longest common substring?

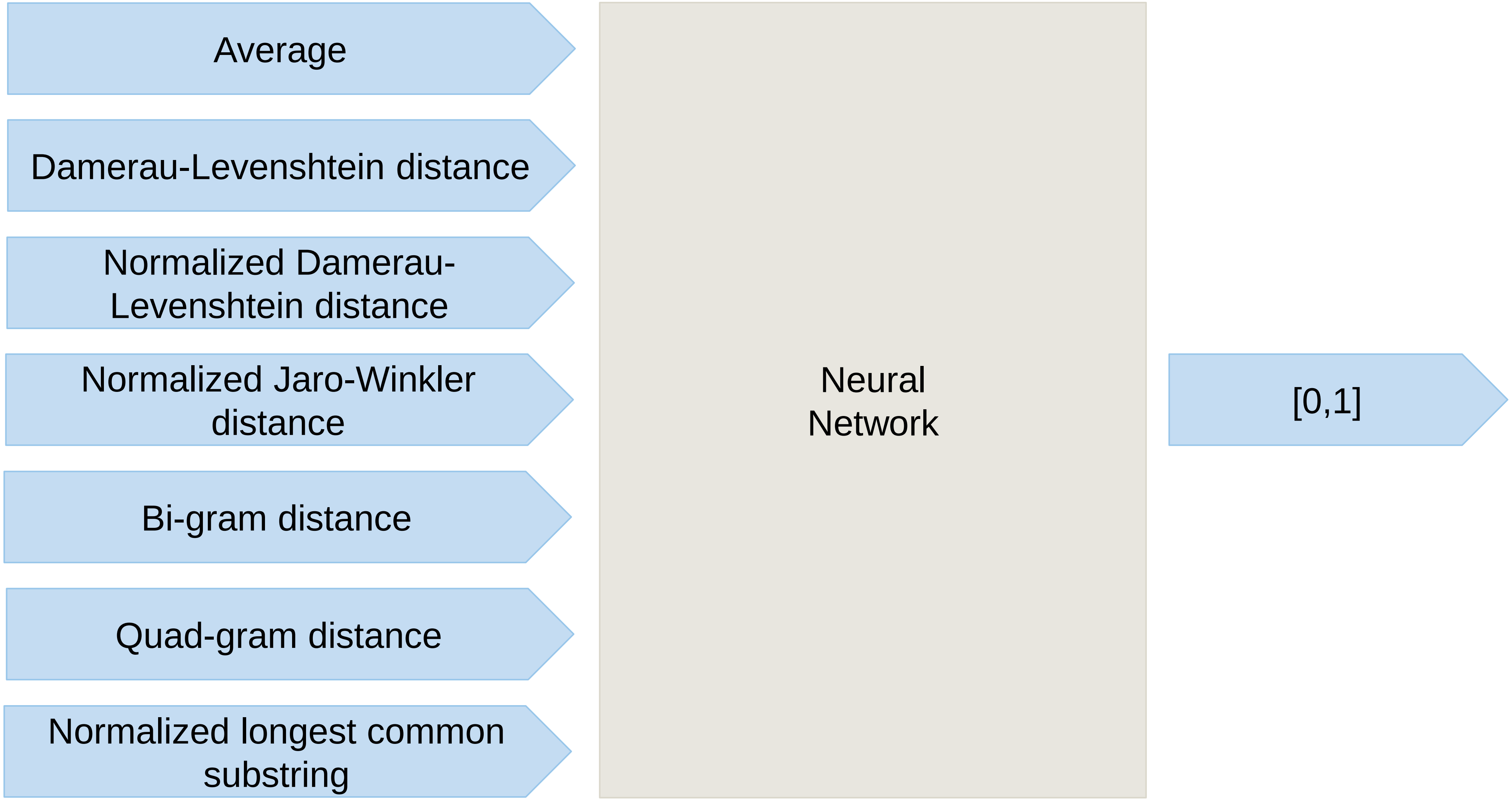
microsoft.com	vs.	microsoft-store.com
muenchen.de	vs.	landkreis-muenchen.de

Simplifications

- Focus on 2LDs
- No IDN Support

Typosquatting Detection (6/7)

Candidate Detection



Average

Damerau-Levenshtein distance

Normalized Damerau-Levenshtein distance

Normalized Jaro-Winkler distance

Bi-gram distance

Quad-gram distance

Normalized longest common substring

Neural Network

[0,1]

Typosquatting Detection (7/7)

Candidate Evaluation

A candidate pair can be dismissed if…

- Both domains are among known popular domains.
- Both domains resolve to the same IP address.
- Both domains’ certificates have significant overlap.

Tie-breaker: Original Registration Date

If in doubt, the older domain is most likely the original one.

Evaluation

Metric	Classifier Score
Accuracy	0.9953
Precision	1.0000
Recall	0.9906
F1 Score	**0.9952**
False Positive Rate	**0.00003750**

Correctly classified	Incorrectly Classified
rnicrosoft.com	wwwtum.de
wwwwikipedia.org	tum-donations.de
tum.de	?
feuerwehr-garching.de	

Thank you for your attention!
Any questions?

B. Sc.

Jan Felix Hoops
felix.hoops@tum.de

Technische Universität München
Faculty of Informatics
Chair of Software Engineering for Business Information Systems

Boltzmannstraße 3
85748 Garching bei München

Tel +49.89.289.
Fax +49.89.289.17136