Sharing FANCI Features: A Privacy Analysis of Feature Extraction for DGA Detection

Benedikt Holmes Arthur Drichel Ulrike Meyer

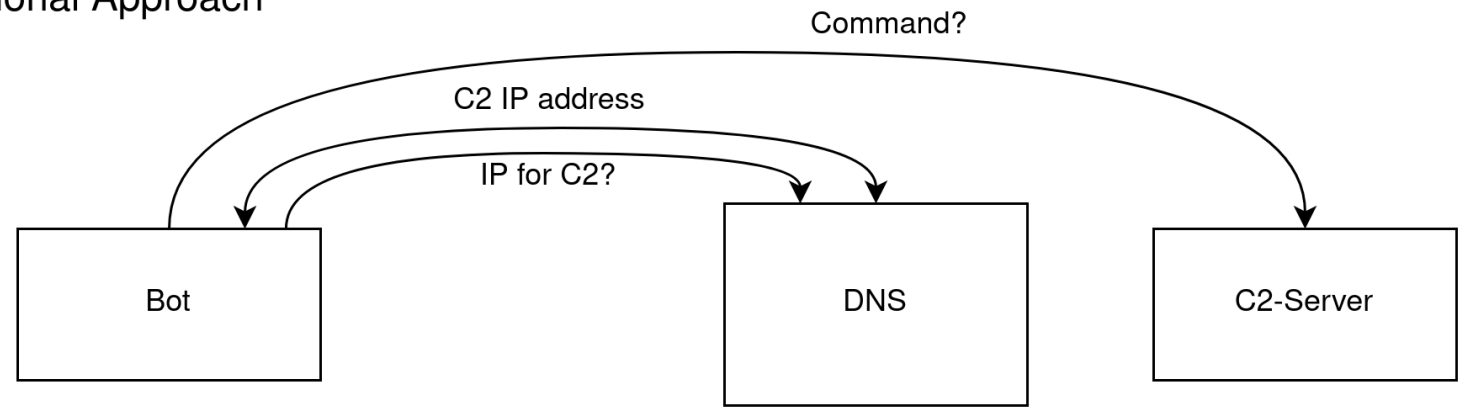
holmes@itsec.rwth-aachen.de

RWTH Aachen University

IT SEC Research Group IT-Security | RWTH AACHEN UNIVERSITY

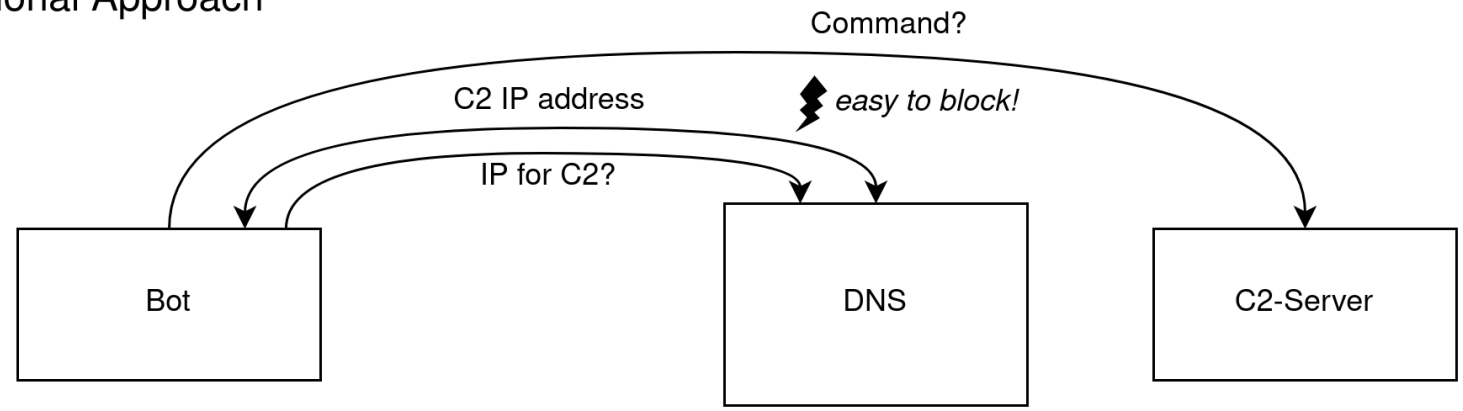
Context - Bot Communication

Traditional Approach

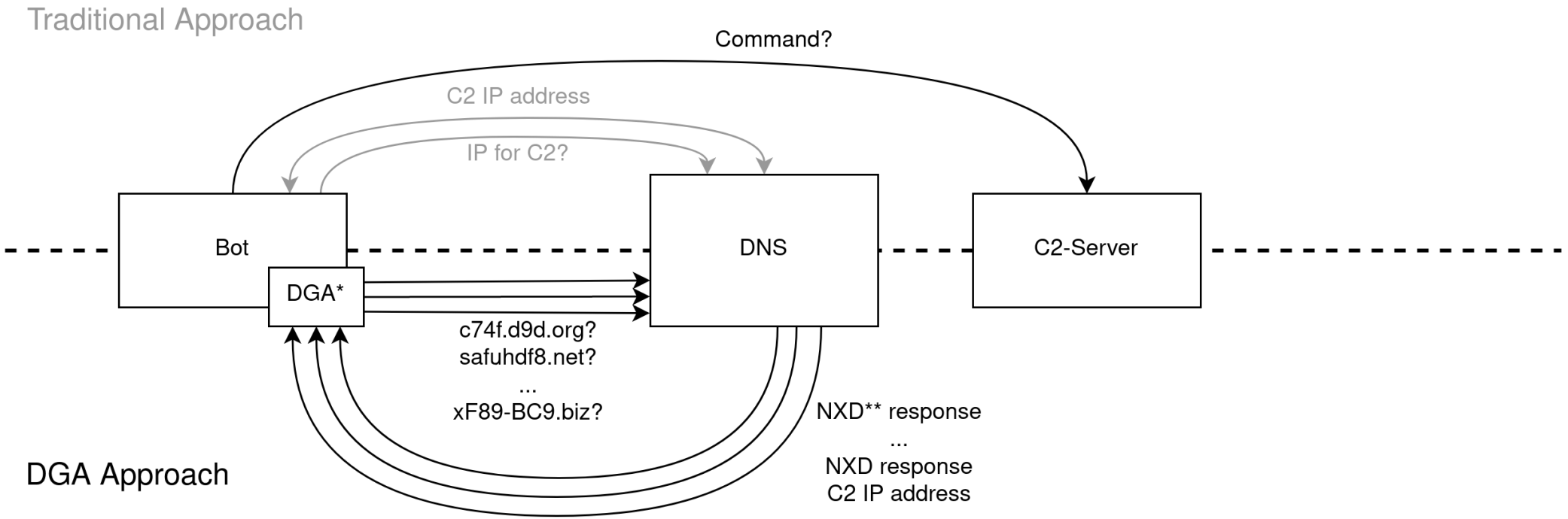


Context - Bot Communication

Traditional Approach

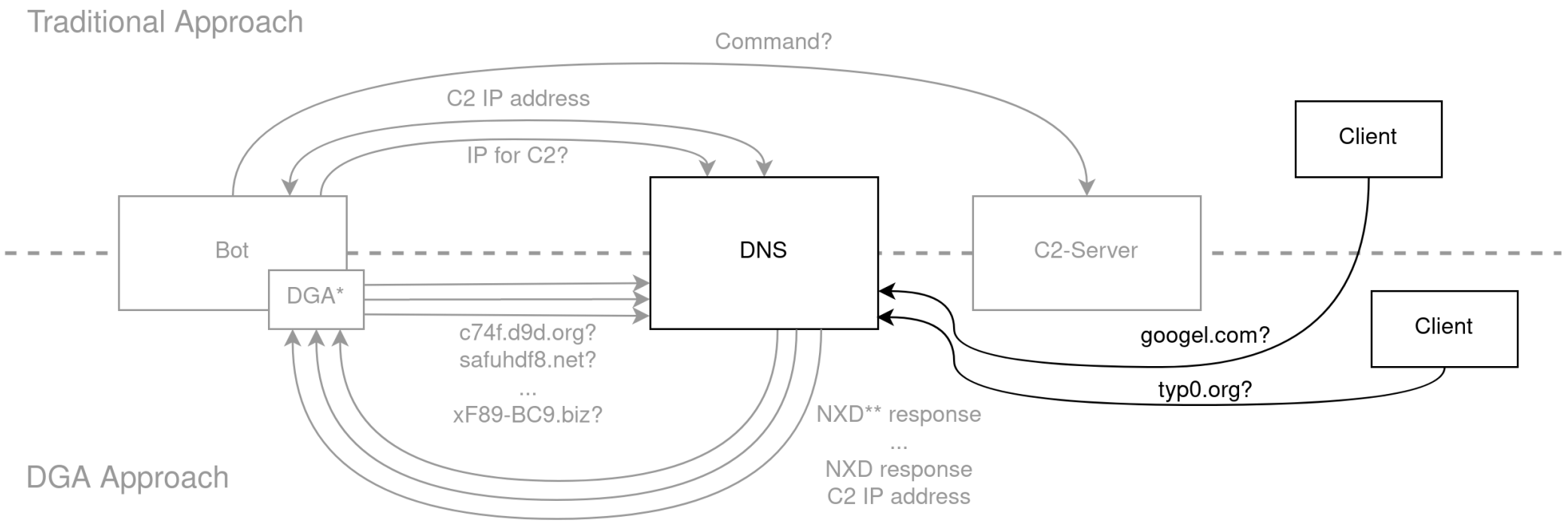


Context - Bot Communication



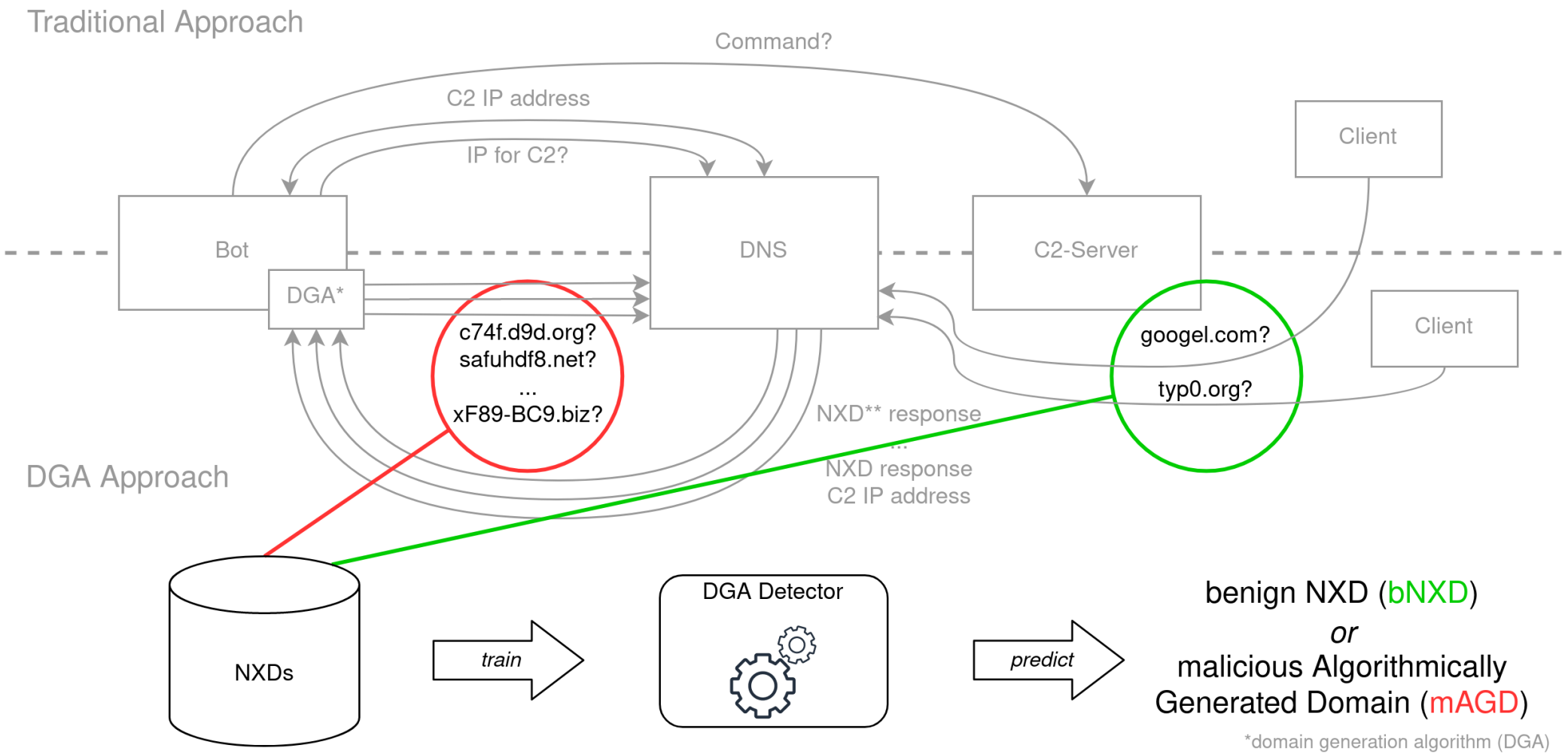
*domain generation algorithm (DGA)
**non-existent domain (NXD)

Context - Bot Communication



*domain generation algorithm (DGA)
**non-existent domain (NXD)

Context - DGA Detection



Challenge - Sharing and Privacy

Improving DGA detection

- Sharing data
- Sharing detection models
- Distributed training approaches

Challenge - Sharing and Privacy

Improving DGA detection

- Sharing data
- Sharing detection models
- Distributed training approaches

Privacy issues when sharing raw data

- Alternative data representation
- Alternative training algorithm

Challenge - Sharing and Privacy

Improving DGA detection

- Sharing data
- Sharing detection models
- Distributed training approaches

Privacy issues when sharing raw data

→ Alternative data representation

- Alternative training algorithm

Challenge - Sharing and Privacy

Improving DGA detection

- Sharing data
- Sharing detection models
- Distributed training approaches

Privacy issues when sharing raw data

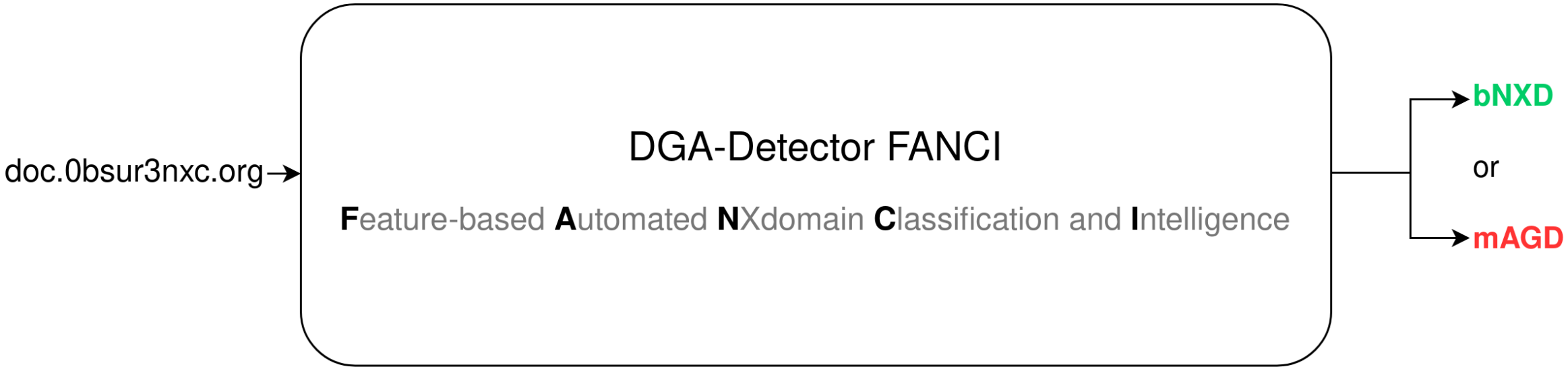
→ Alternative data representation

- Alternative training algorithm

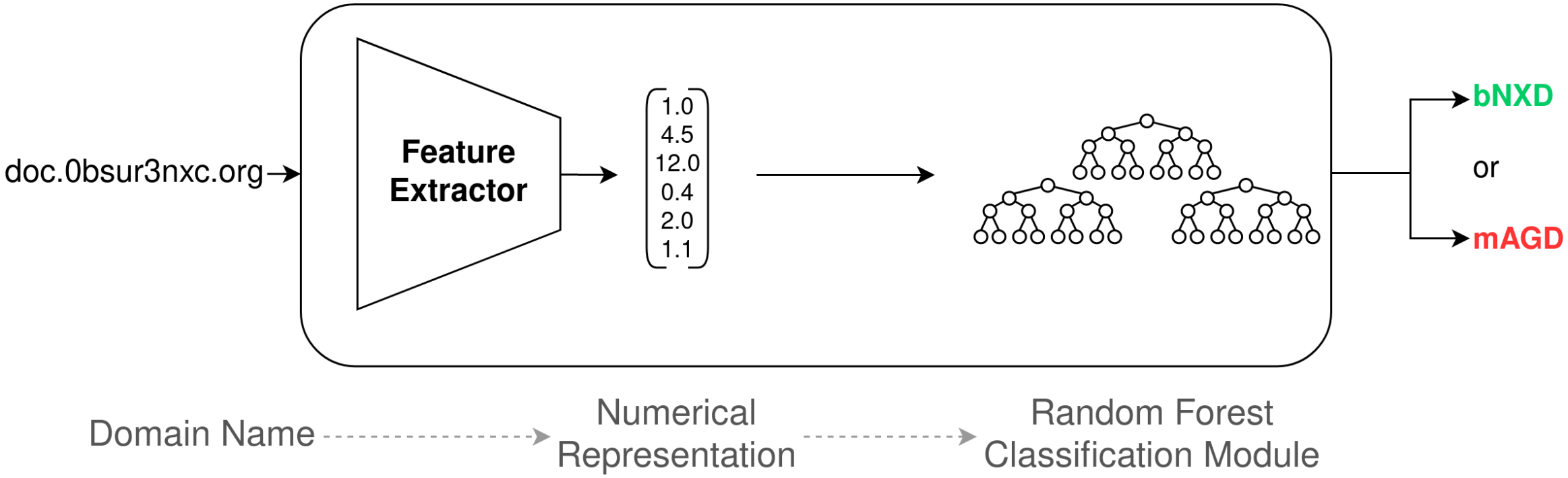
Why are bNXDs privacy sensitive?

- Disclosure of end-user browsing history / behaviour
- Disclosure of usage of misconfigured software
- Frequent bNXDs may be registered for phishing attacks

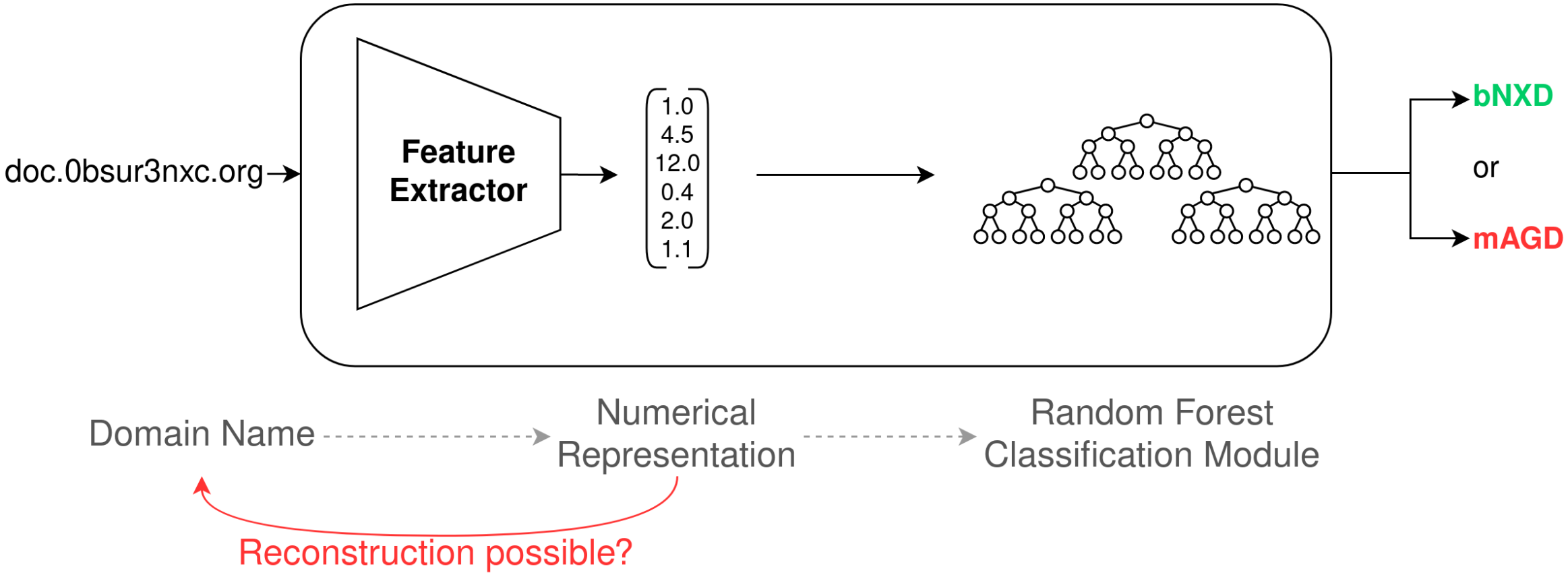
Object under Study: FANCI



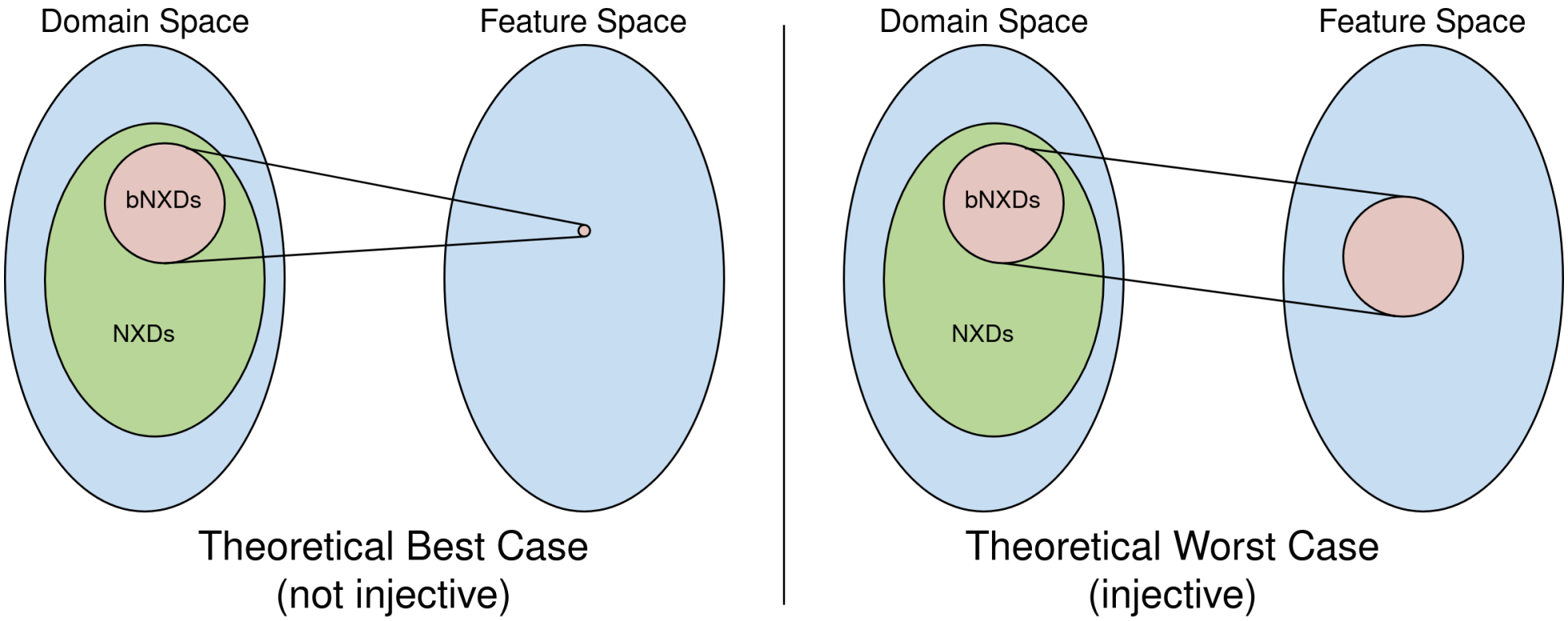
Object under Study: FANCI



Object under Study: FANCI

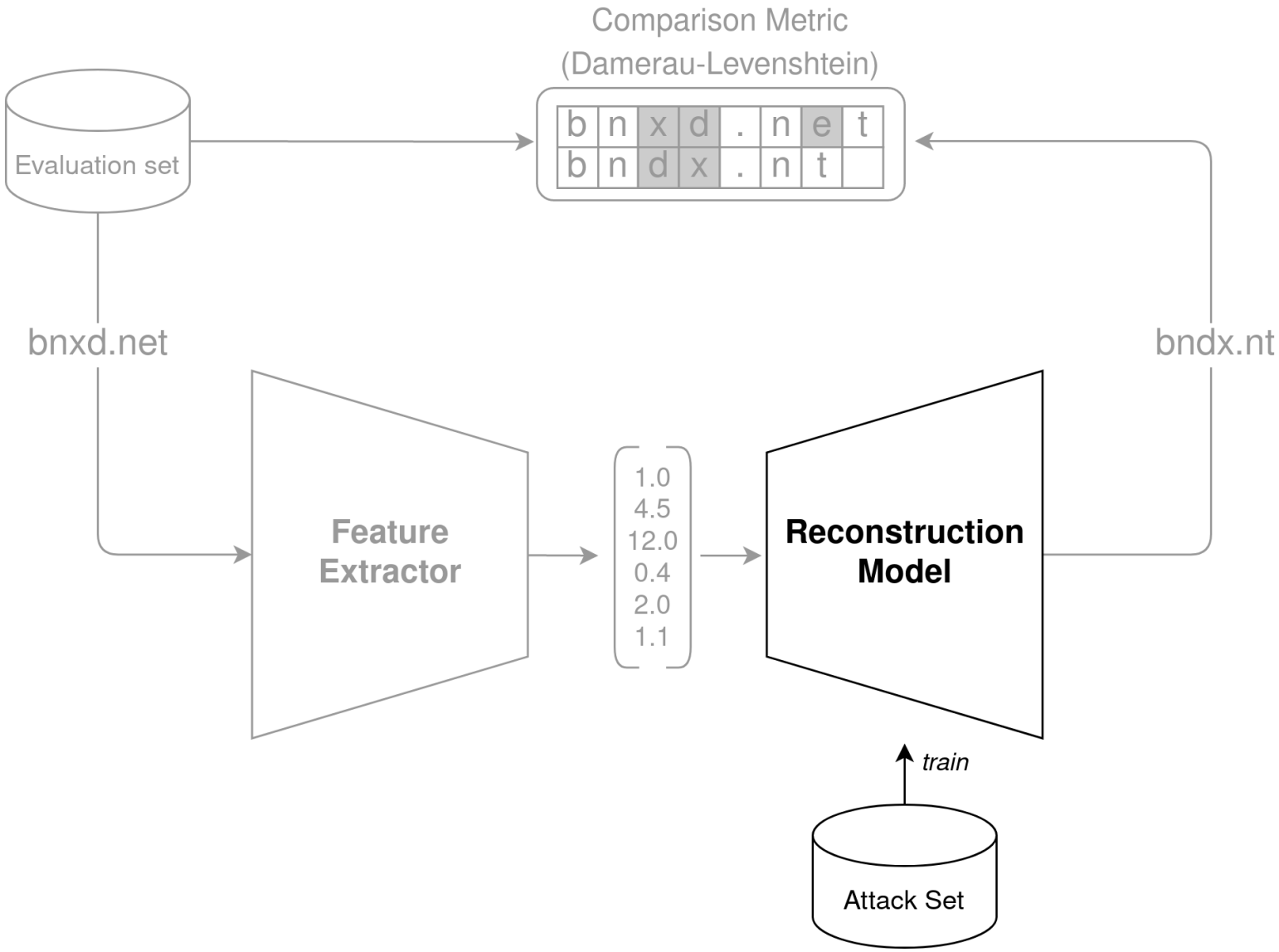


Privacy Aspects of Feature Extraction

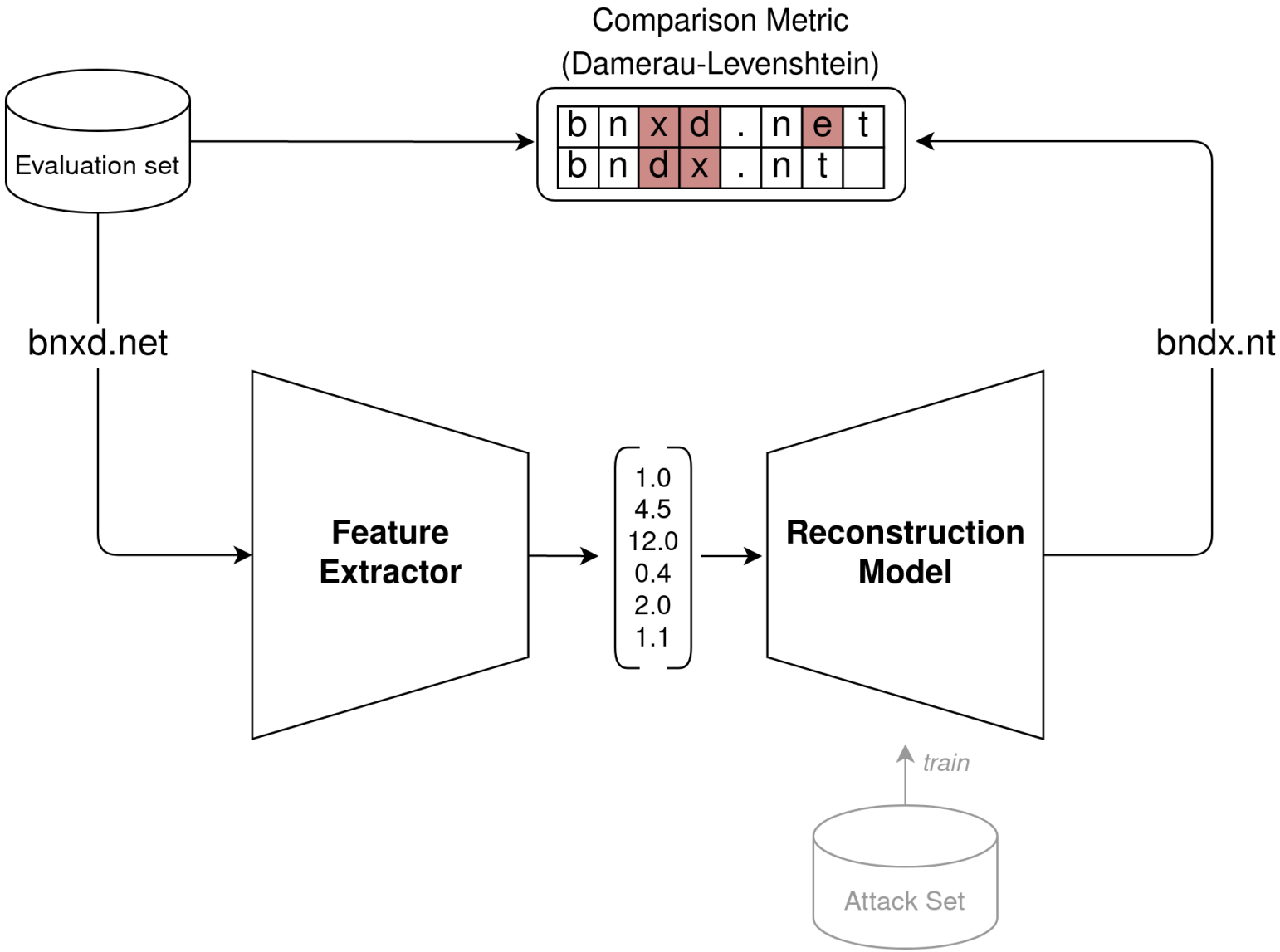


How to quantify suitability of feature representation w.r.t. privacy?

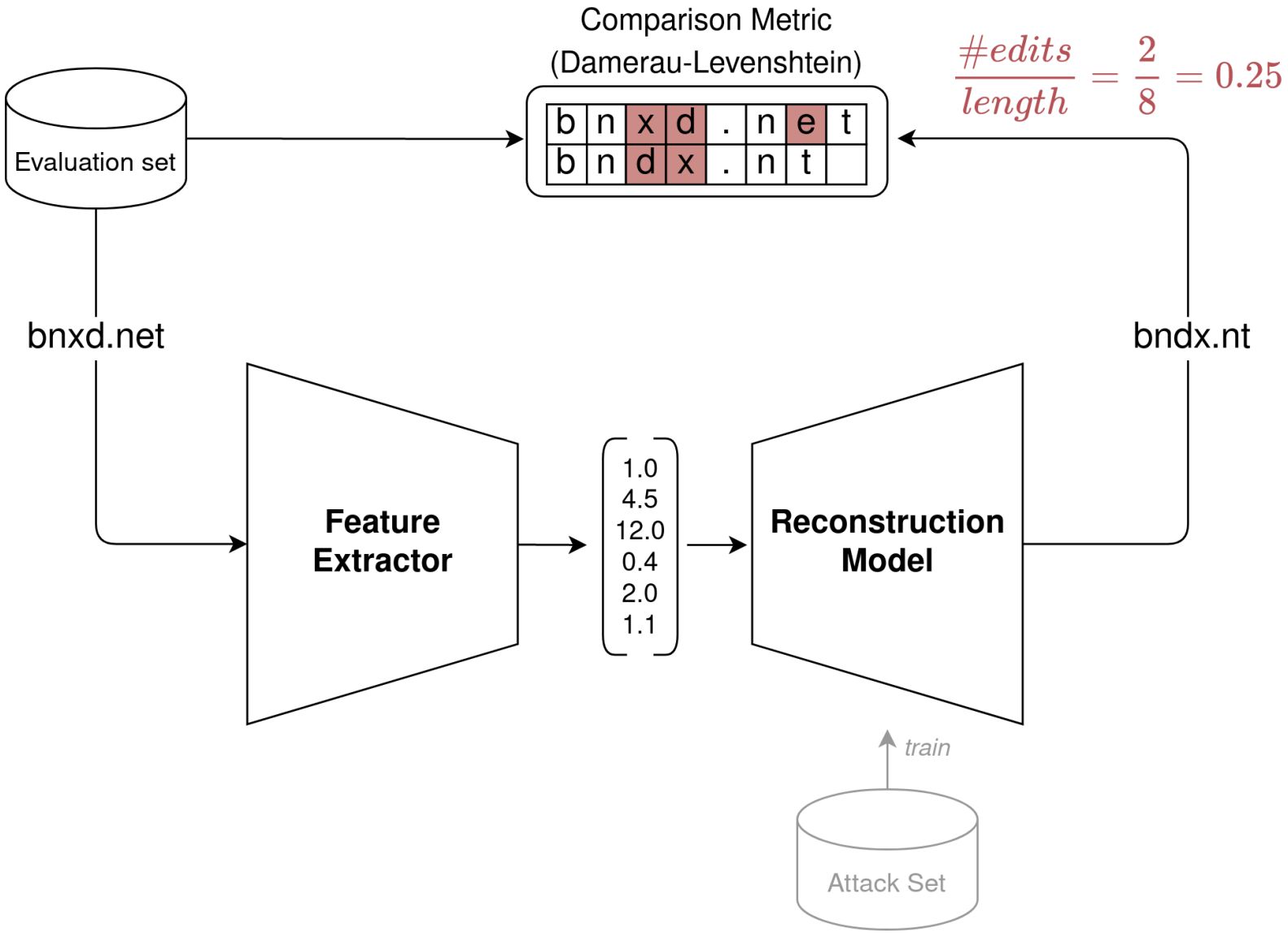
Data-driven Evaluation - Setup



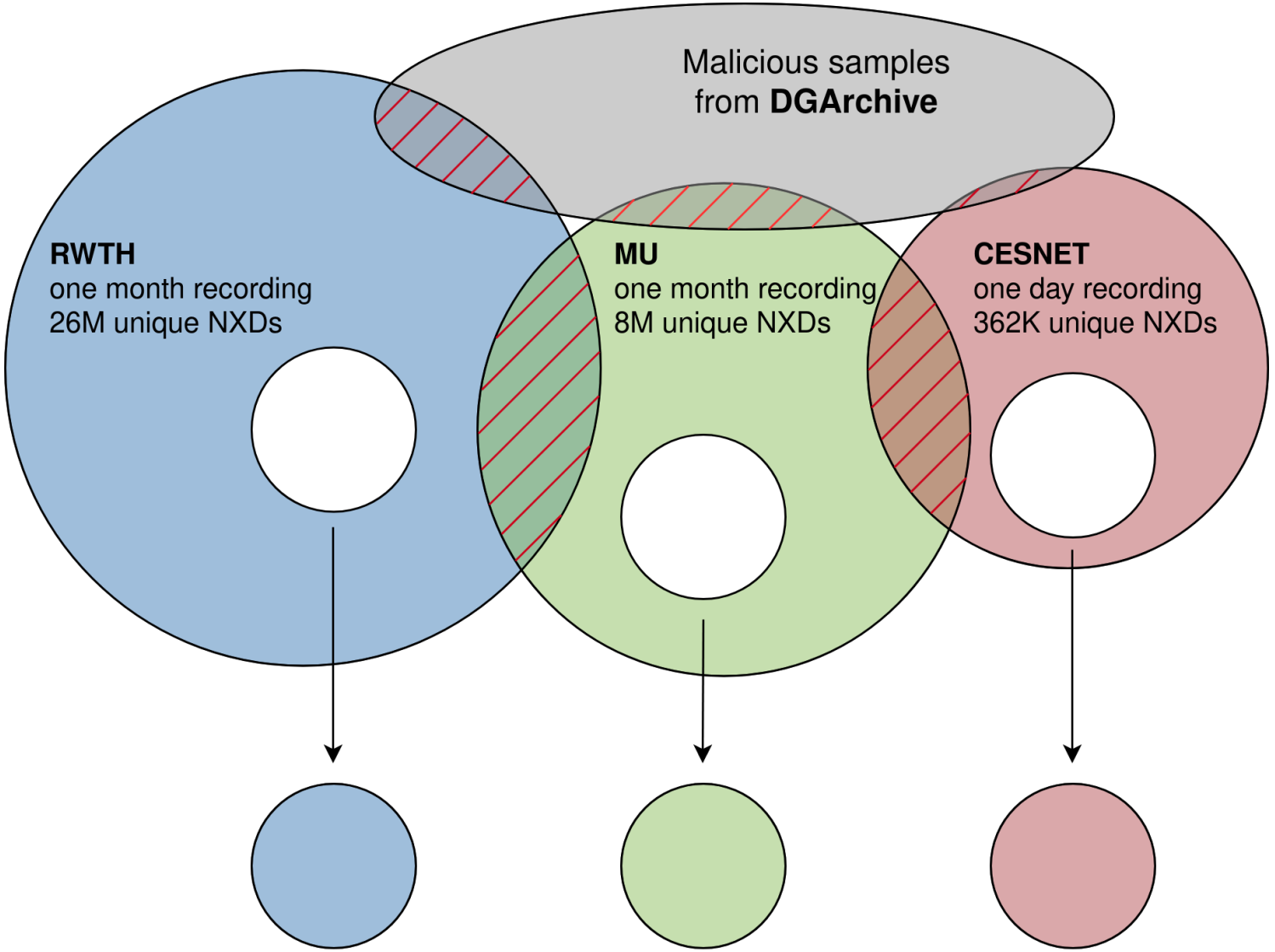
Data-driven Evaluation - Setup



Data-driven Evaluation - Setup



Data-driven Evaluation - Benign NXD Data Sets



Results - Reconstruction Error

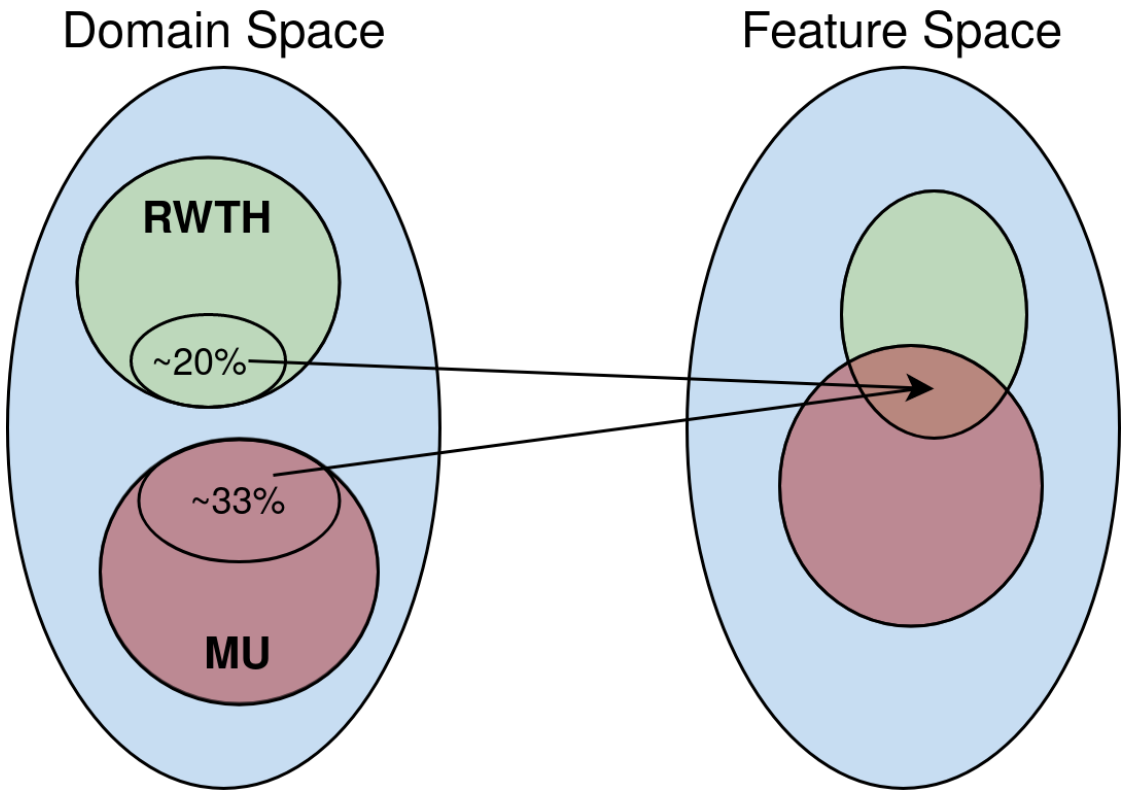
NXD Source		∅ Reconstruction Error
Attack	**Evaluation**	
RWTH	RWTH	0.51
MU	MU	0.53
CESNET	CESNET	0.46

Results - Reconstruction Error

NXD Source		∅ Reconstruction Error
Attack	**Evaluation**	
RWTH	RWTH	0.51
	MU	0.72
	CESNET	0.66
MU	RWTH	0.75
	MU	0.53
	CESNET	0.61
CESNET	RWTH	0.67
	MU	0.65
	CESNET	0.46

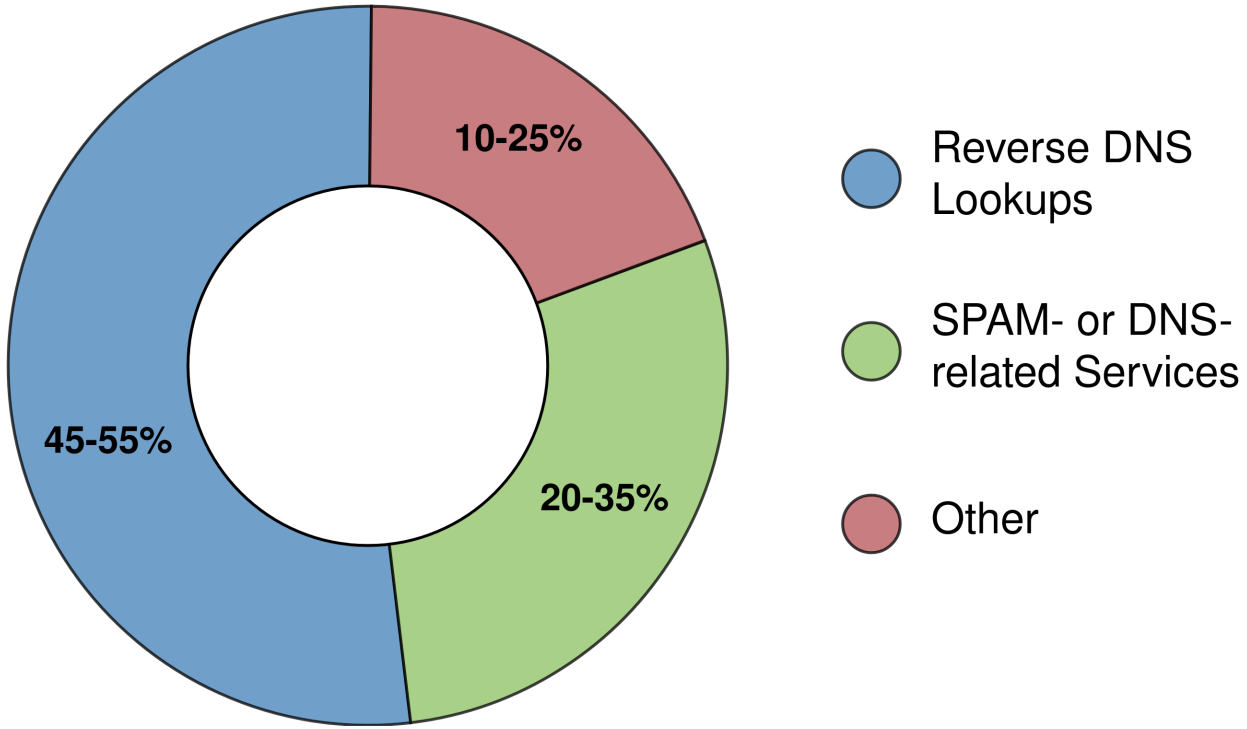
Discussion - Feature Space Overlap

Although comprising unique NXDs, the data sets intersect in the feature space, for instance:



Discussion - Top 10% Best Reconstructions

Average reconstruction performance in top 10% lies at 0.276



Those types of bNXD are not considered privacy critical

Summary & Outlook

FANCI's feature representation constitutes no considerable privacy threat

- We have quantified the risk using three real-world NXD data sets

Sharing scenarios based on this representation could be used in the future

- The privacy guarantee even holds for DGA multi-class classification

The evaluation framework is generally applicable

- Alternative feature extractors can be analyzed in the same manner