NETWORK TRAFFIC CLASSIFICATION BASED ON SINGLE FLOW TIME SERIES ANALYSIS

Authors:
Josef Koumar, *CESNET a.l.e. & CTU in Prague*
Karel Hynek Ph.D, CESNET a.l.e.
Tomáš Čejka Ph.D, CESNET a.l.e.

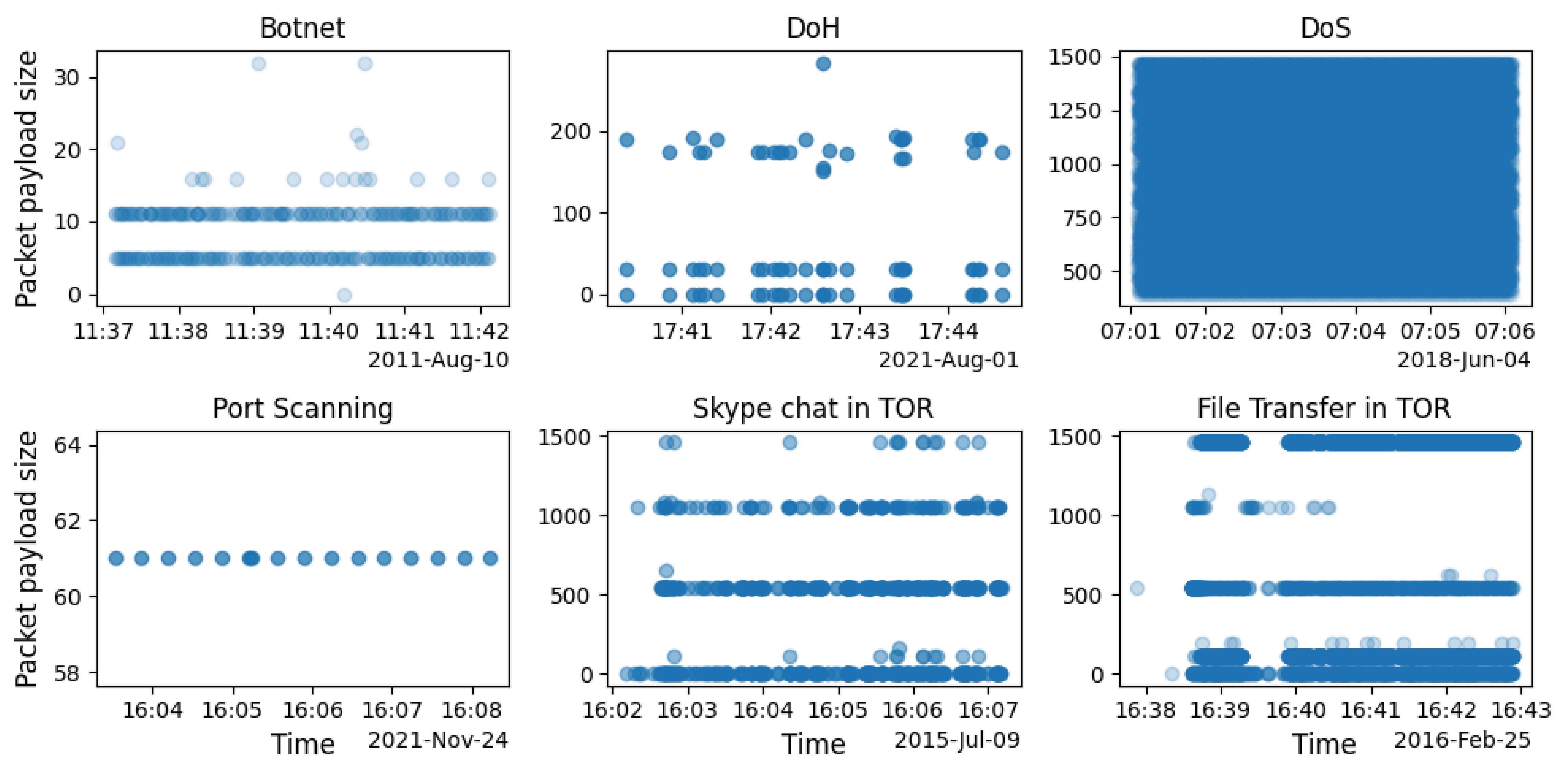
PROBLEM

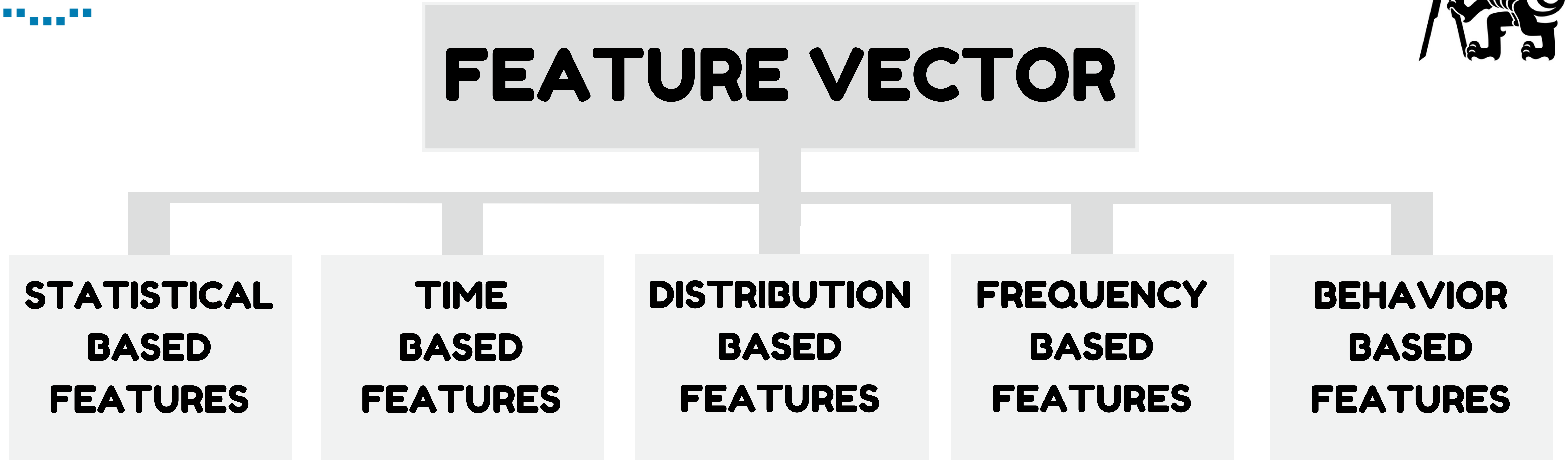
Motivation

- Privacy protections designed to help users also protect attackers from being detected
- Current methods must work with a few pieces of informations from the network traffic

Therefore, we define a Single Flow Time Series, i.e., packet time series of the IP flow for the purpose of description of IP flow by Time Series Analysis.

SINGLE FLOW TIME SERIES





FEATURE VECTOR

STATISTICAL BASED FEATURES

TIME BASED FEATURES

DISTRIBUTION BASED FEATURES

FREQUENCY BASED FEATURES

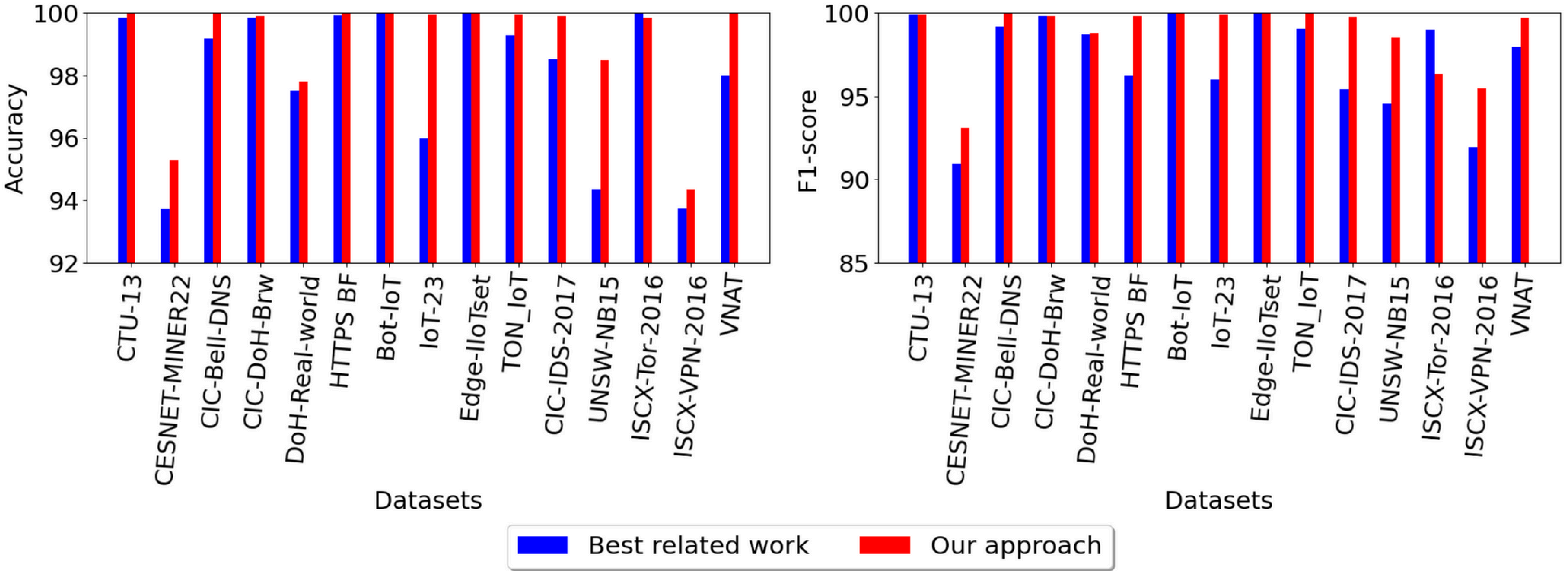
BEHAVIOR BASED FEATURES

- 69 features
- Examples: Mean, Entropy, Time distribution, Hurst exponent, Spectral bandwidth, Spectral crest, Periodicity, Transients, ...

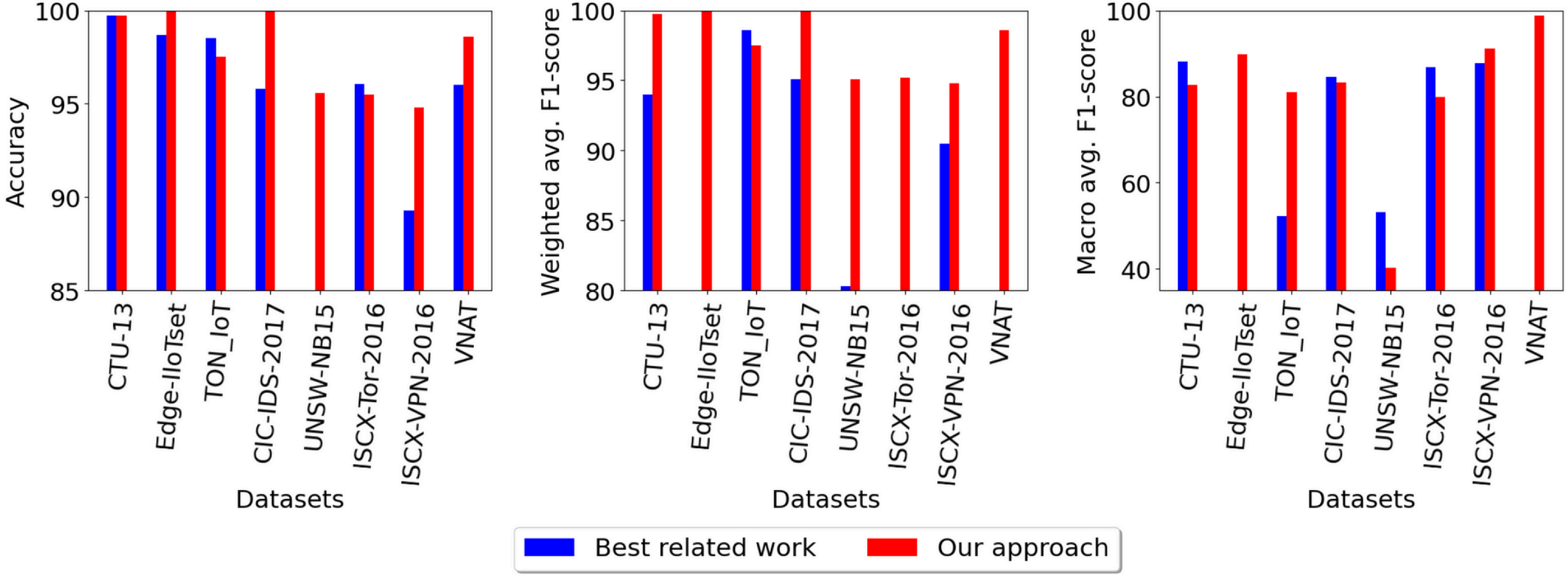
METHODOLOGY

1. **Create dataset** of IP flows extended by the feature vector from the PCAP files
2. Split the dataset of IP flows on the **Train, Validation, and Test** parts in a ratio of 60:20:20
3. Use Train and Validation parts for a **hyperparameters tunning** of XGBoost model
4. Train the **XGBoost model** on the Train part using obtained hyperparameters
5. **One-time test** of the model using the Test part

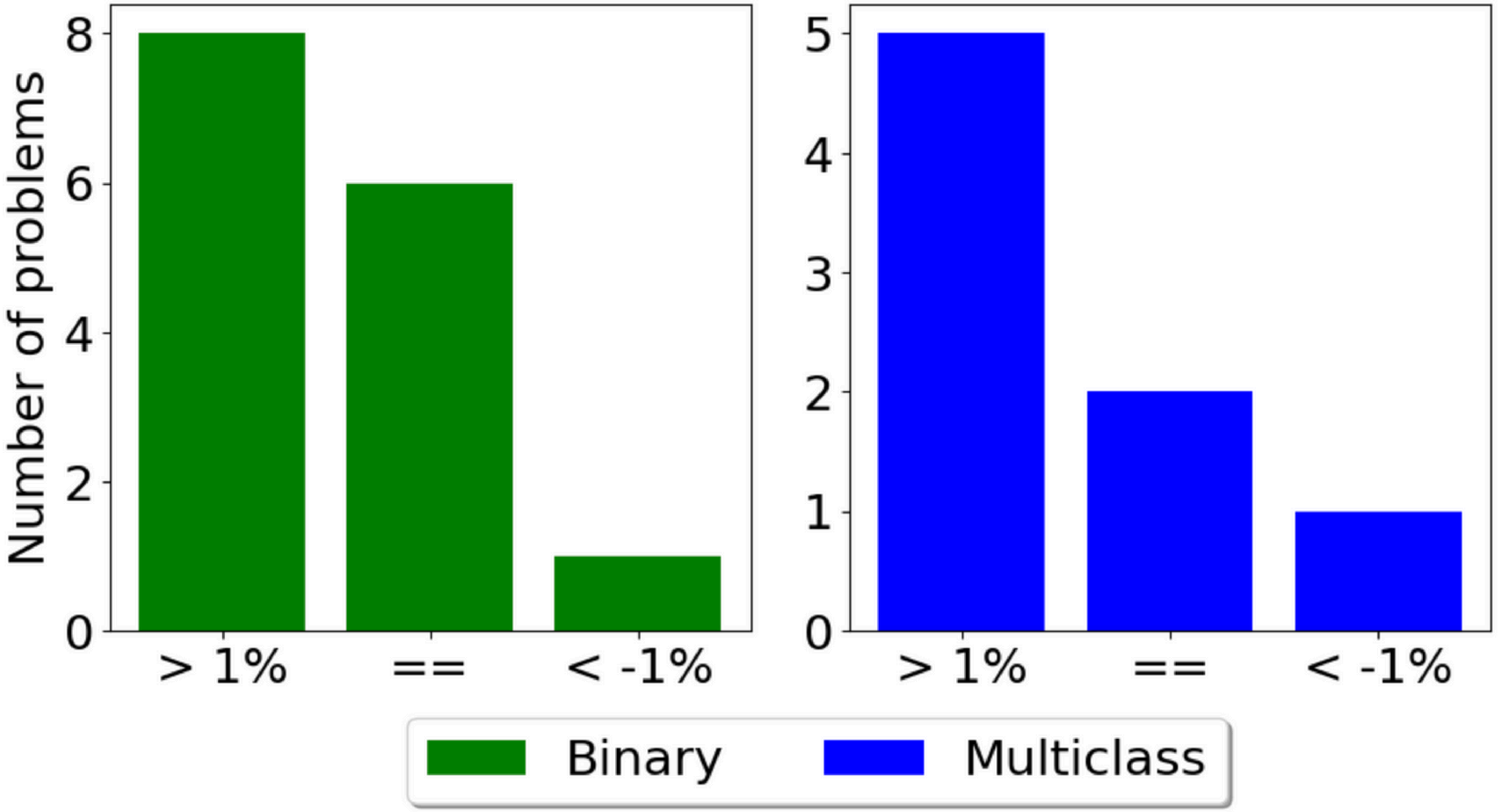
BINARY CLASSIFICATION



MULTICLASS CLASSIFICATION

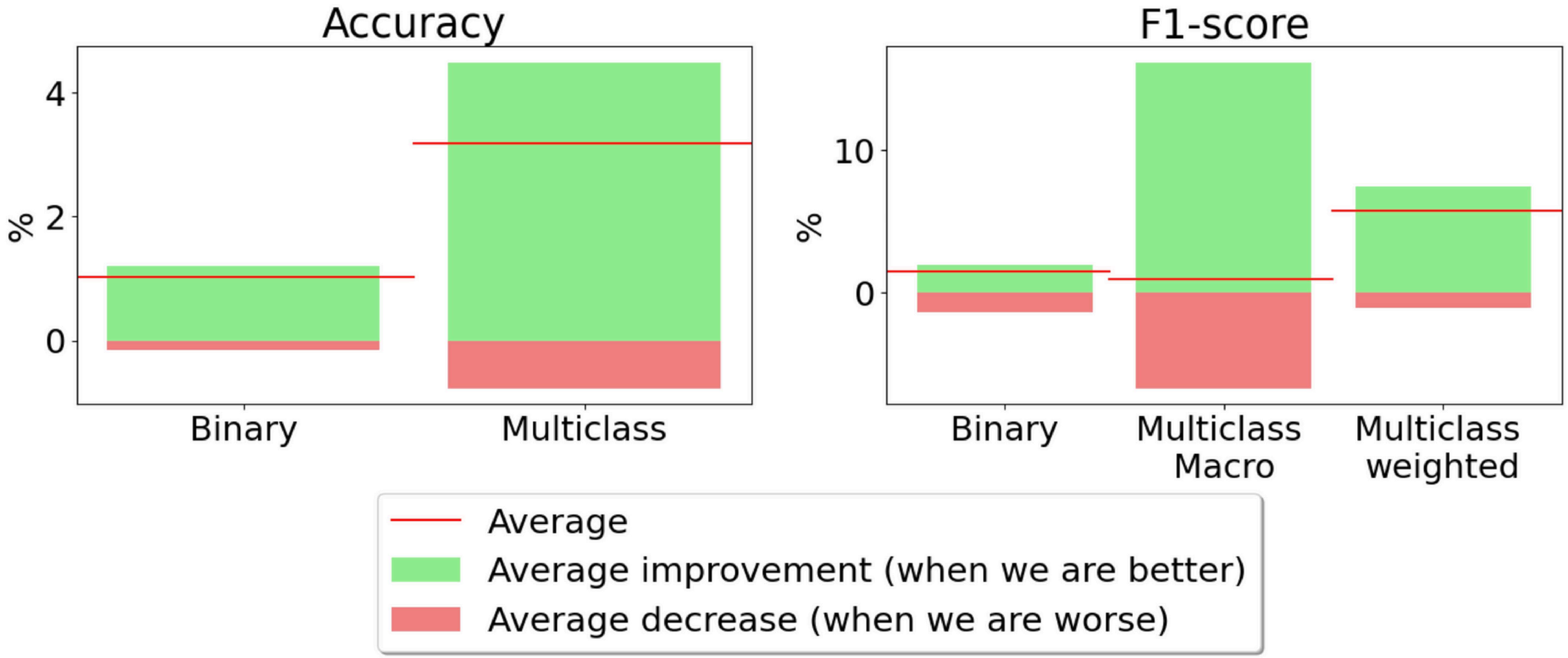


OVERALL COMPARISON



Statistical distribution of problems by comparison with best related work

OVERALL COMPARISON



CONCLUSION

- **Novel approach** using 69 features
- Create datasets of presented features from **15** well-known network datasets
- Created datasets are **publicly available** on Zenodo
- Source codes are **publicly available** on GitHub
- The novel approach achieved **improvement in accuracy and F1-score** then previous best results from relevant works
- Future work: **NetTiSA: Extended IP Flow with Time-series Features for Universal Bandwidth-constrained High-speed Network Traffic Classification** (read our preprint now!)

Ipfixprobe flow exporter:

Created datasets:

Thank you!

CONTACTS

koumajos@fit.cvut.cz
hynekkar@cesnet.cz
cejkat@cesnet.cz