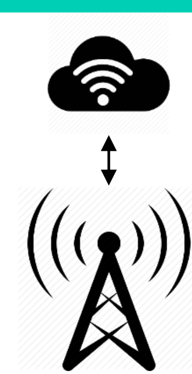
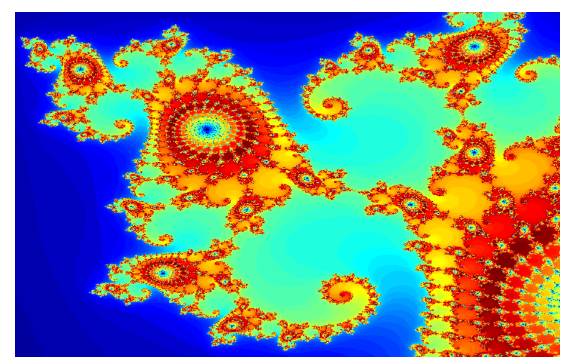
Early online classification of encrypted traffic streams using multi-fractal features

Erik Areström, *Linköping University*
Niklas Carlsson, *Linköping University*

LINKÖPING UNIVERSITY





Motivation and problem

- Early flow classification is important for network operators in order to operate network at high utilization while still providing good quality of experience for the users

Motivation and problem

- Early flow classification is important for network operators in order to operate network at high utilization while still providing good quality of experience for the users
- End-to-end encryption render traditional deep packet inspection techniques useless

Motivation and problem

- Early flow classification is important for network operators in order to operate network at high utilization while still providing good quality of experience for the users
- End-to-end encryption render traditional deep packet inspection techniques useless
- Most flow classification approaches are unable to properly capture the non-linear characteristics of network flows

Motivation and problem

- Early flow classification is important for network operators in order to operate network at high utilization while still providing good quality of experience for the users
- End-to-end encryption render traditional deep packet inspection techniques useless
- Most flow classification approaches are unable to properly capture the non-linear characteristics of network flows
- Problem: Current classification methods are too slow or inaccurate to benefit network operators

Contributions

- A man-in-the-middle based evaluation framework, utilizing the multi-fractal features of encrypted traffic flows to diffrentiate application types

Contributions

- A man-in-the-middle based evaluation framework, utilizing the multi-fractal features of encrypted traffic flows to diffrentiate application types
- Early traffic categorization via tuning of said framwork achieving F1-scores of 0.814 after only 5 seconds, using only multi-fractal features

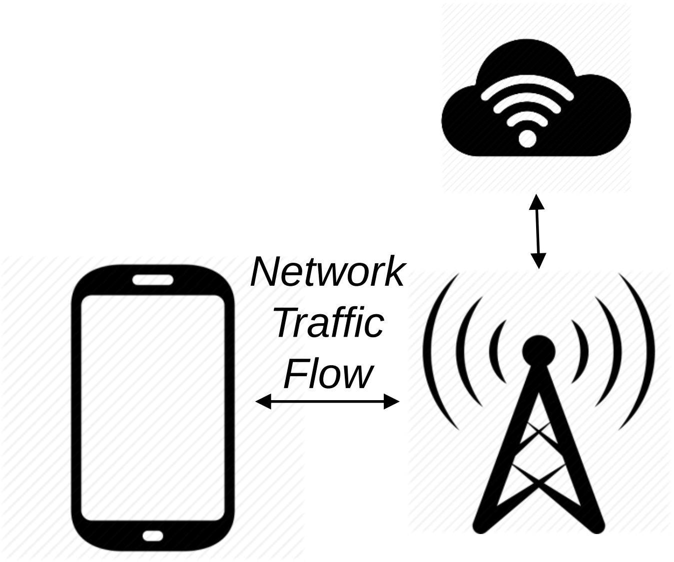
Contributions

- A man-in-the-middle based evaluation framework, utilizing the multi-fractal features of encrypted traffic flows to diffrentiate application types
- Early traffic categorization via tuning of said framwork achieving F1-scores of 0.814 after only 5 seconds, using only multi-fractal features
- In-class categorization of live video versus video on demand delivered from the same services, using only multi-fractal features

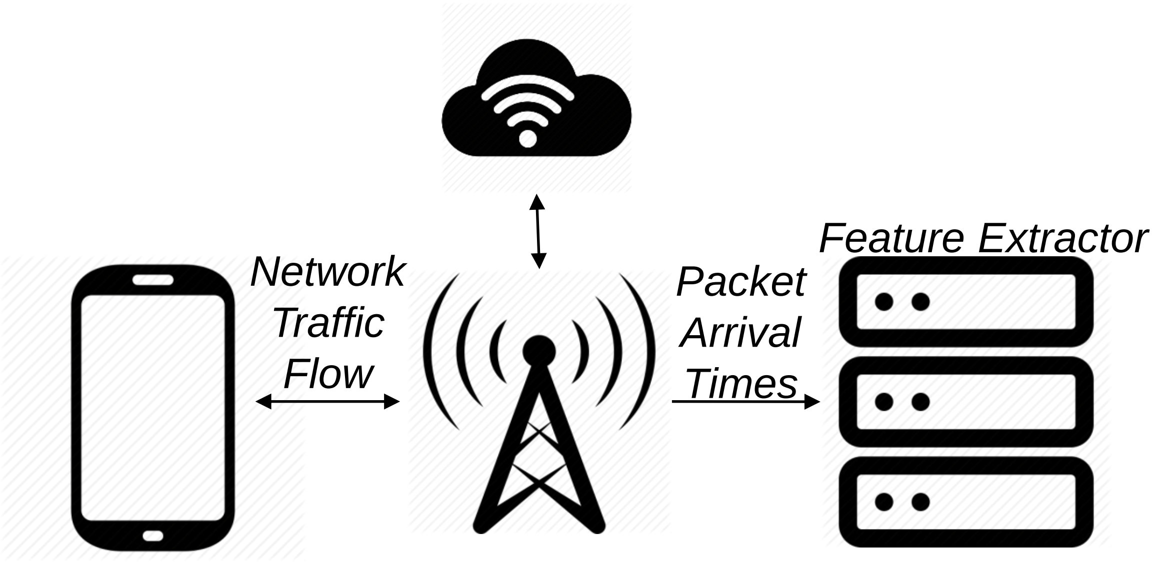
High-level categorization

Application categories	Example service
Video streaming	Youtube
Web browsing	Reddit
Social media	Facebook
Audio communication	Skype
Text communication	Messenger
Bulk download	Google Play

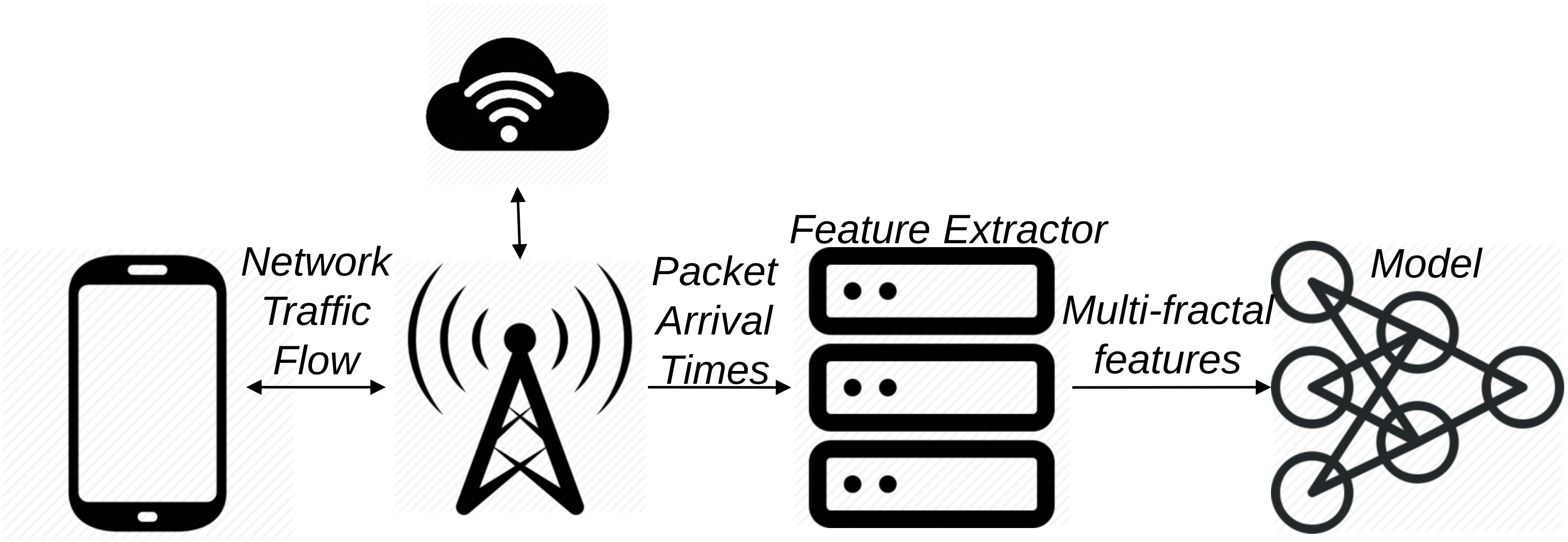
System model



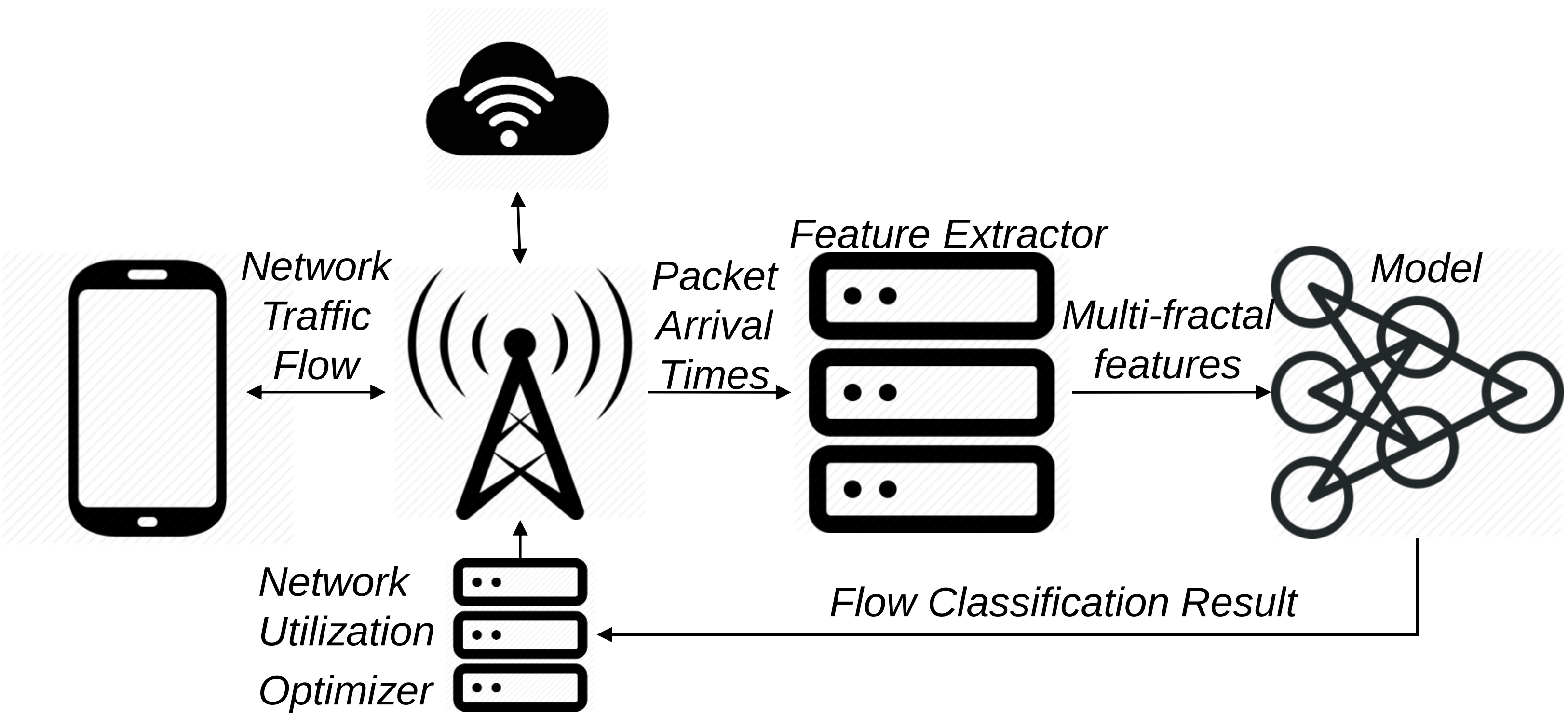
System model



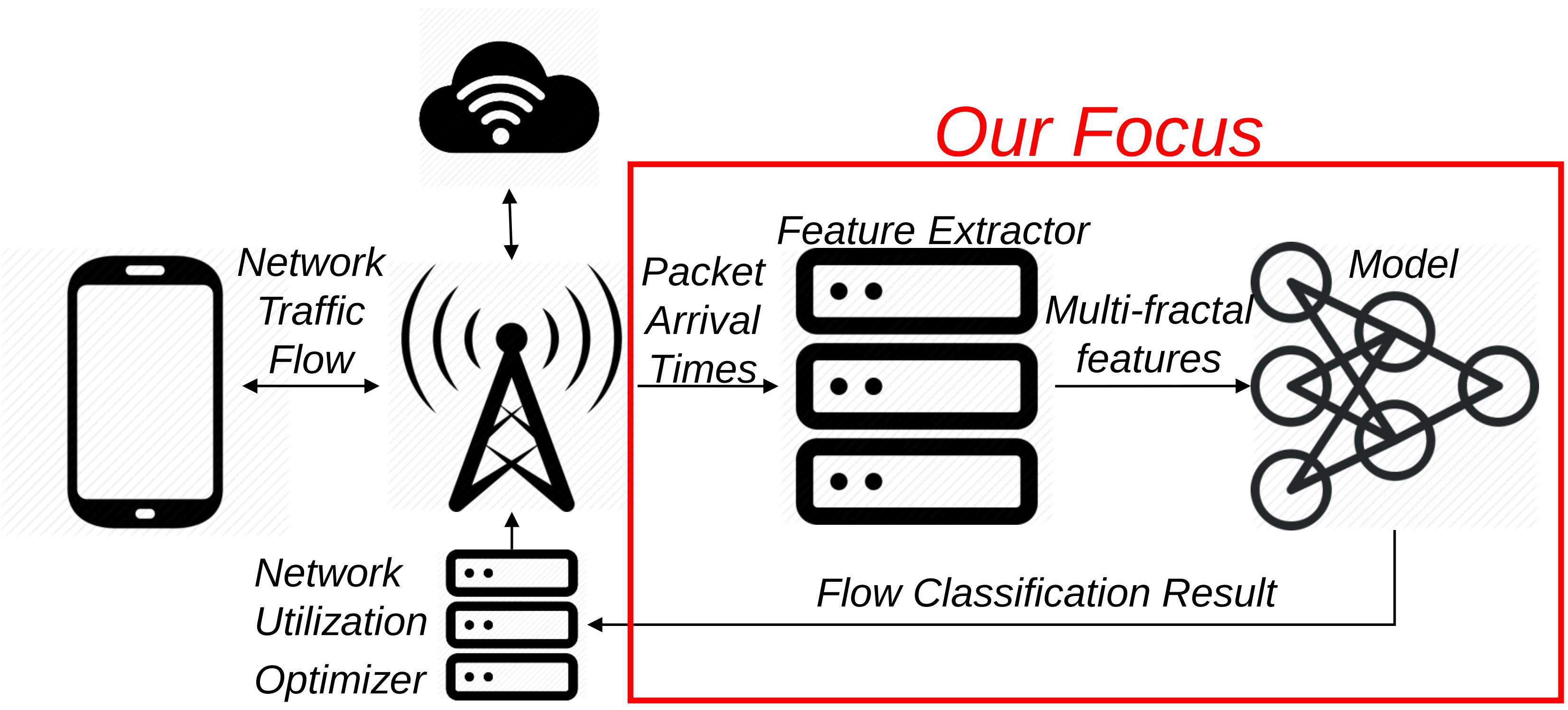
System model



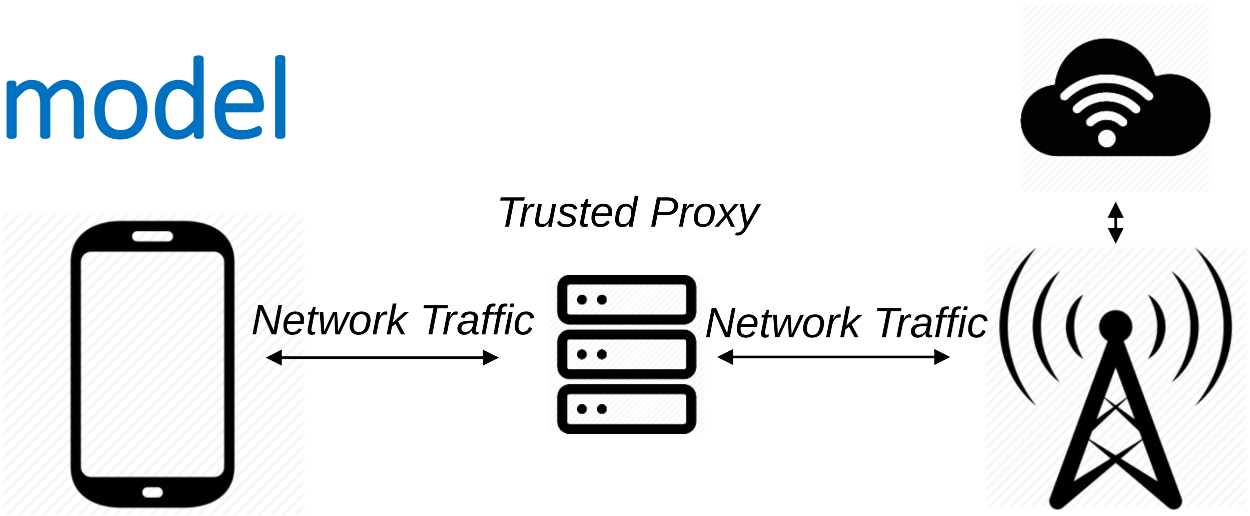
System model



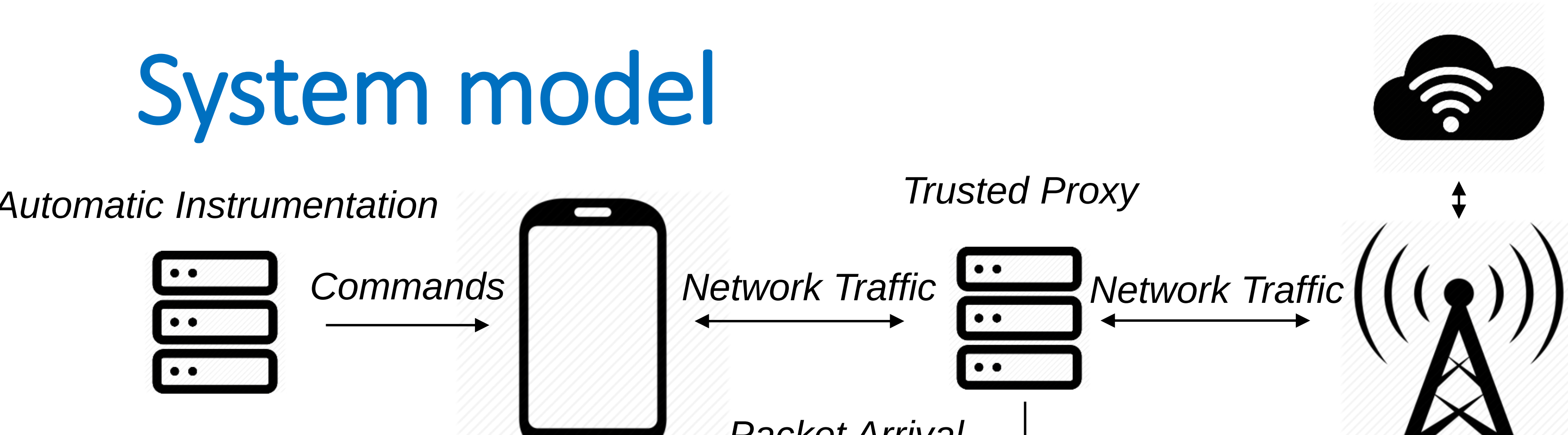
System model



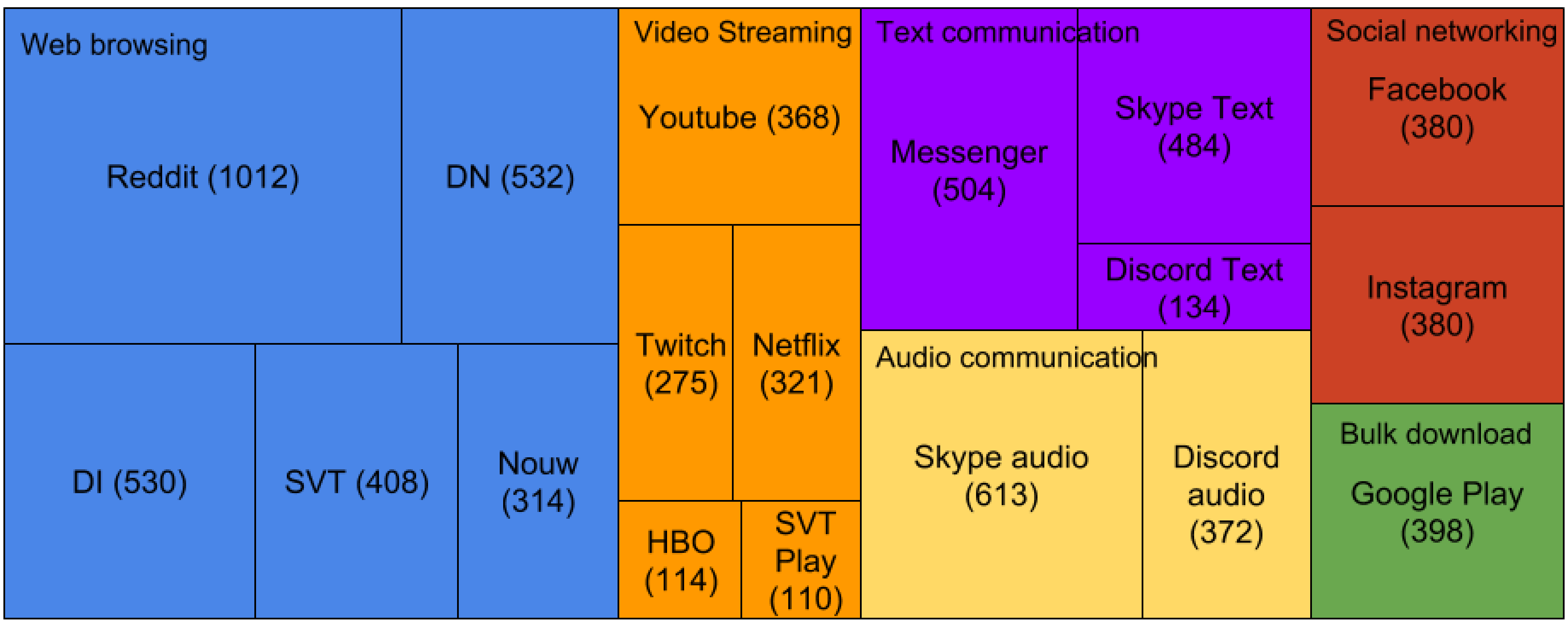
System model



System model



The samples



Feature extraction

Feature extraction

- Given a time series repesenting the arrival of a packet in a timeslot, calculate the wavelet coefficients for different scales of the signal using the Discrete Wavelet Transform

Feature extraction

- Given a time series repesenting the arrival of a packet in a timeslot, calculate the wavelet coefficients for different scales of the signal using the Discrete Wavelet Transform
- Extract the time- or space localized suprema of the coefficents, the so called wavelet leaders

Feature extraction

- Given a time series repesenting the arrival of a packet in a timeslot, calculate the wavelet coefficients for different scales of the signal using the Discrete Wavelet Transform
- Extract the time- or space localized suprema of the coefficents, the so called wavelet leaders
- Form a multi-resolution structure function to estimate the scaling exponents by regression

Feature extraction

- Given a time series repesenting the arrival of a packet in a timeslot, calculate the wavelet coefficients for different scales of the signal using the Discrete Wavelet Transform
- Extract the time- or space localized suprema of the coefficents, the so called wavelet leaders
- Form a multi-resolution structure function to estimate the scaling exponents by regression
- Derive the **Hausdorff dimensions** and corresponding **Holder Exponents** for the signal

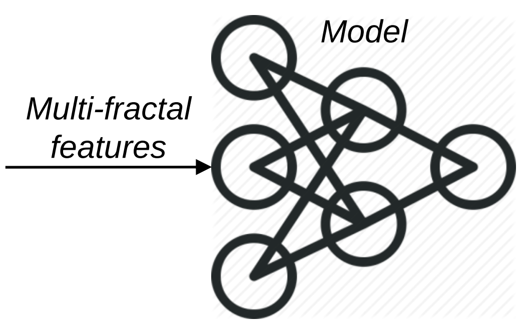
Feature extraction

- Given a time series repesenting the arrival of a packet in a timeslot, calculate the wavelet coefficients for different scales of the signal using the Discrete Wavelet Transform
- Extract the time- or space localized suprema of the coefficents, the so called wavelet leaders
- Form a multi-resolution structure function to estimate the scaling exponents by regression
- Derive the **Hausdorff dimensions** and corresponding **Holder Exponents** for the signal

The multi-fractal features, representing how the observed self-similiarty of the signal changes over time

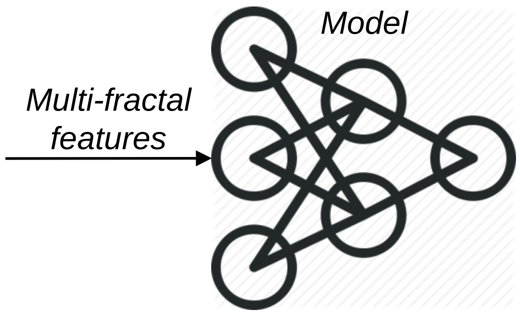
Building the model

- The collection of samples were randomly split into two parts, half the samples were used to build the model

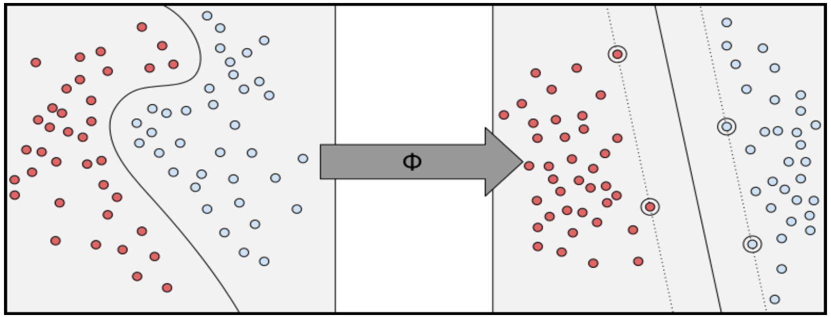


Building the model

- The collection of samples were randomly split into two parts, half the samples were used to build the model
- Multiple Binary Support Vector Machine classifiers were used, fitting the maximun margin separating hyperplane between each class of data



SVM with radial basis kernel function



Evaluation (t = 20 s)

	Audio Com.	Bulk Down.	Text Com.	Social Media	Video	Web	**Precision**
Audio Com.	4847	6	9	11	5	24	0.989
Bulk Down.	5	1949	8	5	4	0	0.989
Text Com.	59	1	5389	7	3	152	0.960
Social Media	2	1	12	3459	49	308	0.903
Video	5	2	10	175	5251	95	0.948
Web	2	1	192	143	178	13441	0.963
Recall	0.985	0.994	0.959	0.910	0.956	0.959	0.960

Targeted class

Output class

Class	F1-score
Audio Communication	0.98
Bulk Download	0.99
Text Communication	0.96

Evaluation (t = 20 s)

Output class \ Targeted class	Audio Com.	Bulk Down.	Text Com.	Social Media	Video	Web	Precision
Audio Com.	4847	6	9	11	5	24	0.989
Bulk Down.	5	1949	8	5	4	0	0.989
Text Com.	59	1	5389	7	3	152	0.960
Social Media	2	1	12	3459	49	308	0.903
Video	5	2	10	175	5251	95	0.948
Web	2	1	192	143	178	13441	0.963
Recall	0.985	0.994	0.959	0.910	0.956	0.959	0.960

Targeted class

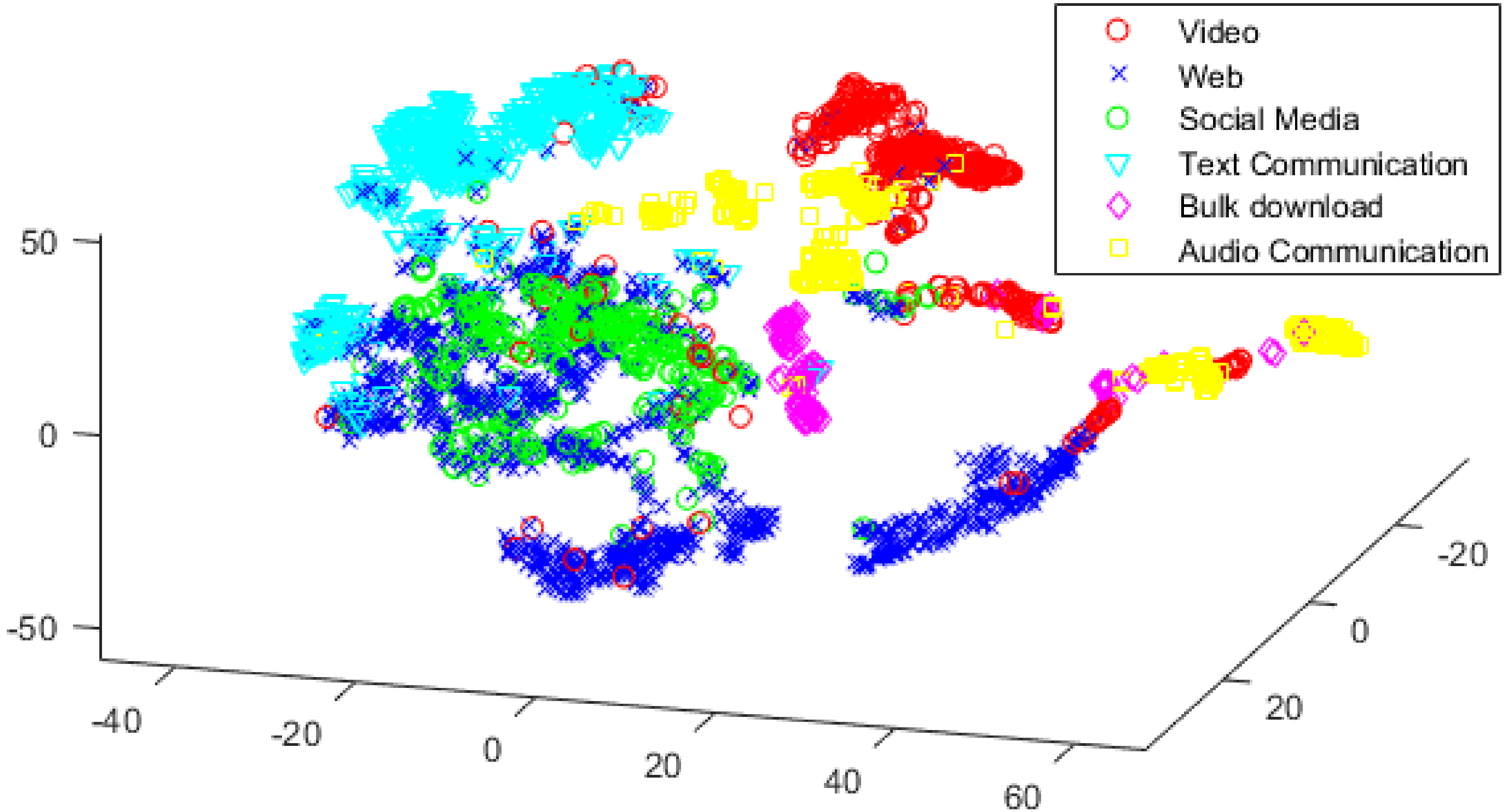
Class	F1-score
Audio Communication	0.98
Bulk Download	0.99
Text Communication	0.96
Social Media	0.90
Video	0.96
Web	0.96

Evaluation (t = 20 s)

Output class \ Targeted class	Audio Com.	Bulk Down.	Text Com.	Social Media	Video	Web	**Precision**
Audio Com.							
Bulk Down.							
Text Com.							
Social Media				3459	49	308	
Video				175	5251	95	
Web				143	178	13441	
Recall							

Class	F1-score
Audio Communication	0.98
Bulk Download	0.99
Text Communication	0.96
Social Media	0.90
Video	0.96
Web	0.96

T-SNE visualization



Early classification

Duration	F1-score	Precision	Recall
20 seconds	0.958	0.958	0.958

Early classification

Duration	F1-score	Precision	Recall
20 seconds	0.958	0.958	0.958
15 seconds	0.892	0.891	0.894
10 seconds	0.844	0.838	0.851

Early classification

Duration	F1-score	Precision	Recall
20 seconds	0.958	0.958	0.958
15 seconds	0.892	0.891	0.894
10 seconds	0.844	0.838	0.851
5 seconds	0.814	0.823	0.805

Early classification

Duration	F1-score	Precision	Recall
20 seconds	0.958	0.958	0.958
15 seconds	0.892	0.891	0.894
10 seconds	0.844	0.838	0.851
5 seconds	0.814	0.823	0.805
2.5 seconds	0.631	0.594	0.673

Early classification

Duration	F1-score	Precision	Recall
20 seconds	0.958	0.958	0.958
15 seconds	0.892	0.891	0.894
10 seconds	0.844	0.838	0.851
5 seconds	0.814	0.823	0.805
2.5 seconds	0.631	0.594	0.673
2 seconds	0.409	0.404	0.415
1 second	0.214	0.202	0.228

Randomly picking one category: 1/6 ≈ 0.167

Impact of added variance in the dataset.

- All packet arrival instances in the evaulation set were perturbed according to a normal distribution:

$$\mathcal{N}(0, \sigma)$$

σ	10	25	50	100	250	500	1000
F1-score	0.952	0.942	0.925	0.927	0.891	0.834	0.695

Impact of added variance in the dataset.

- All packet arrival instances in the evaulation set were perturbed according to a normal distribution:

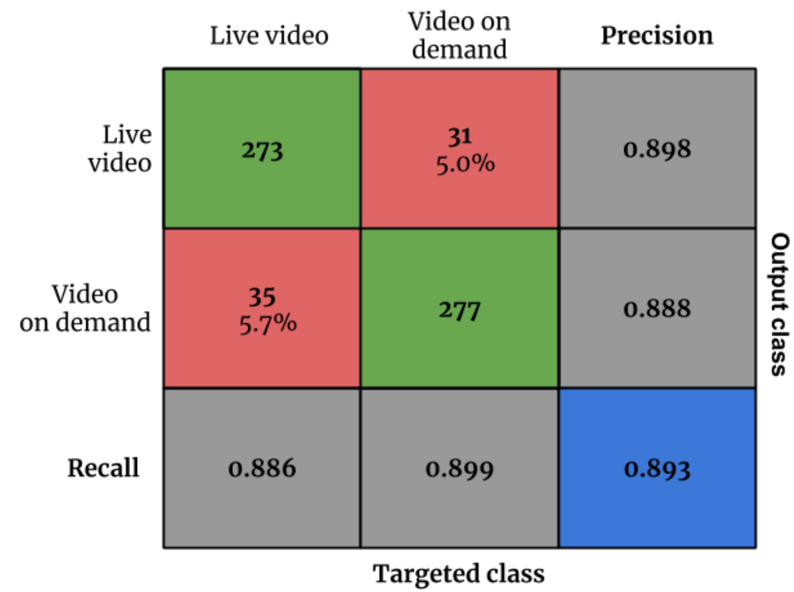
$$\mathcal{N}(0, \sigma)$$

σ	10	25	50	100	250	500	1000
F1-score	0.952	0.942	0.925	0.927	0.891	0.834	0.695

31.8% of the packets arrivals move by more than ± 0.5 seconds

In-class categorization, live vs VoD

- Same IP addresses may be used for both live and VoD content, categorization needs to be done online



Category	Live	Vod
Samples	616	616
Class Composition	Youtube: 214 Twitch: 214 SVT Play: 188	Youtube: 214 Twitch: 214 SVT Play: 188

Conclusion

- The classification method used is able to quickly and effectivly classify encrypted traffic belong to the six most popular traffic types

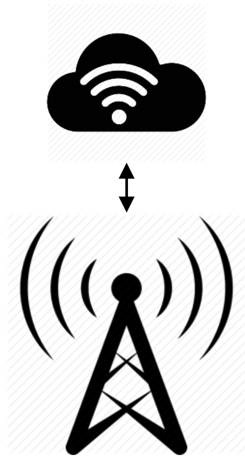
Conclusion

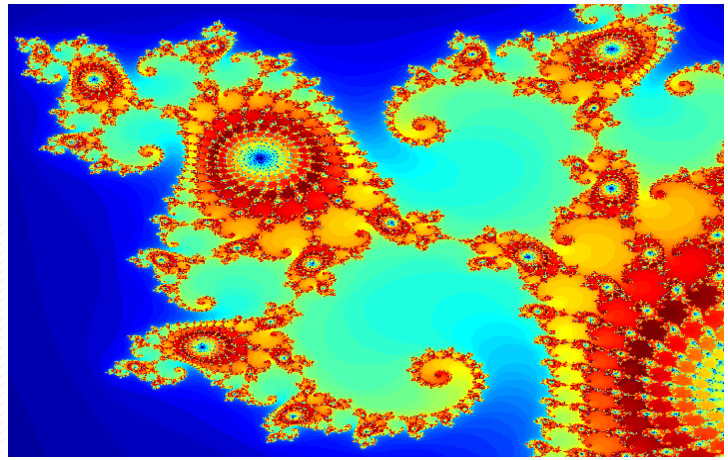
- The classification method used is able to quickly and effectivly classify encrypted traffic belong to the six most popular traffic types
- The method relies only on access to timing information of the packets in a flow and is highly resistant to perturbations of this information

Conclusion

- The classification method used is able to quickly and effectivly classify encrypted traffic belong to the six most popular traffic types
- The method relies only on access to timing information of the packets in a flow and is highly resistant to perturbations of this information
- The method can be applied to distinguish between classes of data belonging to the same services (Vod and live streaming)

Thanks for listening!





Early online classification of encrypted traffic streams using multi-fractal features

LINKÖPING UNIVERSITY

Erik Areström (erik.arestrom@gmail.com)
Niklas Carlsson (niklas.carlsson@liu.se)