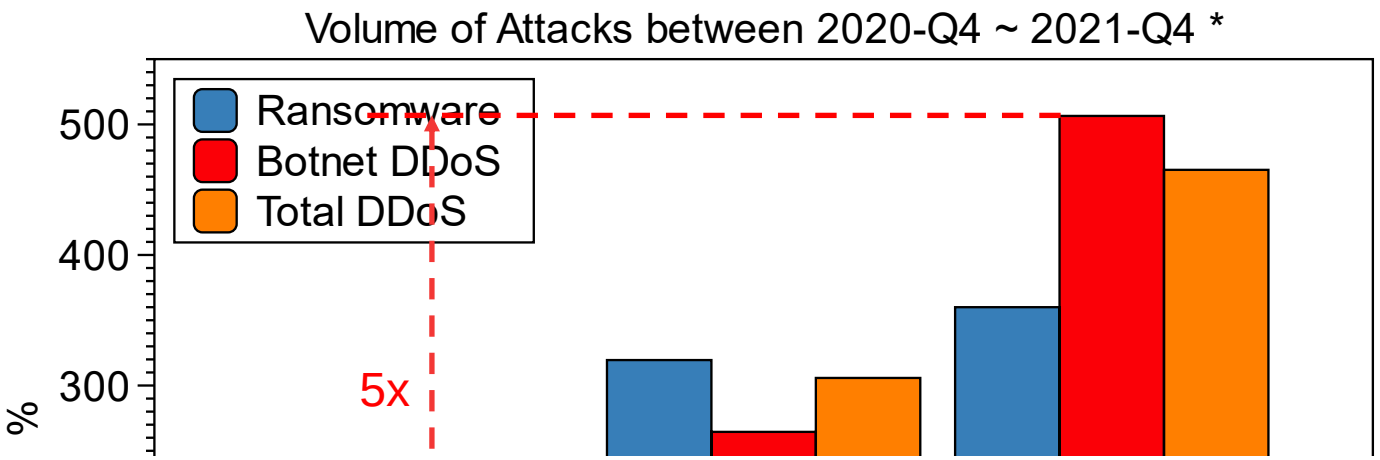
NetSentry: A Deep Learning Approach to Detecting Incipient Large-scale Network Attacks

Haoyu Liu, Paul Patras

The University of Edinburgh

Large-Scale Network Attacks



EUROPE NEWS

Cyberattack hits Norway, pro-Russian hacker group suspected

PUBLISHED WED, JUN 29 2022•11:28 PM EDT

AP

SHARE

KEY POINTS

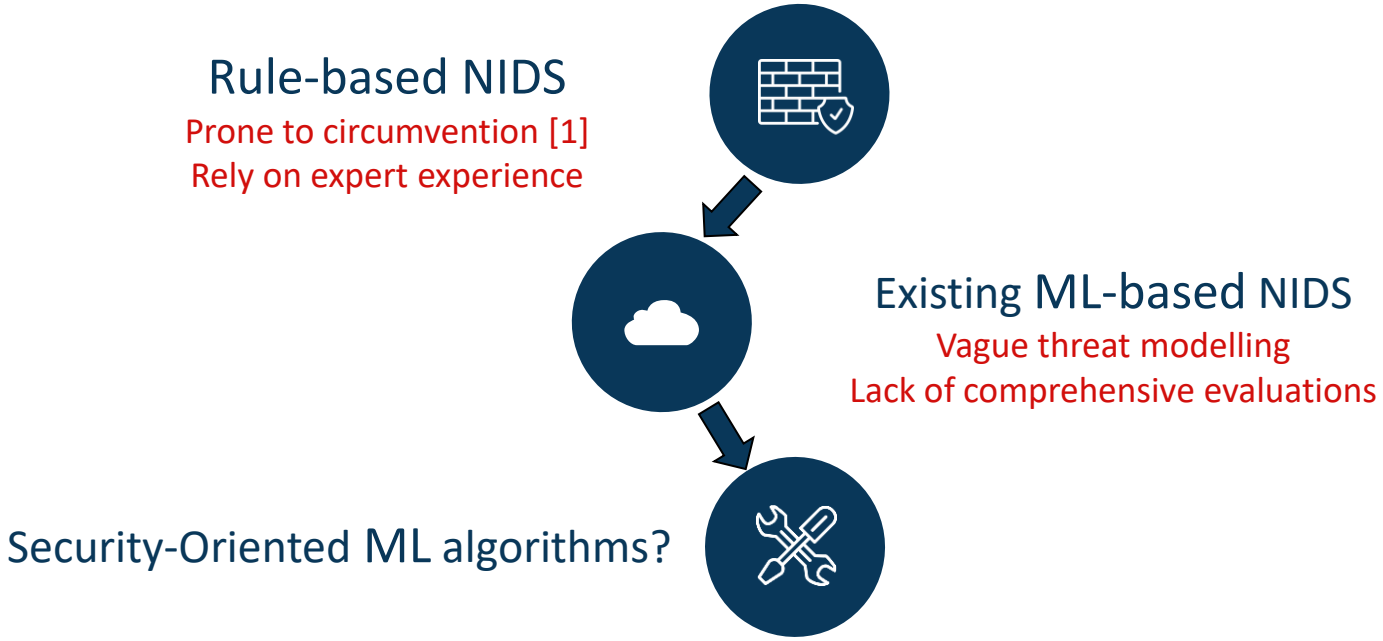
- A cyberattack temporarily knocked out public and private websites in Norway in the past 24 hours, Norwegian authorities said Wednesday.

CNBC TV

Shark Tank WATC

UP NEXT | **Shark Tank** 04:00 pm ET

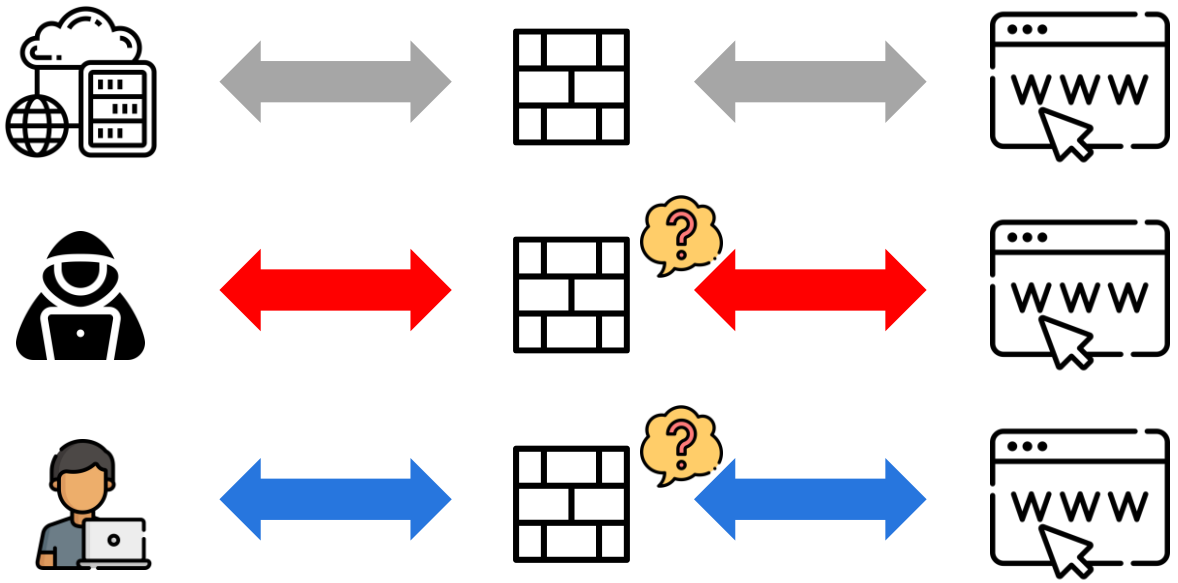
Existing Countermeasures



[1] L. Bilge and T. Dumitraș, “Before we knew it: An empirical study of zero-day attacks in the real world,” in 2012 ACM CCS.

Why Threat Modelling not trivial?

Per-flow classification/clustering/reconstruction

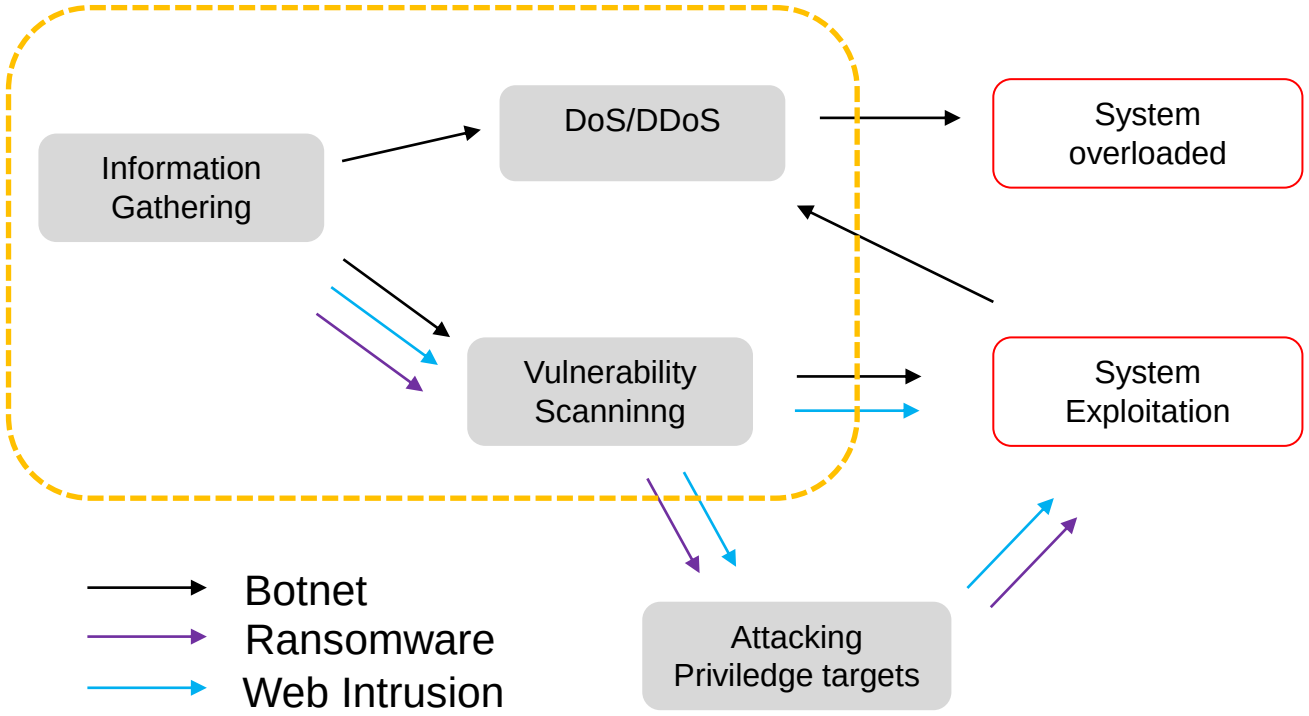


Performance – Cross Evaluation

Algorithms	CSE-CIC-IDS2018 (F1 score)
MLP	0.998
CNN	0.995

[1] *Mirsky, et, al."Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection" In NDSS, 2018*
[2] *Bo, et, al. "Deep autoencoding gaussian mixture model for unsupervised anomaly detection." In ICLR. 2018*
[3] Ruff, et, al. “Deep One-Class Classification.” in *ICML, 2018*

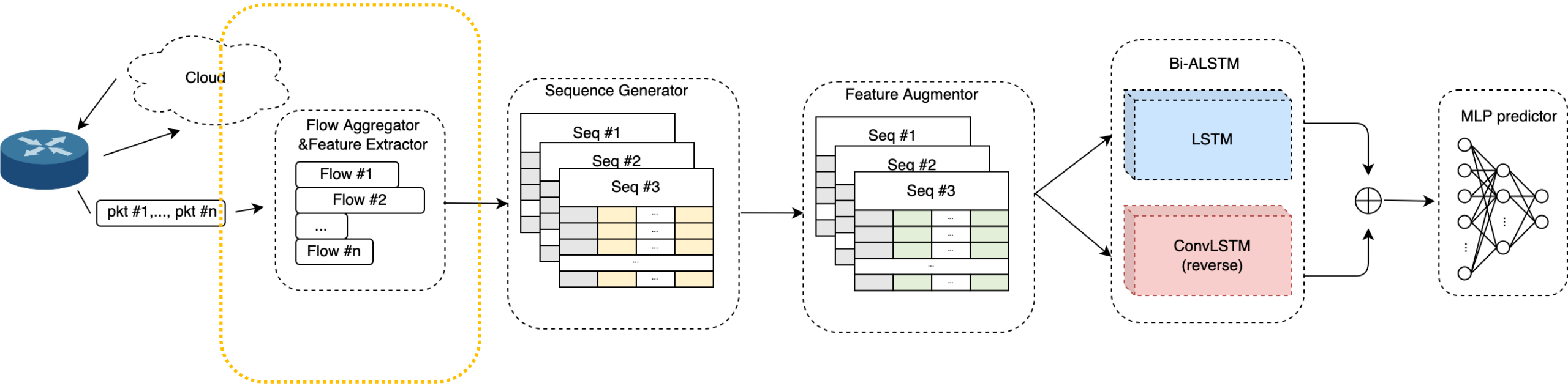
Attack Chains



Insight:

- Similarity of the early stage of intrusion
- Similar traffic -> temporal dependency matters

Defending solution: NetSentry

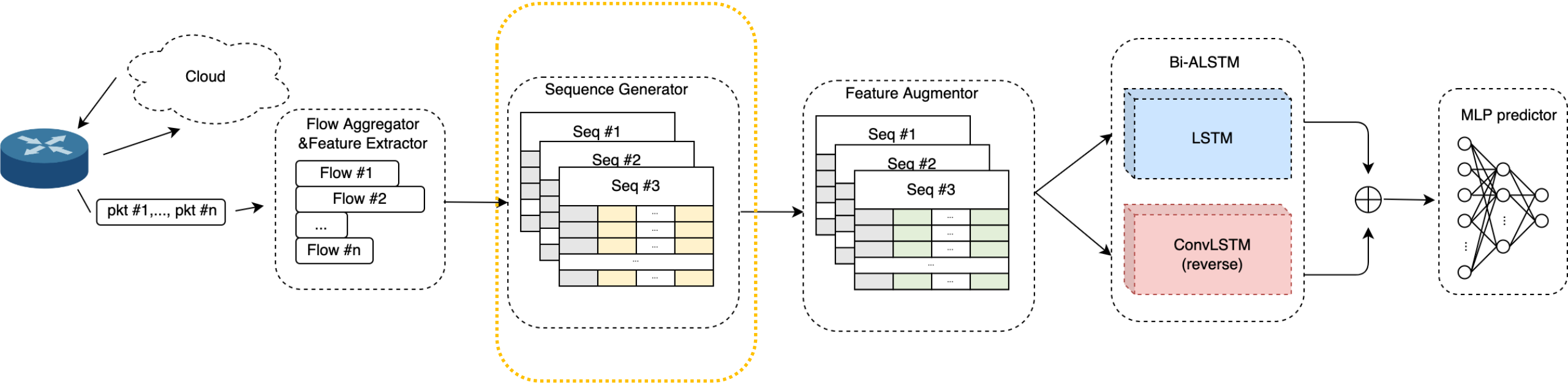


Flow : TCP/ UDP connection

aggregated by (src ip, dst ip, src port, dst port, protocol)

Feature Extractor: CICFlowMeter (CIC-IDS-2017/2018)

Defending solution: NetSentry



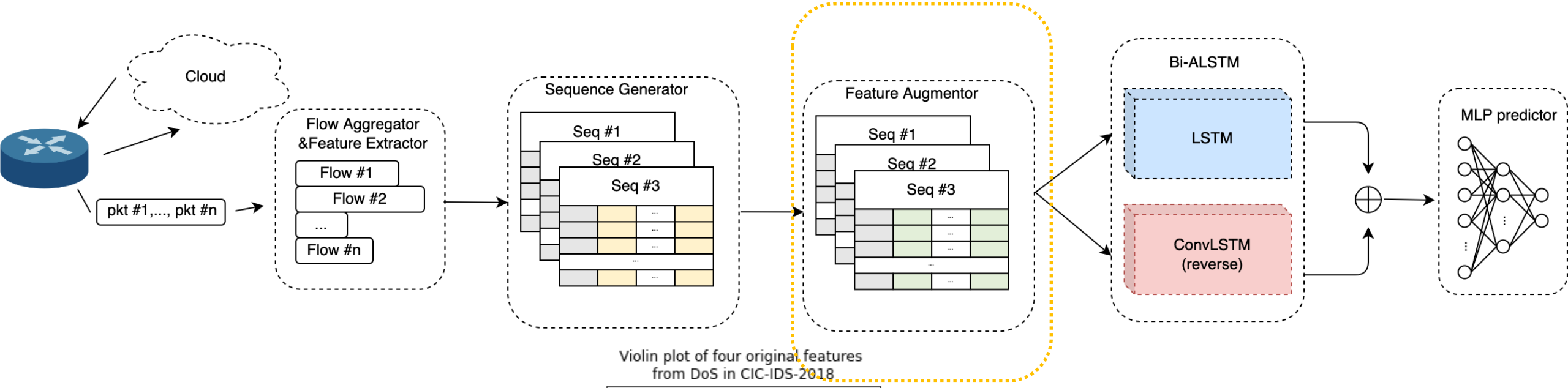
Sequence: Consecutive flows generated by a pair of hosts in a given time interval

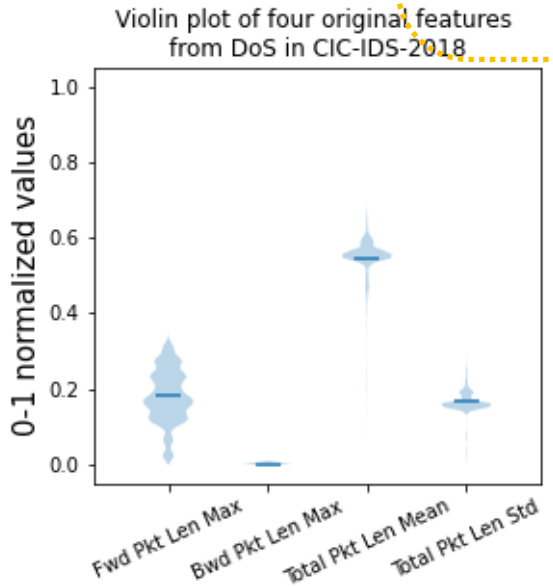
aggregated by: (src ip, dst ip, protocol)

Flexible sequence generation scheme:

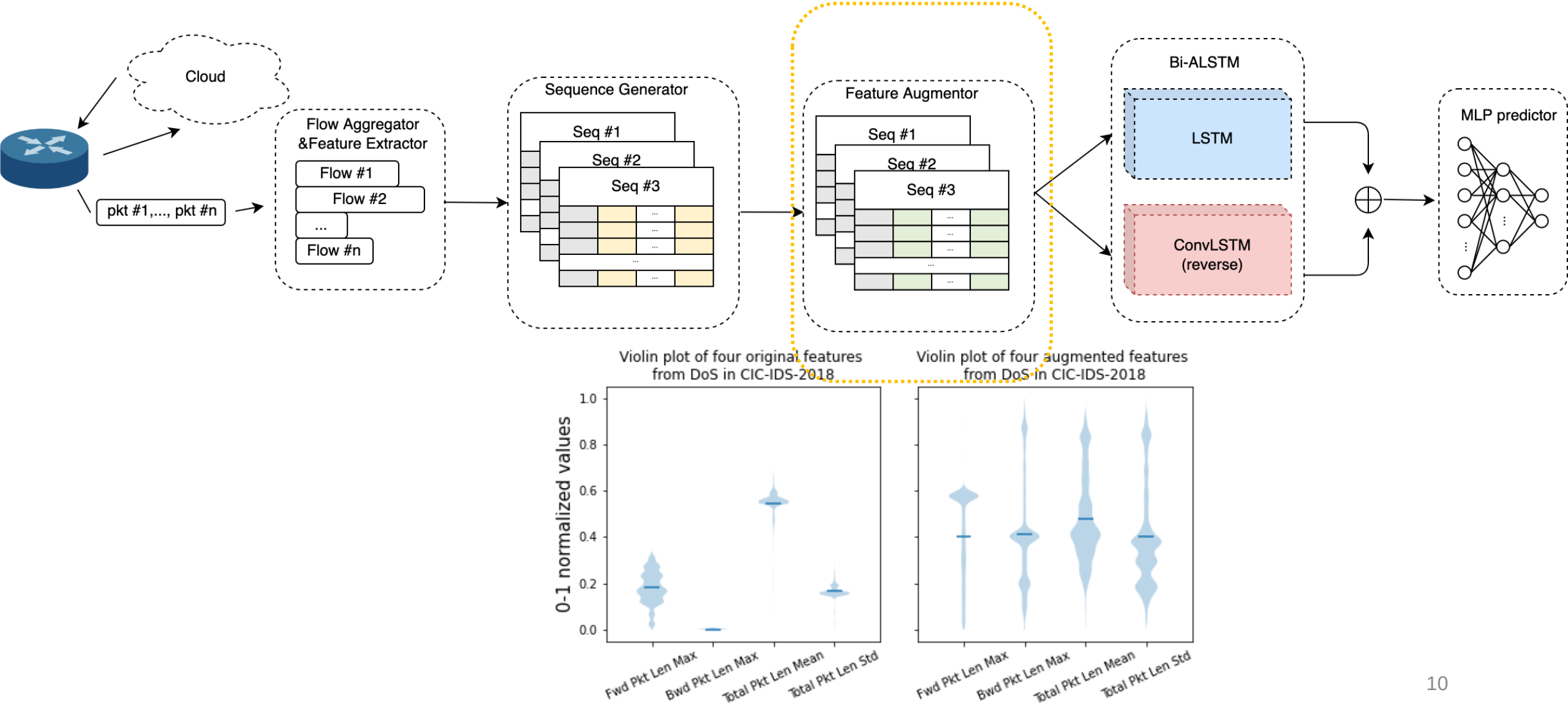
- Sliding window: any sequence reaches a preset length (10) would be passed to subseqeunt module
- Timeout: after a preset time (30s), all sequences would be padded and passed to subsequent module

Why Need Data Augmentation?



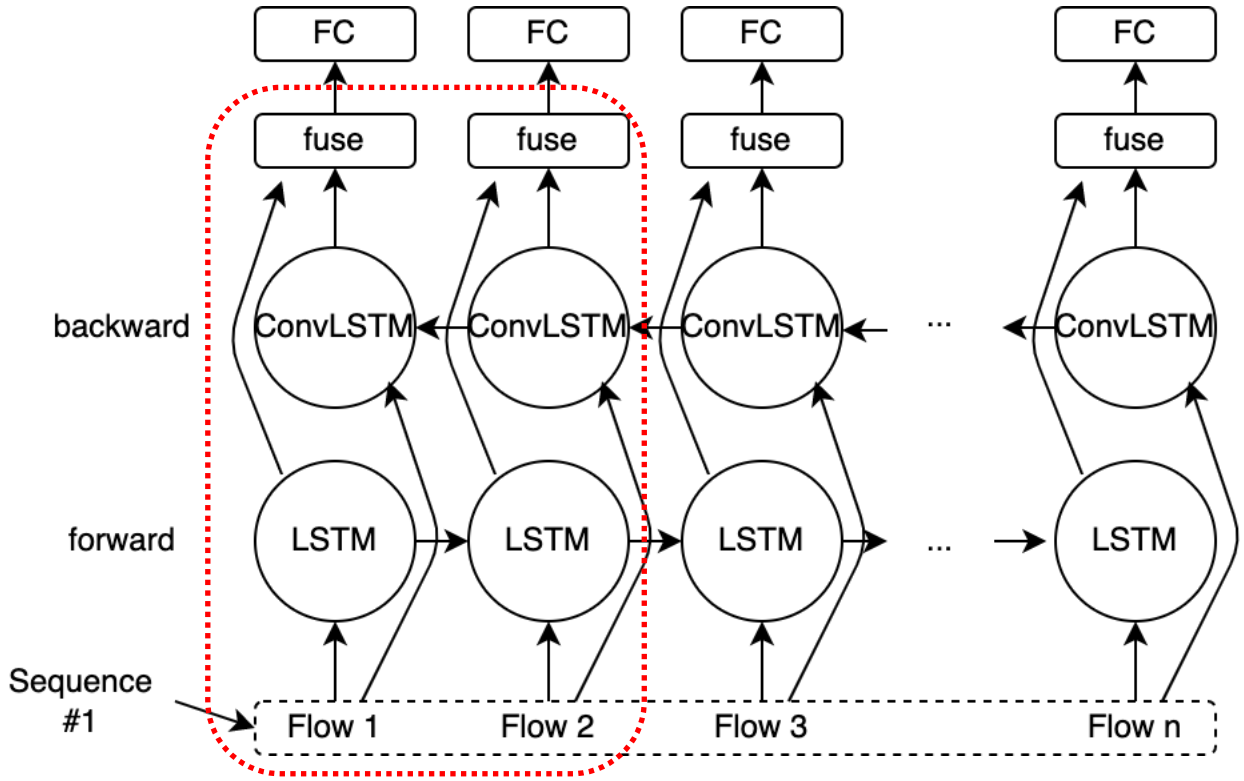


Why Need Data Augmentation?



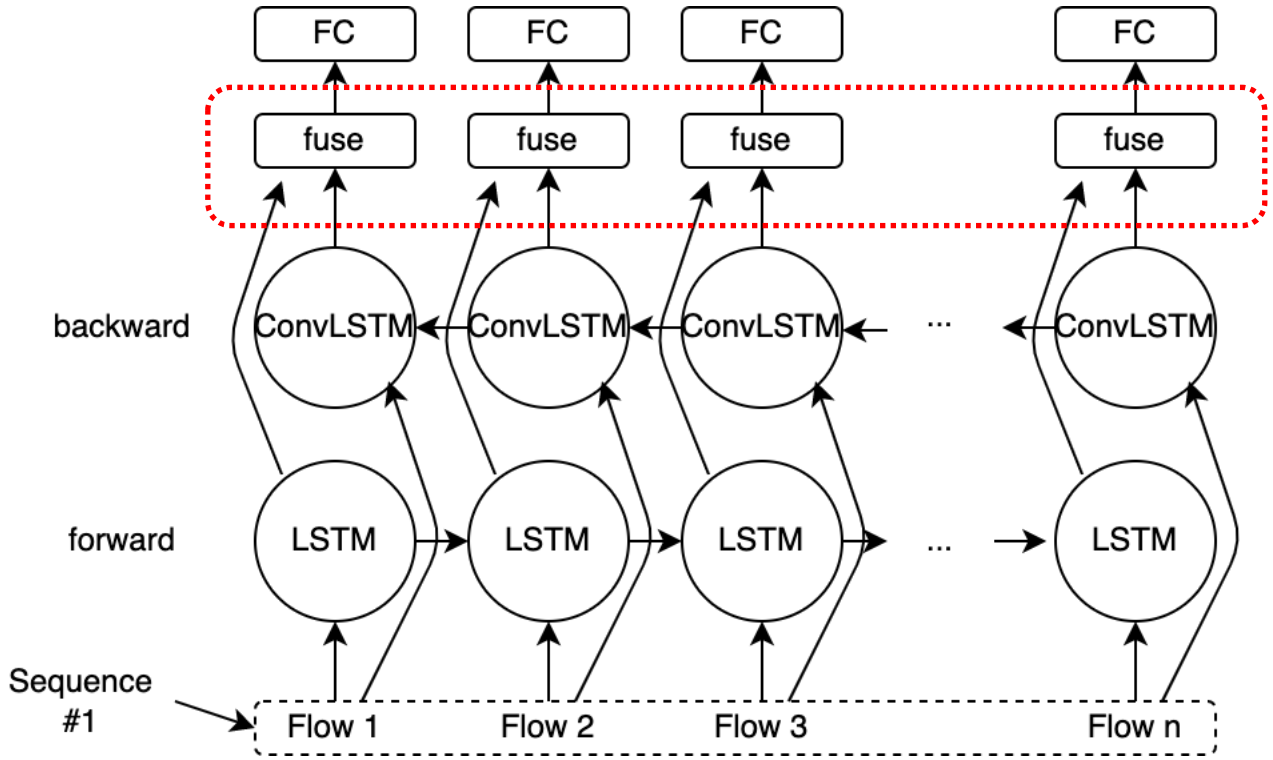
Bidirectional Asymmetric LSTM (Bi-ALSTM)

Why Bidirectional?



Bidirectional Asymmetric LSTM (Bi-ALSTM)

Why Asymmetric?



Evaluations – Bi-ALSTM (without data augmentation)

Algorithms	CSE-CIC-IDS2018 (F1 score)	CIC-IDS-2017 (F1 score)
MLP	0.998	0.544
CNN	0.995	0.696
Autoencoder	0.764	0.428
OC-NN [3]	0.687	0.612
KitNET [1]	0.619	0.401
DAGMM [2]	0.845	0.358
Bi-LSTM	0.998	0.532
CNN-Bi-LSTM	0.998	0.526
Bi-ConvLSTM (ours)	0.997	0.918
Bi-ALSTM(ours)	**0.999**	**0.923**

Evaluations – data augmentation

Algorithms	CSE-CIC-IDS2018 (F1 score)	CIC-IDS-2017 (F1 score)
MLP	0.998	0.544
CNN	0.995	0.696
Autoencoder	0.764	0.428
OC-NN [3]	0.687	0.612
KitNET [1]	0.619	0.401
DAGMM [2]	0.845	0.358
Bi-LSTM	0.998	0.532
CNN-Bi-LSTM	0.999	0.526
Bi-ConvLSTM (ours)	0.997	0.918
Bi-ALSTM(ours)	**0.999**	**0.923**

*: Results in this column are generated by models trained on IDS-2018 with data augmentation

Data Augmentor