**Semi-supervised**

Data set (Keep attacks)

* NSL-KDD:DOS,Probe,R2L,U2R
* UNB ISCX IDS 2012:DDOS and normal
* UNSW\_NB15:9 Attacks

Data Preprocessing -> estimation of network traffic entropy->

online-coclusting -> information gain ratio -> network traffic classification

**Estimate**: use time base sliding window? , To reduce high dimensionally->

One dimension call FSD? , use FSD because zombie host send large number packets

**Online-coclusting**: algorithm split anomalous data in to three cluster. Cluster1 contain only normal traffic, 2 DDOS traffic, 3both.

Simulate with matrix because 3 reasons.

1.Dimensionlity reduction

2.More compressed data

3.Significant reduction

**Network traffic classification**: method detection

* Extra-Trees ensemble classifiers: low compute
* Entire proposed approach: reduce noisy and irrelevant traffic.