sv(M)kmeans - A hybrid feature selection technique for reducing false positives in network anomaly detection

Shubham Saini, Shraey Bhatia, I. Sumaiya Thaseen
Vellore Institute of Technology

Why focus on false positives

- 90% of alerts in an un-tuned IDS are false positives
- High number of false positives dilutes network administrator's concentration on authentic alerts.

Network Anomaly Detection

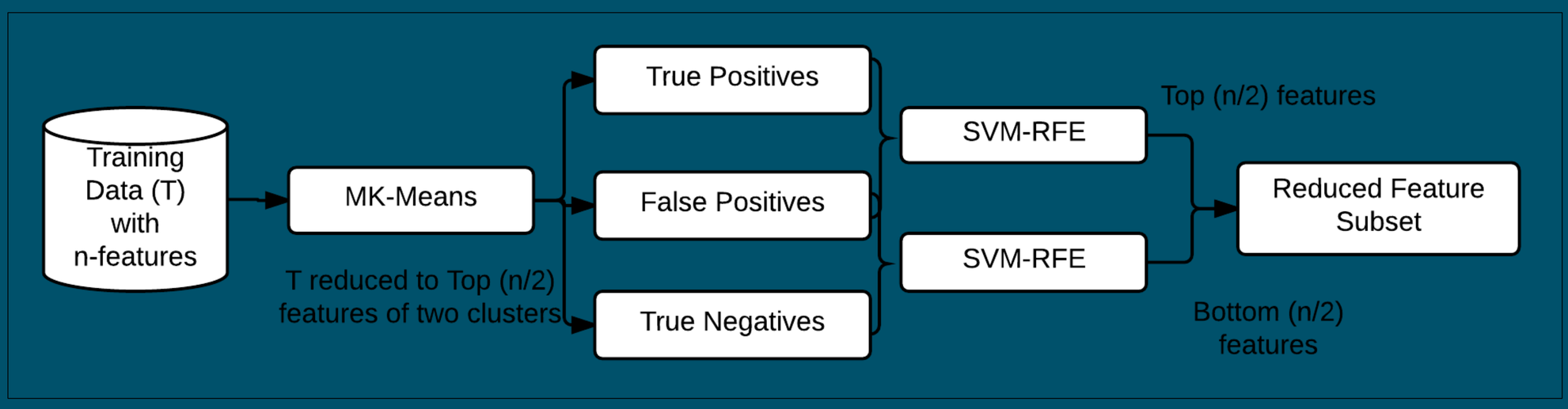
- Detecting network activities with malicious intent or policy violation
- Compare normal system profiles with inbound traffic

Motivation behind sv(M)kmeans

- Simple intuitive idea: group of experts with diverse experience have better chance of arriving at an acceptable solution.
- Reduced feature subset helps in faster execution.

Hybrid clustering + classification approach

MK-Means + SVM-RFE (Wrapper Methods)



Results on NSL-KDD Dataset