Barcelona Neural Networking Center

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Departament d'Arquitectura de Computadors

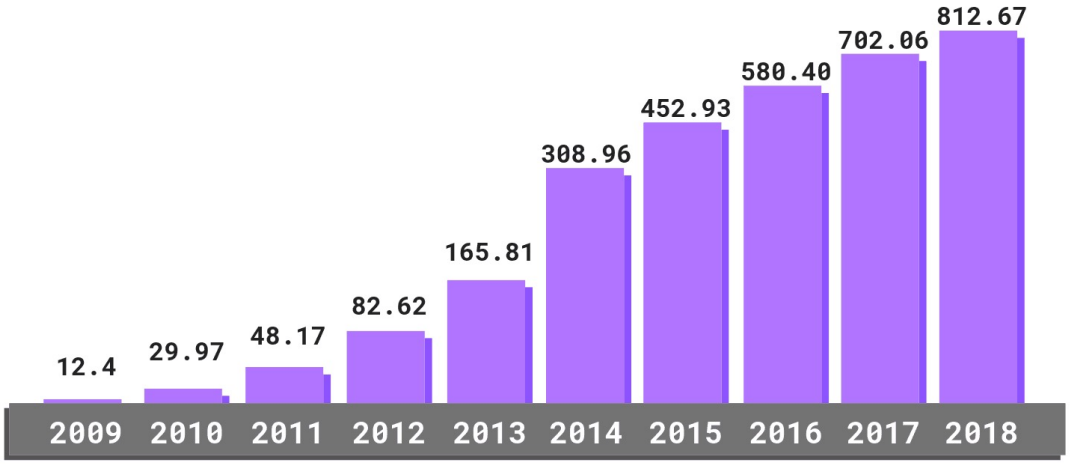
Unveiling the potential of Graph Neural Networks for robust Intrusion Detection

David Pujol-Perich, José Suárez-Varela,
Albert Cabellos-Aparicio, Pere Barlet-Ros

Barcelona Neural Networking Center
Universitat Politècnica de Catalunya (UPC)

Motivation

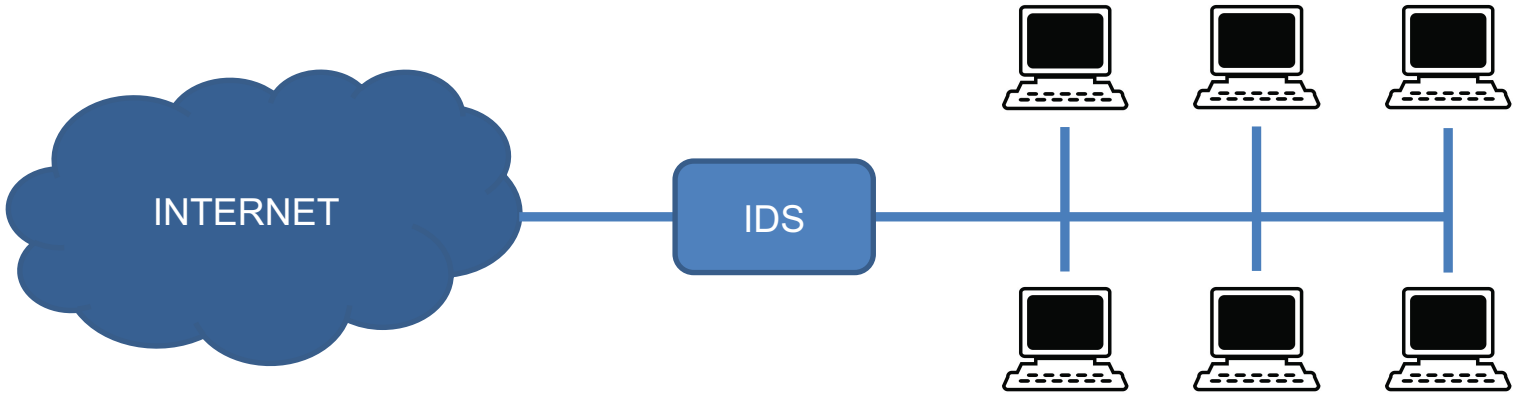
- In recent years the amount of cyberattacks has increased dramatically
- They have a major negative impact in the economy. Concretely, up to $109 billions yearly in the US alone
- They cause frequent data breaches that threatens hundreds of millions of user's privacy



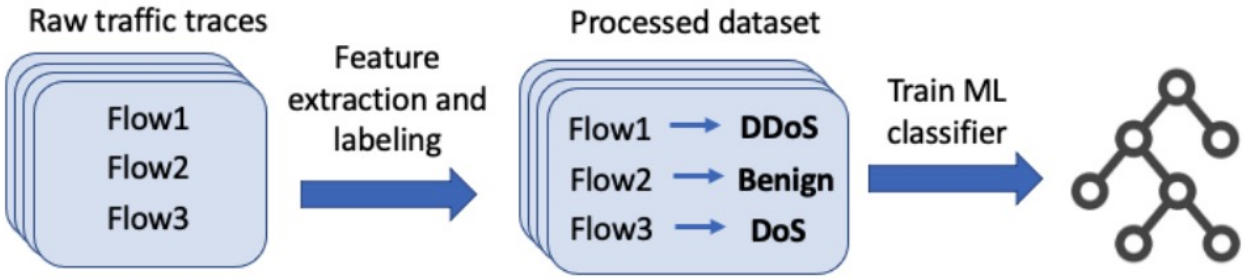
Statistics extracted from [2021 Cyber Security Statistics]

Intrusion detection systems (IDS)

- A great deal of literature has focused on creating Intrusion Detection systems (IDS)
- IDS attempt to **detect** the **malware** traffic meanwhile preserving the benign
- This is thus a classification task of the incoming traffic



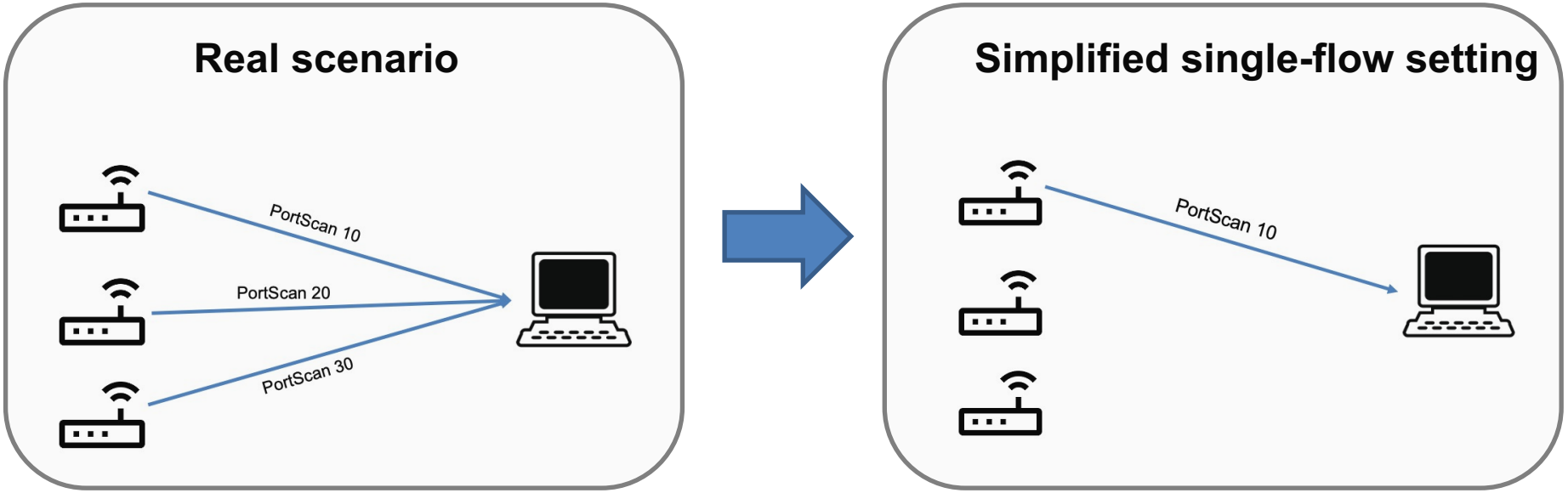
Traditional IDS



- State-of-the-art IDS are often based on traditional ML methods (e.g., MLP, SVM, RF, KNN)
- Many IDS over-simplify the problem by considering each of the flows **independently**
  - These solutions often assume that there are no multi-flow attacks (e.g., DDoS)

Limitations of traditional IDS

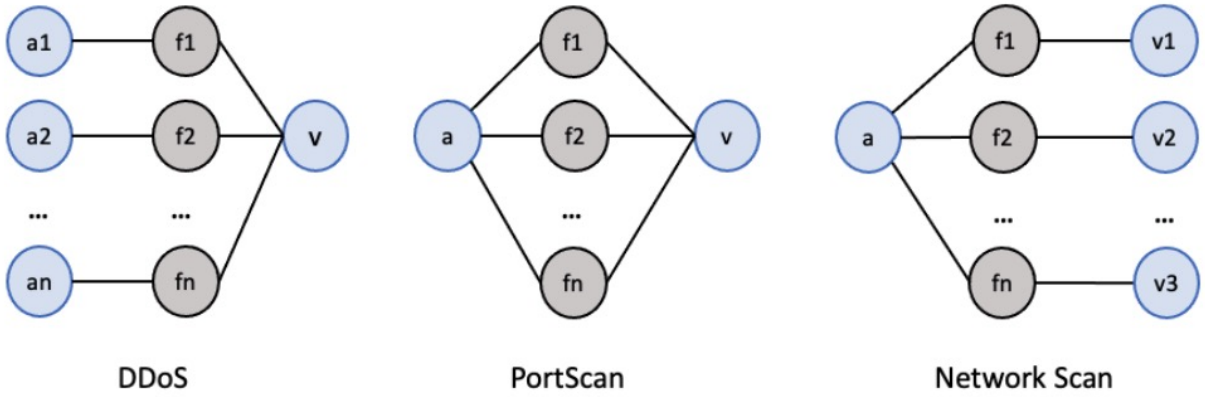
- This assumption is often **not reasonable**



Is it a PortScan **attack** or a **normal traffic**?

Impossible to say

Graph-based IDS



- An IDS may benefit from processing multiple flows simultaniously. Specially for multi-flow attacks (e.g., DDoS)
- To this end, we can model the incoming traffic as a **graph**:
  - Create **attack-characteristic patterns** (different attacks have different graphs)
  - These patterns are **invariant** to specific connection features
  - These are more **robust** to network changes or adversarial attacks

Graph-based IDS

- Traditional ANN can not handle well graph-structured data
- **Graph Neural Networks** adapt perfectly to this setting:
  - They are dynamically assembled based on the input
  - Support different number of nodes and edges
  - Generalize well to unseen graphs
- They have already shown promising results in other fields (e.g., Computer Networks, Power Networks)

F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini. *The graph neural network model*. IEEE Transactions on Neural Networks, 20(1):61–80, 2009

Graph Neural Networks

- Multiple GNN variants
  - Graph Convolutional NN, Gated Graph Sequence NN, Graph Attention Networks, **Message Passing NN**, etc.

- Basic idea
  - Model relational patterns of graph elements (instead of modeling the whole graph)
  - Set of NNs connected according to the input graph

F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini. *The graph neural network model*. IEEE Transactions on Neural Networks, 20(1):61–80, 2009

Message Passing Neural Networks

- Message

$$m_v^{t+1} = \sum_w M_t(h_v^t, h_w^t, e_{vw})$$

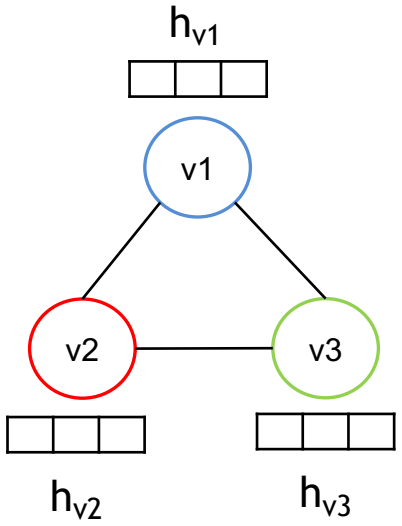
- Update

$$h_v^{t+1} = U_t(h_v^t, m_v^{t+1}), \quad h_v^0 = x_v$$

T times

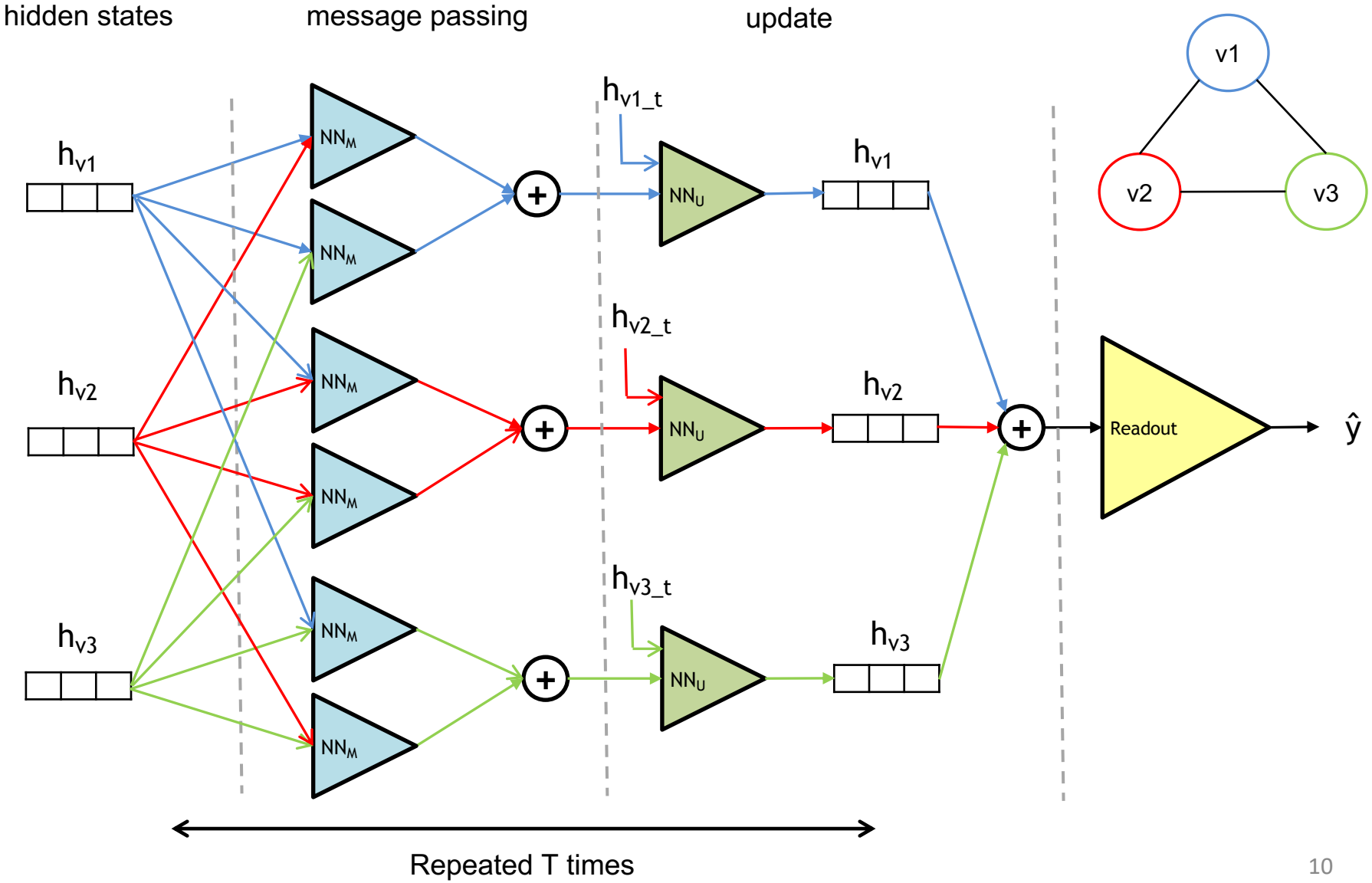
- Readout

$$\hat{y} = R(h), \quad h = \sum_i h_i$$



J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl. *Neural Message Passing for Quantum Chemistry*, Proc. of ICML, Sydney, Australia, Aug. 2017.

Single Message Passing

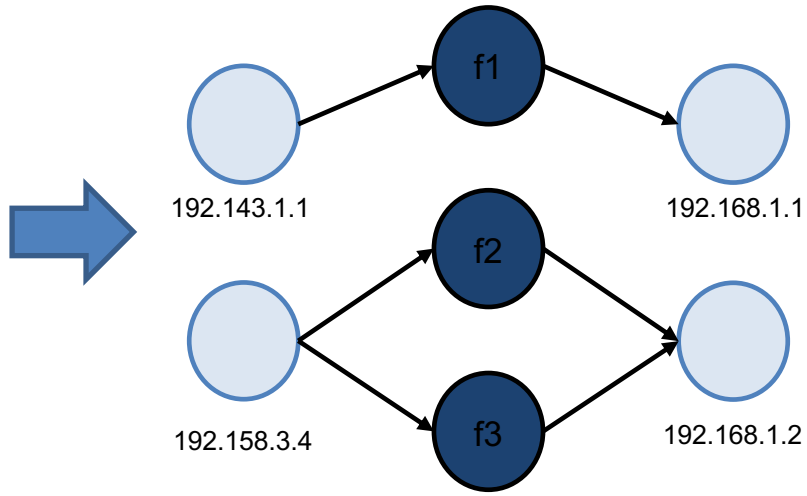


Our proposed GNN-based IDS

Traffic modeling: *Host-Connection graph*

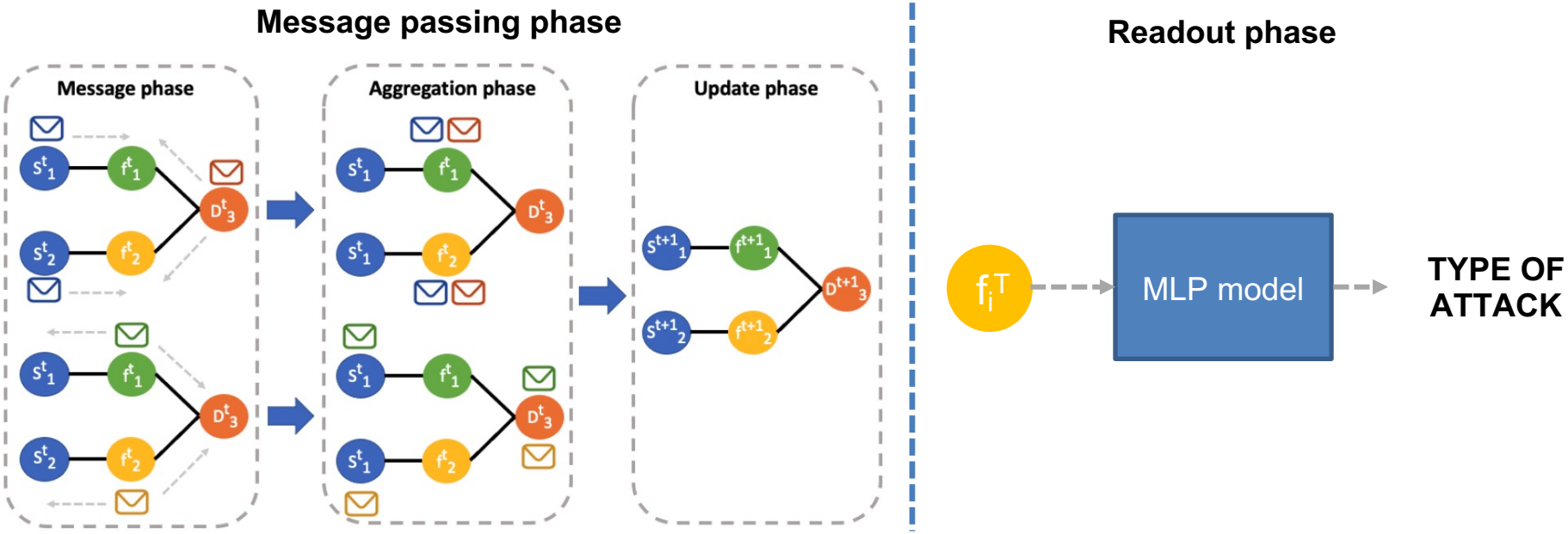
- Connections in a network can be seen as a graph
- We propose the use of the *Host-Connection graph:*
  - Define each of the hosts and each of the connections as a node in the graph
  - We connect them according to the incoming traffic
- This modeling can be practical for many other Cybersecurity problems

Flow	Src IP	Dst IP	Feat 1	…	Feat d
f1	192.143.1.1	192.168.1.1			
f2	192.158.3.4	192.168.1.2			
f3	192.158.3.4	192.168.1.1			



GNN-architecture

- We can train a standard MP architecture that learns by exchanging between the nodes
- Use the resulting embeddings of the flows to make the final prediction



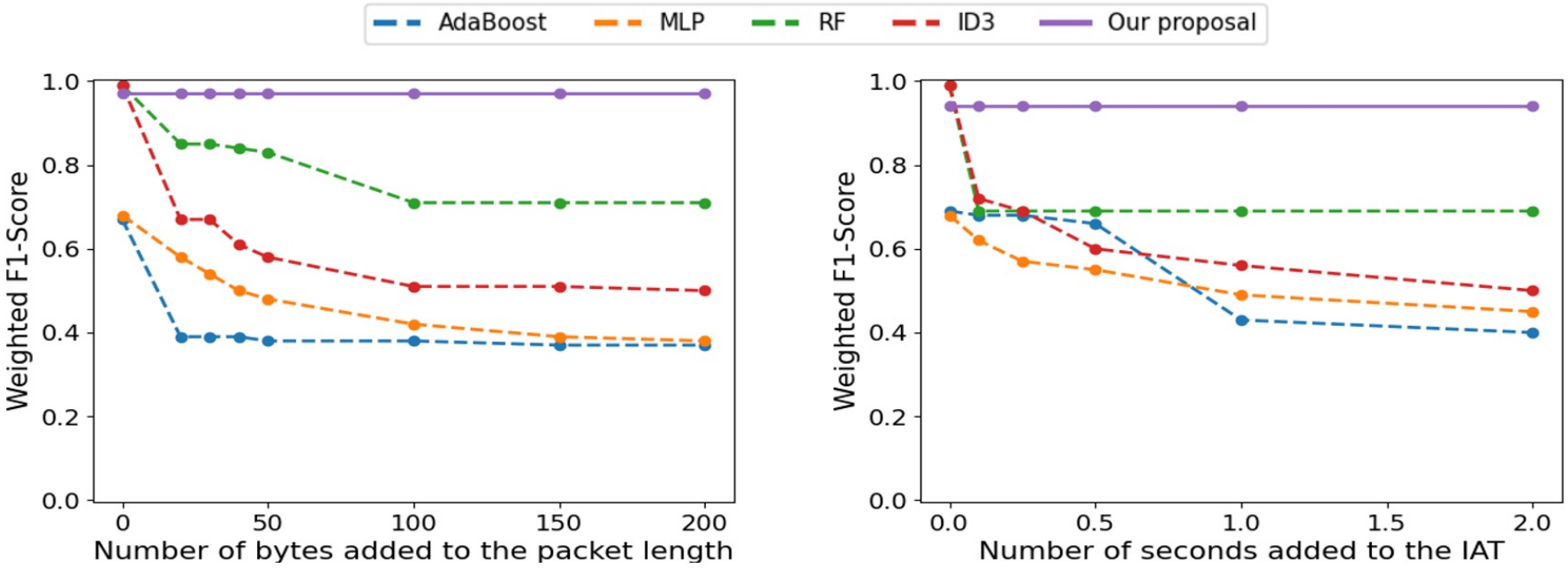
Results

- Results obtained using IDS 2017 dataset

Class label	MLP	AdaBoost	RF	ID3	Our proposal
Benign	0.67	0.68	**0.99**	**0.99**	**0.99**
SSH-Patator	0.0	0.0	**0.99**	**0.99**	0.98
FTP-Patator	0.0	0.0	**0.99**	**0.99**	**0.99**
DoS GoldenEye	0.12	0.0	0.97	0.96	**0.99**
DosHulk	0.63	0.63	**0.99**	**0.99**	**0.99**
DoS slowloris	0.02	0.0	**0.99**	**0.99**	0.98
DoS Slowhttptest	0.01	0.0	**0.98**	**0.98**	0.97
DDoS	0.51	0.0	**0.99**	**0.99**	**0.99**
Web-Brute Force	0.0	0.0	**0.82**	0.76	0.73
Web-XSS	0.0	0.0	0.69	0.65	**0.83**
Bot	0.0	0.0	**0.98**	**0.98**	**0.98**
Port Scan	0.78	0.0	**0.99**	**0.99**	**0.99**

Robustness against adversarial attacks

- We tested the robustness of the models with two simple adversarial attacks:
  - Increase the packet size (bytes)
  - Increase the IAT (second)



Conclusions

- Modeling the incoming traffic as graphs presents critical advantatges with respect to more naive approaches
- Our GNN-based IDS captures estructural properties of the attacks, instead of focusing only on their features
- Our proposed IDS:
  - Performs similarly to the SOTA solutions
  - It is much more robust than SOTA solutions to adversarial attacks.

Interesting resources…

- IGNNITION is a framework for fast prototyping of GNNs
  - Provides support of cybersecurity applications and more (e.g., networking)
  - Develop your GNN without using TensorFlow/PyTorch…

https://ignnition.net

- We have organized two GNN Challenges in 2020 and 2021 with hundreds of participants along with ITU.

https://bnn.upc.edu/challenge/gnnet2021/