Soluzioni di Deep Learning per la Cyber Security

Francesco La Rosa

Università degli Studi di Messina

francesco.larosa@unime.it

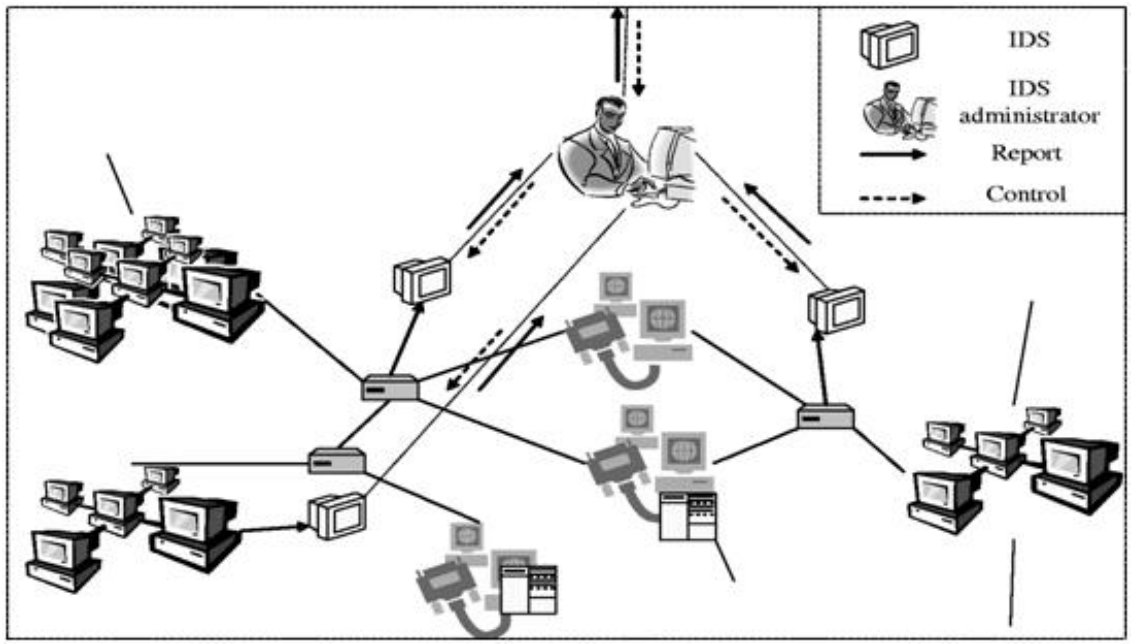
Conferenza GARR 2018 - *Data Revolution*

Cagliari, 03 Ottobre 2018

Agenda

Intrusion Detection System

- L'**identificazione** dei vari **attacchi di rete**, in particolare di quelli mai visti in precedenza, è un problema da risolvere con **urgenza**.
- Un componente fondamentale di un'infrastruttura di cyber security è il sistema di Intrusion Detection (**IDS**)
- Possono essere classificati in due categorie:
  - **Host**-based IDS (HIDS)
  - **Network**-based IDS (NIDS)

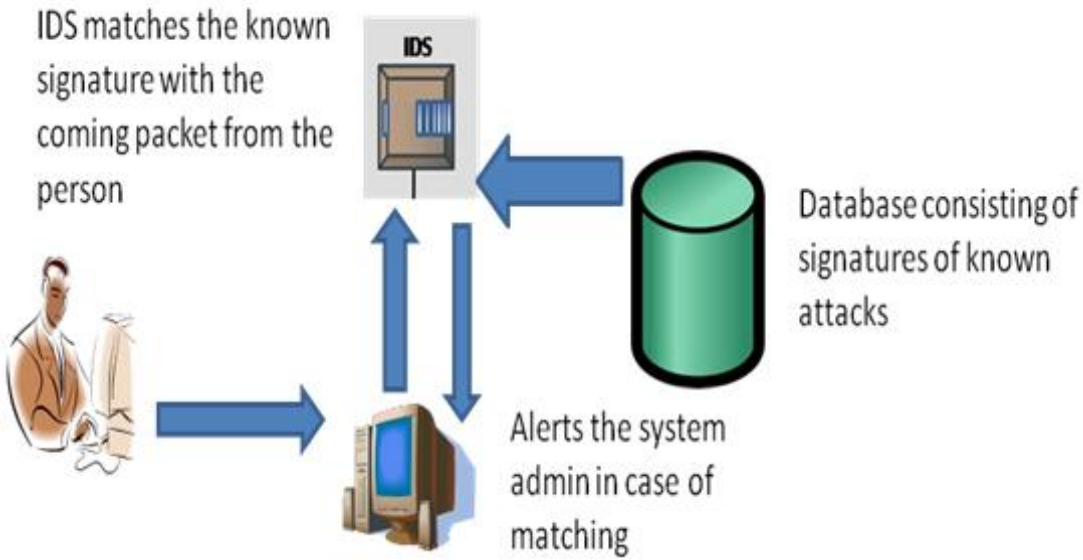


Network Intrusion Detection System

- I Network Intrusion Detection System (NIDS) esistenti possono essere **classificati** in base al tipo di **analisi del traffico di rete** realizzato:
  - Misuse-based
  - Anomaly-based
  - Hybrid

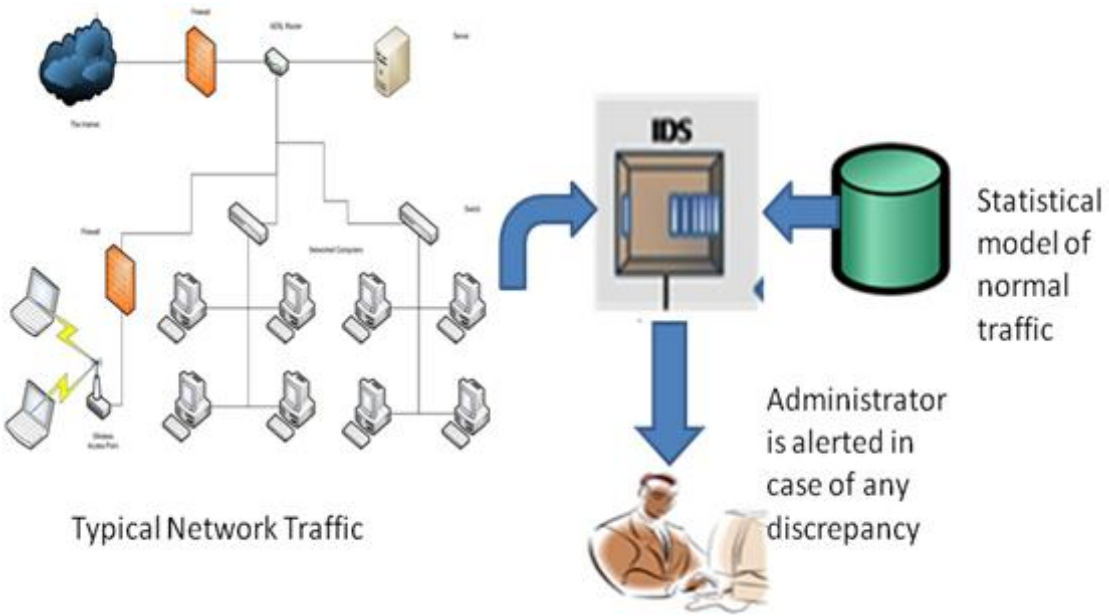
NIDS misuse-based

- Nota anche come detection basata su firma, il rilevamento dell'abuso è basato sul confronto di eventi registrati con schemi predefiniti di attacco.
- Costituisce l'approccio più utilizzato
- Questo sistema di rilevamento delle intrusioni è lento nel generare un modello e presenta un alto costo in caso di cambiamento dello stesso, rendendo difficile rilevare efficacemente nuovi tipi di attacco emergenti su Internet.



NIDS anomaly-based

- Le soluzioni anomaly-based si basano su un **modello** del normale comportamento della rete
- Identificano le anomalie come **deviazioni** dal modello
- Sono interessanti per la loro capacità di rilevare attacchi zero-day
- **Elevato tasso** di falsi allarmi
  - **Comportamenti di sistema** precedentemente non riscontrati possono essere classificati come **anomalie**



Intrusion Detection e Machine Learning

- Diverse sono le **soluzioni proposte** in cui tecniche di **Machine Learning** (ML) sono state applicate al problema dell'intrusion detection [Ford et al., 2014][Buczak et al., 2016]:
  - Support vector machine
  - Algoritmi genetici
  - Fuzzy logic
  - Reti bayesiane
- Il ML si concentra principalmente sulla **classificazione e la discriminazione** in base a caratteristiche note precedentemente apprese per mezzo di dati raccolti per l'addestramento (**learning**)

Machine Learning e Deep Learning

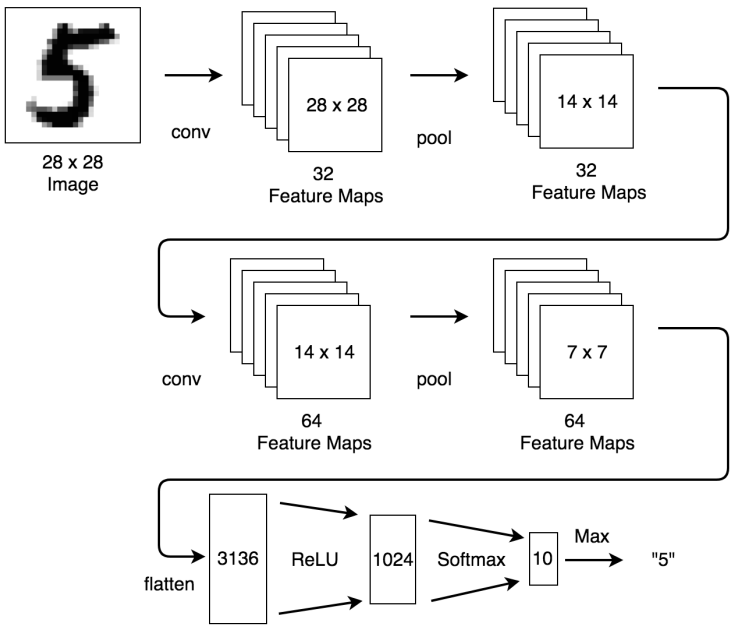
- Il **Deep Learning** (DL) è un campo di ricerca che ricade nell'ambito del **ML**
- La sua motivazione sta nella creazione di una **rete neurale** che simula il comportamento del cervello umano nell'apprendimento analitico
- Analogamente ai metodi di ML, i metodi di DL si possono classificare in **metodi di apprendimento** supervisionato e apprendimento non supervisionato

Deep Learning

- Il DL rende possibile un apprendimento non supervisionato o semi-supervisionato che consente **l'estrazione automatica (ed efficiente) di feature**
- **Algoritmi di DL** comunemente utilizzati:
  - Autoencoders
  - Deep Belief Networks (DBMs)
  - Long Short-Term Memory networks (LSTMs)
  - Convolutional Neural Networks (CNNs)
    [Jia et al., 2017][Wang et al., 2018]

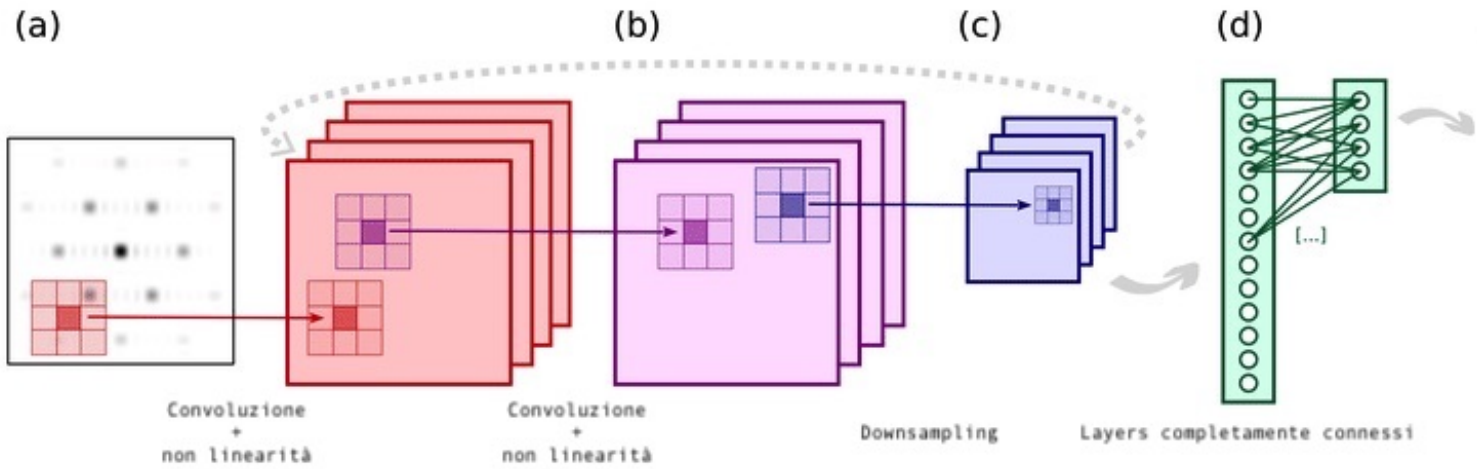
Convolutional Neural Networks - parte 1

- Le Convolutional Neural Networks (CNNs) sono un tipo specializzato di rete neurale adatta al processamento di dati sotto forma di matrice, come serie temporali e **immagini**
- **Una serie temporale** può essere pensata come un vettore monodimensionale **campionato** ad intervalli regolari
- La tipica architettura di una CNN consiste di un **layer di input ed uno di output** completati da diversi **layer nascosti**



Convolutional Neural Networks - parte 2

- I **layer nascosti** possono essere dei:
  - Convolutional layer
  - Pooling layer
  - Layer completamenti connessi



Il classificatore - 1/3

- Classificatore implementato usando **Tensorflow**
- Tutti gli esperimenti sono stati condotti su un server:
  - S.O.: Ubuntu 64-bit 16.04 LTS
  - CPU: Intel Xeon 3.60GHz
  - RAM: 16 GB
  - **GPU NVIDIA GTX 750**

Il classificatore - 2/3

- Convolutional Neural Network
  - **Canali**: 24
  - **Kernel** size: 3*3
  - **Step**: 1
  - Dati raw acquisiti dalla rete di dimensione fissa pari a **28*28** elementi (784 bytes)

Il classificatore - 3/3

- Convolutional Neural Network
  - Convolutional layer
  - **Layer di pooling**: 2
    - Max-pooling
    - Size: 2*2
  - **Rete** completamente connessa: *3 strati nascosti e 5 uscite*
  - Adam **optimizer**:
    - Batch: 60 esempi
    - Funzione di costo: cross-entropy
    - Training time: 0.002
    - Epoche: 50

Misure - Dataset

- Dataset su cui fare benchmark:
  - **DARPA98**
  - KDDCUP99
  - NSLKDD
- Sono dataset ormai datati che seppur non rappresentino in maniera fedele il traffico e i possibili attacchi di una rete moderna, permettono il **confronto tra la soluzione proposta e altre presenti in letteratura**.

Misure - Parametri utilizzati

Accuratezza

$$ACC = \frac{TP+TN}{TP+FP+FN+TN}$$

Detection Rate

$$DR = \frac{TP}{TP+FN}$$

False Acceptance Rate

$$FAR = \frac{FP}{FP+TN}$$

- **TP**: true positive
- **TN**: true negative
- **FP**: false positive
- **FN**: false negative

- **ACC**: Accuracy
- **DR**: Detection Rate
- **FAR**: False Acceptance Rate

Misure - Categorie degli attacchi

Categoria	**Pattern di attacco**
DoS	back, land, neptune, pod, smurf, teardrop
Probe	ipsweep, nmap, portsweep, satan
R2L	ftp-write, guess-passwd, imap, multihop, phf, spy, etc.
U2R	buffer-overflow, loadmodeule, perl, rootkit

Table: Categorie degli attacchi

Misure - Prestazioni del classificatore su DARPA98

Dataset	**Acc%**	**DR%**	**FAR**
DoS	99.62	99.23	0.03
Probe	99.30	83.43	0.02
R2L	99.75	75.1	0.03
U2R	99.98	67.2	0.03
Totale	99.72	97.82	0.08

Table: Le prestazioni del NIDS sul dataset DARPA98

Conclusioni

- **Network Intrusion Detection System** basato su una *Convolutional Neural Network*, già adottata con risultati notevoli in ambiti come quelli della Computer Vision e del Natural Language Processing
- **Idea** di base: "convertire la distribuzione di probabilità insita nel flusso di dati acquisito in una sequenza di *immagini* ed impiegare una CNN per classificarle in base al loro contenuto (traffico di rete)"
- **Sviluppi futuri**:
  - Migliorare le prestazioni della rete (CNN) su **dataset sbilanciati**
  - Migliorare le prestazione del NIDS proposto **combinando** i dati (**raw**) raccolti con **feature tradizionali**

References

F. Jia and L. Kong (2017)

Intrusion detection algorithm based on convolutional neural network

Trans. Beijing Inst. Technol. 37:1271–1275, 12.

V. Ford and A. Siraj (2014)

Applications of machine learning in cyber security

Proceedings of the 27th International Conference on Computer Applications in Industry and Engineering

A. Buczak and E. Guven (2016)

A survey of data mining and machine learning methods for cyber security intrusion detection

IEEE Communications Surveys Tutorials 18(2):1153–1176.

W. Wang, Y. Sheng, J. Wang, X. Zeng, X. Ye, Y. Huang and M. Zhu (2018)

Hast-ids: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection

IEEE Access 6:1792–1806.

Soluzioni di Deep Learning per la Cyber Security

Nicola Cannistrà, Fabio Cordaro, Francesco La Rosa,
Umberto Ruggeri e Riccardo Uccello

Università degli Studi di Messina

Conferenza GARR 2018 - *Data Revolution*

Cagliari, 03 Ottobre 2018

Grazie