



*6th Nov 2020*

**Faculty 1: Chair of IT-Security | Master Thesis: Initial Talk**

---

# **Unsupervised feature extraction from network traffic for content-based anomaly detection in industrial networks**

*B. Sc. Fabian Kopp*

## **Supervisors:**

*M. Sc. Franka Schuster*

*Prof. Dr.-Ing. Andriy Panchenko*

---



- 
- 1 | **Background**
  - 2 | **Related Work**
  - 3 | **Methodology**
  - 4 | **Schedule**
-

- ➔ Cyber attacks like: *industroyer*, *blackengery*, *havex* or *stuxnet* are real threats
- ➔ Industrial field devices are *controlled and monitored* over computer networks
- ➔ Network based anomaly detection is one option to detect *0-day attacks*
- ➔ Feature extraction from network data is the *basis* for all detection algorithms
- ➔ *Unsupervised* machine learning methods can adapt to network protocols automatically

## Network Data

- Series of *packets*
- Packets consist of *layers*
- Layers have *header* and a *payload* field(s)
- Weak order of packets

## Industrial Control Systems (ICS)

- Monitoring & Control functions
- Fixed network topology
- Proprietary Infrastructure

## Intrusion Detection Systems (IDS)

- Used to secure networks
- Host deployment
  - Stack traces, log files, etc
- Network deployment
  - Raw data, aggregated traffic

00000260:	54	50	2f	31	75	31	0d	03	48	6f	73	74	3a	20	78	6e	TP/1.1·Host: xn
00000270:	2d	2d	6d	62	69	75	73	2d	6a	75	61	2e	62	61	6e	64	--mbius-jua.band
00000280:	0d	03	55	73	65	72	2d	61	67	65	6e	74	3a	20	4d	6f	·User-Agent: Mo
00000290:	7a	69	6c	6c	61	2f	35	2e	30	20	28	58	31	31	3b	20	zilla/5.0 (X11;
000002a0:	55	62	75	6e	74	75	3b	2e	4c	69	6e	75	78	28	78	38	Ubuntu; Linux x8
000002b0:	36	51	36	34	3b	20	72	76	3a	38	31	2e	30	31	20	47	6_64; rv:81.0) G
000002c0:	65	63	6b	6f	2f	32	30	31	30	30	31	30	31	20	46	69	ecko/20100101 Fi
000002d0:	72	65	66	6f	78	2f	38	31	2e	30	0d	0a	41	63	63	65	refox/81.0·Acce
000002e0:	70	74	3a	20	74	65	78	74	2f	68	74	6d	6c	2c	61	70	pt: text/html,ap
000002f0:	70	6c	69	63	61	74	69	6f	6e	2f	78	68	74	6d	6c	2b	plication/xhtml+
00000300:	78	6d	6c	2c	61	70	70	6c	69	63	61	74	69	6f	6e	2f	xml,application/
00000310:	78	6d	6c	3b	71	30	2e	39	2c	69	6d	61	67	65	2f		xml;q=0.9,image/
00000320:	77	65	62	70	2c	2a	2f	20	3b	71	30	2e	38	0d	0a		webp,*/*;q=0.8·
00000330:	41	63	63	65	70	74	2d	4c	61	6e	67	75	61	67	65	3a	Accept-Language:
00000340:	20	65	6e	2d	47	2c	65	6e	3b	71	30	2e	35	0d			en-GB,en;q=0.5·
00000350:	0a	41	63	63	65	70	74	2d	45	6e	63	6f	64	69	6e	67	·Accept-Encoding
00000360:	3a	20	67	7a	69	70	2c	20	64	65	66	6c	61	74	65	0d	: gzip, deflate·
00000370:	0a	44	4e	54	3a	20	31	0d	0a	43	6f	6e	6e	65	63	74	·DNT: 1·Connect
00000380:	69	6f	6e	3a	20	6b	65	65	70	2d	61	6c	69	76	65	0d	ion: keep-alive·
00000390:	0a	55	70	67	72	61	64	65	2d	49	6e	73	65	63	75	72	·Upgrade-Insecur
000003a0:	65	2d	52	65	71	75	65	73	74	73	3a	20	31	0d	0a	49	e-Requests: 1·I
000003b0:	66	2d	4d	6f	64	69	66	69	65	6a	2d	53	69	6e	63	65	f-Modified-Since
000003c0:	3a	20	46	72	69	2c	20	31	31	20	53	65	76	20	32	30	: Fri, 11 Sep 20
000003d0:	32	30	20	30	38	3a	34	32	3a	32	37	20	47	4d	54	0d	20 08:42:27 GMT·
000003e0:	0a	49	66	2d	4e	6f	6e	65	2d	4d	61	74	63	68	3a	20	·If-None-Match:
000003f0:	57	21	22	35	63	61	2d	35	61	66	30	35	61	61	32	37	W/"5ca-5af05aa27
00000400:	36	34	39	39	22	0d	0a	43	61	63	68	65	2d	43	6f	6e	6499"·Cache-Con
00000410:	74	72	6f	6c	3a	20	6d	61	78	2d	61	67	65	30	30	0d	trol: max-age=0·
00000420:	0a	0d	0a	cf	7f	90	51	10	1b	08	00	00	00	00	00	00	....._.....B...B

## Anomaly detection (AD)

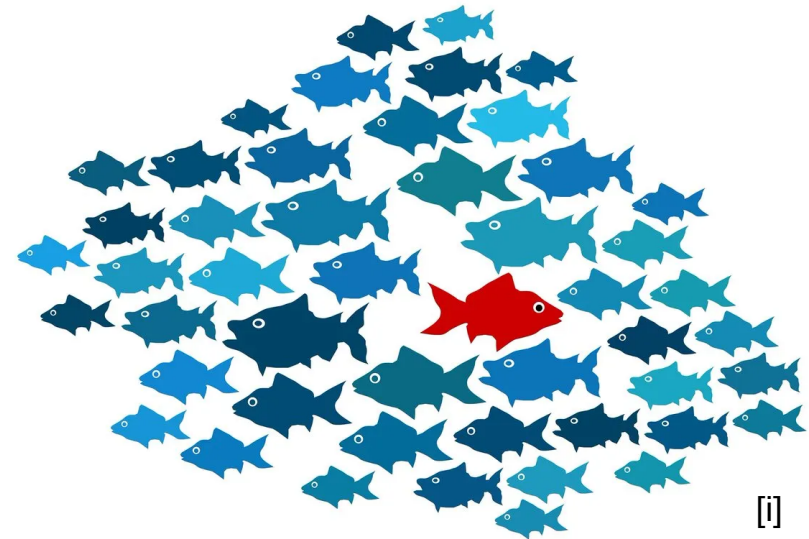
- Unsupervised *binary classification*
- *Learn* normal behavior - alert deviations
- Used in: fraud detection, intrusion detection

## Types of anomalies [1]

- Point → single payload (XSS)
- Collective → multiple payloads (Scan)
- Contextual → order of payloads (0-day)

## Content based AD

- Detects intrusion on the byte level
- Payload and header information as basis



[1]

## Feature extraction

- *Transformation* from input vectors to feature vectors

## Feature selection

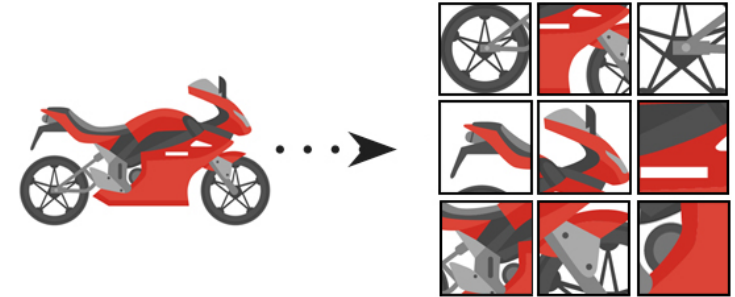
- Determine best subset of features

## Feature engineering

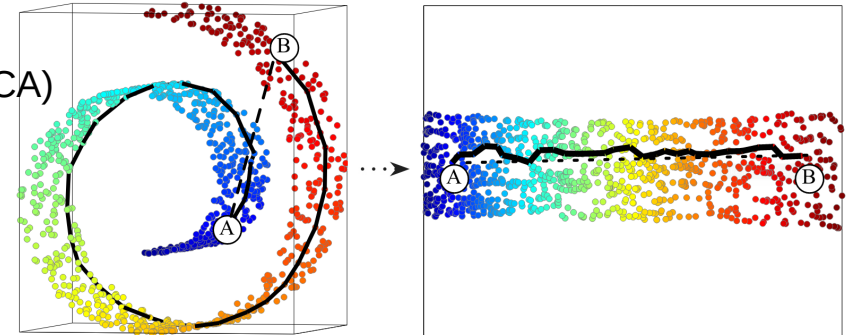
- Consultation of *domain experts*

## Representation learning

- Search for *model* which embeds feature vectors into subset
  - Linear methods
    - Principal component analysis (PCA)
  - Non-Linear methods
    - Kernel-PCA
    - Autoencoder



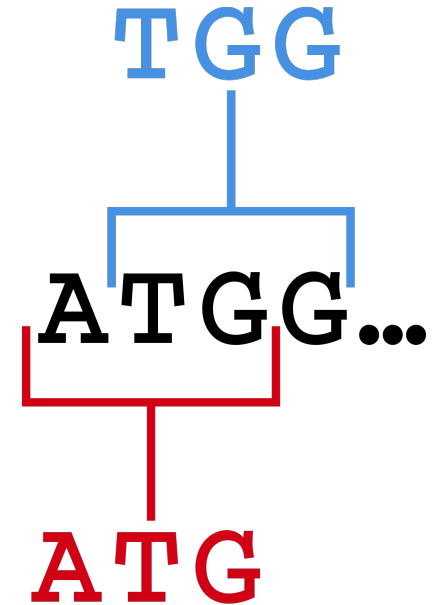
[ii]



[iii]

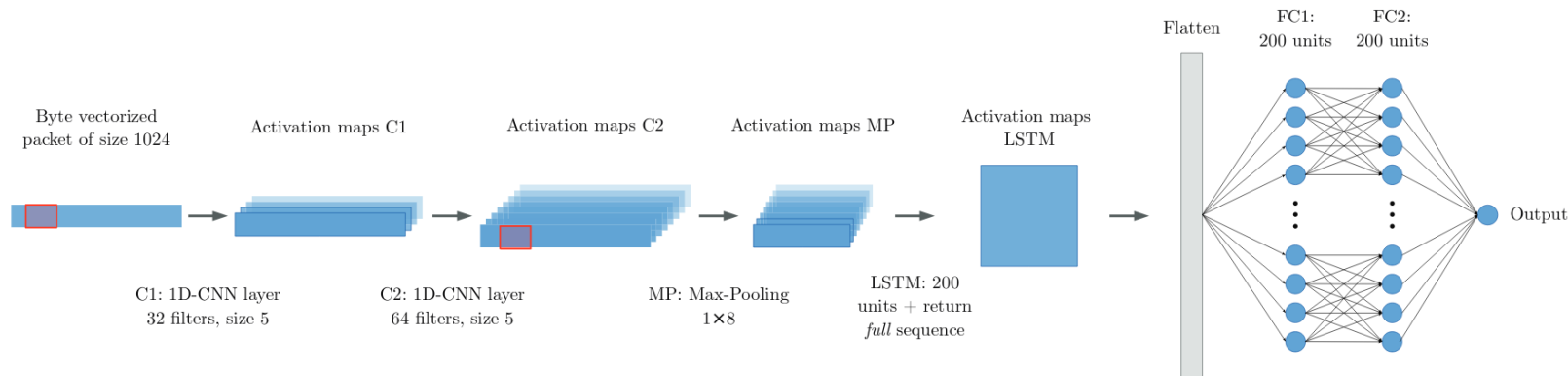
**[2]** ZOE: Content-based Anomaly Detection for Industrial Control Systems (*IEEE/IFIP DSN, 2018*)

- *N-Gram* based feature extraction on application layer payloads
- Prototypical representations specific to individual types of messages
- Filtering rare features using a frequency threshold
- Cluster similarity based intrusion detection
- Takeaways:
  - + Evaluated on ICS related protocols
  - + Unsupervised feature extraction
  - No sequential aspects are considered



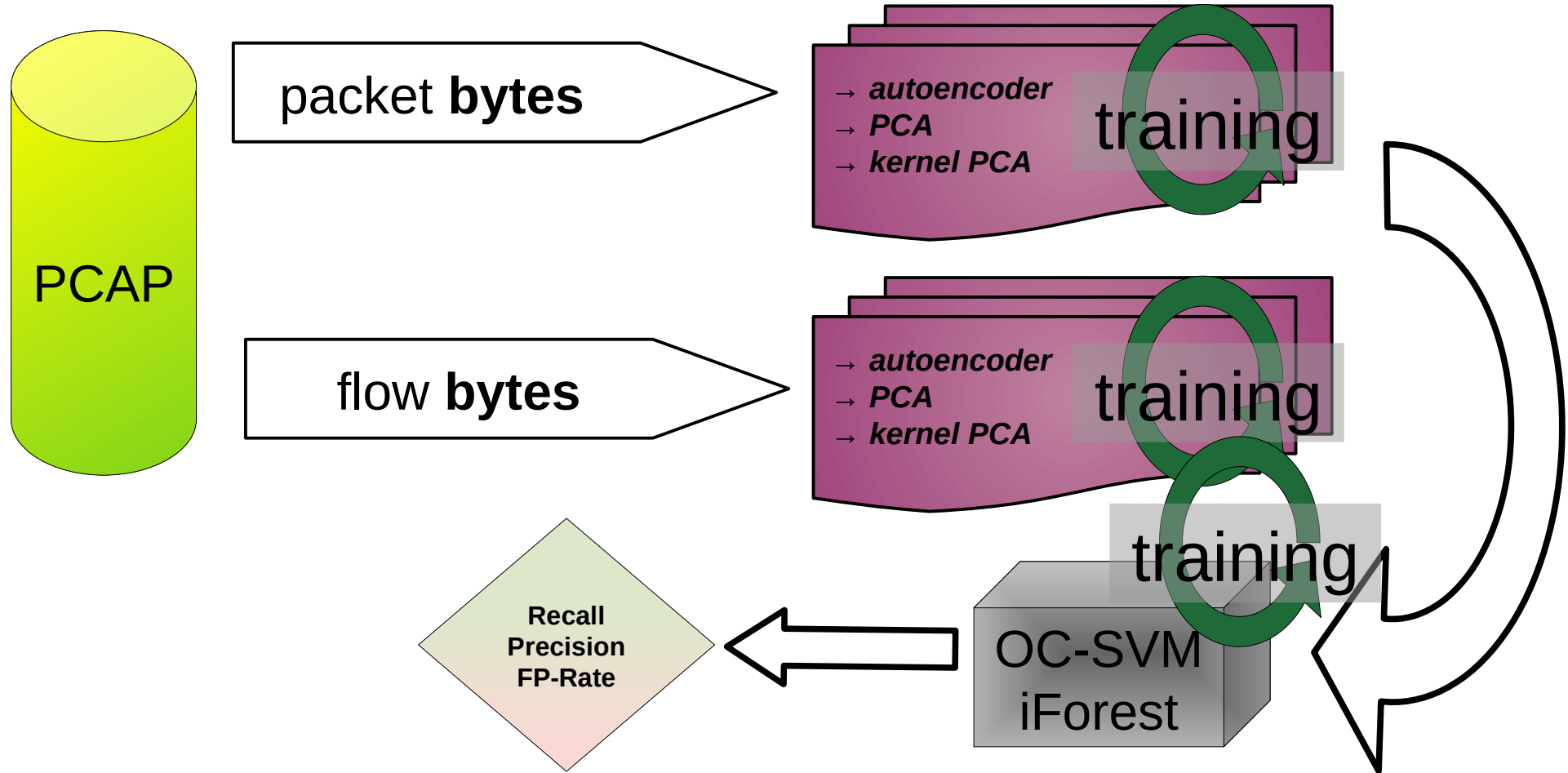
### [3] Deep in the Dark - Deep Learning-based Malware Traffic Detection without Expert Knowledge (*IEEE SPW, 2019*)

- Feature extraction on first  $N$  bytes of every packet / flow
- MAC & IP addresses are sanitized
- Softmax based classification
- Takeaways:
  - End-to-end model
  - Supervised learning
  - Evaluation on ICS unrelated data
  - + Spatial-temporal representation learning on raw traffic



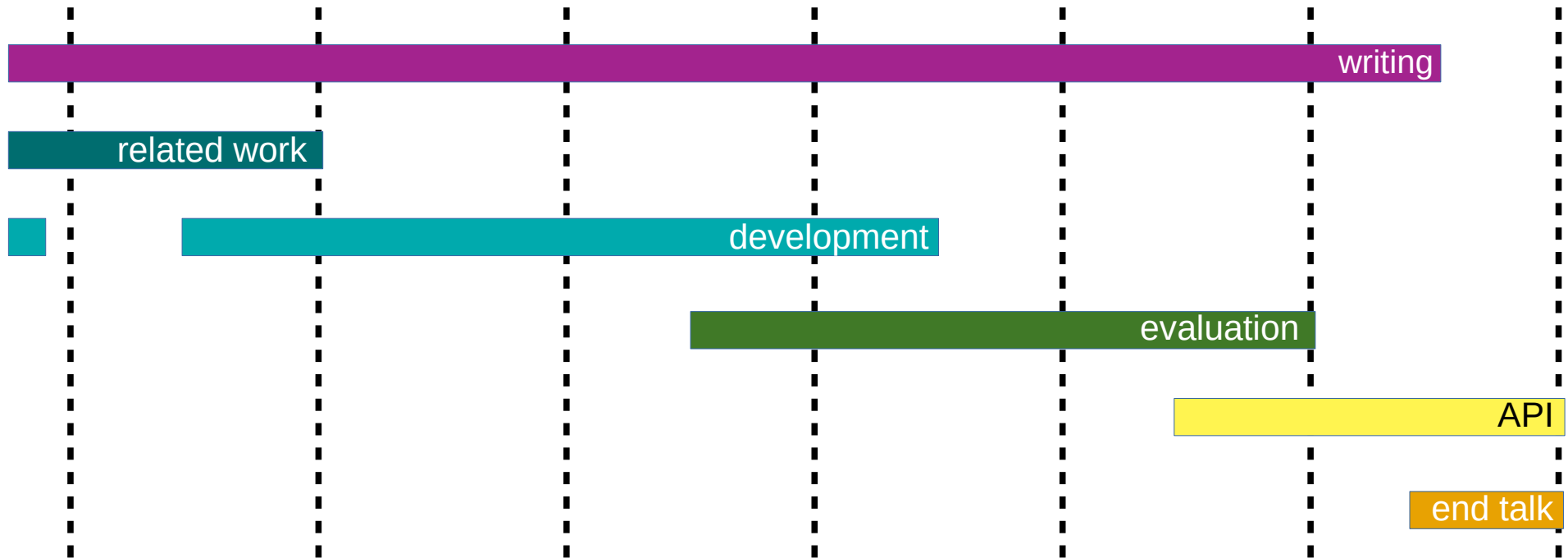


- ➔ **Can AD based on unsupervised feature extraction compete with manual feature extraction?**
  - *Point* anomalies
  - *Contextual* anomalies
  
- ➔ **What kind of byte representation will yield the best results?**
  - *PCAP bytes*
  - *Packet bytes*
  - *Flow bytes*
  
- ➔ **Is the approach *fast* enough ?**
  - Intrusion detection is a real time problem
  
- ➔ **Is the approach capable of extracting relevant features when analyzing variable traffic?**



*initial talk*

*submission*



**Thanks for the attention!**  
*Questions?*

## Sources

**[1]** Chandola, et al. *Anomaly detection: A survey* **ACM computing surveys**, 2009.

**[2]** Wressnegger C., Ansgar K. & Konrad R.

*Zoe: Content-based anomaly detection for industrial control systems.*

IEEE/IFIP: International Conference on Dependable Systems and Networks (**DSN**). **IEEE**, 2018.

**[3]** Marín G., Casas P. & Capdehourat, G.

*Deep in the Dark-Deep Learning-Based Malware Traffic Detection Without Expert Knowledge.*

IEEE: Security and Privacy Workshops (**SPW**). **IEEE**, 2019.

**[i]** [https://i2.wp.com/thedatascientist.com/wp-content/uploads/2019/02/anomaly\\_detection.png](https://i2.wp.com/thedatascientist.com/wp-content/uploads/2019/02/anomaly_detection.png)

**[ii]** [https://miro.medium.com/max/1000/0\\*sQzmiOf8Yb\\_18HX1.png](https://miro.medium.com/max/1000/0*sQzmiOf8Yb_18HX1.png)

**[iii]** Saul, Lawrence K., et al.

*"Spectral methods for dimensionality reduction"* **Semi-supervised learning** 3, 2006.

False Negative (**FN**) – abnormal data that was not detected

True Positive (**TP**) – detected abnormal data

True Negative (**TN**) – correctly ignored normal data

False Positive (**FP**) – normal data wrongly detected

metrics are derived from the confusion matrix

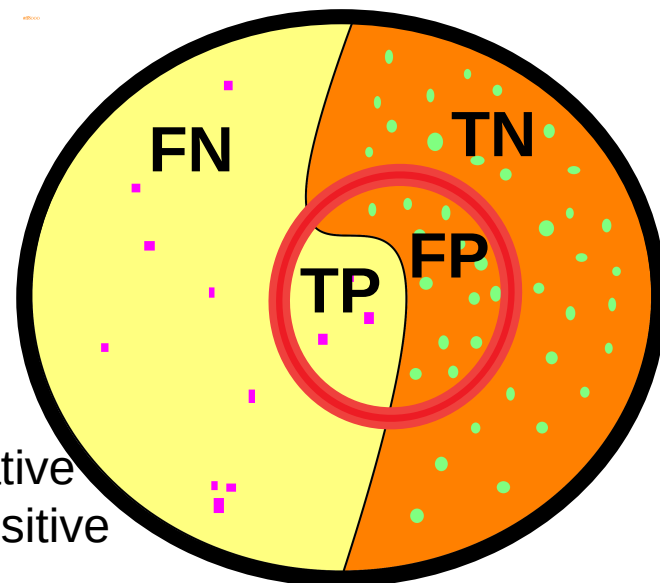
– Recall  $RE = \frac{TP}{FN + TP}$

– Precision  $PR = \frac{TP}{FP + TP}$

– FP-rate  $FPR = \frac{FP}{N + P}$

–  $F_1$ -Score  $F1 = 2 * \left( \frac{RE * PR}{RE + PR} \right)$

- relevant data
- training data
- selected data
- normal – negative
- abnormal – positive



**Problem:** one needs ground truth information for evaluation!

### Concept

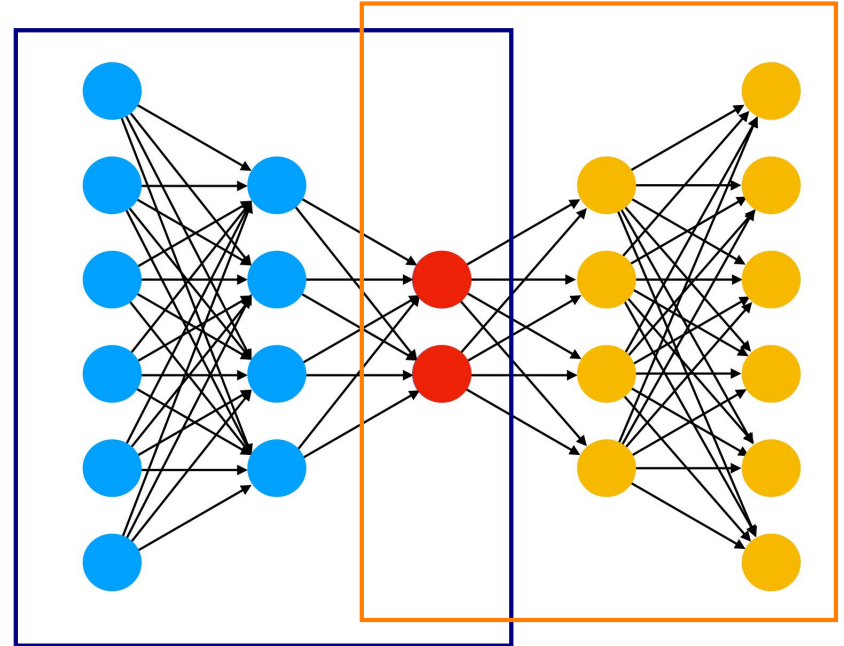
- Learns the identity function  $ID(x) = x$
- *Encoder* network compresses representation  $\rightarrow$  *code*
- *Decoder* network restores sample from *code*
- *Information bottleneck* forces generalization

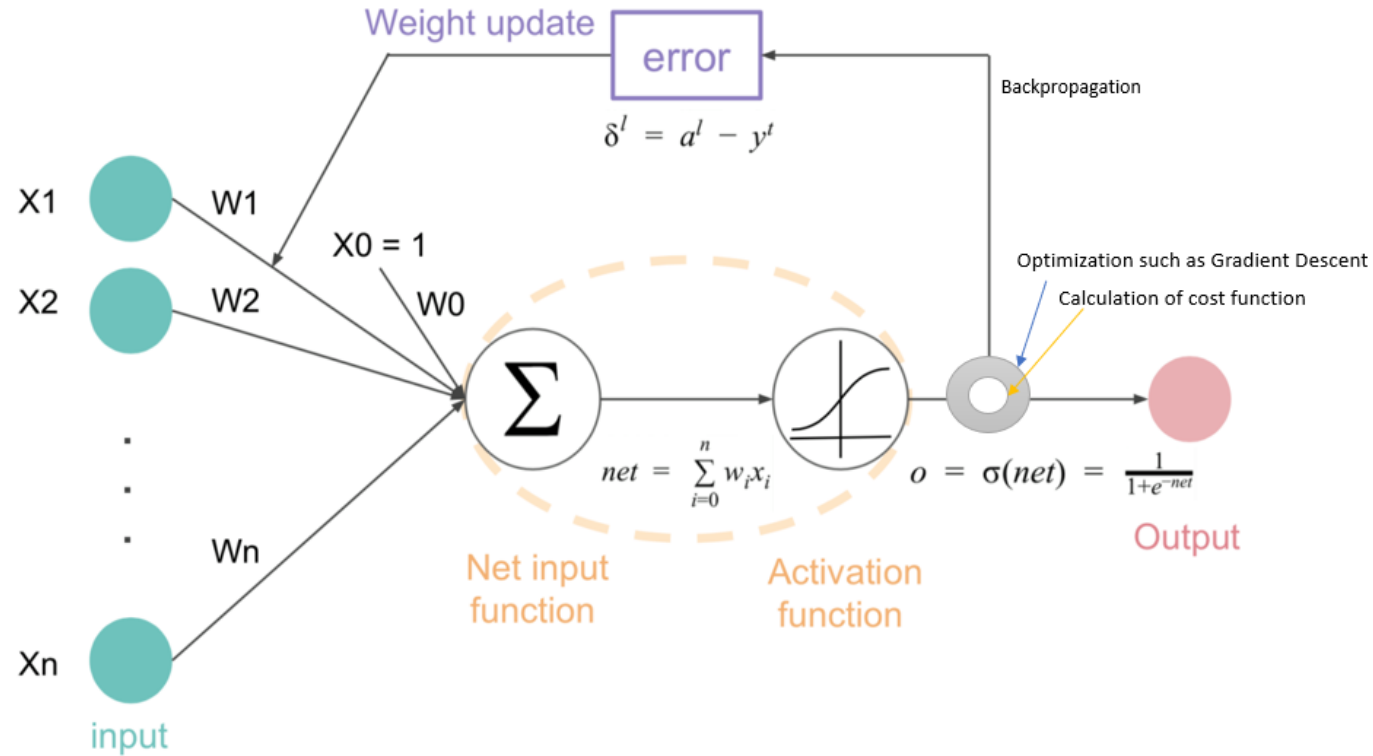
### Trivia

- Learn unsupervised
- Publicized in late 60's [7]
- Reconstruction *not* loss less
- Used in many different architectures (???)
- Neural networks are designed to simulate memory

### Pro's & Con's

- + Out of sample model
- + Gets better with data
- + Works unsupervised
- Blackbox by design

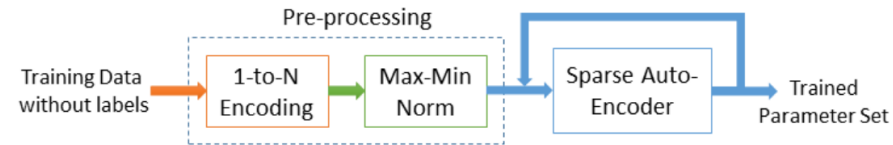




[<https://i.stack.imgur.com/7Ui1C.png>]

### A Deep Learning Approach for Network Intrusion Detection System (*ACM BIONETICS*, 2016)

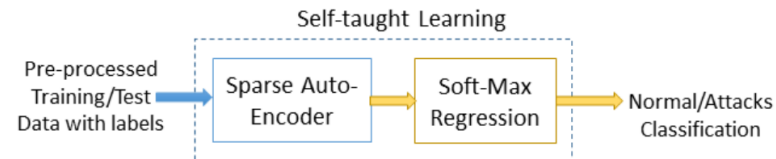
- NSL-KDD dataset (41 features)
- Autoencoder for unsupervised feature learning + Soft-max regression for for classification



(a) Feature Learning from pre-processed data



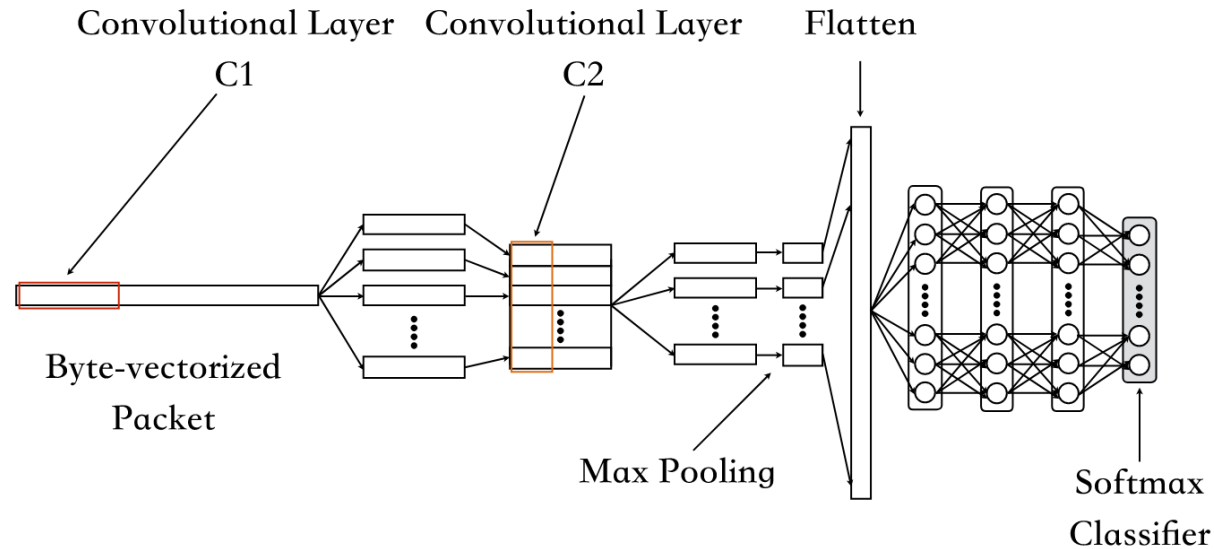
(b) Soft-max Regression classifier training for the derived training data





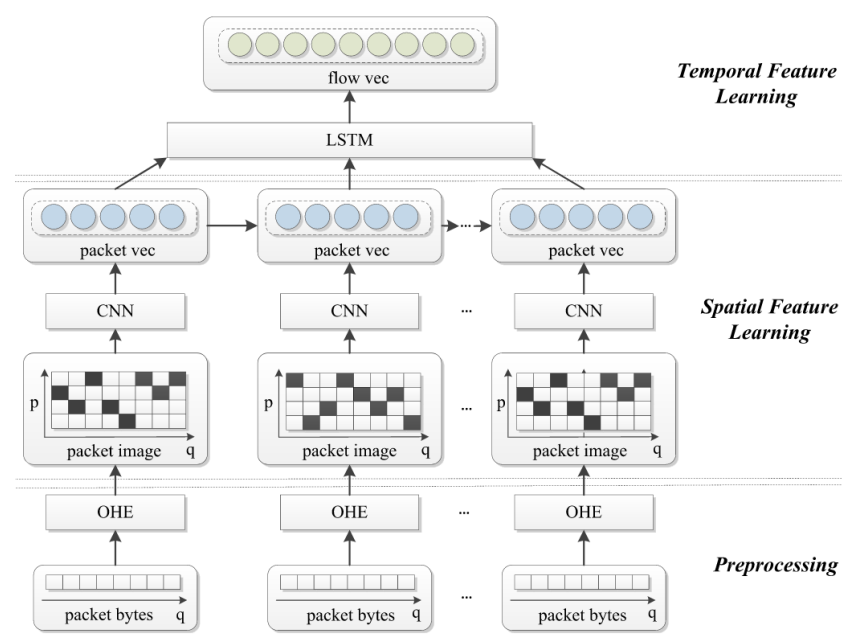
Deep packet: a novel approach for encrypted traffic classification using deep learning (*Soft Computing*, 2020)

- *Traffic characterization* (FTP, P2P, ...) and *application identification* (BitTorrent, Skype, ...)
- Distinguishes between VPN and nonVPN traffic, but fails to classify *tor* traffic
- Comparison between different supervised architectures (SAE and CNN)
- UNB ISCX dataset
- Do not regard any temporal phenomenon



HAST-IDS: Learning Hierarchical Spatial-Temporal Features Using Deep Neural Networks to Improve Intrusion Detection (*IEEE ACCESS*, 2017)

- DARPA1998 and ISCX2012 for evaluation
- Bytes are transformed via a one-hot-encoding
- Soft-max for classification



**[2]** Malware Traffic Classification Using Convolutional Neural Network for Representation Learning (*IEEE ICOIN, 2017*)

- *USTC-TFC2016* data set
- *Spatial* feature extraction (*LeNet-5*) + Soft-Max regression classifier
- Bi-directional packet representation with all layers yields best results

