Cyber Defence Reinforced

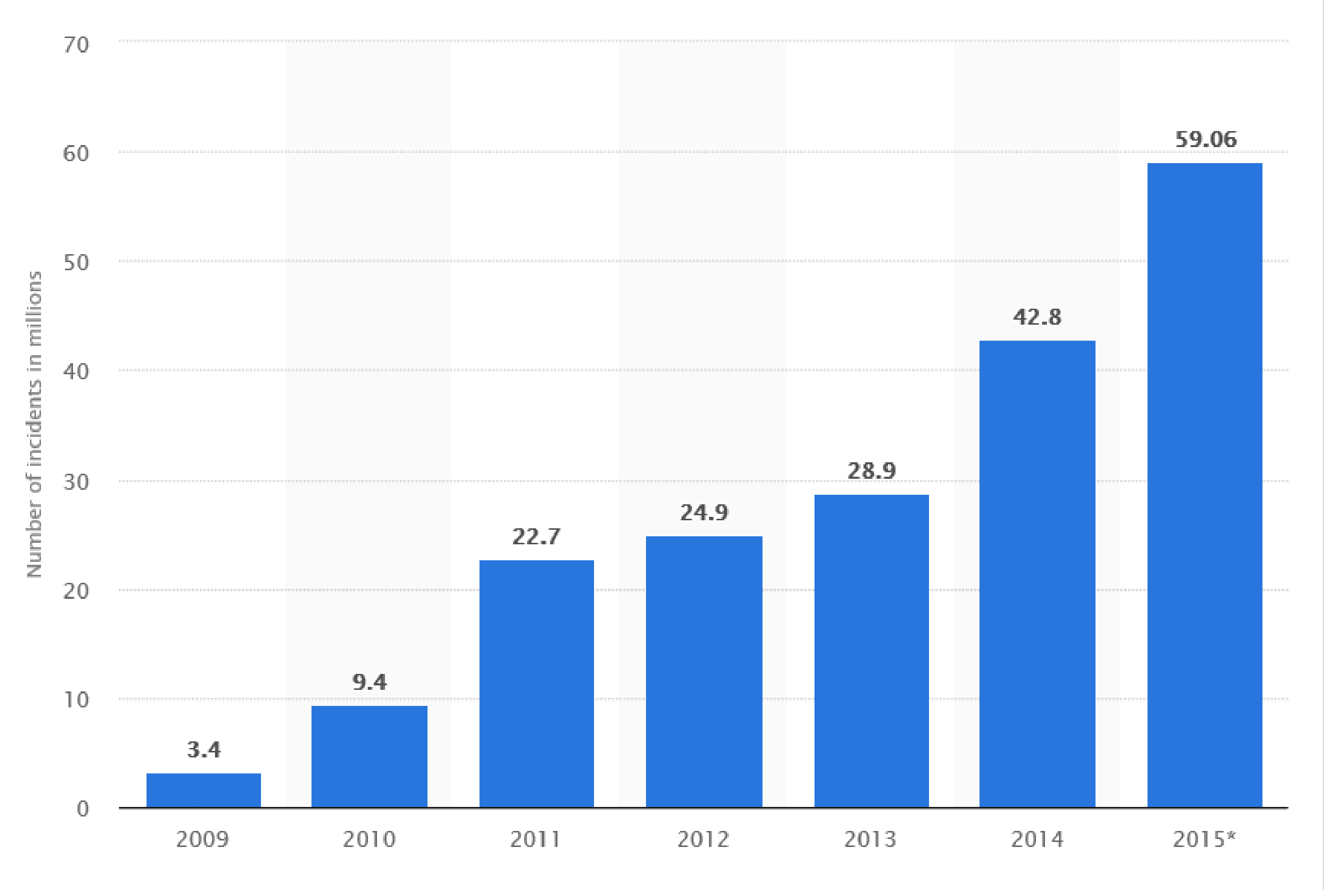
through

Machine Learning and AI

Introduction

- **Frequency** and **Cost** of Cyber Incidents have increased year over year
- According to a 2018 joint report** by CSIS and McAffee, it is estimated that **cybercrime costs** the world's economy about $600 billion a year (~0.8% of global GDP) – up from $445 billion estimated in 2014^
- **Cybercriminal revenues** worldwide is estimated to be at least $1.5 trillion based on a research published in 2018 RSA Conference^^
  - Over 50% of cybercrime revenues are generated in online markets

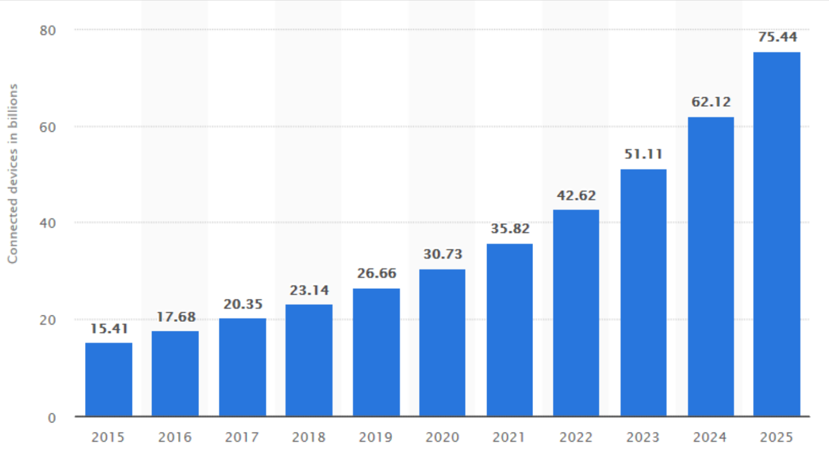
Crime	Annual Revenues
Illegal online markets	$860 Billion
Trade secret, IP theft	$500 Billion
Data Trading	$160 Billion
Crime-ware/CaaS	$1.6 Billion
Ransomware	$1 Billion
Total Cybercrime Revenues	$1.5 Trillion



Global number of cyber security incidents from 2009 to 2015 (in millions)*

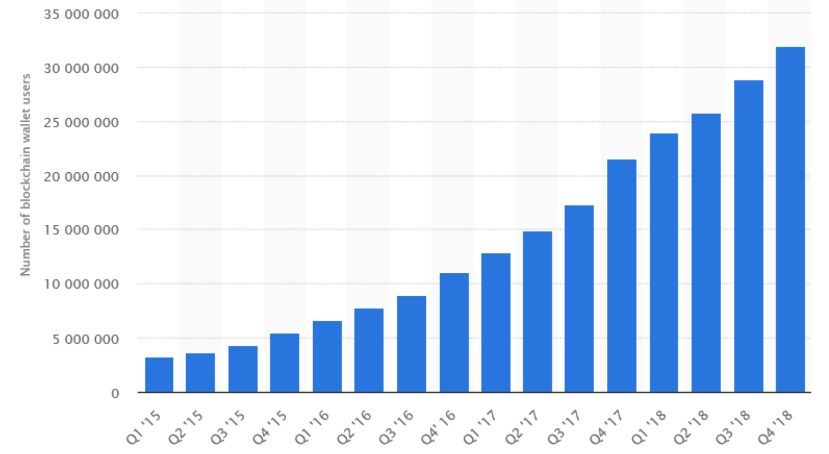
* https://www.statista.com/statistics/387857/number-cyber-security-incidents-worldwide/
** Center for Strategic and International Studies and McAffee, "Economic Impact of Cybercrime— No Slowing Down", February 2018
^ Center for Strategic and International Studies and McAffee, "Net Losses: Estimating the Global Cost of Cybercrime Economic impact of cybercrime II", June 2014
^^ https://www.rsaconference.com/videos/into-the-web-of-profit-tracking-the-proceeds-of-cybercrime

- Increased network bandwidth and number of devices connected to network have create more windows of opportunity
- Cybercriminals being able to conceal their identities and trade using tools like **Tor** and **Bitcoin** hence complicating law enforcement tracking efforts
- Growth of **Cybercrime-as-a-Service** has made it easier to commit cybercrime
- **Cybercrime communities** and knowledge sharing platforms have made information more accessible helping cybercriminals to learn new skills and adopt new tools faster
- High return on investment – according to the 2015 Trustwave Global Security Report, the estimated **ROI of cybercrime is %1,425**



IoT connected devices worldwide (in billions)*

© Statista 2019



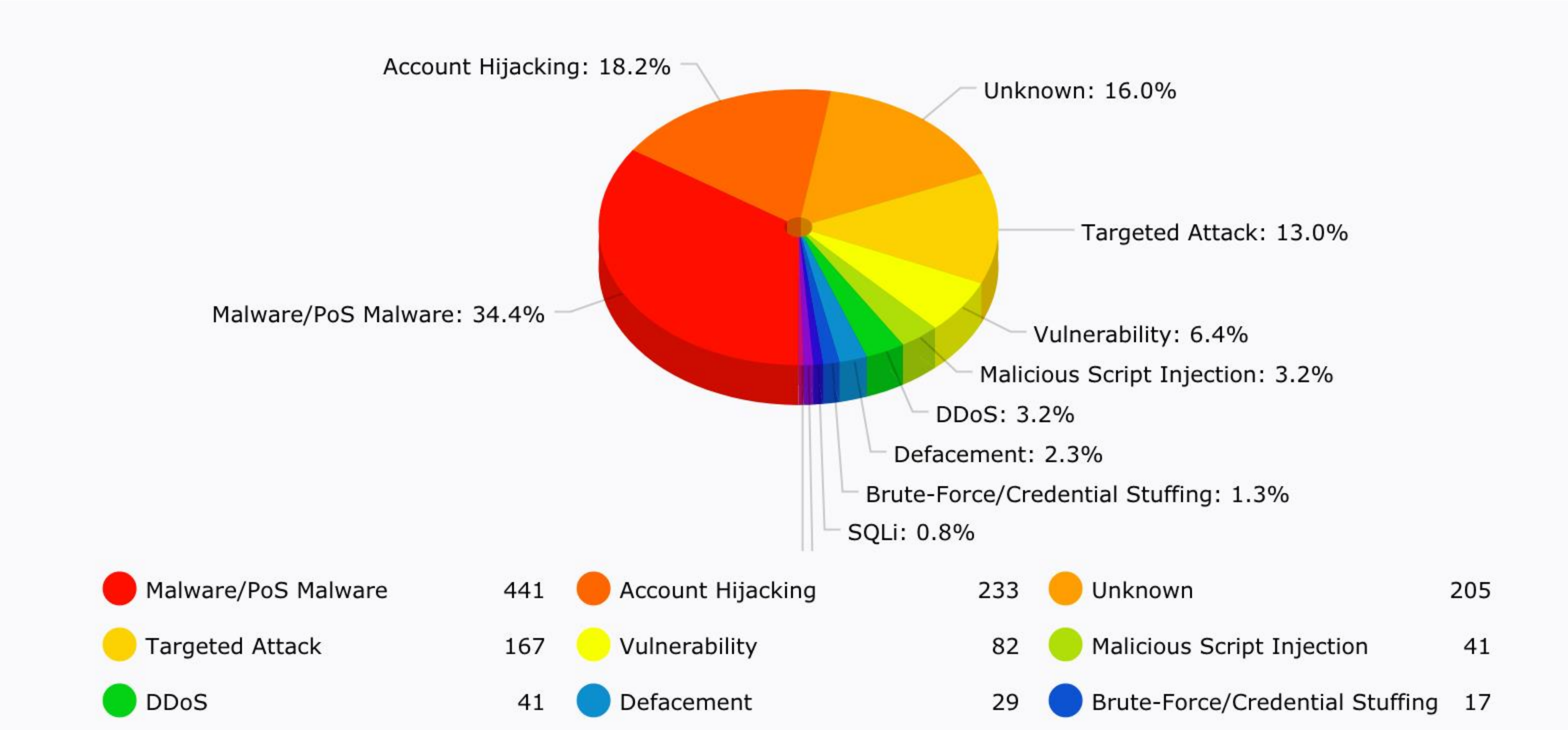
Number of Blockchain wallet users worldwide**

© Statista 2019

* https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/
** https://www.statista.com/statistics/647374/worldwide-blockchain-wallet-users/

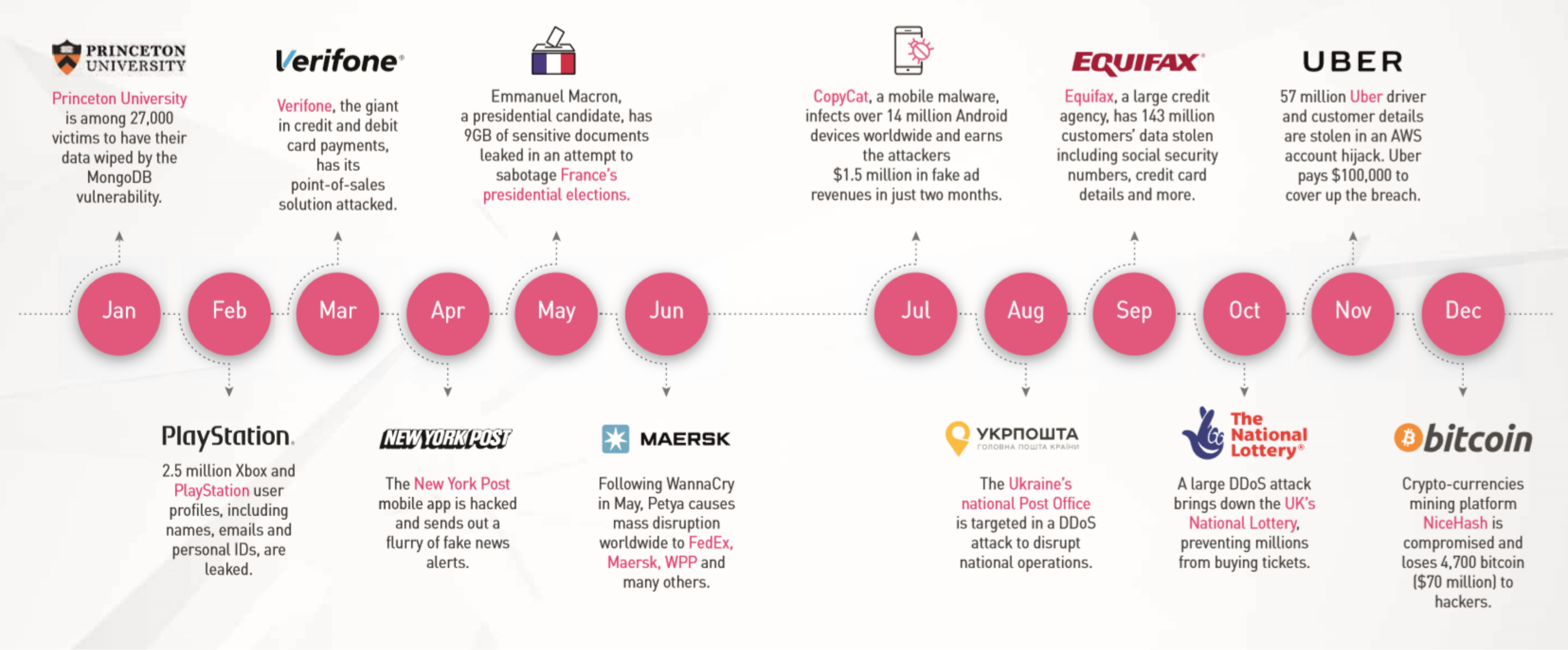
Introduction

Attack Distribution (Top 10 categories in 2018)



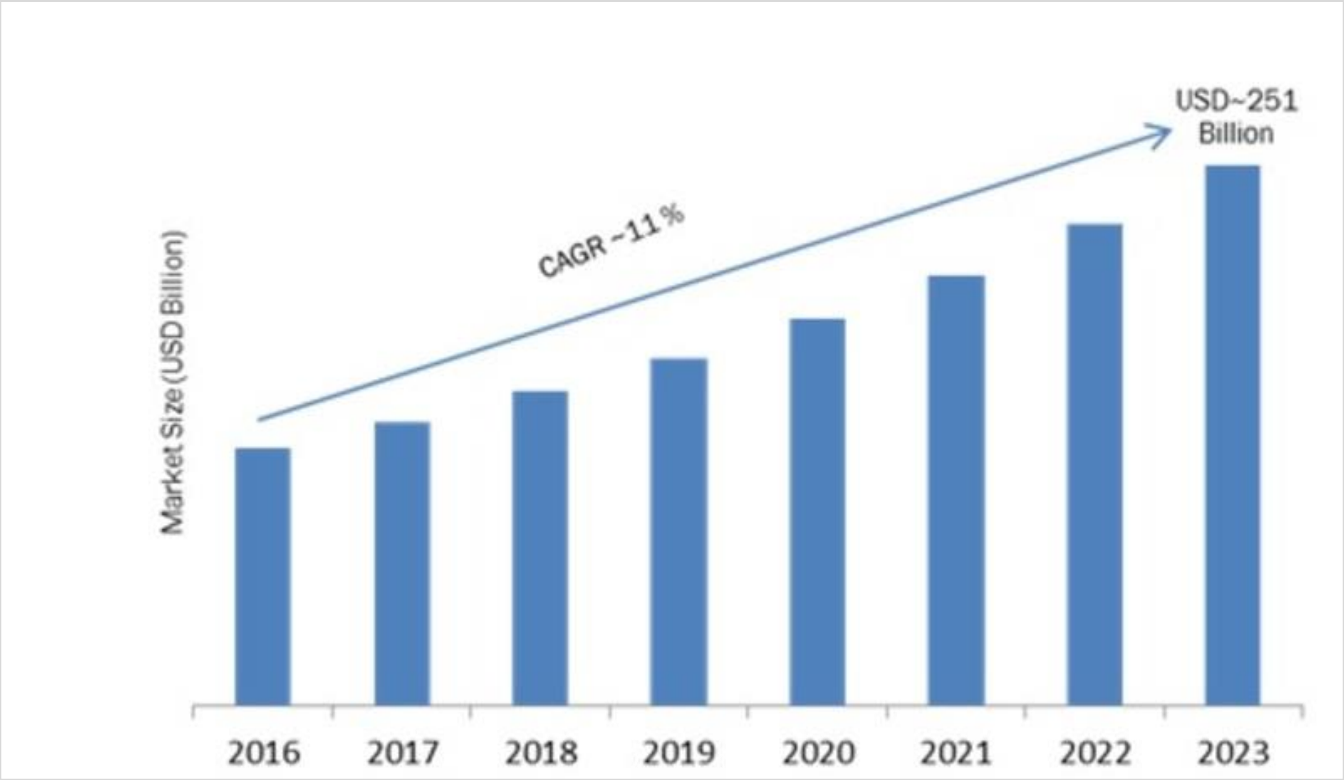
* https://www.hackmageddon.com/2019/01/15/2018-a-year-of-cyber-attacks/

2017 Timeline for Major Cyber Attacks*



* Check Point Research, Security Report, 2018

- The global cyber security market is expected to grow at approx. $251 Billion by 2023, at 11% of CAGR between 2017 and 2023*



* Market Research Future, “Cyber Security Market Research Report- Global Forecast 2023,” January 2019

Introduction

Organization are investing more in their Risk and Security programs to protect their business, including but not limited to:

- Training employees and promoting a risk and security aware culture in the organization
- Developing a unified security architecture that governs how preventive, detective, and corrective security controls are deployed across the organization while ensuring compliance with information security policies and standards
- Setting up a Security Operations Center (SOC) and building an Incident Response (IR) team that looks into events collected from controls to detect and respond to security incidents
    - Examples of controls are:
        - Identity and access management
        - Network security controls such as proxy, firewall, loadbalancer, etc
        - Device scanners and agents such as Antivirus, Intrusion Prevention/Detection System, etc
        - Sandbox
        - Mail server spam filter
        - Data Loss Monitoring/Prevention,…
    - Security logs generated by controls are collected by a Security Information and Event Management (SIEM) solution where the logs are checked against security rules to detect incidents and send alerts to the IR team

Introduction

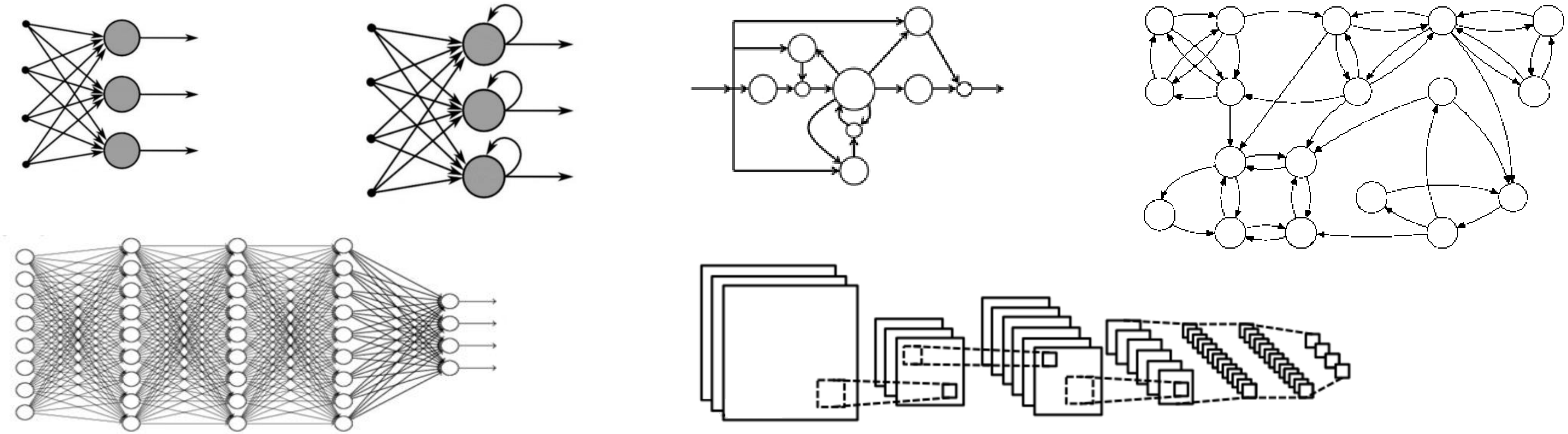
- Conventional controls such as SIEM use rigid rules to detect threats, thus they are only capable of detecting known threats that
  - have been previously seen in other attacks, and
  - are detectable using simple rules.

- Conventional controls are not effective against attacks not seen before
  - A new type of malware
  - Blind spots and vulnerabilities that are still unknown
  - A complex attack, like the threat of an insider who is aware of existing controls and knows how to go under the radar, low and slow,…

- A new breed of security controls is therefore required that can use previously acquired knowledge to solve new problems
  - By definition this is a feature of an intelligent system that is to be realized through AI
  - A light-weight version of this feature is a system's ability to generalize, i.e., the ability to make correct predictions when exposed to data that is not seen before. The ability to generalize is in fact the most important performance measure of a machine learning model.

Learning Machine

- Behavior is learned from data vs. coded instructions
- Extract relevant information from data and exhibit a desired behavior
- What is deemed to be relevant information or desired behavior depends on factors like
    - Architecture
    - Mathematical/Probabilistic assumptions
    - Training strategy
    - Nature of the problem and data which affect the three items above



Learning Paradigms*

- **Supervised Learning**
  Learning from examples of desired behaviour

- **Unsupervised Learning**
  Learning associations / patterns from observations

- **Reinforcement Learning**
  Learning from consequences of actions, i.e., rewards/punishments

* Simon Haykin, “Neural Networks and Learning Machines,” 3rd Edition , Pearson, 2008

Learning Paradigms*

- **Supervised Learning**
  Learning from examples of desired behaviour

- **Unsupervised Learning**
  Learning associations / patterns from observations

Semi-supervised Learning

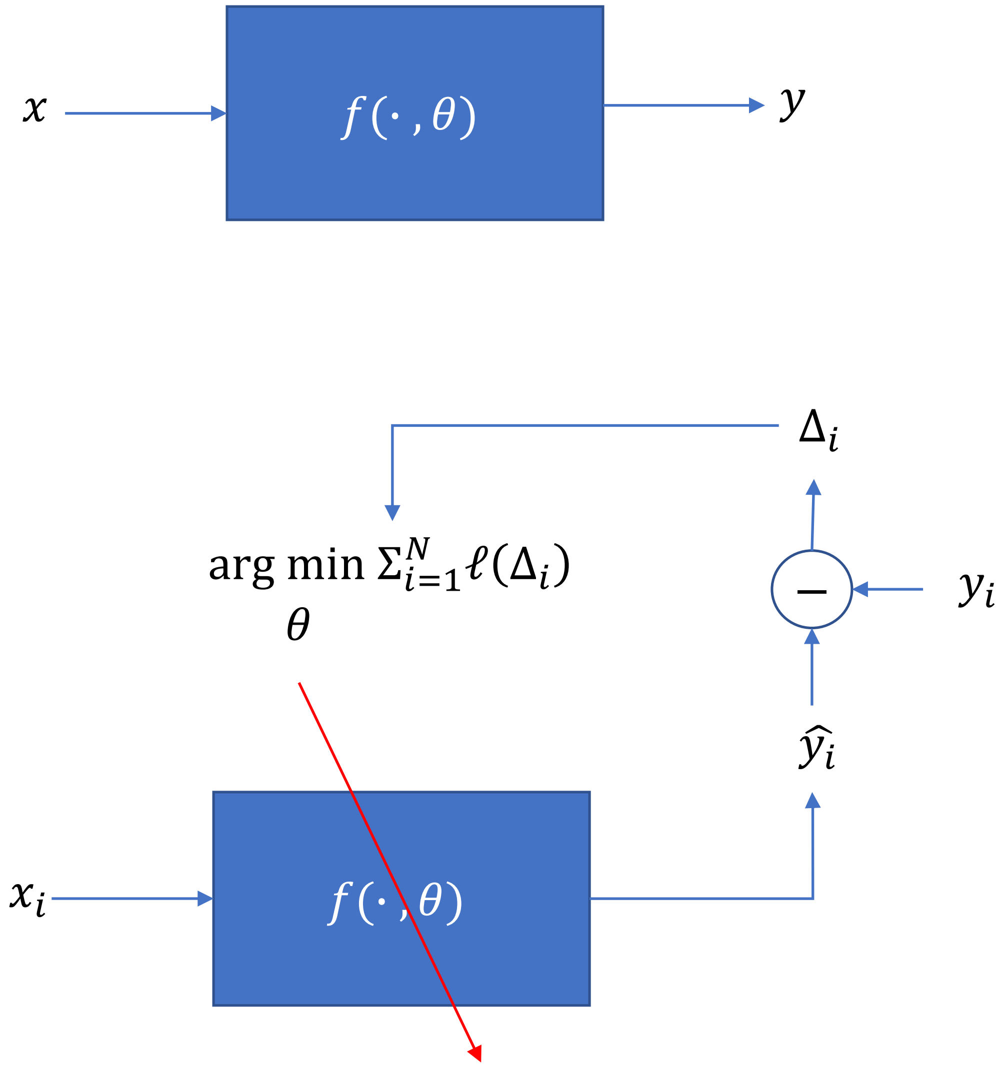
- **Reinforcement Learning**
  Learning from consequences of actions, i.e., rewards/punishments

* Simon Haykin, “Neural Networks and Learning Machines,” 3rd Edition , Pearson, 2008

Supervised Learning

- x : input
- y : target variable
- $f(\cdot, \theta)$: model
- θ : model parameters (to be learned)
- Assume (x_i, y_i) is a training example and $\hat{y}_i = f(x_i, \theta)$
- We now define $\Delta_i = \hat{y}_i - y_i$ as well as a cost function $\Sigma_{i=1}^{N} \ell(\Delta_i)$, which is the total cost over N training examples
- We now go on to compute θ by solving the following optimization problem:

$$\theta^* = \underset{\theta}{\arg\min}\ \Sigma_{i=1}^{N} \ell(\Delta_i)$$

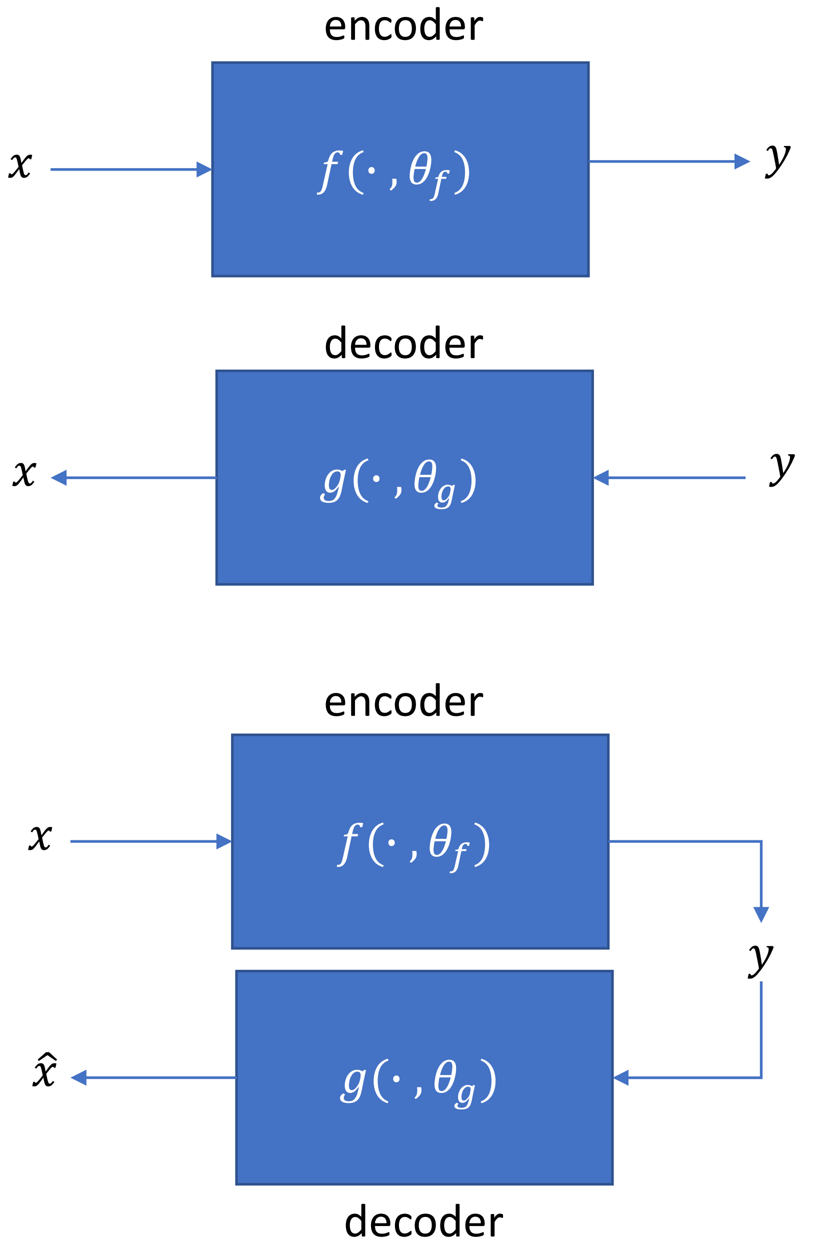


Unsupervised Learning

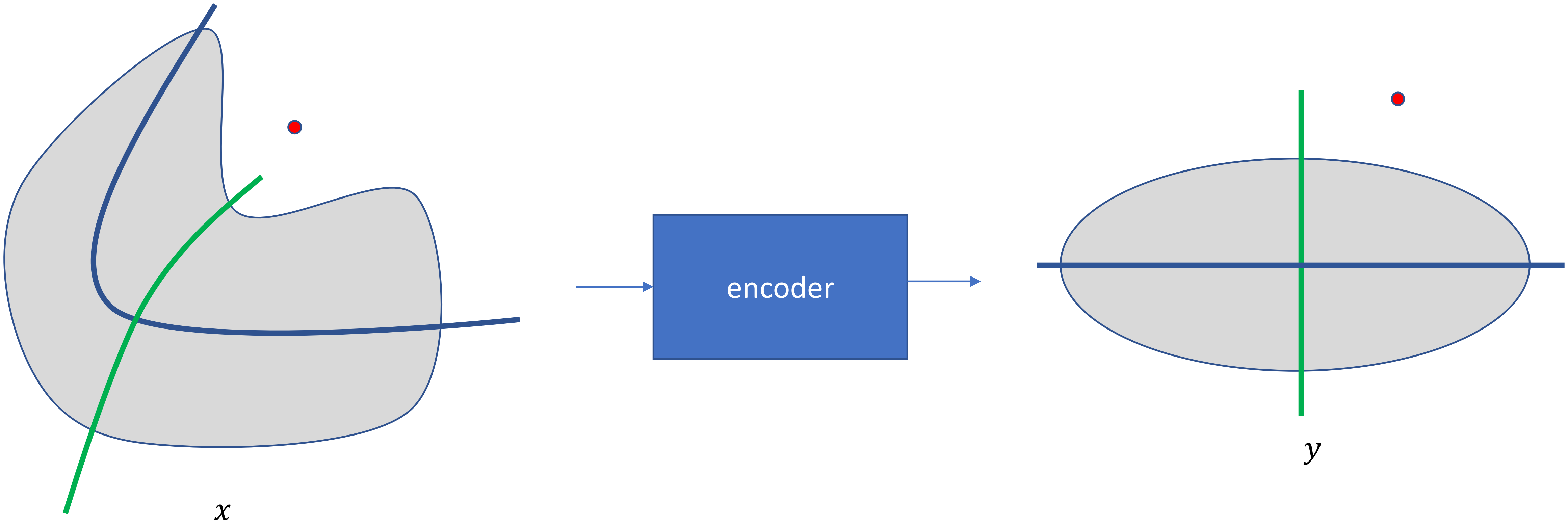
- x : input
- y : latent variable
- $f(\cdot\,,\theta_f)$: encoding (projection)
- θ_f : model parameters (to be learned)
- $g(\cdot\,,\theta_g)$: decoding (reconstruction)
- θ_g : model parameters (to be learned)
- Since there are no labels available, learning is performed using the input x
- We define $\Delta_i = \hat{x}_i - x_i$ as well as a cost function $\Sigma_{i=1}^{N}\ell(\Delta_i)$, which is the total cost over N input samples
- We now go on to compute θ by solving the following optimization problem:

$$\theta^* = \underset{\theta}{\arg\min}\, \Sigma_{i=1}^{N}\ell(\Delta_i)$$

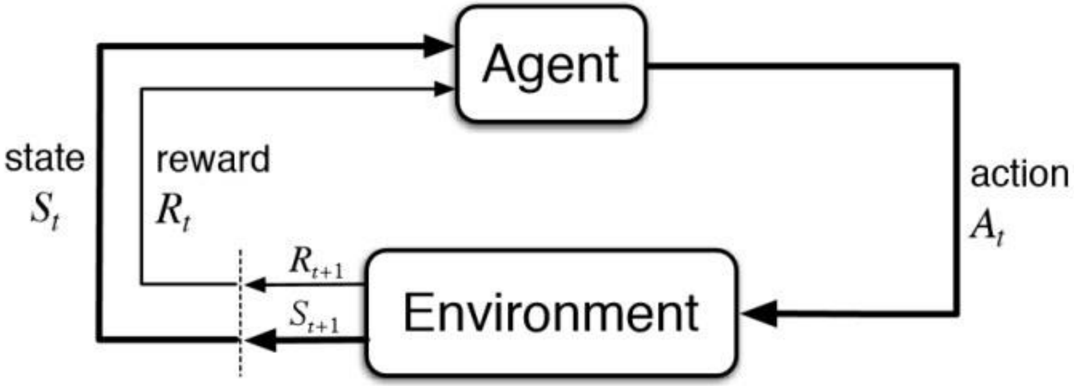
- Often, assumptions are made about statistical characteristics of y to regularize training

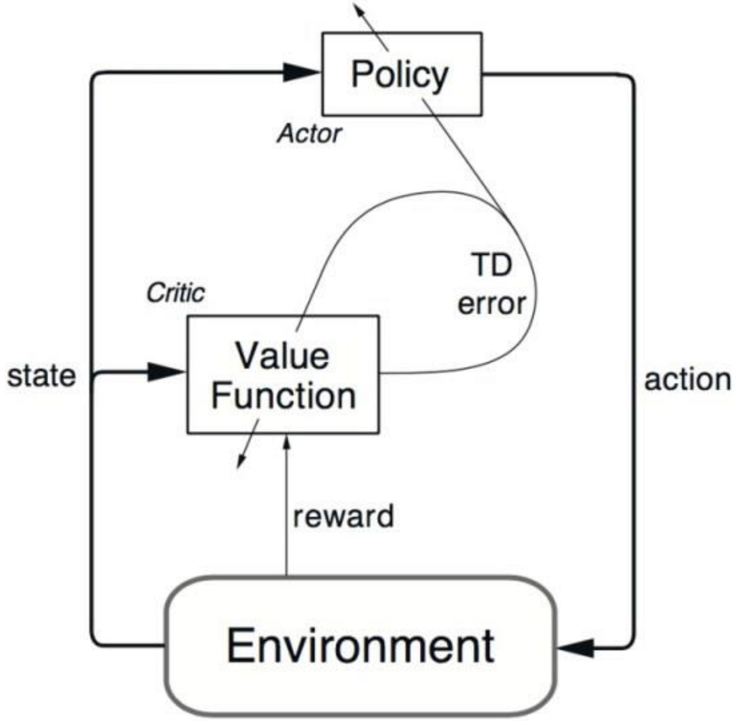


Unsupervised Learning



Reinforcement Learning*





$$V^{\pi}(s) = E[R|s, \pi],$$

$$Q^{\pi}(s, a) = E[R|s, a, \pi],$$

$$Q(s_t, a_t) \leftarrow (1-\alpha) \cdot \underbrace{Q(s_t, a_t)}_{\text{old value}} + \underbrace{\alpha}_{\text{learning rate}} \cdot \overbrace{\left(\underbrace{r_t}_{\text{reward}} + \underbrace{\gamma}_{\text{discount factor}} \cdot \underbrace{\max_a Q(s_{t+1}, a)}_{\text{estimate of optimal future value}} \right)}^{\text{learned value}}$$

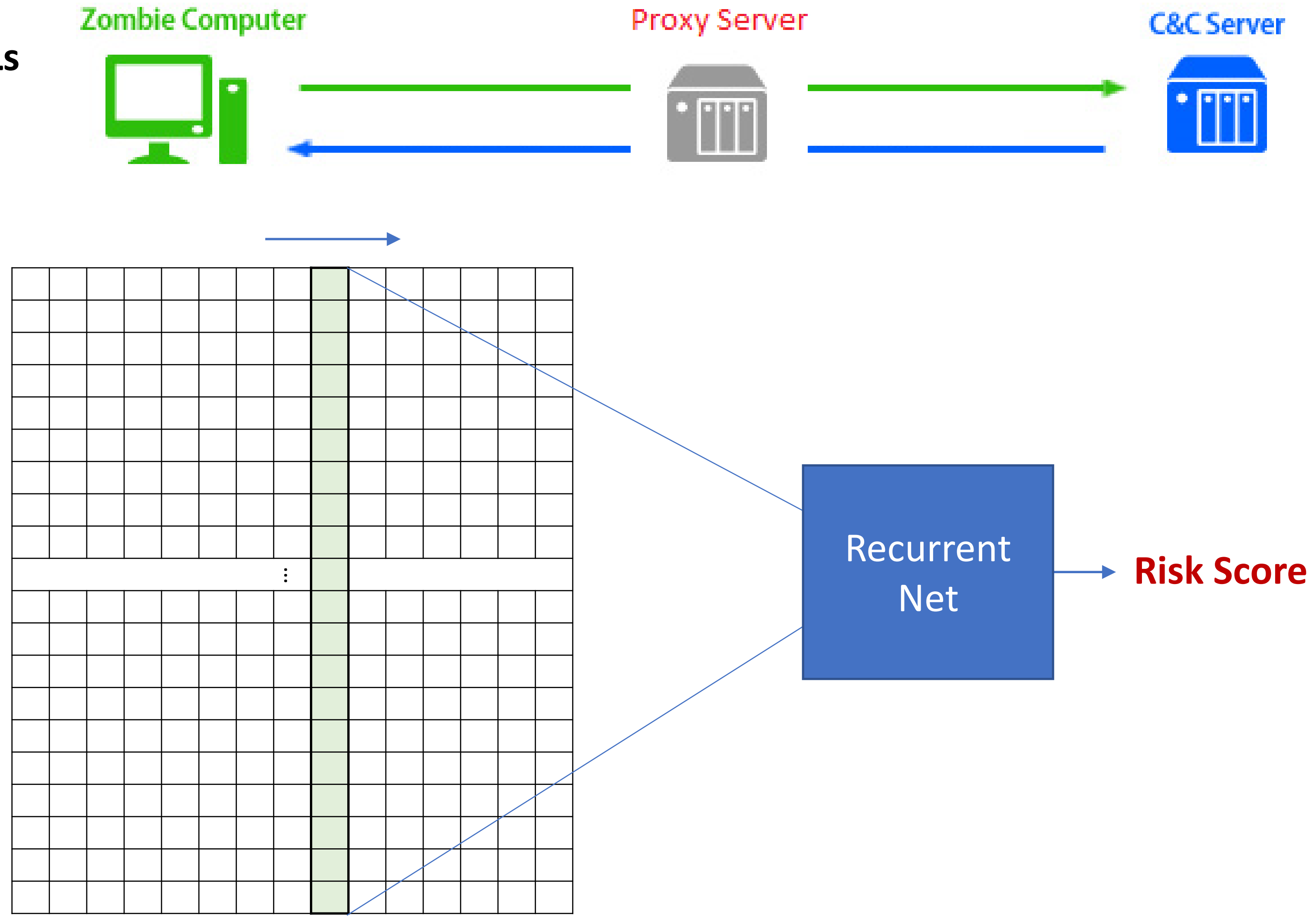
* Richard S. Sutton and Andrew G. Barto, “Reinforcement Learning: An Introduction Second edition,” The MIT Press Cambridge, 2018

Malware Detection – Algorithmically Generated Domains

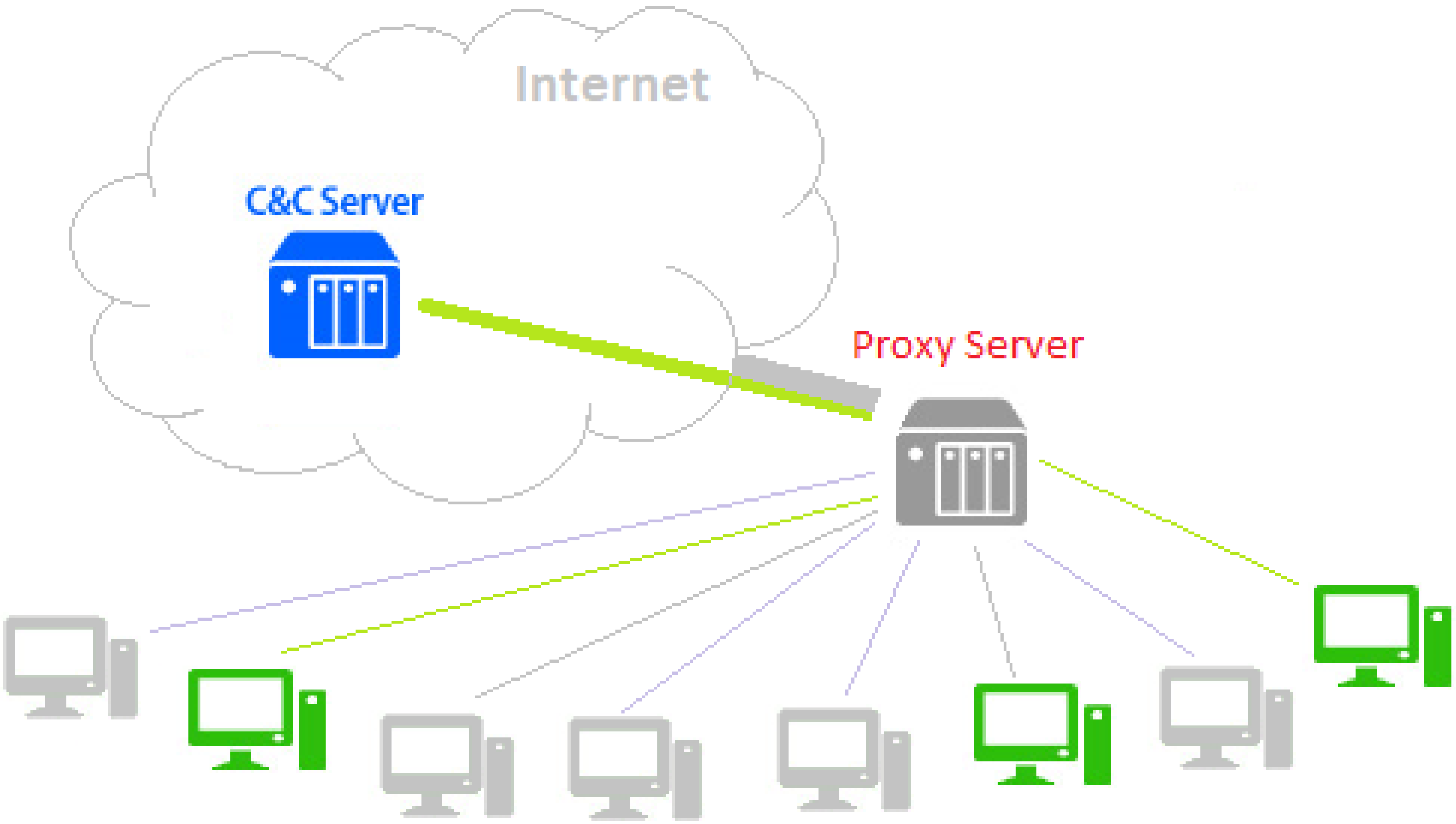
Domain Generation Algorithm (DGA) URLs

1008bnt1iekzdt1fqjb76pijxhr.org
cr22z71id7qxabmf5ne7nrsv8.net
db3xmx3xefvk1r06rk91mkfwvp.com
eiplrprtspxcymrkbcmlzmzxl.ru
fr902710xvla34ydvpj1qyi1vc.com
gmjftbehiynvi.ru
hb9mc2i3hm4vvzfclh1dc33o9.biz
hkkttu1f54fob1blif01z6d5ry.org

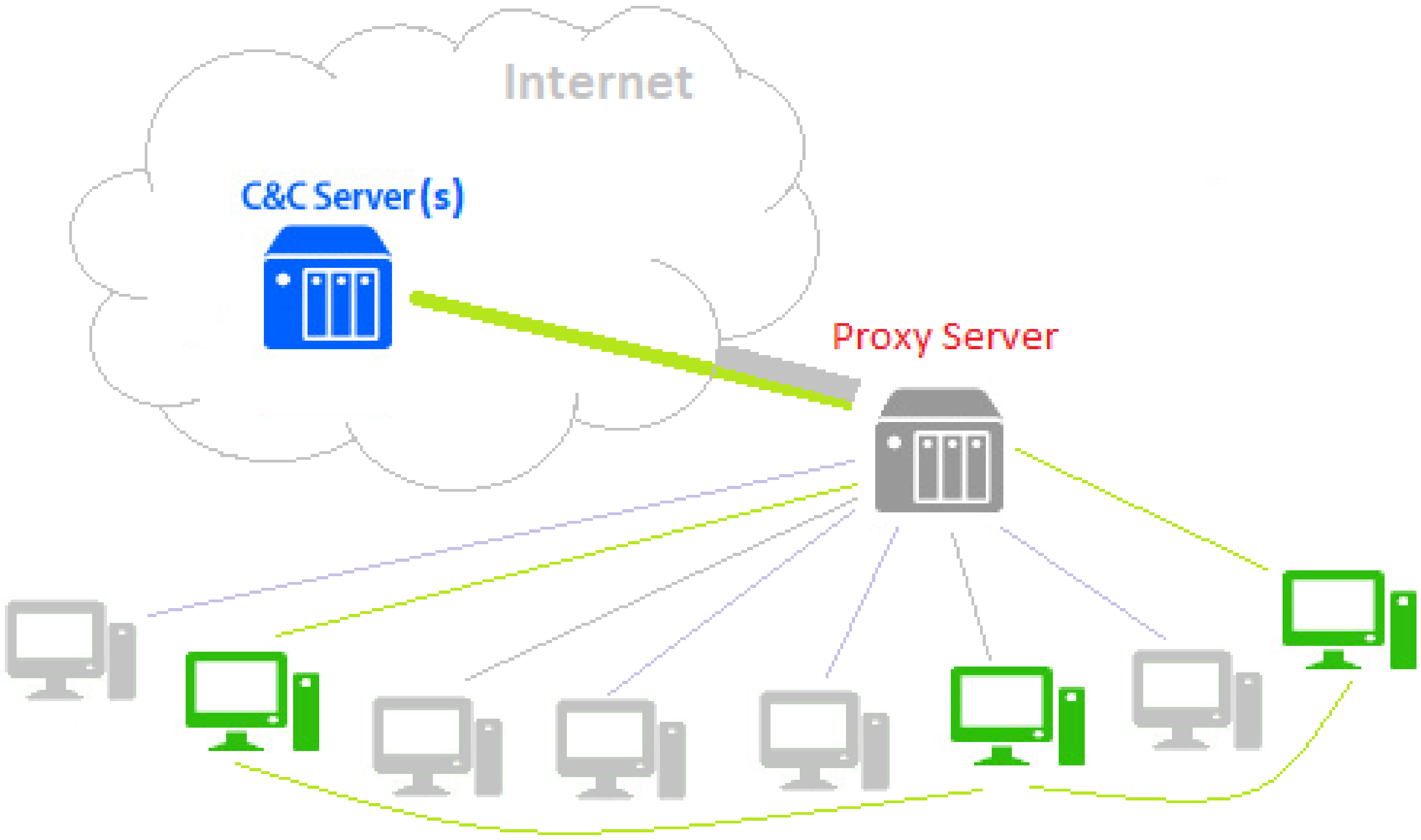
kingwhichtotallyadminis.biz
thareplunjudiciary.net
townsunalienable.net
taxeslawsmockhigh.net
transientperfidythe.biz
inhabitantslaindourmock.cn
thworldthesuffer.biz



Botnet Detection

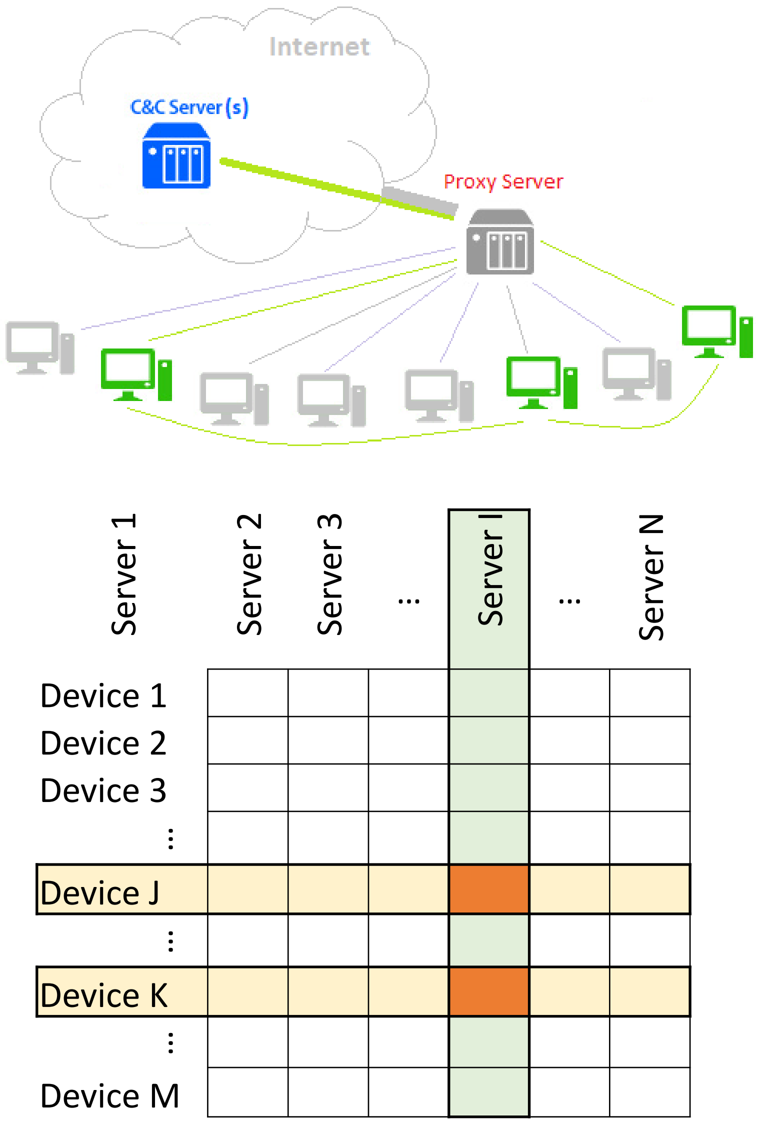


Botnet Detection



Botnet Detection

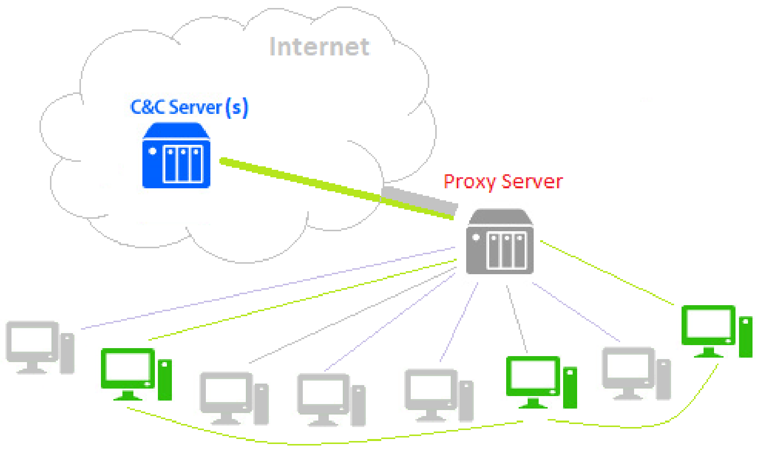
- Traffic between **Inbound Devices** and **Outbound Servers** is analyzed
- For each pair (Device, Server) a **score** is calculated to quantify how robotic the traffic between the pairs are
  - Assumption: robotic behavior is less complex in nature and therefore easier to predict
  - Approach: To calculate the score for a pair, time-series analysis is done on the traffic and predictability (i.e., complexity) of the time-series is quantified
- A sparse matrix is then created with rows and columns being the Devices and Servers, respectively, and predictability scores being the elements of this matrix.

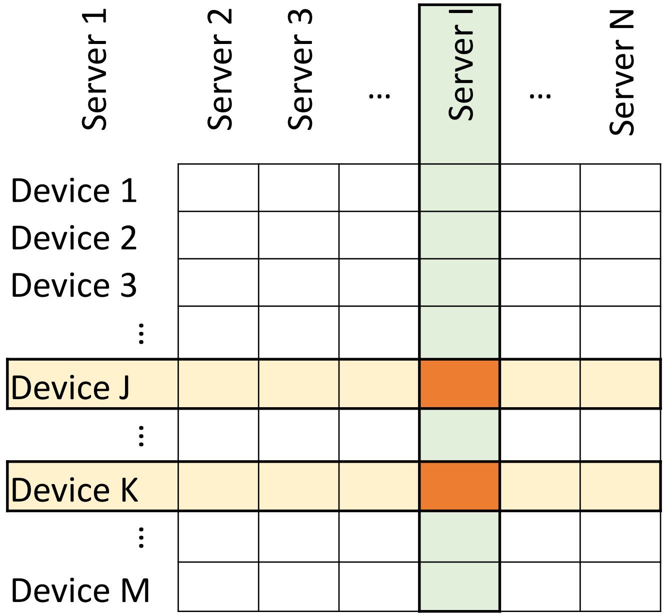


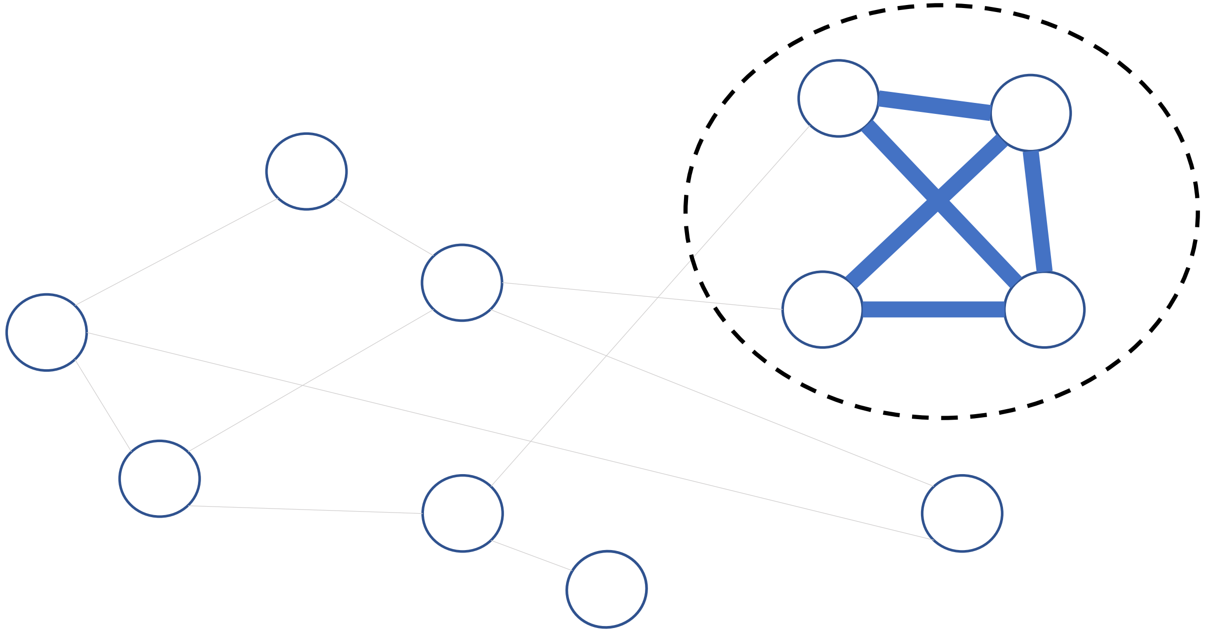
Cybersecurity Use Cases

Botnet Detection

- Next step is to calculate the connection between device pairs, e.g., (Device J, Device K) using their pertinent scores captured in rows J and K, respectively.
- The values calculated are used to create a graph. Strongly connected clusters that are made up of small number of devices are identified within the graph
- Finally, strongly connected clusters are traced back to outbound servers to identify root cause of anomalous behavior

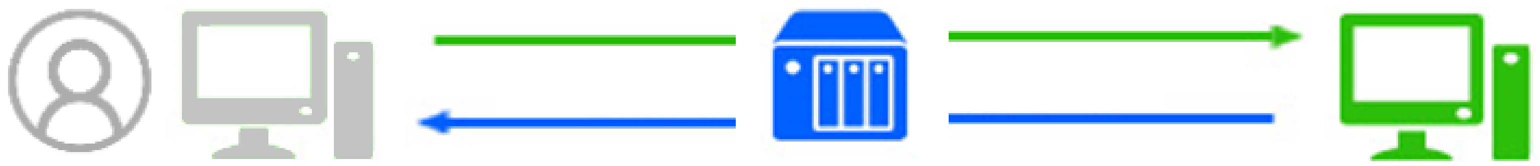






Privileged Access Monitoring

Privileged Access Maneger



User *requests temporary privilege elevation,e.g., for maintenance, etc.*

Activities are captured in the form of session recording (check out / in time, keystroke data, etc.)

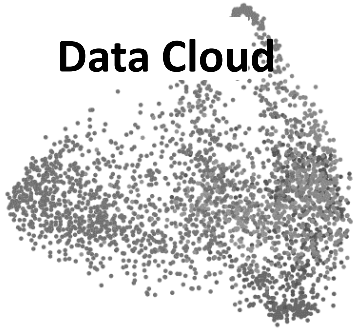
Managed System

Unsupervised

User IP / Device Name
Access: local or remote
Check out time
Session duration
Keystroke data
Service ticket description

First Layer Analysis and Normalization

Data Cloud



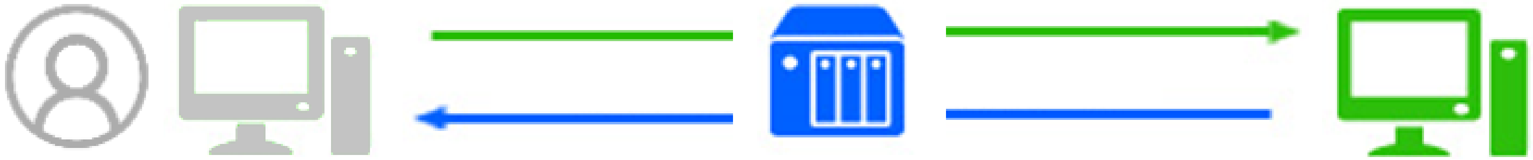
Manifold Learning and Dimensionality Reduction

Anomaly Detection

Suspicious Activity

Privileged Access Monitoring

Privileged Access Maneger



User *requests temporary privilege elevation,e.g., for maintenance, etc.*

Activities are captured in the form of session recording (check out / in time, keystroke data, etc.)

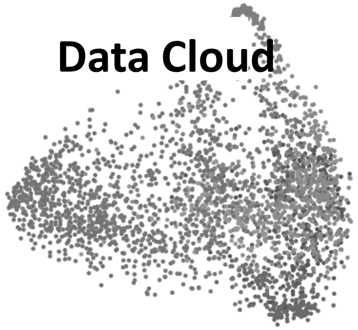
Managed System

Semi-supervised

User IP / Device Name
Access: local or remote
Check out time
Session duration
Keystroke data
Service ticket description

First Layer Analysis and Normalization

Data Cloud



Labelled historical data

Manifold Learning and Dimensionality Reduction

Anomaly Detection / Neighborhood Analysis

Suspicious Activity

Adversarial Reinforcement Learning*

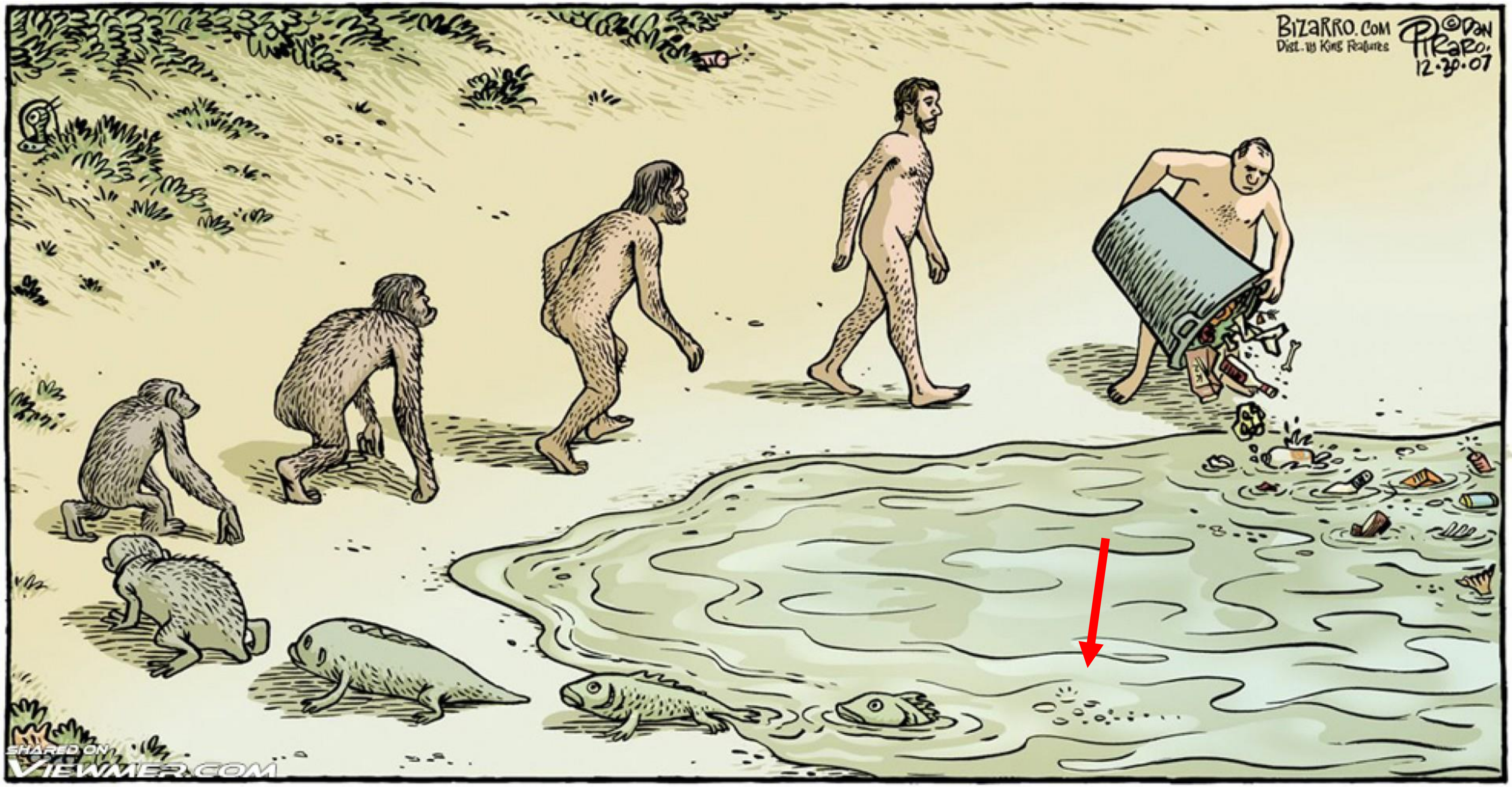
Loading...

R. Elderman, L. J. J. Pater, A. S. Thie, et al., "Adversarial Reinforcement Learning in a Cyber Security Simulation," In Proc. of ICAART, 2017, pp. 559-566

Other Considerations

- Risk-driven approach with targeted use cases
- Use case life-cycle
- Fine tuning of models
- Operationalization of use cases

Concluding Remarks



Questions