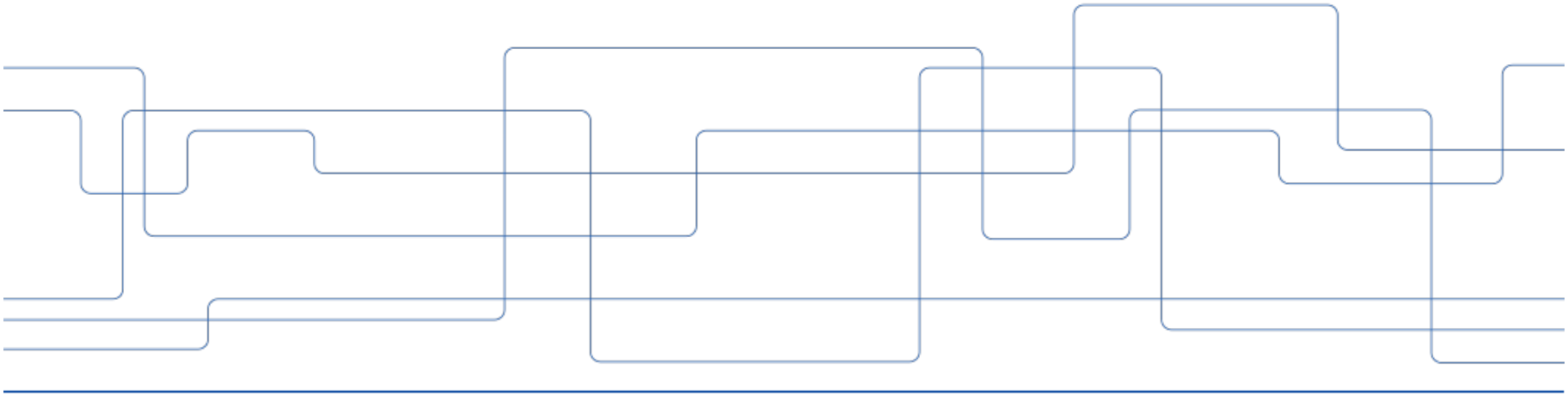
Maskininlärning för cybersäkerhet

Prof Pontus Johnson



Shakey
the Robot

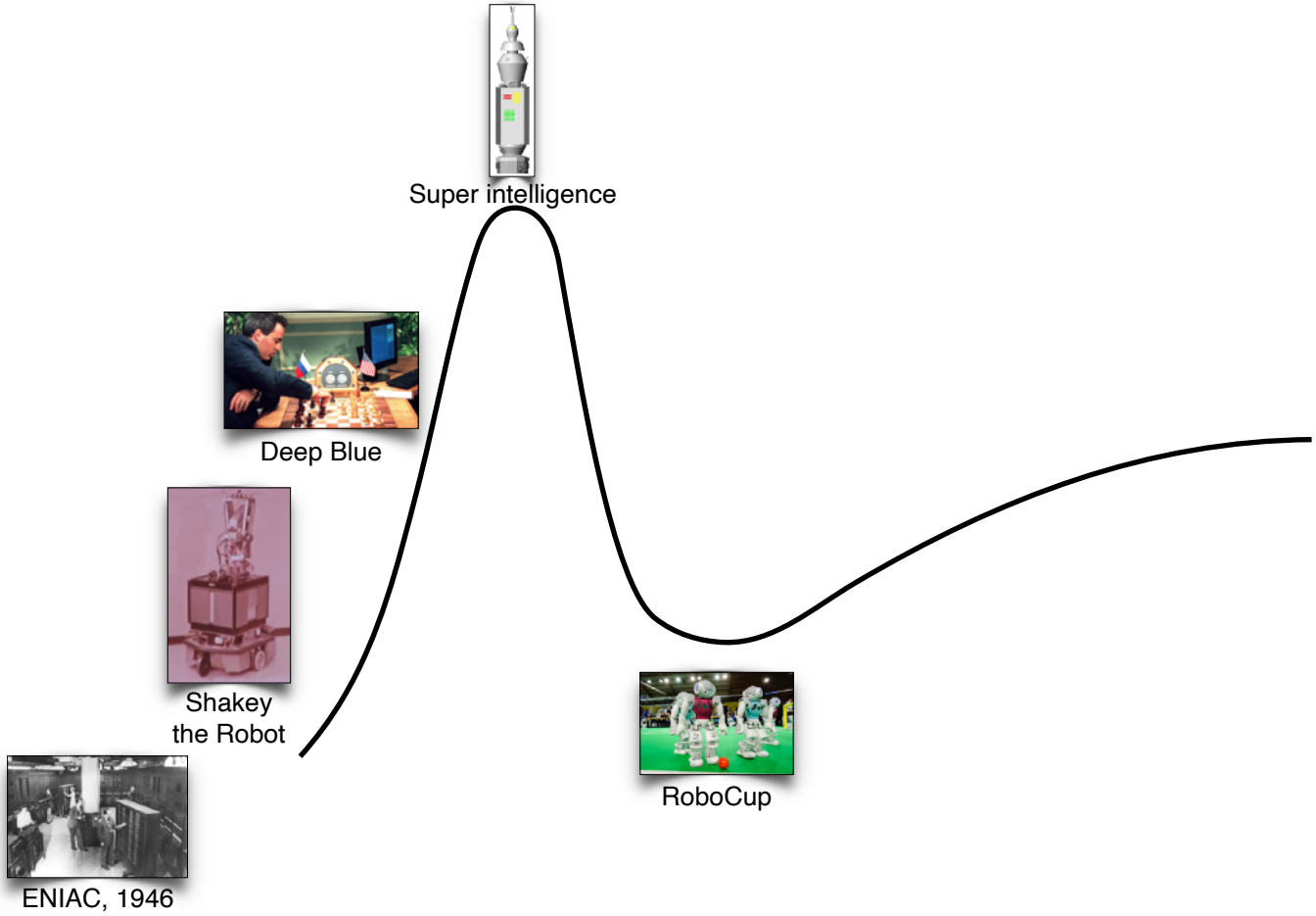
ENIAC, 1946

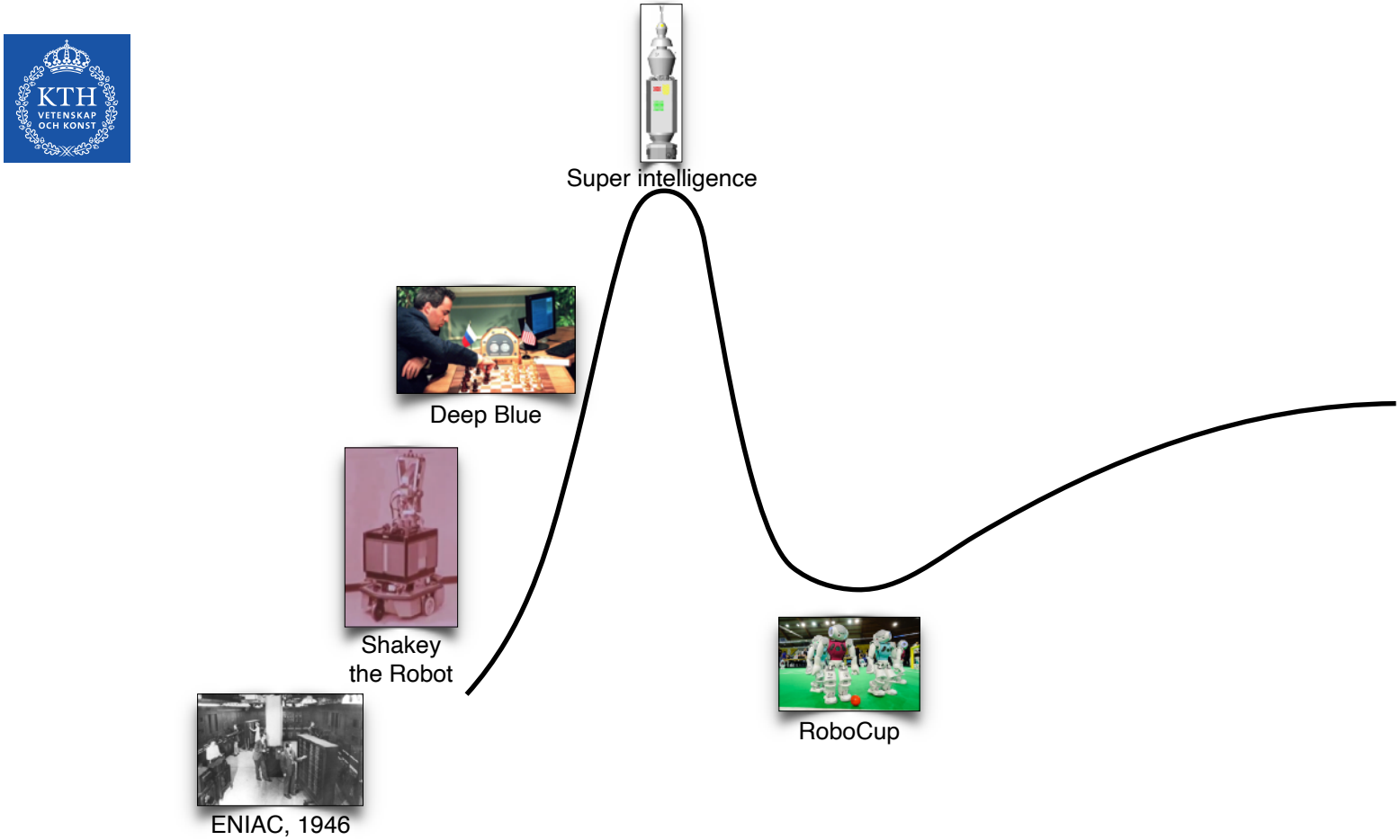
Super intelligence

Deep Blue

Shakey
the Robot

ENIAC, 1946


Super intelligence
Deep Blue
Shakey
the Robot
ENIAC, 1946
RoboCup


Super intelligence
Deep Blue
Shakey
the Robot
ENIAC, 1946
RoboCup

Neural computation 18.7 (2006)

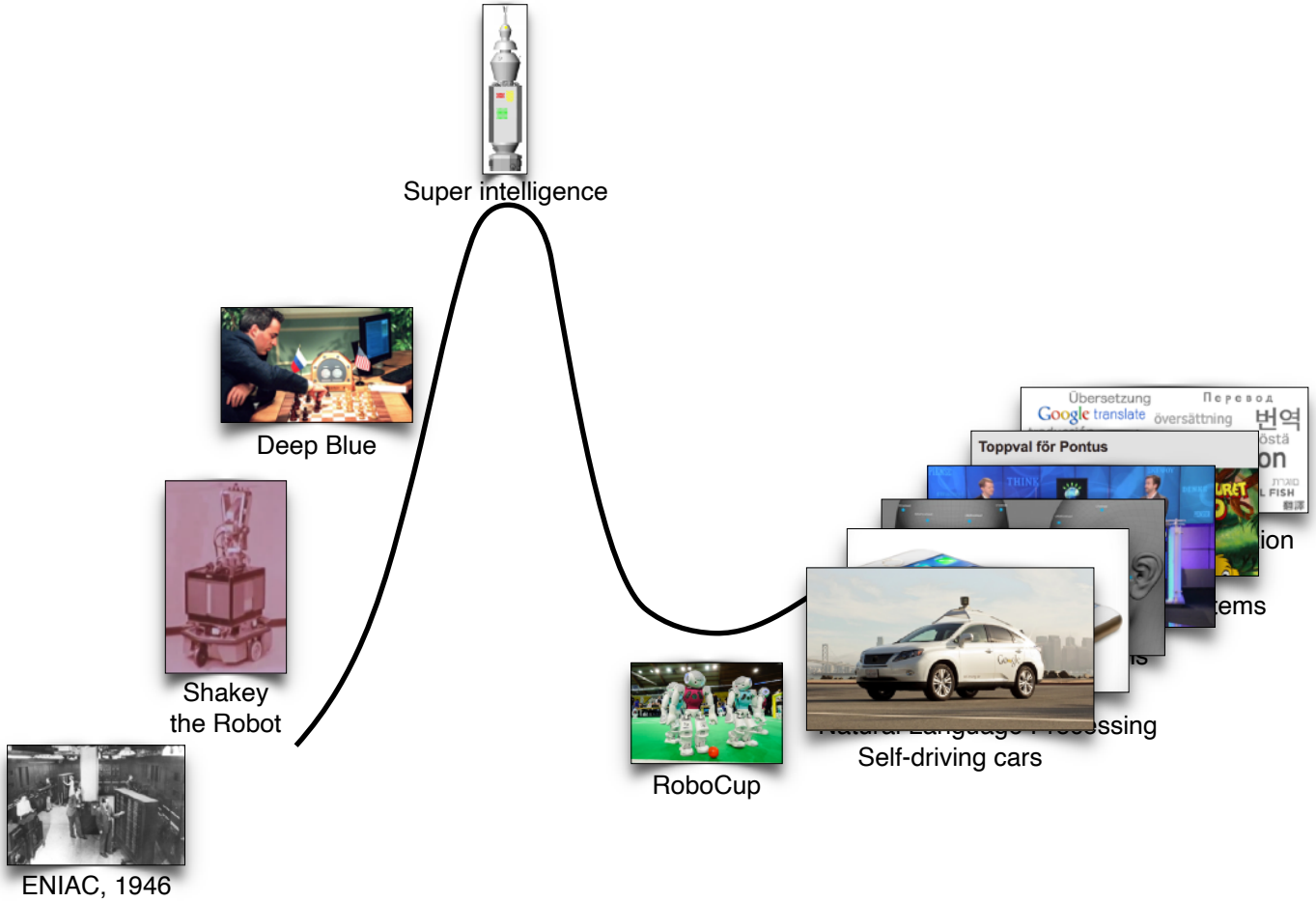
LETTER Communicated by Yann Le Cun

A Fast Learning Algorithm for Deep Belief Nets

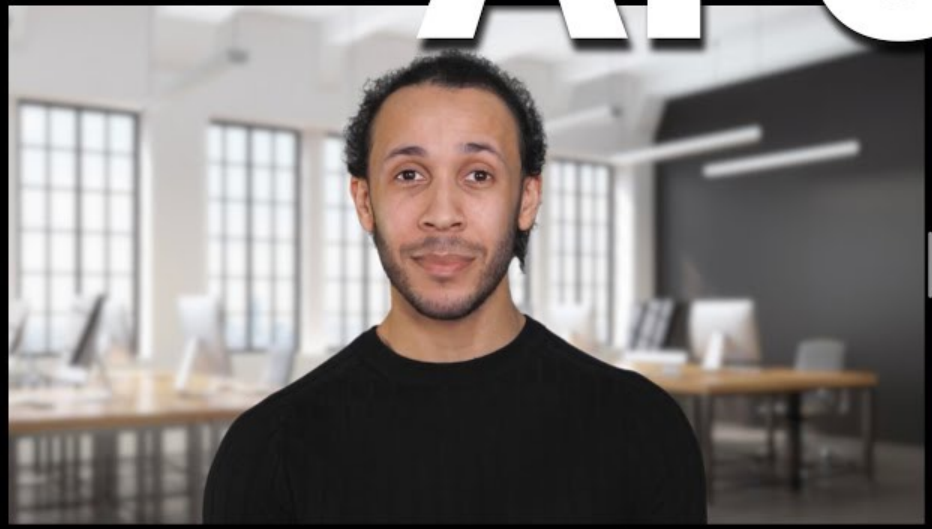
Geoffrey E. Hinton
hinton@cs.toronto.edu
Simon Osindero
osindero@cs.toronto.edu
Department of Computer Science, University of Toronto, Toronto, Canada M5S 3G4

Yee-Whye Teh
tehyw@comp.nus.edu.sg
Department of Computer Science, National University of Singapore, Singapore 117543

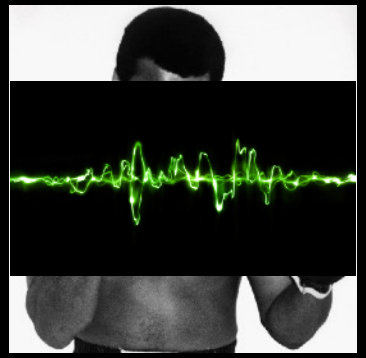
We show how to use "complementary priors" to eliminate the explaining-away effects that make inference difficult in densely connected belief nets

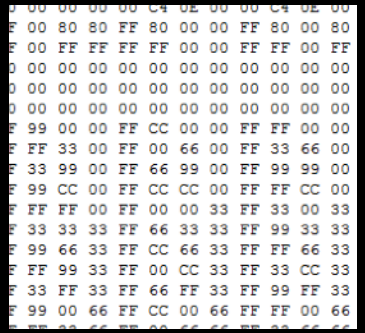

Super intelligence
Deep Blue
Shakey
the Robot
ENIAC, 1946
RoboCup
Self-driving cars
Toppval för Pontus

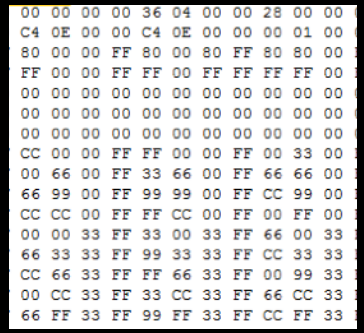
AI Chat

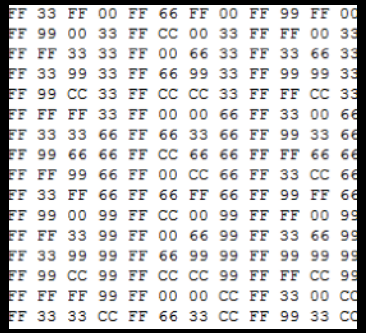


(GPT-3)







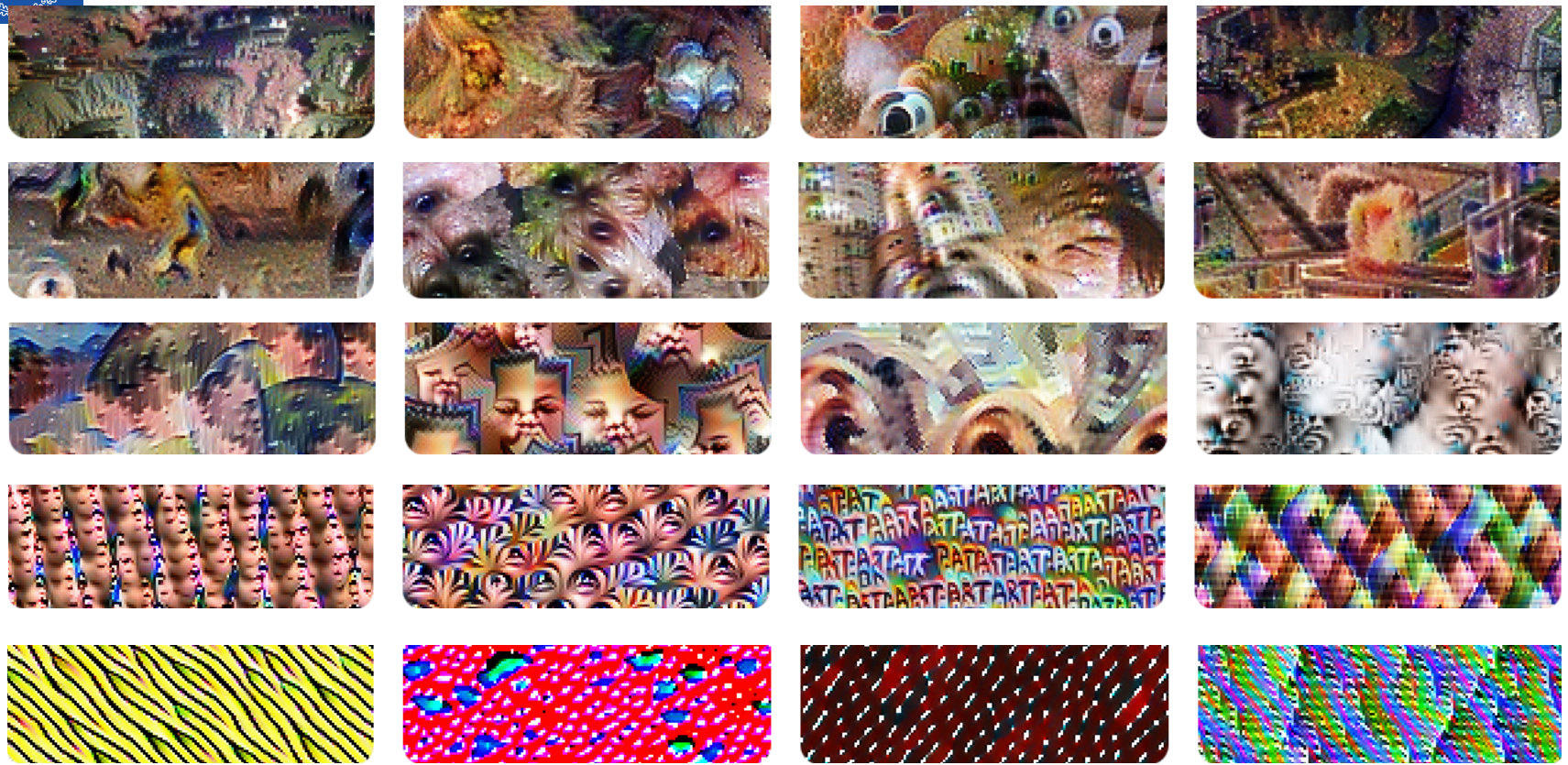


Mohammad Ali

Joe Frazier

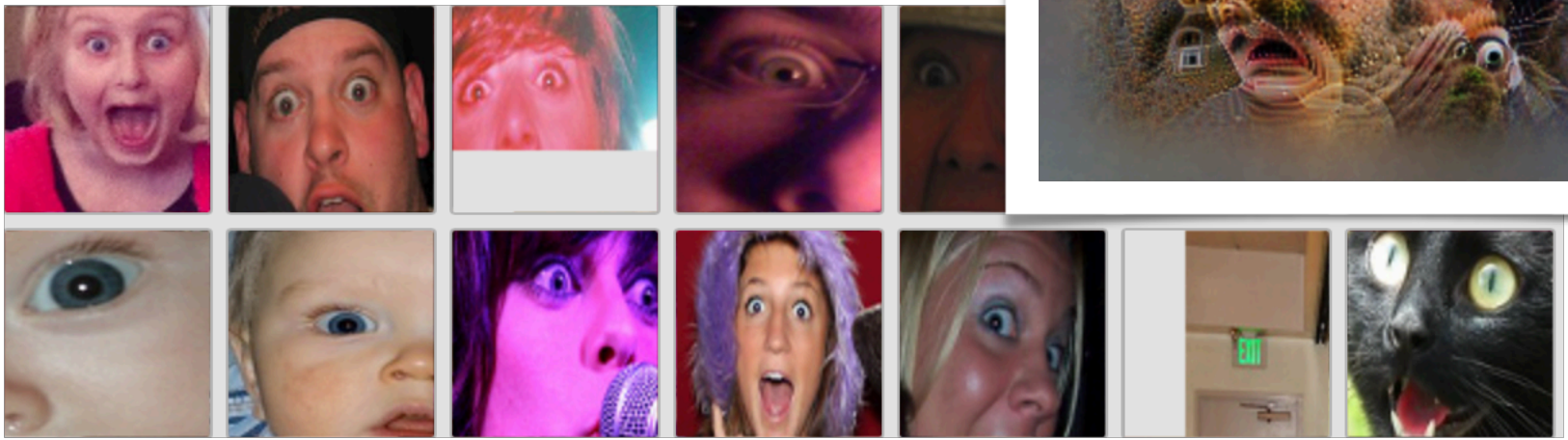
Sugar Ray Robinson

Identifying "Shock" with Deep Neural Networks



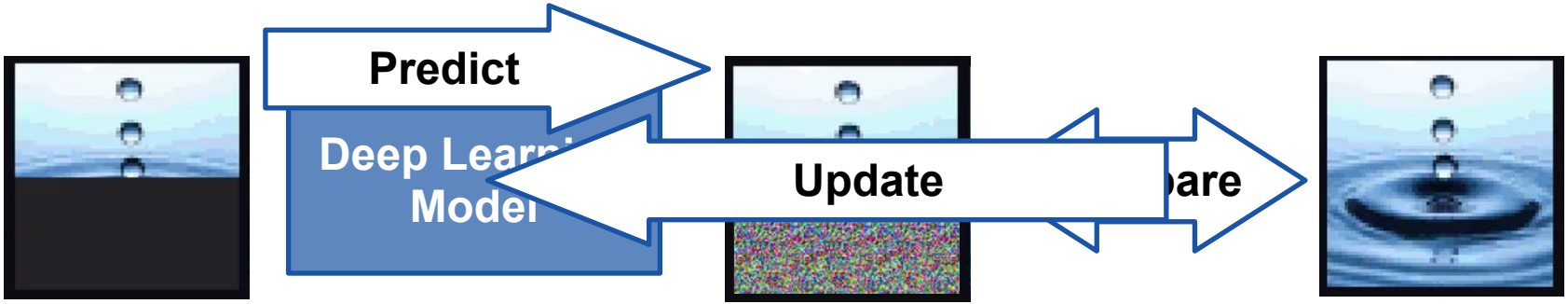
Goh, Gabriel, et al. "Multimodal Neurons in Artificial Neural Networks." Distill 6.3 (2021): e30.

Identifying "Shock" with Deep Neural Networks

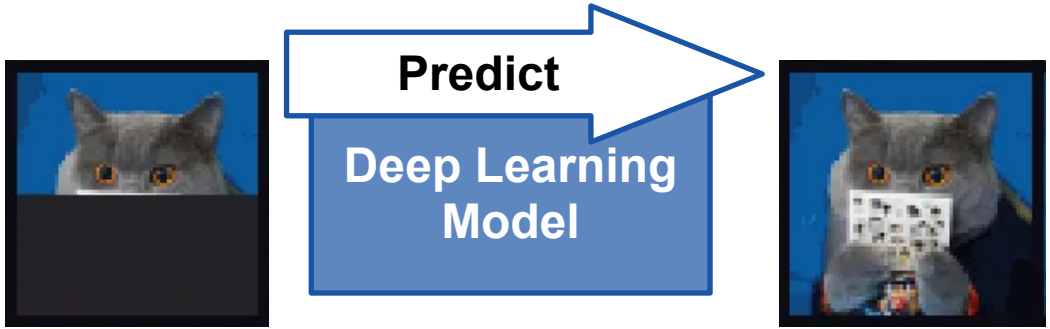


Goh, Gabriel, et al. "Multimodal Neurons in Artificial Neural Networks." Distill 6.3 (2021): e30.

Self-supervised training



Self-supervised learning

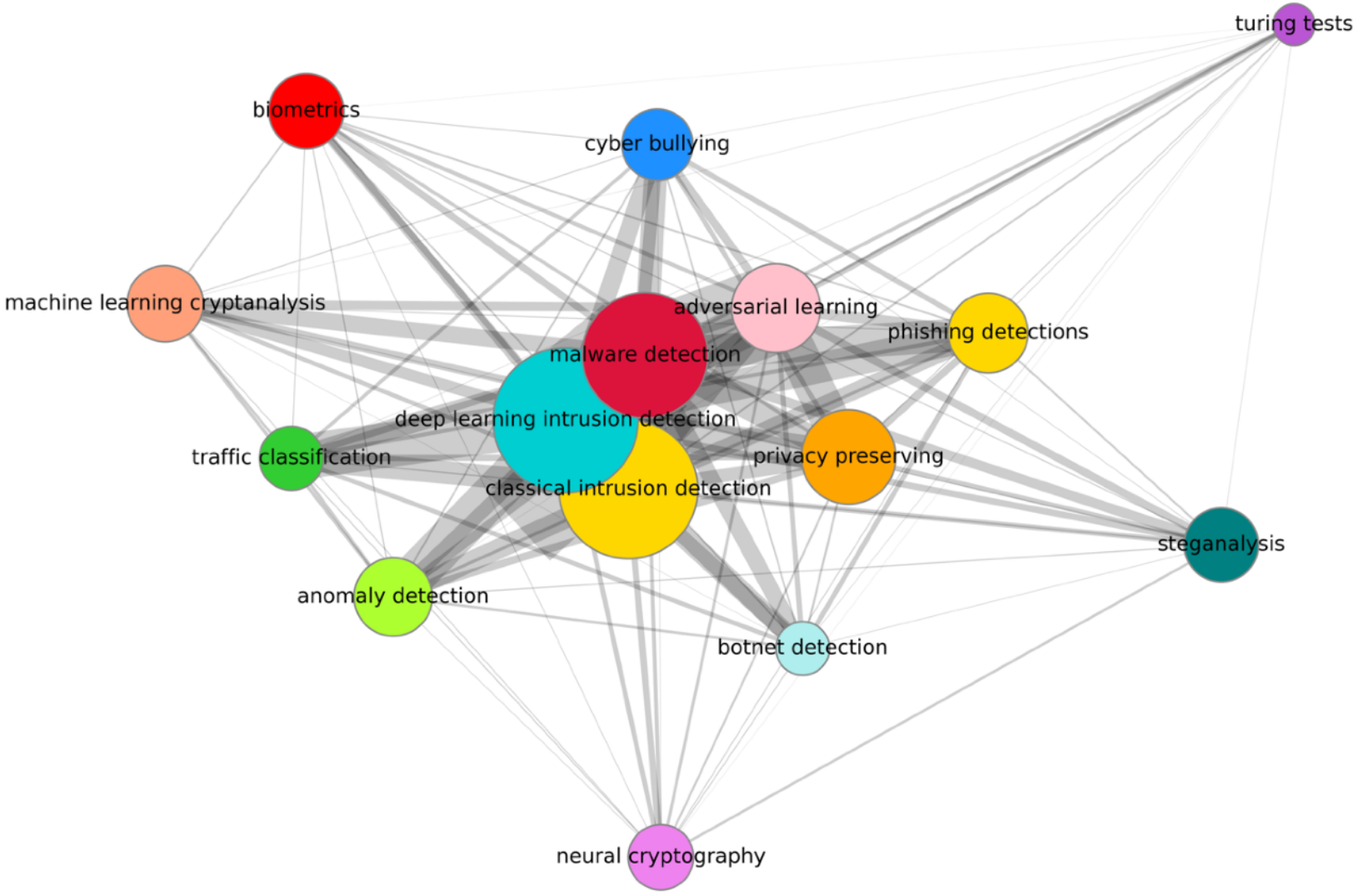


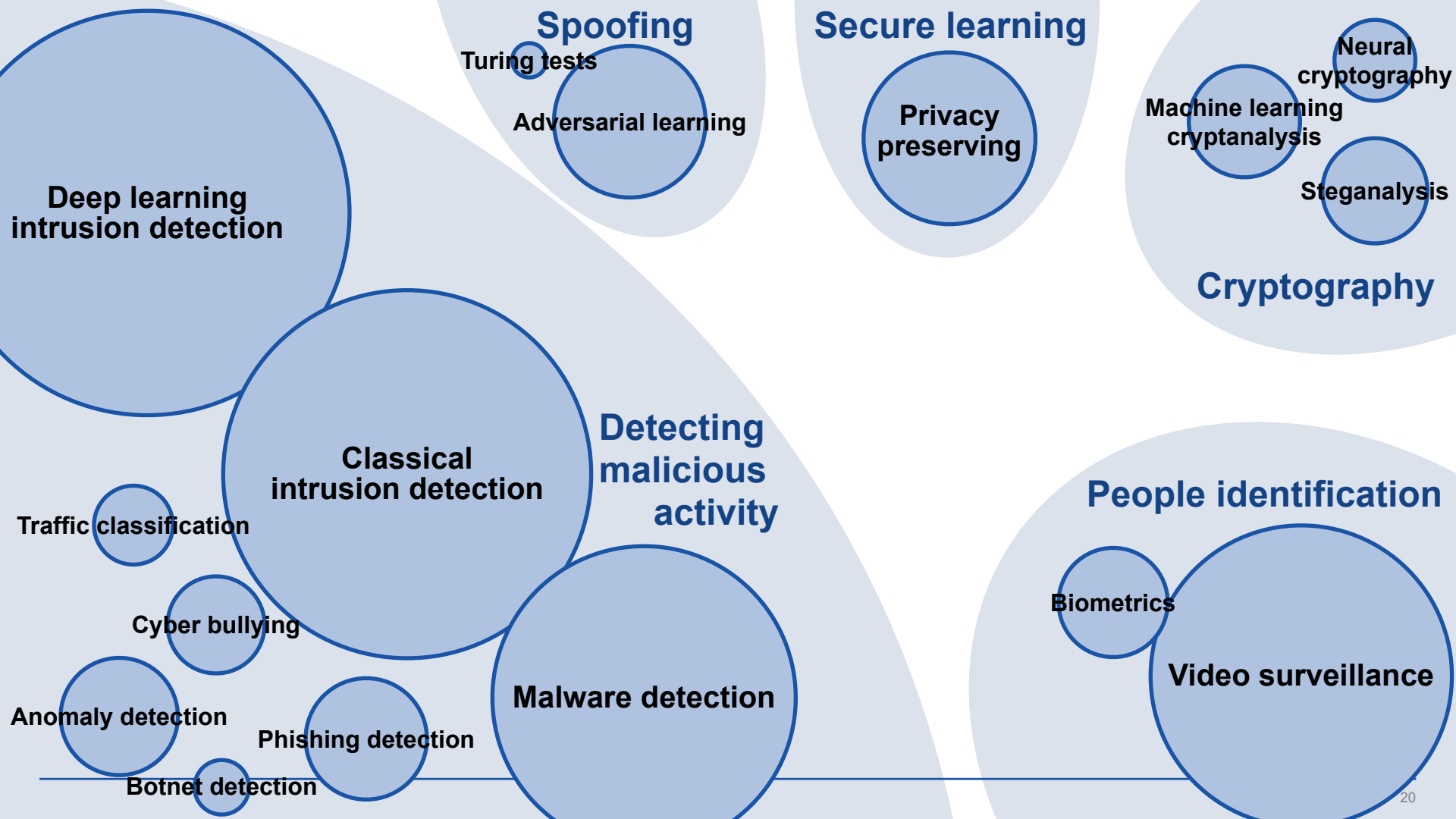
Scientific literature review

- Cyber security and machine learning

- Number of completely captured articles is 25,330
- Number of authors is 60,062
- Number of keywords is 75004
- Number of referenced articles is 366,601

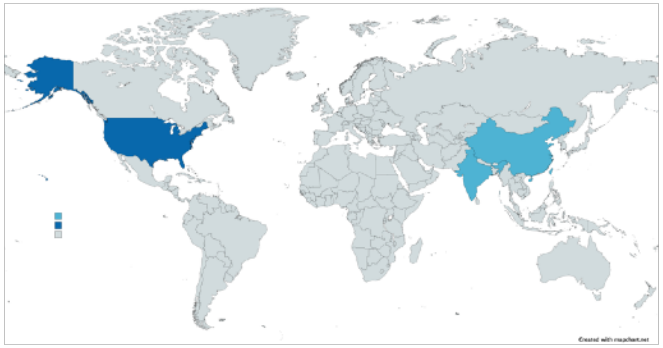
Scopus


turing tests
biometrics
cyber bullying
machine learning cryptanalysis
adversarial learning
phishing detections
malware detection
deep learning intrusion detection
traffic classification
privacy preserving
classical intrusion detection
steganalysis
anomaly detection
botnet detection
neural cryptography


Spoofing
Turing tests
Adversarial learning
Secure learning
Privacy preserving
Neural cryptography
Machine learning cryptanalysis
Steganalysis
Cryptography
Deep learning intrusion detection
Classical intrusion detection
Detecting malicious activity
Traffic classification
Cyber bullying
Anomaly detection
Phishing detection
Botnet detection
Malware detection
People identification
Biometrics
Video surveillance

Classical intrusion detection

intrusion detection systems, intrusion detection, machine learning, anomaly detection, network intrusion detection, machine learning techniques, network security, intrusion detection system, data sets, detection rates, neural networks, machine-learning, feature selection, in-network, support vector machines, artificial neural networks, network attack, false alarm rate, false positive rates, support vector machine

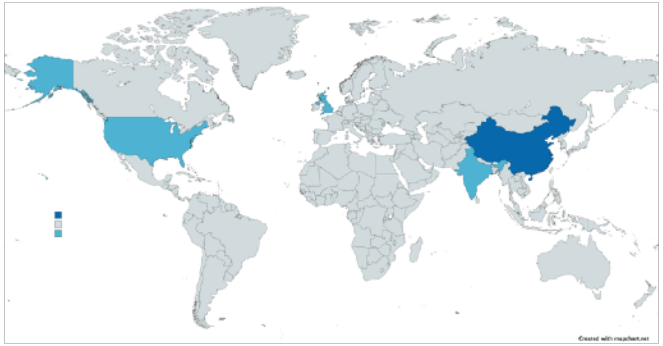


New Mexico Institute of Mining and Technology
Lawrence Berkeley National Laboratory
City University of Hong Kong
SRI International

Denning, An Intrusion-Detection Model, 1987
Sommer, Outside the closed world: On using machine learning for network intrusion detection, 2010
Mukkamala, Intrusion detection using neural networks and support vector machines, 2002
Tsai, Intrusion detection by machine learning: A review, 2009
Lippmann, Improving intrusion detection performance using keyword selection and neural networks, 2000
Wang, A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering, 2010
Wu, The use of computational intelligence in intrusion detection systems: A review, 2010
Debar, A neural network component for an intrusion detection system, 1992

Deep learning intrusion detection

intrusion detection systems, intrusion detection, deep learning, machine learning, network intrusion detection, cyber security, anomaly detection, network security, network intrusion detection systems, machine learning techniques, machine learning methods, intrusion detection system, convolutional neural network, deep neural networks, cyber-attacks, internet of things (iot), auto encoders, security, internet of thing (iot), classification models

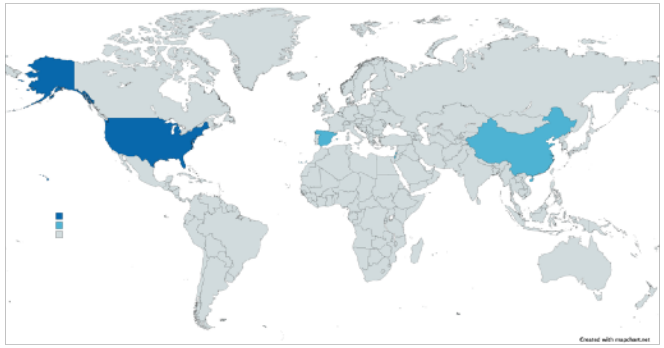


University of Leeds
University of New Brunswick
The Johns Hopkins University
Amrita School of Engineering

Buczak, A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection, 2016
Sharafaldin, Toward generating a new intrusion detection dataset and intrusion traffic characterization, 2018
Yin, A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks, 2017
Shone, A Deep Learning Approach to Network Intrusion Detection, 2018
Tang, Deep learning approach for Network Intrusion Detection in Software Defined Networking, 2016
Kim, Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection, 2016
Vinayakumar, Deep Learning Approach for Intelligent Intrusion Detection System, 2019
Gao, An Intrusion Detection Model Based on Deep Belief Networks, 2015

Malware detection

malware detection, machine learning, android malware, malware, android, deep learning, malware classifications, malwares, machine learning techniques, convolutional neural network, security, android applications, malware classification, machine-learning, classification, false positive rates, malware families, computer security, classification accuracy, static analysis

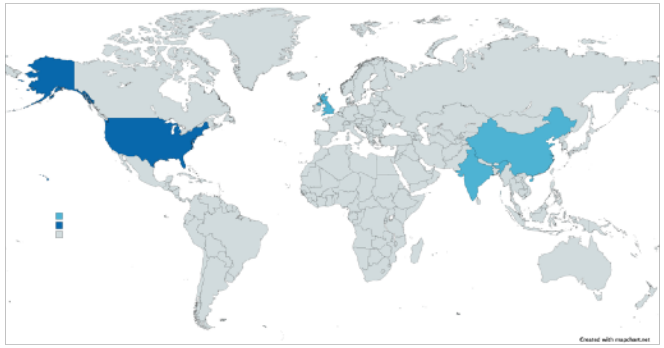


Universidad de Deusto
Ben-Gurion University of the Negev
Queen's University Belfast
Tsinghua University

Saxe, Deep neural network based malware detection using two dimensional binary program features, 2016
Shabtai, Andromaly: A behavioral malware detection framework for android devices, 2012
Biggio, Evasion attacks against machine learning at test time, 2013
Dahl, Large-scale malware classification using random projections and neural networks, 2013
Kolosnjaji, Deep learning for classification of malware system call sequences, 2016
Yuan, Droiddetector: Android malware characterization and detection using deep learning, 2016
Santos, Opcode sequences as representation of executables for data-mining-based unknown malware detection, 2013
Yuan, Droid-Sec: Deep learning in android malware detection, 2015

Phishing detection

machine learning, phishing, machine learning techniques, phishing detections, false positive rates, anti-phishing, machine-learning, phishing detection, generation algorithm, cyber security, sensitive informations, phishing attacks, third party services, classification, deep learning, detection methods, phishing websites, random forests, comparative studies, web page

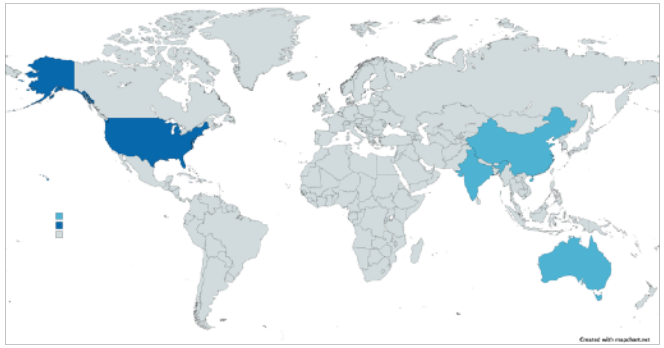


Carnegie Mellon University
Southern Methodist University
University of Washington, Tacoma
University of Northumbria

Xiang, CANTINA+: A feature-rich machine learning framework for detecting phishing web sites, 2011
Abu-Nimeh, A comparison of machine learning techniques for phishing detection, 2007
Mohammad, Predicting phishing websites based on self-structuring neural network, 2014
Sahingoz, Machine learning based phishing detection from URLs, 2019
Blum, Lexical feature based phishing URL detection using online learning, 2010
Tran, A LSTM based framework for handling multiclass imbalance in DGA botnet detection, 2018
Yu, Character Level based Detection of DGA Domain Names, 2018
Almomani, A survey of phishing email filtering techniques, 2013

Insiders, cyber bullies and spammers

machine learning, cyber bullying, cyber security, machine learning techniques, software vulnerabilities, deep learning, insider threat, cyberbullying, data mining, convolutional neural network, network security, on-line social networks, spam detection, vulnerability detection, supervised machine learning, anomaly detection, statistical features, machine-learning, social network security, data analytics

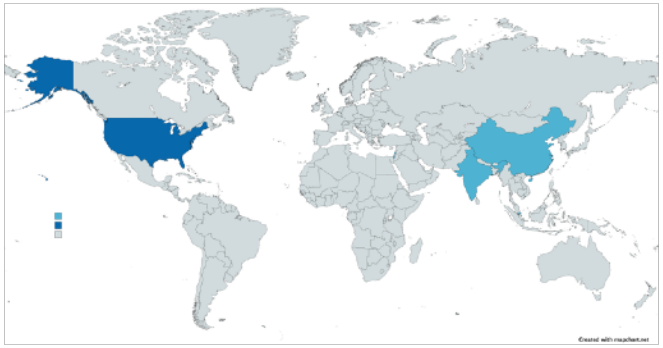


Deakin University
University of Maryland
Swinburne University of Technology
Nanyang Technological University

Liu, Detecting and Preventing Cyber Insider Threats: A Survey, 2018
Wang, Don't follow me - Spam detection in twitter, 2010
Grieco, Toward Large-Scale Vulnerability Discovery using Machine Learning, 2016
Lee, Uncovering social spammers: Social honeypots + machine learning, 2010
Chavan, Machine learning approach for detection of cyber-aggressive comments by peers on social media network, 2015
Zhang, An effective network traffic classification method with unknown flow detection, 2013
Dinakar, Common sense reasoning for detection, prevention, and mitigation of cyberbullying, 2015
Al-Garadi, Cybercrime detection in online communications: The experimental case of cyberbullying detection in the Twitter network, 2016

Cyber-physical anomaly detection

machine learning, anomaly detection, cyber security, false data injection, intrusion detection systems, intrusion detection, machine learning techniques, false data injection attacks, cyber-attacks, smart grid, industrial control systems, sdn, deep learning, network security, bad data detections, machine learning methods, software defined networking (sdn), supervisory control and data acquisition, network intrusion detection, distributed denial of service attack

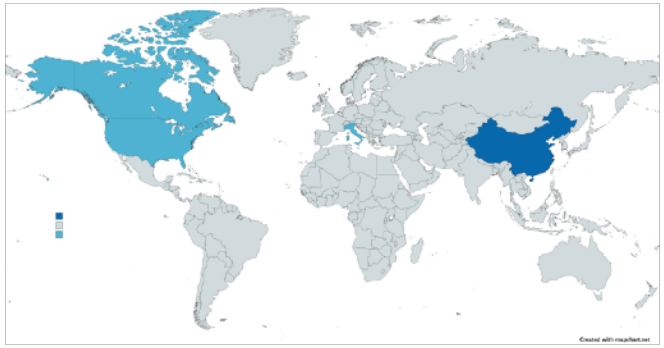


University of Texas at Austin
Singapore University of Technology and Design
Oak Ridge National Laboratory'
KLE Technological University

Sinclair, An application of machine learning to network intrusion detection, 1999
He, Real-Time Detection of False Data Injection Attacks in Smart Grid: A Deep Learning-Based Intelligent Mechanism, 2017
Esmalifalak, Detecting stealthy false data injection using machine learning in smart grid, 2017
Goh, Anomaly detection in cyber physical systems using recurrent neural networks, 2017
Beaver, An evaluation of machine learning methods to detect malicious SCADA communications, 2013
Barki, Detection of distributed denial of service attacks in software defined networks, 2016
Kravchik, Detecting cyber attacks in industrial control systems using convolutional neural networks, 2018
Latah, Towards an efficient anomaly-based intrusion detection for software-defined networks, 2018

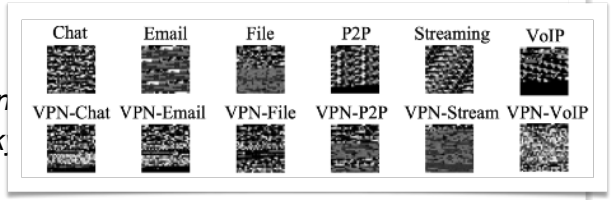
Traffic classification

traffic classification, convolutional neural network, encrypted traffic, machine learning, intrusion detection systems, representation learning, deep learning, encrypted traffic classification, network anomaly detection, network intrusion detection, intrusion detection system, state-of-the-art methods, feature engineerings, network traffic classification, end to end, machine learning techniques, end-to-end, deep neural networks, in-network management, anomalous behavior



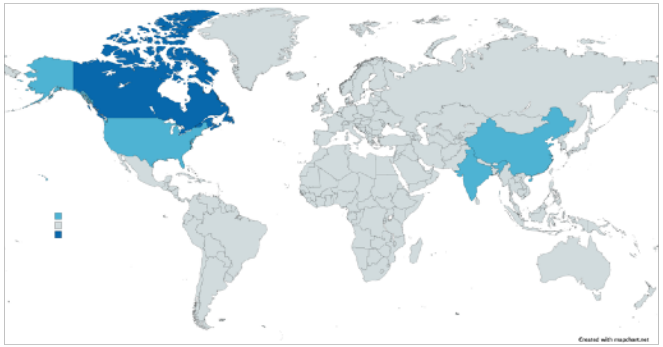
Chinese Academy of Sciences
University of Science and Technology of China
Dalhousie University
Beijing University of Posts and Telecommunications

Wang, Malware traffic classification using convolutional neural network for representation learning, 2017
Wang, End-To-end encrypted traffic classification with one-dimensional convolution neural networks, 2017
Wang, HAST-IDS: Learning Hierarchical Spatial-Temporal Features Using Deep Neural Networks to Improve Intrusion Detection, 2017
Draper-Gil, Characterization of encrypted and VPN traffic using time-related features, 2016
Velan, A survey of methods for encrypted traffic classification and analysis, 2015
Lashkari, Characterization of tor traffic using time based features, 2017
Lotfollahi, Deep packet: a novel approach for encrypted traffic classification using deep learn
Alshammari, Machine learning based encrypted traffic classification: Identifying SSH and Sk

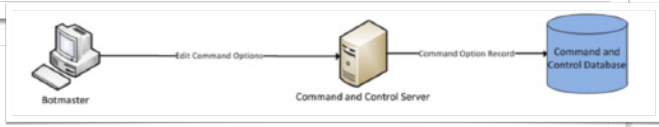


Botnet detection

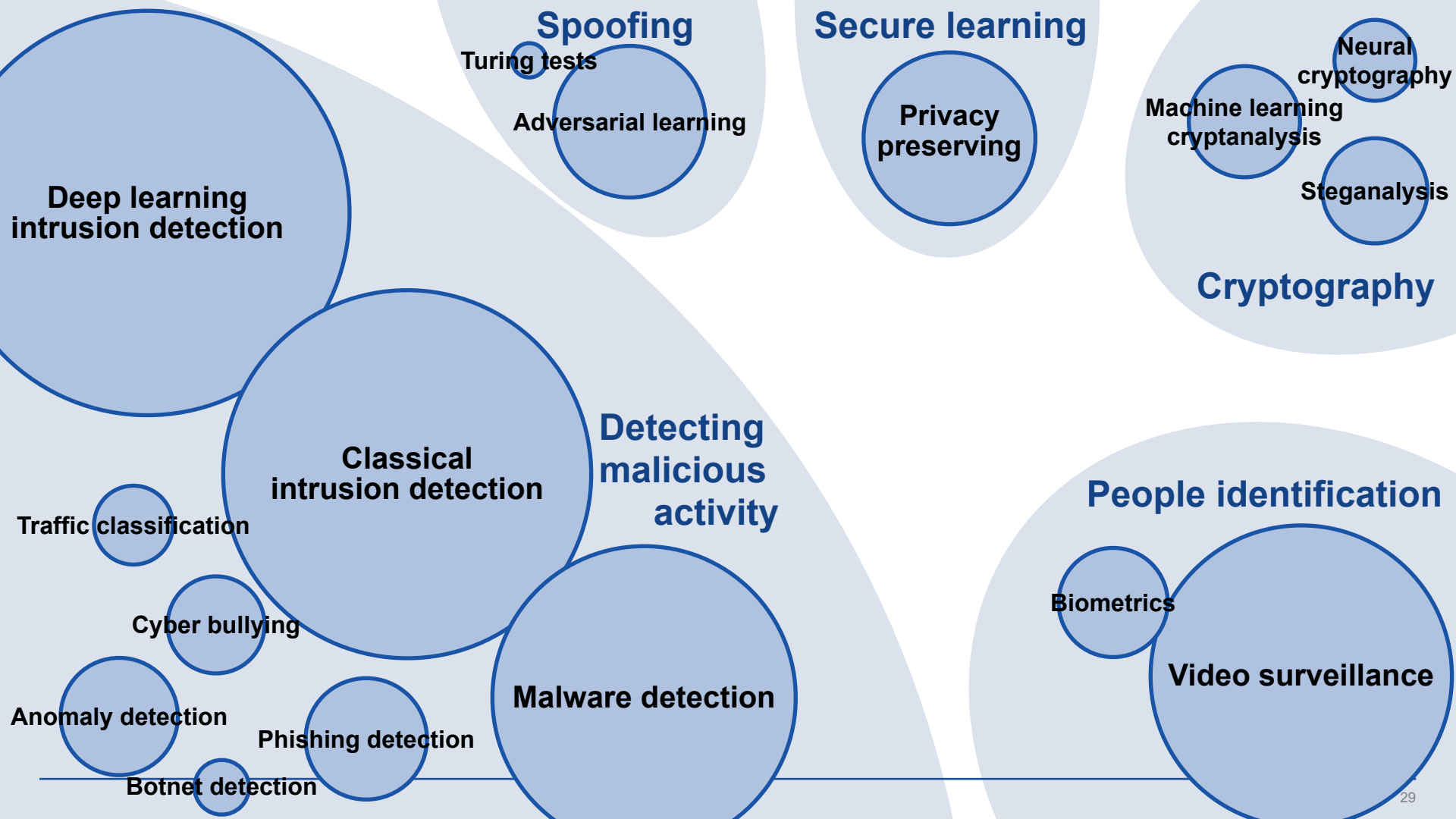
machine learning, botnet detections, network traffic, traffic behavior, machine learning techniques, botnets, botnet, intrusion detection, network flows, botnet detection, experimental evaluation, ddos attack, traffic behavior analysis, complete networks, novelty detection, cyber-crimes, network behaviors, data sets, sensitive informations, early detection



University of New Brunswick
University of Victoria
BBN Technologies
Birla Institute of Technology and Science

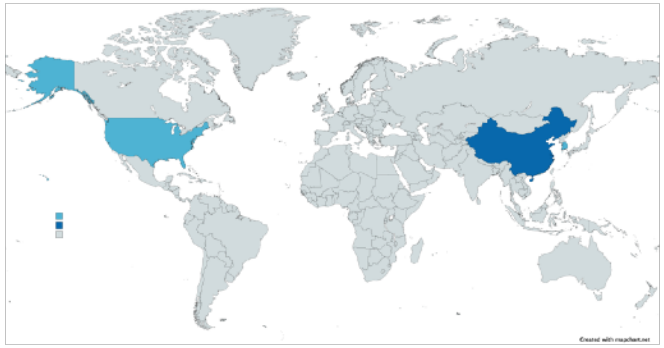


Zhao, Botnet detection based on traffic behavior analysis and flow intervals, 2013
Livadas, Using machine learning techniques to identify botnet traffic, 2006
Saad, Detecting P2P botnets through network behavior analysis and machine learning, 2011
Beigi, Towards effective feature selection in machine learning-based botnet detection approaches, 2014
Ahmed, A new biometrie technology based on mouse dynamics, 2007
Singh, Big Data Analytics framework for Peer-to-Peer Botnet detection using Random Forests, 2014
Veeramachaneni, AI2: Training a Big Data Machine to Defend, 2016
Alauthaman, A P2P Botnet detection scheme based on decision tree and adaptive multilayer neural networks, 2018


Spoofing
Turing tests
Adversarial learning
Secure learning
Privacy preserving
Neural cryptography
Machine learning cryptanalysis
Steganalysis
Cryptography
Deep learning intrusion detection
Classical intrusion detection
Detecting malicious activity
Traffic classification
Cyber bullying
Anomaly detection
Phishing detection
Botnet detection
Malware detection
People identification
Biometrics
Video surveillance

Video surveillance

convolutional neural network, person re identifications, video surveillance, deep learning, person re-identification, feature representation, surveillance video, state-of-the-art performance, discriminative features, convolutional neural networks, state-of-the-art methods, re identifications, surveillance cameras, state-of-the-art approach, anomaly detection, intelligent video surveillance, action recognition, video surveillance systems, convolutional neural networks (cnn), human-action recognition

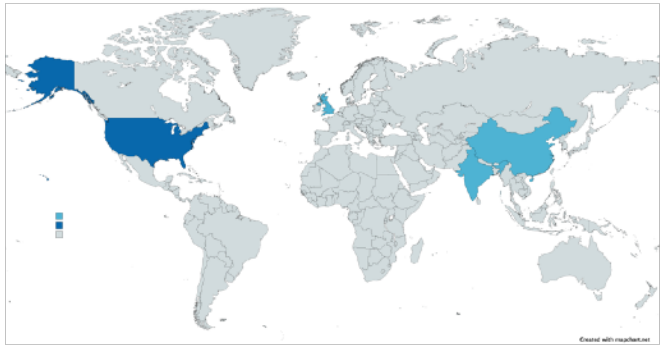


NEC Laboratories America, Inc.
SenseTime Group Limited
Institute of Automation Chinese Academy of Sciences
University of Science and Technology of China

Ji, 3D Convolutional neural networks for human action recognition, 2013
Zhang, An anomaly detection model for network intrusions using one-class SVM and scaling strategy, 2016
Varior, Gated siamese convolutional neural network architecture for human re-identification, 2016
Sultani, Real-World Anomaly Detection in Surveillance Videos, 2018
Ding, Deep feature learning with relative distance comparison for person re-identification, 2015
Liu, A hierarchical intrusion detection model based on the PCA neural networks, 2007
Chen, Beyond triplet loss: A deep quadruplet network for person re-identification, 2017
Zhao, Spindle net: Person re-identification with human body region guided feature decomposition and fusion, 2017

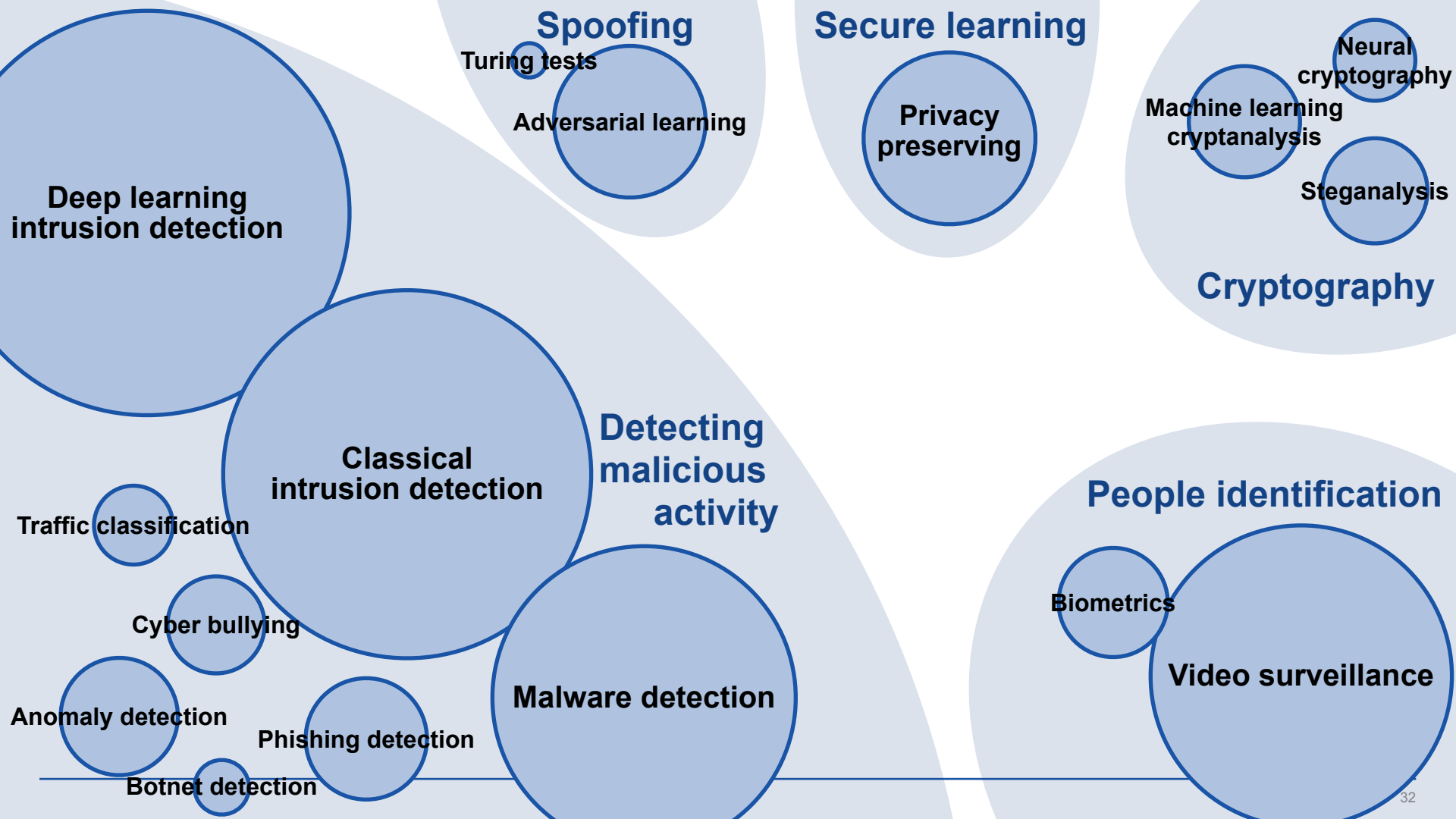
Biometrics

biometrics, keystroke dynamics, user authentication, machine learning, authentication, convolutional neural network, neural networks, authentication systems, biometric authentication, pattern recognition, security, continuous authentications, machine learning techniques, fingerprint liveness detection, liveness detection, biometric authentication system, personal authentication, multimodal biometrics, equal error rate, fingerprint recognition



Universidade Estadual de Campinas
City College of New York
Michigan State University
University of Plymouth

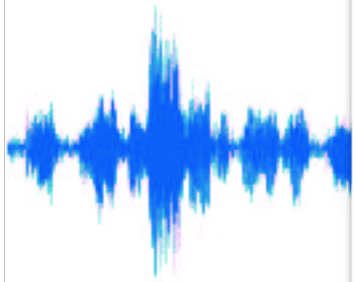
Obaidat, Verification of computer users using keystroke dynamics, 1997
Ahmed, A new biometrie technology based on mouse dynamics, 2007
Nogueira, Fingerprint Liveness Detection Using Convolutional Neural Networks, 2016
Karnan, Biometric personal authentication using keystroke dynamics: A review, 2011
Marcel, Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation, 2007
Lin, Computer-access authentication with neural network based keystroke identity verification, 1997
Brown, Integrated detection and segmentation for hyperspectral imagery using neural networks, 1993
Crawford, Keystroke dynamics: Characteristics and opportunities, 2010


Spoofing
Turing tests
Adversarial learning
Secure learning
Privacy preserving
Neural cryptography
Machine learning cryptanalysis
Steganalysis
Cryptography
Deep learning intrusion detection
Classical intrusion detection
Detecting malicious activity
Traffic classification
Cyber bullying
Anomaly detection
Phishing detection
Botnet detection
Malware detection
People identification
Biometrics
Video surveillance

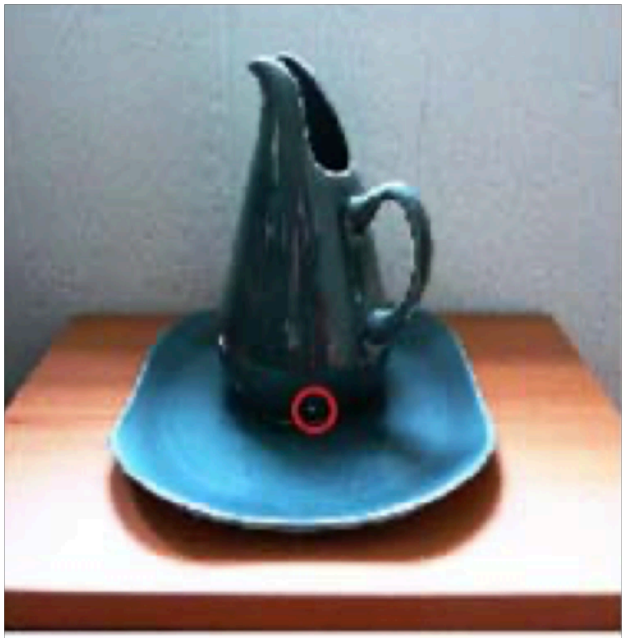
A
A

machi
learnir
securi
evasic
defens
learnir
theory
classif

'Duck'

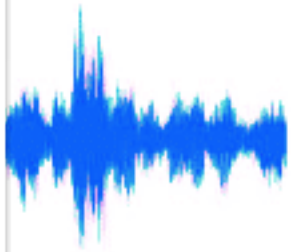


'How are you?'



Teapot(24.99%)
Joystick(37.39%)

'Horse'



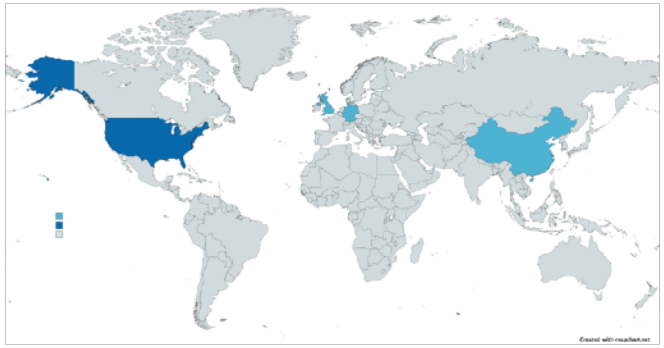
'pen the door'

fornia, Berkeley
Studi di Cagliari
ellon University
State University

Paperno
Biggio, E
Huang, A
Barreno,
Barreno,
Sharif, A
Grosse,
Dong, Boosting Adversarial Attacks with Momentum, 2018

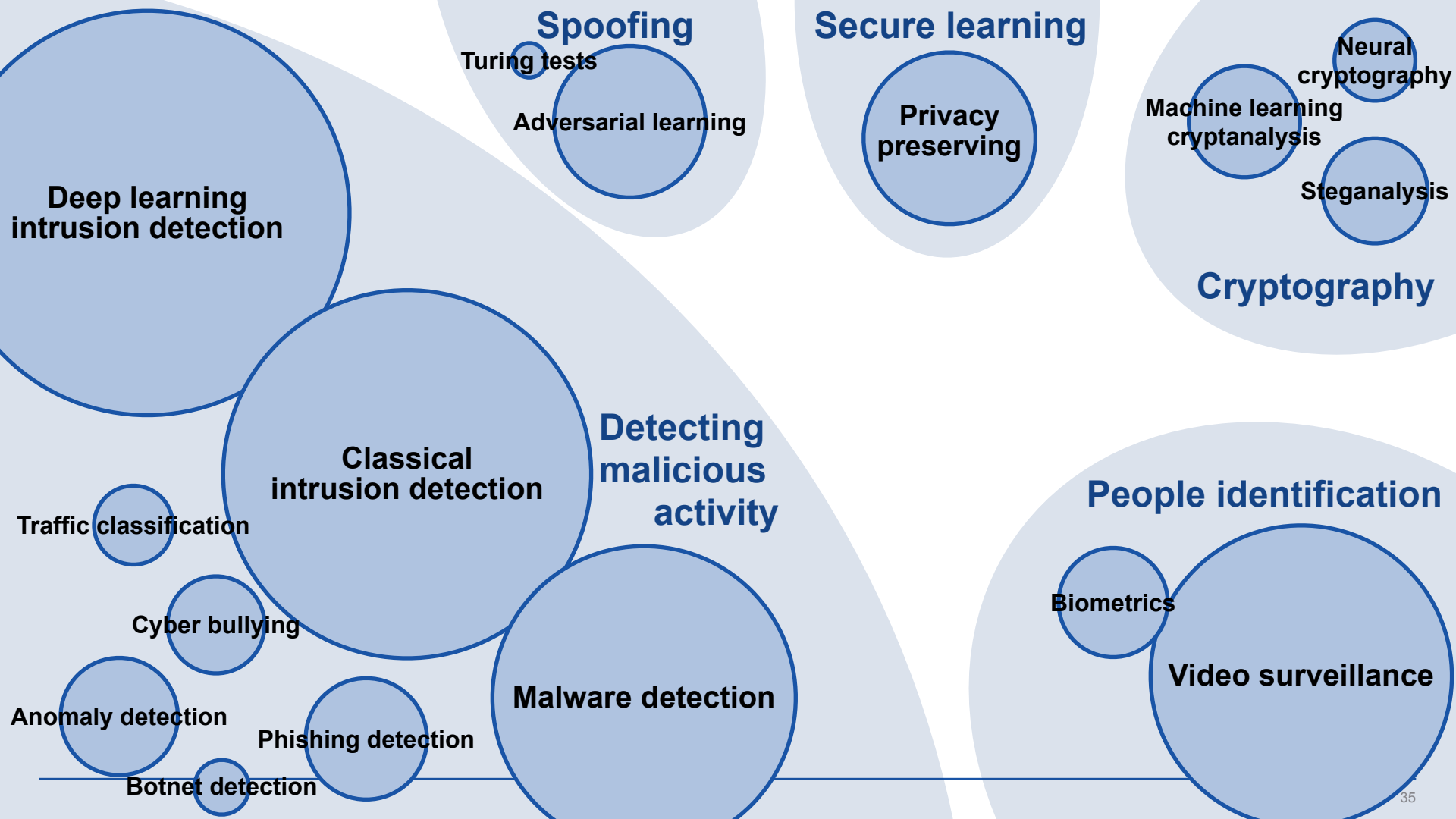
Turing tests

captchas, captcha, turing tests, security, machine learning, novel construction, automated test, win-win, hard problems, deep learning, machine-learning, reverse turing test, text recognition, web registration forms, distorted text, convolutional neural network, knn classifier, svm, k-nn classifier, vision algorithm



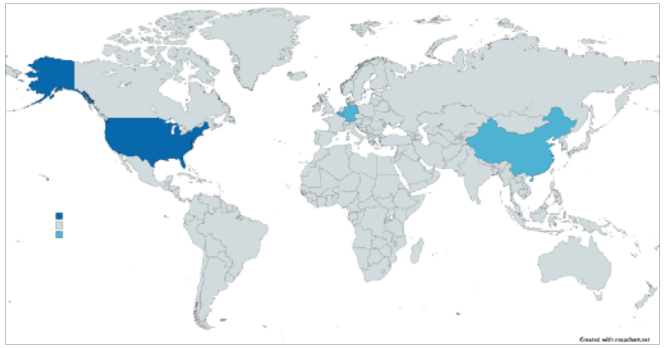
Carnegie Mellon University
Stanford University
Xidian University
IBM Thomas J. Watson Research Center

Von Ahn, CAPTCHA: Using hard AI problems for security, 2003
Von Ahn, Telling humans and computers apart automatically, 2004
Bursztein, Text-based CAPTCHA strengths and weaknesses, 2011
Yan, Breaking visual CAPTCHAs with naïve pattern recognition algorithms, 2007
Rui, Artifacial: Automated reverse turing test using FACIAL features, 2003
Osadchy, No Bot Expects the DeepCAPTCHA! Introducing Immutable Adversarial Examples, 2017
Tam, Breaking audio CAPTCHAs, 2009
Gao, The robustness of hollow CAPTCHAs, 2013


Spoofing
Turing tests
Adversarial learning
Secure learning
Privacy preserving
Neural cryptography
Machine learning cryptanalysis
Steganalysis
Cryptography
Deep learning intrusion detection
Classical intrusion detection
Detecting malicious activity
Traffic classification
Cyber bullying
Anomaly detection
Botnet detection
Phishing detection
Malware detection
People identification
Biometrics
Video surveillance

Machine learning cryptanalysis

machine learning, machine learning techniques, challenge-response pair, physical unclonable functions, evolution strategies, logistic regressions, cryptanalysis, ring oscillator, physical cryptography, malware detection, machine-learning, fpgas and asics, hardware performance counters, new design, computer algorithm, numerical modeling, feed-forward, performance counters, anti virus, antivirus softwares

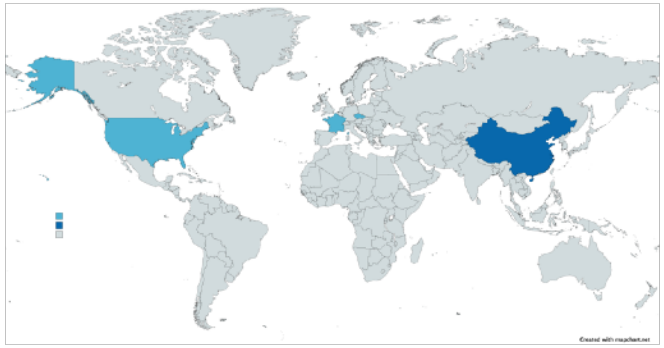


Technical University of Munich
Columbia University
George Mason University
Massachusetts Institute of Technology

Rührmair, Modeling attacks on physical unclonable functions, 2010
Demme, On the feasibility of online malware detection with performance counters, 2013
Ruhrmair, PUF modeling attacks on simulated and silicon data, 2013
Dodis, Fuzzy extractors: How to generate strong keys from biometrics and other noisy data, 2008
Tang, Unsupervised anomaly-based malware detection using hardware features, 2014
Zhang, Color Image Encryption Algorithm Based on TD-ERCS System and Wavelet Neural Network, 2015
Hospodar, Machine learning in side-channel analysis: A first study, 2011
Maghrebi, Breaking cryptographic implementations using deep learning techniques, 2016

Steganalysis

steganalysis, convolutional neural network, deep learning, steganography, steganographic algorithms, convolutional neural networks, feature learning, ensemble classifiers, detection performance, forensics, convolutional neural networks (cnn), gaussian nonlinearity, image database, steganographic system, source mismatch, feature representation, image steganalysis, detection accuracy, training and testing, high-dimensional

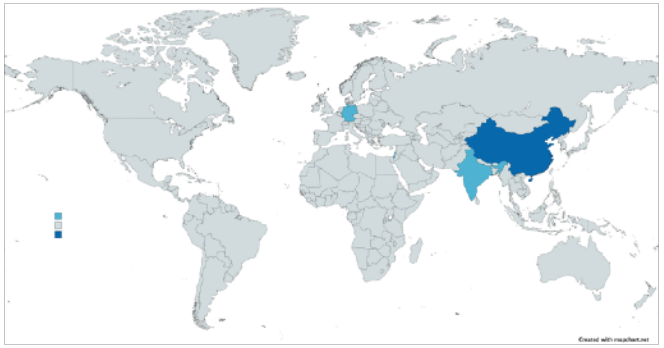


Binghamton University State University of New York
Shenzhen University
Institute of Automation Chinese Academy of Sciences
Sun Yat-Sen University

Qian, Deep learning for steganalysis via convolutional neural networks, 2015
Bas, Break our steganographic system: The ins and outs of organizing BOSS, 2011
Xu, Structural design of convolutional neural networks for steganalysis, 2016
Kodovský, Ensemble classifiers for steganalysis of digital media, 2012
Ye, Deep Learning Hierarchical Representations for Image Steganalysis, 2017
Tan, Stacked convolutional auto-encoders for steganalysis of digital images, 2014
Xu, Deep convolutional neural network to detect J-UNIWARD, 2017
Pibre, Deep learning is a good steganalysis tool when embedding key is reused for different images [...], 2016

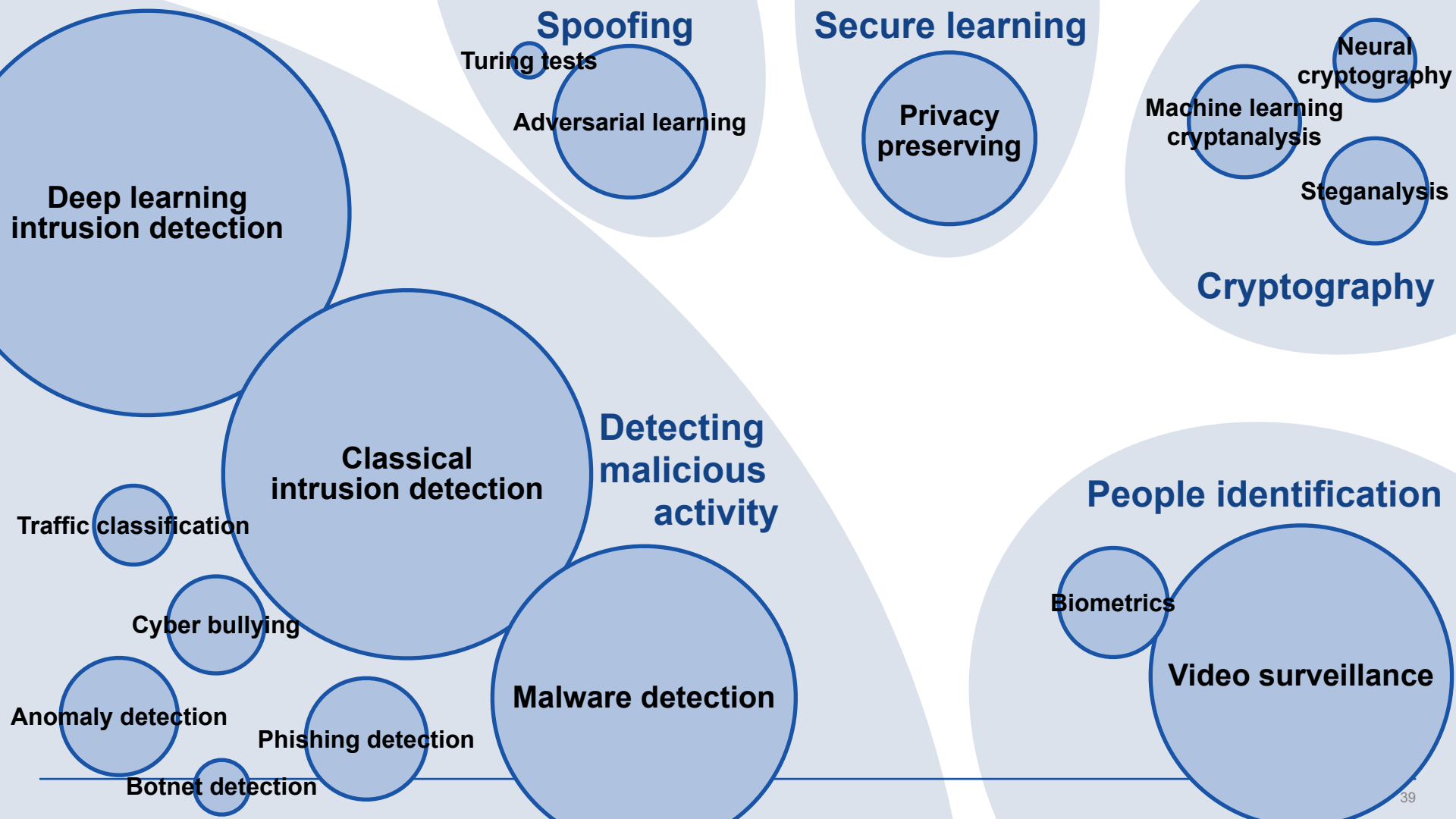
Neural cryptography

neural cryptography, image encryptions, neural network, cryptography, neural networks, mutual learning, chaotic neural network, encryption algorithms, image encryption, image encryption algorithm, algorithms, convolutional neural network, encryption, artificial neural network, cryptographic schemes, synchronization, security, tree parity machine, public keys, key exchange protocols



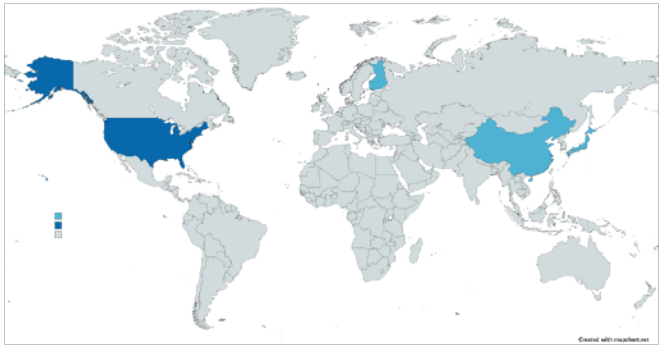
Bar-Ilan University
Weizmann Institute of Science Israel'
City University of Hong Kong
Universität Heidelber

Klimov, Analysis of neural cryptography, 2002
Lian, A block cipher based on chaotic neural networks, 2009
Guo, A New Symmetric Probabilistic Encryption Scheme Based on Chaotic Attractors of Neural Networks, 1999
Kinzel, Neural cryptography, 2002
Rosen-Zvi, Mutual learning in a tree parity machine and its application to cryptography, 2002
Karras, On neural network techniques in the secure management of communication systems, 2003
Ruttor, Genetic attack on neural cryptography, 2006
Biham, Differential crypt analysis of DES-like cryptosystems, 1991


Spoofing
Turing tests
Adversarial learning
Secure learning
Privacy preserving
Neural cryptography
Machine learning cryptanalysis
Steganalysis
Cryptography
Deep learning intrusion detection
Classical intrusion detection
Detecting malicious activity
Traffic classification
Cyber bullying
Anomaly detection
Phishing detection
Botnet detection
Malware detection
People identification
Biometrics
Video surveillance

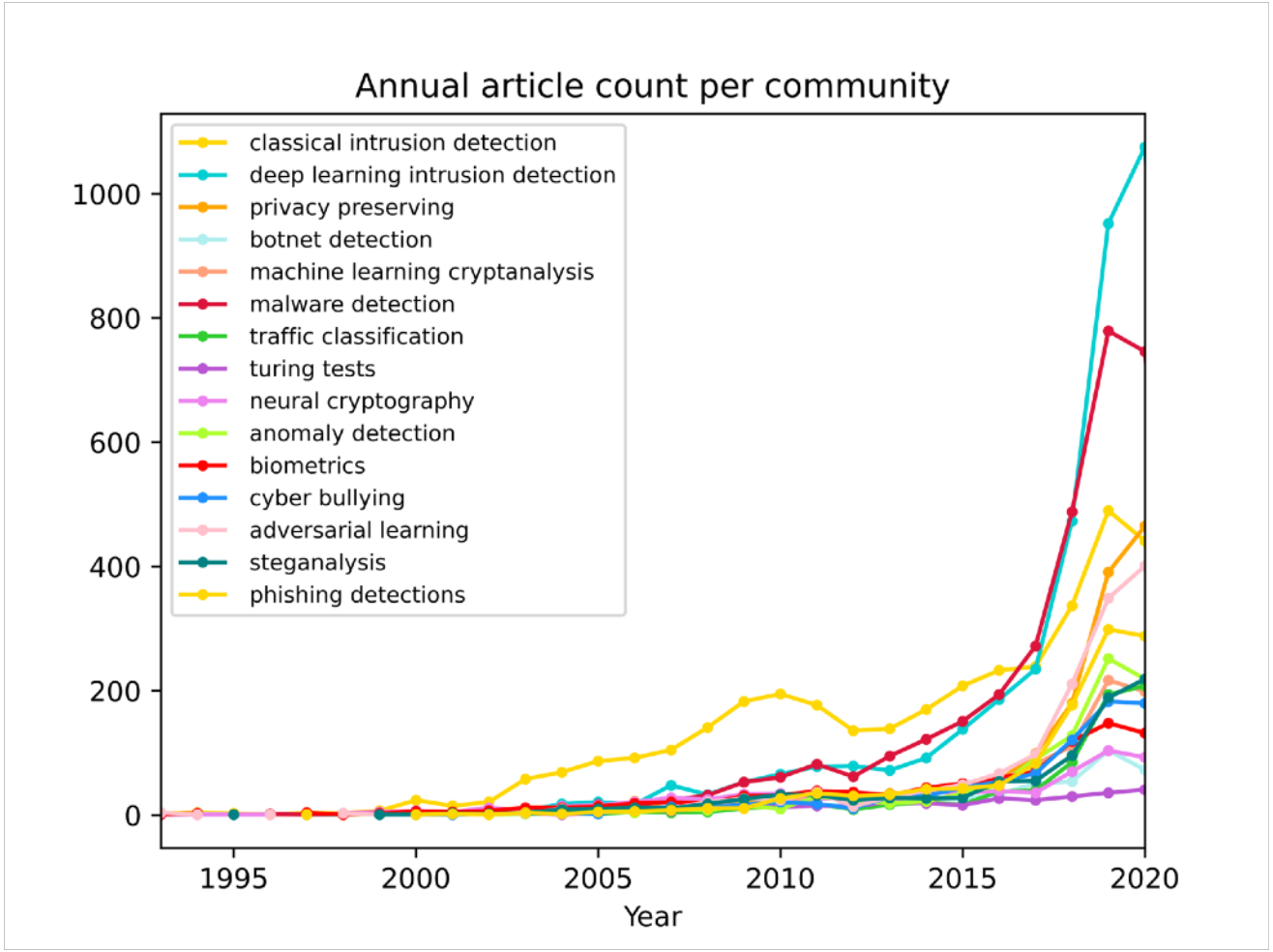
Privacy preserving

privacy preserving, machine learning, privacy, homomorphic encryptions, deep learning, federated learning, machine learning models, communication overheads, neural networks, differential privacies, cloud computing, privacy-preserving protocols, encrypted data, homomorphic encryption, cloud services, high dimensional data, secure aggregation, learning settings, complexity analysis, secure aggregations



Microsoft Research
Google LLC
Japan National Institute of ICT
Guangzhou University

Liu, Oblivious neural network predictions via MiniONN transformations, 2017
Bonawitz, Practical secure aggregation for privacy-preserving machine learning, 2017
Li, Significant Permission Identification for Machine-Learning-Based Android Malware Detection, 2018
Xu, DeepRefiner: Multi-layer Android Malware Detection System Applying Deep Neural Networks, 2018
Juvekar, GAZELLE: A low latency framework for secure neural network inference, 2018
Hitaj, Deep Models under the GAN: Information leakage from collaborative deep learning, 2017
Dowlin, Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy, 2016
Fredrikson, Privacy in pharmacogenetics: An end-to-end case study of personalized warfarin dosing, 2014


Annual article count per community
classical intrusion detection
deep learning intrusion detection
privacy preserving
botnet detection
machine learning cryptanalysis
malware detection
traffic classification
turing tests
neural cryptography
anomaly detection
biometrics
cyber bullying
adversarial learning
steganalysis
phishing detections
1000
800
600
400
200
0
1995
2000
2005
2010
2015
2020
Year

Machine learning for offense

Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence

POMDPs Make Better Hackers: Accounting for Uncertainty in Penetration Testing

Carlos Sarraute
Core Security & ITBA
Buenos Aires, Argentina
carlos@coresecurity.com

Olivier Buffet
INRIA
Nancy, France
buffet@loria.fr

Jörg Hoffmann
Saarland University
Saarbrücken, Germany
hoffmann@cs.uni-saarland.de

Abstract

Penetration Testing is a methodology for assessing network security, by generating and executing possible hacking attacks. Doing so automatically allows for regular and systematic testing. A key question is how to generate the attacks. This is naturally formulated as planning under uncertainty, i.e., under incomplete knowledge about the network configuration. Previous work uses classical planning, and requires costly pre-processes reducing this uncertainty by extensive application of scanning methods. By contrast, we

done in Core Security's "Core Insight Enterprise" tool. We will use the term "attack planning" in that sense.

Lucangeli et al. (2010) encode attack planning into PDDL, and use off-the-shelf planners. This already is useful—in fact, it is currently employed commercially in Core Insight Enterprise, using a variant of Metric-FF (Hoffmann 2003). However, the approach is limited by its inability to handle uncertainty. The pentesting tool cannot be up-to-date regarding all the details of the configuration of every machine in the network, maintained by individual users

DeepSQLi: Deep Semantic Learning for Testing SQL Injection

Muyang Liu
University of Electronic Science and Technology of China
Chengdu, China
muyangl@foxmail.com

Ke Li*
University of Exeter
Exeter, UK
k.li@exeter.ac.uk

Tao Chen
Loughborough University
Loughborough, UK
t.t.chen@lboro.ac.uk

ABSTRACT

Security is unarguably the most serious concern for Web applications, to which SQL injection (SQLi) attack is one of the most devastating attacks. Automatically testing SQLi vulnerabilities is of ultimate importance, yet is unfortunately far from trivial to implement. This is because the existence of a huge, or potentially infinite, number of variants and semantic possibilities of SQL leading to SQLi attacks on various Web applications. In this paper, we propose a deep natural language processing based tool, dubbed `DeepSQLi`, to generate test cases for detecting SQLi vulnerabilities. Through adopting deep learning based neural language model and sequence of words prediction, `DeepSQLi` is equipped with the ability to learn the semantic knowledge embedded in SQLi attacks, allowing it to

18–22, 2020, Virtual Event, USA. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3395363.3397375

1 INTRODUCTION

Web applications have become increasingly ubiquitous and important since the ever prevalence of distributed computing paradigms, such as Cyber-Physical Systems and Internet-of-Things. Yet, they are unfortunately vulnerable to a variety of security threats, among which SQL injection (SQLi) has been widely recognised as one of the most devastating threats. Generally speaking, SQLi is an injection attack that embeds scripts in user inputs to execute malicious SQL statements over the relational database management system (RDBMS) running behind a Web application. As stated in the Akamai report[1], SQLi attacks constituted 65.1% of the cyber attacks

PassGAN: A Deep Learning Approach for Password Guessing*

Briland Hitaj
Stevens Institute of Technology
bhitaj@stevens.edu

Paolo Gasti
New York Institute of Technology
pgasti@nyit.edu

Giuseppe Ateniese
Stevens Institute of Technology
gatenies@stevens.edu

Fernando Perez-Cruz
Swiss Data Science Center, (ETH Zurich and EPFL)
fernando.perezcruz@sdsc.ethz.ch

ABSTRACT

State-of-the-art password guessing tools, such as HashCat and John the Ripper, enable users to check billions of passwords per second against password hashes. In addition to performing straightforward dictionary attacks, these tools can expand password dictionaries using password generation rules, such as concatenation of words (e.g., "password123456") and *leet speak* (e.g., "password" becomes "p4s5w0rd"). Although these rules work well in practice, expanding them to model further passwords is a laborious task that requires specialized expertise.

password hash. Instead of exhaustively trying all possible character combinations, password guessing tools use words from dictionaries and previous password leaks as candidate passwords. State-of-the-art password guessing tools, such as John the Ripper [84] and HashCat [29], take this approach one step further by defining heuristics for password transformations, which include combinations of multiple words (e.g., `iloveyou123456`), mixed letter case (e.g., `iLoVeyOu`), and *leet speak* (e.g., `il0v3you`). These heuristics, in conjunction with Markov models, allow John the Ripper and HashCat to generate a large number of *new* highly likely passwords.

2018 IEEE Symposium on Security and Privacy Workshops

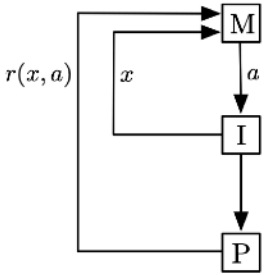
Deep Reinforcement Fuzzing

Konstantin Böttinger
Fraunhofer AISEC
85748 Garching, Germany
konstantin.boettinger@aisec.fraunhofer.de

Patrice Godefroid
Microsoft Research
98052 Redmond, USA
pg@microsoft.com

Rishabh Singh
Microsoft Research
98052 Redmond, USA
rishabh.iit@gmail.com

***Abstract*—Fuzzing is the process of finding security vulnerabilities in input-processing code by repeatedly testing the code with modified inputs. In this paper, we formalize fuzzing as a reinforcement learning problem using the concept of Markov decision processes. This in turn allows us to apply state-of-the-art deep Q-learning algorithms that optimize rewards, which we define from runtime properties of the program under test. By observing the rewards caused by mutating with a specific set of actions performed on an initial program input, the fuzzing agent learns a policy that can next generate new higher-**



2018 17th IEEE International Conference on Machine Learning and Applications

Automated Vulnerability Detection in Source Code Using Deep Representation Learning

Rebecca L. Russell[1*], Louis Kim[1], Lei H. Hamilton[1], Tomo Lazovich[1†],
Jacob A. Harer[1,2], Onur Ozdemir[1], Paul M. Ellingwood[1], Marc W. McConley[1]

[1] *Draper*
[2] *Boston University*

***Abstract*— Increasing numbers of software vulnerabilities are discovered every year whether they are reported publicly or discovered internally in proprietary code. These vulnerabilities can pose serious risk of exploit and result in system compromise, information leaks, or denial of service. We leveraged the wealth of C and C++ open-source code available to develop a large-scale function-level vulnerability detection system using machine learning. To supplement existing labeled vulnerability datasets, we compiled a vast dataset of millions of open-source functions**

II. Related Work

There currently exist a wide variety of analysis tools that attempt to uncover common vulnerabilities in software. Static analyzers, such as Clang [7], do so without needing to execute programs. Dynamic analyzers repeatedly execute programs with many test inputs on real or virtual processors to identify weaknesses. Both static and dynamic analyzers are rule-based

2017 European Intelligence and Security Informatics Conference

Adversarial Machine Learning in Malware Detection: Arms Race between Evasion Attack and Defense

Lingwei Chen, Yanfang Ye*, Thirimachos Bourlai
Department of Computer Science and Electrical Engineering
West Virginia University, Morgantown, WV 26506, USA
lgchen@mix.wvu.edu, yanfang.ye@mail.wvu.edu, thirimachos.bourlai@mail.wvu.edu

***Abstract*—Since malware has caused serious damages and evolving threats to computer and Internet users, its detection is of great interest to both anti-malware industry and researchers. In recent years, machine learning-based systems have been successfully deployed in malware detection, in which different kinds of classifiers are built based on the training samples using different feature representations. Unfortunately, as classifiers become more widely deployed, the incentive for defeating them increases. In this paper, we explore the adversarial machine**

violated by an adversary who may carefully manipulate the input data to exploit specific vulnerabilities of a classifier and thus to compromise its security. In other words, machine learning itself may open the possibility for an adversary who maliciously "mis-trains" a classifier in a malware detection system by simply changing the data distribution or feature importance [3], [4]. When the machine learning system is deployed in a real-world environment, it is of a great interest

Weaponizing data science for social engineering: Automated E2E spear phishing on Twitter

John Seymour and Philip Tully
{jseymour, ptully}@zerofox.com

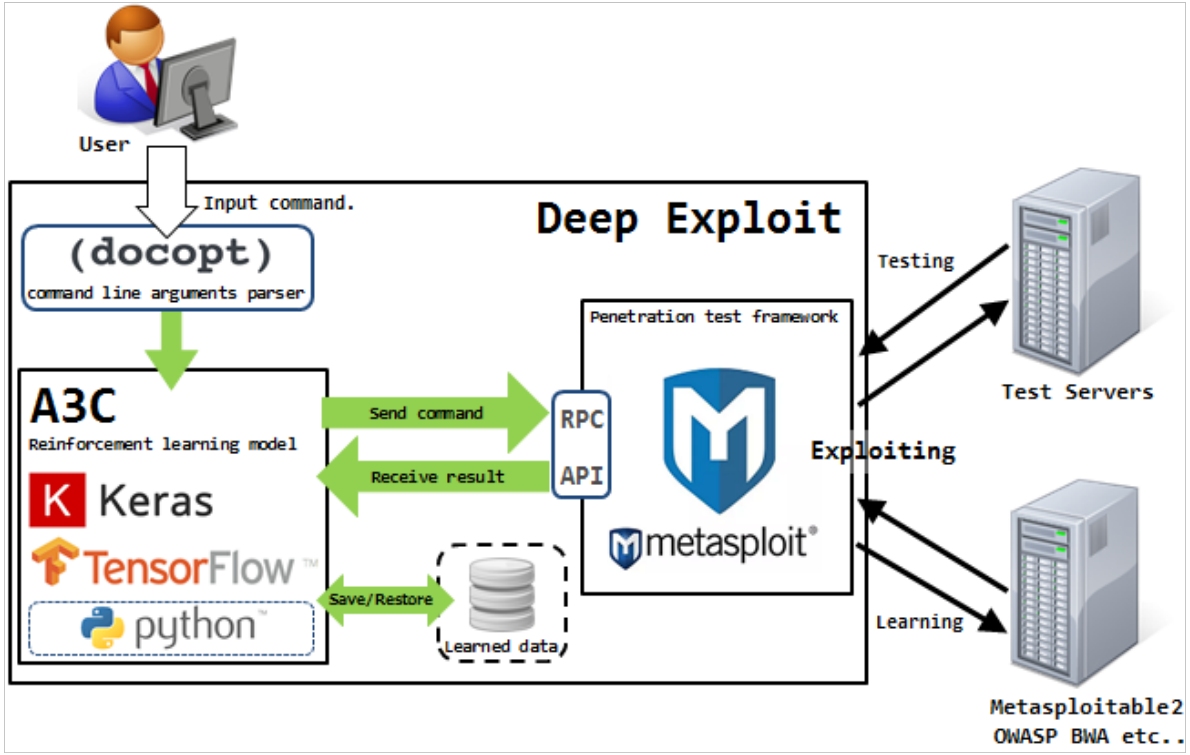
Introduction and Abstract

Historically, machine learning for information security has prioritized defense: think intrusion detection systems, malware classification and botnet traffic identification. Offense can benefit from data just as well. Social networks, especially Twitter with its access to extensive personal data, bot-friendly API, colloquial syntax and prevalence of shortened links, are the perfect venues for spreading machine-generated content.

Radford, Alec, et al. "Language models are unsupervised multitask learners." *OpenAI blog* 1.8 (2019): 9.

SYSTEM PROMPT (HUMAN-WRITTEN)	*In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.*
MODEL COMPLETION (MACHINE-WRITTEN, 10 TRIES)	The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science. Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved. Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

Isao Takaesu, Deep Exploit, 2018



https://github.com/13o-bbr-bbq/machine_learning_security/tree/master/DeepExploit

Deepfakes


WILL FERRELL
JOHN C. REILLY
STEP BROTHERS
They grow up so fast.
july 25th
From the guys who brought you
Talladega Nights


Stallone/Arnold Step Brothers
DEEPFAKE