Yicheng Cai

github.com/ethan-nicholas-tsai | homepage | yccai2024@gmail.com | +86-130-6007-3998

Education

- **Sichuan University** — Sichuan, China
  M.E. in Cybersecurity; GPA: 3.72/4.0 — *Sep. 2021 - Jun. 2024 (Expected)*
- **Sichuan University** — Sichuan, China
  B.E. in Cybersecurity; GPA: 3.87/4.0 ***(Rank: 1/130)*** — *Sep. 2017 - Jun. 2021*

Academic Exchange

- **National University of Singapore** — Singapore
  Visiting Student — *Jul. 2019 - Aug. 2019*
  - Studied the application of smart contract to solve the difficulties in crediting for small upstream firms in supply chains, then designed a poster to demonstrate the results
  - Literature review on the problems existing in supply chain finance and analyzed the relations between entities, including providers, core enterprises, banks, logistics, etc.
  - Conceptualized the idea of credit token and devised a policy to prevent repudiation of crediting, apart from the guaranteed confidentiality, integrity and availability from blockchain
  - Open Source: Poster

Honors and Awards

First Class Merit Based Fellowship (Top 3%)	*Sep. 2021 - Jun. 2024*
Honor Graduate of Sichuan Province (Top 3%)	*2021*
Yongzhuang Top Ten College Student Scholarship (**Top 1%**)	*2021*
Second Class Merit Based Scholarship (Top 4%)	*2019, 2020*
National College Student Information Security Contest (**First Prize, Most Recognized in China**)	*2020*
National Undergraduate Training Program for Innovation and Entrepreneurship (Excellent)	*2020*
China Qulian Blockchain Development Competition (Second Prize)	*2019*
China College Students' "Internet+" Innovation and Entrepreneurship Competition (Bronze Award)	*2019*
National Scholarship (**Top 1%, Most Recognized in China**)	*2018*
First Class Merit Based Scholarship (**Top 1%**)	*2018*

Publications

- **Yicheng Cai**, Haizhou Wang, Huali Ye, Yanwen Jin, Wei Gao. Depression Detection on Online Social Network with Multivariate Time Series Feature of User Depressive Symptoms, in *Expert Systems with Applications* (**ESWA, Impact Factor: 8.5**), 217(119538), 2023.
- **Yicheng Cai**, Haizhou Wang. Spam Movie Review Detection with Multi-View Explicit and Implicit Relations Semantics Fusion. *(Under Review by **IEEE Transactions on Big Data**)*
- Fengyuan Liu, **Yicheng Cai**, Haizhou Wang. Social Bot Detection with Multimodal Feature Representation Learning and Deep Fusion. *(In Preparation)*

Academic Research

- **Depression Detection on Online Social Network with Multivariate Time Series Feature of User Depressive Symptoms**:
  - Built a depression dataset with complete and time-consecutive user tweeting history, containing **3,711** depressed users and **19,526** non-depressed users with a total of **4,854,421** tweets
  - Proposed a novel feature extraction method - user Depressive Symptoms Time Series (DSTS) feature extraction that reveals user depressive symptoms variation in multivariate time series, contributing to a **nearly 3% improvement in F1-Score** compared with the second-best model MFFN
  - Revealed the clinical meaning of features and their contributions in online depression detection, which were **little considered by most existing researchers**
  - Open Source: Code
- **Spam Movie Review Detection with Multi-View Explicit and Implicit Relations Semantics Fusion**:
  - Proposed a novel approach – Multi-View Explicit and Implicit Relations Semantics Fusion model to detect spam movie review, which **outperformed** four state-of-the-art models by **8.7% in F1-Score on average**
  - Introduced a meta-based method to automatically build multi-view implicit relationships review graph, which reveals consistency and inconsistency implicit relationships existing in content, score, and time dimensions of spam movie reviews
  - Introduced a method to build a movie fact graph from internal and external movie review systems, then used heterogeneous graph transformer (HGT) model to extract movie fact embeddings
  - Open Source: Code

- **Key Technologies of Evolutionary Simulation for Intelligent Society Governance**:
  - Arranged tasks for each team member and communicated with leaders of other sub-projects
  - Proposed new research about *Virtual Agents Personification on Social Networks* and wrote an application form about the research for **the General Program of National Natural Science Foundation of China**
  - Conceptualized two frameworks (Macroscopic and Microscopic) to measure and profile topical influence and sentiment contagion of public event stakeholders

Privacy Guardian - Website Privacy leakage Intelligent Perception System:

- Studied the sensitive information detection and threat level quantification on websites, and found **8,312** webpages leaking individuals' privacy in **117** websites **out of 300** websites with approximately **1 million** webpages in total
- Proposed a novel framework to extract sensitive information in unstructured data, which used regular expressions to extract content-based sensitive information with predictable patterns and rule-guided BERT-BiLSTM-CRF model to automatically extract fine-grained context-based sensitive information
- Built a web system to provide comprehensive privacy leakage perception service for users, which especially provided analytical reports of websites under censorship with the sensitive information on webpages highlighted and labeled
- Demo: Video (Youtube)

SDN-based Cyber Attack Recurrence Platform:

- Devised a network topology restoration method to rebuilt the network environment of cyber attack traffic
- Designed a multi-machine interactive replaying algorithm to ensure the packets replayed in correct order (0 miss out of 563 KB/s traffic) and approximate time gaps (-0.02s time difference on average between each packet and the first packet)
- Built a web system to analyze fundamental aspects of network traffic and to retore the topology before replaying the traffic
- Open Source: Code

Patent

- Haizhou Wang, Xinyu Chen, Liang Ke, Yixuan Fang, Sen Wang, **Yicheng Cai**, Wenxian Wang. *A Cantonese Rumor Detection Method Based on Deep Semantic-Aware Graph Convolutional Network*, C.N. Patent Number: ZL 2022 1 0371266.1, July 2022.

Relevant Coursework

Cybersecurity: Network Attack and Defense Technology (A-), Malicious Code Analysis Technology (A-), Multimedia Security (A), Applied Cryptography (A), Data Mining for Cybersecurity (A)
Computer Science: Deep Learning (A), Big Data Analysis and Privacy (A), Computer Organization and Architecture (A+), Computer Networks (A), Assembly Language Programming (A)
Mathematics: Linear Algebra (A+), Calculus (A), Probability Statistics (A+), Number Theory (A)

Skills Summary

- **Languages**: Python, C/C++, JavaScript, Java, Matlab, Latex
- **Frameworks**: Scikit, TensorFlow, Keras, NodeJS, Numpy, Seaborn, Pytorch
- **Platforms**: Linux, Web, Windows