Yicheng Cai

yccai2024@gmail.com | +86-130-6007-3998

Education

- **Sichuan University** — Sichuan, China
  Master of Engineering - Cybersecurity; GPA: 3.72/4.0 — *Sep. 2021 - Jun. 2024 (Expected)*
- **Sichuan University** — Sichuan, China
  Bachelor of Engineering - Cybersecurity; GPA: 3.87/4.0; Rank: 1/130 — *Sep. 2017 - Jun. 2021*

Paper & Patent

- **Depression Detection on Online Social Network with Multivariate Time Series Feature of User Depressive Symptoms**: **Yicheng Cai**, Haizhou Wang, Huali Ye, Yanwen Jin and Wei Gao
- **Spam Movie Review Detection with Multi-View Explicit and Implicit Relations Semantics Fusion**: **Yicheng Cai**, Haizhou Wang (under review)
- **A Cantonese Rumor Detection Method Based on Deep Semantic-Aware Graph Convolutional Network**: Haizhou Wang, Xinyu Chen, Liang Ke, Yixuan Fang, Sen Wang, **Yicheng Cai**, Wenxian Wang, *CN114444516A, 2022.05.06*

Academic Research

- **Depression Detection on Online Social Network with Multivariate Time Series Feature of User Depressive Symptoms**:
  - Studied the depression detection method in terms of characterizing the depressive state variation features of users and establishing the correlation between features and depressive symptoms in clinical
  - 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 in most existing researchers
  - Explored experimentally the various influencing factors that affect the performance of the proposed framework
  - Open Source: Part 1, Part 2, Part 3
- **Spam Movie Review Detection with Multi-View Explicit and Implicit Relations Semantics Fusion**:
  - Studied the spam movie review detection method via mining semantic patterns in both explicit and implicit relations from data on internal and external movie review systems
  - 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
  - Evaluated experimentally the validity, robustness of method and utility of model components, then explored the parameter sensitivity of the proposed model
  - Open Source: Source Code
- **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
  - Summarized the most common sensitive information (according to Personal Information Security Specification and General Data Protection Regulation) into four types, including personal information, network identity information, secret and credential information, and financial information
  - 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
  - Designed a rule-based webpage privacy leakage threat level quantification method, which at first classified sensitive information into four classes in terms of sensitivity and then defined the threat level of a webpage according to amount of sensitive information leakage of all types
  - 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
  - Open Source: Part 1, Part 2; Demo: Video

- **SDN-based Cyber Attack Recurrence Platform**:
  - Studied the cyber simulation and attack recurrence in Software-Defined Networks (SDN)
  - Simulated multiple cyber attack and captured the traffic via TCPDump on every node of a virtual machine network
  - Devised a network topology restoration method to rebuilt the network environment of cyber attack traffic, with the aid of SDN
  - 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: Source Code
- **Credit Token - A Solution to Small Companies' Credit Problems**:
  - 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
- **Research and Application Demonstration of Key Technologies for Experimental Evolution of Intelligent Social Governance**:
  - Arranged tasks for each team member and communicated with leaders of other sub-projects
  - Literature review on methods for representing entities on social media in terms of attributes, relations and behaviors
  - Conceptualized two frameworks (Macroscopic and Microscopic) to measure and profile topical influence and sentiment contagion of public event stakeholders
  - Designed a LinkedIn crawler to collect user profiles and organization members information, which surpassed the anti-spider policies by analyzing the data transmission APIs from HTTP requests and responses
  - Designed a Sina Weibo crawler to collect user profiles, tweets and retweets, which surpassed the anti-spider policies using the same method as the previous one and provided 10 million tweets collection service on a daily basis

Honors and Awards

- First Class Merit Based Fellowship, Sichuan University (Top 3%) - Sep. 2021 - Jun. 2024
- Honor Graduate of Sichuan Province, Education Department of Sichuan Province (Top 3%) - 2021
- Yongzhuang Top Ten College Student Scholarship (Top 1%) - 2021
- Second Class Merit Based Scholarship, Sichuan University (Top 4%) - 2019, 2020
- National College Student Information Security Contest (First Prize, Top 5%), Ministry of Education College Cyberspace Security Professional Teaching Steering Committee - 2020
- National Undergraduate Training Program for Innovation and Entrepreneurship (Excellent) - 2020
- China Qulian Blockchain Development Competition (Second Prize), China Academy of Information and Communications Technology - 2019
- The 5th China College Students' "Internet+" Innovation and Entrepreneurship Competition (Bronze Award), Education Department of Sichuan Province - 2019
- China National Scholarship, Ministry of Education of the People's Republic of China (Top 1%) - 2018
- First Class Merit Based Scholarship, Sichuan University (Top 1%) - 2018

Internships

- **Chengdu Guoxin'an Information Industrial Base** Sichuan, China
  Research Assistant *Jun. 2020 - Dec. 2020*
  - Designed and built a Java Server to receive and display data collected from local hosts, whose front-end was implemented via Spring Boot and Layui framework
  - Designed and built a Python Client that daily checked security compliance of windows hosts and sent the collected data to the server, including patch upgradation, registry, essential host information, etc.
  - Packaged the sever software via Jar package and Docker image and deployed it on a cloud server, while packaged the client software via PyInstaller and installed it on windows hosts
  - Open Source: Source Code

Skills Summary

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