Yifan Tian

yifan.tian.0610@gmail.com | (443) 509-4001

Working Experience

- **Meta** Washington, D.C.
  Security Engineer, Investigations *Aug 2024 – Present*
  - **Sex Trafficking Threat Actor Network Investigation**: Conduct in-depth investigation on threat actors leveraging Meta's platform for human exploitation and sex trafficking purposes
    - **Signal Development**: Develop and integrate 6 signals into malicious actor behavior framework to improve efficiency by 20% and efficacy by 35% respectively for threat actor lead generation and triage
    - **Network Disruption**: Disrupt 2197 assets owned by sex trafficking actor networks in South East Asia
    - **DOJ Sanction**: Respond to multiple DOJ Sanction escalation to remove DOJ-sanctioned and affiliated object from Meta's platforms
  - **Financial Impact Analysis**: Establish a checkpoint to maintain the highest standards of compliance and integrity in Meta's investigative operations by identifying and differentiating revenue from threat actors versus legitimate sources
    - **ETL Pipeline**: Create data pipeline to extract, transform and load data into an easy-to-use table for investigators to understand the potential financial impacts from their operations
    - **High Profile Account Detection**: Provide detection / alert for high profile assets on Meta's platform before final disruption to avoid over-action and improve the enforcement accuracy
  - **Integrity TTP / Killchain Management**: Redesign TTP / Killchain for Meta i3E to demonstrate the evolving threat landscape on integrity violation side, understand the coverage and gap among i3E detection, enforcement and response and direct the under-investigated aspects for roadmapping
    - **Scale**: Redesign 700+ tactics, techniques and procedures and associate signal combinations to procedures and build auto labeling
    - **Automation**: Develop, train and validate decision tree based model with historical investigation data to implement a decision dossier with accuracy of 85%+
  - **HEx Team Metrics Dashboard**: Create team metrics dashboard for human exploitation (HEx) team to demonstrate impact to upper management as well as threat intel for legal, policy and other cross functional teams
    - **Impact**: Implement 20+ metrics derived from HEx investigations and operations, with multiple (2 - 5) visualization for each metric to present an in-detail view of team efforts, success and impact
- **Workday** McLean, VA
  Senior Cybersecurity Engineer *Jul 2023 – Aug 2024*
  - **Detection Rule Curation**: Put some DE works here
    - **Agile Sprint**: Lead team to conduct curation of 10+ detection ideas each week along with detection rule implementation from previous iteration
    - **Threat Intel**: Lead team to discuss latest threat intel from OSINT and in-house TIP, review gaps in detection coverage and propose ideas correspondingly
  - **SOC Collaboration**: Host bi-weekly meeting with SOC team to listen and address Top 5 issues from them
    - **Cross Team Collaboration**: Work as point of contact for SOC team to organize 5 issues / concerns on top of their list for our team to solve in each sprint, collect feedback to improve quality of service and update SLA accordingly
  - **LLM Based Detection Coverage Analysis**: Use LLM / RAG to analyze the detection coverage based on our rule repository and third party vendor policy / control lists
    - **LLM**: Feed 2,000+ rules into Llama and build a rag to scan rules, controls and policies to figure out similar rules and duplicated coverages as well potential gaps in different stages of killchain
    - **Security Platforms**: Manage rules in variuos security platforms including Crowdstrike, Wiz, Cloudflare, Okta, ModSec, Twistlock, GuardDuty etc
  - **Pipeline Health Monitoring**: Build side channel monitoring system to present a health dashboard of pipelines from upstream data sources to Splunk SIEM and XSOAR SOAR to ensure latencies are within SLAs
    - **Automation**: Create server-less functions to run scheduled jobs to query data source endpoints
    - **Monitoring**: Create dashboard with data visualization to show the real-time status of each data pipeline and alert on potential pipeline outage / delay
  - **Rule Not Fired Analysis**: Conduct analysis on detection rules which have not been fired in last 180 days, automatically suggest potential issues including lookup error, data source changes, configuration changes etc
    - **SPL**: Programatically divide rule SPLs into several breakpoints to check performance of each part of the rule, and automatically identify the specific clause suppressing the rule from firing

    - **Jira Integration**: Integrate with Jira to create tickets with full diagnoses details for developers to investigate and debug

- **Abnormal Security** San Fransisco, CA / Remote from Ashburn, VA
  Threat Intelligence Engineer *Sep 2021 – Mar 2023*
  - **Threat Intelligence Portal**: Build and deploy a threat intelligence portal from scratch with React.js, Django and PostgreSQL along with a set of libraries / APIs including Chart.js, Material UI, Pandas, Google APIs, Twilio API, RapidAPI, GPT API, etc to enable threat researchers with real-time scammer engagement, data query, filter and visualization while providing semi-auto message response and mule account collection feature
    - **Scale**: Launch 30,000+ email security investigations and collect 20,000+ mule accounts within 1 year
    - **Automation**: Automate 80% of threat actor communication and 75% of mule account collection
    - **Traffic**: Deliver cutting edge research reports and blog posts to draw 70%+ traffic of the company website
    - **Efficacy**: Help R&D team improve detection efficacy on various attack categories
    - **Impact**: Send hundreds of real-time mule account notification to external partners including 40+ top financial institutes weekly to prevent potential financial losses of $100,000+ and help drive POV to potential customers
  - **Intelligence Augmentation**: Build an ETL pipeline with Python Scrapy crawlers / data processing scripts to collect auxiliary data from both OSINT and member-only resources to augment threat intelligence collection
    - **Comprehension**: Add various data (e.g. domain age, IPs, validated bank codes, etc) to make intelligence collection more comprehensive to enable data analysis in different angles
    - **Budget**: Save $1,000+/yr by crawling public data, reducing API calls, building company-owned services
  - **Internal Intel Sharing**: Build a S3-Databricks pipeline to sync collected threat intelligence and share raw data to internal teams along with dashboards and visualizations using SparkSQL to provide high level threat insight
    - **Timeliness**: Perform hourly update enabling R&D to fine-tune filter rules to prevent latest attacks
    - **Visualization**: Deliver 10+ dashboard / charts to show stakeholders most recent trend of threat landscape
    - **Certification**: Databricks Certified Data Analyst Associate

- **Agari Data** Foster City, CA / Remote from Ashburn, VA
  DevOps Research Engineer / Research Intern *Feb 2020 – Sep 2021 / Jun 2018 – Dec 2019*
  - **Advanced Intelligence Collection**: Build a honeypot with Flask, Bootstrap and S3 to trick scammers with fake wire confirmation and collect in-depth scammer intelligence; Build a Ruby on Rails App to manage each campaign and create an ElasticSearch cluster for efficient large-scale intelligence search
    - **Scale**: Collect 300+ scammer profiles with 1TB+ intelligence in 2 years
    - **Impact**: Uncover how Scattered Canary threat group uses stolen credentials to abuse COVID Pandemic Unemployment Assistant program and get stimulus check, results in 20+ media reports, raises company position in the email security market and prevents potential financial losses of hundreds of millions
  - **Customer Intelligence Portal**: Build a customer portal with React.js, Node.js, Chart.js, REST API, Docker, Kubernetes to provide customers with real-time scammer data analysis, threat actor visualization and API query
    - **Productized**: Create a one of a kind solution for financial institutes suffering from fraudulent account and financial loss by sharing with them timely intelligence and evidence for fraud investigation
    - **Revenue**: Generate an ARR of $200,000+ as the 4th product of the company

- **Embry-Riddle Aeronautical University** Daytona Beach, FL
  Teaching / Research Assistant *Aug 2017 – Dec 2019*
  - **Java & Algorithm**: Teach Java programming, data structures and algorithms, provide tutoring hours & grading
  - **Cybersecurity**: Conduct research including cloud & IoT security, privacy, phishing etc (see publications here)

Selected Research Projects

- **Privacy-preserving authentication for edge-assisted IoDs (JISA'19)**: *Jan 2019 – Oct 2019*
  - **Ad-hoc VANET**: Design an ad-hoc VANET authentication framework for Internet of Drones with the consideration of the high mobility of UAVs with the integration of mobile edge computing (MEC) to further reduce the authentication cost for potential authentication activities prefictively
  - **Privacy**: Enable individual UAV to generate and control secret keys by itself to eliminate the key escrow issue in existing VANETs frameworks as well as provide a study of buffer pseudonym and public key update design to provide privacy protection for UAVs

- **Low-Cost NLOS UAV Invasion Detection (AIAA/IEEE DASC'19)**: *May 2018 – Oct 2019*
  - **Wireless Network**: Collect and analyze UAV (drone) received signal strength indicators with low-cost Raspberry Pi and achieve competitive invasion detection accuracy against expensive radar & acoustic matrix based solutions
  - **Deep Learning**: Build deep neural networks (DBN & LSTM) using Tensorflow & Keras to detect drone invasion

- **Lightweight Privacy-preserving DNN Outsourcing (SecureComm'19/JIoT'20)**: *Sep 2017 – Oct 2019*
  - **Cryptography**: Design a lightweight cryptosystem to support secure DNN data outsourcing to cloud/edge server
  - **CNN**: Apply to well-known models (AlexNet & LeNet) and achieve same accuracy while protecting user privacy
- **Practical Query Integrity Verification for NoSQL Databases (NCA'19)**: *Jan 2019 – Apr 2019*
  - **Data Integrity**: Present a novel integrity verification scheme integrating Pointer Embedded Merkle B+ Tree (PMBT), for NoSQL databases to allow both the data owner and clients to check the integrity of queries
  - **NoSQL**: Design the scheme to be compatible with popular NoSQL databases such as HBase, BigTable, MongoDB
- **Secure Data Filtering for Distributed Data Streams (SmartData'19)**: *Jan 2019 – Apr 2019*
  - **Data Filtering**: Propose a privacy-preserving filtering scheme for distributed data streams outsourced to the public cloud by allowing cloud servers to filter out corresponding data streams directly over encrypted data
  - **Secure Multi-party Computation**: Prevent brute force guessing attacks from public cloud servers by leveraging the multi-cloud design in the scheme
- **Privacy-preserving Truth Discovery with Blockchain (DLoT'19)**: *Apr 2018 – Oct 2018*
  - **Blockchain**: Design Ethereum smart contract to achieve high accuracy in decentralized truth discovery tasks
  - **Privacy**: Integrate differential privacy techniques to protect user data privacy with minor accuracy trade off
- **Privacy-preserving Image Annotation (IEEE CNS'17/EAI ETSS'20)**: *Sep 2016 – Apr 2017*
  - **Image Processing**: Use OpenCV and Scipy to implement an image annotation scheme including feature extraction, vector normalization and dimension deduction
  - **Cryptography**: Protect image feature privacy against public cloud by integrating homomorphic encryption
- **Privacy-Preserving K-means Clustering (IEEE TCC'17)**: *Sep 2016 – Jan 2017*
  - **Privacy**: Design a clustering algorithm over encrypted data to achieve high accuracy while protecting privacy
  - **Map Reduce**: Introduce secure map reduce to suit cloud environment and enable scalability
- **Social Behavior Based Cross SNS Phishing Attacks (IEEE CNS'16)**: *Jan 2016 – May 2016*
  - **Web Crawler**: Use Scrapy and Selenium to implement Python crawlers to collect personal information from public online resources and social network services (SNSs) in ethical style
  - **Social Engineering**: Analyze and propose a SNS linkage cut-off strategy to minimize information leak issue
- **Securing WiFi-Based UAVs From Security Attacks (IEEE Milcom'16)**: *Sep 2015 – Dec 2015*
  - **Penetration Testing**: Use Air-crack and Metasploit to launch multiple hackings (DoS, buffer overflow, de-authentication and ARP poisoning attacks) against commercial Parrot Bebop drone
  - **Vulnerability**: Mitigate potential WiFi-based attacks by analyzing and removing corresponding vulnerabilities

Education

- **Embry-Riddle Aeronautical University** Daytona Beach, FL
  Ph.D. in Computer Science; *Jan 2016 – Dec 2019*
- **Johns Hopkins University** Baltimore, MD
  M.S. in Security Informatics; *Aug 2014 – Dec 2015*

Skills

- **Languages**: Python, Ruby, Java, C\C++, React, Augular, Node, PHP / Hack, SQL, Presto, SPL, Matlab, LaTeX
- **Libraries & Tools**: Tensorflow, Keras, Jyupter, SciPy, NumPy, Scrapy, Selenium, OpenCV, OpenAI, Django, Rails, Flask, Kubernetes, Drone, Vault, helm, Elasticsearch, Kibana, Docker, Nginx, Apache, Git, Metasploit, Nessus, Air-crack, Wireshark
- **Platforms**: AWS, GCP, Splunk, XSOAR, XDR, Crowdstrike, Wiz, Cloudflare, Okta, Prisma, Zscaler, ModSec, Twistlock, GuardDuty, Snyk, Qualys
- **Certifications**: Certified Information Systems Security Professional (CISSP), Splunk Certified Cybersecurity Defense Analyst, Splunk Core Certified Power User, CrowdStrike Certified Falcon Administrator, Certified Scrum Product Owner, Databricks Certified Data Analyst Associate