# Sajjad “JJ” Arshad

+1 (617) 390-3316, sajjad.jj.arshad@gmail.com, https://sajjadium.github.io

## SUMMARY


- **10+ years of software development** experience across multiple programming languages (e.g., **C/C++, Python, Java, Go, Bash, NodeJS**), with a strong foundation in **data structures and algorithms**.
- Demonstrated cross-org coordination skills by leading numerous **cross-team/cross-functional** working groups to develop strategy roadmaps and deliver significant, complex solutions.
- Expert in conducting large-scale **Web security and privacy** measurements, specifically focused on evaluating the effectiveness of Web attacks and analyzing the dynamics of the online advertising ecosystem.
- Proficient in **advanced Linux programming**, including process management, **IPC** (signals, shared memory, semaphores, sockets), and multi-threading, with a strong background in **Linux internals**.
- Developing efficient **large-scale static analysis** algorithms to detect harmful behaviors in cross-platform frameworks within **Android** apps, focusing on languages like Dalvik Bytecode, JavaScript, and native code.
- Expanded my expertise in **GenAI/LLM** by exploring concepts such as **RAG, ReAct, Embeddings, and Agents**, while also gaining practical experience with Gemini API and SAX.
- Possessed a strong understanding of **networking** (e.g., **TCP/IP**, UDP, Raw socket, Switching, Routing), **security protocols** (e.g., **TLS/SSL**, **PKI**, Firewall, IDS/IPS, WAF), and **containerization** (e.g., **Docker**, **Kubernetes**).
- Published over **15 papers in security conferences** like **USENIX Security**, **NDSS**, and **WWW**, with **Web Cache Deception (WCD)** research recognized as PortSwigger's Top Web Hacking Technique of 2019.
- Extensive hands-on experience in **offensive security**, with expertise in **system security** techniques (e.g., **shellcode**, **ROP**, **DEP**, **ASLR**, **heap exploitation**, **sandboxing**, CFI) and tooling (e.g., Ghidra, IDA Pro, Binary Ninja). Proficient in web security, covering topics such as **XSS**, **CSRF**, **SSRF**, Web Cache Deception, OAuth, and CORS. Decent background in **memory safe** system languages such as **Rust**.
- Actively contribute to the cybersecurity community by sharing my expertise through **Android app hacking** workshops at prominent events like DEFCON, BSidesSF, HackerOne, and H@cktivityCon, while also organizing hacking competition like GoogleCTF and ensuring the preservation of knowledge through maintaining CTF archives and writeups.


## EXPERIENCE

**SENIOR SOFTWARE ENGINEER @ Google**, Mountain View, CA — May 2019 - Present

- Leading development of **static analysis & RE tooling** (e.g. Semgrep) to detect Android abuse at scale.
- Owning multiple **backend RPC services** responsible for large-scale abuse detection in Google Play.
- Our Android real-time malware blocking technology received widespread press coverage appearing in outlets such as Bleeping Computer, Tech Radar, Ars Technica, 9to5Google, and The Strait Times.
- Mastered **Google’s CI/CD** toolkits, data storage platforms (e.g., Bigtable, Spanner), microservice frameworks (e.g., gRPC, Protobuf), and load balancing infra to deploy large-scale distributed systems.
- Built a decompiler that allows for the recovery of JavaScript source code from **React Native's Hermes** bytecode.
- Enhanced our Android app review platform by incorporating **LLM**, enabling it to automatically generate concise summaries and offer improved code comprehension tools for the reviewers.

**RESEARCH INTERN @ Mozilla Corporation**, Mountain View, CA — June 2017 - August 2017

- Conducted a measurement study on the adoption of **TLS 1.3** by middleboxes in enterprise networks.
- Developed a Firefox add-on to collect data from Firefox users into Telemetry platform (Spark) for offline analysis.
- Enhanced Firefox browser code base (C++) to enable flexible HTTPS connection configuration for developers.

**RESEARCH INTERN @ Verisign**, Reston, VA — May 2016 - August 2016

- Used ML (e.g. clustering) to analyze large scale HTTP traffic and protect against **DDoS** attacks.
- Acquired in-depth knowledge of **DNS** and related protocols (e.g., DNSSEC, Whois).

**RESEARCH ASSISTANT @ Northeastern University**, Boston, MA — September 2013 - April 2019

- Developed a large-scale and distributed crawling platform using Chromium browser instrumentation.

- Using JVM instrumentation and bytecode analysis (ASM, Soot), we built **static and dynamic analysis tools**, including a fuzzer, to detect algorithmic complexity vulnerabilities in Java, resulting in CVE-2018-1517.
- Conducted a large-scale measurement of Web Cache Deception attack on popular CDNs (Akamai, Cloudflare).
- Large scale measurement of relative path overwrite (RPO) vulnerability using AWS and Common Crawl dataset.
- Instrument Chromium for **large-scale web security measurement** study including detection of malicious web pages/domains, ad injection by browser extensions, malvertising, and outdated JavaScript libraries.
- Conducted extensive research on the ecosystem of **online/web advertising**, its related privacy/security issues (e.g., cookie matching, re-targeted ads), security problems of **Web protocols**, and client-side technologies (e.g., JavaScript, HTML5, browser extensions).
- Built a malware analysis infrastructure based on **Cuckoo sandbox** for large-scale detection of Ransomware.
- Implemented a **content distribution network (CDN)** using **DNS redirection**, user-space threads in Linux, a file system using FUSE library in Linux, and a secure chat system using built-in Java security framework.

---

**PART-TIME SOFTWARE ENGINEER**, Tehran, Iran — September 2004 - August 2013

- Developed and deployed a large-scale **SIEM** log management system using Java, **MySQL**, Hadoop, and MapReduce to efficiently collect and process logs from diverse enterprise network devices, including firewalls, IDSes/IPSes, desktops, and servers.
- Researched security and performance evaluation of **shared Web hosting** solutions for **Apache web server**.
- Developed a machine learning-based botnet detection system in C++ that analyzes network communication patterns between infected hosts and command-and-control (C&C) servers. Utilizing the WEKA tool for clustering, the system identifies anomalous network behavior indicative of botnet activity, providing a novel approach to cybersecurity threat detection.
- Built a robust system capable of processing large volumes of **PCAP** network traffic and transforming them into netflow data, enabling the identification and analysis of malicious attack patterns within the network.
- Created a unified **database** schema to consolidate student and personnel information, effectively centralizing data previously spread across disparate Oracle and MS SQL Server databases within legacy systems.
- Gained a broad range of experience developing **Web applications** in diverse sets of technologies including PHP, J2EE, and ASP.NET, alongside database management using Oracle, MySQL, and MS SQL Server.

## EDUCATION

**Northeastern University**, PhD in Computer Science — April 2019
**Shahid Beheshti University**, MS in Computer Engineering — January 2011
**University of Tehran**, BS in Computer Engineering — August 2008

## PUBLICATIONS

- The Matter of Captchas: An Analysis of a Brittle Security Feature on the Modern Web, **WWW**, 2024
- Cached and Confused: Web Cache Deception in the Wild, **USENIX Security**, 2020
- HotFuzz: Discovering Algorithmic Denial-of-Service Vulnerabilities Through Guided MicroFuzzing, **NDSS**, 2020
- A Longitudinal Analysis of the ads.txt Standard, **ACM IMC**, 2019
- On the Effectiveness of Type-based Control Flow Integrity, **ACSAC**, 2018
- How Tracking Companies Circumvented Ad Blockers Using WebSockets, **ACM IMC**, 2018
- Large-Scale Analysis of Style Injection by Relative Path Overwrite, **WWW**, 2018
- Thou Shalt Not Depend on Me: Analysing the Use of Outdated JavaScript Libraries on the Web, **NDSS**, 2017
- Recommended For You: A First Look at Content Recommendation Networks, **ACM IMC**, 2016
- Identifying Extension-based Ad Injection via Fine-grained Web Content Provenance, **RAID**, 2016
- Tracing Information Flows Between Ad Exchanges Using Retargeted Ads, **USENIX Security**, 2016
- UNVEIL: A Large-Scale, Automated Approach to Detecting Ransomware, **USENIX Security**, 2016
- Include Me Out: In-Browser Detection of Malicious Third-Party Content Inclusions, **FC**, 2016
- Alert Correlation Algorithms: A Survey and Taxonomy, **CSS**, 2013
- A Comprehensive Approach to Abusing Locality in Shared Web Hosting Servers, **IEEE TrustCom**, 2013
- Two Novel Server-Side Attacks against Log File in Shared Web Hosting Servers, **IEEE**, 2012
- Performance Evaluation of Shared Hosting Security Methods, **IEEE TrustCom**, 2012
- An Anomaly-based Botnet Detection Approach for Identifying Stealthy Botnets, **IEEE ICCAIE**, 2011