

# SHOAIB AMJAD KHAN



✉: [khan180@purdue.edu](mailto:khan180@purdue.edu)

in: [linkedin.com/in/shoaib-a-khan](https://www.linkedin.com/in/shoaib-a-khan)

🌐: [shoaib-a-khan.github.io](https://shoaib-a-khan.github.io)

## OBJECTIVE

---

Looking for exciting opportunities involving R&D in the areas of Security, Privacy, Machine Learning and AI at scale.

## EDUCATION

---

- 2016-Present     **Purdue University, Computer Science Department**, W Lafayette, IN  
Ph.D Candidate, research focus on an intersection of *privacy, security, data analytics & machine learning*  
**Advisor:** Prof. Mike Atallah
- 2007–2009     **University of South Florida, Mathematics Department**, Tampa, FL  
M.A. Mathematics, research focus on design of *approximation algorithms for hypergraphs*  
**Advisor:** Prof. Brendan Nagle
- 2001-2005     **National Univ. of Computer & Emerging Sciences, Computer Science Department**, Lahore, Pakistan  
B.S. Computer Science with Honors, final year project on *OCR for Urdu using neural networks*.

## PROFESSIONAL EXPERIENCE

---

- 2018 / 2019     **Microsoft**, Redmond, WA  
PhD. Intern (Summers)
- 2016–Present     **Purdue University, Computer Science Department**, W Lafayette, IN  
Graduate Teaching & Research Assistant
- 2015 - 2016     **Information Technology University, Computer Science Department**, Lahore, Pakistan  
Teaching Fellow
- 2012 - 2015     **National Univ. of Computer & Emerging Sciences, Computer Science Department**, Lahore, Pakistan  
Assistant Professor
- 2009 – 2012     **National Univ. of Computer & Emerging Sciences, Computer Science Department**, Lahore, Pakistan  
Lecturer

## TECHNICAL SKILLS

---

**Languages:** C#, C++/C, Python, Java, JavaScript (ReactJs), SQL, UML, HTML, XML, Assembly Language (Intel x88, x386, Intel 8051/52 MC series), Verilog, Prolog, LaTeX.

**Software Tools:** Visual Studio, VSCode, Eclipse, Microsoft .Net Platform, Git, Atom, Visio, Rational Rose, ERWin.

**Data Science & ML Tools:** Scikit-learn, NumPy, Pandas, Keras, Matplotlib, NLTK.

**Database & Client/Server Technologies:** Oracle 9i, Microsoft Access, Microsoft SQL Server, Apache HTTP Server, CGI.

**Operating Systems:** Microsoft Windows, Linux/Unix, OS X, DOS, Android, iOS.

## PROJECTS DEVELOPED (SELECT)

---

- **Industry Scale (Now Live):**
  - **Outlook Group Discovery Feature:** Design, full stack development and deployment of Group Discovery feature in Outlook Web App. The feature is currently deployed worldwide and is an integral part of Outlook Web UX. It uses ML to continuously improve user-relevant group recommendations based on select attributes.
  - **Yammer People Search:** R & D with Microsoft's Yammer team, focusing on improving "People Search" experience. The feature was deployed worldwide and is now an integral part of Yammer UX.

- **Security & Privacy:**
  - **DNS Cache Poisoning:** Wrote an exploit to carry out Kaminsky's Attack against a local DNS server in a controlled (virtual) environment.
  - **Garbled Computer (in C):** Designed & developed an open source tool for PFE using lightweight crypto primitives.
- **AI & ML:**
  - **Group Recommendation System (in Python):** Feature selection and implementation of a hybrid recommender system using ensemble learning techniques to suggest most relevant groups to a user in a toy social network.
  - **Email Spam Filter (in Python):** Developed and trained a Naïve Bayes classifier using NLTK for count vectorizing NLP dataset. The model had 97% precision and 98% recall.
  - **OCR for Urdu Script Using Neural Nets (in C#):** Studied problems hindering the development of a practical OCR system for Urdu and developed NN solutions to overcome these. I led a team of four.
- **Algorithms & Data Structures:**
  - **Search Engine (in C++):** A Local Search Engine that constituted two modules, i) indexing system and ii) querying system. The focus here was on design, implementation and testing of advanced data structures including B-Trees and Hash Tables to achieve the desired functionality efficiently.
  - **Compression Software in Assembly Language:** Compressor (encryption, decryption, compression & decompression) utility using Assembly Language.
- **Systems & Networks:**
  - **Download Manager using Sockets in C++:** A multithreaded desktop application that had features like accelerate download speed and pause or resume download from a web server on the internet.
  - **Compiler for C-v, a subset of C++ (in C++):** Individually wrote the complete compiler, which had three modules, i) Lexical analyzer, ii) Syntax analyzer, and iii) Intermediate code generator.
  - **Pipelined MIPS Architecture (Verilog):** Architecture design of pipelined MIPS processor in Verilog.
- **Software Engineering & DB:**
  - **Student Registration System (in C++):** An Objected Oriented Design that exclusively focused on issues of OOP e.g. data encapsulation and polymorphism. I led a team of five through the design, implementation & testing phases.
  - **Academic Records and Result Compilation System (Java, JScript, SQL):** Design & development of FE and BE of an online database to manage the records of students at The Punjab University. I was part of a team of five.

## **PROJECTS MANAGED (SELECT)**

---

- **Open Source iPhone to Android App Conversion Tool:** A formal language translation tool converting code written in Objective C to Java code. The tool is extensible and allows for updates in API mapping.
- **EmoTunes:** Using brain wave data from EEG to classify (using KNN's) the present emotional state of the user and generate playlists to either alter or enhance it.
- **An Eye for Blind (OpenCV):** Using a simple VGA camera and a Depth Camera (e.g. Microsoft Kinect), helping the blind to navigate their way around by detecting and identifying obstacles. The project is extensible and aims to include navigation maps to guide the visually impaired.

## **PUBLICATIONS**

---

- S. Khan, M. Atallah, Q.Malluhi, "Garbled Computation: Hiding Software, Data and Computed Values using Light-Weight Primitives", USENIX Security '20, in review.
- Shoaib A. Khan, B. Nagle, "A Hypergraph Regularity Method for Linear Hypergraphs, with Applications", LAP Lambert Academic Publishing, 2011.
- Shoaib A. Khan, "A Hypergraph Regularity Method for Linear Hypergraphs," Master's Thesis, University of South Florida, USA, 2009.

## **HONORS & AWARDS**

---

2016 - Present	<b>Graduate Assistantship</b> by Purdue University.
2007 - 2009	<b>Fulbright Scholarship</b> by United States Education Foundation for Pakistan.
2001 - 2005	<b>Outstanding Talent Scholarship</b> by Punjab Information Technology Board.
2001 - 2005	<b>Dean's Honor's List</b> at National University of Computer & Emerging Sciences.
Fall 2001	<b>University Scholarship for Batch Topper</b> by National University of Computer & Emerging Sciences.

## **REFERENCES**

---

To be furnished on request.