Masood Krohy

Big Data Analytics Consultant

m.kh@ieee.org

Summary

Hi there, I am Masood Krohy (Khosroshahy). Big Data Analytics, including AI / Machine Learning / Deep Learning, is extremely fascinating and I find it very rewarding to take up new hard technical challenges in this field and deliver operationalized solutions typically in short order. I have the scientific background to design mathematical algorithms and the software development expertise to turn those algorithms into scalable, production-quality code on the Hadoop and Spark platforms. I have about 6 years of experience in Data Scientist tasks obtained during the past 3 roles, spanning several industries: recently as the resident Data Scientist at Intact, the Canadian leader in P&C insurance, working on their Big Data telematics projects as well as working on the optimization of their Hadoop/Spark platform (2 years). Previous to this, I was working for Rogers and with their multi-terabyte telecom network data (1 year). Before Rogers, I was at Concordia, doing network security data analysis as part of my PhD thesis work (3 years). Here are few more things to know about me:

• Experience working with relational databases through my recent work with Postgres/PostGIS, through work on SwissQual's SQL Server-based NQDI database, and with Joomla's MySQL database in the context of several web development projects

• Expertise in Machine Learning and Data Mining (implemented a large number of algorithms)

• Equivalent to 6 years of full-time experience in software development (inc. Python, SQL, Java and C++)

• Authored/Co-authored several peer-reviewed papers: 2 ISI journal papers and 6 conference papers

• Certified Associate in Project Management (PMI's CAPM certificate)

• Excellent communication skills with full proficiency in English and French (multiple certificates)

Experience

Data Scientist at Intact Financial Corporation

May 2015 - March 2017 (1 year 11 months)

– Member of Intact Lab R&D team

– Working on the firm's usage-based insurance (UBI) project, involving telematics data analytics (80 billion records/year)

– Utilizing Postgres/PostGIS and several components from the Hadoop/Spark ecosystems

Technical Highlights:
– Architected, implemented, and operationalized a Big Data Mining application on Hadoop (pattern recognition in telematics data)
– Geospatial filtering and analysis of Big Data, implemented using Hive UDFs
– Designed and deployed a distributed, TensorFlow-based infrastructure to enable Deep Learning on Big Data projects
– Designed several Deep Learning models using TensorFlow to pump more intelligence into Intact's telematics program

Management and Coaching:
– Transformed business requirements from the telematics/R&D directors to technical solutions
– Helped the project managers with scoping & tasks assignment during the implementations
– Coached IT specialists during delegation of some operationalized/implemented solutions
– Advised C/D-level executives during several RFI processes with IT/infrastructure vendors
– Organized and delivered an Advanced Data Analytics training to 20+ analysts, senior analysts, and directors in the company (data mining/machine learning with Python/Hive/Spark)

Senior Analyst at Rogers Communications
August 2013 - August 2014 (1 year 1 month)
– Member of OSS Benchmarking team / Service Quality Assurance program
– Refined the multi-carrier wireless network drive testing (benchmarking) program

Technical Highlights:
– Developed a reporting/analytics system to gain insights from multi-terabyte SQL databases
– Used R to run statistical tests on benchmarking data collected from the 3 main cellular network operators in order to identify trends, support marketing claims, etc.
– Conducted root cause analysis of network events to help optimize Rogers' UMTS/LTE networks: voice call drops, coverage issues, low data throughput on LTE/HSPA networks, etc.
– Developed expertise of SwissQual's NQDI, NQWeb, NQView & Diversity Benchmarker

Management and Coaching:
– Transformed business requirements from OSS Benchmarking manager to technical solutions
– Remotely worked with and monitored a national team for data collection (drive tests)
– Provided expert opinion to the legal department in defending Rogers' marketing claims during occasional arbitration processes with its competitors
– Coached marketing executives on how to best utilize the network performance results to craft new marketing messages to the consumers in the Canadian market

– Created various charts/graphs to highlight and communicate potential improvements to the performance of the cellular network to the executives

Ph.D. Candidate at Concordia University

January 2010 - April 2013 (3 years 4 months)

– Developed several Continuous-Time Markov Chain (CTMC) models to predict and estimate botnet size (a Predictive Analytics type analysis using probability theory concepts)
– Wrote numerous scripts in Mathematica and Matlab for computing and analyzing the data
– Compared the results predicted by the developed analytical models with botnet measurement results reported by the Internet security industry (Damballa corporation's data)
– Analyzed the 4G cellular networks and determined a vulnerability in the LTE air interface
– Customized LTE-Sim, a C++-based 4G/LTE simulator, to conduct a simulation to determine the number of botnet nodes needed for a DDoS attack against the cellular network
– PhD Thesis containing analysis and data visualizations: www.masoodkh.com/PhD

High-level experience obtained:
– Gained a great insight into network security issues affecting communication networks
– Developed the capacity to devise solutions to predict and mitigate network security threats
– Enhanced project and time management skills while working on the phases of the project

Produced publications:
– "SComF and SComI Botnet Models: The Cases of Initial Unhindered Botnet Expansion", 25th Annual Canadian Conference on Electrical and Computer Engineering (CCECE12), April 29-May 2, 2012, Montreal, Canada, DOI: 10.1109/CCECE.2012.6334871
– "The SIC Botnet Lifecycle Model: A Step Beyond Traditional Epidemiological Models", Computer Networks (Elsevier), Special Issue on Botnet Activity: Analysis, Detection and Shutdown, Volume 57 (2013), Issue 2, pp. 404–421, DOI: 10.1016/j.comnet.2012.07.020
– "Botnets in 4G Cellular Networks: Platforms to Launch DDoS Attacks Against the Air Interface", 2013 International Conference on Selected Topics in Mobile and Wireless Networking (MoWNeT), 19-21 August 2013, Montréal, Canada.

Researcher/Programmer at Concordia University

September 2009 - December 2009 (4 months)

– Designed and developed a presence-based messaging application in Java (NetBeans IDE)
– Developed a graphical user interface for the client software using the IDE tools
– Project conducted as part of a Ph.D. program course

Lab Instructor/Researcher at ÉCOLE POLYTECHNIQUE DE MONTRÉAL

September 2007 - June 2009 (1 year 10 months)

– Served as lab instructor and teaching assistant for the course "Data transmission and digital

communication networks” (ele4704)
– Conducted experiments with network equipments and supervised students’ experiments
– Graded assignments and conducted exercise sessions for computer network problems
– Worked on a research project regarding the problems of P2P traffic in the access network reported by Bell Canada engineers
– Using the Eclipse IDE, customized and programmed in NS-2, a C++/TCL-based network simulator, to conduct a simulation studying the P2P traffic
–Wrote numerous scripts in Python for parsing, computing and graphing the data/trace files
– Proposed a model to help with the P2P traffic problems and presented it in the following paper: “UARA in Edge Routers: An Effective Approach to User Fairness and Traffic Shaping”, International Journal of Communication Systems (Wiley), 25: 169–184. DOI: 10.1002/dac.1262

Intern at Télécom ParisTech/INRIA
July 2006 - January 2007 (7 months)
– Designed the architecture and determined the necessary parameters of an IEEE 802.11 physical layer and propagation models in a prototype event-based network simulator
– Implemented the designed architecture using C++ in the network simulator
– Documented the project results for integration in the NS-3 network simulator project
– Gained experience in working with and integrating various telecom C++ libraries
– Project conducted for the M.Sc. thesis

Researcher/Programmer at Télécom ParisTech
January 2006 - June 2006 (6 months)
– Modeled Session Initiation Protocol (SIP) and Routing Information Protocol (RIP) using UML in Rhapsody (inc. partial C++ implementation of the protocol logic for code generation)
– Investigated IP Multimedia Subsystem (IMS), and its application servers, for advanced telecom service delivery; the prepared report on IMS helped with the securing of a grant from a major telecom company
– Projects carried out as part of the M.Sc. program

Researcher/Programmer at Iran University of Science and Technology
February 2004 - September 2004 (8 months)
– Implemented the Session Initiation Protocol (SIP) in the J-Sim network simulator
– Investigated the feasibility of using SIP, SDP, and T.38 protocols for real-time fax transmission on the Internet (Fax over Internet Protocol – FoIP)
– Conducted for the B.Sc. final project
– Presented the results in the following paper: “Utilizing DiffServ and SIP Contact Header for Real-time Fax Traffic Engineering”, 18th Annual Canadian Conference on Electrical and Computer Engineering (CCECE05), May 1-4, 2005, Saskatoon, Saskatchewan, Canada

Researcher/Programmer at Iran University of Science and Technology
October 2003 - January 2004 (4 months)

– Analyzed the performance of antenna structures using the HP HFSS software
– Supervised the construction of the antenna having the optimal performance
– Helped with the documentation of the results in three conference papers
– Project carried out as an extension of the work done during Antenna lab in the B.Sc. program

Intern at Fortex
July 2003 - September 2003 (3 months)

– Developed a simulator in C++ for a novel approach in matrix inversion calculation
– Internship done as part of the B.Sc. program

Certifications

Mining Massive Datasets
Coursera March 2015
Machine Learning
Coursera April 2015

Languages

English (Full professional proficiency)
French (Full professional proficiency)
Persian (Native or bilingual proficiency)

Education

Concordia University
Ph.D., Electrical and Computer Engineering (Computer and Telecom Networks), 2013
Télécom ParisTech (ENST)
M.Sc., Networked Computer Systems, 2007
Iran University of Science and Technology
B.Sc., Electrical Eng. - Telecommunications, 2004

Publications

Botnets in 4G Cellular Networks: Platforms to Launch DDoS Attacks Against the Air Interface
2013 International Conference on Selected Topics in Mobile and Wireless Networking (MoWNeT) 2013
Authors: Masood Krohy, Dongyu Qiu, Mustafa K. Mehmet Ali

Botnets are overlay networks built by cybercriminals from hacked smartphones and computers. In this paper, we report a vulnerability of the air interface of 4G cellular networks, the Long Term Evolution (LTE), to Distributed Denial-of-Service (DDoS) attacks launched from botnets. The attack scenario constitutes of a botmaster instructing the botnet nodes to start sending or downloading dummy data in order to overwhelm the air interface, thereby denying service for voice users. Through simulation using a capable LTE simulator,

we determine the number of botnet nodes needed per cell that can effectively render the cellular network unusable. Specifically, we show that a botnet that has spread to only 3% of subscribers is capable of lowering the voice quality from 4.3 to 2.8 in Mean Opinion Score (MOS) scale of 1 to 5 for scheduling strategies designed for real-time flows. On the other hand, a botnet that has managed to spread to 6% of subscribers can cause a MOS value of 1, i.e., a complete outage. The threat identified and the reported results could inspire the implementation of new mechanisms to ensure the security and availability of vital telecommunication services.

The SIC Botnet Lifecycle Model: A Step Beyond Traditional Epidemiological Models

Computer Networks Journal (Elsevier), Special Issue on Botnet Activity: Analysis, Detection and Shutdown February 2013

Authors: Masood Krohy, Mustafa K. Mehmet Ali, Dongyu Qiu

Botnets, overlay networks built by cyber criminals from numerous compromised network-accessible devices, have become a pressing security concern in the Internet world. Availability of accurate mathematical models of population size evolution enables security experts to plan ahead and deploy adequate resources when responding to a growing threat of an emerging botnet. In this paper, we introduce the Susceptible-Infected-Connected (SIC) botnet model. Prior botnet models are largely the same as the models for the spread of malware among computers and disease among humans. The SIC model possesses some key improvements over earlier models: (1) keeping track of only key node stages (Infected and Connected), hence being applicable to a larger set of botnets; and (2) being a Continuous-Time Markov Chain-based model, it takes into account the stochastic nature of population size evolution.

The SIC model helps the security experts with the following two key analyses: (1) estimation of the global botnet size during its initial appearance based on local measurements; and (2) comparison of botnet mitigation strategies such as disinfection of nodes and attacks on botnet's Command and Control (C&C) structure. The analysis of the mitigation strategies has been strengthened by the development of an analytical link between the SIC model and the P2P botnet mitigation strategies. Specifically, one can analyze how a random sybil attack on a botnet can be fine-tuned based on the insight drawn from the use of the SIC model. We also show that derived results may be used to model the sudden growth and size fluctuations of real-world botnets.

UARA in Edge Routers: An Effective Approach to User Fairness and Traffic Shaping

International Journal of Communication Systems April 2011

Authors: Masood Krohy

The ever-increasing share of the peer-to-peer (P2P) traffic flowing in the Internet has unleashed new challenges to the quality of service provisioning. Striving to accommodate the rise of P2P traffic or to curb its growth has led to many schemes being proposed: P2P caches, P2P filters, ALTO mechanisms and re-ECN. In this paper, we propose a scheme named 'UARA: User/Application-aware RED-based AQM' which has a better perspective on the problem: UARA is proposed to be implemented at the edge routers providing real-time near-end-user traffic shaping and congestion avoidance. UARA closes the loopholes exploited by the P2P traffic by bringing under control the P2P users who open and use numerous simultaneous connections. In congestion times, UARA monitors the flows of each user and caps the bandwidth used by

‘power users’ which leads to the fair usage of network resources. While doing so, UARA also prioritizes the real-time traffic of each user, further enhancing the average user quality of experience (QoE). UARA hence centralizes three important functionalities at the edge routers: (1) congestion avoidance; (2) providing user fairness; (3) prioritizing real-time traffic. The simulation results indicate that average user QoE is significantly improved in congestion times with UARA at the edge routers.

Skills & Expertise

Telecommunications
Java
SQL
Project Management
Statistics
Machine Learning
Deep Learning
Data Mining
Data Analytics
TensorFlow
Matlab
C++
Wireless Networking
Linux
Network Security
Mathematica
Python
R
Data Visualization
Hadoop
Technical Writing
Network Traffic Analysis
Network Simulation
VoIP
Microsoft SQL Server
Data Science
Business Intelligence
Network Benchmarking
Octave
Hive
Apache Spark
Hue
Scala

Masood Krohy

Big Data Analytics Consultant

m.kh@ieee.org

Contact Masood on LinkedIn