Paul C. Nichols

Contact Information (858) 245-3914 paul.c.nichols@gmail.com paulcnichols.com

Objective Challenging software engineering position with an amazing team.

Education

University of California, San Diego, La Jolla, California USA

M.S., Computer Science, December 2012 (expected)
Specialization in AI/Machine Learning

UCSD Extension, La Jolla, California USA

Certificate in Data Mining, January, 2010

University of California, San Diego, La Jolla, California USA

B.S., Computer Science, December, 2005

Relevant Coursework Artificial Intelligence (2 quarters), Data Mining, Learning Theory, Multivariate Analysis, Jacobs School of Engineering Management/Entrepreneurism (3 quarters), Software Engineering

Academic Experience Linear Regression, Logistic Regression, Bayesian Networks, Naive Bayes, Hidden Markov Models, Expectation Maximization, Clustering (k-Means), Support Vector Machine, Part of Speech Tagging (Conditional Random Fields), Collaborative Filtering (Matrix Factorization), Topic Modeling (LDA), Neural Networks (Back-Propagation), Profit Modeling, Random Forest, PCA, SVD

Technical Competencies C/C++, Perl, Python, Java, Matlab, R, Javascript, Django, PHP, Node.js, Chrome Plugin, Linux, Hadoop, MapReduce, MySQL, PostgreSQL, EC2, Linode, Virtualization, Visualization (D3, gnuplot), Networking

Professional Experience

Websense - Sr. Software Engineer **January, 2009 - present**

- Designed research infrastructure known as "The Pipeline" to fetch up to 60 Million URLs a day from customer log files and scan against security and content classifiers. Lead a team of 2 other engineers to implement. System spanned 14 IBM blades and utilized optimized C++ HTTP client, numerous perl scripts to move data, a very large distributed cache, and a query interface to search/filter output.
- Created robust, multi-process content scanning service in C++ to serve backend "Threat Seeker Network" (includes Pipeline) and customer facing "ACE Insight" lookup tool. Developed self extracting installer for easy deployment and backend service for datafile updating.
- Created custom NoSQL cache solution (before redis existed) to manage roughly 500 million URLs with only 16G of memory. Caching service ran on 8 blades and could handle 500 requests a second. Implemented in C++ in event driven style, requiring only 32 bytes per entry.
- Designed and created intuitive web-based query editor for domain specific language to search Pipeline output files. Language based on grammar consisting of if-elsif-else flow control, boolean expressions, and a large set of variables with transformation functions. Interface created with HTML, javascript, apache and perl and queries executed in C++ on JSON output files.
- Drove adoption of Hadoop to archive multiple months of customer data. Implemented tools in Java, C++, and perl to easily query stored data. Created several visualizations based on data. Presented to and trained team of 10 researchers on big data ecosystem and technologies.

- Investigated typo-squatting domains based on keyboard distance and look-alike characters. Implemented novel backtracking algorithm to permute most likely target domains for security scanning. Visualized connections between typo-squatting domains with network graph and showed common operators of domains.
- Researched and prototyped novel script classification system to detect variations of common javascript exploits added to compromised sites. Used linear SVM model with strong regularization and 500K lexical features to abstract away from randomized text. Trained using top 1 Million Alexa sites and 1 month's worth of detected javascript exploits. Achieved non-trivial averaged AUC and F1 over 5-fold cross validation. Built web-based lookup tool to classify URLs and display highlighted script by feature weight and bayesian likelihood ratio.
- Built Chrome plugin and backend logging process with node.js to monitor web requests and DOM events while actively browsing. Used to debug web controls and assess coverage of classification product on difficult to replay Web 2.0 style sites.

Nextwave - Software Engineer **September, 2007 - December, 2008**

- Developed diagnostics tool in C++ to manage WiMAX base-station provisioning, configuration, and statistics. Contributed to"handover" feature between mobile devices.
- Worked on porting x86/linux code to arm9/threadx and x86/windows by developing wrappers for signals, pthread rwlocks, fifos/pipes/mqs and select/poll in OS abstraction layer. Gained deep understanding of parallel, network, and linux systems programming by writing framework test cases.
- Drove integration of EAP authentication feature by contributing C/C++ code up and down the software stack from the modem firmware, host software, and wrapper for 3rd party EAP stack. Tested feature by working with base-station emulator group.

B.A.E. Systems, SOCET GXP - Software Engineer **Febuary, 2006 - September, 2007**

- Created database tool in M.S. Access to manage Program & Risk Management Department's records.
- Designed with UML and implemented various GUIs in C++/Qt for SGXP (Military GIS application). Coded import, export and some display capabilities for Mil-Std 2525 graphic overlays in SGXP using 3rd party C++ library. Worked in large team of 50 engineers.

Cardinal Health-Alaris Medical Systems - Intern **June, 2005 - October, 2005**

- Worked on minor features for medical fluid pump in embedded environment. Developed an interface to enter dosing information and a tool to retrieve diagnostic information in C/C++ and some C. Debugged and fixed de-bounce problem with system key pad handling routine.

Scripps Institute of Oceanography - Intern **April, 2004 - May, 2005**

- Designed system to ease data acquisition from scientific equipment such as LIDAR, camera and GPS using 3rd party API and GUIs written in Visual C++ . Wrote numerous data processing scripts. Created DLL to integrate with video processing software to add compression, GPS position data, and timestamp to frame data.

Intel Corporation - Intern **June, 2003 - September, 2003**

HOBBIES & INTEREST

Dogs, Surfing, Swimming (2 time All-American in High School), Music (Classical Guitar, Piano), Cycling, Chess, Hiking