

Tuan-Anh Hoang-Vu

2 MetroTech Center, 10th Floor, 10.034D, Brooklyn, NY 11201 | (917) 514-8879 | hvtuananh@gmail.com

EDUCATION

Ph.D. in Computer Science

Expected May 2017

New York University School of Engineering, Brooklyn, NY

Advisor: Juliana Freire, Ph.D.

GPA: 3.89

2nd place in ACM SIGMOD 2014 Programming Contest

B.S. in Computer Science

Jun 2010

Hanoi University of Science and Technology, Hanoi, Vietnam

Advisor: Thang Q. Huynh, Ph.D.

Thesis: *Opinion-Based Search: Dual Mapping Approach for Online Shopping*

TECHNICAL SKILLS

Python, C/C++, Java, PHP, HTML, JavaScript, MySQL, MongoDB,
Linux, bash, git, nginx, HAProxy, Amazon AWS (EC2, S3, EMR), Hadoop (HDFS, MapReduce), Crawler.

EXPERIENCES

Software Engineering Intern

May 2016 - Aug 2016

Uber Technologies Inc., San Francisco, CA

- Investigated and created PoC for different strategies to handle QoS at application and service levels.
- Developed a program to collect metrics from Mesos/Aurora and Docker services.
- Implemented a feature in Mesos v1 API to allow querying with metric types (MESOS-5731) [under review].

Research Assistant

Sep 2012 - Present

New York University School of Engineering, Brooklyn, NY

- Develop an efficient spatio-temporal textual index that handles spatial, temporal and textual components uniformly, archives high efficiency both in space usage and query evaluation.
- Develop *temporal* entity matching algorithms to integrate various urban datasets to discover underlying owners and ownership patterns.
- Collect news across top news websites in the US using *focused crawler* and associate news topics with tweets via domain adaptation.

Research Intern

Jun 2013 - Aug 2014

New York University Center for Urban Science and Progress, Brooklyn, NY

- Configured and maintained a 1000+ cores Hadoop cluster.
- Developed geocoder and reverse geocoder for NYC addresses.
- Designed a crawler to extract data from NYC Open Data portal.

Software Engineer

Feb 2009 - Jun 2012

Cazoodle Inc., Hanoi, Vietnam

- Developed a comprehensive software suite for creating, testing and managing *visual web crawler*.
- Designed and deployed a crawler scheduling system with 200+ clients that responsible for on-demand and scheduled web crawling.
- Developed an in-house automatic software update system.
- Implemented a lightweight-parsing framework to extract product information from web page on-the-fly.
- Developed a Firefox add-on (*Shopping Caddy*) that fetches offers, reviews and price history for a given product.
- Developed a set of API to be used by *Shopping Caddy* to interact with products database.
- Extracted and aggregated opinions from users' reviews and mapped into products to enable searching for products by opinions.

PUBLICATIONS

1. *A Unified Index for Spatio-Temporal Keyword Queries*. Tuan-Anh Hoang-Vu, Huy T. Vo, and Juliana Freire. In *CIKM 2016*.
2. *The More the Merrier: Efficient Multi-Source Graph Traversal*. Manuel Then, Moritz Kaufmann, Fernando Chirigati, Tuan-Anh Hoang-Vu, Kien Pham, Alfons Kemper, Thomas Neumann, and Huy T. Vo. In *VLDB Vol. 8, 2014*.
3. *Bridging Vocabularies to Link Tweets and News*. Tuan-Anh Hoang-Vu, Aline Bessa, Luciano Barbosa, and Juliana Freire. In *WebDB, 2014*. rev. 170217