

Fatemeh Borran

Chemin des Fleurs de Lys, 43C
CH-1350 Orbe
Switzerland

Mobile: +41-78-6867402
Email: fatemeh.borran@gmail.com
Website: <http://iict-space.heig-vd.ch/fhb/>

Software Engineer (Java, Scala, C/C++)
Big Data Analyst (MapReduce, Hadoop, Spark)
Distributed, Scalable and Fault-Tolerant System Designer

Education

- Feb. 2011 **Ph.D. in Computer Science** Swiss Federal Institute of Technology - Lausanne (EPFL)
“Round-based Consensus Algorithms, Predicate Implementations and Quantitative Analysis”
- Apr. 2006 **B.Sc. and M.Sc. in Computer Science** EPFL
Master thesis: “Efficient Semi-structured Queries in Scala using XQuery Shipping”

Work Experience

- Jan. 2012 - present **Scientific Collaborator** Intelligent Data Analysis Laboratory (IDA)
Institute for Information and Communication Technologies (IICT)
University of Applied Sciences of Western Switzerland (HEIG-VD)
- Apr. 2011 - Dec. 2011 **Postdoctoral Researcher** Distributed Systems Laboratory (LSR), EPFL
- Apr. 2006 - Apr. 2011 **Research and Teaching Assistant** LSR, EPFL
Swiss National Science Foundation (SNSF) grant
- Summer 2005 **Summer Internship** “Comparing Wireless Ad hoc Network Simulators”
National Center of Competence in Research - Mobile Information &
Communication Systems (NCCR-MICS) grant
- Summer 2004 &
Nov. 2004 - Mar. 2005 **Internship** “Decentralized Power Generation Expansion Planning”
Laboratory of Energy Systems (LASEN), EPFL

Projects

- **Current Projects** (Involved in more than 8 applied research projects since January 2012)
 - **CrowdStreams:** design a real-time traffic data monitoring system to estimate/forecast traffic flow in the neighbourhood of a social event, such as Paleo, using the Spark Streaming framework.
 - **Real-time Data Loss Prevention:** design and implementation of a content-based DLP solution for enterprise “data in motion” to construct real-time data leakage detection platform using Apache HBase, Solr and Spark computing for Connectivity Factory module at CrossingTech.
 - **CITATION:** design of a data middleware infrastructure on the Cloud for crowd monitoring project in the context of Internet of Things at IICT and implementation of a REST API using MongoDB and Python.
 - **Efficient Near Duplicate Detection:** design and implementation of an efficient tool to detect similar contents (partial duplications) in large text collections using the MapReduce programming model and evaluation using Apache Hadoop and Spark platforms in the context of security audit.
 - **SITG Catalog Clustering:** automatic clustering of geographical metadata for “Service des Systèmes d’Information et de Géomatique (SSIG) de l’état de Genève” using Apache Mahout in collaboration with innoQ.
 - **Social Media Monitoring:** sentiment analysis, event detection and trend evolution in social networks (e.g., Twitter) using advanced machine learning algorithms for SwissRe Life & Health.
 - **THESAURO:** automatic reorganization of the “Radio Télévision Suisse Romande” thesaurus using association rule mining (FPGrowth in Apache Mahout) in order to enrich audiovisual archives.
 - **GARSV:** semi-automatic text and video alignment tool for the sessions of the “Grand Conseil”.

- **Postdoctoral Projects** (2010 - 2011)
 - **Semi-Synchronous Algorithms:** designing algorithms that their progress is independent of the system timeout (publication).
 - **Fair Atomic Broadcast:** designing algorithms that their progress is independent of the behavior of faulty processes (technical report).
- **Doctoral Projects** (2006 - 2010)
 - Design, implementation and quantitative analysis of fault-tolerant distributed algorithms to provide reliable and highly available systems (two publications).
 - Development of a decentralized Byzantine consensus algorithm that outperforms the existing centralized algorithms in the worst case (four publications).
 - Simulation and performance evaluation of distributed algorithms in the context of wireless mobile ad hoc networks (publication).

Professional Activities

- **Lecturer** for “Creating Applications on the Internet (for Future Engineers)” HEIG-VD since 2012.
- **Teaching assistant** for more than 10 bachelor and master courses at HEIG-VD and EPFL since 2006.
- **Supervising** more than 10 bachelor and master students at HEIG-VD and EPFL since 2006.
- **Given talks in seminal conferences:** SRDS 2008, DISC 2009, ICDCN 2010 and DISC 2011.
- **External reviewer** for SRDS 2008, SRDS 2009, ICDCS 2010, SRDS 2011 and DSN 2011.
- **Reviewer** for IEEE Transactions on Mobile Computing 2012.

Technical Skills

- Programming Languages: Java (+10 years), Scala (+4 years), C/C++ (+3 years)
- Scripting Languages: Bash, Perl, Python (+2 years), Ansible (IT Automation)
- Big Data Analysis (+3 years): MapReduce, Apache Hadoop, Apache Spark, Storm
- Cloud Computing: Amazon EMR, Amazon EC2, OpenStack
- Information Retrieval (+3 years): Apache Lucene, Apache Solr
- Machine Learning (+2 years): Apache Mahout, Weka, Matlab, R
- Databases: MySQL, NoSQL: Apache HBase, MongoDB
- Web Development: HTML5, CSS3, PHP, JavaScript, Play Framework
- Integrated Development Environment: Eclipse, NetBeans, Xcode, Emacs
- Revision Control: CVS, SVN, git
- Operating Systems: Linux (Ubuntu and Red Hat), Mac OS, Microsoft Windows
- Reporting Applications: L^AT_EX, Microsoft Office

Language Skills

- French (fluent; living in Switzerland since 2001 and teaching in French since 2011)
- English (proficient in reading/writing scientific papers and good in speaking)

Publications

- List of publications are available at EPFL Infoscience (more than 10 scientific papers).
- List of citations are available at Google Scholar.

Strong letters of recommendation are available upon request.