

Sethuramalingam Subramaniam

Room No 149, New Boys Hostel,
International Institute of Information Technology,
Gachibowli,
Hyderabad, Andhra Pradesh, India – 500032

E-mail : sethu@research.iiit.ac.in
Mobile : +91 996688 2665
Web : <http://research.iiit.ac.in/~sethu>

Objective

Seeking for a challenging, research and engineering career in the fields of Information Retrieval and Machine Learning.

Education

- Master of Science (By Research) in Computer Science 2007 – 2009
International Institute of Information Technology, Hyderabad
- Thesis : “**Effective Query Translation Techniques for Cross-Language Information Retrieval**”
 - CGPA – **8.33 / 10**
- Bachelor of Engineering in Information Technology 2000 – 2004
Amrita Institute of Technology and Science, Coimbatore, Tamil Nadu, India
- Final Year Project : “**Text categorization using Proximal Support Vector Machines (PSVM)**”
Guided by: Dr. K.P. Soman
 - CGPA – **8.04 / 10**

Research Interests

- Information Retrieval
- Machine Learning
- Natural Language Processing

Thesis Overview

Effective Query translation Techniques for Cross-Language Information Retrieval

Supervised by: Dr. Vasudeva Varma

- Explored the existing ways of transliterating named entities across languages and proposed a mapping based method based of transliteration namely the Compressed Word Format (CWF) algorithm for English-Tamil transliteration
- Proposed a novel way to use the concept of Surface Similarity (phonetic and orthographic similarity) between the Indian languages in the form of fuzzy text search for Indian Language – Indian Language Cross language retrieval.

Experience

Graduate Research Assistant,

July 2007 – Current

Search and Information Extraction Lab (SIEL), IIIT Hyderabad

- Worked on ML algorithms like SVM, Naïve Bayes, Decision Trees, Neural Networks

Systems Engineer,

IBM, Pune

**August 2006 – July 2007
(~1 year)**

- Worked on a data-warehousing project for IBM Financial Systems, US
- Involved in building of data models, and the database design
- Technologies involved – Core Java, DataStage, IBM DB2

Senior Software Engineer,

iGATE Global Solutions, Chennai

**June 2004 – August 2006
(2 years and 2 months)**

- Worked on a development cum maintenance project for General Electronics' subsidiary.
- Involved in building of large scale, enterprise applications using J2EE technologies
- Received client appreciations for commendable job performed in various applications
- Technologies involved – Servlets, JSP, Apache Struts, log4j, JMS, DB2, JDBC, JMS, Websphere, Tomcat

Related Coursework

- Information Retrieval and Extraction
- Machine Learning
- Pattern Recognition
- Natural Language Processing
- NLP Applications: MT&IE

Relevant Course Projects

- **Question Classification Using Support Vector Machines**
Machine Learning, Monsoon 2008
- **Tamil – English CLIR system on the CLEF 2007 dataset**
Information Retrieval and Extraction, Monsoon 2007
- **Nested Named Entity Recognizer for Hindi**
NLP Applications: MT&IE, Spring 2008

Software Tools

- **Sanchay – An Open Source Platform for South Asian languages**
 - Integrated the Transliteration module in Hindi and extended it for Bengali

Publications & Presentations

- Sethuramalingam Subramaniam and Vasudeva Varma. **Effective Query Translation Techniques for Cross-Language Information Retrieval**. Special Issue of the ACM Transactions on Asian Language Information Processing (TALIP) on IR for Indian languages [Submitted]
- Anil Kumar Singh, Sethuramalingam Subramaniam and Tarak Rama. **Multilingual Transliteration and Crosslingual Information Retrieval: The Case of English to Indian Languages**. Special Issue of the ACM Transactions on Asian Language Information Processing (TALIP) on IR for Indian languages [Submitted]
- Praneeth Shishtla, Surya Ganesh V, Sethuramalingam Subramaniam and Vasudeva Varma. **A Language-Independent Transliteration Schema Using Character Aligned Models**. In Proceedings of ACL-IJCNLP 2009 Named Entities Workshop (NEWS 2009), Singapore, August 2009.
- Sethuramalingam Subramaniam, Anil Kumar Singh, Pradeep Dasigi, and Vasudeva Varma. **Experiments in CLIR Using Fuzzy String Search Based on Surface Similarity**. In Proceedings of the 32nd Annual ACM SIGIR Conference. Boston, Massachusetts, July 2009.
- Sethuramalingam S and Vasudeva Varma. **IIIT Hyderabad's CLIR experiments for FIRE-2008**. In Working notes of Forum for Information Retrieval Evaluation (FIRE), Kolkata, India, December 2008.
- Srinivasan Janarthanam, Sethuramalingam S and Udhyakumar Nallasamy. **Named Entity Transliteration for Cross-Language Information Retrieval using Compressed Word Format mapping algorithm**. Proceedings of 2nd International ACM Workshop on Improving Non-English Web Searching (iNEWS08), CIKM, Napa Valley, California, USA, 2008.
- Sethuramalingam S, K.R. Subramaniam, **Text Categorization using SVM**, In Microsoft Techsplash, a National level Technical Symposium held at NIT Suratkal, India, February 2004.

Computer Skills

- Programming Languages : Java, C, Perl
- Toolkits : Lucene, OpenNLP, SVM Light, Weka, GIZA++
- Operating Systems : Linux, Windows

Awards & Honours

- SIGIR & Microsoft Research India student travel grant to attend SIGIR 2009
- Selected as an international student volunteer for SIGIR 2009
- Secured 3rd position in Java Jargon (a Java Programming Contest) in Felicity, IIIT 2008
- Secured 200/200 in mathematics in the higher secondary school examinations
- School topper in academics in the secondary school examinations

References

Dr. Vasudeva Varma

Associate Professor,
Search and Information Extraction Lab (SIEL),
IIIT, Hyderabad, India – 500032
Mail: vv@iiit.ac.in

Anil Kumar Singh

Research Scientist,
Language Technologies Research Centre,
IIIT, Hyderabad, India – 500032
Mail: anil@research.iiit.ac.in