Hui Guo

1701 GORMAN ST., APT 205
RALEIGH, NC 27606
(919) 360-5662
hguo5@ncsu.edu (preferred), ecuguo@gmail.com
https://hguo5.github.io/

Professional Summary

I am currently a PhD candidate in the Department of Computer Science at North Carolina State University. I plan to graduate in early 2021. I am interested in natural language processing, text mining, deep learning, and social computing. I am working on projects that are related to event extraction and inference in texts related to software development, such as app reviews.

Education

JAN 2015 - PRESENT
North Carolina State University, Raleigh, NC - *Current PhD student*
Computer Science. GPA 4.00

AUG 2012 - MAY 2014
East Carolina State University, Greenville, NC - *Master of Science*
Computer Science. GPA 4.00

AUG 2005 - JUL 2009
Tsinghua University, Beijing, China - *Bachelor of Engineering*
Automation Engineering. GPA 3.78

Publications

[C] **Lin: Unsupervised Extraction of Tasks from Textual Communication**, Parth Diwanji, Hui Guo, Anup K. Kalia, and Munindar P. Singh. *COLING*, 2020 (short paper, accepted)

[C] **Caspar: Extracting and Synthesizing User Stories of Problems from App Reviews**, Hui Guo and Munindar P. Singh. *ICSE*, 2020 [link] [pdf]

[C] **Elessar: Ethics in Norm-Aware Agents**, Nirav Ajmeri, Hui Guo, Pradeep K. Murukannaiah, and Munindar P. Singh. *AAMAS*, 2020 [link] [pdf]

[C] **A Framework for Word Segmentation in Images using Density-based Clustering**, Hui Guo and Qin Ding. *CATA*, 2020 [link] [pdf]

[J] **Çorba: Crowdsourcing to Obtain Requirements from Regulations and Breaches**, Hui Guo, Ozgur Kafali, and Munindar P. Singh. *EMSE*, 2019 (and ICSE 2020 Journal First) [link] [pdf]

[W] **Extraction of Natural Language Requirements from Breach Reports Using Event Inference**, Hui Guo, Ozgur Kafali, and Munindar P. Singh. *AIRE*, 2018 [link] [pdf]

[C] **Robust Norm Emergence by Revealing and Reasoning about Context: Socially Intelligent Agents for Enhancing Privacy**, Nirav Ajmeri, Hui Guo, Pradeep K. Murukannaiah, and Munindar P. Singh. *IJCAI*, 2018 [link] [pdf]

[Poster] **Ethics, Values, and Personal Agents: Poster,** Nirav Ajmeri, Hui Guo, Pradeep K. Murukannaiah, and Munindar P. Singh. *HoTSoS*, 2018 [link] [pdf]

[Poster] **Toward Extraction of Security Requirements from Text: Poster,** Hui Guo, Özgür Kafalı, Anne-Liz Jeukeng, Laurie Williams, and Munindar P. Singh. *HoTSoS*, 2018 [link] [pdf]

[J] **Designing Ethical Personal Agents,** Nirav Ajmeri, Hui Guo, Pradeep K. Murukannaiah, and Munindar P. Singh. *IEEE Internet Computing*, 2018 [link] [pdf]

[J] **Teaching Crowdsourcing: An Experience Report,** Hui Guo, Nirav Ajmeri, and Munindar P. Singh. *IEEE Internet Computing*, 2018 [link] [pdf]

[C] **Arnor: Modeling Social Intelligence via Norms to Engineer Privacy-Aware Personal Agents,** Nirav Ajmeri, Pradeep K. Murukannaiah, Hui Guo, and Munindar P. Singh. *AAMAS*, 2017 [link] [pdf]

[C] **miRDiabetes: A microRNA-Diabetes Association Database Constructed by Classification on Literature,** Hui Guo and Qin Ding. *BICoB*, 2016 [link] [pdf]

[J] **The Evolution of Photosynthesis in Chromist Algae through Serial Endosymbioses,** John W. Stiller, John Schreiber, Jipei Yue, Hui Guo, Qin Ding, and Jinling Huang. *Nature Communications*, 2014 [link] [pdf]

Skills

Python, JAVA, and C. TensorFlow, Keras, etc. Eager to learn more.

Awards

- **Computer Science Outstanding Research Award.** 2020. Department of Computer Science, North Carolina State University
- **Computer Science Outstanding TA Award.** 2017. Department of Computer Science, North Carolina State University
- **Robert Clifford Rogers Memorial Award in recognition of Outstanding Graduate Scholars.** 2014. Department of Computer Science, East Carolina University
- **Awards for academic excellence.** 2005 and 2006. Tsinghua University

Projects

JUN 2020 - PRESENT

Understanding Story Structures in App Reviews

- App reviews, especially negative ones, describe app users' interaction stories with the app. These stories follow different structures, and the structures may indicate the type of information app reviews contain. We extract events from app reviews, classify and order them, and combine the related events into coherent stories. We then analyze the common story structures in different types of app reviews. The resulting paper of this project has been submitted to ICSE 2021.

AUG 2018 - PRESENT

Identifying Competitors and Comparisons in App Reviews

- In app reviews for an app, users mention other apps and make comparisons between the apps. Based on app reviews and app descriptions, we identify competitors of a given app. We then extract comparative sentences from app reviews, identify the preferred apps, and tabulate the "votes" to obtain the relative ranking of apps. We are currently finalizing the results and paper.

JUN 2018 - PRESENT

Ordering Event Pairs by Commonsense Inference

- Previous studies order pairs of events by statistically analyzing their appearances in texts. However, humans can order a lot of event pairs by commonsense inference without knowing their context. I focus on such type of event pair ordering. We extract pairs of event phrases from everyday life stories, employ human annotators to order them, and build a new neural network for the automated ordering. The resulting paper of this project has been submitted to EMNLP 2021.

JUN 2019 - AUG 2019

Extracting and Inferring User-Action-App-Problem Pairs from App Reviews

- Users describe their interaction stories in app reviews, especially in those with negative ratings. They describe what they did and how the app reacted. Instead of collecting bug-reporting app reviews in their entirety, we focus on extracting and inferring user-action app-problem pairs from app reviews. The idea is that such type of extraction will yield more informative results for the developers to fix the app problems. Inference on such event pairs can help developers preemptively address such problems. The resulting paper has been accepted by ICSE 2020.

JUN 2018 - DEC 2018

Information Extraction from Legal Document

- I worked at Ernst & Young's Global Innovation office in Palo Alto, CA for six months. I worked on a project that tried to extract key information from legal documents. We used classification and sequence labeling techniques to mark targeted information in the documents. This project involved LSTM networks (Tensorflow), and word and feature embedding methods.

DEC 2017 - APR 2018

Event Prediction for the Analysis of HHS Breach Reports

- HHS breach reports describe incidents in which protected health information has been lost, misused, or impermissibly disclosed, and contain actions that the responsible parties have taken after them. We trained neural network models on the breach reports to infer common actions that follow a certain breach, aiming to give insights to the responsible parties. This project involved techniques and libraries of natural language processing (NLTK), Doc2Vec (deeplearning4j), and LSTM networks (TensorFlow). The resulting paper has been published in AIRE 2018 workshop.

MAY 2017 - DEC 2017

Norm Extraction from HHS Breach Reports and HIPAA Clauses Using Crowdsourcing

- We designed crowdsourcing projects and employed crowd worker to analyze a selected set of HHS breach reports and some related HIPAA clauses. We asked survey questions in terms of norms, and examined the effectiveness of crowdsourcing in the norm extraction task. The projects were deployed on Amazon Mechanical Turk, and surveys were conducted on a JSP website hosted on an Apache Tomcat server. The resulting paper has been published in EMSE in summer 2019.

NOV 2016 - DEC 2017

Enriching Interactions via Context Sharing in Normative Multi-agent Systems

- This project evaluated the performance of socially intelligent personal agents (SIPAs), which share the context of their users when they deviate from social norms, in normative multi-agent systems. We ran simulations of such a SIPA, Ringer, which determines whether or not to ring loud the phone based on users' context when there is an incoming phone call. The simulations were conducted using the Multiagent Simulation Toolkit (MASON) in JAVA. The resulting paper has been published in IJCAI 2018.

OCT 2016 - NOV 2016

Arnor: Modeling Social Intelligence via Norms to Engineer Personal Agents

- We propose Arnor, a framework for engineering socially intelligent personal agents via social norms. We evaluated this framework by comparing the performance of programs based on Arnor against a baseline. We have found that Arnor agents provide improved privacy-preserving social experience to end users. The resulting paper has been published in AAMAS 2017.

SEP 2016 - NOV 2016

Hummed Song Recognition Using Crowdsourcing

- This project was a course project of the Social Computing course in NCSU, for which I was the TA. I designed and organized this assignment. In this assignment, students were required to crowdsource recognition tasks of hummed songs, which I created, using their classmates as crowd workers. This assignment was logistically challenging, but we carried it out with high student satisfaction. We have published an article about this teaching experience in IEEE Internet Computing.

FEB 2014 - MAY 2014

miRDiabetes: A Database of Literature for microRNA-Diabetes Associations

- As my master thesis, I have built a database and a website for papers that study the associations between microRNA and diabetes from PubMed. Based on features extracted from their abstracts, relevant papers are retrieved from PubMed using Logistic Model Tree, a classification technique. The retrieved papers are then manually verified to increase the precision of the literature retrieval process. The resulting paper has been published in BICoB 2016.

AUG 2013 - JUL 2014

Analysis of BLAST Output to Study Evolution of Photosynthesis in Chromist Algae

- We used the Basic Local Alignment Search Tool (BLAST), which output the most similar genomes to a query genome. We examined theses “hits” among several major algal lineages to investigate the evolution of photosynthesis in them. In this project, I was in charge of programmatically parsing the output, maintaining data sets, and querying the database. The resulting paper has been published in Nature Communications.

AUG 2012 - MAY 2015

Various Course Projects

- I have taken numerous graduate-level courses in which high-level projects were required. My projects include: a website built with C# and ASP.NET, an implementation and solution attempt of the Sokoban game in JAVA, cursive handwriting recognition using JAVA and MatLab, index-based subtree search in large-scale datasets, and so on.

Experience

JAN 2017 - PRESENT

North Carolina State University, Raleigh, NC – *Research Assistant*

- I have been working as a research assistant under the advisement of Dr. Munindar P. Singh. My previous projects focused on analysis of HIPAA regulations and HHS breach reports using crowdsourcing, natural language processing tools (e.g. Apache OpenNLP toolkit and NLTK), and machine learning techniques (e.g., classification, word embedding, deep

learning). My current projects focus on event extraction and inference from app reviews, using word embedding and deep learning techniques and tools.

JUN 2018 - DEC 2018

Global Innovation, Ernst & Young, Palo Alto, CA – *AI Research Intern*

- Worked as a research intern in the AI group. My responsibilities included building and experimenting on methods that extracted key information from legal documents. I used machine learning techniques including classification and sequence labeling. The proposed methods included tools and techniques like word embedding, SpaCy, part-of-speech tagging, LSTM network, MLP, TensorFlow, and so on.

JAN 2015 - MAY 2019

North Carolina State University, Raleigh, NC – *Teaching Assistant*

- Worked as a teaching assistant for Data Intensive Computing, Discrete Mathematics, Introduction to Algorithms, Artificial Intelligence, and Social Computing courses. I was awarded the Outstanding TA Award for my work in 2017.

SEP 2014 - DEC 2014

Zhong Xun Hu Lian Co, Shandong, China – *Programmer*

- Worked full-time at this IT company, building websites and writing web pages using HTML, CSS, JavaScript, JQuery, and AJAX.

AUG 2012 - MAY 2014

East Carolina University, Greenville, NC – *Research & Teaching Assistant*

- Worked as a technician under Dr. Qin Ding in a cooperative research between the Department of Biology and Department of Computer Science, in which I was in charge of the creation and maintenance of a MySQL server.
- Worked as a teaching assistant for Database Management and Introduction to JAVA courses.

AUG 2009 - MAY 2012

ITS-UNC, Chapel Hill, NC – *Intern*

- Worked as an intern programmer, creating desktop applications using C# and Microsoft SQL Server.