

HONGYUAN GAO

hongyuag@usc.edu · github.com/hikoship · hgao.net
(213) 245-5576 · 2632 Ellendale Pl., Los Angeles, CA 90007

EDUCATION

University of Southern California M.S. in Computer Science, GPA: 4.0 / 4.0	May 2018 <i>Los Angeles, CA</i>
Shanghai Jiao Tong University B.S. in Computer Science, GPA: 3.5 / 4.0	June 2016 <i>Shanghai, China</i>

TECHNICAL SKILLS

Programming Languages	Python, C/C++, Java, JavaScript/Node.js, PHP
Web Frameworks	Bootstrap, AngularJS, Express.js, Django, Tornado
Tools & Platforms	MySQL, SQLite, Git, Apache, Nginx, Android Studio, Vim

EXPERIENCE

- | | |
|---|--|
| Embedded and Pervasive Computing Center of SJTU
<i>Research Assistant</i> | March 2015 - June 2016
<i>Shanghai, China</i> |
|---|--|
- Devised an efficient chunk-based cache algorithm for data deduplication in backup systems based on spacial locality; improved cache hit ratio by 5.2% and reduced I/O latency by 17.3%.
 - Combined the algorithm with container technique and saved memory consumption by 50.7%.
 - Won A-Class Bachelor's Thesis Award (top 10%) and published a paper in IEEE ICPADS16.
- | | |
|--|--|
| Siemens Manufacturing and Engineering Center Ltd.
<i>Intern, IT Department</i> | January 2015 - February 2015
<i>Shanghai, China</i> |
|--|--|
- Developed a GUI lucky draw system for staff in Python with Tkinter GUI package for the year-end party.

PROJECTS

- | | |
|---|----------------------------|
| Forum Posts Crawler (JavaScript, Python) | December 2015 - April 2016 |
|---|----------------------------|
- Utilized Beautiful Soup to parse HTML files of 1point3acres (A forum for graduate school applicants) to JSON and visualized over 30,000 application results, which got about 1,000 visits from 17 countries within one month.
 - Developed a data filter based on AngularJS and a query API based on Tornado web framework.
 - Created a daemon program to manage and protect the crawler; deployed the daemon on DigitalOcean VPS.
- | | |
|---|-------------------------------|
| The US Congress Database: An Android App (Java, PHP) | November 2016 - December 2016 |
|---|-------------------------------|
- Visualized the US Congress data with a user-friendly interface based on Material Design language.
 - Constructed a powerful API to query thousands of legislators and bills data; hosted the API on AWS.
 - Designed 7 interfaces for users to search the data and save preferences.
- | | |
|---|----------------------|
| Kickoff: A Soccer Statistics Website (JavaScript, PHP) | May 2015 - June 2015 |
|---|----------------------|
- Established a large-scale database for thousands of UEFA soccer teams, players and matches.
 - Developed a crawler to collect matches data from UEFA websites every day.
 - Designed an interface for users to follow their favorite teams and give ratings to players after matches.
 - Proposed an algorithm to select most valuable players according to their performances and ratings.
- | | |
|---|--------------------------|
| PDF Translation Platform (JavaScript, PHP) | January 2015 - June 2015 |
|---|--------------------------|
- Collaborated with a team member and built an online platform for reading and translating public PDF files.
 - Created an intuitive front-end interface for translators to select sections they want to translate.
 - Devised an algorithm to bypass special characters and stored multiple versions of translations in database.
- | | |
|--|-----------------------|
| Large-scale Patents Classifier (Python) | March 2015 - May 2015 |
|--|-----------------------|
- Developed a supervised learning model for classification using LIBLINEAR and LIBSVM library.
 - Led a 3-member group and constructed a multi-thread large-scale patents classifier with thread pool; achieved accuracy of 96.81% by applying Min-Max Modular SVM.
- | | |
|---|---------------------------|
| Netflix Movie Recommendation System (Python) | January 2015 - March 2015 |
|---|---------------------------|
- Predicted missing ratings of users based on given ratings in Netflix Prize Dataset.
 - Reduced the standard deviation of Netflix movie recommendation system to 0.90 (Top 5% in the class) using multiple optimization algorithms such as stochastic gradient descent and collaborative filtering.

PUBLICATIONS

H. Gao et al. An Efficient Cache Algorithm to Accelerate the I/O Processing of Data Deduplication in Storage Systems. In *Proc. of the IEEE ICPADS16*, 2016.