张瑞昌

联系方式	大学：复旦大学，上海 **Email:** zhangruichang112@gmail.com 技术博客：http://zhangruichang.com 个人主页：http://admis.fudan.edu.cn/~rczhang	**求职意向**：软件开发,数据挖掘 **手机**：(+86) 150-0022-6997 **毕业时间**：**2015年6月毕业**

教育背景　　　计算机软件与理论硕士, **复旦大学**,上海　　　　　　　　2012年9月∼ 2015年6月
- 上海市智能信息处理实验室, (**C语言 TA**),导师：周水庚教授

计算机科学与技术学士, **上海大学**,上海　　　　　　　　2008年9月∼ 2012年6月
- GPA: **3.678**/4, 排名: **1**/227

研究兴趣　　　**机器学习，数据挖掘**。翻阅了几本经典的计算机书籍，包括编程之美，编程珠玑，算法导论等，对算法、数据结构较感兴趣

实习　　　新加坡**南洋理工大学，数据挖掘项目**　　　　　　2014年3月∼2014年8月
- 提取合成致死基因对相关的特征，包括蛋白质网络拓扑结构、TCGA基因表达数据等，针对人类已有的合成致死对数量较少，以及没有可靠负样本的情况，采用**PU learning** 对已有的样本进行训练，通过在酵母的数据集上进行验证，来验证提取的特征以及PU learning方法的有效性。

EMC 夏令营, 上海　　　　　　　　　　　　　　　　　2014年8月

项目　　　基于主题模型和聚类算法的宏基因组归类问题研究　　　2013年9月∼2014年3月
- 采用**SKWIC聚类算法**对元基因组序列聚类, SKWIC在K-means基础上，额外考虑了每个feature对于每个cluster的一个权重，使得clustering效果更加显著。整合两个项目成员的Java 和C++ 代码，发布了Windows 和Linux 上运行的两个可执行程序, 文章发表在IEEE 的TCBB 期刊上，在会议上做oral presentation

- https://github.com/zhangruichang/MCluster

- 采用**主题模型-隐含狄利克雷分布(LDA)**，k-mer表示word，sequence表示document。首先从序列中抽取k-mer frequency，然后通过训练LDA 模型将k-mer 空间表示的序列变换为topic空间表示的序列，最后用SKWIC算法进行聚类。文章被会议接收并转投BMC Bioinformatics 期刊。

用户行为收集及分析系统　　　　　　　　　　　　　2012年2月∼2012年6月
- C♯实现的Windows app，监控用户在IE浏览器上的行为，包括鼠标行为(左击，右击，双击等)和标签页行为(打开一个页面，隐藏，关闭这个页面等)。

- 分析用户在网页上的有效浏览时间，并抽取网页的正文，内部实现了**网页正文抽取算法**。算法剔除html标签等正文无关内容，定义行块大小概念表示连续几行html文件的字符数，基于网页正文区域文字信息最密集的依据，将网页正文抽取问题转化为行块分布函数的骤升骤降边界点的确定。

- 生成语义完整的行为序列，序列以打开页面开始，随后包含左击，隐藏，激活等等行为，并最终以关闭页面结束。

- https://github.com/zhangruichang/UserBehaviorCollect

文章　　　**Ruichang Zhang**, Zhanzhan Cheng, Jihong Guan, Shuigeng Zhou*. "Exploiting Topic Modeling to Boost Metagenomic Sequences Binning," *10th International Symposium on Bioinformatics Research and Applications* (ISBRA 2014), vol.8492, LNCS, Springer.(*12th Asia Pacific Bioinformatics Conference* (APBC2014), Poster.)

Ruiqi Liao, **Ruichang Zhang**, Jihong Guan, Shuigeng Zhou*. "A New Unsupervised Binning Approach for Metagenomic Sequences Based on N-grams and Automatic Feature Weighting," *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 11(1): 42-54 (2014).

Hui Liu, **Ruichang Zhang**, Wei Xiong, Jihong Guan, Ziheng Zhuang, Shuigeng Zhou*. "A Comparative Evaluation on Prediction Methods of Nucleosome Positioning," *Briefings in Bioinformatics*, Doi: 10.1093/bib/bbt062, 2013.

Shusi Yu, **Ruichang Zhang**, Jihong Guan*. "Properly and Automatically Naming Java Methods: A Machine Learning Based Approach," *The 8th International Conference on Advanced Data Mining and Applications* ADMA 2012:235-246.

荣誉和竞赛	**威睿奖学金**, *复旦大学*	2014
	EMC 极客挑战赛一等奖, *EMC公司*	2014
	腾讯微爱创新奖学金, *复旦大学*	2013
	硕士一等学业奖学金, *复旦大学*	2013
	上海市优秀毕业生, *上海大学*	2012
	百优团员, *上海大学*	2011
	优秀学生, *上海大学*	2011
	光华奖学金, *上海大学*	2011
	特等学业奖学金, *上海大学*	2011
	一等学业奖学金, *上海大学*	2009, 2010
	二等奖, *首届全国大学生数学竞赛*	2009

技能	**语言**: C++ > C > Java = C♯ =Visual Basic = Matlab，etc
	工具和环境: Linux, LaTeX, Git，etc
	外语: 英语, CET4 578，CET6 554

Ruichang Zhang

Contact	**University: Fudan University**, Shanghai **Job Intension**: Software Development, Data Mining

University: Fudan University, Shanghai   **Job Intension**: Software Development, Data Mining

Email: zhangruichang112@gmail.com   **Phone**: (+86) 150-0022-6997

Blog: `http://zhangruichang.com`   **Graduation Date**: Jun. 2015

Personal Page: `http://admis.fudan.edu.cn/~rczhang`

Education

Fudan University, Master Degree in Computer Science   Sep.2012~Jun.2015
- Shanghai Key Lab of Intelligent Information Processing, **TA of C language programming**
- Advisor: Prof. Zhou Shuigeng

Shanghai University, Bachelor Degree in Computer Science   Sep.2008~ Jun.2012
- GPA: **3.678**/4, Rank: **1**/227

Research and Interests

Machine learning, data mining. Interested in data structure and algorithm, read some parts of several classical books, such as Beauty of Programming, Programming Pearls, Introduction to Algorithm, etc.

Internship and Experience

Nanyang Technological University, Singapore   Mar.2014~Aug.2014
- **Data mining project**. Extract features related to synthetic lethality pairs(SLs), including PPI topological features and TCGA gene expression data. As synthetic lethality in human cancers are limited, and reliable negative samples are not available, exploit **PU learning** to train positive samples and test on abundant SLs in yeast to validate the effectiveness of features and PU learning method.

EMC Summer Campus, Shanghai   Aug.2014

Projects

Metagenomic Sequences Binning Based on Topic Model   Sep.2013 ~Mar.2014
- Exploit **SKWIC algorithm** to cluster metagenomic reads. SKWIC is a variant of K-means, which takes the weight of each feature to each cluster into consideration, thus boosting the performance of clustering. Integrate C++ and Java codes of two members, and released two executables on Windows and Linux, paper is published on IEEE journal *TCBB*, and give a presentation on a conference.
- `https://github.com/zhangruichang/MCluster`
- Exploit **topic model-LDA**, k-mer represents word, sequence represents document. Firstly, extracting k-mer frequencies from sequences, then transforming reads from k-mer space to topic space, and finally clustering reads with SKWIC algorithm. Paper was accepted by a conference and submitted to *BMC Bioinformatics*.

User Behavior Collecting and Analysing System   Feb.2012~Jun.2012
- Windows app using C♯, monitor user's actions on IE browser, including mouse actions(left-click, right-click, double-click, etc) and tab actions(opening a tab, hiding, activating and closing the tab, etc)
- Analyze user's real browsing time on a page, and extract the content, including a **web page content extraction algorithm** inside. The algorithm first removes html tags which are not related to content, and defines the concept of line block size as the number of characters in several continuous lines, and extract the densest area as approximately content is the densest area in the html file.

- Generate complete sequences of actions, a sequence begins with opening a tab, followed by left-click, hiding, activating and finally ends with closing this tab
- https://github.com/zhangruichang/UserBehaviorCollect

Publication	**Ruichang Zhang**, Zhanzhan Cheng, Jihong Guan, Shuigeng Zhou*. "Exploiting Topic Modeling to Boost Metagenomic Sequences Binning," *10th International Symposium on Bioinformatics Research and Applications* (ISBRA 2014), vol.8492, LNCS, Springer.(*12th Asia Pacific Bioinformatics Conference* (APBC2014), Poster.)
	Ruiqi Liao, **Ruichang Zhang**, Jihong Guan, Shuigeng Zhou*. "A New Unsupervised Binning Approach for Metagenomic Sequences Based on N-grams and Automatic Feature Weighting," *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 11(1): 42-54 (2014).
	Hui Liu, **Ruichang Zhang**, Wei Xiong, Jihong Guan, Ziheng Zhuang, Shuigeng Zhou*. "A Comparative Evaluation on Prediction Methods of Nucleosome Positioning," *Briefings in Bioinformatics*, Doi: 10.1093/bib/bbt062, 2013.
	Shusi Yu, **Ruichang Zhang**, Jihong Guan*. "Properly and Automatically Naming Java Methods: A Machine Learning Based Approach," *The 8th International Conference on Advanced Data Mining and Applications* ADMA 2012:235-246.

Honors	**VMware Award**, *Fudan University*	2014
	1st Prize in EMC Guru Contest, *EMC Corp.*	2014
	Tencent Innovation Award, *Fudan University*	2013
	First Class Scholarship, *Fudan University*	2013
	Shanghai Outstanding Graduate Student, *Shanghai University*	2011
	Excellent League Member, *Shanghai University*	2011
	Excellent Student, *Shanghai University*	2011
	Guanghua Scholarship, *Shanghai University*	2011
	Special Scholarship, *Shanghai University*	2011
	First Scholarship, *Shanghai University*	2009, 2010
	Second Prize, *National Mathematics Contest for Undergraduate Student*	2009

Skills	**Computer Languages**: C++>C>C♯ = Visual Basic = Java = Matlab,etc
	Tools && **Environments**: Linux, LaTeX, Git, etc
	Foreign Language: English, CET4 578, CET6 554