

## 求职意向

利用机器学习, 人工智能, 数据挖掘, 知识图谱等技术, 服务真实有趣的应用场景。

## 教育及工作背景

- 2015.10 – 至今 **机器学习/大数据研究员**, 华为, 深圳
- 2014.10 – 2015.3 **访问学者**, 计算基础理论实验室, 信息学院, 爱丁堡大学, 英国  
导师: 樊文飞教授 (ACM Fellow)
- 2011.8 – 2015.10 **哲学博士**, 计算机系, 香港浸会大学, 香港  
导师: 蔡冠球教授, 徐建良教授
- 2010.8 – 2010.10 **实习后台开发工程师**, 腾讯, 深圳
- 2007.9 – 2011.8 **工程学士**, 计算机科学与工程学院, 华南理工大学, 广州  
(GPA – 86.3/100, 3.7/4.0)

## 项目经历(部分)

- 2017.4 – 至今 **基于机器学习/数据挖掘的收入洞察算法研发**。  
华为的在线计费系统 (Online Charging System OCS) 为全球近 200 个运营商局点提供收入管理服务, 其中收入洞察 (Revenue Insights) 将在未来成为不可或缺的增值服务。我们主要基于机器学习及数据挖掘 (LookALike, MNL, XGBoost, DNN 等), 刻画用户画像 (特别是用户消费行为画像), 为用户推荐最合适的套餐的同时, 预测套餐未来的销售收入, 给出适当的套餐优化建议, 最终使得套餐配置达到最优, 提升运营商整体收入及用户满意度等。
- 2016.6 – 2017.4 **基于机器学习/数据挖掘的计费系统数据搬迁算法研发**。  
在老系统更新/升级过程中, 由于新老系统数据模型的差异, 导致数据搬迁 (Data Migration) 的难度高, 工作量大。如华为 OCS 的搬迁涉及用户数据和资费数据, 每年需搬迁约 10+ 局点, 单局点平均耗费 2000 人天。为了解决搬迁难问题, 我们采用机器学习 (XGBoost, DNN 等), 预测新老系统数据模型字段的匹配规则 (即 Schema Mapping), 提升用户数据搬迁效率; 并利用海量话单日志, 自研老系统套餐资费推演算法, 提升资费数据搬迁效率。总体上, 单局点可提升 20%+ 搬迁效率, 为公司节省大量人力成本。相关成果将在今年投稿国际顶级数据挖掘会议, 并将输出专利。
- 2015.10 – 2016.6 **基于在线学习的推荐点击率预测算法研发**。合作方: **Eric Xing** (卡耐基梅隆大学)  
基于华为当前 IPTV 视频推荐系统的三层架构 (即离线, 近线及在线层), 针对 IPTV 用户行为偏好变化快的特点, 将原有在线层的 Google FTRL 推荐点击率预测算法 (业界最为著名的在线学习 Online Learning 算法之一) 进行深度优化, 提出带有时间衰减的自适应在线预测算法 TDAP。其 AUC 相对 FTRL 提升 5%, 并有效提升运营商 IPTV 视频推荐点击率。相关成果成功发表于 SIGKDD'16, 并已输出专利。
- 2014.10 – 2015.3 **基于图主键的知识图谱查询及清洗技术研究**。合作方: **樊文飞** (爱丁堡大学), **Luna Dong** (Google Research)  
我们率先定义了图主键的概念 (Keys for Graphs), 不同于传统关系型数据的主键, 图主键在 Knowledge Fusion 和 Social Network Reconciliation 等新兴领域有大量潜在的应用。我们用图主键解决了 Entity Matching 在知识图谱上的挑战。针对 MapReduce 和 Vertex-centric Asynchronous 两类分布式计算模型, 我们分别提出相对应的分布式并行算法, 并取得了很好的清洗及查询效果。相关成果成功发表于 PVLDB'15。

2011 – 2014 可信的大规模图数据查询优化技术研究。

保护用户数据的隐私和正确性是当前数据安全领域亟需解决的问题。针对大规模图数据的查询场景(如社交网络,知识图谱查询),我们提出了可信的子图查询(Subgraph Isomorphism Query)和可达性查询(Graph Reachability Query)技术,以保证用户图数据的安全。我们进一步设计了一系列优化技术以解决大规模图数据下的查询性能问题。相关成果均发表于计算机领域顶级杂志和会议上(TKDE, TSC, ICDE)。

2013 安全网盘 iSafebin 研发。合作方:香港 Privatera 团队

参与一款名为“iSafebin”的安全网盘系统的研发。主要负责设计并开发 RESTful 接口,具体包括账户安全管理,文件安全存储和安全传输等。相关技术包括 AES, RSA, Node.js 以及 MongoDB。

2010.8 – 2010.10 腾讯实习后台开发工程师。

参与互联网服务部门(Internet Service Department)的内部 Web 平台后台优化,主要包括日常运维平台和内部数据分享平台。相关技术包括 PHP, MySQL, JavaScript 和 Ajax 等。

## 在校经历(部分)

2013 – 2014 学生代表, 计算机系, 香港浸会大学

作为理学院研究生学生代表出席理学院议会。组织计算机系各类学生活动,如 Postgraduate Day 和 LunchBite 等,并作为博士生代表参加理学院暑期博士生夏令营。

2009 – 2010 学生会主席, 计算机科学与工程学院, 华南理工大学。

组织多场大型校园活动,如 ACM 大学生程序设计大赛,国际机器人大赛,广东省高校杯,IT 文化节等。与校内外学生组织和企业建立了长期的友好合作关系。在此期间,我的综合能力(包括统筹,规划,交流,团队意识,责任心等)得到了充分的磨练。这是我一生宝贵的财富。

## 荣誉奖项(部分)

2016 华为公司金牌员工(得奖比例约 1%)

2016 华为公司未来之星奖

2016 华为公司优秀发明人奖

2011 – 2015 香港浸会大学计算机系博士生全额奖学金

2015 亚坤博士生奖学金

2014 香港浸会大学计算机系海外学术访问奖学金

2013 香港浸会大学计算机系第 16 届研究生最佳论文奖

2012 – 2014 多次获得香港浸会大学计算机系最佳助教奖

2008 – 2011 多次获得华南理工大学一等奖学金,腾讯科技创新奖学金

2009 华南理工大学计算机创新创业大赛本科组第一名

2007 – 2011 多次获得华南理工大学优秀学生干部,优秀共青团员,三好学生

## 学术研究成果

TKDE, TSC(SIGKDD, PVLDB, ICDE)均为计算机科学领域顶级杂志(会议)

On going Z. Fan, Z. Li. “Taming Data Migration for Online Charging System”.

2017 X. Lin, J. Xu, H. Hu, Z. Fan. “Reducing Uncertainty of Probabilistic Top-k Ranking via Pair-wise Crowdsourcing”. *IEEE Transactions on Knowledge and Data Engineering (TKDE'17)*. Full Paper.

2016 Y. Tan, Z. Fan, G. Li, F. Wang, Z. Li, S. Liu, Q. Pan, Eric Xing, Qirong Ho. “Scalable Time-Decaying Adaptive Prediction Algorithm”. *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'16)*. Full Paper.

B. Choi, Z. Fan, Q. Chen, J. Xu and S. Bhowmick. “Structure-Preserving Subgraph Isomorphism Query Services (Abstract Extension)”. *IEEE International Conference on Data Engineering (ICDE'16)*. Poster Paper.

- 2015 W. Fan, **Z. Fan**, C. Tian, X. Dong. "Keys for Graphs". *International Conference on Very Large Data Bases (PVLDB'15)*. Full Paper.
- Z. Fan**, B. Choi, J. Xu and S. Bhowmick. "Asymmetric Structure-Preserving Subgraph Queries for Large Graphs". *IEEE International Conference on Data Engineering (ICDE'15)*. Full Paper.
- Z. Fan**, B. Choi, Q. Chen, J. Xu and S. Bhowmick. "Structure-Preserving Subgraph Isomorphism Query Services". *IEEE Transactions on Knowledge and Data Engineering (TKDE'15)*. Full Paper.
- X. Xie, **Z. Fan**, B. Choi, P. Yi, S. Bhowmick and S. Zhou. "PIGEON: Progress Indicator for Subgraph Queries". *IEEE International Conference on Data Engineering (ICDE'15)*. Demo Paper.
- 2014 Y. Peng, **Z. Fan**, B. Choi, J. Xu and S. Bhowmick. "Authenticated Subgraph Similarity Search in Outsourced Graph Databases". *IEEE Transactions on Knowledge and Data Engineering (TKDE'14)*. Full Paper.
- P. Yi, **Z. Fan**, and S. Yin. "Privacy-Preserving Reachability Query Services for Sparse Graph". *IEEE International Conference on Data Engineering (ICDE'14) GDM Workshop*.
- S. Yin, **Z. Fan**, P. Yi, B. Choi, J. Xu and S. Zhou. "Privacy-Preserving Reachability Query Services". *International Conference on Database Systems for Advanced Applications (DAS-FAA'14)*. Full Paper.
- 2013 **Z. Fan**, Y. Peng, B. Choi, J. Xu and S. Bhowmick. "Towards Efficient Authenticated Subgraph Query Service in Outsourced Graph Databases". *IEEE Transactions on Services Computing (TSC'13)*. Full Paper.

## 专利

- 2017 **Z. Fan** and F. Wang. "一种智能分析语义的方法及装置 (中国)".
- 2016 **Z. Fan** and Y. Tan. "一种自动选择模型优化策略的方法及装置 (中国)".
- 2015 B. Choi and **Z. Fan**. "Structure-Preserving Subgraph Queries (美国)".

## 学术审稿委员会

- 2017 IEEE Transactions on Knowledge and Data Engineering (TKDE)  
软件学报, 核心期刊

## 学术演讲

- 2016 **Z. Fan**. "Scalable Time Decaying Adaptive Prediction Algorithm". SIGKDD, 旧金山, 美国.
- 2015 **Z. Fan**. "Structure-Preserving Subgraph Query Services". ICDE, 首尔, 韩国.
- 2014 **Z. Fan**. "Querying a Graph Database You can Trust". 华傲数据, 深圳, 中国
- 2014 **Z. Fan**. "Privacy-Preserving Graph Reachability Query Services". ICDE, 芝加哥, 美国
- 2013 **Z. Fan**. "Querying a Graph Database You can Trust". 京都大学, 京都, 日本

## 专业技能

- 语言 英语, 广东话, 普通话
- 编程语言 C/C++, Python, Scala, SQL
- 编程框架 Spark, TensorFlow, MxNet, Hadoop, MySQL