

Dejun Teng

Phone: +86 13011777619 Email: teng@sdu.edu.cn

School of Computer Science and Technology, Shandong University, Qingdao, China

Current Titles and Affiliations

Assistant Professor, School of Computer Science and Technology, Shandong University

Research Focus

Computer Systems, Big Data, Spatial Data Management, GIS

Education and Training

- **Stony Brook University** 08/2019 – 12/2021
 - Ph.D. in Computer Science (graduating) GPA: 4.0/4.0
- **The Ohio State University** 08/2014 – 05/2018
 - Graduate researcher in Computer Science & Engineering GPA: 3.9/4.0
- **Emory University** 08/2012 – 05/2014
 - M.S. in Computer Science GPA: 3.7/4.0
- **Xi'an Jiaotong University** 08/2008 – 07/2012
 - B.E. in Software Engineering GPA: 84/100

Publications

- [EDBT22] **Dejun Teng**, Furqan Baig, Vo Hoang, Yanhui Liang, Jun Kong and Fusheng Wang: 3DPro: Querying Complex Three-Dimensional Data with Progressive Compression and Refinement. 25th International Conference on Extending Database Technology (EDBT 2022). pp. 104-117, March 29-April 1, 2022. Edinburgh, UK. (CCF B)
- [TSAS22] **Dejun Teng**, Yanhui Liang, Hoang Vo, Jun Kong, and Fusheng Wang. "Efficient 3D Spatial Queries for Complex Objects." *ACM Transactions on Spatial Algorithms and Systems (TSAS)* 8, no. 2 (2022): 1-26.
- [SIGSPATIAL21] **Dejun Teng**, Akshay Nehe, Prajeeth Emanuel, Furqan Baig, Jun Kong, and Fusheng Wang. "GPU-based real-time contact tracing at scale." 29th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL 2021), ACM, pp. 1–10, (24% acceptance rate)
- [MDM21] **Dejun Teng**, Furqan Baig, Qiheng Sun, Jun Kong, and Fusheng Wang. "IDEAL: a Vector-Raster Hybrid Model for Efficient Spatial Queries over Complex Polygons." *IEEE 22nd International Conference on Mobile Data Management (MDM 2021)*. IEEE, pp. 99–108, (CCF C, 26% acceptance rate)
- [DAPD19A] **Dejun Teng**, Jun Kong, and Fusheng Wang. "Scalable and flexible management of medical image big data." *Distributed and Parallel Databases* 37, no. 2 (2019): 235-250, (CCF C)
- [TOS18] **Dejun Teng**, Lei Guo, Rubao Lee, Feng Chen, Yanfeng Zhang, Siyuan Ma, and Xiaodong Zhang. "A Low-cost Disk Solution Enabling LSM-tree to Achieve High Performance for Mixed Read/Write Workloads." *ACM Transactions on Storage (TOS)* 14, no. 2 (2018): 1-26, (CCF A, only 17 regular papers in the 2018 issue)
- [ICDCS17] **Dejun Teng**, Lei Guo, Rubao Lee, Feng Chen, Siyuan Ma, Yanfeng Zhang, and Xiaodong Zhang. "LSbM-tree: Re-enabling buffer caching in data management for mixed reads and writes." In *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, pp. 68-79. IEEE, 2017, (CCF B, 18.5% acceptance rate)
- [JPHS22] Kayley Abell-Hart, Sina Rashidianb, **Dejun Teng**, Richard N Rosenthal, and Fusheng Wang. "Where Opioid Overdose Patients Live Far From Treatment: Geospatial Analysis of Underserved Populations in New York State." In *JMIR Public Health and Surveillance (JPHS)*, JMIR, 2022, (accepted IF 3.5)
- [ICDE21] Baig Furqan, **Dejun Teng**, Jun Kong, and Fusheng Wang. "Spear: Dynamic spatio-temporal query processing over high velocity data stream." *IEEE 37th International Conference on Data Engineering (ICDE 2021)*: pp. 2279–2284, (CCF A)
- [LI2020] Roy Mousumi, Fusheng Wang, Hoang Vo, **Dejun Teng**, George Teodoro, Alton B. Farris, Eduardo Castillo-Lion, Miriam B. Vos, and Jun Kong. "Deep-learning-based accurate hepatic steatosis quantification for histological assessment of liver biopsies." *Laboratory Investigation* (2020), *Nature Publishing Group*: 1-17, (IF 5.8)
- [FBD20] Baig Furqan, Chao Gao, **Dejun Teng**, Jun Kong, and Fusheng Wang. "Accelerating Spatial Cross-Matching on CPU-GPU Hybrid Platform with CUDA and OpenACC." *Frontiers Big Data* 3 (2020): 14.
- [DAPD19B] Vo Hoang, Jun Kong, **Dejun Teng**, Yanhui Liang, Ablimit Aji, George Teodoro, and Fusheng Wang. "MaRelA: a cloud MapReduce based high performance whole slide image analysis framework." *Distributed and Parallel Databases* 37, no. 2 (2019): 251-272. (CCF C)
- [DMAH16] Vo Hoang, Jun Kong, **Dejun Teng**, Yanhui Liang, Ablimit Aji, George Teodoro, and Fusheng Wang. "Cloud-Based Whole Slide Image Analysis Using MapReduce." In *VLDB Workshop on Data Management and Analytics for Medicine and Healthcare*, pp. 62-77. Springer, Cham, 2016.

Research Experiences

- **3DPro: Querying Complex Three-dimensional Data with Progressive Compression and Refinement**
 - We proposed a novel progressive spatial refinement method on supporting efficient three-dimensional data management and implemented 3DPro, a 3D data management system that adapts progressive refinement for efficient spatial querying. Mentor: Prof. Fusheng Wang. [TSAS22, EDBT22]
- **GPU-based Real-time Contact Tracing at Scale**

- We propose the GLINT system that achieves tens of millions scale real-time contact tracing queries with GPU. Mentor: Prof. Fusheng Wang. [SIGSPATIAL21]
- **IDEAL: a Vector-Raster Hybrid Model for Efficient Spatial Queries over Complex Polygons**
 - We proposed IDEAL, a Hybrid Vector Raster Model, by rasterizing a polygon and preserving relationship information between the rasterization pixels and the polygon at the pixel level. IDEAL achieves efficient 2D spatial query over complex polygons. Mentor: Prof. Fusheng Wang. [MDM21, TKDE22(under review)]
- **LSbM-tree: Re-enabling buffer caching in data management for mixed reads and writes**
 - We introduced LSbM-tree as a new variant of LSM-tree which re-enables high-speed caching during intensive writes. We used an on-disk merge buffer to mitigate the cache invalidation issue caused by merges, while preserving the merits of the LSM-tree. Mentor: Prof. Xiaodong Zhang. [TOS18, ICDCS17]
- **Scalable Data Management for Big Medical Imaging Data**
 - We implemented two distributed DICOM (Digital Imaging and Communications in Medicine) image management systems, RLDB with XML query enabled (DB2), and NoSql DB (MongoDB). We then comprehensively compared their performances on importing and querying DICOM images and revealed the pros and cons of those two systems. Mentor: Prof. Fusheng Wang. [DAPD19A]
- **PAIS: Large Scale Spatial Data Management for Pathology Imaging**
 - PAIS (Pathology Analytical Imaging Standards) is a large-scale pathology image and spatial data management system. We implemented the backend spatial data storage with DB2 distributed spatial database and analyzed those spatial data with machine learning algorithms. We also developed a web portal for visualizing the images, spatial objects, and analytical results. Mentor: Prof. Fusheng Wang. [DAPD19B, DMAH16]

Work Experiences

- | | |
|---|-------------------|
| ● Algorithm Engineer, Alibaba Cloud, Beijing, China | 12/2021 – 03/2022 |
| ● Algorithm Engineer, Huawei Technologies Co., Ltd., Shen Zhen, China | 05/2018 – 06/2019 |
| ● Summer Intern, Alibaba Inc., Seattle, WA, USA | 05/2015 – 08/2015 |
| ● Summer Intern, Pitney Bowes Inc., Shelton, CT, USA | 05/2014 – 08/2014 |

Honors

- Siyuan Scholarship, scholarship for top 20% student, 2008 – 2011

Computer Literacy

- **Programming**
 - C/C++, Java, Scala, Python, Shell, JavaScript, LLVM IR, Valgrind IR, Assembly Language
- **Data management**
 - DB2, MySQL, MongoDB, LevelDB, RocksDB, MonetDB, Hadoop, Spark

Language Proficiency

- **English** (listening, speaking, reading, and writing)
 - TOEFL iBT: 101 (Oral 23)
 - GRE: Verbal 151(52%) Quantitative 170(96%) Analytical Writing 3.0(17%)
- **Chinese** (listening, speaking, reading, and writing)

Services

- **PC member**
 - VLDB DMAH 2021
- **Manuscript reviewers**
 - Springer DAPD; ACM TSAS; ACM TODS
- **External reviewer**
 - SIGSPATIAL 2019, 2020, 2021; SIGMOD 2021, 2022; eSCIENCE 2021; IEEE Big Data 2019