[76] Cordella LP, Foggia P, Sansone C. A (sub) graph isomorphism algorithm for matching large graphs. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2004,26(10):1367–1372.

[77] Shapiro LG, Haralick RM. Structural descriptions and inexact matching. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1981,(5):504–519.

[78] Tsai WH, Fu KS. Subgraph error-correcting isomorphisms for syntactic pattern recognition. IEEE Trans. on Systems, Man and Cybernetics, 1983,(1):48–62.

[79] Cordella LP, Foggia P, Sansone C, Vento M. Subgraph transformations for the inexact matching of attributed relational graphs. In: Jolion JM, ed. Proc. of the Graph Based Representations in Pattern Recognition. Vienna: Springer-Verlag, 1998. 43–52.

[80] Zou L, Chen L, Özsu MT. Distance-join: Pattern match query in a large graph database. Proc. of the VLDB Endowment, 2009, 2(1):886–897.

[81] Umeyama S. An eigendecomposition approach to weighted graph matching problems. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1988,10(5):695–703.

[82] Christmas WJ, Kittler J, Petrou M. Structural matching in computer vision using probabilistic relaxation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1995,17(8):749–764.

[83] Giugno R, Shasha D. Graphgrep: A fast and universal method for querying graphs. Object Recognition Supported by User Interaction for Service Robots, 2002,(8)2:112–115.

[84] Yan X, Yu PS, Han J. Graph indexing: A frequent structure-based approach. In: Valduriez P, ed. Proc. of the 2004 ACM SIGMOD Int’l Conf. on Management of Data. New York: ACM, 2004. 335–346.

[85] He H, Singh AK. Closure-tree: An index structure for graph queries. In: Weikum G, ed. Proc. of the 22nd Int’l Conf. on Data Engineering, ICDE 2006. Washington: IEEE Computer Society, 2006. 38.

[86] Cheng J, Ke Y, Ng W. FG-index: Towards verification-free query processing on graph databases. In: Ooi BC, ed. Proc. of the 2007 ACM SIGMOD Int’l Conf. on Management of Data. New York: ACM, 2007. 857–872.

[87] Zhao P, Yu JX, Yu PS. Graph indexing: Tree+delta>=graph. In: Klas W, ed. Proc. of the 33rd Int’l Conf. on Very Large Data Bases. Vienna: VLDB Endowment, 2007. 938–949.

[88] Zhang S, Hu M, Yang J. Treepi: A novel graph indexing method. In: Ioannidis Y, ed. Proc. of the 23rd IEEE Int’l Conf. on Data Engineering. Washington: IEEE, 2007. 966–975.

[89] Williams DW, Huan J, Wang W. Graph database indexing using structured graph decomposition. In: Damien K, ed. Proc. of the 23rd IEEE Int’l Conf. on Data Engineering. Washington: IEEE, 2007. 976–985.

[90] Zou L, Chen L, Yu JX. A novel spectral coding in a large graph database. In: Teubner J, ed. Proc. of the 11th Int’l Conf. on Extending Database Technology: Advances in Database Technology. New York: ACM, 2008. 181–192.

[91] Shang H, Zhang Y, Lin X. Taming verification hardness: An efficient algorithm for testing subgraph isomorphism. Proc. of the VLDB Endowment, 2008,1(1):364–375.

[92] He H, Singh AK. Graphs-at-a-time: Query language and access methods for graph databases. In: Shasha D, ed. Proc. of the 2008 ACM SIGMOD Int’l Conf. on Management of Data. New York: ACM, 2008. 405–418.

[93] Zhang S, Li S, Yang J. GADDI: Distance index based subgraph matching in biological networks. In: Kersten M, Novikov B, Teubner J, eds. Proc. of the 12th Int’l Conf. on Extending Database Technology: Advances in Database Technology. New York: ACM, 2009. 192–203.

[94] Zhao P, Han J. On graph query optimization in large networks. Proc. of the VLDB Endowment, 2010,3(1-2):340–351.

[95] Di Natale R, Ferro A, Giugno R. Sing: Subgraph search in non-homogeneous graphs. BMC Bioinformatics, 2010,11(1):96.

[96] Bonnici V, Ferro A, Giugno R, Pulvirenti A, Shasha D. Enhancing graph database indexing by suffix tree structure. In: Dijkstra TMH, Tsivtsivadze E, Marchiori E, Heskes T, eds. Proc. of the Pattern Recognition in Bioinformatics, PRIB 2010. Lecture Notes in Computer Science, Heidelberg: Springer-Verlag, 2010. 195–203.

[97] Yuan D, Mitra P. Lindex: A lattice-based index for graph databases. The VLDB Journal, 2013,22(2):229–252.

[98] Gouda K, Hassaan M. Compressed feature-based filtering and verification approach for subgraph search. In: Guerrini G, ed. Proc. of the 16th Int’l Conf. on Extending Database Technology. New York: ACM, 2013. 287–298.

[99] Zheng W, Zou L, Lian X. SQBC: An efficient subgraph matching method over large and dense graphs. Information Sciences, 2014,261:116–131.

[100] Fan W, Li J, Ma S. Graph pattern matching: From intractable to polynomial time. Proc. of the VLDB Endowment, 2010,3(1-2):264–275.

[101] Fan W, Wang X, Wu Y. Incremental graph pattern matching. ACM Trans. on Database Systems (TODS), 2013,38(3):18.

[102] Fan W. Graph pattern matching revised for social network analysis. In: Deutsch A, ed. Proc. of the 15th Int'l Conf. on Database Theory. New York: ACM, 2012. 8–21.

[103] Fan W, Wang X, Wu Y. Diversified top-*k* graph pattern matching. Proc. of the VLDB Endowment, 2013,6(13):1510–1521.

[104] Udrea O, Pugliese A, Subrahmanian VS. GRIN: A graph based RDF index. In: Holte RC, Howe A, eds. Proc. of the 22nd AAAI Conf. on Artificial Intelligence. Menlo Park: AAAI Press, 2007. 1465–1470.

[105] Bröcheler M, Pugliese A, Subrahmanian VS. DOGMA: A disk-oriented graph matching algorithm for RDF databases. In: Alani H, Kagal L, Fokoue A, Groth P, Biemann C, Parreira JX, Aroyo L, Noy N, Welty C, Janowicz K, eds. Proc. of the Int'l Semantic Web Conf. Heidelberg: Springer-Verlag, 2009. 97–113.

[106] Atre M, Chaoji V, Zaki MJ. Matrix Bit loaded: A scalable lightweight join query processor for RDF data. In: Rappa M, ed. Proc. of the 19th Int'l Conf. on World Wide Web. New York: ACM, 2010. 41–50.

[107] Yuan P, Liu P, Wu B, *et al*. TripleBit: A fast and compact system for large scale RDF data. Proc. of the VLDB Endowment, 2013, 6(7):517–528.

[108] Aluç G, Ozsu MT, Daudjee K. Chameleon-DB: A workload-aware robust RDF data management system. In: Rappa M, ed. Proc. of the 19th Int'l Conf. on World Wide Web. New York: ACM, 2010. 41–50.

[109] Yu JXJ, Cheng J. Graph reachability queries: A survey. In: Aggarwal C, Wang H, eds. Advances in Database Systems-Managing and Mining Graph Data. Boston: Springer-Verlag, 2010. 181–215.

[110] Mendelzon AO, Wood PT. Finding regular simple paths in graph databases. SIAM Journal on Computing, 1995,24(6):1235–1258.

[111] Florescu D, Levy AY, Suciu D. Query containment for conjunctive queries with regular expressions. In: Mendelson A, Paredaens J, eds. Proc. of the 17th ACM SIGACT-SIGMOD-SIGART Symp. on Principles of Database Systems. New York: ACM, 1998. 139–148.

[112] Calvanese D, De Giacomo G, Lenzerini M, Vardi MY. Reasoning on regular path queries. ACM SIGMOD Record, 2003, 83–92.

[113] Barceló P, Libkin L, Lin AW. Expressive languages for path queries over graph-structured data. ACM Trans. on Database Systems (TODS), 2012,37(4):31.

[114] Pérez J, Arenas M, Gutierrez C. nSPARQL: A navigational language for RDF. In: Sheth A, eds. Lecture Notes in Computer Science. Heidelberg: Springer-Verlag, 2008. 66–81.

[115] Libkin L, Vrgoĉ D. Regular path queries on graphs with data. In: Deutsch A, ed. Proc. of the 15th Int'l Conf. on Database Theory. New York, 2012. 74–85.

[116] Koschmieder A, Leser U. Regular path queries on large graphs. In: Ailamaki A, Bowers S, eds. Lecture Notes in Computer Science. Heidelberg: Springer-Verlag, 2012. 177–194.

[117] Fan WF, Li JZ, Ma S, Tang N, Wu YH. Adding regular expressions to graph reachability and pattern queries. In: Proc. of the 27th IEEE Int'l Conf. on Data Engineering. Washington: IEEE Computer Society, 2011. 39–50.

[118] Gubichev A, Bedathur SJ, Seufert S. Sparqling kleene: Fast property paths in RDF-3X. In: Boncz P, ed. Proc. of the 1st Int'l Workshop on Graph Data Management Experiences and Systems. New York: ACM, 2013. 14–21.

[119] Dey S, Cuevas-Vicenttín V, Köhler S. On implementing provenance-aware regular path queries with relational query engines. In: Guerrini G, ed. Proc. of the Joint EDBT/ICDT 2013 Workshops. New York: ACM, 2013. 214–223.

[120] Wang X, Ling J, Wang JH, Wang KW, Feng ZY. Answering provenance-aware regular path queries on RDF graphs using an automata-based algorithm. In: Chung CW, Border A, Shim K, eds. Proc. of the 23rd Int'l Conf. on World Wide Web. New York: ACM, 2014. 395–396.

[121] Aggarwal CC, Wang H. Managing and Mining Graph Data (Advances in Database Systems). New York: Springer Science-Business Media, 2010. 1–620.

[122] Plantikow S. Summary chart of cypher, PGQL, and G-core. Neo4j. 2018. https://gql.today/wp-content/uploads/2018/05/ytz-030r1-Summary-Chart-of-Cypher-PGQL-GCore-1.pdf

[123] Wikipedia. The free encyclopedia. LargeTripleStores. 2018. https://www.w3.org/wiki/LargeTripleStores

[124] Eclipse RDF4J. RDF4J. 2018. http://rdf4j.org/

[125] OpenLink Software. OpenLink Virtuoso. 2018. https://virtuoso.openlinksw.com/
[126] Franz Inc. AllegroGraph. 2018. https://franz.com/agraph/allegrograph/
[127] Ontotext. GraphDB. 2018. http://graphdb.ontotext.com/
[128] Blazegraph by Systap, LLC. Blazegraph. 2018. https://www.blazegraph.com/
[129] Stardog Union. Stardog—The knowledge graph platform for the enterprise. 2018. http://www.stardog.com/
[130] JanusGraph Authors. JanusGraph—Distributed graph database. 2018. http://janusgraph.org/
[131] Spmallette. Titan—Distributed graph database. 2018. http://titan.thinkaurelius.com/
[132] Callidus Software Inc. OrientDB-multi-model database. 2018. http://orientdb.com/
[133] CayleyGraph. Cayley. 2018. https://cayley.io/
[134] Amazon Web Services, Inc. Amazon Neptune—Fast, reliable graph database build for cloud. 2018. https://aws.amazon.com/neptune/
[135] Star ArangoDB. ArangoDB—Native multimodel database. 2018. https://arangodb.com/
[136] Microsoft Azure. Microsoft Azure Cosmos DB. 2018. https://docs.microsoft.com/en-us/azure/cosmos-db/introduction
[137] DataStax. DataStax enterprise graph. 2018. https://www.datastax.com/products/datastax-enterprise-graph
[138] Sparsity Technologies. Sparksee—Scalable high-performance graph database. 2018. http://www.sparsity-technologies.com/#sparksee
[139] TigerGraph. TigerGraph—The first native parallel graph. 2018. https://www.tigergraph.com/
[140] Stanford University. Stanford large network dataset collection. 2018. http://snap.stanford.edu/data/
[141] Insight Centre for Data Analytics. The linked open data cloud. 2018. https://lod-cloud.net/
[142] University of Mannheim. DBpedia. 2018. http://wiki.dbpedia.org/About
[143] UniProt Consortium. UniProt. 2018. http://www.uniprot.org/
[144] AKSW. LinedGeoData. 2018. http://linkedgeodata.org/About
[145] Ghemawat S, Gobioff H, Leung ST. The Google file system. In: Scott ML, Peterson L, eds. Proc. of the 19th ACM Symp. on Operating Systems Principles. New York: ACM, 2003. 29–43.
[146] Chang F, Dean J, Ghemawat S. Bigtable: A distributed storage system for structured data. ACM Trans. on Computer Systems (TOCS), 2008,26(2):4.
[147] Dean J, Ghemawat S. MapReduce: Simplified data processing on large clusters. Communications of the ACM, 2008,51(1): 107–113.
[148] Malewicz G, Austern MH, Bik AJC. Pregel: A system for large-scale graph processing. In: Elmagarmid A, ed. Proc. of the 2010 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM, 2010. 135–146.
[149] Low Y, Bickson D, Gonzalez J. Distributed GraphLab: A framework for machine learning and data mining in the cloud. Proc. of the VLDB Endowment, 2012,5(8):716–727.
[150] Harth A, Umbrich J, Hogan A, Decker S. YARS2: A federated repository for querying graph structured data from the Web. In: Aberer K, eds. Lecture Notes in Computer Science. Heidelberg: Springer-Verlag, 2007. 211–224.
[151] Choi H, Son J, Cho YH. SPIDER: A system for scalable, parallel/distributed evaluation of large-scale RDF data. In: Cheung D, ed. Proc. of the 18th ACM Conf. on Information and Knowledge Management. New York: ACM, 2009. 2087–2088.
[152] Rohloff K, Schantz RE. High-performance, massively scalable distributed systems using the MapReduce software framework: The SHARD triple-store. In: Tilevich E, ed. Proc. of the Programming Support Innovations for Emerging Distributed Applications. New York: ACM, 2010. 4.
[153] Ladwig G, Harth A. CumulusRDF: Linked data management on nested key-value stores. In: Grau BC, ed. Proc. of the 7th Int'l Workshop on Scalable Semantic Web Knowledge Base Systems. Bonn, 2011. 30.
[154] Shao B, Wang H, Li Y. Trinity: A distributed graph engine on a memory cloud. In: Ross K, ed. Proc. of the 2013 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM, 2013. 505–516.
[155] Zeng K, Yang J, Wang H. A distributed graph engine for Web scale RDF data. PVLDB, 2013,6(4):265–276.
[156] Huang J, Abadi DJ, Ren K. Scalable SPARQL querying of large RDF graphs. Proc. of the VLDB Endowment, 2011,4(11): 1123–1134.

[157] Lee K, Liu L. Scaling queries over big RDF graphs with semantic hash partitioning. Proc. of the VLDB Endowment, 2013,6(14): 1894–1905.
[158] Zhang X, Chen L, Tong Y. EAGRE: Towards scalable I/O efficient SPARQL query evaluation on the cloud. In: Proc. of the ICDE 2013 IEEE Int'l Conf. on Data Engineering. Washington: IEEE, 2013. 565–576.
[159] Gurajada S, Seufert S, Miliaraki I. TriAD: A distributed shared-nothing RDF engine based on asynchronous message passing. In: Dyreson C, ed. Proc. of the 2014 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM, 2014. 289–300.
[160] Hammoud M, Rabbou DA, Nouri R. DREAM: Distributed RDF engine with adaptive query planner and minimal communication. Proc. of the VLDB Endowment, 2015,8(6).
[161] Jones ND. An introduction to partial evaluation. ACM Computing Surveys (CSUR), 1996,28(3):480–503.
[162] Fan W, Wang X, Wu Y. Performance guarantees for distributed reachability queries. Proc. of the VLDB Endowment, 2012,5(11): 1304–1316.
[163] Ma S, Cao Y, Huai J. Distributed graph pattern matching. In: Mille A, ed. Proc. of the 21st Int'l Conf. on World Wide Web. New York: ACM, 2012. 949–958.
[164] Fan W, Wang X, Wu Y. Distributed graph simulation: Impossibility and possibility. Proc. of the VLDB Endowment, 2014,7(12).
[165] Peng P, Zou L, Özsu MT, Chen L, Zhao D. Processing SPARQL queries over distributed RDF graphs. The VLDB Journal—The Int'l Journal on Very Large Data Bases, 2016,25(2):243–268.
[166] Wang X, Wang JH, Zhang XW. Efficient distributed regular path queries on RDF graphs using partial evaluation. In: Mukhopadhyay S, ed. Proc. of the 25th ACM Int'l Conf. on Information and Knowledge Management. New York: ACM, 2016. 1933–1936.
[167] Papailiou N, Tsoumakos D, Konstantinou I. H2RDF+: An efficient data management system for big RDF graphs. In: Dyreson C, ed. Proc. of the 2014 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM, 2014. 909–912.
[168] Schätzle A, Przyjaciel-Zablocki M, Neu A, Lausen G. Sempala: Interactive SPARQL query processing on hadoop. In: Mika P, *et al*., eds. Lecture Notes in Computer Science. Cham: Springer-Verlag, 2014. 164–179.
[169] Lai L, Qin L, Lin X. Scalable subgraph enumeration in mapreduce. Proc. of the VLDB Endowment, 2015,8(10):974–985.
[170] Lai L, Qin L, Lin X. Scalable distributed subgraph enumeration. Proc. of the VLDB Endowment, 2016,10(3):217–228.
[171] Bi F, Chang LJ, Lin XM, Qin L, Zhang WJ. Efficient subgraph matching by postponing Cartesian products. In: Proc. of the SIGMOD Conf. 2016. 1199–1214.
[172] Schätzle A, Przyjaciel-Zablocki M, Skilevic S. S2RDF: RDF querying with SPARQL on spark. Proc. of the VLDB Endowment, 2016,9(10):804–815.
[173] Wang X, Xu Q, Chai LL, Yang YJ, Chai YP. Efficient distributed query processing on large scale RDF graph data. Ruan Jian Xue Bao/Journal of Software, 2019,30(3):498–514 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/5696.htm [doi: 10.13328/j.cnki.jos.005696]
[174] He L, Shao B, Li Y. Stylus: A strongly-typed store for serving massive RDF data. Proc. of the VLDB Endowment, 2017,11(2): 203–216.
[175] Peng P, Zou L, Chen L. Adaptive distributed RDF graph fragmentation and allocation based on query workload. IEEE Trans. on Knowledge and Data Engineering, 2019,31(4):670–685.
[176] Xin YQ, Wang X, Jin D, Wang S. Distributed efficient provenance-aware regular path queries on large RDF graphs. In: Liu CF, Zou L, Li JX, eds. Proc. of the Database Systems for Advanced Applications. London: Springer-Verlag, 2018. 766–782.
[177] Punnoose R, Crainiceanu A, Rapp D. SPARQL in the cloud using Rya. Information Systems, 2015,48:181–195.
[178] S1CK. Cypher for Apache Spark. 2019. https://github.com/opencypher/cypher-for-apache-spark
[179] LDBC. LDBC Task force: Property graphs data model. 2019. http://www.ldbcouncil.org
[180] Guo Y, Pan Z, Heflin J. LUBM: A benchmark for OWL knowledge base systems. Web Semantics: Science, Services and Agents on the World Wide Web, 2005,3(2-3):158–182.
[181] Schmidt M, Hornung T, Lausen G. SP2Bench: A SPARQL performance benchmark. In: Proc. of the 25th IEEE Int'l Conf. on Data Engineering. Washington: IEEE, 2009. 222–233.
[182] Bizer C, Schultz A. The Berlin sparql benchmark. Int'l Journal on Semantic Web and Information Systems (IJSWIS), 2009,5(2): 1–24.

[183] Morsey M, Lehmann J, Auer S, Ngomo ACN. DBpedia SPARQL benchmark—Performance assessment with real queries on real data. In: Aroyo L, *et al*., eds. Lecture Notes in Computer Science. Heidelberg: Springer-Verlag, 2011. 454–469.

[184] Aluç G, Hartig O, Özsu MT, Daudjee K. Diversified stress testing of RDF data management systems. In: Mika P, *et al*., eds. Lecture Notes in Computer Science. Cham: Springer-Verlag, 2014. 197–212.

[185] Armstrong TG, Ponnekanti V, Borthakur D. LinkBench: A database benchmark based on the Facebook social graph. In: Ross K, ed. Proc. of the 2013 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM, 2013. 1185–1196.

[186] LDBC. Social network benchmark (SNB). 2019. http://ldbcouncil.org/benchmarks/snb

[187] LDBC. Semantic publishing benchmark (SPB) v2.0. 2019. http://ldbcouncil.org/benchmarks/spb

[188] The GQL Manifesto team. The GQL Manifesto. 2018. https://gql.today/

[189] Speicher S, Arwe J, Malhotra A. Linked data platform 1.0. W3C Recommendation, 2015,26.

附中文参考文献:

[16] 邹磊,彭鹏.分布式RDF数据管理综述.计算机研究与发展,2017,54(6):1213–1224.

[17] 于戈,谷峪,鲍玉斌.云计算环境下的大规模图数据处理技术.计算机学报,2011,34(10):1753–1767.

[18] 于静,刘燕兵,张宇.大规模图数据匹配技术综述.计算机研究与发展,2015,52(2):391–409.

[19] 王童童,荣垂田,卢卫.分布式图处理系统技术综述.软件学报,2018,29(3):569–586. http://www.jos.org.cn/1000-9825/5450.htm [doi: 10.13328/j.cnki.jos.005450]

[20] 官赛萍,靳小龙,贾岩涛.面向知识图谱的知识推理研究进展.软件学报,2018,29(10):74–102. http://www.jos.org.cn/1000-9825/5551.htm [doi: 10.13328/j.cnki.jos.005551]

[21] 许嘉,张千桢,赵翔.动态图模式匹配技术综述.软件学报,2018,29(3):663–688. http://www.jos.org.cn/1000-9825/5444.htm [doi: 10.13328/j.cnki.jos.005444]

[173] 王鑫,徐强,柴乐乐,杨雅君,柴云鹏.大规模RDF图数据上高效率分布式查询处理.软件学报,2019,30(3):498–514. http://www.jos.org.cn/1000-9825/5696.htm [doi: 10.13328/j.cnki.jos.005696]

王鑫(1981－),男,天津人,博士,副教授,CCF高级会员,主要研究领域为知识图谱数据管理,图数据库,大规模知识处理.

彭鹏(1987－),男,博士,助理教授,CCF专业会员,主要研究领域为分布式知识图谱管理.

邹磊(1981－),男,博士,教授,博士生导师,CCF高级会员,主要研究领域为图数据库,知识图谱,大数据系统.

冯志勇(1965－),男,博士,教授,博士生导师,CCF杰出会员,主要研究领域为知识工程,服务计算.

王朝坤(1976－),男,博士,副教授,博士生导师,CCF专业会员,主要研究领域为图和社交数据管理,大数据系统.