

P.O.Box 8718, Beijing 100080, China	Journal of Software, Dec. 2005,16(12):2063-2079
E-mail: jos@iscas.ac.cn	ISSN 1000-9825, CODEN RUXUEW, CN 11-2560/TP
http://www.jos.org.cn	Copyright © 2005 by The Editorial Department of <i>Journal of Software</i>

XML数据索引技术

孔令波, 唐世渭, 杨冬青, 王腾蛟, 高 军

[Full-Text PDF](#) [Submission](#) [Back](#)

孔令波¹, 唐世渭^{1,2}, 杨冬青¹, 王腾蛟¹, 高 军¹

¹(北京大学 计算机科学技术系,北京 100871)

²(北京大学 视觉与听觉信息处理国家重点实验室,北京 100871)

作者简介: 孔令波(1974—),男,山东日照人,博士生,主要研究领域为关系数据库实现技术,XML数据处理技术,数据挖掘;

唐世渭(1939—),男,教授,博士生导师,CCF高级会员,主要研究领域为数据库,半结构化数据,Web数据集成,数据挖掘;杨冬青(1945—),女,教授,博士生导师,CCF高级会员,主要研究领域为数据库,数据仓库,Web数据集成,移动数据挖掘;王腾蛟(1973—),男,博士,副教授,CCF高级会员,主要研究领域为数据库,数据仓库,Web数据集成,数据挖掘;高军(1975—),男,博士,副教授,主要研究领域为数据库,数据仓库,半结构化数据,Web数据集成,移动数据挖掘.

联系人: 孔令波 Phn: +86-10-62755440, E-mail: lbkong@db.pku.edu.cn, http://www.pku.edu.cn

Received 2004-12-07; Accepted 2005-08-24

Abstract

XML has become the de facto standards for data representation and exchange on Web applications, such as digital library, Web service, and electronic business, etc. Indexing technique is still significant for efficient XML data processing. This paper discusses the actualities of the recent researches on XML indexing. It classifies the techniques into two categories, node-record-style index with three subcategories, and structural-summary-style index. It analyzes the virtue and deficiency of the related schemes based on the considerations for query processing efficiency and data modification supporting. And hereby it proposes three issues for future XML indexing researches, including internal structure retrieval, multi-dimensional processing on node paths, efficient modification-validating support and the index amalgamation for satisfying both querying and IR on XML data.

Kong LB, Tang SW, Yang DQ, Wang TJ, Gao J. XML indices. *Journal of Software*, 2005,16(12):2063-2079.

DOI: 10.1360/jos162063

<http://www.jos.org.cn/1000-9825/16/2063.htm>

摘要

对XML数据建立有效的索引,是左右XML数据处理性能的重要因素.深入地讨论了目前XML索引技术的研究现状,将XML索引技术分为两大类:节点记录类索引(本身还可以分为3个小的类型)和结构摘要类索引.根据XML数据查询处理效率以及XML数据修改对XML索引的要求,讨论了相关XML索引方法的优点和不足,并归结出XML索引后续研究的3个方向:XML结构信息的获取,路径信息的多维处理,数据修改合法性的有效支持,以及涉及能够同时有效满足XML查询和信息获取的索引.

基金项目: Supported by the National High-Tech Research and Development Plan of China under Grant Nos.2002AA4Z3440, 2005AA4Z3070 (国家高技术研究发展计划(863))

References:

[1] Meng XF, Zhou LX, Wang S. State of the art and trends in database research. *Journal of Software*, 2004,15(12):1822-1836 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/12/1822.htm>

[2] Barashev D, Novikov B. Indexing XML to support path expressions. In: Manolopoulos Y, Návrat P, eds. Proc. of the 6th East European Conf. on Advances in Databases and Information Systems (ADBIS). Bratislava: Springer-Verlag, 2002. 1-10.

- [3] Li QZ, Moon B. Indexing and querying XML data for regular path expressions. In: Apers PMG, Atzeni P, Ceri S, Paraboschi S, Ramamohanarao K, Snodgrass RT, eds. Proc. of the 27th Int'l Conf. on Very Large Data Bases (VLDB). Roma: Morgan Kaufmann Publishers, 2001. 361-370.
- [4] Cooper BF, Sample1 N, Franklin MJ, Hjaltason GR, Shadmon M. A fast index for semistructured data. In: Apers PMG, Atzeni P, Ceri S, Paraboschi S, Ramamohanarao K, Snodgrass RT, eds. Proc. of the 27th Int'l Conf. on Very Large Data Bases (VLDB). Roma: Morgan Kaufmann Publishers, 2001. 341-350.
- [5] Zhang C. Relational databases for XML indexing [Ph.D. Thesis]. Wisconsin: University of Wisconsin-Madison, 2002.
- [6] Bruno N, Koudas N, Srivastava D. Holistic twig joins: Optimal XML pattern matching. In: Franklin MJ, Moon B, Ailamaki A, eds. Proc. of the 2002 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Madison: ACM Press, 2002. 310-321.
- [7] Amer-Yahia S, Cho S, Lakshmanan LVS, Srivastava D. Minimization of tree pattern queries. In: Aref WG, ed. Proc. of the 2001 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Santa Barbara: ACM Press, 2001. 497-508.
- [8] Jiang HF, Wang W, Lu HJ, Yu JX. Holistic Twig joins on Indexed XML documents. In: Freytag JC, Lockemann PC, Abiteboul S, Carey MJ, Selinger PG, Heuer A, eds. Proc. of the 29th Int'l Conf. on Very Large Data Bases (VLDB). Berlin: Morgan Kaufmann Publishers, 2003. 273-284.
- [9] Chen ZY, Jagadish HV, Korn F, Koudas N. Counting twig matches in a tree. In: Young DC, ed. Proc. of the 17th Int'l Conf. on Data Engineering (ICDE). Heidelberg: IEEE Computer Society, 2001. 595-604.
- [10] Lee DW, Srivastava D. Counting relaxed twig matches in a tree. In: Lee YJ, Li JZ, Whang KY, Lee DH, eds. Proc. of the 9th Int'l Conf. on Database Systems for Advances Applications (DASFAA). LNCS 2973, Springer-Verlag, 2004. 88-99.
- [11] Jagadish HV, Al-Khalifa S. TIMBER: A native XML database. The VLDB Journal, 2002,11(4):274-291.
- [12] Fiebig T, Helmer S, Kanne CC. Anatomy of a native XML base management system. The VLDB Journal, 2002,11(4):292-314.
- [13] Miklau G, Siciu D. Containment and equivalence for a fragment of XPath. Journal of the ACM, 2004,51(1):2-45.
- [14] Tatarinov I, Ives ZG, Halevy AY, Weld DS. Updating XML. In: Aref WG, ed. Proc. of the 2001 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Santa Barbara: ACM Press, 2001. 413-424.
- [15] Papakonstantinou Y, Vianu V. Incremental validation of XML documents. In: Calvanese D, Lenzerini M, Motwani R, eds. Proc. of the 9th Int'l Conf. (ICDT). LNCS 2572, Siena: Springer-Verlag, 2003. 47-63.
- [16] Barbosa D, Mendelzon AO, Libkin L, Mignet L. Efficient incremental validation of XML documents. In: Tittsworth F, ed. Proc. of the 20th Int'l Conf. on Database Engineering (ICDE). Boston: IEEE Computer Society, 2004. 671-682.
- [17] Abrao MA, Bouchou B, Ferrari MH. Incremental constraint checking for XML documents. In: Bellahsene Z, Milo T, Rys M, Suci D, Unland R, eds. Proc. of the 2nd Int'l XML Database Symp. (Xsym), Database and XML Technologies. LNCS 3186, Springer-Verlag, 2004. 112-127.
- [18] Kaushik R, Bohannon P, Naughton JF, Shenoy P. Updates for structure indexes. In: Bressan S, Chaudhri AB, Lee ML, Yu JX, Lacroix Z, eds. Proc. of the 28th Int'l Conf. on Very Large Data Bases (VLDB). LNCS 2590, Hong Kong: Morgan Kaufmann Publishers, 2002. 239-250.
- [19] Yi K, He H, Stanoi I, Yang J. Incremental maintenance of XML structural indexes. In: Weikum G, K?nig AC, De?loch S, eds. Proc. of the ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Paris: ACM Press, 2004. 491-502.
- [20] Deutsch A, Fernandez M, Suci D. Storing semistructured data with STORED. In: Haas LM, Tiwary A, eds. Proc. of the ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD'99). Philadelphia: ACM Press, 1999. 431-442.
- [21] Shanmugasundaram J, Tufte K, He G. Relational databases for querying XML documents: Limitations and opportunities. In: Atkinson MP, Orłowska ME, Valduriez P, Zdonik SB, Brodie ML, eds. Proc. of the 25th Int'l Conf. on Very Large Data Bases (VLDB). Edinburgh: Morgan Kaufmann Publishers, 1999. 302-314.

- [22] Kanne CC, Moerkotte G. Efficient storage of XML data. In: Young DC, ed. Proc. of the 16th Int'l Conf. on Data Engineering (ICDE). San Diego: IEEE Computer Society, 2000. 198.
- [23] Klettke M, Meyer H. XML and object-relational database systems-enhancing structural mappings based on statistics. In: Suciu D, Vossen G, eds. Proc. of the Int'l Workshop on the Web and Databases (WebDB). LNCS 1997, Dallas: Springer-Verlag, 2000. 151-170.
- [24] Schmidt AR, Kersten ML, Windhouwer M, Wass F. Efficient relational storage and retrieval of XML documents. In: Suciu D, Vossen G, eds. Proc. of the Int'l Workshop on the Web and Databases (WWW). LNCS 1997, Amsterdam: Springer-Verlag, 2000. 137-150.
- [25] Bohannon P, Freire J, Roy P, Siméon J. From XML schema to relations: A cost-based approach to XML storage. In: Agrawal R, Dittrich K, Ngu AHH, eds. Proc. of the 18th Int'l Conf. on Data Engineering (ICDE). San Jose: IEEE Computer Society, 2002. 64-92.
- [26] Hwang JH, Nguyen VT, Ryu KH. A new indexing structure to speed up processing XPath queries. In: Zhou LZ, Ooi BC, Meng XF, eds. Proc. of the 10th Int'l Conf. on Database Systems for Advanced Applications (DASFAA). LNCS 3453, Trondheim: Springer-Verlag, 2005. 900-906.
- [27] Grust T. Accelerating XPath location steps. In: Franklin MJ, Moon B, Ailamaki A, eds. Proc. of the 2002 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Madison: ACM Press, 2002. 109-120.
- [28] DeHaan D, Toman D, Consens MP. A comprehensive XQuery to SQL translation using dynamic interval encoding. In: Halevy AY, Ives ZG, Doan A, eds. Proc. of the 2003 ACM SIGMOD Intl. Conf. on Management of Data (SIGMOD). San Diego: ACM Press, 2003. 623-634.
- [29] Grust T, Teubner J. Accelerating XPath evaluation in any RDBMS. ACM Trans. on Database Systems, 2004,29(1):91-131.
- [30] Amagasa T, Yoshikawa M, Uemura S. QRS: A robust numbering scheme for XML documents. In: Dayal U, Ramamritham K, Vijayaraman TM, eds. Proc. of the 19th Int'l Conf. on Data Engineering (ICDE). Bangalore: IEEE Computer Society, 2003. 705-707.
- [31] O'Neil P, O'Neil E, Pal S, Cseri I, Schaller G. ORDPATHs: Insert-Friendly XML node labels. In: Weikum G, K?nig AC, De?loch S, eds. Proc. of the ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Paris: ACM Press, 2004. 903-908.
- [32] Tatarinov I, Viglas SD. Storing and querying ordered XML using a relational database system. In: Franklin MJ, Moon B, Ailamaki A, eds. Proc. of the 2002 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Madison: ACM Press, 2002. 204-215.
- [33] Cohen E, Kaplan H, Milo T. Labeling dynamic XML trees. In: Popa L, ed. Proc. of the 21st ACM SIGACT-SIGMOD-SIGART Symp. on Principles of Database Systems (PODS). Madison: ACM Press, 2002. 271-281.
- [34] Han ZM, Fu NY. Efficiently coding and querying XML document. In: Bhalla S, ed. Proc. of the 4th Int'l Workshop on Databases in Networked Information Systems (DNIS). LNCS 3433, Springer-Verlag, 2005. 54-69.
- [35] Han ZM, Xi CT, Le JJ. Efficiently coding and indexing XML document. In: Zhou LZ, Ooi BC, Meng XF, eds. Proc. of the 10th Int'l Conf. on Database Systems for Advanced Applications (DASFAA). LNCS 3453, Beijing: Springer-Verlag, 2005. 138-150.
- [36] Wu XD, Lee ML, Hsu W. A prime number labeling scheme for dynamic ordered XML trees. In: Proc. of the 20th Int'l Conf. on Database Engineering (ICDE). Boston: IEEE Computer Society, 2004. 66-78.
- [37] Runapongsa K. Methods for efficient storage and indexing in XML databases [Ph.D. Thesis]. Michigan: University of Michigan, 2003.
- [38] Seo CY, Lee SW, Kim HJ. An efficient inverted index technique for XML documents using RDBMS. Information and Software Technology, 2003,45(1):11-22.
- [39] Florescu D, Kossmann D, Manolescu I. Integrating keyword search into XML query processing. <http://www9.org/w9cdrom/index.html>
- [40] Wu YQ, Patel JM, Jagadish HV. Estimating answer sizes for XML queries. In: Jensen CS, Jeffery KG, Pokorny J, Saltenis S, Bertino E, B?hm K, Jarke M, eds. Proc. of the 8th Int'l Conf. on Extending Database Technology (EDBT 2002). Prague: Springer-Verlag, 2002. 590-608.

[41] Florescu D, Kossman D. A performance evaluation of alternative mapping schemes for storing XML in a relational database. Technical Report, No. 3680, INRIA, France, 1999.

[42] Al-Khalifa S, Jagadish HV, Koudas N, Patel JM, Srivastava D, Wu Y. Structural joins: A primitive for efficient XML query pattern matching. In: Agrawal R, Dittrich K, Ngu AHH, eds. Proc. of the 18th Int'l Conf. on Data Engineering (ICDE). San Jose: IEEE Computer Society, 2002. 141-152

[43] Barbosa D, Freire J, Mendelzon AO. Designing information-preserving mapping schemes for XML. In: Böhm K, Jensen CS, Haas LM, Kersten ML, Larson P, Ooi BC. eds. Proc. of the 31st Int'l Conf. on Very Large Data Bases (VLDB). Trondheim: ACM Press, 2005. 109-120.

[44] Kriegel HP, PStke M, Seidl T. Managing intervals efficiently in object-relational databases. In: Abbadi AE, Brodie ML, Chakravarthy S, Dayal U, Kamel N, Schlageter G, Whang KY, eds. Proc. of the 26th Int'l Conf. on Very Large Data Bases (VLDB). Cairo: Morgan Kaufmann, 2000. 407-418.

[45] Jiang HF, Lu HJ, Wang W, Ooi BC. XR-Tree: Indexing XML data for efficient structural joins. In: Dayal U, Ramamritham K, Vijayaraman TM, eds. Proc. of the 19th Int'l Conf. on Data Engineering (ICDE). Bangalore: IEEE Computer Society, 2003. 253-264.

[46] Zhang C, Naughton J, DeWitt D, Luo Q, Lohman G. On supporting containment queries in relational database management systems. In: Aref WG, ed. Proceedings of the 2001 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Santa Barbara: ACM Press, 2001. 425-436.

[47] Chien SY, Vagena Z, Zhang D, TsoTRAS VJ, Zaniolo C. Efficient Structural Joins on Indexed XML Documents. In: Bressan S, Chaudhri AB, Lee ML, Yu JX, Lacroix Z, eds. Proc. of the 28th Int'l Conf. on Very Large Data Bases (VLDB). LNCS 2590, Hong Kong: Morgan Kaufmann, 2002. 263-274.

[48] Wang W, Jiang HF, Lu HJ, Yu JX. Containment join size estimation: Models and methods. In: Halevy AY, Ives ZG, Doan A, eds. Proc. of the 2003 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). San Diego: ACM Press, 2003. 145-156.

[49] Wang J, Meng XF, Wang S. Structural join of XML based on range partitioning. Journal of Software, 2004,15(5):720-729 (in Chinese with English). <http://www.jos.org.cn/1000-9825/15/720.htm>

[50] Lu JH, Ling TW, Chan CY, Chen T. From region encoding to extended dewey: On efficient processing of XML twig pattern matching. In: Böhm K, Jensen CS, Haas LM, Kersten ML, Larson P, Ooi BC. eds. Proc. of the 31st Int'l Conf. on Very Large Data Bases (VLDB). Trondheim: ACM Press, 2005. 193-204.

[51] Harding PJ, Li QZ, Moon B. XISS/R: XML indexing and storage system using RDBMS. In: Freytag JC, Lockemann PC, Abiteboul S, Carey MJ, Selinger PG, Heuer A, eds. Proc. of the 29th Int'l Conf. on Very Large Data Bases (VLDB). Berlin: Morgan Kaufmann Publishers, 2003. 1073-1076.

[52] Silberstein A, He H, Yi K, Yang J. BOXes: Efficient maintenance of order-based labeling for dynamic XML data. In: Stephanie Kawada, ed. Proc. of the 21st Int'l Conf. on Data Engineering (ICDE). Tokyo: IEEE Computer Society, 2005. 285-296.

[53] Kratky M, Pokorny J, Snasel V. Indexing XML data with UB-trees. In: Manolopoulos Y, Návrát P, eds. Proc. of the 6th East European Conf. on Databases and Information Systems (ADBIS). Bratislava: Springer-Verlag, 2002. 155-164.

[54] Rao P, Moon B. PRIX: Indexing and querying XML using präfer sequence. In: Titsworth F, ed. Proc. of the 20th Int'l Conf. on Database Engineering (ICDE). Boston: IEEE Computer Society, 2004. 288-300.

[55] Wang HX, Park S, Fan W, Yu PS. ViST: A dynamic index method for querying XML data by tree structure. In: Halevy AY, Ives ZG, Doan A, eds. Proc. of the 2003 ACM SIGMOD Intl. Conf. on Management of Data (SIGMOD). San Diego: ACM Press, 2003. 110-121.

[56] Bayer R. XML Databases: Modeling and multidimensional indexing. Invited talk at DEXA Conf. München, 2001. <http://link.springer.de/link/service/series/0558/bibs/2113/21130001.htm>

[57] Jiang HF, Lu HJ, Wang W, Yu JX. Path materialization revisited: An efficient storage model for XML data. In: Zhou XF, ed. Proc. of the 13th Australasian Database Conf. on Database Technologies 2002 (ADC). Melbourne: Australian Computer Society, 2002. 85-94.

[58] Wang HX, Meng XF. On the sequencing of tree structures for XML indexing. In: Stephanie Kawada, ed. Proc. of the 21st Int'l Conf. on Data Engineering (ICDE). Tokyo: IEEE Computer Society, 2005. 372-383.

[59] Yoshikawa M, Amagasa T. XRel: A path-based approach to storage and retrieval of XML documents using relational databases. ACM Trans. on Internet Technology, 2001,1(1):110-141.

[60] Chen Y, Davidson SB, Zheng YF. BLAS: An efficient XPath processing system. In: Weikum G, K?nig AC, De?loch S, eds. Proc. of the ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Paris: ACM Press, 2004. 47-58.

[61] Goldman R, Widom J. Dataguides: Enabling query formulation and optimization in semistructured databases. In: Jarke M, Carey MJ, Dittrich KR, Lochovsky FH, Loucopoulos P, Jausfeld MA, eds. Proc. of the 23rd Int'l Conf. on Very Large Data Bases (VLDB). Athens: Morgan Kaufmann, 1997. 436-445.

[62] Milo T, Suciu D. Index structures for path expressions. In: Beeri C, Buneman P, eds. Proc. of the 1999 Int'l Conf. on Database Theory (ICDT). LNCS 1540, Jerusalem: Springer-Verlag, 1999. 277-295.

[63] Kaushik R, Sheony P, Bohannon P, Gudes E. Exploiting local similarity for efficient indexing of paths in graph structured data. In: Agrawal R, Dittrich K, Ngu AHH, eds. Proc. of the 18th Int'l Conf. on Data Engineering (ICDE). San Jose: IEEE Computer Society, 2002. 129-140.

[64] Chung C, Min J, Shim K. APEX: An adaptive path index for XML data. In: Franklin MJ, Moon B, Ailamaki A, eds. Proc. of the 2002 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Madison: ACM Press, 2002. 121-132.

[65] Kaushik R, Bohannon P, Naughton JF, Korth HF. Covering indexes for branching path queries. In: Franklin MJ, Moon B, Ailamaki A, eds. Proc. of the 2002 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Madison: ACM Press, 2002. 133-144.

[66] Polyzotis N, Garofalakis M. Statistical synopses for graph-structured XML databases. In: Franklin MJ, Moon B, Ailamaki A, eds. Proc. of the 2002 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Madison: ACM Press, 2002. 358-369.

[67] Chen Q, Lim A, Ong KW. D(k)-index: An adaptive structural summary for graph-structured data. In: Halevy AY, Ives ZG, Doan A, eds. Proc. of the 2003 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). San Diego: ACM Press, 2003. 134-144.

[68] He H, Yang J. Multiresolution indexing of XML for frequent queries. In: Tittsworth F, ed. Proc. of the 20th Int'l Conf. on Data Engineering (ICDE). Boston: IEEE Computer Society, 2004. 683-694.

[69] Wang W, Wang HZ, Lu HJ, Jiang HF, Lin XM, Li JZ. Efficient processing of XML path queries using the disk-based F&B index. In: B?hm K, Jensen CS, Haas LM, Kersten ML, Larson P, Ooi BC. eds. Proc. of the 31st Int'l Conf. on Very Large Data Bases (VLDB). Trondheim: ACM Press, 2005. 145-156.

[70] Barg M, Wong RK. Structural proximity searching for large collections semi-structured data. In: Paques H, Liu L, Grossman D, eds. Proc. of the ACM Conf. on Information and Knowledge Management (CIKM 2001). Atlanta: ACM Press, 2001. 175-182.

[71] Cohen S, Mamou J, Kanza Y, Sagiv Y. Xsearch: A semantic search engine for xml. In: Freytag JC, Lockemann PC, Abiteboul S, Carey MJ, Selinger PG, Heuer A, eds. Proc. of the 29th Int'l Conf. on Very Large Data Bases (VLDB). Berlin: Morgan Kaufmann Publishers, 2003. 45-56.

[72] Curtmola E, Amer-Yahia S, Brown P, Fern?ndez M. GalaTex: A conformant implementation of the XQuery FullText language. In: Florescu D, Pirahesh H, eds. Proc. of the 2nd Int'l Workshop on XQuery Implementation, Experience, and Perspectives (XIME-P). Baltimore: ACM Press, 2005. 1024-1025.

[73] Amer-Yahia S, Botev C, Shanmugasundaram J. TeXQuery: A FullText search extension to XQuery. In: Feldman SI, Uretsky M, Najork M, Wills CE, eds. Proc. of the 13th Conf. on World Wide Web (WWW). Manhattan: ACM Press, 2004. 583-594.

[74] Amer-Yahia S, Lakshmanan LV, Pandit S. FlexPath: Flexible structure and full-text querying for XML. In: Weikum G, K?nig AC, De?loch S, eds. Proc. of the ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Paris: ACM Press, 2004. 83-94.

[75] Fuhr N, Gro?johann K. XIRQL: A query language for information retrieval in XML documents. In: Croft WB, Harper DJ, Kraft DH, Zobel J,

eds. Proc. of the 24th Annual Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR). New Orleans: ACM Press, 2001. 172-180.

[76] Balmin A, Papakonstantinou Y, Hristidis V. A system for keyword proximity search on XML databases. In: Freytag JC, Lockemann PC, Abiteboul S, Carey MJ, Selinger PG, Heuer A, eds. Proc. of the 29th Int'l Conf. on Very Large Data Bases (VLDB). Berlin: Morgan Kaufmann Publishers, 2003. 1069-1072.

[77] Weigel F, Meuss H, Schulz KU, Bry F. Content and structure in indexing and ranking XML. In: Amer-Yahia S, Gravano L, eds. Proc. of the 7th Int'l Workshop on the Web and Databases (WebDB). Maison de la Chimie: ACM Press, 2004. 67-72.

[78] Xu Y, Papakonstantinou Y. Efficient keyword search for smallest LCAs in XML databases. In: Proc. of the ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Baltimore: ACM Press, 2005.

[79] Guo L, Shao F, Botev C, Shanmugasundaram J. XRANK: Ranked keyword search over XML documents. In: Halevy AY, Ives ZG, Doan A, eds. Proc. of the 2003 ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). San Diego: ACM Press, 2003. 16-27.

[80] Kane B, Su H, Rundensteiner EA. Consistently updating XML documents using incremental constraint check queries. In: Chiang RHL, Lim EP, eds. Proc. of the 4th ACM CIKM Int'l Workshop on Web Information and Data Management (WIDM). McLean: ACM Press, 2002. 1-8.

[81] Bouchou B, Halfeld M, Alves F. Updates and incremental validation of XML documents. In: Lausen G, Suciú D, eds. Proc. of the 9th Int'l Conf. on Data Base Programming Languages (DBLP). LNCS 2921, Potsdam: Springer-Verlag, 2003. 216-232.

[82] Thompson H. xsv: Schema validator. 2002. <http://www.w3c.org/2001/03/webdata/xsv>

附中文参考文献:

[1] 孟小峰,周龙骧,王珊.数据库技术发展趋势.软件学报,2004,15(12):1822-1836. <http://www.jos.org.cn/1000-9825/15/1822.htm>

[49] 王静,孟小峰,王珊.基于区域划分的XML结构连接.软件学报,2004,15(5):720-729. <http://www.jos.org.cn/1000-9825/15/720.htm>