A parallel dimensionality reduction for time-series data and some of its applications

Thanh H.C.

Department of Informatics, Hanoi University of Science, VNUH, 334 Nguyen Trai Rd., Hanoi, Viet Nam

Abstract: The subsequence matching in a large time-series database has been an interesting problem. Many methods have been proposed that cope with this problem in an adequate extent. One of the good ideas is reducing properly the dimensionality of time-series data. In this paper, we propose a new method to reduce the dimensionality of high-dimensional time-series data. The method is simpler than existing ones based on the discrete Fourier transform and the discrete cosine transform. Furthermore, our dimensionality reduction may be executed in parallel. The method is used to time-series data matching problem and it decreases drastically the complexity of the corresponding algorithm. The method preserves planar geometric blocks and it is also applied to minimum bounding rectangles as well. Copyright © 2011 Inderscience Enterprises Ltd.

Author Keywords: Database; Dimensionality reduction; Matching problem; MBR; Minimum bounding rectangle; Time-series data

Index Keywords: Database; Dimensionality reduction; Matching problems; MBR; Minimum bounding rectangle; Time-series data; Cosine transforms; Database systems; Discrete cosine transforms; Discrete Fourier transforms; Geometry; Data reduction

Authors with affiliations:

• Thanh, H.C., Department of Informatics, Hanoi University of Science, VNUH, 334 Nguyen Trai Rd., Hanoi, Viet Nam

References:

• Agrawal, R., Faloutsos, C., Swami, A., Efficient similarity search in sequence databases (1993) Proceedings of the 4th International Conference on Foundations of Data Organization and Algorithms, pp. 69-84. , USA

- Cornen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C., (2001) Introduction to Algorithms, , The, MIT Press
- Faloutsos, C., Ranganathan, M., Manolopoulos, Y., Fast subsequence matching in time-series databases (2001) Proceedings of the International Conference on Management of Data, pp. 419-429. , ACM SIGMOD
- Keogh, E., Chakrabarti, K., Pazzani, M., Mehrotra, S., Dimensionality reduction for fast similarity search in large time-series database (2000) Journal of Knowledge and Information Systems, 3 (3), pp. 263-286
- Keogh, E., Chakrabarti, K., Mehrotra, S., Pazzani, M., Locally adaptive dimensionality reduction for indexing large time-series databases (2001) Proceedings of the International Conference on Management of Data, pp. 151-162. , ACM SIGMOD
- Moon, Y.S., An MBR-safe transformation for high-dimensional MBRs in similar sequence matching (2007) Proceedings of the International Conference on Database systems for Advanced Applications, , Thailand
- Moon, Y.S., Kim, J., A theoretical study on MBR-safe transformations (2007) Proceedings of the 12th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, , Italy
- Moon, Y.S., Whang, K.Y., Han, W.S., General match: A subsequence matching method in time-series databases based on generalized windows (2002) Proceedings of the International Conference on Management of Data, pp. 382-393. , ACM SIGMOD
- Thanh, H.C., Transforming sequential processes of a net system into concurent ones (2007) International Journal of Knowledge-based and Intelligent Engineering Systems, 11 (6), pp. 391-397
- Thanh, H.C., Parallel dimensionality reduction transformation for time-series data (2009) ACIIDS 2009, IEEE Computer Society, pp. 104-108. , Ngoc Thanh Nguyen, Huynh Phan Nguyen and Adam Grzech (Eds.)