

EINLADUNG

Zeit: Donnerstag, 11.11.2010, 16.30 Uhr

Ort: AH I, Ahornstr. 55

Referent: Michael E. Houle
Visiting Professor
National Institute of Informatics
Tokyo, Japan

Titel: **The Relevant-Set Correlation Model for Data Clustering**

Abstract:

This presentation introduces a model for clustering, the Relevant-Set Correlation (RSC) model, that requires no direct knowledge of the nature or representation of the data. Instead, the RSC model relies solely on the existence of an oracle that accepts a query in the form of a reference to a data item, and returns a ranked set of references to items that are most relevant to the query. The quality of cluster candidates, the degree of association between pairs of cluster candidates, and the degree of association between clusters and data items are all assessed according to the statistical significance of a form of correlation among pairs of relevant sets and/or candidate cluster sets. The RSC significance measures can be used to evaluate the relative importance of cluster candidates of various sizes, avoiding the problems of bias found with other shared-neighbor methods that use fixed neighborhood sizes.

A scalable clustering heuristic based on the RSC model, GreedyRSC, is also presented. GreedyRSC has many desirable characteristics, including the ability to determine the number of clusters naturally and automatically, without requiring the user to supply data-dependent parameter values. GreedyRSC can be implemented so as to make efficient use of disk storage, allowing the processing of data sets that are too large to fit in main memory. Experimental results are provided that demonstrate the effectiveness of the heuristic for very large, high-dimensional datasets, including image, protein sequence, and text, using the SASH approximate similarity search structure as the oracle.

Es laden ein: Die Dozenten der Informatik

Fachgruppe Informatik

Formlos überreicht mit der Bitte um Kenntnisnahme und Aushang

Prof. Dr. B. Vöcking, Informatik 1
Prof. Dr. P. Rossmanith, Informatik 1
Prof. Dr. Ir. J.-P. Katoen, Informatik 2
Prof. (em.) Dr. K. Indermark, Informatik 2
Prof. Dr. J. Giesl, Informatik 2
Prof. Dr. E. Ábrahám, Informatik 2
Prof. (em.) Dr. M. Nagl, Informatik 3
Prof. Dr. B. Rumpe, Informatik 3
Prof. Dr. H. Lichter, Informatik 3
Prof. Dr. O. Spaniol, Informatik 4
Prof. Dr. K. Wehrle, Informatik 4
Prof. Dr. M. Jarke, Informatik 5
Prof. G. Lakemeyer, Informatik 5
Prof. Dr. H. Ney, Informatik 6
Prof. Dr. W. Thomas, Informatik 7 (2 Exemplare)
Prof. (em.) Dr. W. Oberschelp, Informatik 7
Prof. Dr. E. Grädel, Informatik 7
Prof. Dr. L. Kobbelt, Informatik 8
Prof. Dr. B. Leibe, Informatik 8
Prof. Dr. Th. Seidl, Informatik 9
Prof. Dr. U. Schroeder, Informatik 9
Prof. Dr. J. Borchers, Informatik 10
Prof. Dr. S. Kowalewski, Informatik 11
Prof. C. Bischof, Ph.D., Informatik 12
Prof. Dr. U. Naumann, Informatik 12
Prof. Dr. F. Wolf, GRS
Prof. P. Bientinesi, Ph.D., AICES
Prof. Dr. T. Kuhlen, Informatik 12
Prof. Dr. U. Meyer, UMIC, IT-Security
Prof. Dr. J. Gross, UMIC, Mobile Network Performance
Prof. Dr. T. Berlage, Fraunhofer Institut f. Angewandte Informationstechnik
Prof. W. Prinz, Ph.D., Fraunhofer Institut f. Angewandte Informationstechnik
Prof. Dr. T. Rose, Fraunhofer Institut f. Angewandte Informationstechnik
Prof. Dr. J. Roßmann, Lehrstuhl f. Mensch-Maschine-Interaktion
Prof. Dr. M. Wiegner, Lehrstuhl I f. Mathematik
Prof. Dr. E. Triesch, Lehrstuhl II f. Mathematik für Ingenieure
Prof. Dr. W. Dahmen, Lehrstuhl f. Mathematik und Institut f. Geometrie und Praktische Mathematik
Prof. Dr. A. Reusken, Lehrstuhl f. Numerische Mathematik
Prof. Dr. A. Krieg, Lehrstuhl A f. Mathematik
Prof. Dr. W. Plesken, Lehrstuhl B f. Mathematik
Prof. Dr. H. Th. Jongen, Lehrstuhl C f. Mathematik
Prof. Dr. G. Hiß, Lehrstuhl D f. Mathematik
Prof. Dr. G. Nebe, Lehr- und Forschungsgebiet Mathematik (Algebra)
Prof. Dr. E. Zerz, Lehr- und Forschungsgebiet Mathematik
Prof. Dr. M. Herty, Lehr- und Forschungsgebiet Mathematik
Prof. Dr. U. Kamps, Lehrstuhl f. Statistik
Prof. Dr. H.-H. Bock, Lehr- und Forschungsgebiet Angewandte Statistik
Prof. Dr. A. Steland, Lehrstuhl f. Stochastik
Prof. Dr. E. Cramer, Lehr- und Forschungsgebiet Angewandte Stochastik
Prof. Dr. D. Bothe, MathCCES, Pauwelsstraße
Prof. Dr. R. Mathar, Lehrstuhl f. theoretische Informationstechnik
Prof. Dr. T. Noll, Lehrstuhl f. Allgemeine Elektrotechnik und Datenverarbeitungssysteme
Prof. Dr. Th. Bemmerl, Lehrstuhl f. Betriebssysteme
Prof. Dr. J.-R. Ohm, Lehrstuhl f. Elektrische Nachrichtentechnik
Prof. Dr. P. Vary, Lehrstuhl u. Institut f. Nachrichtengeräte u. Datenverarbeitung
Prof. Dr. B. Rembold, Lehrstuhl u. Institut f. Hochfrequenztechnik
Prof. Dr. T. Aach, Lehrstuhl f. Bildverarbeitung
Prof. Dr. G. Ascheid, Lehrstuhl f. Integrierte Systeme d. Signalverarbeitung
Prof. Dr. B. Walke, Lehrstuhl f. Kommunikationsnetze
Prof. Dr. W. Michaeli, Institut f. Kunststoffverarbeitung
Prof. Dr. D. Abel, Institut f. Regelungstechnik
Prof. Dr. G. Hirt, Institut f. Bildsame Formgebung
Prof. Dr. M. Bastian, Lehrstuhl f. Wirtschaftsinformatik