

Personal identification No. (PIN)

Family Name of Applicant

42579

Han

A. Five Most Significant Research Contributions

1. A comprehensive study of data mining methods [TKDE'96, DMBK'00, GDM'01] and the development of the DBMiner system [KDD'96] (<http://db.cs.sfu.ca/DBMiner>), based on our proposed on-line analytical mining (OLAM) methodology [SIGMOD Record'98] for integrated OLAP-mining. The system has been licensed to industry and used in industries and universities.
2. In-depth studies of scalable data mining methods, including frequent pattern-growth [SIGMOD'00], constraint-based mining query optimization [SIGMOD'98, SIGMOD'99, Computer'99], meta-rule guided, cube-based mining [KDD'97], multi-level mining of association rules [TKDE'97/VLDB'95], etc.
3. Investigation of methods for spatial data mining [CACM'98], including spatial clustering [VLDB'94,ICDT'01], spatial association [SSD'95], spatial classification [SDH'98] and spatial data cube construction [TKDE'99], and the development of GeoMiner system.
4. Investigation of methods for time-series mining [ICDE'99], object-oriented database mining [DKE'98], multimedia mining, Weblog mining, and Web mining [KDD'95], and the development of MultiMediaMiner and WeblogMiner [ADL'98].
5. Further studies on deductive and object-oriented databases, including evaluation of declarative recursive programs [IS'98, TKDE'95, TKDE'94], indexing for object-oriented and spatial databases [TKDE'98, VLDBJ'96, IJGIS'95] and the development of the LogicBase system (<http://db.cs.sfu.ca/LogicBase>).

B. Other Research Contributions

B.1. Selected articles in refereed journals since 1996:

1. [TKDE'01c] K. Wang, Y. He and J. Han, "*Pushing Support Constraints into Association Mining*", accepted by IEEE Transactions on Knowledge and Data Engineering, August 2001.
2. [TKDE'01b] A. K. H. Tung, H. Lu, J. Han, and L. Feng, "*Breaking the Barrier of Transactions: Mining Inter-Transaction Association Rules*", accepted by IEEE Transactions on Knowledge and Data Engineering (with minor revision), Mar. 2001.
3. [TKDE'01a] R. Ng and J. Han, "*CLARANS: A Method for Clustering Objects for Spatial Data Mining*", accepted by IEEE Transactions on Knowledge and Data Engineering (with minor revision), Dec. 2000.
4. [TOIS'00] H. Lu, J. Han, and L. Feng, "*Beyond Intra-Transaction Association Analysis: Mining Multi-Dimensional Inter-Transaction Association Rules*", ACM Transactions on Information Systems, 18(4): 423-454, 2000.
5. [JIIS'00] D. Cheung, C. Hwang, A. Fu, and J. Han, "*Efficient Rule-Based Attributed-Oriented Induction for Data Mining*", Journal of Intelligent Information Systems, 15(2): 175-200, 2000.
6. [TKDE'00] N. Stefanovic, J. Han, and K. Koperski, "*Object-Based Selective Materialization for Efficient Implementation of Spatial Data Cubes*," IEEE Transactions on Knowledge and Data Engineering, 12(6): 938-958, 2000.
7. [Computer'99] J. Han, L. V. S. Lakshmanan, and R. T. Ng, "*Constraint-Based, Multidimensional Data Mining*", COMPUTER (special issues on Data Mining), 32(8): 46-50, 1999.
8. [TKDE'99] J. Han and Y. Fu, "*Discovery of Multiple-Level Association Rules from Large Databases*", IEEE Transactions on Knowledge and Data Engineering, 11(5):798-805, 1999.
9. [TKDE'99b] J. Han, Z. Xie, and Y. Fu, "*Join Index Hierarchy: An Indexing Structure for Efficient Navigation in Object-Oriented Databases*", IEEE Transactions on Knowledge and Data Engineering, 11(2):321-337, 1999.
10. [DKE'98] J. Han, S. Nishio, H. Kawano, and W. Wang, "*Generalization-Based Data Mining in Object-Oriented Databases Using an Object-Cube Model*", Data and Knowledge Engineering, 25(1-2):55-97, 1998.

Personal identification No. (PIN)

Family Name of Applicant

42579

Han

11. [IS'98] J. Han, L. Liu and T. Lu, "*Evaluation of Declarative N-Queens Recursion: A Deductive Database Approach*", Information Sciences, 105:69-100, 1998.
12. [SIGMOD Record'98] J. Han, "*Towards On-Line Analytical Mining in Large Databases*", ACM SIGMOD RECORD, 27(1):97-107, 1998.
13. [TKDE'96] M.-S. Chen, J. Han, and P.S. Yu, "*Data Mining: An Overview from Database Perspective*", IEEE Trans. Knowledge and Data Engineering, 8(6):866-883, 1996.
14. [TKDE'96b] J. Han, Y. Huang, N. Cercone, and Y. Fu, "*Intelligent Query Answering by Knowledge Discovery Techniques*", IEEE Trans. Knowledge and Data Engineering, 8(3): 373-390, 1996.
15. [VLDBJ'96] B. C. Ooi, J. Han, H. Lu, and K. L. Tan, "*Index-Nesting — An Efficient Approach to Indexing in Object-Oriented Databases*", The VLDB Journal, 5(3):215-228, 1996.

B.1(b) Books and Edited Conference Proceedings

1. [DMBK'00] J. Han and M. Kamber, **Data Mining: Concepts and Techniques**, (Foreword by Jim Gray), Morgan Kaufmann, August 2000 (550 + xxiv pages).
2. [GDM'01] H. Miller and J. Han (eds.), **Geographic Data Mining and Knowledge Discovery**, Taylor and Francis, (to be published in 2001).

B.3. Selected articles in refereed conference proceedings and monographs since 1994:

1. [ICDM'01] W. Li, J. Han, and J. Pei, "*CMAR: Accurate and Efficient Classification Based on Multiple Class-Association Rules*", Proc. 2001 Int. Conf. on Data Mining (ICDM'01), San Jose, CA, Nov. 2001.
2. [ICDM'01] J. Pei, J. Han, H. Lu, S. Nishio, S. Tang, and D. Yang, "*H-Mine: Hyper-Structure Mining of Frequent Patterns in Large Databases*", Proc. 2001 Int. Conf. on Data Mining (ICDM'01), San Jose, CA, Nov. 2001.
3. [CIKM'01] H. Pinto, J. Han, J. Pei, K. Wang, Q. Chen, and U. Dayal, "*Multi-Dimensional Sequential Pattern Mining*", Proc. 2001 Int. Conf. on Information and Knowledge Management (CIKM'01), Atlanta, GA, Nov. 2001.
4. [VLDB'01] G. Dong, J. Han, J. Lam, J. Pei, and K. Wang, "*Mining Multi-Dimensional Constrained Gradients in Data Cubes*", Proc. 2001 Int. Conf. on Very Large Data Bases (VLDB'01), Rome, Italy, Sept. 2001.
5. [DaWaK'01] S. H. S. Chee, J. Han, and K. Wang, "*RecTree: An Efficient Collaborative Filtering Method*", Proc. 2001 Int. Conf. on Data Warehouse and Knowledge Discovery (DaWaK'01), Munich, Germany, Sept. 2001.
6. [KDD'01] W. Jin, K.H. Tung and J. Han, "*Mining Top-n Local Outliers in Large Databases*", Proc. 2001 Int. Conf. on Knowledge Discovery in Databases (KDD'01), San Fransisco, California, Aug. 2001, pp. 293-298.
7. [SIGMOD'01] J. Han, J. Pei, G. Dong, and K. Wang, "*Efficient Computation of Iceberg Cubes with Complex Measures*", Proc. 2001 ACM-SIGMOD Int. Conf. on Management of Data (SIGMOD'01), Santa Barbara, CA, May 2001, pp. 1-12. (Also, ACM SIGMOD RECORD, 30(3), 2001).
8. [ICDE'01] J. Pei, J. Han, B. Mortazavi-Asl, H. Pinto, Q. Chen, U. Dayal, and M.-C. Hsu, "*PrefixSpan: Mining Sequential Patterns Efficiently by Prefix-Projected Pattern Growth*", Proc. 2001 Int. Conf. on Data Engineering (ICDE'01), Heidelberg, Germany, April 2001, pp. 215-224.
9. [ICDE'01] A. K. H. Tung, J. Hou, and J. Han, "*Spatial Clustering in the Presence of Obstacles*", Proc. 2001 Int. Conf. on Data Engineering (ICDE'01), Heidelberg, Germany, April 2001, pp. 359-367.
10. [ICDE'01] J. Pei, J. Han, and L. V. S. Lakshmanan, "*Mining Frequent Itemsets with Convertible Constraints*", Proc. 2001 Int. Conf. on Data Engineering (ICDE'01), Heidelberg, Germany, April 2001, pp. 433-332.

Personal identification No. (PIN)

42579

Family Name of Applicant

Han

11. [ICDT'01] A. K. H. Tung, J. Han, L. V. S. Lakshmanan, and R. T. Ng, "*Constraint-Based Clustering in Large Databases*", Proc. 2001 Int. Conf. on Database Theory (ICDT'01), London, U.K., Jan. 2001. (15 pages)
12. [VLDB'00] K. Wang, Y. He and J. Han, "*Mining Frequent Itemsets Using Support Constraints*", Proc. 2000 Int. Conf. on Very Large Data Bases (VLDB'00), Cairo, Egypt, Sept. 2000, pp. 43-52.
13. [DEXA'00] E. D. Kim, J. M. W. Lam, and J. Han, "*AIM: Approximate Intelligent Matching for Time Series Data*", Proc. 2000 Int. Conf. on Data Warehouse and Knowledge Discovery (DaWaK'00), Greenwich, U.K., Sept. 2000, pp. 347-357.
14. [KDD'00] J. Han, J. Pei, B. Mortazavi-Asl, Q. Chen, U. Dayal, M.-C. Hsu, "*FreeSpan: Frequent Pattern-Projected Sequential Pattern Mining*", Proc. 2000 Int. Conf. on Knowledge Discovery and Data Mining (KDD'00), Boston, MA, August 2000, pp. 355-359.
15. [KDD'00] J. Pei and J. Han "*Can We Push More Constraints into Frequent Pattern Mining?*", Proc. 2000 Int. Conf. on Knowledge Discovery and Data Mining (KDD'00), Boston, MA, August 2000, pp. 350-354.
16. [SIGMOD'00] J. Han, J. Pei, and Y. Yin, "*Mining Frequent Patterns without Candidate Generation*", Proc. 2000 ACM-SIGMOD Int. Conf. on Management of Data (SIGMOD'00), Dallas, TX, May 2000, pp. 1-12.
17. [PAKDD'00] J. Pei, J. Han, B. Mortazavi-Asl, and H. Zhu, "*Mining Access Pattern efficiently from Web logs*", Proc. 2000 Pacific-Asia Conf. on Knowledge Discovery and Data Mining (PAKDD'00), Kyoto, Japan, April 2000, pp. 396-407.
18. [PAKDD'00] A. K. H. Tung, J. Hou, and J. Han, "*COE: Clustering with Obstacles Entities, A Preliminary Study*", Proc. 2000 Pacific-Asia Conf. on Knowledge Discovery and Data Mining (PAKDD'00), Kyoto, Japan, April 2000, pp. 165-168.
19. [ICDE'00] O. R. Zaïane, J. Han, and H. Zhu, "*Mining Recurrent Items in Multimedia with Progressive Resolution Refinement*", Proc. 2000 Int. Conf. on Data Engineering (ICDE'00), San Diego, CA, Feb. 2000, pp. 461-470.
20. [KDD'99] A. K. H. Tung, H. Lu, J. Han, and L. Feng, "*Breaking the Barrier of Transactions: Mining Inter-Transaction Association Rules*", Proc. 1999 Int. Conf. on Knowledge Discovery and Data Mining (KDD'99), San Diego, CA, Aug. 1999.
21. [SSD'99] X. Zhou, D. Truffet, and J. Han, "*Efficient Polygon Amalgamation Methods for Spatial OLAP and Spatial Data Mining*", Proc. 6th Int. Symp. on Large Spatial Databases (SSD'99), Hong Kong, July 1999, pp. 167-187.
22. [SIGMOD'99] L. V. S. Lakshmanan, R. Ng, J. Han and A. Pang, "*Optimization of Constrained Frequent Set Queries with 2-Variable Constraints*", Proc. 1999 ACM-SIGMOD Conf. on Management of Data (SIGMOD'99), Philadelphia, PA, June 1999, pp. 157-168.
23. [ICDE'99] J. Han, G. Dong and Y. Yin, "*Efficient Mining of Partial Periodic Patterns in Time Series Database*", Proc. 1999 Int. Conf. on Data Engineering (ICDE'99), Sydney, Australia, March 1999, pp. 106-115.
24. [HBK'99] J. Han, "*Characteristic Rules*", and "*DBMiner*", in Handbook of Data Mining and Knowledge Discovery, W. Klösgen and J. Zytkow (eds.), Oxford University Press, 1999.

B.5. Technology transfer

- MPR Teltech Ltd. industry license: The DBMiner data mining system (January 1996).
- The DBMiner data mining system. Technology transfer to DBTech Software Inc. (March 1997).

B.6. Participation in industrially relevant R & D activities

Personal identification No. (PIN)

42579

Family Name of Applicant

Han

- Project leader and principal investigator, NCE/IRIS-3:HAN-06 (Building, Querying, Analyzing, and Mining Data Warehouses on the Internet), working with IBM Canada (1998–2002).
- Principal investigator, NCE/GEOID:3, (Cartographic Interface for Multi-Dimensional Exploration of Environmental, Health Indicators on the World-Wide Web”, (1998–2002).
- Working (1994-1998) with Boeing, Tandem Computers, HP Lab, Hughes Aircraft Research Lab, Seagate Software, IBM Canada Pacific Development Centre, ISM-BC on data warehousing and/or data mining.

C. Contributions to the Training of Highly Qualified Personnel:

- Current graduate students (as senior-supervisor): 4 Ph.D. and 3 M.Sc. students.
4 Ph.D. students are: Wen Jin, Eddie Kim, Jian Pei, and Zhong Zhang. 3 M.Sc. students are: Eugene Belchev, Hamayoun Afshar, and Yvonne Zheng. They are all working on data mining, data warehousing, and/or Web technology.
- Graduate degrees conferred (as sole senior-supervisor): 20 M.Sc. and 8 Ph.D. degrees conferred since 1995: Yongjian Fu: Ph.D.'96 (Univ. Missouri); Nebojsa Stefanovic: M.Sc.'97 (Seagate Software); Wan Gong: M.Sc.'97 (Seagate Software); Betty Xia: M.Sc.'97 (Empress software, Toronto); Shan Cheng: M.Sc.'98 (Statistics Canada, Ottawa); Gabor Melli: M.Sc.'98 (DataSage, Boston); Y. Jenny Tam: M.Sc.'98 (Simba Tech., Vancouver); Hua Zhu: M.Sc.'98 (DBMiner Tech.); Osmar R. Zaiane: Ph.D.'99 (U. Alberta); Kris Koperski: Ph.D.'99 (Mathsoft Inc., Seattle); Wei Wang: M.Sc.'99 (Seagate Software); Qing Chen: M.Sc.'99 (DBMiner Tech.); Jin Li: M.Sc.'99 (Seagate Software); Jean Fen-ju Hou: M.Sc.'99 (CyberSource, California); Sonny H. S. Chee: M.Sc.'00 (Pivotal Technology), Xuebin Lu: M.Sc.'00 (software engineer, Vancouver); Anthony Tung: Ph.D.'01 (National U. of Singapore), Haiming Huang: M.Sc.'01 (software engineer, Vancouver); Julia Itskevitch: M.Sc.'01 (software engineer, Pittsburg); Joyce Man Lam: M.Sc.'01 (software engineer, Vancouver); George Wenmin Li: M.Sc.'01 (software engineer, Vancouver); Nancy Liao: M.Sc.'01 (software engineer, Vancouver); Runying Mao: M.Sc.'01 (software engineer, Microsoft); Behzad Mortazavi-Asl: M.Sc.'01 (software engineer, Vancouver); Helen Pinto: M.Sc.'01 (software engineer, Alberta Research Council); Zhaoxia Wang: M.Sc.'01 (CS instructor, Vancouver); and Yiwen Yin: M.Sc.'01 (software engineer, Vancouver).
- Visiting scholars supervised: 7 (from Australia, Canada, China, Japan, Germany, and Singapore).

D. Other Evidence of Impact and Contributions:

Awards/Honors (since 1994):

- Listed in *Who's Who in the World* (1993-2000), *Who's Who in the West* (1994-2000), *Who's Who in Science and Engineering* (1994-2000), etc. Excellent paper award, *8th JSAI Conf.*, Tokyo, Japan (1994).

Journal editorial boards:

- Subject editor, *IEEE Transactions on Knowledge and Data Engineering*. (1996-2000)
- Subject editor, *Journal of Data Mining and Knowledge Discovery*, Kluwer Publishers. (1998-present)
- Subject editor, *Int. Journal of Intelligent Information Systems*, Kluwer Publishers. (1998-present)

Refereeing and reviewing:

- Refereeing: Many journals, including *ACM Trans. Database Systems (TODS)*, *IEEE Trans. Knowledge and Data Eng. (TKDE)*, *Data Mining and Knowledge Discovery (DAMI)*, *IEEE Trans. Software Eng. (TSE)*, *Communications of ACM (CACM)*, *COMPUTER*, *The VLDB Journal (VLDBJ)*, etc. and over 30 international conferences/workshops.
- Evaluation of: Research grant applications: *NSERC* ('90-'00), *U.S./NSF Review Panel* ('97), *B.C. Science Council* ('94-'95), *B.C. Advanced Systems Inst.* ('96), *Hong Kong Research Council* ('95-'00), *Australia Research Council* ('97-'98), *Austria Research Council* ('98). Tenure/promotion: 5 in Canada, 5 in U.S., 1 in Germany, and 1 in Singapore. Ph.D./M.Sc. theses: (as external examiner): 12.

Personal identification No. (PIN)

42579

Family Name of Applicant

Han

Conference and workshop organization:

- **Conference/workshop co-chairman, vice-chairman, or co-organizer (since 1995)**

ACM-SIGKDD (executive committee member), 2002 SIAM Data Mining Conf. (SIAMDM'02) PC co-chairman, 2002 Int. Conf. on Knowledge Discovery and Data Mining (KDD'01) student awards chairman, 2002 Int. Conf. Data Engineering (ICDE'02) (vice PC chairman). 2001 Int. Conf. on Knowledge Discovery and Data Mining (KDD'01) best paper awards chairman, 2001 SIAM Data Mining Conf. (SIAMDM'01) PC co-chairman, 2001 Pacific-Asia Conf. on Knowledge Discovery and Data Mining (PAKDD'01) conference co-chairman, 2000 ACM SIGMOD Int. Conf. on Management of Data (SIGMOD'00) (demo and exhibits chairman), 1999 Int. Conf. Knowledge Discovery and Data Mining (KDD'99) (tutorial chairman), 1997 Int. Conf. Data Engineering (ICDE'97) (panel co-ordinator), 1996 Int. Conf. Knowledge Discovery and Data Mining (KDD'96) (program committee co-chairman), SIGMOD'96 and SIGMOD'97 Workshops on Research Issues on Data Mining and Knowledge Discovery (DMKD'96 & '97) (workshop co-organizer), Int. Workshop on Integration of Knowledge Discovery and Deductive and Object-Oriented Databases (KDOOD'95) (workshop co-organizer), 1995 Int. Conf. Data Engineering (ICDE'95) (vice PC chairman).

- **Program committee member (since 1995)**

ACM SIGMOD Int. Conf's Management of Data (SIGMOD'2002,'99/96), Int. Conf's Data Mining and Knowledge Discovery (KDD'96-99,'01, '02), Int. Conf. Very Large Databases (VLDB'2000/96), Int. Conf. Extending Data Base Technology (EDBT'2000), Int. Conf's Data Engineering (ICDE'02,99/98/95), Int. Symp's Large Spatial Databases (SSD'99/97), SIGMOD Workshops Research Issues on Data Mining and Knowledge Discovery (DMKD'98-01), Int. Conf. Information and Knowledge Management (CIKM'99-00'01, 95-97), and many others.

Seminars, invited talks, keynote speeches and conference tutorials:

- **Tutorials for international conf./symp./workshops (since 1995)**

2000 Int. Database Systems Conf. (IDS'2000) (Hong Kong); 1998 Pacific-Asia Conf. Knowledge Discovery and Data Mining (PAKDD-98), (Melbourne, Australia); 1997 Int. Symp. Large Spatial Databases (SSD'97) (Berlin, Germany), 1997 Int. Conf. Data Engineering (ICDE'97) (Birmingham, England), 1996 Int. Conf. Cooperative Information Systems (CoopIS'96) (Brussels, Belgium), 1996 ACM SIGMOD Int. Conf. Management of Data (SIGMOD'96) (Montreal, Canada), 1995 Int. Conf. Data Engineering (ICDE'95) (Taipei, Taiwan).

- **Invited speeches for international conf./symp./workshops (since 1995)**

(Invited tutor) UCLA Extension Short Course: Data Mining (L.A., California, 1998—1999, 2002); (tutor) Int. Conf. Knowledge Discovery and Data Mining (KDD-01) (San Francisco, 2001); (Invited tutor) IBM Chair Short Course on Data Mining (Antwerp, Belgium, 1999); (Invited tutor) IBM China Summer Course on Data Mining and Data Warehousing (Shanghai, China, 1998); 1998 Int. Database Engineering and Applications Symp. (IDEAS-98), (Cardiff, England); 1998 IEEE Advances in Digital Libraries Conf. (ADL-98) (Santa Barbara, CA, 1998); 1998 Pacific-Asia Conf. Knowledge Discovery and Data Mining (PAKDD-98) (Melbourne, Australia, 1998); 7th IFIP 2.6 Working Conference on Database Semantics (DS-7), Searching for Semantics: Data Mining, Reverse Engineering, etc. (Leysin, Switzerland, 1997); NASA-CESDIS Workshop on Data Mining, Data Warehousing, and High Performance Computing (Greenbelt, MD, 1997); Microsoft Research and Univ. of Washington, Summer Institute of Data Mining (Seattle, WA, 1997); Data Mining and Beyond (Portland, Oregon, 1996); CIKM'95 (Baltimore, Maryland, 1995).