

{tag}

{/tag}

International Journal of Computer Applications

© 2014 by IJCA Journal

Volume 89 - Number 5

Year of Publication: 2014

Authors:

Do Van Nguyen

Koichi Yamada

Muneyuki Unehara

10.5120/15495-4286

{bibtex}pxc3894286.bib{/bibtex}

## Abstract

The original rough set theory deals with precise and complete data, while real applications frequently contain imperfect information. A typical imperfect data studied in rough set research is the missing values. Though there are many ideas proposed to solve the issue in the literature, the paper adopts a probabilistic approach, because it can incorporate other types of imperfect data including imprecise and uncertain values in a single approach. The paper first discusses probabilities of attribute values assuming different type of attributes in real applications, and proposes a generalized method of probability of matching. This probability is then used to define valued tolerance/similarity relations and to develop new rough set models based on the valued tolerance/similarity relations. An algorithm for deriving decision rules based on the rough set models is also studied and proposed.

**Refer**

**ences**

- D. Barbar'a, H. Garcia-Molina, and D. Porter. The management of probabilistic data.

IEEE Trans. on Knowl. and Data Eng. , 4(5):487–502, October 1992.

- Roger Cavallo and Michael Pittarelli. The theory of probabilistic databases. In Proceedings of the 13th International Conference on Very Large Data Bases, VLDB '87, pages 71–81, San Francisco, CA, USA, 1987. Morgan Kaufmann Publishers Inc.

- Peter Clark and Tim Niblett. The cn2 induction algorithm. Mach. Learn. , 3(4):261–283, March 1989.

- Yucai Feng, Wenhai Li, Zehua Lv, and Xiaoming Ma. Probabilistic approximation under incomplete information systems. In Zhongzhi Shi, K. Shimohara, and David Dagan Feng, editors, Intelligent Information Processing, volume 228 of IFIP, pages 73–80. Springer, 2006.

- Jerzy W. Grzymala-Busse. Characteristic relations for incomplete data: A generalization of the indiscernibility relation. 3700:58–68, 2005.

- Jerzy W. Grzymala-Busse. A rough set approach to data with missing attribute values. In Guoyin Wang, James F. Peters, Andrzej Skowron, and Yiyu Yao, editors, RSKT, volume 4062 of Lecture Notes in Computer Science, pages 58–67. Springer, 2006.

- Jerzy W. Grzymala-Busse. Rule induction. In Maimon and Rokach , pages 249–265.

- Jerzy W. Grzymala-Busse and Witold J. Grzymala-Busse. Handling missing attribute values. In Maimon and Rokach

- , pages 33–51.

- Jerzy W. Grzymala-Busse and Ming Hu. A comparison of several approaches to missing attribute values in data mining. In Revised Papers from the Second International Conference on Rough Sets and Current Trends in Computing, RSCTC '00, pages 378–385, London, UK, UK, 2001. Springer-Verlag.

- Jerzy W. Grzymala-Busse and Wojciech Rzasca. Definability and other properties of approximations for generalized indiscernibility relations. T. Rough Sets, 11:14–39, 2010.

- Marzena Kryszkiewicz. Rough set approach to incomplete information systems. Inf. Sci. , 112(1-4):39–49, December 1998.

- Marzena Kryszkiewicz. Rules in incomplete information systems. Inf. Sci. , 113(3-4):271–292, 1999.

- Nagamachi M. Kansei engineering: a new ergonomics consumer-oriented technology for product development. International Journal of Industrial Ergonomics, 15:3–10, 1995.

- Nagamachi M. Kansei engineering as a powerful consumer-oriented technology for product development. Fundamental Reviews in Applied Ergonomics, 33(3):289–294, 2002.

- Oded Maimon and Lior Rokach, editors. Data Mining and Knowledge Discovery Handbook, 2nd ed. Springer, 2010.

- A. Motro. Uncertainty management information systems: From needs to solutions. Kluwer Acad. , 1997.

- Do Van Nguyen, Koichi Yamada, and Muneyuki Unehara. Knowledge reduction in incomplete decision tables using probabilistic similarity-based rough set model. In 12th International Symposium on Advanced Intelligent Systems (ISIS 2011), pages 147–150, 2011.

- Do Van Nguyen, Koichi Yamada, and Muneyuki Unehara. Rough set model based on parameterized probabilistic similarity relation in incomplete decision tables. In Soft Computing and Intelligent Systems (SCIS) and 13th International Symposium on Advanced Intelligent Systems (ISIS), 2012 Joint 6th International Conference on, pages 577–582, 2012.

- Do Van Nguyen, Koichi Yamada, and Muneyuki Unehara. Extended tolerance relation to define a new rough set model in incomplete information systems. *Advances in Fuzzy Systems*, 2013, 2013. Article ID 372091.
- Do Van Nguyen, Koichi Yamada, and Muneyuki Unehara. On probability of matching in probabilistic based rough set definitions. In *IEEE-SMC2013*, pages 449–454, Manchester, The UK, 2013.
- Z. Pawlak. Rough sets. *International Journal of Computer and Information Sciences*, 11:341–356, 1982.
- Z. Pawlak. *Rough Sets. Theoretical Aspects of Reasoning about Data*. Kluwer Acad. , 1991.
- Jerzy Stefanowski and Alexis Tsoukis. On the extension of rough sets under incomplete information. In Ning Zhong, Andrzej Skowron, and Setsuo Ohsuga, editors, *RSFDGrC*, volume 1711 of *Lecture Notes in Computer Science*, pages 73–81. Springer, 1999.
- Jerzy Stefanowski and Alexis Tsoukis. Incomplete information tables and rough classification. *Computational Intelligence*, 17(3):545–566, 2001.
- Guoyin Wang. Extension of rough set under incomplete information systems. In *Fuzzy Systems, 2002. FUZZ-IEEE'02. Proceedings of the 2002 IEEE International Conference on*, volume 2, pages 1098–1103, 2002.
- Xibei Yang, Xiaoning Song, and Xiaohua Hu. Generalisation of rough set for rule induction in incomplete system. *IJGCRSIS*, 2(1):37–50, 2011.
- Y. Y. Yao. On generalizing rough set theory. In Guoyin Wang, Qing Liu, Yiyu Yao, and Andrzej Skowron, editors, *RSFDGrC*, volume 2639 of *Lecture Notes in Computer Science*, pages 44–51. Springer, 2003.
- Y. Y. Yao. Three-way decisions with probabilistic rough sets. *Information Sciences*, 2010(3):341–353, 2010.

### Index Terms

Computer Science

Information Sciences

### Keywords

Imperfect Information Systems    Probability of Matching    Approximation Space  
Rough Sets  
Decision Rules

