

ŻYCIORYS

dr hab. inż. Szymon Jaroszewicz

Rok urodzenia 1974
E-mail s.jaroszewicz@itl.waw.pl, sj@cs.umb.edu

Edukacja, zatrudnienie

2010	Habilitacja, Instytut Podstaw Informatyki, PAN
2006 –	Instytut Łączności, adiunkt
2004 – 2009	Politechnika Szczecińska, adiunkt
2003	Doktorat, University of Massachusetts w Bostonie. Tytuł pracy: “Information Theoretical and Combinatorial Methods in Data Mining”
1999 – 2003	Studia doktoranckie, kierunek informatyka, University of Massachusetts w Bostonie
1998 – 1999	stypendium Fulbrighta
1998	Dyplom: magister inżynier informatyk, Instytut Informatyki Politechniki Szczecińskiej
1993 – 1998	Studia magisterskie, Politechnika Szczecińska, kierunek informatyka

Stypendia, nagrody, wyróżnienia

2007	Nagroda za najlepszy referat na 11th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD)
2005	Stypendium dla młodych naukowców Fundacji na Rzecz Nauki Polskiej
2002	Grant wewnętrzny na University of Massachusetts w Bostonie
2000	Stypendium im. Randall G. Malbone na wydziale matematyki i informatyki University of Massachusetts w Bostonie
1998	Stypendium Fulbrighta na rok akademicki 1998/99: praca naukowa na University of Massachusetts w Bostonie
1997, 1996	Stypendia Ministra Edukacji Narodowej za wybitne osiągnięcia naukowe na lata akademickie 1996/97 i 1997/98

Działalność organizacyjna

- Recenzent czasopism: Data Mining and Knowledge Discovery, ACM Transactions on Database Systems, IEEE Transactions on Knowledge and Data Engineering, Data and Knowledge Engineering, IEEE Transactions on Information Theory, Transactions on Rough Sets, KAIS, JRPIT
- Członek komitetów programowych konferencji: PAKDD’06, Discovery Science’06, PKDD’06, PAKDD’07, ACM SAC’07 Data Mining Track, PKDD’07, PAKDD’08, ACM SAC’08 Data Mining Track, PKDD’08 (area chair), CIKM’08, ICDM’08, PAKDD’09, ACM SAC’09 Data Mining Track, KDIR’09, ACM CSE’09, CIKM’09, Local Patterns to Global Models Workshop’09
- Członek Kolegium Redaktorskiego czasopisma Data Mining and Knowledge Discovery (od 2009)

Praca naukowa związana z Data Mining

- Data-mining atrybutów numerycznych, współpraca z B. Goethalsem i T. Caldersem z Uniwersytetu Antwerpskiego
- Algorytmy oparte na próbkowaniu, współpraca z prof. T. Schefferem z Uniwersytetu Humboldta w Berlinie
- Analiza danych genetycznych z mikromacierzy i SAGE, badania prowadzone wspólnie z Decision Support Group, Brigham and Women's Hospital w Bostonie
- Miary jakości i wybór interesujących reguł asocjacyjnych

Informacje dodatkowe

- **Jezyki:** angielski (bardzo dobra znajomość), francuski, niemiecki (podstawowa znajomość)

Lista publikacji

- [1] S. Jaroszewicz. Using interesting sequences to interactively build hidden markov models. *Data Mining and Knowledge Discovery*, 2010. accepted.
- [2] S. Jaroszewicz. Discovering interesting patterns in numerical data with background knowledge. In Y.S. Koh and N. Rountree, editors, *Rare Association Rule Mining and Knowledge Discovery: Technologies for Infrequent and Critical Event Detection*, pages 118–130. IGI Global, 2010.
- [3] S. Jaroszewicz, T. Scheffer, and D.A. Simovici. Scalable pattern mining with bayesian networks as background knowledge. *Data Mining and Knowledge Discovery*, 18(1):56–100, 2009.
- [4] S. Jaroszewicz. Interactive HMM construction based on interesting sequences. In *Proc. of Local Patterns to Global Models (LeGo'08) Workshop at the 12th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD'08)*, pages 82–91, Antwerp, Belgium, 2008.
- [5] S. Jaroszewicz. Minimum variance associations — discovering relationships in numerical data. In *The Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, pages 172–183, Osaka, Japan, 2008.
- [6] T. Calders and S. Jaroszewicz. Efficient auc optimization for classification. In *11th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD'07)*, pages 42–53, Warsaw, Poland, 2007. **Best paper award.**
- [7] S. Jaroszewicz, L. Ivantysynova, and T. Scheffer. Schema matching on streams with accuracy guarantees. *Intelligent Data Analysis*, 12(3):253–270, 2008.
- [8] S. Jaroszewicz and M. Korzeń. Approximating representations for large numerical databases. In *7th SIAM International Conference on Data Mining (SDM'07)*, pages 521–526, Minneapolis, MN, 2007.
- [9] S. Jaroszewicz, L. Ivantysynova, and T. Scheffer. Accurate schema matching on streams. In *4th International Workshop on Knowledge Discovery from Data Streams at the 10th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD'06)*, pages 3–12, 2006.

- [10] T. Calders, B. Goethals, and S. Jaroszewicz. Mining rank-correlated sets of numerical attributes. In *Proc. of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'06)*, 2006.
- [11] S. Jaroszewicz. Polynomial association rules with applications to logistic regression. In *Proc. of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'06)*, 2006.
- [12] S. Jaroszewicz and M. Korzeń. Comparison of information theoretical measures for reduct finding. In *Proc. of the 8th International Conference on Artificial Intelligence and Soft Computing (ICAISC'06)*, LNAI 4029, pages 518–527, Zakopane, June 2006. Springer-Verlag.
- [13] D. A. Simovici and S. Jaroszewicz. A new metric splitting criterion for decision trees. *Parallel Algorithms Appl.*, 21(4):239–256, 2006.
- [14] D. A. Simovici and S. Jaroszewicz. Generalized conditional entropy and a metric splitting criterion for decision trees. In *10th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining (PAKDD'06)*, pages 35–44, Singapore, 2006.
- [15] S. Jaroszewicz and T. Scheffer. Fast discovery of unexpected patterns in data, relative to a bayesian network. In *11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2005)*, pages 118–127, Chicago, IL, August 2005.
- [16] D. Simovici and S. Jaroszewicz. A new metric splitting criterion for decision trees. *International Journal of Parallel, Emergent and Distributed Systems*, 21(4):239–256, August 2006.
- [17] S. Jaroszewicz and W. Kosiński. Machine learning for speech recognition researchers. In *Proceedings of the Speech Analysis, Synthesis and Recognition Applications of Phonetics Conference*, Kraków, Poland, September 2005. Publication on CD-ROM.
- [18] M. Korzeń and S. Jaroszewicz. Finding reducts without building the discernibility matrix. In *Proceedings of the Fifth International Conference on Intelligent Systems Design and Applications (ISDA'05)*, pages 450–455, Wrocław, Poland, 2005.
- [19] D. Simovici and S. Jaroszewicz. A metric approach to building decision trees based on Goodman-Kruskal association index. In *PAKDD 2004*, LNAI 3056, pages 181–190, Sydney, Australia, May 2004. Springer-Verlag.
- [20] S. Jaroszewicz, D. A. Simovici, and I. Rosenberg. Measures on Boolean polynomials and their applications in data mining. *Discrete Applied Mathematics*, 144(1–2):123–139, November 2004.
- [21] Szymon Jaroszewicz and Dan Simovici. Interestingness of frequent itemsets using bayesian networks as background knowledge. In *10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2004)*, pages 178–186, Seattle, WA, August 2004.
- [22] S. Jaroszewicz, D. A. Simovici, W. P. Kuo, and L. Ohno-Machado. The Goodman-Kruskal coefficient and its applications in genetic diagnosis of cancer. *IEEE Transactions on Biomedical Engineering*, 51(7):1095–1102, July 2004.
- [23] Szymon Jaroszewicz. *Information Theoretical and Combinatorial Methods in Data Mining*. PhD thesis, University of Massachusetts Boston, December 2003.

- [24] Dan A. Simovici and Szymon Jaroszewicz. Generalized conditional entropy and decision trees. In *Journées francophones d'Extraction et de Gestion de Connaissances (EGC 2003)*, pages 369–380, Lyon, France, January 2003.
- [25] S. Jaroszewicz and D. A. Simovici. Support approximations using Bonferroni-type inequalities. In *6th European Conference on Principles of Data Mining and Knowledge Discovery (PKDD 2002)*, pages 212–223, Helsinki, Finland, August 2002.
- [26] S. Jaroszewicz and D. A. Simovici. Pruning redundant association rules using maximum entropy principle. In *Advances in Knowledge Discovery and Data Mining, 6th Pacific-Asia Conference, PAKDD'02*, pages 135–147, Taipei, Taiwan, May 2002.
- [27] I. Rosenberg, D. A. Simovici, and S. Jaroszewicz. On functions defined on free boolean algebras. In *IEEE International Symposium on Multiple-Valued Logic ISMVL'02*, Boston, MA, May 2002.
- [28] S. Jaroszewicz, D. A. Simovici, and I. Rosenberg. An inclusion-exclusion result for boolean polynomials and its applications in data mining. In *Workshop on Discrete Mathematics and Data Mining (DM&DM), Second SIAM International Conference on Data Mining*, Arlington, VA, April 2002.
- [29] D. A. Simovici and S. Jaroszewicz. An axiomatization of partition entropy. *IEEE Transactions on Information Theory*, 48(7):2138–2142, July 2002.
- [30] S. Jaroszewicz and D. A. Simovici. A general measure of rule interestingness. In *5th European Conference on Principles of Data Mining and Knowledge Discovery (PKDD 2001)*, pages 253–265, 2001.
- [31] S. Jaroszewicz and D. A. Simovici. Data mining of weak functional decompositions. In *Proc. of the 30th International Symposium on Multiple-Valued Logic (ISMVL 2000)*, pages 77–82, 2000.
- [32] D. A. Simovici and S. Jaroszewicz. On information-theoretical aspects of relational databases. In C. Calude and G. Paun, editors, *Finite versus Infinite*. Springer Verlag, London, 2000.
- [33] S. Jaroszewicz and D. A. Simovici. On axiomatization of conditional entropy of functions between finite sets. In *Proceedings of the 29th International Symposium on Multiple-Valued Logic*, pages 24–28, Freiburg, Germany, 1999.
- [34] S. Jaroszewicz. Minimization of incompletely specified multiple-valued functions in reed-muller domain. In *Proc. of the 19th International Scientific Symposium for Students and Young Research Employees*, pages 121–126, Zielona Góra, Poland, 1997. (in Polish).
- [35] S. Jaroszewicz, V. Shmerko, and S. Yanushkevich. Exact irredundant searching for a minimal reed-muller expansion for an incompletely specified mvl function. In *Proc. of the International Conference on Application of Computer Systems*, pages 65–74, Szczecin, Poland, 1996.
- [36] A. D. Zakrevskij, S. N. Yanushkevich, and S. Jaroszewicz. Minimization of reed-muller expansions for systems of incompletely specified mvl functions. In *Proc. of the International Conference on Methods and Models in Automatics and Robotics (MMAR'96)*, pages 1085–1090, Międzyzdroje, Poland, 1996.