

Benjamin C M Fung

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/1640224/publications.pdf>

Version: 2024-02-01

118
papers

7,496
citations

117453

34
h-index

74018

75
g-index

124
all docs

124
docs citations

124
times ranked

5046
citing authors

#	ARTICLE	IF	CITATIONS
1	Privacy-preserving data publishing. ACM Computing Surveys, 2010, 42, 1-53.	16.1	1,207
2	A decision tree method for building energy demand modeling. Energy and Buildings, 2010, 42, 1637-1646.	3.1	470
3	A systematic procedure to study the influence of occupant behavior on building energy consumption. Energy and Buildings, 2011, 43, 1409-1417.	3.1	461
4	Hierarchical Document Clustering Using Frequent Itemsets. , 2003, , .		288
5	Anonymizing Classification Data for Privacy Preservation. IEEE Transactions on Knowledge and Data Engineering, 2007, 19, 711-725.	4.0	261
6	Security and privacy challenges in smart cities. Sustainable Cities and Society, 2018, 39, 499-507.	5.1	220
7	Differentially private data release for data mining. , 2011, , .		215
8	A review of the-state-of-the-art in data-driven approaches for building energy prediction. Energy and Buildings, 2020, 221, 110022.	3.1	212
9	Privacy-preserving trajectory data publishing by local suppression. Information Sciences, 2013, 231, 83-97.	4.0	186
10	Asm2Vec: Boosting Static Representation Robustness for Binary Clone Search against Code Obfuscation and Compiler Optimization. , 2019, , .		184
11	Anonymizing sequential releases. , 2006, , .		170
12	Differentially private transit data publication. , 2012, , .		156
13	Correlated network data publication via differential privacy. VLDB Journal, 2014, 23, 653-676.	2.7	151
14	Detecting breaking news rumors of emerging topics in social media. Information Processing and Management, 2020, 57, 102018.	5.4	132
15	Direct Discovery of High Utility Itemsets without Candidate Generation. , 2012, , .		130
16	Publishing set-valued data via differential privacy. Proceedings of the VLDB Endowment, 2011, 4, 1087-1098.	2.1	123
17	A novel methodology for knowledge discovery through mining associations between building operational data. Energy and Buildings, 2012, 47, 430-440.	3.1	117
18	Mining writeprints from anonymous e-mails for forensic investigation. Digital Investigation, 2010, 7, 56-64.	3.2	109

#	ARTICLE	IF	CITATIONS
19	Mining High Utility Patterns in One Phase without Generating Candidates. IEEE Transactions on Knowledge and Data Engineering, 2016, 28, 1245-1257.	4.0	98
20	A Hybrid Framework for Sentiment Analysis Using Genetic Algorithm Based Feature Reduction. IEEE Access, 2019, 7, 14637-14652.	2.6	97
21	Handicapping attacker's confidence: an alternative to k-anonymization. Knowledge and Information Systems, 2007, 11, 345-368.	2.1	96
22	A novel approach of mining write-prints for authorship attribution in e-mail forensics. Digital Investigation, 2008, 5, S42-S51.	3.2	95
23	Advances and challenges in building engineering and data mining applications for energy-efficient communities. Sustainable Cities and Society, 2016, 25, 33-38.	5.1	90
24	A methodology for identifying and improving occupant behavior in residential buildings. Energy, 2011, 36, 6596-6608.	4.5	88
25	A unified data mining solution for authorship analysis in anonymous textual communications. Information Sciences, 2013, 231, 98-112.	4.0	84
26	Centralized and Distributed Anonymization for High-Dimensional Healthcare Data. ACM Transactions on Knowledge Discovery From Data, 2010, 4, 1-33.	2.5	81
27	Anonymity for continuous data publishing. , 2008, , .		80
28	Anonymizing healthcare data. , 2009, , .		74
29	Development of building energy saving advisory: A data mining approach. Energy and Buildings, 2018, 172, 139-151.	3.1	73
30	Extracting knowledge from building-related data " A data mining framework. Building Simulation, 2013, 6, 207-222.	3.0	65
31	Learning Stylometric Representations for Authorship Analysis. IEEE Transactions on Cybernetics, 2019, 49, 107-121.	6.2	61
32	Walking in the crowd. , 2009, , .		57
33	Privacy-preserving data publishing for cluster analysis. Data and Knowledge Engineering, 2009, 68, 552-575.	2.1	53
34	Mining criminal networks from unstructured text documents. Digital Investigation, 2012, 8, 147-160.	3.2	53
35	Secure Two-Party Differentially Private Data Release for Vertically Partitioned Data. IEEE Transactions on Dependable and Secure Computing, 2014, 11, 59-71.	3.7	49
36	Anonymizing trajectory data for passenger flow analysis. Transportation Research Part C: Emerging Technologies, 2014, 39, 63-79.	3.9	47

#	ARTICLE	IF	CITATIONS
37	Anonymity meets game theory: secure data integration with malicious participants. VLDB Journal, 2011, 20, 567-588.	2.7	43
38	Systematic approach to provide building occupants with feedback to reduce energy consumption. Energy, 2020, 194, 116813.	4.5	43
39	BinClone: Detecting Code Clones in Malware. , 2014, , .		42
40	Privacy-preserving trajectory stream publishing. Data and Knowledge Engineering, 2014, 94, 89-109.	2.1	42
41	Service-Oriented Architecture for High-Dimensional Private Data Mashup. IEEE Transactions on Services Computing, 2012, 5, 373-386.	3.2	40
42	Privacy-preserving heterogeneous health data sharing. Journal of the American Medical Informatics Association: JAMIA, 2013, 20, 462-469.	2.2	40
43	e-mail authorship verification for forensic investigation. , 2010, , .		39
44	ϵ -Differential Privacy for Collaborative Data Publishing. IEEE Transactions on Knowledge and Data Engineering, 2014, 26, 2520-2533.	4.0	37
45	Publishing Sensitive Transactions for Itemset Utility. , 2008, , .		36
46	Quantifying the costs and benefits of privacy-preserving health data publishing. Journal of Biomedical Informatics, 2014, 50, 107-121.	2.5	35
47	Development of a ranking procedure for energy performance evaluation of buildings based on occupant behavior. Energy and Buildings, 2019, 183, 659-671.	3.1	35
48	Privacy-preserving data mashup. , 2009, , .		34
49	Product customization of tablet computers based on the information of online reviews by customers. Journal of Intelligent Manufacturing, 2015, 26, 97-110.	4.4	31
50	Kam1n0. , 2016, , .		31
51	Wordnet-Based Criminal Networks Mining for Cybercrime Investigation. IEEE Access, 2019, 7, 22740-22755.	2.6	31
52	SafePath: Differentially-private publishing of passenger trajectories in transportation systems. Computer Networks, 2018, 143, 126-139.	3.2	30
53	Medical Concept Embedding with Multiple Ontological Representations. , 2019, , .		30
54	A Visualizable Evidence-Driven Approach for Authorship Attribution. ACM Transactions on Information and System Security, 2015, 17, 1-30.	4.5	29

#	ARTICLE	IF	CITATIONS
55	A cross-national study of apparel consumer preferences and the role of product-evaluative cues. Asia Pacific Journal of Marketing and Logistics, 2017, 29, 796-812.	1.8	29
56	Young Chinese Consumersâ€™ Choice between Product-Related and Sustainable Cuesâ€™The Effects of Gender Differences and Consumer Innovativeness. Sustainability, 2020, 12, 3818.	1.6	29
57	E-mail authorship attribution using customized associative classification. Digital Investigation, 2015, 14, S116-S126.	3.2	28
58	Malware classification and composition analysis: A survey of recent developments. Journal of Information Security and Applications, 2021, 59, 102828.	1.8	28
59	Frequent grams based embedding for privacy preserving record linkage. , 2012, , .		26
60	Secure Distributed Framework for Achieving ϵ -Differential Privacy. Lecture Notes in Computer Science, 2012, , 120-139.	1.0	23
61	Subject-based semantic document clustering for digital forensic investigations. Data and Knowledge Engineering, 2013, 86, 224-241.	2.1	22
62	Integrating Private Databases for Data Analysis. Lecture Notes in Computer Science, 2005, , 171-182.	1.0	20
63	Privacy protection for RFID data. , 2009, , .		20
64	Scalable code clone search for malware analysis. Digital Investigation, 2015, 15, 46-60.	3.2	19
65	A study of apparel consumer behaviour in China and Taiwan. International Journal of Fashion Design, Technology and Education, 2018, 11, 22-33.	0.9	16
66	Anonymizing data with quasi-sensitive attribute values. , 2010, , .		15
67	Towards discovering criminal communities from textual data. , 2011, , .		15
68	Analyzing topics and authors in chat logs for crime investigation. Knowledge and Information Systems, 2014, 39, 351-381.	2.1	15
69	Investigating the relation between lexical smells and change- and fault-proneness: an empirical study. Software Quality Journal, 2017, 25, 641-670.	1.4	15
70	Revealing determinant factors for early breast cancer recurrence by decision tree. Information Systems Frontiers, 2017, 19, 1233-1241.	4.1	15
71	Preserving RFID data privacy. , 2009, , .		14
72	m-Privacy for Collaborative Data Publishing. , 2011, , .		14

#	ARTICLE	IF	CITATIONS
73	Arabic Authorship Attribution. ACM Transactions on Asian and Low-Resource Language Information Processing, 2019, 18, 1-51.	1.3	14
74	Mining known attack patterns from security-related events. PeerJ Computer Science, 0, 1, e25.	2.7	14
75	Exploring the Meanings of <i>Hijab</i> through Online Comments in Canada. Journal of Intercultural Communication Research, 2016, 45, 214-232.	0.3	13
76	Toward Tweet-Mining Framework for Extracting Terrorist Attack-Related Information and Reporting. IEEE Access, 2021, 9, 115535-115547.	2.6	13
77	A framework for privacy-preserving cluster analysis. , 2008, , .		12
78	Secure Outsourced Frequent Pattern Mining by Fully Homomorphic Encryption. Lecture Notes in Computer Science, 2015, , 70-81.	1.0	12
79	Parallel Eclat for Opportunistic Mining of Frequent Itemsets. Lecture Notes in Computer Science, 2015, , 401-415.	1.0	12
80	Incremental Mining of High Utility Patterns in One Phase by Absence and Legacy-Based Pruning. IEEE Access, 2019, 7, 74168-74180.	2.6	12
81	Learning Phenotypes and Dynamic Patient Representations via RNN Regularized Collective Non-Negative Tensor Factorization. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33, 1246-1253.	3.6	12
82	Towards a Knowledge-Based Recommender System for Linking Electronic Patient Records With Continuing Medical Education Information at the Point of Care. IEEE Access, 2019, 7, 15955-15966.	2.6	12
83	Service-Oriented Architecture for Privacy-Preserving Data Mashup. , 2009, , .		11
84	Mining Criminal Networks from Chat Log. , 2012, , .		11
85	Privacy-preserving data mashup model for trading person-specific information. Electronic Commerce Research and Applications, 2016, 17, 19-37.	2.5	11
86	Opportunistic mining of top-n high utility patterns. Information Sciences, 2018, 441, 171-186.	4.0	11
87	Improving interpretations of topic modeling in microblogs. Journal of the Association for Information Science and Technology, 2018, 69, 528-540.	1.5	11
88	Heterogeneous data release for cluster analysis with differential privacy. Knowledge-Based Systems, 2020, 201-202, 106047.	4.0	11
89	Season-Based Occupancy Prediction in Residential Buildings Using Machine Learning Models. E-Prime, 2021, 1, 100003.	2.1	11
90	Trade-off between accuracy and fairness of data-driven building and indoor environment models: A comparative study of pre-processing methods. Energy, 2022, 239, 122273.	4.5	11

#	ARTICLE	IF	CITATIONS
91	I-MAD: Interpretable malware detector using Galaxy Transformer. Computers and Security, 2021, 108, 102371.	4.0	10
92	D-Mash: A Framework for Privacy-Preserving Data-as-a-Service Mashups. , 2014, , .		9
93	Preserving privacy and frequent sharing patterns for social network data publishing. , 2013, , .		8
94	Differentially private multidimensional data publishing. Knowledge and Information Systems, 2018, 56, 717-752.	2.1	8
95	Differentially private data publishing for arbitrarily partitioned data. Information Sciences, 2021, 553, 247-265.	4.0	8
96	\$\$\$extsf {SecDM}}\$\$\$: privacy-preserving data outsourcing framework with differential privacy. Knowledge and Information Systems, 2020, 62, 1923-1960.	2.1	7
97	An Advanced Noise Reduction and Edge Enhancement Algorithm. Sensors, 2021, 21, 5391.	2.1	7
98	An Efficient Hybrid Hierarchical Document Clustering Method. , 2008, , .		6
99	Using RFID tags to improve pilgrimage management. , 2009, , .		6
100	A cross-national study of young female consumer behaviour, innovativeness and apparel evaluation: China and India. Journal of the Textile Institute, 2020, 111, 334-344.	1.0	6
101	Enabling Secure Trustworthiness Assessment and Privacy Protection in Integrating Data for Trading Person-Specific Information. IEEE Transactions on Engineering Management, 2021, 68, 149-169.	2.4	6
102	Using data mining to analyse fashion consumersâ€™ preferences from a cross-national perspective. International Journal of Fashion Design, Technology and Education, 2014, 7, 42-49.	0.9	5
103	Fusion: Privacy-Preserving Distributed Protocol for High-Dimensional Data Mashup. , 2015, , .		5
104	Distinguishing between fake news and satire with transformers. Expert Systems With Applications, 2022, 187, 115824.	4.4	5
105	Joint Learning of Phenotypes and Diagnosis-Medication Correspondence via Hidden Interaction Tensor Factorization. , 2018, , .		5
106	Automatic Fall Risk Detection Based on Imbalanced Data. IEEE Access, 2021, 9, 163594-163611.	2.6	5
107	Factors influencing consumer choice: a study of apparel and sustainable cues from Canadian and Indian consumersâ€™ perspectives. International Journal of Fashion Design, Technology and Education, 2021, 14, 151-161.	0.9	4
108	Detecting High-Engaging Breaking News Rumors in Social Media. ACM Transactions on Management Information Systems, 2021, 12, 1-16.	2.1	4

#	ARTICLE	IF	CITATIONS
109	Anonymizing location-based RFID data. , 2009, , .		3
110	Service-oriented architecture for sharing private spatial-temporal data. , 2011, , .		2
111	Verification of data pattern for interactive privacy preservation model. , 2011, , .		2
112	Anonymizing Social Network Data for Maximal Frequent-Sharing Pattern Mining. Lecture Notes in Social Networks, 2015, , 77-100.	0.8	2
113	Preface to the Fourth IEEE Workshop on Privacy Aspects of Data Mining. , 2013, , .		0
114	Privacy-Preserving Medical Reports Publishing for Cluster Analysis. , 2014, , .		0
115	A Data Mining Analysis of Cross-Regional Study of Apparel Consumption. Lecture Notes in Networks and Systems, 2022, , 23-32.	0.5	0
116	Authorship Attribution With Few Training Samples. International Series on Computer Entertainment and Media Technology, 2020, , 75-87.	0.7	0
117	Writeprint Mining For Authorship Attribution. International Series on Computer Entertainment and Media Technology, 2020, , 57-74.	0.7	0
118	Messaging Forensics In Perspective. International Series on Computer Entertainment and Media Technology, 2020, , 23-36.	0.7	0