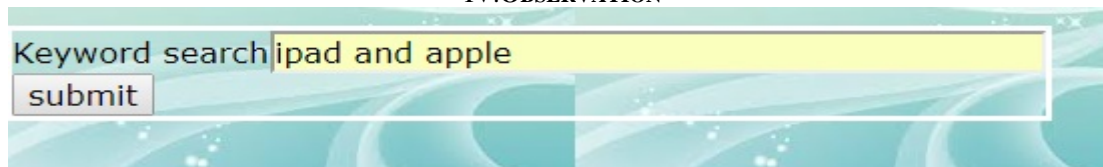


IV.OBSERVATION



The query string is:ipad and apple

ipad and apple Inc, ipad and apple technology, ipad with apple inc., usb, touchpad, ipad, laptop, microsoft, apple, intel, mango, apple, grape,

The conceptualization process

The given terms are:ipad,apple

The term ipadconcepts

device:1 company:0 fruit:0 Selected concept for ipad: device

The term appleconcepts

device:0 company:1 fruit:1 Selected concept for apple: company,fruit

Co-occurrence statistics for apple

The Search results ipad and apple Inc, ipad and apple technology, ipad with apple inc.,

V.CONCLUSIONS

We have carried out a comprehensive study of techniques, algorithms and frameworks used for intend based diversification, data pre- processing, outlier detection, data summarization and fuzzy keyword set creation. The modern hybrid diversification frameworks are able to take intelligent decisions at very earlier stage to reduce the unwanted and redundant search task over various data forms. Automatic text classification algorithms have been introduced for data pre-processing, CONCEPT creation and conceptualization. This approach has paved the way for developing mathematical and statistical model for feature construction and subset creation. The study redefines "outlier" as a bound for the keyword set in the high dimensional data space. This enlightened the need of further coding for developing high quality clusters. The survey lists out some summarization frame works, and studied the efficient use of historical summaries and time lines for facets and nuggets creation. Along with the state-of-art methods such as edit distance, wild card mapping, LS joins, gram based techniques. The Fuzzy bag-of-words clustering and feature vector classification for pattern matching and fuzzy keyword set creation provide more relevancies to the search.

REFERENCES

- [1] M. Hasan, A. Mueen, V. Tsotras, and E. Keogh, "Diversifying Query Results on Semi-structured Data," Proceedings of the 21st ACM International Conference on Information and Knowledge Management, pp. 2099–2103, 2012.
- [2] M. R. Vieira, H. L. Razente, M. C. Barioni, M. Hadjieleftheriou, D. Srivastava, C. Traina, and V. J. Tsotras, "On Query Result Diversification," Data Engineering (ICDE), IEEE , pp. 1163–1174, 2011.
- [3] E. Vee, U. Srivastava, J. Shanmugasundaram, P. Bhat, and S. A. Yahia, "Efficient Computation of Diverse Query Results," Data Engineering, ICDE , pp. 228–236, 2008.
- [4] M. Drosou and E. Pitoura, "Diversity over Continuous Data," IEEE Data Eng. Bull., vol. 32, no. 4, pp. 49–56, 2009.
- [5] J. Li, C. Liu, and J. X. Yu, "Context-based Diversification for Keyword Queries over XML Data," IEEE Transactions on Knowledge and Data Engineering, vol. 27, no. 3, pp. 660–672, 2015.
- [6] E. J. Ruiz, V. Hristidis, and P. G. Ipeirotis, "Facilitating Document Annotation using Content and Querying Value," IEEE Transactions on Knowledge and Data Engineering, vol. 26, no. 2, pp. 336–349, 2014.
- [7] N. Sarkas, N. Bansal, G. Das, and N. Koudas, "Measure-driven Keyword Query Expansion," Proceedings of the VLDB Endowment, vol. 2, no. 1, pp. 121–132, 2009.
- [8] A. Angel and N. Koudas, "Efficient Diversity Aware Search," Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data, pp. 781–792, 2011.
- [9] R. L. Santos, C. Macdonald, and I. Ounis, "Exploiting Query Reformulations for Web Search Result Diversification," Proceedings of the 19th International Conference on World Wide Web, pp. 881–890, 2010.
- [10] E. Demidova, P. Fankhauser, X. Zhou, and W. Nejdl, "DivQ: Diversification for Keyword Search over Structure Databases," Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 331–338, 2010.
- [11] R. L. Santos, J. Peng, C. Macdonald, and I. Ounis, "Explicit Search Result Diversification through Sub-queries," ECIR, pp. 87–99, 2010.
- [12] D. Panigrahi, A. Das Sarma, G. Aggarwal, and A. Tomkins, "Online Selection of Diverse Results," Proceedings of the Fifth ACM International Conference on Web Search and Data Mining, pp. 263–272, 2012.
- [13] H. Chen and D. R. Karger, "Less is More: Probabilistic Models for Retrieving fewer Relevant Documents," Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 429–436, 2006.
- [14] C. Sun, C.-Y. Chan, and A. K. Goenka, "Multiway SLCA-Based Keyword Search in XML Data," Proceedings of the 16th International Conference on World Wide Web, pp. 1043–1052, 2007.
- [15] J. Zhou, W. Wang, Z. Chen, J. X. Yu, X. Tang, Y. Lu, and Y. Li, "Top-Down XML Keyword Query Processing," IEEE Transactions on Knowledge and Data Engineering, vol. 28, no. 5, pp. 1340–1353, 2016.

- [16] H. Wu and Z. Tang, "An Efficient Algorithm for Meaningful SLCA in XML Keyword Search," *Web Information Systems and Mining*, pp. 280–283, 2009.
- [17] Y. Xu and Y. Papakonstantinou, "Efficient Keyword Search for Small est LCAs in XML Databases," *Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data*, pp. 527–538, 2005.
- [18] J. Li, C. Liu, R. Zhou, and B. Ning, "Processing xml Keyword Search by Constructing Effective Structured Queries." *Apweb/waim*, vol. 5446, pp. 88–99, 2009.
- [19] Y. Li, C. Yu, and H. Jagadish, "Schema Free Xquery," *Proceedings of the Thirtieth International Conference on Very Large Data Bases*, Vol. 30, pp. 72–83, 2004.
- [20] J. Li, C. Liu, R. Zhou, and W. Wang, "Top-k Keyword Search over Probabilistic XML Data," *Data Engineering (ICDE)*, pp. 673–684, 2011.
- [21] I. J. Chiang, C. C. H. Liu, Y. H. Tsai, and A. Kumar, "Discovering Latent Semantics in Web Documents using Fuzzy Clustering," *IEEE Transactions on Fuzzy Systems*, vol. 23, no. 6, pp. 2122–2134, 2015.
- [22] G. Bhalotia, A. Hulgeri, C. Nakhe, S. Chakrabarti, and S. Sudarshan, "Keyword Searching and Browsing in Databases using BANKS," *Data Engineering*, pp. 431–440, 2002.
- [23] D. Kim, H. Wang, and A. H. Oh, "Context-Dependent Conceptualization," *IJCAI*, pp. 2654–2661, 2013.
- [24] Y. Song, H. Wang, Z. Wang, H. Li, and W. Chen, "Short Text Conceptualization using a Probabilistic Knowledgebase," *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, vol. 3, pp. 2330–2336, 2011.
- [25] W. Wu, H. Li, H. Wang, and K. Q. Zhu, "Probase: A Probabilistic Taxonomy for Text Understanding," *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, pp. 481–492, 2012.
- [26] Q. Chen, M. Zhang, and B. Xue, "Feature Selection to Improve Generalization of Genetic Programming for High-Dimensional Symbolic Regression ," *IEEE Transactions on Evolutionary Computation*, 2017.
- [27] K. Nag and N. R. Pal, "A Multiobjective Genetic Programming-based Ensemble for Simultaneous Feature Selection and Classification," *IEEE Transactions on Cybernetics*, vol. 46, no. 2, pp. 499–510, 2016.
- [28] D. P. Muni, N. R. Pal, and J. Das, "Genetic Programming for Simultaneous Feature Selection and Classifier Design," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 36, no. 1, pp. 106–117, 2006.
- [29] M. Dash and H. Liu, "Consistency-based Search in Feature Selection," *Artificial Intelligence*, vol. 151, no. 1-2, pp. 155–176, 2003.
- [30] M. K. Dalal and M. A. Zaveri, "Automatic Text Classification of Sports Blog Data," *Computing, Communications and Applications Conference (ComComAp)*, 2012, pp. 219–222, 2012.
- [31] B. Xue, M. Zhang, W. N. Browne, and X. Yao, "A Survey on Evolutionary Computation Approaches to Feature Selection," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 4, pp. 606–626, 2016.
- [32] F. Liu, F. Liu, and Y. Liu, "A Supervised Framework for Keyword Extraction from Meeting Transcripts," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 3, pp. 538–548, 2011.
- [33] C. Hou, F. Nie, H. Tao, and D. Yi, "Multi-view Unsupervised Feature Selection with Adaptive Similarity and View Weight," *IEEE Transactions on Knowledge and Data Engineering*, 2017.
- [34] V. Hautamki, I. Krinen, and P. Franti, "Outlier Detection Using k-Nearest Neighbour Graph," *Proceedings of the 17th International Conference on Pattern Recognition, ICPR* , vol. 3, pp. 430–433, 2004.
- [35] M. Radovanovic, A. Nanopoulos, and M. Ivanovic, "Reverse Nearest Neighbors in Unsupervised Distance-based Outlier Detection," *vol. 27, no. 5*, pp. 1369–1382, 2015.
- [36] T. Matsumoto and E. Hung, "Accelerating Outlier Detection with Uncertain Data using Graphics Processors," *Advances in Knowledge Discovery and Data Mining*, pp. 169–180, 2012.
- [37] X. Yang, B. Wang, and C. Li, "Cost-based Variable Length Gram Selection for String Collections to Support Approximate Queries Efficiently," *Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data*, pp. 353–364, 2008.
- [38] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy Keyword Search over Encrypted Data in Cloud Computing," *INFOCOM, 2010 Proceedings IEEE*, pp. 1–5, 2010.
- [39] K. Deng, X. Li, J. Lu, and X. Zhou, "Best Keyword Cover Search," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 1, pp. 61–73, 2015.
- [40] T. Barbu, "An Automatic Unsupervised Pattern Recognition Approach," *Proceedings of the Romanian Academy*, vol. 7, no. 1, pp. 73–78, 2006.
- [41] G. Chatzimilioudis, C. Costa, a. L. W. C. Zeinalipour Yazdi, Demetrios, and E. Pitoura, "Distributed In-Memory Processing of All k Nearest Neighbor Queries," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 4, pp. 925–938, 2016.
- [42] J. Carbonell and J. Goldstein, "The use of MMR, Diversity-based Reranking for Reordering Documents and Producing Summaries," *Proceedings of the 21st annual international ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 335–336, 1998.
- [43] C. X. Zhai, W. W. Cohen, and J. Lafferty, "Beyond Independent Relevance: Methods and Evaluation Metrics for Subtopic Retrieval," *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, pp. 10–17, 2003.
- [44] J. Wang and J. Zhu, "Portfolio Theory of Information Retrieval," *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 115–122, 2009.
- [45] R. Agrawal, S. Gollapudi, A. Halverson, and S. Jeong, "Diversifying Search Results," *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, pp. 5–14, 2009.
- [46] F. Radlinski and S. Dumais, "Improving Personalized Web Search using Result Diversification," *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 691–692, 2006.
- [47] J. Liu and D. Yan, "Answering Approximate Queries over XML Data," *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 2, pp. 288–305, 2016.
- [48] S. Qi, D. Wu, and N. Mamoulis, "Location Aware Keyword Query Suggestion based on Document Proximity," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 1, pp. 82–97, 2016.
- [49] D. Sathiaraj and E. Triantaphyllou, "On Identifying Critical Nuggets of Information during Classification Tasks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 6, pp. 1354–1367, 2013.
- [50] Z. Liu, P. Sun, and Y. Chen, "Structured Search Result Differentiation," *Proceedings of the VLDB Endowment*, vol. 2, no. 1, pp. 313–324, 2009.
- [51] Y. Chen, W. Wang, Z. Liu, and X. Lin, "Keyword Search on Structured and Semi Structured Data," *Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data*, pp. 1005–1010, 2009.
- [52] Z. Yu, H. Wang, X. Lin, and M. Wang, "Understanding Short Texts through Semantic Enrichment and Hashing," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 2, pp. 566–579, 2016.
- [53] C. L. Clarke, M. Kolla, G. V. Cormack, O. Vechtomova, A. Ashkan, S. B'utcher, and I. MacKinnon, "Novelty and Diversity in Information Retrieval Evaluation," *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 659–666, 2008.
- [54] A. Angel and N. Koudas, "Efficient Diversity Aware Search," *Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data*, pp. 781–792, 2011.
- [55] A. Halevy, M. Franklin, and D. Maier, "Principles of Data Space Systems," *Proceedings of the twenty-fifth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, pp. 1–9, 2006.
- [56] S. Gollapudi and A. Sharma, "An Axiomatic Approach for Result Diversification," *Proceedings of the 18th International Conference on World Wide Web*, pp. 381–390, 2009.
- [57] C. Xu, D. Tao, and C. Xu, "A survey on Multi-view Learning," *ArXiv Preprint ArXiv:1304.5634*, 2013.
- [58] W. Li and C. X. Chen, "Efficient Data Modeling and Querying System for Multi-dimensional Spatial Data," *Proceedings of the 16th*

- ACM SIGSPATIAL International Conference on Advances in Geographic Systems, 2008.
- [59] S. Murthy, CA, "Bridging Feature Selection and Extraction: Compound Feature Generation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 4, pp. 757–770, 2017.
- [60] Z.-Q. Wang, X. Sun, D.-X. Zhang, and X. Li, "An Optimal SVM-based Text Classification Algorithm," *Machine Learning and Cybernetics*, 2006 International Conference on, pp. 1378–1381, 2006.
- [61] V. Berisha, A. Javadi, K. R. Hammet, D. V. Anderson, and A. Gray, "Making Decisions about Unseen data: Semi-supervised Learning at Different Levels of Specificity," *Signals, Systems and Computers (ASILOMAR)*, pp. 75–79, 2010.
- [62] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data Clustering: A Review," *ACM Computing Surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999.
- [63] T. Kurita, "An Efficient Agglomerative Clustering Algorithm for Region Growing," *IAPR Workshop on Machine Vision*, 1994.
- [64] W. Hua, Z. Wang, H. Wang, K. Zheng, and X. Zhou, "Understand Short Texts by Harvesting and Analyzing Semantic Knowledge," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, pp. 499–512, 2017.
- [65] Y. Kim and S. Lee, "SVM-based Web Content Mining with Leaf Classification Unit from DOM-tree," *Knowledge and Smart Technology*, pp. 359–364, 2017.
- [66] S. Debnath, P. Mitra, N. Pal, and C. L. Giles, "Automatic Identification of Informative Sections of Web Pages," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 9, pp. 1233–1246, 2005.
- [67] Y.-T. Wen, J. Yeo, W.-C. Peng, and S.-W. Hwang, "Efficient Keyword Aware Representative Travel Route Recommendation," *IEEE Transactions on Knowledge and Data Engineering*, 2017.
- [68] D. Karaboga and C. Ozturk, "A Novel Clustering Approach: Artificial Bee Colony (ABC) Algorithm," *Applied Soft Computing*, vol. 11, no. 1, pp. 652–657, 2011.
- [69] C.-Y. Chen, S.-C. Hwang, and Y.-J. Oyang, "An Incremental Hierarchical Data Clustering Algorithm Based on Gravity Theory," *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 237–250, 2002.
- [70] H.-P. Kriegel and M. Pfeifle, "Density-based Clustering of Uncertain Data," *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*, pp. 672–677, 2005.
- [71] P. J. Rousseeuw and A. M. Leroy, "Robust Regression and Outlier Detection," *Wiley Inter Science Paper Back Series*, vol. 589, 2005.
- [72] J. Lin, D. Etter, and D. DeBarr, "Exact and Approximate Reverse Nearest Neighbor Search for Multimedia Data," *Proceedings of the 2008 SIAM International Conference on Data Mining*, pp. 656–667, 2008.
- [73] F. Korn and S. Muthukrishnan, "Influence Sets based on Reverse Nearest Neighbor Queries," *ACM Sigmod Record*, vol. 29, no. 2, pp. 201–212, 2000.
- [74] F. Angiulli, S. Basta, S. Lodi, and C. Sartori, "GPU Strategies for Distance-Based Outlier Detection," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 11, pp. 3256–3268, 2016.
- [75] C. Bohm, R. Noll, C. Plant, and B. Wackersreuther, "Density-based Clustering Using Graphics Processors," *Proceedings of the 18th ACM Conference on Information and Knowledge Management*, pp. 661–670, 2009.
- [76] R. Wu, B. Zhang, and M. Hsu, "Clustering Billions of Data Points Using GPUs," *Proceedings of the Combined Workshops on Un Conventional High Performance Computing Workshop Plus Memory Access Workshop*, pp. 1–6, 2009.
- [77] M. Alshawabkeh, B. Jang, and D. Kaeli, "Accelerating the Local Outlier Factor Algorithm on a GPU for Intrusion Detection Systems," *Proceedings of the 3rd Workshop on General-Purpose Computation on Graphics Processing Units*, pp. 104–110, 2010.
- [78] F. Azmandian, A. Yilmazer, J. G. Dy, J. A. Aslam, and D. R. Kaeli, "GPU-Accelerated Feature Selection for Outlier Detection using the Local Kernel Density Ratio," *Data Mining (ICDM)*, pp. 51–60, 2012.
- [79] X. Liu, C. Deng, B. Lang, D. Tao, and X. Li, "Query-adaptive Reciprocal Hash Tables for Nearest Neighbor Search," *IEEE Transactions on Image Processing*, vol. 25, no. 2, pp. 907–919, 2016.
- [80] V. Singh, B. Zong, and A. K. Singh, "Nearest Keyword Set Search in Multi-Dimensional Datasets," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 3, pp. 741–755, 2016.
- [81] R. Feldbauer and A. Flexer, "Centering Versus Scaling for Hubness Reduction," *International Conference on Artificial Neural Networks*, pp. 175–183, 2016.
- [82] M. Ernst and G. Haesbroeck, "Comparison of Local Outlier Detection Techniques in Spatial Multivariate Data," *Data Mining and Knowledge Discovery*, vol. 31, no. 2, pp. 371–399, 2017.
- [83] E. Schubert, A. Zimek, and H.-P. Kriegel, "Local Outlier Detection Reconsidered: A Generalized View on Locality with Applications to Spatial, Video, and Network Outlier Detection," *Data Mining and Knowledge Discovery*, vol. 28, no. 1, pp. 190–237, 2014.
- [84] C. C. Aggarwal, A. Hinneburg, and D. A. Keim, "On the Surprising Behavior of Distance Metrics in High Dimensional Space," *ICDT*, pp. 420–434, 2001.
- [85] D. Navile and G. Ravikumar, "Outlier Detection in High Dimension Data Based On Multimodality And Neighbourhood Size Using KNN Method," *International Journal for Engineering and Computer Science*, vol. 5, 2016.
- [86] Z. Dou, Z. Jiang, S. Hu, J.-R. Wen, and R. Song, "Automatically Mining Facets for Queries from their Search Results," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 2, pp. 385–397, 2016.
- [87] C. Li, A. Sun, J. Weng, and Q. He, "Tweet Segmentation and Its Application to Named Entity Recognition," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 2, pp. 558–570, 2015.
- [88] Z. Wang, L. Shou, K. Chen, G. Chen, and S. Mehrotra, "On Summarization and Timeline Generation for Evolutionary Tweet Streams," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 5, pp. 1301–1315, 2015.
- [89] T. Takahashi, R. Tomioka, and K. Yamanishi, "Discovering Emerging Topics in Social Streams via Link-Anomaly Detection," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 1, pp. 120–130, 2014.
- [90] S. Yan and X. Wan, "SRRank: Leveraging Semantic Roles for Extractive Multi-document Summarization," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 12, pp. 2048–2058, 2014.
- [91] F. M. F. Wong, C. W. Tan, S. Sen, and M. Chiang, "Quantifying Political Learning from Tweets and Retweets," *ICWSM*, vol. 13, pp. 640–649, 2013.
- [92] J. Zhang, V. S. Sheng, J. Wu, and X. Wu, "Multi-Class Ground Truth Inference in Crowd sourcing with Clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 4, pp. 1080–1085, 2016.
- [93] D. Vandić, S. Aanen, F. Frasincar, and U. Kaymak, "Dynamic Facet Ordering for Faceted Product Search Engines," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 5, pp. 1004–1016, 2017.
- [94] L. Breiman, "Some Properties of Splitting Criteria," *Machine Learning*, vol. 24, no. 1, pp. 41–47, 1996.
- [95] C.-D. Tan, F. Min, M. Wang, H.-R. Zhang, and Z.-H. Zhang, "Discovering Patterns with Weak-Wildcard Gaps," *IEEE Access*, vol. 4, pp. 4922–4932, 2016.
- [96] R. Zhao and K. Mao, "Fuzzy Bag-of-Words Model for Document Representation," *IEEE Transactions on Fuzzy Systems*, 2017.
- [97] J. Wang, X. Yang, B. Wang, and C. Liu, "Ls-join: Local Similarity Join on String Collections," *IEEE Transactions on Knowledge and Data Engineering*, 2017.
- [98] B. Gu, Z. Li, X. Zhang, A. Liu, G. Liu, K. Zheng, L. Zhao, and X. Zhou, "The Interaction Between Schema Matching and Record Matching in Data Integration," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 1, pp. 186–199, 2017.
- [99] Y. Zheng, Z. Bao, L. Shou, and A. K. Tung, "INSPIRE: A framework for Incremental Spatial Prefix Query Relaxation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 7, pp. 1949–1963, 2015.
- [100] W.-Y. Lin, F. Wang, M.-M. Cheng, S.-K. Yeung, P. H. Torr, M. N. Do, and J. Lu, "Code: Coherence based Decision Boundaries for Feature Correspondence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.

- [101] S. Ji, G. Li, C. Li, and J. Feng, "Efficient Interactive Fuzzy Keyword Search," Proceedings of the 18th International Conference on World Wide Web, pp. 371–380, 2009.
- [102] K. Venugopal and R. Buyya, "Mastering C++", Tata McGraw-Hill Education, 2013.
- [103] V. Levenshtein, "Binary Codes Capable of Correcting Spurious Insertions and Deletions of Ones," Problems of information Transmission, vol. 1, no. 1, pp. 8–17, 1965.
- [104] L. Zhao, T. Lin, K. Zhou, S. Wang, and X. Chen, "Pseudo 2D String Matching Technique for High Efficiency Screen Content Coding," IEEE Transactions on Knowledge and Data Engineering, vol. 18, no. 3, pp. 339–350, 2016.
- [105] C. Li, B. Wang, and X. Yang, "VGRAM: Improving Performance of Approximate Queries on String Collections using Variable-length Grams," Proceedings of the 33rd International Conference on Very Large Data Bases, pp. 303–314, 2007.
- [106] F. M. Anuar, R. Setchi, and Y.-K. Lai, "Semantic Retrieval of Trademarks Based on Conceptual Similarity," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 46, no. 2, pp. 220–233, 2016.
- [107] A. Bartoli, A. De Lorenzo, E. Medvet, and F. Tarlao, "Inference of Regular Expressions for Text Extraction from Examples," IEEE Transactions on Knowledge and Data Engineering, vol. 28, no. 5, pp. 1217–1230, 2016.
- [108] Y. Mitani, F. Ino, and K. Hagihara, "Parallelizing Exact and Approximate String Matching via Inclusive Scan on a GPU," IEEE Transactions on Parallel and Distributed Systems, vol. 28, no. 7, pp. 1989–2002, 2017.
- [109] C.-H. Lin, J.-C. Li, C.-H. Liu, and S.-C. Chang, "Perfect Hashing Based Parallel Algorithms for Multiple String Matching on Graphic Processing Units," IEEE Transactions on Parallel and Distributed Systems, 2017.