

A INTEGRAÇÃO DO MODELO PROBABILÍSTICO DE TÓPICOS COM O MÉTODO MULTICRITÉRIO: A ESCOLHA DE UMA ESTRATÉGIA DE POLICIAMENTO EM FUNÇÃO DA DEMANDA CRIMINAL LOCAL

Marcio Pereira Basilio

Departamento de Engenharia de Produção – UFF
Rua Passo da Pátria, 156 Bl. D Sala 306
marciopbasilio@gmail.com

Valdecy Pereira

Departamento de Engenharia de Produção – UFF
Rua Passo da Pátria, 156 Bl. D Sala 306
valdecy.pereira@gmail.com

Resumo

A pesquisa objetivou desenvolver uma metodologia para descoberta de conhecimento em banco de dados dos serviços de atendimento de emergência, com base nos relatos das ocorrências policiais atendidas, com a finalidade de gerar informação para subsidiar os órgãos encarregados de cumprir a lei no planejamento das ações de investigação e combate de ações criminais. O modelo desenvolvido divide-se em duas partes. A primeira constitui a metodologia de descoberta de conhecimento com a utilização de técnicas de mineração de texto, através da utilização da técnica LDA (Latent Dirichlet Allocation) para obtenção de tópicos sobre a criminalidade. A segunda parte corresponde técnica multicritério ELECTRE IV que com os resultados dos tópicos sobre a criminalidade criará um ranking das estratégias de combate ao crime nas localidades estudadas. A avaliação e validação dos tópicos foi feita por especialistas da área. Como resultados pode-se relatar que o método utilizado permitiu a identificação dos delitos mais comuns ocorridos no período de 01 de janeiro a 31 de dezembro de 2016, nas áreas estudadas. A análise dos tópicos identificados permitiu reafirmar que os crimes não ocorrem de forma linear em uma determinada localidade, no presente estudo 40% dos delitos identificados na Área Integrada de Segurança Pública nº 5 (AISP 5 – Região do Centro da Cidade do Rio de Janeiro), não possuíam correspondência com a Área Integrada de Segurança Pública nº 19 (AISP 19 – Bairro de Copacabana), bem como, 33% dos delitos da AISP 19 não foram identificados na AISP 5. Como limitação pode-se registrar que os dados coletados representam a dinâmica social dos bairros do centro e da zona sul da cidade do Rio de Janeiro no período específico de janeiro de 2013 a dezembro de 2016. O que implica dizer que os resultados não podem ser generalizados para áreas com características diferentes. A metodologia desenvolvida contribui de forma complementar na identificação de práticas delituosas e suas características a partir dos relatos das ocorrências policiais arquivadas nos bancos de dados dos serviços de emergências. O conhecimento gerado permite aos especialistas dos órgãos encarregados de fazer cumprir a lei avaliar, reformular e construir estratégias diferenciadas para o combate de crimes em determinada localidade. Como implicações sociais pode-se inferir que com a escolha das estratégias adequadas ao combate da criminalidade local, o modelo proposto proporcionará um

aumento da sensação de segurança por meio da redução efetiva dos delitos. A originalidade da pesquisa consiste na integração de técnicas de mineração de textos e de multicritério.

Palavras-chaves: Mineração de texto; Latent Dirichlet Allocation, Multicritério

Referência Bibliográfica

- Agnew, R. (2016). A theory of crime resistance and susceptibility. *Criminology*, 54(2), pp. 181-211. doi: 10.1111/1745-9125.12104.
- Akilan, A. (2015). Text mining: Challenges and future directions. *2nd International Conference on Electronics and Communication Systems, ICECS 2015* (pp. 1679-1683). Kumbakonam: Institute of Electrical and Electronics Engineers Inc.
- Aliguliyev, R. M. (2009). A new sentence similarity measure and sentence based extractive technique for automatic text summarization. *Expert Systems with Applications*, 36(4), pp. 7764-7772. doi:10.1016/j.eswa.2008.11.022.
- Alwidian , S., Bani-Salameh, H., & Alslaity, A. (2015). Text data mining: A proposed framework and future perspectives. *International Journal of Business Information Systems*, 18(2), pp. 127-140.
- Araújo Júnior, R., & Tarapanoff, K. (2006). Precision rate in the information retrieval process: the use of text mining. *Ciência da Informação*, 35(3), pp. 236-247. <https://dx.doi.org/10.1590/S0100-19652006000300023>.
- Ball, L. (2016, April). Automating social network analysis: A power tool for counter-terrorism. *Security Journal*, 29(2), pp. 147–168. <https://doi.org/10.1057/sj.2013.3>.
- Banu, G. R., & Chitra, V. (2015). A Survey of Text Mining Concepts. *International Journal of innovations in Engineering and Technology*, 5(2), pp. 121-127.
- Basilio, M. P., Pereira, V., & Costa, H. G. (2017). Review of the literature on multicriteria methods applied in the field of public security. *Universal Journal of Management*, 5(12), pp. 549-562. DOI:10.13189/ujm.2017.051202.
- Bayley, D. H. (1994). *Police for the future*. New York: Oxford University Press.
- Bezerra, C. A., & Guimarães, A. J. (2014, abr/jun). Mineração de texto aplicada às publicações científicas sobre gestão do conhecimento no período de 2003 a 2012. *Perspectivas em Ciência da Informação*, 19(2), pp. 131-146.
- Bisogno, E., Dawson-Faber, J., & Jandl, M. (2015). The International Classification of Crime for Statistical Purposes: A new instrument to improve comparative criminological research. *European Journal of Criminology*, 12(5), pp. 535-550. DOI:10.1177/1477370815600609.

- Blei, D. M. (2011). *Introduction to probabilistic topic models*. Communications of the ACM.
- Blei, D. M. (2012, April). Probabilistic Topic Models. *communications of the acm*, 55(4), pp. 77-84. doi:10.1145/2133806.2133826.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, pp. 993-1022.
- Bouranta, N., Siskos, Y., & Tsotsolas, N. (2015). Measuring police officer and citizen satisfaction: comparative analysis. *Policing*, 38(4), pp. 705-721. <https://dx.doi.org/10.1108/PIJPSM-01-2015-0008>.
- Braga, A. A. (2001). The Effects of Hot Spots Policing on Crime. *The Annals of the American Academy of Political and Social Science*, 578(1), pp. 104-125.
- Braga, A. A. (2005). Hot spots policing and crime prevention: A systematic review of randomized controlled trials. *Journal of Experimental Criminology*(1), pp. 317–342.
- Braga, A. A., Weisburd, D. L., Waring, E. J., Mazerolle, L. A., Spelman, W., & Gajewski, F. (1999). Problem-oriented policing in violent crime places: A randomized controlled experiment. *Criminology*, 37(3), pp. 541-579.
- Brown, D. E. (2016). Text Mining the Contributors to Rail Accidents. *IEEE Intelligent Transportation Systems Society*, 17(2), pp. 346 - 355. 10.1109/TITS.2015.2472580.
- Calvo, H., Godoy-Calderón, S., Moreno-Armendáriz, M., & Martínez-Hernández, V. M. (2017). Forecasting, clustering and patrolling criminal activities. *Intelligent Data Analysis*, 21(3), pp. 697-720. DOI:10.3233/IDA-170883.
- Camacho-Collados, M., Liberatore, F., & Angulo, J. M. (2015). A multi-criteria Police Districting Problem for the efficient and effective design of patrol sector. *European Journal of Operational Research*, 246(2), pp. 674-684. <https://dx.doi.org/10.1016/j.ejor.2015.05.023>.
- Capuano, E. A. (2009). The cognitive power of artificial neural networks model ART1 for information retrieval. *Ciência da Informação*, 38(1), pp. 9-30. <https://dx.doi.org/10.1590/S0100-19652009000100001>.
- Chen, Y.-L., Liu, Y.-H., & Ho, W.-L. (2013, February). A text mining approach to assist the general public in the retrieval of legal documents. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY*, 64(2), pp. 280–290. DOI:10.1002/asi.22767.
- Chen, Z., Huang, Y., Tian, J., Liu, X., Fu, K., & Huang, T. (2015). Joint model for subsentence-level sentiment analysis with Markov logic. *Journal of the Association for Information Science and Technology*, 66(9), pp. 1913-1922.
- Correia, A., & Gonçalves, A. (2017). Topics discovery in text mining. *5th World Conference on Information Systems and Technologies, WorldCIST*. 569, pp. 251-256. Porto Santo Island, Madeira: Springer Verlag.

- Cunha, C. (2001). *Nova gramática do português contemporâneo* (3. ed.). Rio de Janeiro: Nova Fronteira.
- Davis, T. A., & Hu, Y. (2011). The University of Florida Sparse Matrix Collection. *ACM Transactions on Mathematical Software*, 38(1).
- Day, M. Y., & Lee, C. (2016). Deep learning for financial sentiment analysis on finance news providers. *Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (pp. 1127-1134). San Francisco: IEEE.
- De Paula Silva, C. J., Moura, A. C., Paiva, P. C., Ferreira, R. C., Silvestrini, R. A., Vargas, A. M., . . . Ferreira E Ferreira, E. (2015). Maxillofacial injuries as markers of interpersonal violence in belo Horizonte-Brazil: Analysis of the socio-spatial vulnerability of the location of victim's residences. *PLoS ONE*, 10(8), p. Art. n. e0134577. <https://dx.doi.org/10.1371/journal.pone.0134>.
- Dey, L., Sameera Bharadwaja, H., Meera, G., & Shroff, G. (2013). Email analytics for activity management and insight discovery. *Proceedings - 2013 IEEE/WIC/ACM International Conference on Web Intelligence*. 1, pp. 557-564. Atlanta: IEEE.
- Dijkstra, L. J., Yakushev, A. V., Duijn, P. A., Boukhanovsky, A. V., & Sloot, P. M. (2014). Inference of the Russian drug community from one of the largest social networks in the Russian Federation. *Qual Quant*, 48, pp. 2739–2755. DOI 10.1007/s11135-013-9921-6.
- Ding, . S., Fung, B., Iqbal, F., & Cheung, W. W. (2017). Learning Stylometric Representations for Authorship Analysis. *IEEE Transactions on Cybernetics*.
- Duff, I. S., Grimes, R. G., & Lewis, J. G. (1989). Sparse matrix test problems. *ACM Transactions on Mathematical Software (TOMS)*, 15(1), pp. 1-14. Doi:10.1145/62038.62043.
- Elkan, C. P. (2014, February 12). *Text mining and topic models*. Retrieved 06 10, 2017, from <http://cseweb.ucsd.edu/~elkan/250B/topicmodels.pdf>
- Emu, I. H., Ahmed, A. U., Islam, M. M., Al Mamun, M. S., & Uddin, M. A. (2017). An efficient approach for keyphrase extraction from english document. *International Journal of Intelligent Systems and Applications*, 9(12), pp. 59-66. doi:10.5815/ijisa.2017.12.06.
- Faleiros, T. D., & Lopes, A. D. (2016). *MODELOS PROBABILÍSTICOS DE TÓPICOS: DESVENDANDO O LATENT DIRICHLET ALLOCATION*. Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação. São Carlos: USP.
- Fan, W., Wallace, L., & Zhang, Z. (2006, sept). Tapping the power of text mining. *Communications of the ACM*, 49(9), pp. 77-82.
- Feldman , R., Dagan, I., & Hirsh, H. (1998). Mining text using keyword distributions. *Journal of Intelligent Information Systems*, 10(3), pp. 281-300.
- Feldman, R., & Dagan, I. (1995). Knowledge Discovery in Textual Databases (KDT). In U. Fayyad, & R. Uthurusamy (Ed.), *THE FIRST INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING*, (pp. 112-117). Montreal.

- Feldman, R., & Sanger, J. (2007). *The text mining handbook: Advanced approaches in analyzing unstructured data*. Cambridge, UK: Cambridge University Press.
- Figueiredo, C. J., & Mota, C. M. (2016). A classification model to evaluate the security level in a city based on GIS-MCDA. *Mathematical Problems in Engineering*, 2016(Art. n° 3534824), p. <https://dx.doi.org/10.1155/2016/3534824>.
- Fox, E. A., Andrews, C., Fan, W., Jiao, J., Kassahun, A., Lu, S.-C., . . . Boutwell, L. (2008, March 1). A Digital Library for Recovery, Research, and Learning From April 16, 2007, at Virginia Tech. *Traumatology*, 14(1), pp. 64-84. <https://doi.org/10.1177/1534765608315632>.
- Gajzer, M. (2010). Text and data mining techniques in aspect of knowledge acquisition for decision support system in construction industry. *Technological and Economic Development of Economy*, 16(2), pp. 219-232. <https://doi.org/10.3846/tede.2010.14>.
- Gálvez, R. H., & Gravano, A. (2017). Assessing the usefulness of online message board mining in automatic stock prediction systems. *Journal of Computational Science*, 19, pp. 43-56. <http://dx.doi.org/10.1016/j.jocs.2017.01.001>.
- Goldstein, H. (1990). *Problem-oriented policing*. New York: McGraw-Hill.
- Gowri, S., Anandha Mala, G. S., & Divya, G. (2014, August 20). Enhancing the digital data retrieval system. *Journal of Theoretical and Applied Information Technology*, 66(2), pp. 481-489.
- Griffiths, T., & Steyvers, M. (2004). Finding scientific topics. *PNAS*, 101(1), pp. 5228–5235.
- Gupta, M., Chandra, B., & Gupta, M. P. (2014). A framework of intelligent decision support system for Indian police. *Journal of Enterprise Information Management*, 27(5), pp. 512-540. <https://dx.doi.org/10.1108/JEIM-10-2012-0073>.
- Gurgel, A. M., & Mota, C. M. (2013). A multicriteria prioritization model to support public safety planning. *Pesquisa Operacional*, 33(2), pp. 251-267. <https://dx.doi.org/10.1590/S0101-74382013000200007>.
- Gurusamy, S., Manjula, D., & Geetha, T. (2002). Text Mining in ‘Request for Comments Document Series’. *Proceedings - Language Engineering Conference*, (pp. 147-155).
- Haberman, C. P. (2017, May). Overlapping Hot Spots? Examination of the Spatial Heterogeneity of Hot Spots of Different Crime Types. *Criminology and Public Policy*, 16(2), pp. 633-660. <https://doi.org/10.1111/1745-9133.12303>.
- Hao, J., & Dai, H. (2016). Social media content and sentiment analysis on consumer security breaches. *Journal of Financial Crime*, 23(4), pp. 855-869. <https://doi.org/10.1108/JFC-01-2016-0001>.
- Huang, J. (2016). Patent network analysis of cloud computing by text mining. *Journal of Technology*, 31(2), pp. 127-146.

- Ibekwe-SanJuan, F. (2006). A task-oriented framework for evaluating theme detection systems: A discussion paper. *5th International Conference on Language Resources and Evaluation, LREC 2006* (pp. 2485-2490). Genoa: ELRA.
- Kaur, J., Yusof, M. F., Boursier, P., & Ogier, J. M. (2010). Automated scientific document retrieval. *2nd International Conference on Computer and Automation Engineering, ICCAE 2010* (pp. 732-736). Singapore: IEEE.
- Kostoff, R. N., Block, J. A., Stump, J. A., & Johnson, D. (2008). Literature-related discovery (LRD): Potential treatments for Raynaud's Phenomenon. *Technological Forecasting and Social Change*, 75(2), pp. 203-214. doi:10.1016/j.techfore.2007.11.005.
- Ku, C.-H., & Leroy, G. (2014, October). A decision support system: Automated crime report analysis and classification for e-government. *Government Information Quarterly*, 31(4), pp. 534-544. <https://doi.org/10.1016/j.giq.2014.08.003>.
- Lee, J. H., & Segev, A. (2012). Knowledge maps for e-learning. *Computers & Education*, 59, pp. 353-364. doi:10.1016/j.compedu.2012.01.017.
- Liberatore, F., & Camacho-Collados, M. (2016). A Comparison of Local Search Methods for the Multicriteria Police Districting Problem on Graph. *Mathematical Problems in Engineering*, 2016(Art. n° 3690474), p. <https://dx.doi.org/10.1155/2016/3690474>.
- Liu, S.-H., Liao, H.-L., Pi, S.-M., & Hu, J.-W. (2011, June). Development of a Patent Retrieval and Analysis Platform – A hybrid approach. *Expert Systems with Applications*, 38(6), pp. 7864-7868. <https://doi.org/10.1016/j.eswa.2010.12.114>.
- Mayr, P., Frommholz, I., Cabanac, G., Chandrasekaran, M. K., Jaidkal, K., Kan, M., & Wolfram, D. (2017). Introduction to the special issue on bibliometric-enhanced information retrieval and natural language processing for digital libraries (BIRNDL). *International Journal on Digital Libraries*, pp. 1-5. doi:0.1007/s00799-017-0230-x.
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), pp. 1093-1113. doi:10.1016/j.asej.2014.04.011.
- Mittermayer, M. A. (2004). Forecasting intraday stock price trends with text mining techniques. *Proceedings of the Hawaii International Conference on System Sciences*. 37, pp. 1029-1038. United States: IEEE.
- Morais, E., & Ambrósio, A. (2007). *Mineração de Textos*. Technical Report - INF_005/07, Universidade Federal de Goiás, Instituto de Informática.
- Moulton, E. (2013). Myth and reality: interpreting the dynamics of crime trends: lies, damn lies and (criminal) statistics. *Police Practice and Research*, 14(3), pp. 219-227. <https://doi.org/10.1080/15614263.2012.674300>.
- Müller, H. M., Kenny, E. E., & Sternberg, P. W. (2004). Textpresso: An ontology-based information retrieval and extraction system for biological literature. *PLoS Biology*, 2(11).

- Netsuwan, T., & Kesorn, K. (2017). Unify Framework for Crime Data Summarization using RSS Feed Service. *Walailak Journal of Science & Technology*, 14(10), pp. 769-781.
- Nishanth, K. J., Ravi, V., Ankaiah, N., & Bose, I. (2012). Soft computing based imputation and hybrid data and text mining: The case of predicting the severity of phishing alerts. *Expert Systems with Applications*, 39, pp. 10583–10589. <http://dx.doi.org/10.1016/j.eswa.2012.02.138>.
- Nissan, E. (2017). Digital technologies and artificial intelligence's present and foreseeable impact on lawyering, judging, policing and law enforcement. *AI & Soc*, 32(3), pp. 441–464. <https://doi.org/10.1007/s00146-015-0596-5>.
- Okuhara, T., Ishikawa, H., Okada, M., Kato, M., & Kiuchi, T. (2018). Contents of Japanese pro- and anti-HPV vaccination websites: A text mining analysis. *Patient Education and Counseling*, 101(3), pp. 406-413. 10.1016/j.pec.2017.09.014.
- Özdemir, Ü., & Güneröğlu, A. (2015). Strategic approach model for investigating the cause of maritime accidents [Stratejik bir model yaklaşımı ile deniz kaza sebeplerinin araştırılması]. *Promet - Traffic – Traffico*, 27(2), pp. 113-123. <https://dx.doi.org/10.7307/ptt.v27i2.1461>.
- Poelmans, J., Elzinga, P., Viaene, S., & Dedene, G. (2011, April). Formally analysing the concepts of domestic violence. *Expert Systems with Applications*, 38(4), pp. 3116-3130. <https://doi.org/10.1016/j.eswa.2010.08.103>.
- Poelmans, J., Elzinga, P., Viaene, S., Hulle, M. M., & Dedene, G. (2009, November). Gaining insight in domestic violence with Emergent Self Organizing Maps. *Expert Systems with Applications*, 36(9), pp. 11864–11874. <https://doi.org/10.1016/j.eswa.2009.04.027>.
- Polpinij, J. (2013). Ontology-based knowledge discovery from unstructured text. *International Journal of Information Processing and Management*, 4(4), pp. 21-31. doi:10.4156/ijipm.vol4.issue4.3.
- Prado, H., & Ferneda, E. (2008). *Emerging technologies of text mining: techniques and applications*. Hershey: Information Science Reference.
- Pramanik, M., Lau, R. Y., Yue, W. T., Ye, Y., & Li, C. (2017). Big data analytics for security and criminal investigations. *WIREs Data Mining Knowl Discov*, 7(4), pp. 1-19. doi:10.1002/widm.1208.
- Przeszlowski, K., & Crichlow, V. (2018). An exploratory assessment of community-oriented policing implementation, social disorganization and crime in America. *Social Sciences*, 7(3), p. 35.
- Rakotonirainy, A., Chen, S., Scott-Parker, B., Loke, S. W., & Krishnaswamy, S. (2015). A Novel Approach to Assessing Road-Curve Crash Severity. *Journal of Transportation Safety & Security*, 7(4), pp. 358-375. <https://doi.org/10.1080/19439962.2014.959585>.
- Rao, P. R., & Lalitha Devi, S. (2018). Enhancing multi-document summarization using concepts. *Sadhana - Academy Proceedings in Engineering Sciences*, 43(2), p. 27.

- Renu, R. S., & Mocko, G. M. (2016). Computing similarity of text-based assembly processes for knowledge retrieval and reuse. *Journal of Manufacturing Systems*, 39(1), pp. 101-110. doi:10.1016/j.jmsy.2016.03.004.
- Rosenbaum, D. (1988). Community crime prevention: A review and synthesis of the literature. *Justice Quarterly*, 5(3), pp. 323-395. doi:10.1080/07418828800089781.
- Roy, B., & Hugonnard, J. C. (1982). Classement des prolongements de métro en banlieue parisienne (Présentation d'une méthode multicritère originale). *Cahiers du CERO*, 23, 153-171.
- Saha, P., Bose, I., & Mahanti, A. (2016, April). A knowledge based scheme for risk assessment in loan processing by banks. *Decision Support Systems*, 84, pp. 78-88. <https://doi.org/10.1016/j.dss.2016.02.002>.
- Santos, R. B. (2014). The Effectiveness of Crime Analysis for Crime Reduction: Cure or Diagnosis? *Journal of Contemporary Criminal Justice*, 30(2), pp. 147-168. doi:10.1177/1043986214525080.
- Sapori, L. F. (2007). *Segurança pública no Brasil: desafios e perspectivas*. Rio de Janeiro: Editora FGV.
- Sherman, L. W., Gartin, P. R., & Buerger, M. E. (1989). Hot Spots of predatory crime: routine activities and the criminology of place. *Criminology*, 27(1), pp. 27-56. DOI:10.1111/j.1745-9125.1989.tb00862.x.
- Sherman, L. W., MacKenzie, D. L., Gottfredson, D. C., Eck, J., Reuter, P., & Bushway, S. D. (1998). *Preventing crime: what works, what doesn't, what's promising*. U.S. Department of Justice.
- Sherman, L., & Weisburd, D. (1995). General deterrent effects of police patrol in crime hot spots: A randomized controlled trial. *Justice Quarterly*, 12, pp. 625-648.
- Skolnick, J. H., & Bayley, D. H. (1986). *The new blue line: Police innovation in six American cities*. New York: Free Press.
- Snyders, E., & Landman, K. (2018, February). Perceptions of crime hot-spots and real locations of crime incidents in two South African neighbourhoods. *Security Journal*, 31(1), pp. 265-284. <https://doi-org.ez24.periodicos.capes.gov.br/10.1057/s41284-017-0099-9>.
- Soares, M., Prati, R. C., & Monard, M. C. (2008). *PreTexT: a reestruturação da ferramenta de pré-processamento de textos*. São Carlos: ICMC.
- Song, J., Song, T. M., Seo, D.-C., & Jin, J. H. (2016). Data Mining of Web-Based Documents on Social Networking Sites That Included Suicide-Related Words Among Korean Adolescents. *Journal of Adolescent Health*, 59(6), pp. 668-673. doi:10.1016/j.jadohealth.2016.07.025.

- Tseng, Y.-H., Lin, C.-J., & Lin, Y.-I. (2007). Text mining techniques for patent analysis. *Information Processing & Management*, 43(5), pp. 1216-1247. <https://doi.org/10.1016/j.ipm.2006.11.011>.
- van Banerveld , M., Kechadi , M., & Le-Khac , N. (2016). A Natural Language Processing Tool for White Collar Crime Investigation. In A. Hameurlain , J. Küng , R. Wagner , T. Dang , & N. Thoai , *Transactions on Large-Scale Data- and Knowledge-Centered Systems XXIII* (Vol. 9480, pp. 1-22). Lecture Notes in Computer Science. Springer, Berlin, Heidelberg.
- Venčkauskas, A., Damaševičius, R., Marcinkevičius, R., & Karpavičius, A. (2015). Problems of authorship identification of the national language electronic discourse. *1st International Conference on Information and Software Technologies, ICIST 2015*. 538, pp. 415-432. Druskininkai: Springer Verlag.
- Wang , H., & Hong, M. (2015). Distance Variance Score: An Efficient Feature Selection Method in Text Classification. *Mathematical Problems in Engineering*, pp. 1-10. <http://dx.doi.org/10.1155/2015/695720>.
- Wang, S.-H., Ding, Y., Zhao, W., Huang, Y.-H., Perkins, R., Zou, W., & Chen, J. J. (2016, March 19). Text mining for identifying topics in the literatures about adolescent substance use and depression. *BMC Public Health*, pp. 2-8. <https://doi.org/10.1186/s12889-016-2932-1>.
- Weinborn, C., Ariel, B., Sherman, L. W., & O' Dwyer, E. (2017). Hotspots vs. harmspots: Shifting the focus from counts to harm in the criminology of place. *Applied Geography*, 86, pp. 226-244. <http://dx.doi.org/10.1016/j.apgeog.2017.06.009>.
- Weisburd, D., & Amram, S. (2014). The law of concentrations of crime at place: the case of Tel Aviv-Jaffa. *Police Practice and Research: An International Journal* , 15(2), pp. 101-114. <https://doi.org/10.1080/15614263.2013.874169>.
- Weisburd, D., & Eck, J. E. (2004). What Can Police Do to Reduce Crime, Disorder, and Fear? *The Annals of the American Academy of Political and Social Science*, 593(1), pp. 42-65.
- Weisburd, D., McElroy, J., & Hardyman, P. (1988). Challenges to supervision in community policing: Observations on a pilot project. *American Journal of Police*, 7(2), pp. 29-50.
- Zaeem, R. N., Manoharan, M., Yang, Y., & Barber, K. S. (2017, March). Modeling and analysis of identity threat behaviors through text mining of identity theft stories. *Computers & Security*, 65, pp. 50-63. <https://doi.org/10.1016/j.cose.2016.11.002>.