

Citations of DeBot

For example, DeBot (Chavoshi, Hamooni, and Mueen 2016) can be described as a special case of our approach based on a sophisticated temporal hashing scheme preserving dynamic time warping distance

Pacheco, D., Hui, P.-M., Torres-Lugo, C., Truong, B. T., Flammini, A., & Menczer, F. (2021). Uncovering Coordinated Networks on Social Media: Methods and Case Studies. Proceedings of the International AAAI Conference on Web and Social Media, 15(1), 455-466. <https://doi.org/10.1609/icwsm.v15i1.18075>

Recent approaches in computer science turn towards unsupervised machine learning and group-based detection methods . . . by detecting cross-user activity correlations (Chavoshi et al., 2016) . . .

Martini, F., Samula, P., Keller, T. R., & Klinger, U. (2021). Bot, or not? Comparing three methods for detecting social bots in five political discourses. Big Data & Society, 8(2). <https://doi.org/10.1177/20539517211033566>

The rationale for clustering stems from previous research in human and bot behaviors, in that humans have been proven to exhibit much more behavioral heterogeneity than automated accounts [5,11]. - [5] is DeBot.

Michele Mazza, Stefano Cresci, Marco Avvenuti, Walter Quattrociocchi, and Maurizio Tesconi. 2019. RTbust: Exploiting Temporal Patterns for Botnet Detection on Twitter. In Proceedings of the 10th ACM Conference on Web Science (WebSci '19). Association for Computing Machinery, New York, NY, USA, 183-192. <https://doi.org/10.1145/3292522.3326015>

A well-known model that detects SMBs by using unsupervised machine learning techniques is DeBot (Chavoshi, Hamooni, Mueen, 2016a, Chavoshi, Hamooni, Mueen, 2016b).

Mariam Orabi, Djedjiga Mouheb, Zaher Al Aghbari, Ibrahim Kamel, Detection of Bots in Social Media: A Systematic Review, Information Processing & Management, Volume 57, Issue 4, 2020, 102250, ISSN 0306-4573, <https://doi.org/10.1016/j.ipm.2020.102250>.

An alternative strategy of detecting bots has been to develop technologies based on new machine learning algorithms (Chavoshi, Hamooni, and Mueen, 2016)

Stukal, D., Sanovich, S., Bonneau, R., & Tucker, J. (2022). Why Botter: How Pro-Government Bots Fight Opposition in Russia. American Political Science Review, 116(3), 843-857. doi:10.1017/S0003055421001507

Different from these supervised approaches, Chavoshi et al. use a warped correlation finder to detect bot accounts that have correlated activities [245].

Artificial intelligence for social good: A survey ZR Shi, C Wang, F Fang - arXiv preprint arXiv:2001.01818, 2020 - arxiv.org

Debot is a bot detection service that generates daily reports of bot activity, and stems from the work of [10][9]. It comes with an API which we were able to query to obtain over 700,000 accounts that the service detected as bots.

Juan Echeverriĉja, Emiliano De Cristofaro, Nicolas Kourtellis, Ilias Leontiadis, Gianluca Stringhini, and Shi Zhou. 2018. LOBO: Evaluation of Generalization Deficiencies in Twitter Bot Classifiers. In Proceedings of the 34th Annual Computer Security Applications Conference (ACSAC '18). Association for Computing Machinery, New York, NY, USA, 137-146. <https://doi.org/10.1145/3274694.3274738>

Chavoshi et al. [26] designed the DeBot system using an unsupervised learning approach and proposed a new hash-mapping technique that could quickly group a large number of associated users

B. Wu, L. Liu, Y. Yang, K. Zheng and X. Wang, Using Improved Conditional Generative Adversarial Networks to Detect Social Bots on Twitter, in IEEE Access, vol. 8, pp. 36664-36680, 2020, doi: 10.1109/ACCESS.2020.2975630.

Chavoshi et al. [29] discovered a set of Turkish bots with correlated deletion activity to hide bot like behavior.

T. Elmas, R. Overdorf, A. F. Özkalay and K. Aberer, "Ephemeral Astroturfing Attacks: The Case of Fake Twitter Trends," 2021 IEEE European Symposium on Security and Privacy (EuroS&P), Vienna, Austria, 2021, pp. 403-422, doi: 10.1109/EuroSP51992.2021.00035.

DeBot, for example, searches for users who have correlated tweet times (Chavoshi, Hamooni, and Mueen 2016). Placing their approach within our framework, the action and action type is simply to tweet, giving a one-view coordination network.

Magelinski, T., Ng, L., & Carley, K. (2022). A Synchronized Action Framework for Detection of Coordination on Social Media. Journal of Online Trust and Safety, 1(2). <https://doi.org/10.54501/jots.v1i2.30>

As an unsupervised learning approach, Nikan Chavoshi et al. [15]. proposed a dynamic time warping method that detects bots based on the correlations among accounts. Their results outperformed supervised learning.

Iizuka R, Toriumi F, Nishiguchi M, Takano M, Yoshida M (2022) Impact of correcting misinformation on social disruption. PLoS ONE 17(4): e0265734. <https://doi.org/10.1371/journal.pone.0265734>

The state-of-the-art in unsupervised bot detection is DeBot [8], which uses dynamic time warping to identify accounts with synchronized tweeting patterns.

Chen, Z., Tanash, R.S., Stoll, R., Subramanian, D. (2017). Hunting Malicious Bots on Twitter: An Unsupervised Approach. In: Ciampaglia, G., Mashhadi, A., Yasseri, T. (eds) Social Informatics. SocInfo 2017. Lecture Notes in Computer Science(), vol 10540. Springer, Cham. https://doi.org/10.1007/978-3-319-67256-4_40

Chavoshi et al. [20] has leveraged the semantic and temporal similarity of accounts to identify bots in an unsupervised fashion, creating the DeBot model which we will compare against in our results section.

Beskow, D.M., Carley, K.M. (2020). You Are Known by Your Friends: Leveraging Network Metrics for Bot Detection in Twitter. In: Tayebi, M.A., Glässer, U., Skillicorn, D.B. (eds) Open Source Intelligence and Cyber Crime. Lecture Notes in Social Networks. Springer, Cham. https://doi.org/10.1007/978-3-030-41251-7_3

[DeBot] is an unsupervised method to detect bots using a temporal pattern that is obtained from bot activities.

Boyeon Jang, Sihyun Jeong, Chong-kwon Kim, Distance-based customer detection in fake follower markets, Information Systems, Volume 81, 2019, Pages 104-116, ISSN 0306-4379, <https://doi.org/10.1016/j.is.2018.12.001>.

This list was used in the Twitter API to feed DeBot, which is the tool we used to search for automatic accounts. DeBot is an unsupervised bot classifier developed by Chavoshi et al. [15], [20]. We chose DeBot over other tools (such as Botometer [21]) because it allowed us to capture data by analyzing suspicious accounts in real time based on their behavior and filter them by tags, which reduced processing time afterwards. Also, we were able to get access to DeBot source code, allowing us to add some specific features for this research.

D. Rofrío et al., Presidential Elections in Ecuador: Bot Presence in Twitter, 2019 Sixth International Conference on eDemocracy & eGovernment (ICEDEG), Quito, Ecuador, 2019, pp. 218-223, doi: 10.1109/ICEDEG.2019.8734426.

Citations of MASS

... we show that MASS is 20 times faster than brute force search on the UMAP embedding space.

G. Franch, G. Jurman, L. Coviello, M. Pendesini, and C. Furlanello, MASS-UMAP: Fast and Accurate Analog Ensemble Search in Weather Radar Archives, Remote Sensing, vol. 11, no. 24, 2019.

[MASS] is known to greatly speed up the calculation of distances between sequences.

P. Avogadro, and M. Dominoni, Topological Approach for Finding Nearest Neighbor Sequence in Time Series, in Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management - KDIR, SciTePress, 2019, pp. 233-244.

For computing the similarity, we use Mueen's ultra-fast Algorithm for Similarity Search (MASS), because MASS is algorithmically efficient and free of hyperparameters.

S. Wilhelm and J. Kasbauer, Exploiting Smart Meter Power Consumption Measurements for Human Activity Recognition (HAR) with a Motif-Detection-Based Non-Intrusive Load Monitoring (NILM) Approach, Sensors, vol. 21, no. 23, 2021.

The strong point of MASS is the much faster indexing, ...

D. Piatov, S. Helmer, A. Dignos, and J. Gamper, Interactive and space-efficient multi-dimensional time series subsequence matching, Information Systems, vol. 82, pp. 121-135, 2019.

The adoption of MASS essentially accelerates the extraction process compared to using pruning techniques based on brute force approach.

J. Zuo, K. Zeitouni, and Y. Taher, Incremental and Adaptive Feature Exploration over Time Series Stream, in 2019 IEEE International Conference on Big Data (Big Data), 2019, pp. 593-602.

This approach has the advantages of being easy to implement with the ready-to-use implementations of MASS.

D. Janka, F. Lenders, S. Wang, A. Cohen, and N. Li, Detecting and locating patterns in time series using machine learning, Control Engineering Practice, vol. 93, p. 104-169, 2019.

... the exact algorithm such as MASS by Mueen et al. can be used to efficiently retrieve the exact correlations.

H. Qiu, H. T. Lam, F. Fusco, and M. Sinn, Learning Correlation Space for Time Series, CoRR, vol. abs/1802.03628, 2018.

The distance profile for a single subsequence can be calculated very fast using the MASS algorithm which only needs $O(n \log(n))$ time, ...

Leonard Clauß, Humboldt University of Berlin.

In this paper we utilized the MASS algorithm to speed up the original SetFinder algorithm.

G. Yoshimura, A. Kanemura, and H. Asoh, Enumerating hub motifs in time series based on the matrix profile. 5th Workshop on Mining and Learning from Time Series (MiLeTS'19), 2019

The Mueen's ultra-fast Algorithm for Similarity Search (MASS) strategy is adopted to speed up the ED calculation to ensure the efficiency of the proposed scheme.

J. Hu and N. Chen, Enhanced Feature Summarizing for Effective Cover Song Identification, IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 27, no. 12, pp. 2113-2126, 2019.

... we calculate the distances between this pattern and all subsequences in the series using MASS.

D. De Paepe and S. Van Hoecke, Mining Recurring Patterns in Real-Valued Time Series using the Radius Profile, in 2020 IEEE International Conference on Data Mining (ICDM), 2020, pp. 984-989.

After constructing a motif dictionary, we further apply the MASS algorithm to each sensory time series.

S. Lin, X. Wu, and N. V. Chawla, motif2vec: Semantic-aware Representation Learning for Wearables' Time Series Data, in 2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA), 2021, pp. 1–10.

... conventional MASS-ts and Euclidean distance calculation methods showed similar high accuracy;

H.-J. Lee, Y. Kwon, and S.-Y. Ihm, Dual-ISM: Duality-Based Image Sequence Matching for Similar Image Search, Applied Sciences, vol. 12, no. 3, 2022.

The standard deviation for MASS and UDTW are zero, since they are deterministic retrieval algorithms.

V. Gupta, S. Bedathur, and A. De, Learning Temporal Point Processes for Efficient Retrieval of Continuous Time Event Sequences, 2022.

Like its matrix profile counterparts for pair-wise segment similarity computation, MASS has fast time performance.

J. Liu, J. Li, and L. Liu, FastOPM—A practical method for partial match of time series, Pattern Recognition, vol. 130, p. 108808, 2022.

... the number of neighbors closer to the candidate subsequences than the defined maximum distance is calculated with the MASS algorithm ...

L. Grunerbel, F. Heinrich, D. Dieboldler, and M. Richter, Wearable Decubitus Prophylaxis Tool Based on Machine Learning Methods, in 2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), 2022, pp. 730–734

... MASS method is utilized to construct the full distance matrix of the subsequence with length ...

X. Zheng, X. Zou, X. Ren, C. Ji, and M. Zhang, ECG Prediction Based on Bidirectional Time Series Chain Discovery Algorithm, in 5th International Conference on Crowd Science and Engineering, ser. ICCSE 2021.

We also use MASS directly for subsequence-similarity measurement, in addition to a similar feature-based method that we develop.

N. Singh, Sifting sound - interactive extraction, exploration, and expressive recombination of large and heterogeneous audio collections, Master's thesis, Massachusetts Institute of Technology, Boston, MA, 2020.