Proceedings of the
Seidenberg Annual Research Day

December 17 2024

SARD 2024

SEIDENBERG ANNUAL RESEARCH DAY

Hosted by
Seidenberg School of Computer Science and Information Systems, Pace University
New York, New York

edited by Li-Chiou Chen
Sung-Hyuk Cha

Proceedings of the Seidenberg Annual Research Day 2024

Hosted by:

The Seidenberg School of Computer Science and Information Systems

Published by Pace University Press

Available online
http://csis.pace.edu/∼scha/SARD2024

Table of Contents

Openning Ceremony

Invited Talks

Poster Competition

Commemorating the Legacy of Professor Narayan S. Murthy (1943-2024)

Li-Chiou Chen
Dean of Seidenberg School of Computer Science & Information Systems

The Seidenberg School of Computer Science and Information Systems at Pace University mourns the loss of Professor Narayan S. Murthy, a luminary in the field of computer science and a cherished member of our community. Professor Murthy, who passed away in November 2024, left behind a legacy of dedication, scholarship, and innovation that had profoundly shaped our institution and the lives of countless students and colleagues.

Dr. Murthy was one of the founding members of the Seidenberg School and served as a faculty member for an extraordinary 38 years. For 28 of those years, he led the Computer Science Department on the Westchester campus as its chair, steering its growth and establishing its reputation for academic excellence. His leadership and vision had left an indelible mark on our school.

Dr. Murthy's academic journey was as diverse as it was distinguished. He held a PhD and MS in Mathematics, along with an MS in Computer Science, from the University of Rhode Island. Prior to this, he earned an MS in Mathematics and a BS in Physics and Mathematics from the University of Mysore, India. His multidisciplinary background as a computer scientist, mathematician, and physicist enriched his teaching and research, inspiring those around him to approach problems with curiosity and rigor.

A prolific scholar, Dr. Murthy made significant contributions to a wide range of fields. His notable works include advancements in optimal complete binary search trees, min-max sorting algorithms, mode 1 notch stress intensity factors, photoelastic determination, and multiple precision arithmetic. He authored and co-authored numerous textbook chapters, conference papers, and articles published in respected journals such as the International Journal of Environmental, Cultural, Economic and Social Sustainability and the Journal of Computing Sciences in Colleges.

Dr. Murthy's dedication extended beyond academia into securing resources for future generations of scholars. He was instrumental in obtaining numerous grants from organizations such as Thinkfinity, the Verizon Foundation, and the National Security Agency. These grants enabled the development of innovative programs and research initiatives, fostering a culture of discovery and growth.

As we dedicate the Seidenberg Annual Research Day (SARD) 2024 to Professor Murthy's memory, we honor not only his academic achievements but also his unwavering commitment to education and the community he served. His contributions had left a legacy that will continue to inspire and guide us.

Professor Narayan S. Murthy will be deeply missed, but his influence will endure in the Seidenberg School and beyond. We celebrate his life and work, grateful for the profound impact he has had on all of us.

Welcome to the Seidenberg Annual Research Day 2024

Jonathan Hill
Provost and Executive Vice President of Academic Affairs

It is with great pleasure that we welcome you to the Seidenberg Annual Research Day (SARD) 2024, an esteemed conference hosted by the Seidenberg School of Computer Science and Information Systems at Pace University. This event has established itself as a cornerstone of academic excellence, celebrating both pioneering research and exceptional student contributions.

SARD 2024 offers a vibrant platform for faculty, students, and professionals to engage in meaningful dialogue, share groundbreaking ideas, and explore advancements in Computer Science, Data Science, and Information Technologies. By featuring a diverse range of participants, including seasoned researchers and emerging scholars, this conference fosters an inclusive environment that bridges academic and professional communities.

Expert talks by distinguished faculty members aim to inspire and challenge our attendees, ensuring accessibility and relevance for a broad audience, particularly students. This year's invited speakers are Dr. Soheyla Amirian, Dr. Jonathan Williams, and Dr. Yegin Genc. We extend our sincere gratitude to them for their invaluable contributions.

The proceedings of SARD 2024 include 15 outstanding abstracts selected from over 16 submissions. These abstracts represent original research and thought-provoking problems pertinent to Computer Science and Information Systems. We commend all contributors for their dedication and innovation and thank them for enriching this conference with their work.

We hope this event serves as a catalyst for new ideas, collaborations, and learning opportunities. Thank you for your participation, and we look forward to your active engagement in the sessions ahead.

Welcome to SARD 2024!

Invited Talk I
Fair and Responsible AI in Medical Imaging Informatics

Soheyla Amirian
Department of Computer Science, Pace University New York, NY
samirian@pace.edu

Artificial intelligence (AI) has already made a big leap in healthcare, with AI-powered medical imaging at the forefront of innovation. However, the rapid uptake and adoption of AI brings challenges in ensuring fairness, equity, and responsibility in its applications. This talk highlights challenges and solutions for ensuring fairness and responsibility in AI-powered medical imaging informatics. Focusing on widely accessible and cost-effective imaging modalities such as plain radiographs (X-rays), we -together- will explore computational strategies to first detect and then mitigate biases and enhance AI algorithm impartiality. This underscores the importance of responsible AI practices to ensure that medical imaging technologies are safe, equitable, and beneficial for all.

Dr. Soheyla Amirian is currently an Assistant Professor at the Seidenberg School of Computer Science and Information Systems, Pace University. She leads the Applied Machine Intelligence Initiatives & Education (AMIIE) Laboratory, collaborating with a multidisciplinary team of faculty, students, and investigators to design, build, validate, and deploy AI algorithms in various real-world applications, including public health, imaging informatics, and AI-powered education. Dr. Amirian earned her BSc, MSc, and PhD in Computer Science, specializing in AI and deep learning computer vision. She has received numerous accolades, including the 2019 International Conference on Computational Science and Computational Intelligence (CSCI) Outstanding Achievement award, the 2021 UGA Outstanding Teaching Assistant award, the NVIDIA GPU award, and the ACM Richard TAPIA Conference Scholarship in 2020 and 2022. Additionally, she was a finalist for the 2020 NCWIT Collegiate Award and has authored over 25 peer-reviewed publications. Dr. Amirian has organized several conferences and tutorials on computational intelligence, such as ISVC and IEEE ICHI, and has served as a Program Committee member at IEEE ICHI and Co-Chair of Research Tracks at the World Congress in Computer Science, Computer Engineering, and Applied Computing (CSCE) and the International Conference on Computational Science & Computational Intelligence (CSCI). She was a faculty fellow at the Institute for Artificial Intelligence and a faculty lecturer at the School of Computing, University of Georgia for three years. In 2023, she received the IEEE Atlanta Section Outstanding Educator Award. Most recently, she was awarded a grant from the NIH/National Institute on Aging (NIA).

Invited Talk II
Designerly Knotting in Beginning Design Education

Jonathan Williams
Department of Information Technology, Pace University New York, NY
jwilliams6@pace.edu

When positioned within the design domain, designers can distinguish their own form of knot tying, designerly knotting, that is different from other professional or common applications of knot tying. When designerly knots are placed into a learning environment, designerly knots function as constructionist learning objects within beginning design education practice. This research presents two instructional cases in which knots were used in the context of beginning design education. The first case describes the use of designerly knotting to teach design to academically gifted high school students in a summer residential program; the second case describes the use of designerly knotting to teach design to adult learners as part of a career training program. This research provides evidence that designerly knotting activities can stand alone, or stand alongside, other beginning design activities and exercises to teach beginning design.

Dr. Jonathan Williams is a Clinical Assistant Professor at Pace University in the Seidenberg School of Computer Science and Information Systems (New York, NY). Jonathan's practice of design is an integral part to his agenda for practice-based design research. Jonathan's own design practice spans multiple practice areas, including information visualization, instructional design, and design strategy.

Jonathan has worked broadly in adult education, having designed technical courses at Harvard Business School, General Assembly, and edX/2U, Inc. Jonathan has held previous academic appointments as an Assistant Professor of Management & Information Technology at St. Francis College (Brooklyn, NY), Clinical Lecturer at Iona University (New Rochelle, NY), Adjunct Assistant Professor at New York University's School of Professional Studies (New York, NY), an Adjunct Assistant Professor of Health Administration at New York University's Robert F. Wagner Graduate School of Public Service (New York, NY), a Part Time Lecturer at Parsons School of Design (New York, NY), and a Part Time Lecturer at Rutgers University (Newark, NJ).

Jonathan holds a Doctor of Design (D.Des) from North Carolina State University, Master of Science (MS) in Strategic Design and Management from Parsons School of Design, and a Bachelor of Science (BS) in Mathematical Statistics from Wake Forest University.

Invited Talk III
Role of AI Explanations in Augmented Decision Making

Yegin Genc
Department of Information Technology, Pace University New York, NY
ygenc@pace.edu

Recent studies have revealed several metacognitive processes when human decisions are augmented with AI recommendations. These processes help decision makers evaluate the validity of AI recommendations (system-monitoring), and the validity of their own knowledge and beliefs (self-monitoring.) This work examines the effects of AI explanations - that are typically used to evaluate the correctness of AI recommendations, on augmented decision making from the perspective of these cognitive mechanisms.

Dr. Yegin Genc is the Chairperson of the Information Technology Department at Pace University and Co-Director of the Computational Intelligence Lab, where he promotes collaborative research and experiential learning. Prior to joining Pace University, Dr. Genc served as a Professor of Computer Information Systems and Business Analytics at the School of Business, Manhattan College. From 2004 to 2012, he worked as Director and Manager of Professional Services at Interneer Inc., gaining extensive industry experience. Dr. Genc holds a B.S. in Mechanical Engineering from Istanbul Technical University (2001), an M.S. in Information Technology from the University of Central Missouri (2003), and a Ph.D. in Information Management from Stevens Institute of Technology (2014).

Mobile Money Security: End-to-End Encryption over Legacy Infrastructure

Fabrice Nice Neza, and Anthony Joseph,
Department of Computer Science, Pace University New York, NY
{fn16449n, ajoseph2}@pace.edu

Mobile Network Operators (MNO) infrastructure is rapidly expanding in many Sub-Saharan African (SSA) countries, where mobile payment services are becoming increasingly popular. However, the existing model remains outdated and ill-suited for today's dynamic, relatively secure, and application-driven environments. Unstructured Supplementary Services Data (USSD) technology continues to be widely deployed in SSA due to factors like handset compatibility, user experience, cost-effectiveness, and ease of deployment for MNOs. Surveys indicate that approximately 65-75% of the population in SSA countries rely on USSD for mobile payments. However, the MNO infrastructure was primarily developed for connectivity rather than security, leaving the e-money ecosystem vulnerable to cyber-attacks. End-to-end encryption (E2EE) is often suggested as a key cybersecurity solution. This research demonstrates that tokenization, combined with identity-based encryption techniques, is an effective method for securing financial transactions over USSD. Tokenization replaces sensitive financial data with a unique identifier (token) that can be used in place of the original data, reducing the risk of data breaches. Furthermore, identity-based encryption techniques, though effective, are not inherently end-to-end. To ensure a secure communication path while minimizing the need for significant changes to existing technology and infrastructure, thus achieving end-to-end data security, we propose combining tokenization, end-to-end encryption techniques with Internet Protocol Security (IPSec) and Transport Layer Security (TLS), which provide a secure communication path from the mobile device to the server. This integration ensures that the entire transaction process remains encrypted, protecting user data throughout its transmission. However, the implementation introduces additional maintenance and management costs, which would require governments and other key stakeholders to introduce market-friendly laws, regulations, and policies that promote skill development and enable MNOs to achieve a positive return on investment.

References

[1] Malek, B., 2014. *Method and System for a Certificate-less Authenticated Encryption Scheme Using Identity Based Encryption*, US Patent No. 8, pp. 694 771.

[2] Iwasokun, G. et al., 2018. *Encryption and Tokenization-Based System for Credit Card Information Security*. International Journal of Cyber-Security and Digital Forensics, Vol. 7, pp. 283-293.

[3] Ramadan, M. et al., 2016. *EEE-GSM: End-to-end encryption scheme over GSM system*. International Journal of Security and its Applications, Vol. 10, pp. 229-240.

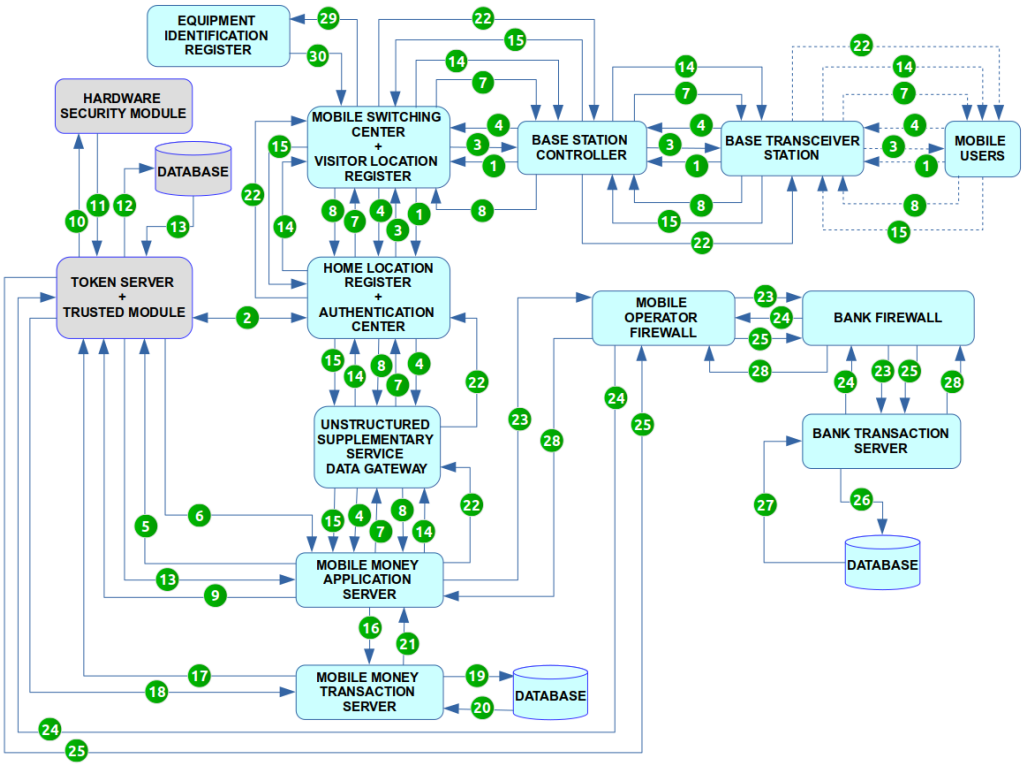


Figure 1: Encryption and tokenization scheme flow diagram

Legend:

1. MS initiates the USSD session and identifies the trusted module (TM) by the TM identity string (IdTM)
2. Exchange of the plurality of public parameters (PK) between the trusted module (TM) and HLR
3. The PK associated with or generated by TM is distributed by the HLR to the mobile user through the MSC/VLR
4. The MS transmits the ciphertext (C) to the MoMoAS over the mobile network (BTS-BSC-MSC/VLR-HLR/AuC-USSD GW)
5. The MoMoAS requests the PK and its own prvMoMoAS from TM
6. The TM provides the PK to MoMoAS
7. MoMoAS generates pubMS, decrypts (C), and displays a menu at MS along with the request to submit the MoMoAN
8. The MoMoAN is submitted to the MoMoAS
9. The MoMoAS uses the TS's API to securely request a token related to the MoMoAN
10. The TS uses the HSM to create the token
11. The HSM replies the TS
12. The token, MoMoAN, and possibly the couple MSISDN-IMSI are stored in the token data store (database)
13. The TS securely returns the token to the MoMoAS
14. The token is sent back through USSD GW and HLR to MS where it is securely stored
15. The mobile user securely initiates the action to complete the transaction
16. The MoMoAS securely sends a request to the TrS
17. The TrS securely sends the token to the TS to retrieve the actual MoMoAN
18. The MoMoAN is securely returned to the TrS
19. The transaction information is securely sent to the Database to be validated
20. The database validates the info and completes the transaction
21. The transaction status is securely sent back to the MoMoAS
22. A notification status is sent back to MS through HLR with a flag to terminate the original USSD session
23. If the TrS is located outside the MNO network, the MoMoAS sends requests to the MtrS through an IPSec link
24. The MtrS securely initiates the request to the MNO's TS to retrieve the actual MoMoAN
25. The MoMoAN is securely returned to the MtrS
26. The transaction information is securely sent to the Database to be validated
27. The database validates the info and completes the transaction
28. The transaction status is sent back to the MoMoAS and then to the mobile user via (22)
29. Mobile Switching Center sends a request to verify the MS [doesn't come into the financial transaction MS-MMoAS]
30. The Equipment Identification Register sends a reply [doesn't come into the financial transaction MS-MMoAS]

Knee Osteoarthritis Classification Using Bone Distance

Rahul Singh and Juan Shan,
Department of Computer Science, Pace University New York, NY
{rs63891n, jshan}@pace.edu

Osteoarthritis (OA) is the most prevalent form of arthritis affecting the knee joint, with severity commonly assessed using the Kellgren-Lawrence (KL) scale. Traditional methods for evaluating OA severity rely on direct measurement of cartilage thickness, which is time-consuming and challenging in clinical practice. This study introduces an innovative approach to classify knee OA severity by using femur-tibia bone distances as a proxy biomarker for cartilage thickness. Bone masks were segmented from MRI images using a U-Net model. For each case, 160 slices were available, a custom algorithm identified approximately 28 valid slices per case by focusing on slices with complete and well-defined femur and tibia structures while excluding irrelevant or incomplete slices. On each valid slice, the boundaries of the femur and tibia bones were identified using external contour detection techniques. These contours were sorted by area, with the largest contour corresponding to the femur and the second largest to the tibia. Distance measurements were restricted to the gap region between the femur and tibia bones, where the Euclidean distance was calculated for all point pairs along their respective boundary contours. The average distance across all gap-specific points was computed for each slice, yielding a single representative distance value per slice. These distances were aggregated into a feature vector for each patient, with dimensions of approximately 28 features per patient. The final dataset consisted of 197 rows (patients) and 28 columns (features). This feature matrix was used to train a Random Forest Classifier for binary classification of OA severity, distinguishing non-OA cases (KL 0–2) from OA cases (KL 3–4). The model was evaluated using 10-fold cross-validation, achieving a mean accuracy of 65% and a ROC AUC score of 60%. These results demonstrate the potential of this method to reduce reliance on direct cartilage measurements, making OA classification more efficient and accessible. Future work will focus on enhancing the accuracy by refining feature extraction and exploring advanced machine learning models.

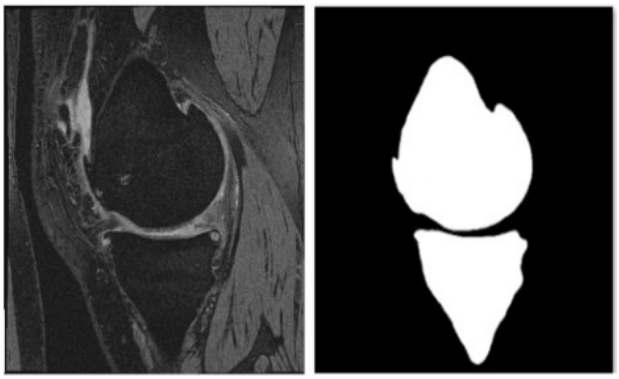


Figure 1: (a) A raw MRI image and (b) Bone mask segmented by U-net

References

[1] R. Ponnusamy, M. Zhang, Y. Wang, X. Sun, M. Chowdhury, J. B. Driban, T. McAlindon, and J. Shan, *Automatic segmentation of bone marrow lesions on MRI using a deep learning method*, Bioengineering 2024, vol. 11, no. 4: 374

[2] R. Almajalid, J. Shan, M. Zhang, G. Stonis, and M. Zhang, *Knee Bone Segmentation on Three-Dimensional MRI*, 2019 18th IEEE International Conference on Machine Learning and Applications (ICMLA), New York City, NY, USA, 2019

[3] A. D. Desai, A. M. Schmidt, E. B. Rubin, C. M. Sandino, M. S. Black, V. Mazzoli, K. J. Stevens, R. Boutin, C. Ré, G. E. Gold, B. Hargreaves, and A. Chaudhari, *SKM-TEA: A Dataset for Accelerated MRI Reconstruction with Dense Image Labels for Quantitative Clinical Evaluation*, in Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks 2021

Small Scale Interoperable Social Media: A Design Framework For Decentralization

Akshay Vijay Nair and Sotirios Skevoulis
Department of Computer Science, Pace University New York, NY
{an66160n, sskevoulis}@pace.edu

The pitfalls of using centralized social media platforms for individuals and institutions have been well documented and analyzed. Those pitfalls led to the rise of decentralized social media platforms. Initially, the technical challenges surrounding the developments and maintenance were high. That however has changed significantly with open-source solutions, widely accepted open protocols, the lowering cost of cloud services for infrastructure, and newly available LLM-based technologies. This paper explores the design for a social media system for internal institutions and organizations.

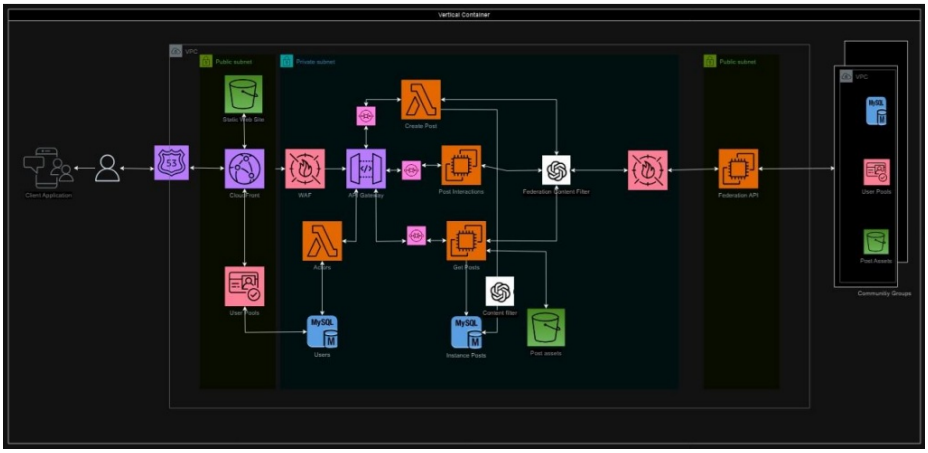


Figure 1: : Software architecture example with LLM integrations

References

[1] https://www.w3.org/TR/activitypub/

[2] Deepak Kumar, Yousef AbuHashem, and Zakir Durumeric, *Watch Your Language: Investigating Content Moderation with Large Language Models*, arXiv:2309.14517 [cs.HC], https://doi.org/10.48550/arXiv.2309.14517

Comparative study of neural nets and graphical models for causality and interpretability on the same dataset

Yusuf Ahmed and Krishna M. Bathul
Department of Computer Science, Pace University New York, NY
{ya89495n, kbathula}@pace.edu

As machine learning systems increasingly inform critical decisions, the need for causal explanations will become crucial. While deep learning neural networks (DL-NNs) have achieved remarkable performance in a broad spectrum of domains, their black-box nature raises concerns among users and stakeholders who require explainable artificial intelligence (XAI). Probabilistic Graphical Models (PGMs) can be used as a possible alternative, that offers strong theoretical foundations for causality and interpretability. This study presents a systematic comparison between DL-NNs and PGMs, evaluating both approaches on identical datasets using causal and interpretable metrics. In examining existing literature and conducting empirical analysis, this research attempts to investigate the tradeoffs between predictive power and explainability. This research will hopefully provide a basic and accountable framework for understanding how PGM and DL-NN models are used in applications requiring more causal reasoning and transparency by comparing their results and metrics that may offer advantages of using one model over the other in some use-cases.

References

[1] Arrieta, Alejandro Barredo, et al., *Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI* 2021

[2] Russell, S. J., and Norvig, P., *Artificial Intelligence: A Modern Approach* , 3rd ed., Prentice Hall, Upper Saddle River, New Jersey, 2009

[3] Bishop, Christopher. *Pattern Recognition and Machine Learning*, 2006

[4] Sithakoul, Samuel, et al., *BEExAI: Benchmark to Evaluate Explainable AI.* 2024.

[5] Dwivedi, Vijay Prakash, et al., *Benchmarking Graph Neural Networks.* 2020.

Ideal Roommate Finder Platform

Bhakti Shelke and Krishna M. Bathul
Department of Computer Science, Pace University New York, NY
{bs58507, kbathula}@pace.edu

The Ideal Roommate Finder project leverages collaborative filtering to match international students with compatible roommates. By analysing user preferences and ratings, the platform recommends suitable matches based on shared interests, lifestyles, and cultural backgrounds. Innova Test is employed to collect and analyze user data. A comprehensive questionnaire gathers information about each user's desired living arrangements, lifestyle habits, cultural background, and budget. The platform's user interface features an intuitive questionnaire, a visualization of the matching process, a profile comparison tool, and a secure messaging system. By utilizing collaborative filtering and Innova Test, the Ideal Roommate Finder project aims to improve matching accuracy, enhance user experience, increase user satisfaction, and build a stronger community. This innovative approach addresses the challenges faced by international students in finding suitable roommates, providing an efficient and effective solution.

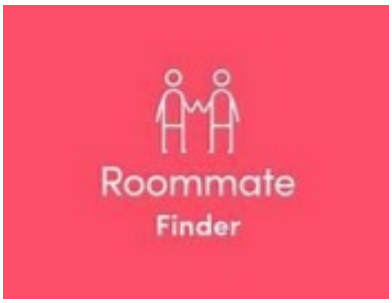


Figure 1: Roommate Finder Logo

References

[1] E. Elviwani, A. Putera Utama Siahaan and L, Fitriana, *Performance-based Stable Matching using Gale-Shapley Algorithm*, July. 2018

[2] M. Lamba, R. Goswami, M. Vinay and M. Lamba, *Roommate Compatibility Detection Through Machine Learning Techniques*, pp. 1-4, 2020

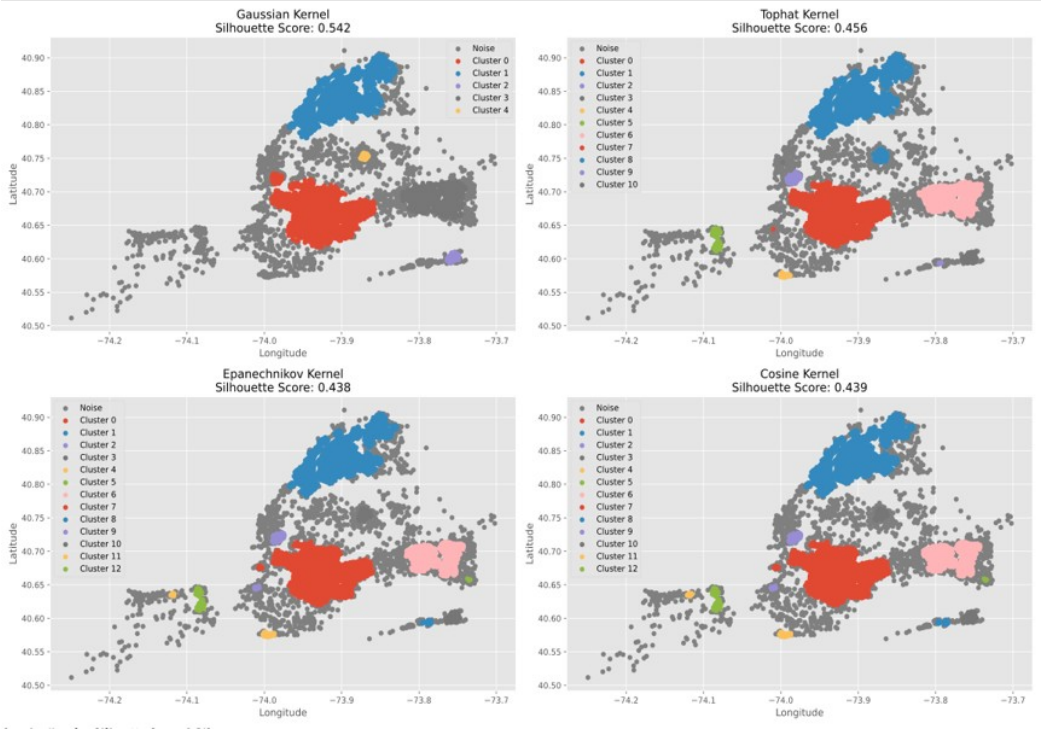
Kernel Density Based Clustering with Noise

Rohan Niranjan Kalpavruksha, Roshan Niranjan Kalpavruksha, and Sung-Hyuk Cha
Seidenberg School Of Computer Science and Information Systems, Pace University
New York, NY
{rk25464n, rk68465n, scha}@pace.edu

This research introduces a novel clustering methodology that leverages Kernel Density Estimation (KDE) to overcome the limitations of traditional density-based algorithms like DBSCAN. Existing methods often struggle with fixed parameter settings, such as minimum points and epsilon, leading to challenges in handling clusters with varying densities. Furthermore, these techniques rely heavily on the number of samples within a radius, often ignoring the actual density of the data points. These issues result in reduced accuracy and difficulty in distinguishing meaningful clusters in complex datasets.

The proposed approach focuses on density-driven clustering, utilizing KDE to estimate the density distribution of data points through various kernel functions, such as Gaussian, triangular, and rectangular. By identifying high-density regions and selecting the top 80% of points, the method excludes low-density outliers as noise, significantly improving cluster precision. Clusters are formed by grouping points within the kernel range, ensuring flexibility and adaptability to varying densities. Comparative analysis demonstrates superior performance, with improved silhouette scores and robustness to noise, compared to DBSCAN and other conventional techniques.

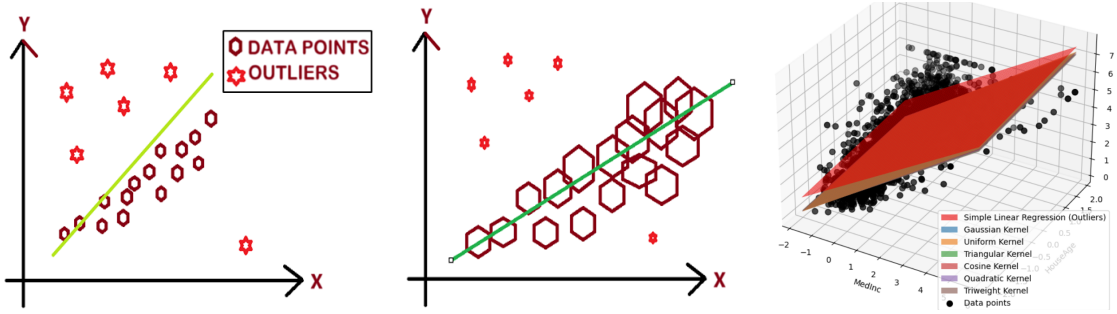
This research highlights the potential of kernel density as a powerful tool for enhancing data analysis, particularly in datasets with uneven distributions, such as social media sentiment analysis, customer segmentation in e-commerce, and medical data with noise-prone measurements.



Kernel Density Based Linear Regression

Roshan Niranjan Kalpavruksha, Rohan Niranjan Kalpavruksha, and Sung-Hyuk Cha
Seidenberg School Of Computer Science and Information Systems, Pace University
New York, NY
{rk68465n, rk25464n, scha}@pace.edu

This research proposes a kernel density-based linear regression (KDLR) method to overcome the limitations of traditional linear regression when dealing with noisy and unevenly distributed datasets. Standard regression techniques assume all data points are equally significant, making them vulnerable to outliers and missing critical patterns in real-world data. KDLR addresses these challenges by using kernel density estimation to assign higher weights to data points in dense regions, while de-emphasizing sparse or noisy points. This dynamic weighting approach enhances the model's robustness and ensures better pattern recognition. Evaluations on datasets, such as the California housing data with varying levels of outliers, using metrics such as mean squared error and R^2-score, demonstrate that KDLR achieves superior accuracy compared to traditional regression. By prioritizing dense regions, the method captures meaningful patterns while reducing the influence of noise. Applications include stock price predictions in finance, patient data modelling in healthcare, and climate modelling in environmental studies, where robust handling of anomalies and noise is crucial. This research establishes KDLR as a transformative tool for predictive analytics, offering a powerful combination of precision, resilience, and adaptability. Its demonstrated ability to manage noisy data and highlight critical patterns positions it as a versatile solution for both academic and industrial applications, with the potential to redefine best practices in regression modelling.



(a) Ordinary LR (b) univariate KDLR (c) multidimensional case

References

[1] Scott, D. W., *Multivariate Density Estimation: Theory, Practice, and Visualization*, Wiley 2015

[2] Hastie, T., Tibshirani, R., & Friedman, J., *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, 2009

Exponentially Weighted Moving Linear Regression

Jagadesh Varma Nadimpalli and Sung-Hyuk Cha
Department of Computer Science, Pace University New York, NY
{jn80343n, scha}@pace.edu

The Exponentially Weighted Moving Linear Regression (EWMLR) is a novel approach that extends the widely used Exponentially Weighted Moving Average (EWMA) by integrating linear regression to enhance forecasting accuracy. While EWMA effectively emphasizes recent observations, it struggles to capture underlying linear trends often present in time series data. EWMLR addresses this limitation by combining exponential weighting with linear regression, enabling it to model both recent changes and broader trends. To evaluate its performance, EWMLR was tested on Bitcoin stock data, a time series characterized by high volatility and discernible trends. The results demonstrate that EWMLR consistently outperforms EWMA in predictive accuracy, particularly in capturing the dynamic patterns of Bitcoin's price movements. This improvement highlights EWMLR's potential as a valuable tool for financial analysis, demand forecasting, and real-time decision-making in dynamic environments.

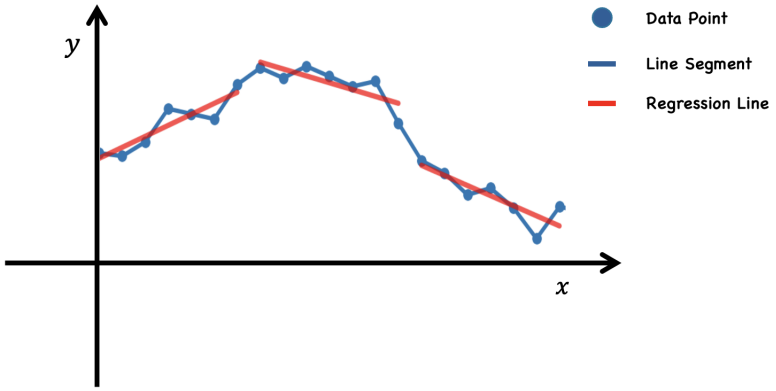


Figure 1: Working example of Exponentially Weighted Moving Linear Regression (EWMLR)

References

[1] S. W. Roberts, "Exponential Moving Average," *Encyclopedia of Operations Research and Management Science*, 2nd ed. Wiley, 2009.

[2] J. Huang and L. Yang, "Application of linear regression models in time series analysis for stock market prediction," *Quantitative Finance*, vol. 18, no. 5, pp. 783-798, 2022.

No Epsilon Needed: A Density-Based Clustering Method Using Euclidean Distances

Tyler Helmrich,
Computer Science Department, Pace University, New York, NY, USA
th44707n@pace.edu

When it was first presented in 1996 DBSCAN (Density-Based Spatial Clustering of Applications with Noise) kicked off a paradigm shift in clustering identification, from primarily proximity-based methods such as k-medoids, to a focus on point density. The density-based approach has had many advantages, better computational efficiency, "natural" cluster identification as opposed to relying on domain knowledge for cluster count, ability to identify oddly shaped clusters that proximity-based approaches easily miss, and ability to discard noise points. Despite these advantages the density technique still has downsides. The standard method of computing point density relies on finding the number of points within a fixed radius called Epsilon, denoted as ϵ. The first issue with ϵ is that being fixed causes it to struggle to find clusters of varying density, too large and more dense clusters can easily become attached to less dense ones, too small and less dense clusters won't be identified at all, due to this contradiction finding the goldilocks ϵ can be impossible. Second, while techniques exist to approximate a decent ϵ, users often fall back on hand picking an ϵ, shifting the domain knowledge requirement from number of clusters to size of ϵ. This paper proposes a new density-based algorithm for identifying clusters that resolves the issues of the ϵ method by abandoning ϵ altogether. The new approach instead generates a circle centered over each point whose radius is the euclidean distance to it's nearest neighbor. The overlap of these circles is used to create a graph, each distinct graph then represents a cluster. These cluster graphs preserve all the benefits of the standard approach while also being able to identify clusters of variable density without requiring any domain knowledge from the user.

References

[1] Ester, M., Kriegel, H. P., Sander, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining* (KDD'96). AAAI Press. 10.5555/3001460.3001507

[2] R. Ng, and J. Han (1994) Efficient and effective clustering method for spatial Data Mining, In *Proceedings of 20th International Conference on very Large Scale databases* (VLDB '94). 10.5555/645920.672827

Training a model to classify open-source contributors on GitHub

Fabian Hiller
Department of Computer Science, Pace University New York, NY
fh53510n@pace.edu

A significant proportion of contemporary software is dependent on open-source software, which is often developed by independent contributors. For those responsible for maintaining an open-source project, it can be beneficial to be able to identify which contributors may require additional support, such as funding, mentorship, or a direct involvement in the core team. This research project presents a comprehensive study on the development and training of a machine learning model designed to classify new and existing open-source contributors on GitHub based on their activity, contributions, and engagement patterns. The methodology employed utilizes a diverse dataset comprising over 100,000 instances of GitHub user interactions to identify discrete contributor types. The efficacy of the model will be assessed based on its capacity to accurately predict contributor categories and its practical utility in real-world GitHub repositories. Potential future enhancements may include the expansion of the feature set to encompass more nuanced indicators of contribution quality and impact.

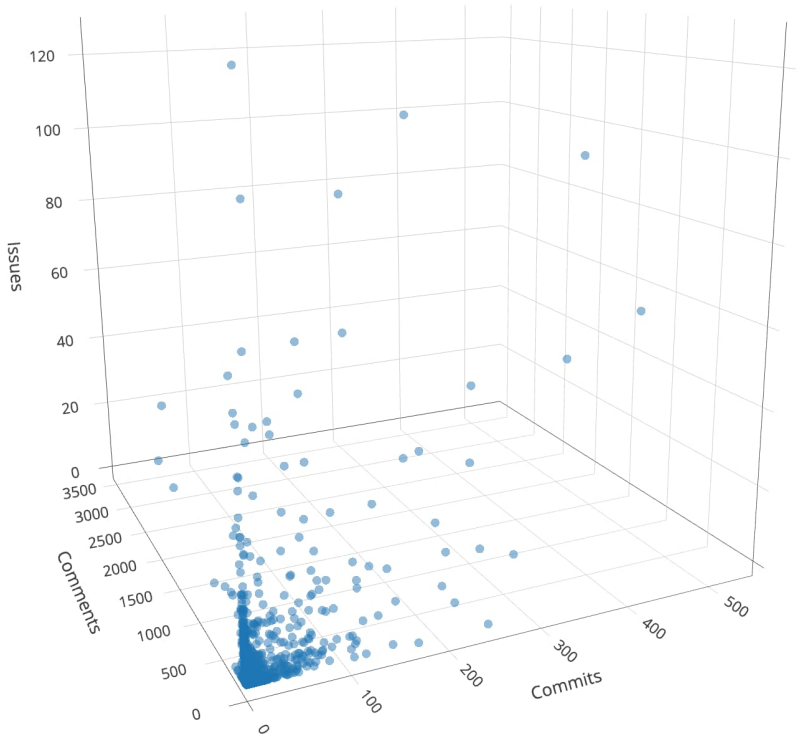


Figure 1: 3D scatter plot of the dataset used

Learning how to envision artificial intelligence futures using the PLACARD pattern

M. Tedeschi[1], P. Ricaurte[2], S. Ayloo[3], R. López-Seseneś[3], J. Corneli[4], C. J. Danoff[5] and S. Belich[3],
[1]Seidenberg School of Computer Science and Information Systems, Pace University, NY, USA
[2]Tecnologico de Monterrey Mexico
[3]Computer Systems Technology, New York City College of Technology USA
[4]Hyperreal Enterprises Limited United Kingdom Oxford Brookes University UK
[5]Mr. Danoff's Teaching Laboratory USA

At EuroPLoP 2024 Mary Tedeschi led the "AI Future Envisioning with PLACARD" focus group in Germany. Three conference attendees joined in the room while Sridevi, Paola, and Charles cofacilitated remotely via a web conference. The participants were introduced to a Futures Studies technique with the goal of capturing envisionments of Artificial Intelligence (AI) going forward. To set an atmosphere a technology focused card game was used to make the session more interactive. To close everyone co-created a Project Action Review to recap of the event to capture learnings that has been summarized in this paper. The Focus Group was structured based on lessons learned over six earlier iterations.

References

[1] Mary Tedeschi, Paola Ricaurte, Sridevi Ayloo, Joseph Corneli, Charles Jeffrey Danoff, and Sergio Belich, *AI Future Envisioning with PLACARD*, In Proceedings of the 29th European Conference on Pattern Languages of Programs, People, and Practices, Association for Computing Machinery, 2024 https://doi.org/10.1145/3698322.3698444

Efficiency Prediction of the 5-Aminotetrazole Inhibitor in the Corrosion of AA6065-AZ31 Alloys Using Electrochemical Noise and Artificial Neural Networks

Arianna Parrales[1], Sung-Hyuk Cha[2], J. M. Angeles[1], R. López-Sesenes[3], and J. A. Hernández[1],
[1]Centro de Investigación en Ingeniería y Ciencias Aplicadas, UAEM, Morelos, México.
[2]Seidenberg School of Computer Science and Information Systems, Pace University, NY, USA
[3]Facultad de Ciencias Químicas e Ingeniería, UAEM, Morelos, México
arianna.parrales@uaem.mx

Corrosion is a natural process that deteriorates metallic materials, compromising their resistance and functionality while impacting safety, the environment, and the economy. To mitigate this issue, chemical inhibitors are often used to form protective films on metal surfaces, although their effectiveness requires thorough evaluation. Electrochemical Noise (EN) is a widely used technique for assessing metal corrosion by monitoring their spontaneous potential and current fluctuations [1]. Despite its utility, interpreting EN signals remains challenging due to their highly nonlinear and random nature, which limits broader applications [2]. This study introduces an innovative approach using artificial neural networks (ANN) to model and predict EN behavior. The study model EN-ANN focuses on the AA6065-AZ31 aluminum-magnesium alloy, evaluating its response to different concentrations of the inhibitor 5-Aminotetrazole in a saline environment.

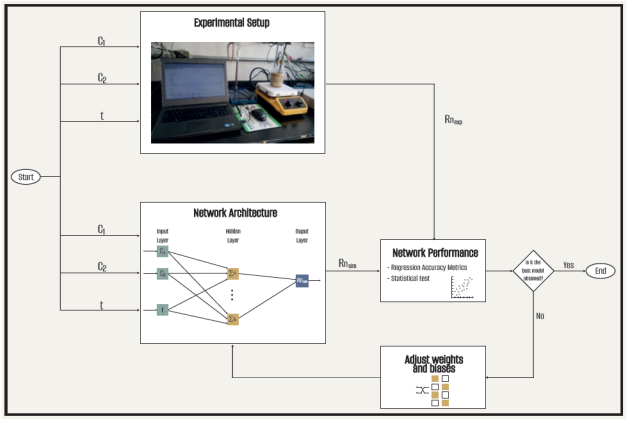


Figure 1: Figure 1. Schematic diagram of ANNs training process for electrochemical noise modeling.

References

[1] I.B. Obot, I.B. Onyeachu, A. Zeino, S.A. Umoren, *Electrochemical noise (EN) technique: review of recent practical applications to corrosion electrochemistry research*, Journal of Adhesion Science and Technology. 33 (2019) 1453 - 1496. https://doi.org./10.1080/01694243.2019.1587224.

[2] C.A. Loto, *Electrochemical noise evaluation and data statistical analyses of stressed aluminium alloy in NaCl solution*, Alexandria Engineering Journal. 57 (2018) 1313 - 1321. https://doi.org/10.1016/j.aej.2017.05.009.

Edstutia VR Tutor Training System

Ariana Contes, Ashley Peleg, Raines Whitehill, and Tom M. Schmidt
Seidenberg School of Computer Science and Information Systems
Pace University, NY, USA
{acontes, apeleg, rwhitehill}@pace.edu

The Edstutia VR Tutor Training System is an innovative virtual reality (VR) application designed to revolutionize tutor training through immersive simulations. By integrating AI-driven avatars that replicate diverse student behaviors and real-world teaching challenges, this system provides a dynamic and interactive platform for educators to hone their skills. The core features include customizable VR training scenarios, performance analytics, and session recording for comprehensive feedback. These functionalities aim to enhance tutor preparedness, foster adaptability, and improve communication techniques in a safe, controlled environment. The system prioritizes user accessibility while laying the foundation for future advancements, such as enhanced scenario customization and AI-driven feedback mechanisms. Although the system has the potential for performance analytics integration, it would require a separate, dedicated system to implement effectively. This project represents a significant step toward redefining educational training, leveraging cutting-edge VR and AI technologies to meet the evolving needs of educators.

Enhanced Detection of DoS and DDoS Attacks in IoT Networks using Image Based Analysis

Bhakti Shastri and Li-Chiou Chen
Seidenberg School of Computer Science and Information Systems
Pace University, NY, USA
{bhakti.o.shastri,lchen}@pace.edu

The rapid growth of the IoT ecosystem has resulted in a significant increase in Distributed Denial of Service (DDoS) attacks and Denial of Service (DoS) attacks, reducing service performance. These sophisticated threats cannot be effectively addressed by traditional methods like firewalls and intrusion detection systems (IDS) because of their inherent limitations. In recent years, deep-learning models in particular, convolutional neural networks-have drawn a lot of interest due to their remarkable image-processing capabilities. This paper presents an innovative deep-learning methodology that reconfigures network traffic and detection mechanisms via ResNet50, utilizing RGB images for the identification of DDoS and DoS attacks. This paper introduces a novel deep-learning technique that converts the network traffic into images to detect DDoS and DoS attacks using ResNet50. We purpose novel technique of feature selection, normalization and transform methods for training and testing. This method outperforms the traditional methods in achieving high recall, better F1-measure, and a precision rate of 93.02% with the CIC-DDoS2019 dataset. The results demonstrate the scalability and dependability of image-based detection, establishing a foundation for adaptive security in IoT systems. By simplifying complex transformations, this method effectively retains high accuracy in recognizing malicious patterns while enhancing operational efficiency.

References

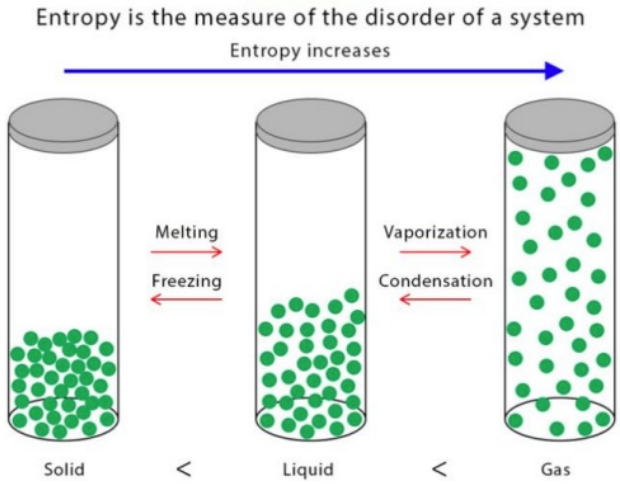
[1] D. Sharafaldin, Iman & Habibi Lashkari, Arash & Sahib, Iqball & Ghorbani, Ali. (2019). Developing Realistic Distributed Denial of Service (DDoS) Attack Dataset and Taxonomy. 1-8.

[2] S. F. Hussain, S. G. Abbas, M. Husnain, U. U. Fayyaz, F. Shahzad, and G. A. Shah, "IoT DoS and DDoS Attack Detection using ResNet," in 2020 IEEE 23rd International Multitopic Conference (INMIC), 2020, pp. 1–6.

On Parameterized Entropy

Sidratul Muntaha and Arpit Shah
Computer Science Department, Pace University
New York, NY
{sm92721n, as61268n}@pace.edu

Entropy was first introduced in the mid nineteenth century by Clausius. In various fields of study entropy is used, but it was initially used in thermodynamics. But as the years passed and there were more advancements made to different fields, entropy was used as a significant topic. Entropy is currently used vastly in machine learning, signal processing and statistical inference. In today's age, Shannon entropy is the most widely used entropy. Despite of being other forms of entropy such as Gini's, Renyi and Tsallis entropies, Shannon's entropy gives us the best results when calculated. The purpose of our paper is to try different parameterized entropy and check if it incorporates with the generalized entropies as well as whether it gives us better results than Shannon entropy. By using parameterized entropy, we plan to have a more flexible adaptation of Shannon's. With its complex history and diverse interpretation, as well as the uncertainty of concept of entropy, it has been one of the most crucial topics for any application in different scientific field. Throughout paper, we plan to find out if there is any better solution to finding entropy used different parameters in different datasets.



References

[1] Hei Tat Tse, *The Definition And Application Of Entropy*, 2022

[2] Marko Popovic, *Research in entropy wonderland: A review of the entropy concept*, 2018

[3] F. Flores Camacho, N. Ulloa Lugo, and H. Covarrubias Martinez, *The concept of entropy, from its origins to teachers*, 2015

Adaptive Decision Tree Learning for Intrusion Detection Systems

Joseph Acampora
Seidenberg School of Computer Science and Information Systems
Pace University, NY, USA
jacampora@pace.edu

The continuously changing cybersecurity threat landscape causes the effectiveness of a machine learning model for intrusion detection to decrease over time, unless such a model can incorporate the characteristics of emergent threats in its predictions. Techniques for incorporating identifying information for new threats can include applying continuous-learning algorithms and information importance weighting based on domain knowledge in addition to classic information impurity measurements. This research demonstrates how the Hoeffding Tree algorithm, which is distinct in its ability to continuously perform supervised learning from a stream of labeled data, can effectively classify malicious events based on new training samples from a simulated real-time threat feed. Included here is a demonstration of successful classification for malicious events that have been dynamically altered to avoid detection or events whose characteristics are within reasonable deviation from learned values. The effectiveness of the Hoeffding Tree classifications is contrasted with a decision tree lacking any capability for adapting its model.

References

[1] C. T. Thanh, "A Study of Machine Learning Techniques for Cybersecurity," 2021 15th International Conference on Advanced Computing and Applications (ACOMP), Ho Chi Minh City, Vietnam, pp. 54–61, https://doi.org/10.1109/ACOMP53746.2021.00014

[2] Gazdar, Tahani, "A New IDS for Smart Home Based on Machine Learning," 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN), Al-Khobar, Saudi Arabia, 2022, pp. 393–400, https://doi.org/10.1109/CICN56167.2022.10008310

[3] J. Aiken and S. Scott-Hayward, "Investigating Adversarial Attacks against Network Intrusion Detection Systems in SDNs," 2019 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN), Dallas, TX, USA, 2019, pp. 1–7, https://doi.org/10.1109/NFV-SDN47374.2019.9040101

[4] D. Han et al., "Evaluating and Improving Adversarial Robustness of Machine Learning-Based Network Intrusion Detectors," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2632–2647, Aug. 2021, https://doi.org/10.1109/JSAC.2021.3087242

[5] M. Usama, M. Asim, S. Latif, J. Qadir, and Ala-Al-Fuqaha, "Generative Adversarial Networks For Launching and Thwarting Adversarial Attacks on Network Intrusion Detection Systems," 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC), Tangier, Morocco, 2019, pp. 78–83, https://doi.org/10.1109/IWCMC.2019.8766353

[6] Muallem, A., et al., "TDDEHT: Threat Detection Using Distributed Ensembles of Hoeffding Trees on Streaming Cyber Datasets," MILCOM 2018 - 2018 IEEE Military Communications Conference (MILCOM), pp. 1–6, 2018.

[7] Alsolami, T., Alsharif, B., and Ilyas, M., "Enhancing Cybersecurity in Healthcare: Evaluating Ensemble Learning Models for Intrusion Detection in the Internet of Medical Things," *Sensors (Basel, Switzerland)*, vol. 24, no. 18, 2024, https://doi.org/10.3390/s24185937

[8] D. G. Corrêa, F. Enembreck, and C. N. Silla, "An Investigation of the Hoeffding Adaptive Tree for the Problem of Network Intrusion Detection," 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 2017, pp. 4065–4072, https://doi.org/10.1109/IJCNN.2017.7966369

[9] LevelBlue - Open Threat Exchange, "LevelBlue Open Threat Exchange." https://otx.alienvault.com/

[10] CIC-IDS2017, "University of New Brunswick est. 1785." https://www.unb.ca/cic/datasets/ids-2017.html

Index of Participants

A - Authors CC - Conference Chairs DG - Distinguished Guest PC - Program Committee