2nd INTERNATIONAL CONFERENCE ON
Artificial Intelligence
31 May - 1 June 2026

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Cross-Dataset Empirical Evaluation of NSGA-II for Multi-Objective Feature Selection in Intrusion Detection Systems

Mahdis Rahmani*, Fereshteh-Azadi Parand
Master of Science Student in Computer Science at Allameh Tabataba'i University
Associate Professor of Computer Science at Allameh Tabataba'i University
mahdis_rahmani@atu.ac.ir

Abstract

Intrusion Detection Systems (IDS) are crucial for network security, but high-dimensional IDS datasets challenge model training. Redundant or irrelevant features degrade classifier performance and increase computational overhead, making feature selection essential. Existing single-objective methods often struggle to balance feature count and accuracy.

This study evaluates the Non-dominated Sorting Genetic Algorithm II (NSGA-II) for multi-objective feature selection. The framework is tested on nine benchmark datasets: NSL-KDD, UNSW-NB15, CIC-IDS2017, CSE-CIC-IDS2018, UKM-IDS20, UNR-IDD, N-BaIoT, BoT-IoT, and X-IIoTID, spanning traditional, modern, and IoT environments. Results show NSGA-II **reduces the feature space by an average of 72.71%** while maintaining or improving detection accuracy. These findings demonstrate NSGA-II to be a generalizable optimization framework for high-dimensional IDS feature selection.

Research method

The study employs an NSGA-II wrapper-based framework for multi-objective feature selection. Within this framework, NSGA-II iteratively generates and evaluates feature subsets by training base classifiers (k-NN, Random Forest, or Logistic Regression) to simultaneously minimize classification error and feature count. This evolutionary loop produces an optimal Pareto front, which then undergoes final performance evaluation using an unseen test set.

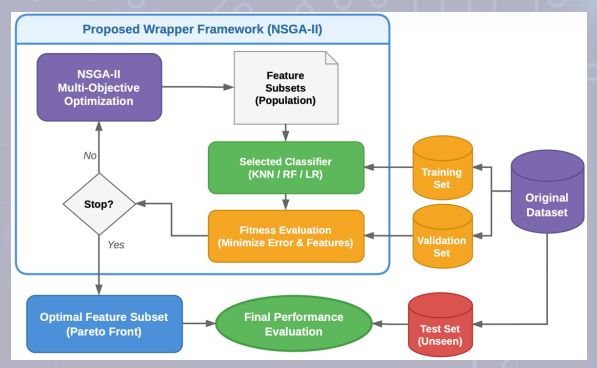

Fig. 1. The proposed NSGA-II wrapper-based feature selection framework

conclusion

The framework successfully balances model complexity and predictive performance across diverse network environments. Among the base learners, Random Forest produced the most stable and accurate Pareto fronts. The algorithm particularly excelled at the extremes, isolating a single viable feature in the UNR-IDD dataset (a **96.43% reduction**) and achieving a peak **accuracy improvement of +1.61%** in the X-IIoTID dataset.

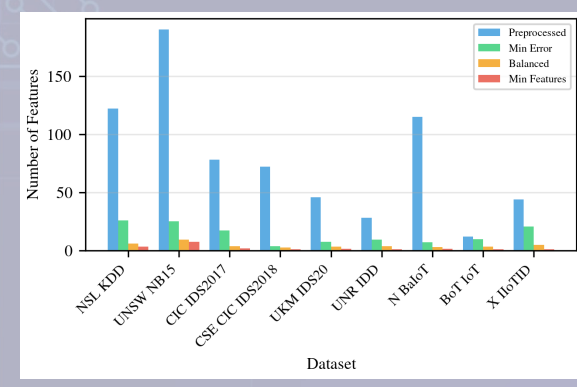

Fig. 2. NSGA-II feature selection results compared to initial preprocessed dataset dimensionality

The purpose of the research

This research evaluates the Non-dominated Sorting Genetic Algorithm II (NSGA-II) for multi-objective feature selection in Intrusion Detection Systems (IDS), aiming to balance the conflicting goals of minimizing feature count and maximizing detection accuracy.

The contributions of this paper are as follows:

1) **A comprehensive empirical evaluation** of NSGA-II across nine benchmark datasets encompassing traditional, modern, and IoT environments, thereby demonstrating its cross-domain generalizability.
2) **A systematic assessment** of three base machine learning classifiers (k-NN, LR, RF) integrated within the NSGA-II wrapper framework.
3) **Significant reductions in computational overhead**, with feature space reductions from 8.33% to 96.43% while preserving or even improving detection accuracy.

Practical or simulation results

Across all **27 experimental configurations**, NSGA-II achieved an **average feature reduction of 72.71%**. The resulting optimized subsets not only maintained an **average detection accuracy of 94.04%,** slightly exceeding the baseline, but also **improved classification accuracy in 66.7% of the test configurations.**

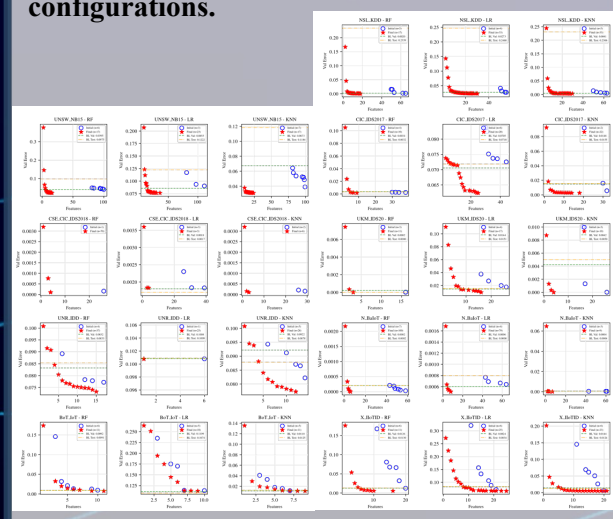

Fig. 3. Pareto Fronts across all datasets and classifiers

References

- K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," IEEE Trans. Evol. Comput., vol. 6, no. 2, pp. 182-197, Apr. 2002.
- C. Khammassi and S. Krichen, "A NSGA2-based feature selection approach for intrusion detection," in Proc. 2020 IEEE/ACS 17th Int. Conf. Comput. Syst. Appl. (AICCSA), 2020, pp. 1-6.
- M. Roopak, G. Y. Tian, and J. Chambers, "Multi-objective-based feature selection for DDoS attack detection in IoT networks," IET Netw., vol. 9, no. 3, pp. 120-127, May 2020.
- A. J. Rabash, M. Z. A. Nazri, A. Shapii, and M. K. Hasan, "Non-dominated sorting genetic algorithm-based dynamic feature selection for intrusion detection system," IEEE Access, vol. 11, pp. 125080-125093, 2023.
- Y. Meidan et al., "N-BaIoT Network-based detection of IoT botnet attacks using deep autoencoders," IEEE Pervasive Comput., vol. 17, no. 3, pp. 12-22, Jul. 2018.