GRADUATE SCHOOL OF ENGINEERING & SCIENCE
AGÜ

# ELECTRICAL AND COMPUTER ENGINEERING

A Ph.D. Thesis

**Hilal Hacılar**

hilal.hacilar@agu.edu.tr

0000-0002-5811-6722

Thesis Advisor

**Assoc. Prof. Dr. Burcu Bakır Güngör**

burcu.gungor@agu.edu.tr

22.08.2024

## Machine Learning Based Network Anomaly Detection

**abstract** I Intelligent technologies have led to a significant rise in internet users and applications. However, this rise in internet usage has also brought serious security challenges. Organizations rely on Network Intrusion Detection systems (NIDS) to protect sensitive data from unauthorized access and theft. To enhance the capabilities of IDS, Machine Learning (ML) and Deep Learning (DL) techniques have been increasingly integrated into these systems. In this context, anomaly-based network intrusion detection surpasses other detection mechanisms significantly in several instances. These systems analyze network traffic to detect suspicious activities, such as attempted breaches or cyberattacks. However, existing studies lack a thorough assessment of class imbalances, feature selection and extraction methods, hyperparameter optimization, and classification performance for different types of network intrusions: wired, wireless, and Software Defined Networking (SDN). Additionally, existing methods may achieve high accuracy; they may suffer from high training times, low detection rate (DR), and computational complexity. By combining metaheuristics and neural networks, it is possible to solve complex optimization problems that are challenging to solve using conventional methods. To address these challenges, this thesis study first evaluates different network intrusion datasets, such as wired, wireless, and SDN, together, considering class imbalance, feature selection, and hyperparameter optimization tasks. Secondly, it proposes a novel hybrid approach combining Deep Autoencoder (DAE) and Artificial Neural Network (ANN) models trained by a parallel Artificial Bee Colony (ABC) algorithm with Bayesian hyperparameter optimization.

**keywords** Network Intrusion Detection systems (NIDS), Network Anomaly Detection, Machine Learning (ML), Deep Learning (DL), Metaheuristics.

**özet** Akıllı teknolojiler, internet kullanıcılarının ve uygulamalarının önemli ölçüde artmasına neden olmuştur. Ancak, internet kullanımındaki bu artış ciddi güvenlik sıkıntılarını da beraberinde getirmiştir. Kuruluşlar, hassas verileri yetkisiz erişim ve hırsızlıktan korumak için Ağ Saldırı Tespit Sistemlerine (NIDS) güvenmektedir. IDS'nin yeteneklerini artırmak amacıyla, makine öğrenimi (ML) ve derin öğrenme (DL) teknikleri giderek daha fazla bu sistemlere entegre edilmektedir. Bu bağlamda, anomali tabanlı ağ saldırı tespiti, birçok durumda diğer tespit mekanizmalarını önemli ölçüde geride bırakmaktadır. Bu sistemler, ağ trafiğini analiz ederek, saldırı girişimleri veya siber saldırılar gibi şüpheli faaliyetleri tespit etmektedir. Ancak, mevcut çalışmalar, kablolu, kablosuz ve Yazılım Tanımlı Ağlar (SDN) gibi farklı türde ağ saldırıları için sınıf dengesizlikleri, özellik seçimi ve çıkarma yöntemleri, hiperparametre optimizasyonu ve sınıflandırma performansı konularında kapsamlı bir değerlendirmeden yoksundur. Ayrıca, mevcut yöntemler yüksek doğruluk elde edebilirken, yüksek eğitim süreleri, düşük tespit oranları ve hesaplama karmaşıklığı gibi sorunlar yaşayabilirler. Metaheuristikler ve sinir ağlarını birleştirerek, geleneksel yöntemlerle çözülmesi zor olan karmaşık optimizasyon problemlerini çözmek mümkündür. Bu zorlukları ele almak için, bu tez çalışması ilk olarak, kablolu, kablosuz ve SDN gibi farklı ağ saldırı veri setlerini sınıf dengesizliği, özellik seçimi ve hiperparametre optimizasyonu görevlerini dikkate alarak birlikte değerlendirmektedir. İkinci olarak, Bayes hiperparametre optimizasyonu ile paralel yapay arı kolonisi algoritması tarafından eğitilen Derin Otomatik Kodlayıcı ve ANN modellerini birleştiren yeni bir hibrit yaklaşım önermektedir.

**anahtar kelime** Ağ Saldırı Tespit Sistemleri (NIDS), Ağ Anomali Tespiti, Makine Öğrenimi (ML), Derin Öğrenme (DL), Metaheuristikler.