RADemics

Social Media Fraud and Fake Profile Detection Using AI and Behavioral Analysis

Shubhangi Abhijit Solanke, Saraswathi Natarajan
PVG'S College of Engineering, Dhaanish Ahmed institute of Technology

Social Media Fraud and Fake Profile Detection Using AI and Behavioral Analysis

[1]Shubhangi Abhijit Solanke, PVG'S College of Engineering, Technology and Management, Pune, Maharashtra, India. shubhmalas@gmail.com

[2]Saraswathi Natarajan, Assistant Professor, Department of Science and Humanities-Mathematics, Dhaanish Ahmed institute of Technology, Coimbatore, Tamil Nadu, India. saraswathi31994@gmail.com

Abstract

The exponential growth of social media platforms has transformed global digital communication while simultaneously increasing vulnerabilities associated with fake profiles, identity impersonation, phishing attacks, misinformation propagation, financial scams, and coordinated cybercriminal activities. Rapid advancements in cyber threat strategies, automated bot networks, and AI-generated synthetic identities have created significant challenges for traditional fraud detection mechanisms within online social ecosystems. Intelligent cybersecurity frameworks supported through Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), and behavioral analytics provide advanced solutions for identification and prevention of malicious social media activities. This book chapter presents a comprehensive investigation of AI-driven fraud detection methodologies focused on behavioral pattern analysis, anomaly recognition, predictive cybersecurity analytics, social graph intelligence, Natural Language Processing (NLP), and hybrid deep learning architectures for fake profile identification. Analytical discussion highlights the role of supervised learning, unsupervised learning, Graph Neural Networks (GNNs), Convolutional Neural Networks (CNNs), and sequential behavioral modeling techniques for detecting coordinated malicious operations and evolving cyber threats across social networking platforms. Critical challenges involving data imbalance, adversarial attacks, privacy preservation, scalability limitations, and explainability within AI-based fraud detection systems receive detailed examination alongside emerging solutions involving Explainable Artificial Intelligence (XAI), federated learning, and blockchain-assisted identity verification mechanisms. Integration of behavioral biometrics, temporal activity analysis, and multimodal learning approaches demonstrates substantial potential for improving detection accuracy, adaptive threat intelligence, and real-time cybersecurity monitoring within modern digital communication environments. The presented discussion contributes toward development of scalable, intelligent, and trustworthy social media security frameworks capable of strengthening digital trust and protecting online communities against sophisticated fraudulent activities and coordinated cyber threats.

Keywords: Social Media Fraud, Fake Profile Detection, Artificial Intelligence, Behavioral Analysis, Deep Learning, Cybersecurity

Introduction

Social media platforms have revolutionized digital communication by enabling rapid information exchange, virtual collaboration, online marketing, and social interaction across global communities [1]. Platforms such as Facebook, Instagram, X (formerly Twitter), LinkedIn, and TikTok generate massive volumes of user-generated data through posts, comments, multimedia sharing, and interactive communication activities [2]. Continuous technological advancements in mobile computing, cloud infrastructure, and internet accessibility have accelerated the growth of social networking ecosystems worldwide [3]. Alongside these developments, social media environments have increasingly attracted cybercriminal activities involving fake profiles, phishing attacks, misinformation campaigns, spam dissemination, identity impersonation, and financial fraud operations [4]. Large-scale digital connectivity and anonymous online interactions provide favorable conditions for malicious actors seeking exploitation of platform vulnerabilities and user trust relationships [5]. Rapid evolution of cyber threat strategies has created serious concerns regarding digital security, online privacy, and reliability of social communication systems across personal, organizational, and governmental environments.

Fake profiles represent one of the most significant threats within modern social networking platforms due to widespread involvement in deceptive digital activities and cybercrime operations [6]. Fraudulent accounts frequently utilize fabricated identities, stolen photographs, synthetic media content, and manipulated personal information for imitation of genuine users and establishment of false credibility [7]. Cybercriminal groups employ such profiles for phishing attacks, malware distribution, financial scams, social engineering operations, political propaganda, and artificial engagement generation [8]. Automated bot networks and coordinated fake account communities intensify the complexity of fraudulent activities through synchronized posting behavior, mass communication campaigns, and manipulation of online discussions [9]. Traditional rule-based security frameworks and manual moderation systems encounter substantial limitations while addressing rapidly evolving fraud patterns within dynamic social networking ecosystems [10]. Increasing sophistication of AI-generated identities, deepfake technologies, and machine-generated communication structures has created urgent demand for intelligent cybersecurity mechanisms capable of detecting malicious activities with high accuracy and operational scalability.