الجمهورية الشعبية الديمقراطية الجزائرية
People's Democratic Republic of Algeria
وزارة التعليم العالي و البحث العلمي
Ministry of Higher Education and Scientific Research
المدرسة العليا للإعلام الآلي 8 ماي 1945 - سيدي بلعباس
Higher School of Computer Science
8 Mai 1945 - Sidi Bel Abbes

Master's Thesis

To obtain the diploma of **Master's Degree**

Field of Study: **Computer Science**

Specialization: **IASD**

Theme

Solution For Fake Profile Detection In Social Media

Presented by
Benounene Abdelrahmane

Defended on: **09, 2025**
In front of the jury composed of

Mr. [Jury Member Name]	President of the Jury
Ms. Saidi Imene	Thesis Supervisor
Mr. Mahammed Nadir	Co-Supervisor
Mr. [Jury Member Name]	Examiner

Academic Year: 2024/2025

Abstract

The proliferation of fake social media profiles poses a critical threat to digital trust, enabling misinformation, fraud, and societal manipulation. This thesis conducts a comprehensive analysis of detection paradigms through a novel taxonomy categorizing text-based, hybrid, and image-based approaches. By synthesizing 17 seminal studies (2017-2023), we benchmark performance against real-world constraints like adversarial evasion and platform variability. Our findings reveal that image-driven methods—particularly optimized gradient features—achieve state-of-the-art accuracy (99.98%) but face generalization gaps under interface changes. Text-centric models prove vulnerable to low-content profiles, while hybrid approaches sacrifice scalability for robustness. The study identifies urgent cross-paradigm challenges: synthetic data bias, "cyborg" account detection failures, and platform lock-in. These insights yield actionable pathways for adaptive multimodal frameworks, advocating standardized benchmarks and adversarial hardening. This work bridges theoretical advances with operational realities to fortify digital ecosystems against evolving impersonation threats.

Keywords: Fake account detection, Social media security, Text-based detection, Hybrid detection, Image-based detection, Taxonomy, Comparative analysis, State-of-the-art review

الملخص

يشكل انتشار الملفات الشخصية المزيفة على وسائل التواصل الاجتماعي تهديدًا خطيرًا للثقة الرقمية، مما يتيح التضليل والاحتيال والتلاعب المجتمعي. تُجري هذه الأطروحة تحليلاً شاملاً لنماذج الكشف من خلال تصنيف جديد يصنف المقاربات القائمة على النصوص، والمقاربات الهجينة، والمقاربات القائمة على الصور. من خلال تجميع 71 دراسة أساسية (7102-3202)، نقوم بقياس الأداء مقابل قيود العالم الحقيقي مثل التهرب من الخصوم وتباين المنصات. تكشف النتائج التي توصلنا إليها أن الأساليب التي تعتمد على الصور - خاصةً ميزات التدرج المحسّنة - تحقق دقة فائقة (89.99%) ولكنها تواجه ثغرات في التعميم في ظل تغييرات الواجهة. وتثبت النماذج التي تركز على النص أنها عرضة للتأثر بالملامح منخفضة المحتوى، بينما تضحي النُهج الهجينة بقابلية التوسع من أجل المتانة. تحدد الدراسة التحديات الملحة عبر النماذج: التحيز في البيانات الاصطناعية، وفشل الكشف عن حسابات "السايبورغ"، وانغلاق المنصة. تسفر هذه الرؤى عن مسارات قابلة للتنفيذ للأطر التكيفية متعددة الوسائط، وتدعو إلى وضع معايير موحدة وتقوية الخصومة. يربط هذا العمل بين التقدم النظري والواقع التشغيلي لتحصين النظم الإيكولوجية الرقمية ضد تهديدات انتحال الشخصية المتطورة.

الكلمات المفتاحية:الذكاء الحاسوبي الذكاء الحاسوبي، الملفات الشخصية المزيفة الكشف عن الملفات الشخصية المزيفة، تطبيع الصور، التحليل البارامترية، علم ما وراء الطبيعة، التعلم الآلي

Résumé

La prolifération de faux profils de médias sociaux représente une menace critique pour la confiance numérique, permettant la désinformation, la fraude et la manipulation sociétale. Cette thèse effectue une analyse complète des paradigmes de détection par le biais d'une nouvelle taxonomie catégorisant les approches basées sur le texte, les approches hybrides et les approches basées sur l'image. En synthétisant 17 études fondamentales (2017-2023), nous comparons les performances aux contraintes du monde réel telles que l'évasion des adversaires et la variabilité des plateformes. Nos résultats révèlent que les méthodes basées sur l'image - en particulier les caractéristiques de gradient optimisées - atteignent une précision de pointe (99,98 %) mais sont confrontées à des lacunes de généralisation en cas de changement d'interface. Les modèles centrés sur le texte s'avèrent vulnérables aux profils à faible contenu, tandis que les approches hybrides sacrifient l'évolutivité à la robustesse. L'étude identifie les défis urgents qui se posent à tous les paradigmes : le biais des données synthétiques, les échecs de détection des comptes « cyborg » et l'enfermement dans une plateforme. Ces observations ouvrent des voies d'action pour des cadres multimodaux adaptatifs, préconisant des repères normalisés et un renforcement de la résistance à l'adversité. Ce travail fait le lien entre les avancées théoriques et les réalités opérationnelles afin de fortifier les écosystèmes numériques contre les menaces d'usurpation d'identité en constante évolution.

Keywords : Intelligence informatique, Faux profils détection, Normalisation d'image, Analyse paramétrique, Métaheuristique, Apprentissage automatique