Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and
Mathematics
Department of Computer Science

SDLA: Synchronous and Decentralized Live Analysis Digital Forensic Model for Edge IoT Environments

**A Thesis Submitted to the Council of the College of
Computer Science and Mathematics
University of Mosul
as a Partial Fulfillment of Requirements
for the Degree of Doctor of Philosophy in
Computer Science**

By

Karam Muhammed Mahdi Salih

Supervised by

Dr. Najla Badie Ibraheem Khalil

2024 A.D. **1445 A.H.**

Abstract

A critical field of study called digital forensics looks for and analyzes digital evidence to find and stop cybercrimes. It involves the systematic examination of digital devices, networks, and information in order to locate, save, and analyze evidence for use in court. In order to improve the accuracy, performance, and reliability of forensic analyses, a digital forensics model integrates advanced technologies and methodologies into an organized framework that directs and speeds up the process of investigation. These models are designed to meet the evolving requirements in modern digital environments.

With an emphasis on edge Internet of Things (IoT) devices, this thesis seeks to create and apply a new digital forensic model called Synchronous and Decentralized Live Analysis (SDLA) digital forensic model, which is distinguished by its Decentralized, Synchronous, Heterogeneous, Privacy-Preserving, Scalable, and Trusted features. The primary objectives include Decentralizing tasks for local processing and peer-to-peer communication, creating real-time communication protocols for synchronized device analysis, creating neural network architectures that can manage heterogeneous data, utilizing federated learning techniques to preserve privacy, enabling scalable parallel processing, and establishing trust through access control and blockchain technology.

The main works done by SDLA model are demonstrated through improvements in Edge IoT digital forensics, enhanced analysis of IoT sensor data, and robust evidence preservation. An Improved Edge IoT dataset for non-identical federated learning with 4 datasets is developed. In addition, a realistic IoT environment is created for real-time data collection, mirroring real edge IoT scenarios through hardware deployment represented by ESP32. This makes it easier to obtain real-time, authentic data for forensic analysis. Furthermore, the CNN-Autocoder model offers an interesting system in the classification of attack scenarios in non-identical federated learning environments. Using autoencoder architectures and convolutional neural networks (CNNs), this model improves the efficiency and accuracy of recognizing and classifying various cyberattack types in heterogeneous Internet of Things datasets. Growing on this basis, the FedForensic system proposal presents deep learning algorithm architectures (RNN and LSTM) and forensic techniques designed to efficiently analyze and report digital evidence from edge IoT devices. This model analyzes packets using digital forensics tools, collect the information, then generates reports automatically suitable for legal proceedings and speeds up the analysis process by providing insights into possible cyber threats and breaches. Additionally, the Alpha-FedAvg algorithm's development meets the need for adaptable federated learning strategies in edge IoT

environments by guaranteeing that data distributions and network conditions change in real-time. Lastly, the CustodyChainGuardian (CCG) system's implementation strengthens the trustworthiness and reliability of forensic investigations in edge IoT environments by utilizing blockchain technology and a multi-layered architecture for secure storage and management to guarantee the integrity and traceability of digital evidence throughout the forensic process.

The effectiveness and efficiency of the SDLA model are proven across a range of models and system configurations through testing and validation, providing a comprehensive solution for digital forensic investigations in edge IoT environments. RNN-based AlphaFed-Avg model performs well in the FedForensic IoT subsystem, demonstrating its adaptability with four datasets from Edge-IIoTset with an accuracy of 93.05%. Performance is increased to 99.84% when LSTM architecture is used. The model maintains 97.45% accuracy even with new environment's real-time data, demonstrating its capacity to adapt to changing situations. Additionally, the CNN-Autocoder model, with an accuracy of 92.6%, contributes to the diverse range of attacks classification in IoT forensics. In the CCG subsystem, the London protocol shows the lowest ultimate cost among Ethereum protocols highlighting the financial consequences. Additionally, by incorporating immutability, transparency, trust, and privacy, our CCG model guarantees robustness.

وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم علوم الحاسوب

SDLA
نموذج التحليلات الجنائية الشرعية الرقمية للتحليل المباشر المتزامن واللامركزي لبيئات إنترنت الأشياء

اطروحة مقدمة
الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة دكتوراه فلسفة في
علوم الحاسوب

من قبل

كرم محمد مهدي صالح

بإشراف
ا.م.د. نجلاء بديع إبراهيم خليل

١٤٤٥ هـ ٢٠٢٤ م

المستخلص

تعد الأدلة الجنائية الشرعية الرقمية مجال دراسي بالغ الأهمية حيث يبحث عن الأدلة الرقمية ويحللها للعثور على الجرائم الإلكترونية وإيقافها. ويتضمن الفحص المنهجي للأجهزة والشبكات والمعلومات الرقمية من أجل تحديد الأدلة وحفظها وتحليلها لاستخدامها في المحكمة. من أجل تحسين دقة وأداء وموثوقية تحليلات الأدلة الجنائية الشرعية الرقمية، يدمج نموذج الأدلة الجنائية الشرعية الرقمية التقنيات والمنهجيات المتقدمة في إطار منظم يوجه عملية التحقيق ويسرعها. حيث يتم تصميم هذه النماذج لتلبية المتطلبات المتطورة في البيئات الرقمية الحديثة، مثل تعقيد أجهزة إنترنت الأشياء الطرفية ومتطلبات الأساليب اللامركزية التي تحافظ على الخصوصية لجمع الأدلة وتحليلها.

مع التركيز على أجهزة إنترنت الأشياء (IoT)، تسعى هذه الأطروحة إلى إنشاء وتطبيق نموذج جديد الأدلة الجنائية الشرعية الرقمية يسمى نموذج التحليل المباشر المتزامن واللامركزي (SDLA)، والذي يتميز بخصائصه اللامركزية، والمتزامنة، وتعامله مع البيانات الغير المتجانسة، والحفاظ على الخصوصية، والوثوقية. تشمل الأهداف الأساسية تحقيق اللامركزية في مهام المعالجة المحلية والتواصل من نظير إلى نظير، وإنشاء بروتوكولات اتصال في الوقت الفعلي لتحليل الأجهزة المتزامنة، وإنشاء هيكليات شبكة عصبية يمكنها إدارة البيانات غير المتجانسة، واستخدام تقنيات التعلم الموحد للحفاظ على الخصوصية، وتمكين المعالجة المتوازية القابلة للتطوير. وبناء الثقة من خلال التحكم في الوصول وكذلك تقنية blockchain.

يتم إظهار مساهمات نموذج SDLA من خلال التحسينات في الأدلة الجنائية الشرعية الرقمية لـ Edge IoT، والتحليل المعزز لبيانات مستشعر إنترنت الأشياء، والحفاظ بشكل قوي على الأدلة. تم تطوير مجموعة بيانات محسنة لإنترنت الأشياء للتعلم الموحد غير المتطابق. بالإضافة إلى ذلك، تم إنشاء بيئة إنترنت الأشياء الواقعية لجمع البيانات في الوقت الفعلي، مما يعكس سيناريوهات إنترنت الأشياء الحقيقية من خلال استخدام ESP32. وهذا يجعل من السهل الحصول على بيانات حقيقية في الوقت الحقيقي لتحليل الأدلة الجنائية الشرعية الرقمية. علاوة على ذلك، يقدم نموذج CNN–Autocoder نظامًا مثيرًا للاهتمام في تصنيف سيناريوهات الهجوم في بيئات التعلم الموحدة غير المتطابقة. باستخدام بنيات التشفير التلقائي والشبكات العصبية التلافيفية (CNNs)، يعمل هذا النموذج على تحسين كفاءة ودقة التعرف على أنواع الهجمات الإلكترونية المختلفة وتصنيفها في مجموعات بيانات إنترنت الأشياء غير المتجانسة. وعلى هذا الأساس، يقدم مقترح نظام FedForensic بنيات خوارزمية التعلم العميق (RNN وLSTM) وتقنيات الأدلة

الجنائية الشرعية الرقمية المصممة لتحليل الأدلة الرقمية والإبلاغ عنها بكفاءة من أجهزة إنترنت الأشياء الحافة. يقوم هذا النظام بتحليل الحزم باستخدام أدوات الأدلة الجنائية الرقمية، وجمع المعلومات، ثم إنشاء تقارير مناسبة تلقائيًا للإجراءات القانونية وتسريع عملية التحليل من خلال توفير رؤى حول التهديدات والانتهاكات السيبرانية المحتملة. بالإضافة إلى ذلك، يلبي تطوير خوارزمية Alpha-FedAvg الحاجة إلى استراتيجيات تعلم موحدة قابلة للتكيف في بيئات إنترنت الأشياء المتطورة من خلال ضمان تغير توزيع البيانات وظروف الشبكة في الوقت الفعلي. وأخيرًا، يعمل تطبيق نظام CustodyChainGuardian على تعزيز مصداقية وموثوقية التحقيقات الجنائية في بيئات إنترنت الأشياء المتطورة من خلال استخدام تقنية blockchain وبنية متعددة الطبقات للتخزين والإدارة الآمنة لضمان سلامة الأدلة الرقمية وإمكانية تتبعها طوال عملية التحقيق الجنائي.

تم إثبات فعالية وكفاءة نموذج SDLA عبر مجموعة من النماذج وتكوينات النظام من خلال الاختبار والتحقق من الصحة، مما يوفر حلاً شاملاً لتحقيقات الطب الشرعي الرقمي في بيئات إنترنت الأشياء المتطورة. يعمل نموذج AlphaFed-Avg المستند إلى RNN بشكل جيد في النظام الفرعي FedForensic IoT، مما يدل على قدرته على التكيف مع أربع مجموعات بيانات من Edge-IIoTset بدقة تبلغ 93.05%. يتم زيادة الأداء إلى 99.84% عند استخدام بنية LSTM. ويحافظ النموذج على دقة تبلغ 97.45% حتى مع البيانات في الوقت الفعلي للبيئة الجديدة، مما يدل على قدرته على التكيف مع المواقف المتغيرة. بالإضافة إلى ذلك، يساهم نموذج CNN-Autocoder، بدقة تصل إلى 92.6%، في مجموعة متنوعة من تصنيف الهجمات في التحليل الجنائي لإنترنت الأشياء. في نظام CCG الفرعي، يُظهر بروتوكول لندن أقل تكلفة نهائية بين بروتوكولات الإيثريوم مما يسلط الضوء على العواقب المالية. بالإضافة إلى ذلك، من خلال دمج الثبات والشفافية والثقة والخصوصية، يضمن نموذج CCG الخاص بنا المتانة.