iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.81456

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

ONIONTRACEX: A Dark Web Intelligence Framework

Abhishek P Nair, Akhila V, Al Amaan S M, Aldrin Joshua, Sindhya Mathew
Department Computer Science and Engineering (Cybersecurity) Rajadhani Institute of Engineering and Technology Thiruvananthapuram, India

Abstract: ***Anonymous networks have enabled the growth of hidden online services that are difficult to monitor using conventional techniques. These environments host diverse activities, often involving financial transactions, illicit trade, and decentralized communication. Existing analysis methods typically focus on isolated components such as crawling or classification, which limits their ability to capture relationships between services.***
This work introduces ONIONTRACEX, a multi-layer intelligence framework designed to perform automated analysis of hidden services by integrating keyword-driven discovery, semantic content interpretation, financial transaction linkage, and temporal availability tracking. The system incorporates a transformer-based model to interpret contextual meaning from collected content, enabling identification of related services beyond the initial search query.
Experimental observations show that when a specific keyword is used as an entry point, a significant portion of the discovered services belongs to different functional categories. This indicates the presence of interconnected service ecosystems rather than isolated domains. The framework also combines multiple indicators into a unified scoring mechanism to prioritize entities based on observed behavior.
The proposed approach improves intelligence coverage and provides a structured view of hidden service interactions, supporting deeper analysis of complex online environments.
Keywords: dark web, threat intelligence, blockchain analysis, risk scoring, content classification

I. INTRODUCTION

Hidden services operating on anonymity networks present unique challenges for monitoring and analysis. Unlike surface web systems, these environments lack centralized indexing and exhibit frequent structural changes. As a result, identifying and analysing relevant services becomes difficult using traditional approaches.

Most existing solutions focus on single functionalities such as data collection or classification. However, such approaches often fail to capture the relationships between services, financial activity, and operational behaviour. In practice, hidden services are not isolated; they interact through shared infrastructure, transaction networks, and overlapping content.

This paper proposes ONIONTRACEX, a framework designed to address these limitations by combining multiple analytical components into a unified system. The framework processes data through stages of discovery, extraction, semantic interpretation, and correlation.

A key aspect of the proposed approach is the ability to expand analysis beyond initial search inputs. By incorporating semantic understanding into the workflow, the system identifies related services that may not directly match the original query, enabling broader intelligence discovery.

II. RELATED WORK

Previous studies have explored automated classification of hidden services, phishing detection, and machine learning-based analysis of dark web content. While these approaches have contributed to improving visibility into hidden networks, they are generally limited to specific tasks.

Most existing systems do not integrate financial transaction analysis, vendor relationship modeling, and temporal behavior tracking within a unified framework. As a result, their ability to provide comprehensive intelligence remains limited.

The proposed framework extends prior to work by combining multiple analytical dimensions into a cohesive system, enabling deeper insight into the structure and behavior of dark web services.

III. SYSTEM ARCHITECTUERE

The ONIONTRACEX framework is designed as a layered architecture that processes data through multiple stages. The system includes the following components:

- Keyword intelligence module
- Adaptive crawling module
- Data extraction module
- Semantic classification module
- Financial analysis module
- Vendor relationship module
- Temporal analysis module
- Risk scoring module
- Storage and visualization components

Figure 1. ONIONTRACEX multi-layer intelligence architecture

The architecture allows continuous data flow from initial discovery to final intelligence generation, enabling efficient correlation of information across different layers.

IV. METHADOLOGY

A. Keyword Based Discovery

The system begins with a set of user-defined keywords that act as entry points for the crawling process. The crawler explores hidden services by following links and expanding the search space dynamically.

B. Semantic Driven Cross Category Discovery

A key contribution of the system is its ability to identify services that are not directly related to the initial keyword. This is achieved through the integration of a transformer based classification model that analyzes the semantic content of each page.

When a keyword such as “mixer” is provided, the system does not restrict its output to mixing services alone. Instead, it identifies related services that fall under different categories such as fraud, counterfeit, and weapons.

Experimental observations indicate that approximately 40 percent of the discovered services belong to categories different from the initial query. This demonstrates that dark web services are interconnected and cannot be analyzed in isolation.

This capability improves the effectiveness of the system by expanding intelligence coverage and revealing hidden relationships between services.

C. Financial Analysis

The system extracts cryptocurrency wallet addresses from collected data and uses them to build transaction relationships. These relationships are analyzed to identify patterns such as high transaction activity and potential links between services.

D. Vendor Relationship Modeling

Services are grouped based on shared characteristics such as identifiers, structural similarities, and recurring patterns. This allows the system to identify clusters of related services that may be operated by the same entity.

E. Temporal Analysis

The availability of services is monitored over time to understand their operational behavior. This includes tracking periods of activity, inactivity, and reappearance.

V. RISK SCORING MODEL

ONIONTRACEX assigns a risk score to each dark web service by combining multiple indicators derived from different analysis layers. The score integrates financial activity, semantic classification, vendor relationships, and temporal behavior to provide a comprehensive assessment.

The composite risk score is defined as:

$$R(s) = \alpha F_s + \beta C_s + \gamma V_s + \delta L_s$$

Where each component represents a specific aspect of the service, including transaction activity, content category, network connections, and availability patterns. The weighting parameters control the influence of each factor in the final score.

This multi-dimensional approach enables more effective prioritization of services compared to single-factor methods, allowing the system to identify high-risk entities based on combined behavioral signals.

VI. IMPLEMENTATION

The system is developed using Python, with asynchronous crawling enabled through the Tor network to efficiently collect data from hidden services. Extracted data is stored in a structured database for further processing and analysis.

Machine learning models are integrated for semantic classification, allowing automated categorization of services based on content. Additional modules handle financial data extraction, vendor relationship mapping, and temporal tracking of service availability.

A web-based visualization interface is provided to present the results in the form of graphs and structured views, enabling easier interpretation of relationships and risk indicators.The system successfully demonstrates multi dimensional analysis of dark web services.

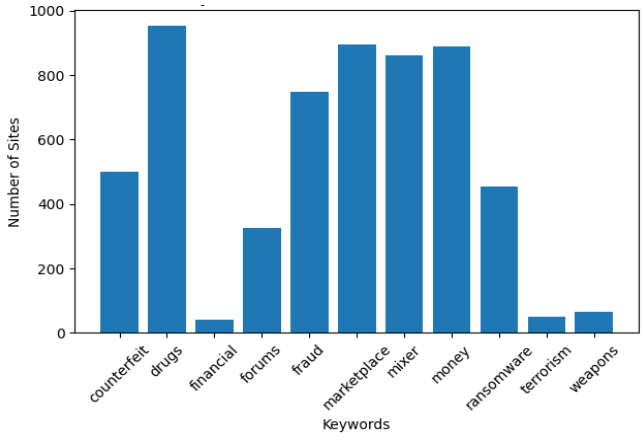


Figure 2. Keyword distribution from December 2025 to March 2026

The results show that certain categories such as drugs, marketplace, and financial related services are more prominent.

Additionally, the system identifies a significant portion of services that belong to categories different from the original keyword, confirming the effectiveness of semantic driven discovery.

This highlights the advantage of the proposed approach over traditional keyword-based systems.

VII. CONCLUSION

This paper presented ONIONTRACEX, a unified framework for dark web intelligence that integrates multiple analytical components into a single system.

The system demonstrates the ability to uncover hidden relationships between services, perform financial analysis, and track temporal behavior.

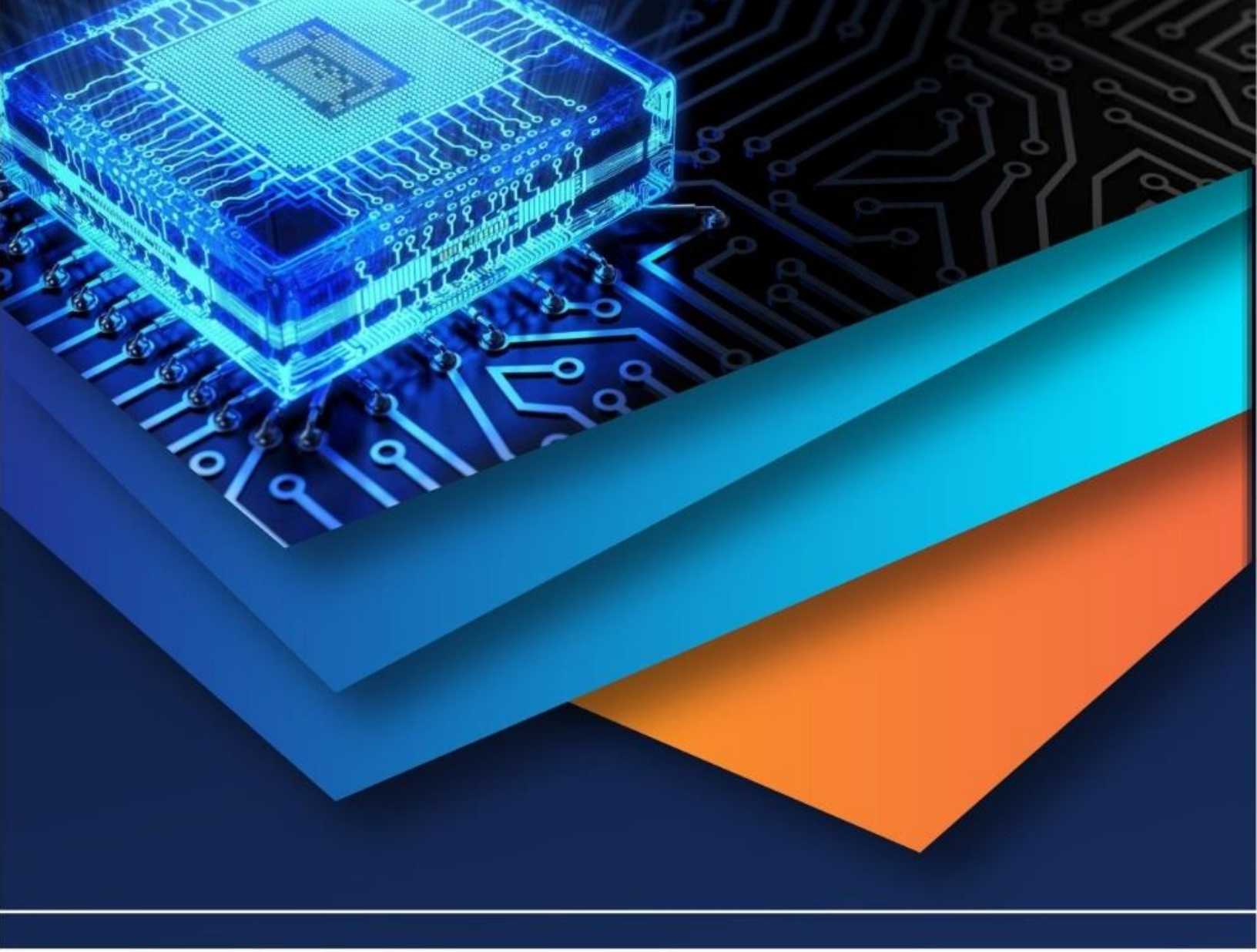
Future work includes improving real-time analysis and enhancing classification models for better accuracy.

The integration of AI-driven semantic analysis further strengthens the system's capability to identify cross-domain relationships within dark web ecosystems.

Additionally, the proposed framework provides a scalable foundation for advanced threat intelligence and forensic investigations.

REFERENCES

[1] D. Kim, Y. Park, and S. Kim, "A measurement study on Tor hidden services via keyword-based dark web collection framework," IEEE Access, vol. 11, 2023.
[2] M. Owen and N. Savage, "Empirical analysis of Tor hidden services," IET Information Security, vol. 10, no. 3, pp. 113–118, 2016.
[3] K. Soska and N. Christin, "Measuring the longitudinal evolution of the online anonymous marketplace ecosystem," in Proc. USENIX Security Symposium, 2015, pp. 33–48.
[4] M. Fleder, M. S. Kester, and S. Pillai, "Bitcoin transaction graph analysis," arXiv preprint arXiv:1502.01657, 2015.
[5] D. Ron and A. Shamir, "Quantitative analysis of the full Bitcoin transaction graph," in Proc. Financial Cryptography and Data Security, 2013, pp. 6–24.
[6] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in Proc. NAACL-HLT, 2019.
[7] Z. Yang et al., "XLM-R: Unsupervised cross-lingual representation learning at scale," arXiv preprint arXiv:1911.02116, 2019.
[8] Y. Zhao et al., "DarkBERT: A language model for the dark web," arXiv preprint, 2023.
[9] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," in Proc. USENIX Security Symposium, 2004.
[10] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.
[11] A. Biryukov, D. Khovratovich, and I. Pustogarov, "Deanonymisation of clients in Bitcoin P2P network," in Proc. ACM CCS, 2014.
[12] M. Conti, S. Kumar, C. Lal, and S. Ruj, "A survey on security and privacy issues of Bitcoin," IEEE Communications Surveys & Tutorials, vol. 20, no. 4, pp. 3416–3452, 2018.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY