A Generic Architecture for Building a Domain Name Reputation System

Carmen Ionela ROTUNĂ[1,2]*, Alexandru GHEORGHIȚĂ[1,2], Ionut SANDU[1], Mihail DUMITRACHE[1,3,4], Meda UDROIU[1,2], Dragoș SMADA[1]

[1] National Institute for Research and Development in Informatics, 8-10 Mareșal Averescu Avenue, Bucharest, 011455, Romania
carmen.rotuna@ici.ro (**Corresponding author*), ionut.sandu@ici.ro, dragos.smada@ici.ro, alexandru.gheorghita@ici.ro, meda.udroiu@ici.ro, mihail.dumitrache@ici.ro

[2] Politehnica University of Bucharest, 313 Splaiul Independenţei, Bucharest, 060042, Romania

[3] University of Bucharest, Faculty of Letters, 5-7 Edgar Quinet Street, Bucharest, 010017, Romania

[4] Academy of Romanian Scientists, 3 Ilfov Street, 030167, Bucharest, Romania

Abstract: Due to the significant increase in the number of newly registered domains (nearly 100 million in a year), and to the rise in the number of malicious domains that pose a threat to DNS servers, there is a need for an automated monitoring solution for evaluating the reputation of a domain name. This paper aims to present the architecture of an automatic monitoring platform, which could dynamically establish the reputation level for each .ro domain, and also for other domains. The primary objective of this paper is to enhance the trustworthiness of .ro domains against malicious activities in the Internet space through automatic monitoring and by offering solutions for data protection. If the reputation level of a .ro domain is determined, its future owners (including authorities, public institutions, private companies, natural persons, etc.) could have a clear understanding of the degree of trust associated with that domain, thereby creating a safer online environment.

Keywords: Domain, DNS, Generic architecture, Domain reputation system, TLD, Security, Domain Registry.

1. Introduction

In today's digitally connected world, online transactions and interactions have become an integral part of people`s daily lives. With the increase in the number of online users, the need for trustworthy and secure systems has become more important than ever before. A Domain Name Reputation System is a type of system that helps to assess the reputation of domains in order to prevent malicious activities and improve online security.

This paper proposes a generic architecture for a domain name reputation system that can be customized to suit different domain types (generic, country code domains etc.) and provide a scalable and adaptable solution for online security. This architecture uses a combination of Machine Learning algorithms and domain-specific features to assess the reputation of the analysed domains accurately. The proposed architecture was also evaluated through experiments on real-world data and its effectiveness in detecting malicious activities was demonstrated.

A domain name reputation system is designed to evaluate the reputation of a particular domain name based on various factors, such as its history, behaviour, and content. The purpose of such a system is to help identify potentially malicious or spammy domains, and to provide a measurement of trustworthiness for domains that are deemed safe and reliable.

The cornerstone of an efficient approach to mitigating threats lies in conducting threat assessments, devising strategies for mitigating them, and establishing appropriate levels of risk. (Cîrnu et al., 2018).

There are several approaches that can be implemented when designing a domain name reputation system. One common approach is to use Machine Learning algorithms to analyse the behaviour of domains and to identify patterns of behaviour that are indicative of malicious or spammy activity. This can involve analysing factors such as the frequency and types of emails sent from a domain, the age and registration details of the domain, and the content of the website associated with the domain (Banciu et al., 2016).

Another approach lies in using a collaborative filtering method, where user ratings and feedback are employed to evaluate the reputation of a domain. This approach can be particularly effective in identifying new and emerging threats, as users are often the first to detect suspicious behaviour and can provide feedback on the

https://doi.org/10.24846/v32i2y202304

quality and trustworthiness of a domain (Banciu et al., 2019).

Top-level domains (TLDs) make up the highest level of the Domain Name System (DNS) hierarchy. (Goerzen, 2004) (Dooley and Rooney, 2017). They are the last segment of a domain name, appearing after the final dot, and are used for indicating the type and purpose of a domain. There are two main types of TLDs, namely country code top-level domains (ccTLDs) that correspond to a specific country or geographic location and generic top-level domains (gTLDs) that are not associated with a specific country (.net, .com etc.) (Akumiah, 2016).

The main objectives of this paper are to identify the requirements for domain reputation system development, identify key domain Registry parameters, identify external services for domain reputation determination, develop a generic architecture for domain reputation REGRep that will be used to develop a reusable solution for determining domain reputation which could be profiled and implemented by any ccTLD Registry or its partners (Registrars).

The remainder of this paper is structured as follows. Section 2 presents a review of the main methods used so far to determine the reputation of a domain, through blacklists and whitelists, using the passive data collected. Section 3 refers to the functionality of the .ro domain registry system and to the main concepts and rules according to which domains are operated. Section 4 sets forth the methodology used in research and the requirements and operational criteria for establishing the reputation of a domain. Section 5 illustrates the generic architecture of the domain reputation system. Finally, Section 6 includes the conclusion of this paper.

2. State of the Art

Domain takeover is a major concern for organizations, as attackers may exploit them for malicious purposes such as malware hosting and credential harvesting, leading to reputational damage (OWASP, 2020). From a technical perspective, DNS operates as a hierarchical system of name servers, utilizing a globally distributed database that contains details about each domain. (Zou et al., 2016).

The relevant DNS data is stored across multiple DNS servers, allowing for quick retrieval whenever a user initiates a query. If not adequately secured, the Domain Name System (DNS) can be vulnerable to exploitation by malicious individuals. Malware perpetrators are aware of the importance of DNS accessibility and actively seek ways to disrupt DNS uptime and the servers that maintain it (Scalzo, 2017).

Currently there are several techniques that can be used for improving the accuracy and effectiveness of a domain name reputation system. These include:

- Real-time monitoring, which involves continuously monitoring the behaviour of domains in real time and updating their reputation scores accordingly. This can help to quickly identify and mitigate new threats as they emerge (Vevera et al., 2018);

- Blacklists and whitelists are lists of domains that are known to be either malicious or safe. By using these lists as a starting point, a domain name reputation system can quickly classify domains as safe or risky;

- IP reputation can help to identify malicious activity that may be associated with a particular IP address.

Overall, a state-of-the-art domain name reputation system will likely incorporate a combination of these techniques, along with advanced Machine Learning algorithms and real-time monitoring capabilities. By doing so, such a system could provide an accurate and reliable measurement of the reputation of a given domain name, thereby helping to protect users from potential threats and ensuring a safe and secure online environment.

Exposure is a system designed to detect malicious domains by analysing DNS data. It employs a large-scale, passive DNS analysis technique to identify domains that are involved in malicious activities. Exposure uses this technique to identify domains that are associated with known malware, phishing campaigns, or other malicious activities. The system works by first collecting DNS data from various sources, such as recursive DNS servers, malware analysis platforms, and threat intelligence feeds. It then uses Machine Learning algorithms to analyze this data and identify patterns that are indicative of malicious activity (Bilge et al., 2011).

Notos is a reputation system for DNS that identifies malicious domains based on their unique characteristics by using passive DNS query data and analysing the network and zone features of domains. The data is used to develop models of trustworthy domains and malicious domains. (Antonakakis et al., 2010). The system works by analysing the behaviour of domains over time and identifying patterns that are indicative of malicious activity. Notos considers a wide range of factors, including the network topology and flow of traffic, the registration and expiration dates for domains, and the use of certain DNS record types.

Kopis utilizes a passive approach to observe DNS traffic at higher levels of the DNS hierarchy. By analysing worldwide DNS query resolution patterns, it can precisely identify malicious domains. In contrast to earlier DNS reputation systems, like Notos and Exposure, which depend on monitoring traffic from local recursive DNS servers, Kopis offers a novel perspective and incorporates new traffic features that capitalize on the global visibility achieved by observing network traffic (Antonakakis et al., 2011).

Fast Flux Service Networks (FFSNs) are a type of networks used by cybercriminals to evade detection and maintain the availability of a malicious infrastructure. FFSNs are characterized by many constantly changing IP addresses, which can make them difficult to detect if traditional network analysis techniques are used. However, passive DNS analysis can be an effective way to identify FFSNs and detect malicious activities. Machine Learning algorithms can be used for detecting FFSNs based on passive DNS data. By training Machine Learning models on large datasets of known FFSNs, researchers can develop models that can accurately detect FFSNs in real-time (Caglayan et al., 2009; Lombardo et al., 2018).

The topology-based flow model proposed by Mishsky & Gal-Oz (2015) provides a promising approach to computing domain reputation. By leveraging network topology and flow analysis, the model has the potential to provide a more accurate and effective protection against malicious domains.

Risk Analytics (2023) provides information about malicious domains, and it also offers a list of domains that are known to host malware.

Fukushima et al. (2011) proposes the development of a blacklisting system with the ability to analyse the features of malicious websites using their domain information, such as Autonomous System (AS), IP address block, IP address, domain, and registrar and proposes a blacklisting scheme that combines IP address blocks and registrars with a low reputation, which are frequently utilized by attackers.

However, creating and maintaining these blacklists is difficult, as it leads to errors and omissions. To address this, a Machine Learning model based on deep neural architecture was proposed by Lison and Mavroeidis (2017) which automatically detects whether domain names and IP addresses are malicious. Because it was trained on an extensive passive DNS database, the model achieves a high detection rate of 95%. The use of Blacklisting only to determine the reputation of a domain name proves to be highly inefficient in identifying both known and newly generated malicious URLs. Moreover, it relies on human input and proves to be a time-consuming process, especially in real-time environments (Vinayakumar et al., 2018).

The complexity of identifying the classification of malicious domains can be addressed by utilizing contemporary adaptive and learning methodologies. Intelligent unsupervised classification of domains can be achieved using metaheuristic-based search algorithms like Cuckoo Search. Object-oriented engineering can be used for implementing the proposed model, with future extensions being a possibility (Sarkar et al., 2013).

MaldomDetector is a system designed to identify malicious domain names that are algorithmically generated using Machine Learning techniques. It employs a Machine Learning model that analyses various features of a domain name, such as its length, entropy, and string distribution, to determine whether it is likely to be algorithmically generated. The model is trained on a large dataset of known malicious and benign domain names to learn the patterns and characteristics of algorithmically generated domain names (Almashhadani, 2020).

Most of existing domain reputation systems do not provide a dynamic real-time monitoring ML based solution which can be used by domain ccTLDs or their registrars to scan domain names daily so as to enable the detection of compromised domains names.

The system proposed within this research paper uses historical domain information combined with new domain information from TLD Registry database with external tools for determining a reputation scoring for a .ro domain. The solution is proactive and allows, through Machine Learning techniques, the creation of a "malicious" pattern of a domain. This action enables the selection and quarantine of any suspicious domain for its rehabilitation or elimination.

In comparison with the existing domain reputation systems (Notos, Kopis, Exposure) the proposed system architecture enables real-time, dynamic domain reputation scoring.

Table 1 illustrates the features of the existing systems in comparison with the ones of the system proposed within this research paper.

Exposure, Notos, Kopis and MaldomDetector are passive, commercial domain reputation systems that are not provided by a domain Registry. On the other hand, TLDRep has the advantage of having all the necessary data sourced directly from a domain Registry, which allows it to obtain a more accurate and nuanced domain reputation score. Additionally, TLDRep offers a continuous, and real-time, active DNS analysis, which makes it a powerful tool for detecting and mitigating potential threats to network security.

TLDRep provides real-time analysis of domain data, enabling the quick identification and response to potential threats, it actively scans and analyzes DNS data to identify potential threats and vulnerabilities, it has access to a comprehensive database of TLD domain Registry data, thereby enabling a more accurate and comprehensive domain reputation scoring, it continuously scans domains for potential threats and alerts users in real time and uses Machine Learning algorithms to improve its domain reputation scoring and analysis, which makes it more effective over time. Overall, TLDRep's combination of real-time analysis, active DNS analysis, access to TLD Registry data, continuous scanning and alerting, and use of Machine Learning makes it a powerful and comprehensive tool for monitoring and protecting domain reputation.

3. The Romanian Domain Registry System

The Romanian TLD Registry manages the .ro top-level domain names for Romania. The Registry is operated by the National Institute for Research and Development in Informatics (ICI). The Registry was established in 1992 and is responsible for the management, administration, and technical operation of the .ro TLD.

Some of the key features and services offered by the Romanian TLD Registry include:

- Domain name registration: the Registry provides domain name registration services for the .ro TLDs through accredited Registrars;
- DNS management: The Registry is responsible for managing the domain name system (DNS) for the .ro TLD, including the operation and maintenance of the .ro name servers;
- Domain name dispute resolution: The Registry provides a domain name dispute resolution service for disputes related to .ro domain names;

Table 1. Features of existing domain reputation systems in comparison with those of TLDRep

System	Real-Time Analysis	Active DNS Analysis	Uses TLD Domain Registry Database	Continuous Domain Scanning and Alerting	Machine Learning use
Exposure	No	No	No	No	Yes
Notos	No	No	No	No	Yes
Kopis	No	No	No	No	Yes
MaldomDetector	No	No	No	No	Yes
TLDRep	Yes	Yes	Yes	Yes	Yes

- Domain name system security: The Registry operates a domain reputation system to ensure the security and stability of the .ro TLDs. The system monitors domain registrations, DNS queries, and other network activities to identify and prevent malicious activities;

- Collaboration: The Registry collaborates with other TLD registries, security organizations, and law enforcement agencies to share threat intelligence and coordinate the response to security incidents. (Lockheed Martin Corporation, 2019);

- Data protection: The Registry ensures that the personal data of .ro domain owners and users is protected in compliance with relevant data protection regulations.

Overall, the Romanian TLD Registry plays a critical role in managing the .ro domain names and ensuring the security and stability of the Romanian domain name system.

For a TLD name such as .ro to be accessible on the Internet, it needs to be entered into the root nameservers by IANA, the authority that approves the entry of a TLD name into root zone (IANA, 2020).

For each DNS zone there is only one delegated Registry, therefore there is only one Registry for any ccTLD or gTLD. Registries must comply with internationally adopted regulations (Stăicuț, 1995).

The natural or legal person who requests the registration of a domain can be considered the "owner" (Registrant) of the respective domain name. Registrants are the organizations or individuals that own and register domain names. For registering a domain name an applicant contacts a Registrar who, on behalf of the Registrant, interacts with the Registry that manages the database of registered domain names.

Registrars are companies, partners of the Registry, which in some cases also offer other services, such as web hosting and email services. The Registry propagates the changes to the zone database (write zone) and later it is used to populate the DNS servers and the WHOIS server.

Domain names are generally registered for several years, depending on Registry rules. At the end of a domain name's registration period, the holder (registrant) may request to renew it or it may become available for someone else to register. A domain may be transferred from one Registrant to another upon request.

An important feature of the .ro Registry is the structure of its database according to the Registry-Registrar-Registrant model. Registrars communicate with the Registry on behalf of the Registrant (domain owners). This approach ensures that each domain object that is registered in the database is owned by a Registrar (called "Designated Registrar") and no other Registrar can change it. A registrant is permitted to change the Designated Registrar of a domain through the transfer process.

Figure 1 illustrates the model adopted on a large scale by most organizations that administer domains, a model also used in the management of .ro domains. A user that registers a .ro domain becomes a Registrant, the holder of a domain name, and can update, renew, trade or transfer a domain, update nameservers, update contact data, obtain trade/transfer keys, add DNSSEC records and delete a domain. All these operations are carried out on the Registrar Portal, a partner of the .ro Registry which uses the Registry API services. The .ro Registry domain updates are propagated to the Zone Database and Nameservers.

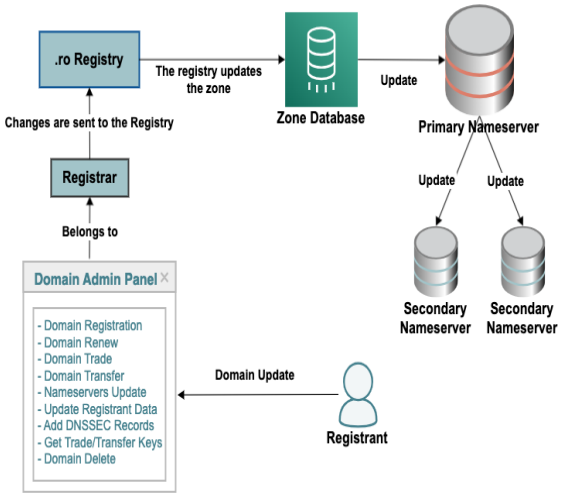

Figure 1. .ro Registry overview

4. The Requirements for a Domain Reputation System

4.1 Research Methodology

The aim of this subchapter is to describe the methodology that was followed in order to conduct this research, which consists of the following phases:

1. The exploratory phase, which comprised extensive research, with the aim of acquiring a consistent knowledge base on the process of developing the architecture of a domain reputation system. Within the exploratory phase, the existing approaches towards building the architecture of a domain reputation system were analysed from several perspectives;

2. Identify the requirements and criteria for the development of the domain reputation system architecture and solution;

3. Identification of the main architecture artifacts that were used for the development of the reputation system`s architecture and the relations between them;

4. Selection of tools that were used for developing the generic architecture model;

5. The architecture development phase included the design of the system architecture. In order to carry out this task, the output of the previous phases was used. The output could also contain gaps that would be handled during this development phase;

6. Validation phase included the validation of the architecture in a real scenario, within the scope of developing a domain reputation system that would be tested and validated on a large set of domains registered by the Romanian TLD Registry;

7. Developing the solution: The fourth step would involve developing the solution by designing, building, and testing the domain reputation system according to the architecture blueprint. This would involve selecting the appropriate technology platforms and tools, implementing the system components, and validating the system according to the requirements;

8. Deploying and maintaining the system: The final step of the methodology would be to deploy the domain reputation system to production and maintain it over time. This involves monitoring the performance of the system, addressing any issues or bugs that may arise, and updating the system as it is necessary to keep up with changing requirements and emerging threats.

Phases six to eight above could be detailed in a future research paper as the current research focuses on the development of the generic architecture for a domain name reputation system.

4.2 Requirements for Developing a Domain Reputation System

A domain reputation system for a domain Registry should meet certain requirements in order to be effective. The key requirements are:

- Accurate data collection: The system must be able to accurately collect data related to domain names, including their registration date, the IP addresses associated with them, and any historical data related to their use;

- Real-time monitoring: The system should be able to monitor domain names in real time to detect any malicious or suspicious activity as soon as possible;

- Machine Learning algorithms: A domain reputation system should use Machine Learning algorithms to analyse data and identify patterns of behaviour that may indicate fraudulent or abusive use;

- Collaborative filtering: The system should use collaborative filtering techniques to identify relationships between domains, such as those domains belonging to the same owner or being used for similar purposes;

- Threat intelligence integration: A domain reputation system should integrate with threat intelligence feeds to provide an additional context for domain name behavior;

- User-friendly interface: The system should have a user-friendly interface that allows domain registrars and law enforcement agencies to easily access and interpret the data;

- Privacy protection: The system should protect the privacy of domain registrants while still providing sufficient data for analysis;

- Timely response: The system should provide timely responses to domain registrars and

law enforcement agencies when suspicious activity is detected, so that appropriate actions could be taken to prevent abuse;

- Scalability: The system should be able to handle large volumes of data and be scalable in order to accommodate increasing numbers of domain names and users;

- Robust security: The system should feature strong security measures for protection against unauthorized access and data breaches (Holland, 2019).

4.3 Criteria for Developing a Domain Reputation System

The parameters that determine a good reputation for a domain name can vary depending on the context, the general factors that are commonly considered for TLDRep are the following:

From internal sources:

- Age of the domain: a domain that has been registered for a longer period can have a better reputation as it suggests that an organization may be more well-established and reputable;

- Owner history: a domain that has had a clean history, with no frequent transfers, is more likely to have a good reputation;

- Domain name relevance: a domain that has had a clean history without frequent transfers is more likely to have a good reputation;

- Hosts history can provide valuable information which can be explored to determine a domain`s reputation;

- Trade history can provide insight into how often the owner of the domain is changed;

- NS history provides the nameservers` change history and when nameservers are frequently changed, for example daily, this can be indicative of a compromised domain name;

- Domain Status: if the domain has one of the following statuses, that is Hold, Registrant Transfer Prohibited, Update Prohibited, Renew Prohibited or Delete Prohibited this is indicative of a domain with a bad reputation;

- Registry Data Analysts evaluate the existing parameters of a domain and user feedback, and the results would be registered as an indicator for determining the reputation of a domain;

- Policy violation shows how often a domain name has been subject to domain-name disputes which are resolved by agreement, legal action, or arbitration.

From external sources:

- Reputation management services: a domain that is listed with domain reputation services such as Google Safe Browsing or Norton Safe Web can have a good reputation as this suggests that that domain was deemed safe after being reviewed and verified by trusted third-party services;

- Absence of phishing activities: a domain that is not involved in phishing activities such as sending unsolicited emails, posting unwanted advertisements or participating in phishing scams, could maintain a good reputation;

- SSL certificate: a domain with a valid SSL certificate can indicate that the domain owner takes security seriously and cares about the privacy and security of domain users;

- DNS records: a domain with properly configured DNS records and no indications of DNS hijacking is more likely to be trusted;

- Registrar Reputation: there are registrars that also register domains with a bad reputation and if a domain is registered by one such registrar this can be indicative of the fact that the respective domain name has a bad reputation;

- Domain popularity: if a domain is frequently visited and has a high volume of traffic, this can indicate that it is trustworthy and useful;

- Domain scoring is provided by external domain scoring services, which are widely used for determining whether a domain is trustworthy or not.

TLDRep dedicated sources

Also, the TLDRep has some specific data inut sources which are used to determine the reputation of a domain, such as:

- DNS record history registered by TLDRep;

- Number of Redirects;

- Service Vulnerability evaluates the domain reputation by using External Sources of information;

- Registry Information provides domain information from the internal database records of the Registry;

- SSL use.

Figure 2 shows the data input sources for a domain name reputation system and the connections between them.

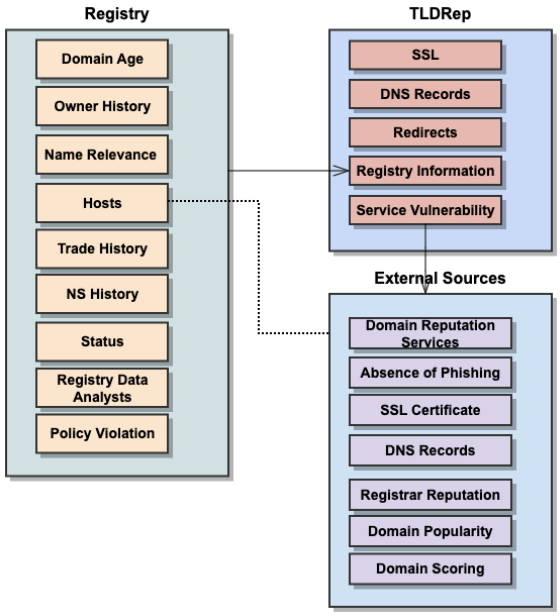

Figure 2. The criteria for establishing a domain name reputation system

5. A Generic Architecture for a Domain Name Reputation System

Overall, the architecture of a domain name reputation system involves a complex system of data collection, analysis, and feedback loops to identify and prevent spam, phishing, and other malicious activities associated with a given domain name. Some of the key features of the proposed system include:

- Continuous monitoring of domain-related criteria identified in subchapter 4.3 to detect any domains that may be associated with malicious activity;

- Machine Learning algorithms to analyse the data and identify potential compromised domains;

- Leveraging a variety of threat intelligence sources to stay up to date on the latest threats and malware and providing customizable reports that can be tailored to meet the specific needs of different domain registries or registrars.

By using the proposed generic architecture, the domain registries, registrars and other organizations can customize and implement a solution dedicated to their specific needs.

Figure 3 illustrates the generic architecture of the proposed TLDRep reputation system.

The domain reputation automated monitoring architecture is designed to provide a comprehensive analysis of a domain's reputation by collecting data from multiple sources and analysing it by using internal and external tools that provide a score as a result. In parallel, a Machine Learning algorithm will evaluate the same domain and it will try to extract patterns in malicious domains that will be further used for improving the analysis speed of the system. In order to achieve the intended objectives, this architecture includes several interconnected components. The proposed solution takes multiple domains as an input, which are then passed on to the filter. The filter contains a fast database that stores the information on whether a domain was analysed before and on the outcomes of the analysis. The filter also contains a whitelist for known, trustworthy domains that should not be included in the analysis process. The filtered domains are then passed on to an array of crawlers for data gathering. The crawler uses data collected from various sources such as Registry Data and data retrieved by Internal tools, and External tools. After the data is successfully collected, the crawler puts the domain in a processing queue.

The processing queue includes the domains and their corresponding collected meta-data that will be processed by a worker from an array of workers. The worker's job is to apply reputation algorithms to the collected data peculiar to a domain and output a result in the form of a score. The result is then placed in a result queue where it waits to be processed by the core application and stored in a permanent database. The result queue also provides data for the ML algorithms to evaluate and correlate the obtained results. In

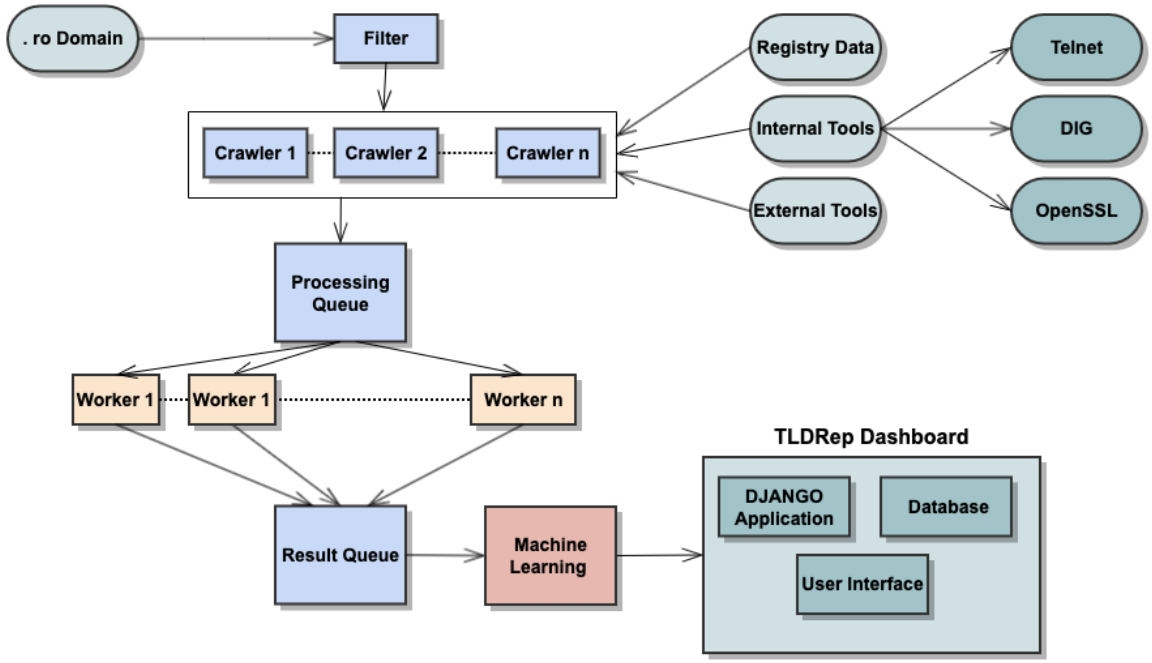

Figure 3. The generic architecture of the proposed TLDRep Domain Name Reputation System

this way, ML algorithms can be trained to better identify patterns and validate results obtained by using worker output, thus improving its accuracy in determining a malicious domain. Domain analysis results using the TLDRep solution will be stored in a database and displayed in the TLDRep Dashboard. The dashboard provides a summary of the domain's reputation and its meta-data for manual verification, including information on its history, domain age, DNS information, and other relevant factors that may impact on a domain`s reputation. Overall, the domain reputation automated monitoring solution provides a powerful tool for monitoring the reputation of domains, which can be used by domain registrars, website owners, and security professionals to identify potential threats and take appropriate actions for protecting their online assets.

This solution provides real-time monitoring, which enables users to quickly respond to any changes in the domain's reputation and take the necessary measures to mitigate any potential risks.

6. Conclusion

The purpose of this paper is to ensure that the trustworthiness of different domains is preserved in the case of malicious activities in the Internet space by automatically monitoring them and offering solutions to keep data safe. The highly dynamic nature of the domain name ecosystem and the proliferation of malicious domains that represent a real and immediate threat to the privacy and security of people and companies, are the main causes that require finding an automated solution for establishing the reputation level of a domain and further monitoring its changes throughout its lifetime.

By establishing the level of reputation of the owned domain/domains, their current and future owners (authorities, public institutions, private companies, individuals, etc.) will have a correct perspective on the trustworthiness of the owned domain, thus creating a safer internet space.

In comparison with other existing domain reputation systems (Notos, Kopis and Exposure) the proposed system architecture enables real-time, dynamic domain reputation scoring which offers the possibility to examine not only newly registered domains but also domains that are being transferred or re-registered by other individuals that might not have the same intentions as the previous owner (either good or bad).

Based on the proposed architecture, an experimental technical solution will be developed

that will act on the anonymized data set from the RoTLD domain Registry. By studying the patterns of legitimate and malicious domains, TLDRep can generate reputation scores for new domains based on its learned patterns. The generic architecture presented in this paper will be used for the detection and monitoring of malicious registered domains and for the development of a technical solution. By using the TLDRep solution which will be developed as a next step, domain name registries or registrars could gain valuable insights into potential compromised domains and take proactive steps to mitigate the emerging threats.

Acknowledgements

This research work was supported by a grant of the Romanian Ministry of Research, Innovation and Digitalization, Nucleu Programme, project code: PN 2338 02 01, project name: "Architecture – platform of an intelligent system for monitoring Internet domains by developing a dynamic reputation establishment system (TLDRep)" and by a grant of the Romanian Academy of Sciences through the competition "AOSR-TEAMS-II" - Edition 2023-2024 - "Digital transformation in the sciences", project name: "Digital transformation tools for eGovernment through the use of .ro domains".

REFERENCES

Akumiah, E. (2016) *ccTLD Best Practices*. https://www.slideshare.net/gsorkpor/cctld-best-practices [Accessed 25th February 2023].

Almashhadani, A. O., Kaiiali, M., Carlin, D. & Sezer, S. (2020) MaldomDetector: A system for detecting algorithmically generated domain names with Machine Learning. *Computers & Security*. 93, 101787. doi: 10.1016/j.cose.2020.101787.

Antonakakis, M., Perdisci, R., Dagon, D., Lee, W. & Feamster, N. (2010) Building a dynamic reputation system for DNS. In: *Proceedings of the 19th USENIX Security Symposium*, *SEC`10, 11-13 August, 2010, Washington, USA.* pp. 273-290.

Antonakakis, M., Perdisci, R., Lee, W., Vasiloglou II, N. & Dagon, D. (2011) Detecting malware domains at the upperDNS hierarchy. In: *Proceedings of the 20th USENIX Security Symposium*, *SEC`11, August 10-12, 2011, San Francisco, USA*.

Banciu, D. & Dumitrache, M. (2016) Managing a cloud-based documents system. In: *Proceedings of the 2nd International Scientific Conference SAMRO 2016 - "News, challenges and trends in management of knowledge-based organizations"*, *October 14-16, 2016, Păltiniș, Romania*. Bucharest, Technical Publishing House. pp. 13-19.

Banciu, D., Petre, I. & Dumitrache, M. (2019) Electronic system for assessing and analysing digital competences in the context of Knowledge Society. In: *Proceedings of the 11th International Conference on Electronics, Computers and Artificial Intelligence*, *ECAI 2019, June 27-29, 2019, Pitesti, Romania*. IEEE. doi:10.1109/ECAI46879.2019.9042151.

Bilge, L., Kirda, E., Kruegel, C. & Balduzzi, M. (2011) *EXPOSURE: Finding Malicious Domains Using Passive DNS Analysis*. https://sites.cs.ucsb.edu/~chris/research/doc/ndss11_exposure.pdf. [Accessed 5th February 2023].

Caglayan, A. Toothaker, M., Drapeau, D., Burke, D. & Eaton, G. (2009) Real-Time Detection of Fast Flux Service Networks. In: *Proceedings of the 2009 Cybersecurity Applications & Technology Conference for Homeland Security, CATCH 2009, March 3-4, 2009, Washington, USA.* IEEE. pp. 285-292.

Cîrnu, C. E., Rotună, C. I., Vevera, A. V. & Boncea, R.. (2018) Measures to mitigate cybersecurity risks and vulnerabilities in service-oriented architecture. *Studies in Informatics and Control*. 27(3), 359-368. doi: 10.24846/v27i3y201811.

Dooley, M. & Rooney, T. (2017) Introduction to the Domain Name System (DNS). In: Dooley, M. & Rooney, T. (eds.) *DNS Security Management*. IEEE Press Wiley, pp.17-29.

Fukushima, Y., Hori, Y. & Sakurai, K. (2011) Proactive Blacklisting for Malicious Web Sites by Reputation Evaluation Based on Domain and IP Address Registration. In: *2011 IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications, November 16-18, 2011, Changsha, China*. IEEE. pp. 352-361.

Goerzen, J. (2004). Domain Name System. In: Rhodes, B. & Goerzen, J. (eds.) *Foundations of Python Network Programming*. New York, APress Media, LLC, pp. 65-85.

Holland, B. (2019) *TLD Operator Perspective on the Changing Cyber Security Landscape*. https://www.cigionline.org/articles/tld-operator-perspective-changing-cyber-security-landscape/ [Accessed: 10th February 2023].

IANA (2020) *Root Zone Database*. https://www.iana.org/domains/root/db [Accessed: 10th February 2023].

Lison, P. & Mavroeidis, V. (2017) Neural reputation & models learned from passive DNS data. In: *IEEE International Conference on Big Data, BIG-DATA 2017, December 11-14, 2017, Boston, USA*. IEEE. pp. 3662-3671.

Lockheed Martin Corporation (2019) *Cyber Kill Chain®*. https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html [Accessed 10th February 2023].

Lombardo, P., Saeli, S., Bisio, F., Bernardi, D. & Massa, D. (2018) Fast flux service network detection via data mining on passive DNS traffic. In: *Proceedings of the 21st International Conference on Information Security*, *ISC 2018, September 9-12, 2018, Guildford, UK*. Springer. pp. 463-480. https://arxiv.org/pdf/1804.06316.pdf [Accessed 10th January 2023])

Risk Analytics. (2023) https://riskanalytics.com/ [Accessed: 17th January 2023].

Mishsky, I. & Gal-Oz, N. (2015) A Topology Based Flow Model for Computing Domain Reputation. In: Samarati, P. (ed.) *Proceedings of the 29th Annual IFIP WG11.3 Working Conference on Data and Applications Security and Privacy, DBSec 2015, July 13-15, 2015, Fairfax, USA*. Springer. pp. 277-292.

OWASP (2020) *Open Source Foundation for Application Security*. https://www.owasp.org [Accessed 28th January 2023].

Sarkar, M., Banerjee, S. & Hassanien, A. (2013) Searching DNS for malicious domain registration: identification through hybrid cuckoo search metaphor and object-oriented implementation. *International Journal of Reasoning-based Intelligent Systems*. 5(4), 280-289. doi: 10.1504/IJRIS.2013.058773.

Scalzo F. (2017) *DNS-based threats: DNS reflection and amplification attacks*. https://blog.verisign.com/security/dns-based-threats-dns-reflection-amplification-attacks/ [Accessed 17th January 2023].

Stăicuț, E. (1995) Domain Name Systems, InterNic and RIPE Procedures. In: *Proceedings of the NATO Advanced Networking Workshop, The First CENet Workshop on Networks Technology, The road to Global connectivity*, *September 15-24, 1995, Warszawa, Poland*. A CENet Publication.

Vevera, A. V. & Udroiu, A. M. (2018) Using blockchain technology in designing cybersecurity solutions. In: *Proceedings of the International Scientific Conference Strategies XXI*, *November 27-28, Bucharest, Romania*. Bucharest, "Carol I" National Defence University Publishing House. pp. 324-332.

Vinayakumar, R., Soman, K. & Prabaharan, P. (2018) Evaluating deep learning approaches to characterize and classify malicious URL's. *Journal of Intelligent & Fuzzy Systems.* 34(3), 1333-1343. doi: 10.3233/JIFS-169429.

Zou, F., Zhang, S., Pei, B., Pan, L., Li, L. & Li, J. (2016) Survey on Domain Name System Security. In: *IEEE First International Conference on Data Science in Cyberspace (DSC)*, *Changsha, China*. pp. 602-607.