Mini Project Survey Report
On

URL CONVERSION APPLICATION

Submitted by

2105340 - Vura Surya Suprathik
2105359 - Arpita Pal
2105397 - Raj Nandani

School of Computer Engineering

KALINGA INSTITUTE OF INDUSTRIAL TECHNOLOGY (KIIT)

Deemed to be University U/S 3 of UGC Act, 1956

Table of Contents

A. Initial Findings for the Domain

-

Initial Findings Regarding the Domain

Tasks Assigned	Findings
Relevant Sub - Domains	I. Cloud Based URL shortening: II. Healthcare Data Sharing with Encrypted URLs Platform (GCP) III. E-commerce Integration for URL Shortening
Issues Faced	I. Cloud Based URL shortening: A. Long urls B. increasing demand and scalability . C. to migrate existing data and ensure a smooth transition for users. II. Healthcare Data Sharing with Encrypted URLs A. conventional methods of sharing URLs B. Limited Scope of Encryption: III.E-commerce Integration for URL Shortening: A. Malware Detection B. Long and unwieldy product URLs C. Data Encryption

Proposed Solutions	I. Cloud Based URL shortening: migrate URL conversion application to the cloud, specifically leveraging Amazon Web Services (AWS) or GCP. II.Healthcare Data Sharing with Encrypted URLs The healthcare provider adopts a URL conversion application that encrypts patient data before embedding it in URLs for sharing. Patient records are encoded using strong encryption algorithms and stored securely on the server. Access to encrypted URLs is restricted to authorized users with proper authentication credentials. III.E-commerce Integration for URL Shortening: The e-commerce platform integrated a URL conversion application into its system, allowing users to generate shortened URLs directly from the platform's dashboard. The application provided analytics to track click-through rates, conversion rates, and revenue generated from each shortened link.
Tools/Methodology Used	I. Cloud Based URL SQLite database engine for the database operations, Amazon EC2, Google Cloud Storage, and Big Query II.Health care Data Sharing with Encrypted URLs Secure Multiparty Computation (SMC) , Homomorphic Encryption, Identity Access Management (IAM),Libraries/ Frameworks such as SEAL, HElib III.E-commerce Integration for URL Shortening: Data Collection , Data Analysis , Comparison and Insights , Google Analytics, Amazon Cloud Watch, or Microsoft Azure Monitor ,Agile Framework Prototype Development

Chapter 1

Introduction

In today's interconnected world, navigating the vast landscape of the internet requires a toolbox of digital skills. However, challenges arise when encountering URLs (Uniform Resource Locators) that are incompatible with our devices, limited by format restrictions, or shrouded in potential security risks. This is where URL conversion applications emerge as unsung heroes, bridging the digital divide and empowering users with greater control over their online experience.URL conversion applications offer a multifaceted approach to tackling these challenges. One key function is the ability to transform mobile app URLs into web- viewable formats. This functionality empowers users who lack access to smart phones or tablets to still access content designed primarily for mobile applications. Imagine a student researching a topic and encountering a link to an educational app unavailable on their desktop computer. A URL converter could bridge this gap, rendering the app's content viewable within a web browser, ensuring seamless access to valuable information regardless of device limitations.Beyond mobile app to web conversion, URL conversion applications tackle format incompatibility. The internet is a dynamic ecosystem, and content can be presented in various formats – flash being a prime example. Many devices lack the capability to render flash content, effectively rendering the information inaccessible. Here, URL converters step in, transforming these incompatible formats into more accessible alternatives like HTML or text, ensuring users can consume the desired content without technical barriers.Security concerns are another prevalent challenge in the digital age. Shortened URLs, often used on social media platforms, can mask malicious websites. URL conversion applications with security analysis features can address this concern. By analyzing the target URL, these tools can identify potential security risks, alerting users before they click on a link that could compromise their data or device. This proactive approach fosters a safer online environment, empowering users to navigate the web with greater confidence.The benefits of URL conversion applications extend beyond technical solutions. They contribute to increased accessibility, allowing users with limited technological expertise or outdated devices to participate fully in the digital world. Furthermore, by enhancing security, these applications empower users to make informed decisions about the websites they visit, fostering a more secure online experience

Chapter 2

Problem Statement & Objectives

Cloud Based URL shortening:

Problem Statement:

The current landscape of URL sharing often involves lengthy and cumbersome website addresses. These long URLs can be difficult to remember, share, and integrate into various communication channels like social media posts, text messages, or emails. Additionally, managing and analyzing click-through data on these long URLs can be challenging.

There is a need for a **cloud-based URL shortening service** that offers a user-friendly and efficient solution for shortening long URLs while maintaining essential functionalities.

Objective:

The objective of this project is to develop a cloud-based URL shortening service that provides the following functionalities:

1. **URL Shortening:** Users should be able to easily paste a long URL into the application and receive a significantly shortened, unique version of the original URL. Users should have the option to customize the shortened URL with a preferred alias (within reason and availability) for better branding or memorability.
2. **Click Tracking & Analytics:** The service should track clicks on shortened URLs and provide users with basic analytics on click-through rates and potentially even anonymized user data like location or device type (depending on privacy regulations).
3. **Security & Reliability:** The shortened URLs should be secure and redirect users to the intended destination without introducing malware or security risks. The cloud-based infrastructure should be reliable and offer high uptime for consistent service availability. The service should be designed to scale efficiently to accommodate a growing user base and increasing traffic volume.

Healthcare Data Sharing with Encrypted URLs

Problem Statement:

Securely sharing sensitive healthcare data is crucial in the medical field, but traditional methods often pose challenges:

- **Unsecured File Sharing:** Sharing healthcare data through email or file transfer services can be risky due to potential breaches or unauthorized access.
- **Limited Control:** Once data is shared, senders lose control over who can access it and how it's used.
- **Compliance Issues:** Data sharing needs to comply with regulations like HIPAA (Health Insurance Portability and Accountability Act) to protect patient privacy.

These issues highlight the need for a more secure and controlled method for healthcare data sharing. URL conversion applications can play a vital role in addressing this challenge.

Objective:

A URL conversion application can be leveraged to create a secure healthcare data sharing system that achieves the following objectives:

1. **Encrypted URL Generation:** The application would integrate with existing health care data storage systems. When a health care professional needs to share data, the application would encrypt the data and generate a unique, time-limited URL that points to the encrypted data. The URL would be shareable with specific recipients only. Authorization mechanisms could be implemented through password protection or integration with existing health care provider login systems.
2. **Automatic Decryption:** Upon accessing the URL with proper authorization, the recipient would be directed to a secure environment where the data can be decrypted and viewed. Once the time limit expires, the URL and access to the data would automatically deactivate.
3. **Audit Logging:** The application would maintain a log of all URL generation, access attempts, and data downloads. This log would be crucial for compliance purposes and monitoring data usage.

By integrating URL conversion functionalities with encryption and access control mechanisms, this system would offer a secure and controlled approach to healthcare data sharing.

E-commerce Integration for URL Shortening:

Problem Statement

In the context of integrating e-commerce functionality with URL shortening services in cloud computing environments, the proliferation of malware presents a significant challenge. The use of shortened URLs for promotional campaigns, product links, and user engagement exposes e-commerce platforms and their users to the risk of malware attacks. Malicious actors exploit URL shortening services to conceal harmful links, leading to potential security breaches, data theft, and financial losses. Despite advancements in cloud security measures, the dynamic nature of malware threats poses an ongoing concern for e-commerce businesses seeking to leverage URL shortening for marketing and sales activities.

Objectives:

1. **Mitigate Malware Risks**: Develop robust mechanisms to detect and prevent malware-laden URLs from being generated or disseminated through the URL shortening service, thus safeguarding e-commerce platforms and their users from malicious attacks.

2. **Enhance Security Measures:** Implement advanced security measures within the URL shortening application to fortify against malware infiltration, including input validation, URL filtering, encryption, and real-time malware scanning.

3. **Maintain Service Reliability**: Ensure the reliability and availability of the URL shortening service by minimizing downtime caused by malware-related incidents. Implement robust monitoring and response mechanisms to promptly detect and mitigate malware outbreaks.

By addressing these objectives, e-commerce businesses can effectively manage the risks associated with integrating URL shortening services in cloud computing environments, thereby safeguarding their platforms, protecting user data, and fostering trust among customers.

Chapter 3

Literature Survey

Name	Paper	Method Used	Tools Used	Explanation
Cloud Based URL shortening	Sankhala, R., Kharbanda, M., Yadav, A., Suthar, P., & Kaur, P. (2022). Sukshma - A URL Shortening Service Project. (IJCRT.ORG)].	SQLite database engine for the database operations	1. Framework 2. Database-engine 3. Data Flow Architecture	In this paper, we studied a brief overview of how the URL shortening service works. The proposed shortening has features such as, Putting the domain name in the shortened URL that gives the user trust to use the service, providing the user with the option of editing the URL when needed, maintaining the record of shortened URLs and their data.
	Berube, David. (2007). Shortening URLs with shorturl. 10.1007/978-1-4302-0193-9_30.	1.Shorturl Library 2.Integration with Shortening Services	1. Primary Tool: shorturl Library (Ruby) 2. External URL Shortening Services.	This paper likely catered to developers working with e-commerce or web applications in the late 2000s. It would have provided them with a tool (the "shorturl" library) to simplify interactions with popular URL shortening services at that time.
	Le-Khac, Nhien-An & Kechadi, Tahar. (2015). Security Threats of URL Shortening: A User's Perspective. International Journal of Advances in Computer Networks. 3. 213. 10.7763/JACN.2015.V3.169.	1. User Survey Participants 2.Survey Design	1. Survey Software 2. Data Analysis Software	This research aimed to understand how users perceive the security risks of URL shortening and their level of awareness about potential threats. The findings from the study could be valuable in raising awareness about URL shortening security and potentially informing the development of safer shortening services or user education initiatives.
Healthcare Data Sharing with Encrypted URLs	Kumar, A & Sujith, Mogalapalli & Sai, Kosuri & Rajesh, Galla & Yashwanth, Devulapalli. (2020). Secure Multiparty computation enabled E-Healthcare system with Homomorphic encryption. IOP Conference Series: Materials Science and Engineering. 981. 022	Secure Multiparty Computation (SMC) , Homomorphic Encryption	Libraries/ Frame works such as SEAL and HElib	This paper presents a promising approach for secure healthcare data sharing. While there are computational challenges, advancements in cryptography and computing power can lead to more efficient implementations in the future.

	A. Yamada, Y. Miyake, K. Takemori, A. Studer and A. Perrig, "Intrusion Detection for Encrypted Web Accesses," 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07), Niagara Falls, ON, Canada, 2007, pp. 569-576, doi: 10.1109/AINAW.2007.212.	Data Analysis Building a Baseline Machine Learning	Data Collection Software , Data Analysis Tools	The paper aims to develop a system that can detect intrusions (malicious attempts to access the web server) even when the traffic is encrypted. By analyzing characteristics like data size and timing patterns, the system can identify unusual activities that deviate from normal user behavior.
	Wang X, Zhang A, Xie X, Ye X. Secure-aware and privacy-preserving electronic health record searching in cloud environment. Int J Commun Syst. 2019; 32:e3925. https://doi.org/10.1002/dac.3925	1.Data Encryption 2.Privacy-Preserving Techniques	Cryptographic Libraries , Development Environment	The paper aims to develop a secure and privacy- preserving EHR search system for cloud environments. By using encryption and potentially additional privacy techniques, the system allows authorized users to search for relevant medical information without compromising patient data confidentiality.
E-commerce Integration for URL Shortening	Page, Sophie & Jourdan, Guy-Vincent & Bochmann, Gregor & Flood, Jason & Onut, Iosif-Viorel. (2018). Using URL shorteners to compare phishing and malware attacks. 1-13. 10.1109/ECRIME.2018.8376215.	Data Collection , Data Analysis , Comparison and Insights	E.g.,1.Data Source Leverage:API, public Data Feeds3.Data Analysis Techniques:Statistical Software, Data Visualization Tools	The paper "Using URL Shorteners to Compare Phishing and Malware Attacks" by the Anti- Phishing Working Group (APWG) investigates how attackers utilize URL shortening services differently for phishing and malware attacks.Overall, the paper sheds light on how attackers leverage URL shorteners differently for phishing and malware attacks. This knowledge can be valuable for improving cybersecurity measures and hindering malicious activities.
	Nguyen, Thi Thuy Linh, and Thi Thuy Trang Mai. "The impact of customer behavior on marketing automation in e-commerce industry." (2023).	1.Marketing automation platform data 2. Surveys and User Research	E.g., E-commerce platform analytics tools , Data analysis software	The researchers aim to understand how to use customer behavior data to improve marketing automation strategies in e-commerce. This can help online stores target their marketing efforts more effectively and increase sales.
	Tupia-Astoray, Anthony, and Laberiano Andrade-Arenas. "Implementation of an e-Commerce System for the Automation and Improvement of Commercial Management at a Business Level." International Journal of Advanced Computer Science and Applications 12.1 (2021).	1. Agile Framework 2. Prototype Development	E.g., Software Development Tools such as Programming Languages and Testing Tools.	The research aimed to demonstrate the feasibility and benefits of implementing an e- commerce system for automating and improving commercial management in businesses. By developing a prototype, the researchers can showcase how such a system can streamline processes, improve customer experience, and potentially increase sales.

Chapter 4

Methodology Used

(2105359)
Cloud Based URL shortening
Issue focused on - Long URLs
Methodology - Requirement Analysis , Design and Architecture
SQLITE database engine

Requirement Analysis:

User Needs: Identify what users expect from the URL shortening service. This could include features like easy submission of long URLs, analytic to track click-through rates, and a user-friendly interface.

Stakeholder Requirements: Consider the requirements of stakeholders such as administrators who manage the service, developers who maintain it, and potentially advertisers or partners who may integrate with it.

System Capabilities: Assess the technical capabilities of the system, including factors like scalability, security requirements, integration with other systems, and performance considerations.

Design and Architecture:

Overall Architecture: Define the high-level architecture of the system, including the main components such as the URL shortening engine, database, user interface, and analytics module.

Component Design: Detail the design of individual components, specifying their responsibilities, interfaces, and interactions with other components.

Technology Selection: Choose the appropriate technologies and frameworks for implementing each component, considering factors like scalability, security, developer familiarity, and community support.

Scalability and Performance: Design the system to be scalable and performant, ensuring that it can handle a large volume of URL shortening requests and provide timely responses to users.

Security Measures: Incorporate security measures such as input validation, encryption of sensitive data, access controls, and protection against common security threats like injection attacks and unauthorized access.

Choice of Database Engine (SQLite):

- SQLite is selected as the database engine for managing data operations within the URL shortening service.
- SQLite is a lightweight, serverless, self-contained, and zero-configuration relational database management system.
- It's well-suited for embedded systems, mobile applications, and small to medium-scale websites due to its simplicity and efficiency.

Data Storage:

Long URLs submitted by users are stored in the SQLite database along with their corresponding short URLs.
Additional metadata such as creation date, expiration date (if applicable), and usage statistics may also be stored.

Short URL Generation:

When a long URL is submitted for shortening, the SQLite database is used to generate a unique short URL.
This process may involve checking the database to ensure that the generated short URL is not already in use.

Data Retrieval:

When a user requests a short URL, the SQLite database is queried to retrieve the corresponding long URL.
If the short URL is valid and exists in the database, the corresponding long URL is retrieved and redirected to the user.

Security Considerations:

Security measures, such as input validation and sanitization, may be implemented within the SQLite database operations to prevent injection attacks and ensure data integrity.

Scalability and Performance:

The performance and scalability of the SQLite database for handling the anticipated load of URL shortening requests are evaluated and optimized as necessary.
Techniques such as indexing and query optimization may be employed to enhance performance.

(2105340)

Health care Data Sharing with Encrypted URLs

Issue focused on - conventional method of sharing URL

Methodology - Implementation and performance Evaluation , Secure Multiparty Computation (SMC) , Homomorphic Encryption

Implementation and Performance Evaluation:

Implementation builds the system using your chosen SMPC protocol and HE scheme. This involves:
Picking the programming tools and environment.
Implementing the HE scheme to encrypt medical data.
Integrating the SMPC protocol for secure computations on encrypted data.
Designing user interfaces and functionalities for patients, doctors, and the server.

Performance Evaluation measures how well the system works. This involves:
Defining how well the system performs (speed, accuracy).
Running tests with different data sizes and usage patterns.
Analyzing the results to see if the system is efficient and accurate.
Comparing it to existing systems (if applicable).

Secure Multiparty Computation (SMC):

SMC allows multiple parties to jointly compute a function over their private inputs while keeping those inputs confidential. In the context of a curl conversion application, SMC can be utilized to securely process encrypted URLs containing healthcare data.

Application:

- The curl conversion application running in the cloud acts as a computational node in the SMC protocol.
- When a request containing an encrypted URL is received, the application collaborates with other authorized parties to perform computations on the encrypted data.
- Through SMC protocols, the application ensures that computations are performed securely without revealing the plaintext healthcare data to any single party, including the cloud service provider.

Homomorphic Encryption:

Homomorphic encryption allows computations to be performed directly on encrypted data without decrypting it first. This enables privacy-preserving data processing while maintaining confidentiality. In the context of a curl conversion application, homomorphic encryption can be applied to encrypt URLs containing healthcare data.

Application:

- Before transmitting healthcare URLs to the curl conversion application, they are encrypted using homomorphic encryption techniques.
- The encrypted URLs are then processed by the application without the need for decryption.
- Computations, such as converting the encrypted URL to its corresponding plaintext form, are performed directly on the encrypted data within the cloud environment.

(2105397)

E-commerce Integration for URL Shortening

Issue focused on - Malware Detection

Methodology - Data Collection , Data Analysis , Comparison and Insights

Data Collection:

The researchers gathered a dataset of over 7,000 malicious short URLs reported as phishing or malware attacks over a two-year period (2016-2017).

Each short URL likely pointed to a longer, malicious URL that was ultimately taken down.

Data Analysis:

The researchers analyzed the short URLs in the dataset, focusing on the following aspects:

Shortener Service: Identifying the URL shortening service used to create the short link (e.g., Bitly, TinyURL). Shortener service APIs likely provided data for this analysis.

Reported Date: The date the malicious short URL was reported, potentially obtained from user reports or automated detection systems.

Click-Through Rate (CTR): This metric indicates the percentage of users who clicked on the shortened link, likely estimated through analytics provided by the URL shortening service (if available).

Unique Domains: Analyzing how many unique domains (excluding subdomains) the shortened URLs pointed to. This helps assess the attacker's strategy - using the same domain for multiple attacks might suggest a less sophisticated operation.

Comparison and Insights:

By comparing the data for phishing and malware attacks, the researchers aimed to identify potential behavioral patterns.

For example, they might have investigated:

Timeliness: Do phishing attacks have a shorter lifespan (time between creation and report) compared to malware attacks?

Click Rates: Are phishing attacks more successful in attracting clicks compared to malware attacks?

Domain Reuse: Do malware attackers tend to reuse the same domains more frequently than phishing attackers?

Overall, the paper utilizes click analytics data from URL shortening services to compare the lifecycles and user engagement patterns of phishing and malware attacks. This approach provides valuable insights into attacker behavior and can potentially aid in developing better detection and prevention strategies.

Chapter 6

Conclusion

Cloud computing plays a significant role in several aspects of e-commerce integration, URL shortening, and healthcare data sharing with encrypted URLs. Cloud computing offers on-demand scalability, allowing e-commerce platforms to handle surges in traffic during peak seasons or marketing campaigns. This ensures smooth operation and prevents website crashes due to high load. E-commerce businesses can avoid the upfront costs of purchasing and maintaining physical servers by utilizing cloud resources. They only pay for the resources they use, making it a flexible and cost-efficient solution. Cloud-based URL shortening services can be accessed from anywhere with an internet connection, making them ideal for e-commerce businesses with a global audience.Cloud infrastructure ensures high availability of the URL shortening service. Even if one server experiences an issue, others can take over, minimizing downtime and ensuring users can access shortened URLs consistent Cloud platforms offer data analytics tools that can track click-through rates on shortened URLs. This valuable data helps e-commerce businesses understand marketing campaign effectiveness and optimize their strategies. Cloud providers offer robust security features like encryption and access controls, crucial for safeguarding sensitive healthcare data. Encryption protects data at rest and in transit, while access controls ensure only authorized personnel can view patient information.Cloud platforms can help healthcare organizations comply with regulations like HIPAA (Health Insurance Portability and Accountability Act) by providing features that ensure data privacy and security. Cloud storage facilitates collaboration among healthcare professionals by allowing them to securely access and share encrypted patient data from any location. This improves patient care coordination and treatment planning.Cloud computing is a major driving force behind advancements in e-commerce integration, URL shortening, and healthcare data sharing. It empowers businesses with scalability, cost-efficiency, and robust security features, ultimately contributing to a more streamlined and secure digital experience for users across various sectors.

References

1. Hu, Mingqi, Yanli Ren, and Cien Chen. 2023. "Privacy-Preserving Medical Data-Sharing System with Symmetric Encryption Based on Blockchain" Symmetry 15, no. 5: 1010. https://doi.org/10.3390/sym15051010

2. Jamshidi, Pooyan & Ahmad, Aakash & Pahl, Claus. (2014). Cloud Migration Research: A Systematic Review. IEEE Transactions on Cloud Computing. 1. 142 - 157. 10.1109/TCC.2013.10.

3. https://www.sciencedirect.com/topics/computer-science/cloud-migration

4. Kumar, A & Sujith, Mogalapalli & Sai, Kosuri & Rajesh, Galla & Yashwanth, Devulapalli. (2020). Secure Multiparty computation enabled E-Healthcare system with Homomorphic encryption. IOP Conference Series: Materials Science and Engineering. 981. 022079. 10.1088/1757-899X/981/2/022079.

5. Page, Sophie & Jourdan, Guy-Vincent & Bochmann, Gregor & Flood, Jason & Onut, Iosif-Viorel. (2018). Using URL shorteners to compare phishing and malware attacks. 1-13. 10.1109/ECRIME.2018.8376215.

6. Sankhala, R., Kharbanda, M., Yadav, A., Suthar, P., & Kaur, P. (2022). Sukshma - A URL Shortening Service Project. [INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS - IJCRT (IJCRT.ORG)].

7. Berube, David. (2007). Shortening URLs with shorturl. 10.1007/978-1-4302-0193-9_30.

8. Le-Khac, Nhien-An & Kechadi, Tahar. (2015). Security Threats of URL Shortening: A User's Perspective. International Journal of Advances in Computer Networks. 3. 213. 10.7763/JACN.2015.V3.169.

9. A. Yamada, Y. Miyake, K. Takemori, A. Studer and A. Perrig, "Intrusion Detection for Encrypted Web Accesses," 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07), Niagara Falls, ON, Canada, 2007, pp. 569-576, doi: 10.1109/AINAW.2007.212.

10. Wang X, Zhang A, Xie X, Ye X. Secure-aware and privacy-preserving electronic health record searching in cloud environment. Int J Commun Syst. 2019; 32:e3925. https://doi.org/10.1002/dac.3925

11. Nguyen, Thi Thuy Linh, and Thi Thuy Trang Mai. "The impact of customer behavior on marketing automation in e-commerce industry." (2023).

12. Tupia-Astoray, Anthony, and Laberiano Andrade-Arenas. "Implementation of an e-Commerce System for the Automation and Improvement of Commercial Management at a Business Level." International Journal of Advanced Computer Science and Applications 12.1 (2021).