URL shortener for web consumption: an extensive and impressive security algorithm

Saikat Gochhait[1], Yogesh Singh Rathore[2], Irina Leonova[3], Mahima Shanker Pandey[2],
Bal Krishna Saraswat[4], Santosh Kumar Maurya[5], Hare Ram Singh[6], Nidhi Bansal[4]

[1]Symbiosis Institute of Digital and Telecom Management, Constituent of Symbiosis International Deemed University, Pune, India
[2]Computer Science and Application Department, School of Engineering and Technology, Sharda University, Greater Noida, Uttar Pradesh, India
[3]Faculty of Social Sciences, Lobachevsky University, Nizhny Novgorod, Russia
[4]Computer Science and Engineering Department, School of Engineering and Technology, Sharda University, Greater Noida, Uttar Pradesh, India
[5]Computer Science and Engineering Department, School of Engineering and Technology, IIMT University Meerut, Uttar Pradesh, India
[6]Computer Science and Engineering Department, GNIOT, Greater Noida, Uttar Pradesh, India

Article Info

ABSTRACT

URL stands for uniform resource locator are the addresses of the unique resources on the internet. We all need URLs to access any type of resource on the internet, such as any web page, and document. Sometimes URLs can be long, irrelative and unattractive and unable to send sometimes via email. So, for this, we proposed a URL shortener web application based on the Python-Django platform which is fast and makes your long URLs in the shortest form which you can share on social media platforms. It makes all the messy, unattractive URLs short and shareable. Writing paper proposed a premium section in our application that gives access to the customizable URLs and analytics of your shorten URLs. Customizable URLs are the URLs you create by your own keywords. By creating a premium profile with the application, you can create your own URLs by using your own keywords. We have considered security a major part of the application that prevents the short URLs from being hacked or redirected to any advertising website or content. We store all the data related to the URL to show you the best view of your analytics and update it regularly. Main contribution in this field that for web application that provides users with a fast, secure and shortest URL for their using long URLs. Comparatively to other services that are currently available, the application provides superior security, availability, and confidentiality.

Corresponding Author:

Mahima Shanker Pandey
Computer Science and Application Department, School of Engineering and Technology, Sharda University
Greater Noida, Uttar Pradesh, India
Email: mahimashanker@gmail.com

1. INTRODUCTION

With the vast database for the uniform resource locator (URL) links has been situated within the predefined storages, complex to fetch the right one within one click by first attempt. If an organization uses very long and complex URLs, then their customers are likely to be confused by the link they are given in the electronic communication. Since social media platforms like Twitter require 140-character messages, URL compression becomes increasingly important. Our project offers URL shortening services for converting long URLs into shorter ones [1]-[3]. When the user wants to shorten the long URL, they need to submit the long

URL which actually casues of the main problem for accessing the exact solution for the particular URL then ofcourse user will get the a short URL (by proposing the method) that hardly exceeds 25 characters. The user can then include that short URL instead of the long URL in Twitter messages or in an email. The basic need: URL shorteners are useful for a number of things, including redirect management, customization, tracking and analytics, easier sharing, enhanced readability, and space conservation. Example:-long URL:- "https://www.flipkart.com/tyy/4io/~cs-b6vm0rg6g5/pr?sid=tyy,4io&collection-tab-name=Samsung+Galaxy+ F12&param=5987&offer=nb:mp:09c4a20e25&fm=neo%2Fmerchandising&iid=M_0349fea5-ee26-4173- b80d-deb9df9d4b87_4.LY3JEX76VHSY&ssid=5gw1rj3kdc0000001635655861688&otracker=hp_omu_Top %2BOffers_3_4.dealCard.OMU_LY3JEX76VHSY_4&otracker1=hp_omu_PINNED_neo%2Fmerchandisin g_Top%2BOffers_NA_dealCard_cc_3_NA_view-all_4&cid=LY3JEX76VHSY".   After   applying   the proposal: short URL:-"http://gitb.in/lrAYUk". By utilizing blackboard's architecture, we propose a secure URL shortening service which takes a long web address and creates a shorter one that won't impede posting. As a result, the proposed secure URL service will allow the user to trust the service more, knowing it will be available, secure and trustworthy.

Around the world, we have a tendency to the net continues to alter the method we connect with others, organize things and share information. The web is changing into a lot of and more necessary in our daily lives because it contains a larger impact on individual customers and huge powers [4]-[6]. In 2019, the amount of net users worldwide was 3.97 billion. This suggests that nowadays quite half the world' population is connected to the globe wide web. However, whereas the digital population is visibly increasing in several components of the world, net access and convenience will vary wide from region to region. On-line population around the world as of 2020, China remains the globe' leading on-line market, followed by Republic of India and also the United States. Overall, East Asia is that the region with the foremost net users within the world, associate degreed geographic area is the region with the very best net penetration in the world. Compared to countries like Iceland and Denmark, wherever internet penetration is on the brink of 100%, China' online penetration remains comparatively low at 58%. This black eye is explained by the slow progress of digital infrastructure in remote areas. This is often an obstacle that's still seen in several components of the world. Mobile devices accelerate digital connections whereas international net access is on the rise, property quality is superb in countries wherever infrastructure is being developed. As of 2020, Singapore has the quickest average fastened broadband internet speed within the world, while Republic of Korea has the fastest average cellular speed. The mobile internet has become a lot of and more common and popular in recent years as smartphones became more accessible and cheaper than ever before. As net users' bit by bit switch to mobile devices to surf the web on the go, the mobile net currently accounts for quite 50% of all web traffic worldwide. Net use and favorite on-line activities because the net became an essential tool for information, communication and entertainment, the typical daily net time per person is increasing year by year. In 2019, users spent a calculable one hundred seventy minutes online, totally on their smartphones per day. The foremost common mobile internet activities embrace the employment of instant electronic communication services, video streaming platforms and social media. On average, net users pay quite a hundred and forty minutes on a daily basis on social media, and Facebook remains the world' leading social network. The increase of e-commerce and digital recognition round the world Over the last few decades, we have a tendency to be the net has conjointly modified the method we obtain associate degreed sell merchandise round the world. As lookers shop a lot of and more on-line, international e-commerce sales in 2020 exceeded $ 4.2 billion [7]-[10]. With more and more online transactions being processed via mobile devices, e-commerce is changing into a progressively profitable industry. However as digital attainment grows, so will online privacy, fraud, and hacking concerns, shown by Figure 1.

2.   METHOD

By using Gitb.in, enter your long URL in the input field and click on generate button after that click by the algorithm entered URL will be verified if the URL is invalid in any form such as there is missing https before the URL or missing domain name (.com, .in, and .Edu) then it will show you error with it is type of error and ask you to re-enter valid URL and if it is a valid URL then it will assign a unique and random keyword to that particular URL and store it in the database and redirect you to a new page where you will get the short URL for your entered URL with the copy button just right there so that you simply copy that shortened URL by clicking on the copy button [11]-[14]. Then when you enter it in the URL bar it fetches the original URL from the database table by that assigned keyword to your URL and fast redirect it to that original URL. Example:-http://<host>/<path>. A short URL is uniquely identified by the string *path>, where *host> is the USS's hostname.
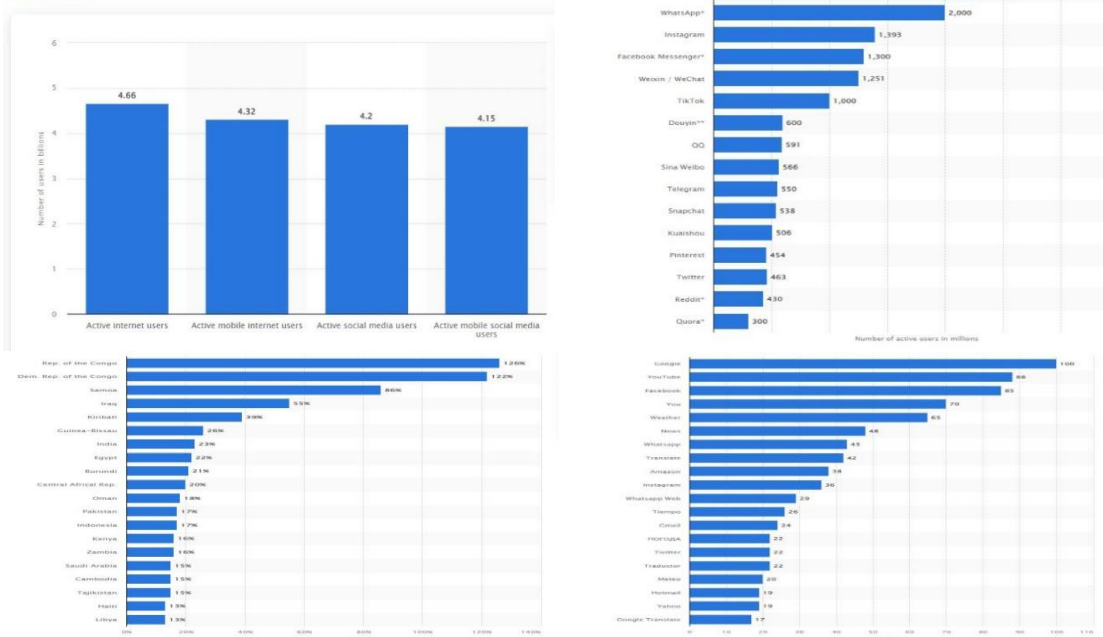
Figure 1. An analysis of survey data on mobile and social media internet usage

2.1. Algorithm

As new contribution, the proposed steps are following with the valid words are here:

```
−  Step 1: get the long URL. Entered by the user
   Original URL=GET ["original URL"]
−  Step 2: validation of the URL
   if (original URL==valid)
   {
   Action=go to the step 3
   }         else
   {
   Action=redirect to the home page with errors
   }
−  Step 3: get the original URL and assign a keyword and save both
   in a table in database
   "https://mail.google.com/mail/u/0/#inbox" k4y5ft
−  Step 4: redirect to the shortened page and give shortened URL as
   output.
   "https://www.gitb.in/k4y5ft"
   After entering this URL bar.
−  Step 5: fetch the original URL using keyword k4y5ft from the
   database.
   Original URL: "https://mail.google.com/mail/u/0/#inbox"
−  Step 6: redirect to the URL.
−  Step 7: STOP
```

2.1.1. Experimental setup

URL shortener with Python tool. Through this you'll have a fully functional API-driven web application that creates shortened URLs that forward to target URLs. Also Figure 2 is showing the communication. Some functionalities of this: i) create a FastAPI that creates shortened URL; ii) run this on web server; and iii) investigate security of the auto-generated URL.

To elaborate the proposed process, the original link address is brought up and some new letters are asked to be applied to the predefined words of the alphabet [15]-[18]. The proposed algorithm can generate new identifiers from a self-designed database. And finally, a new keyword is created and it can do all the activities related to that site without any threat. Also Figure 3 is showing the workflow of proposed model.
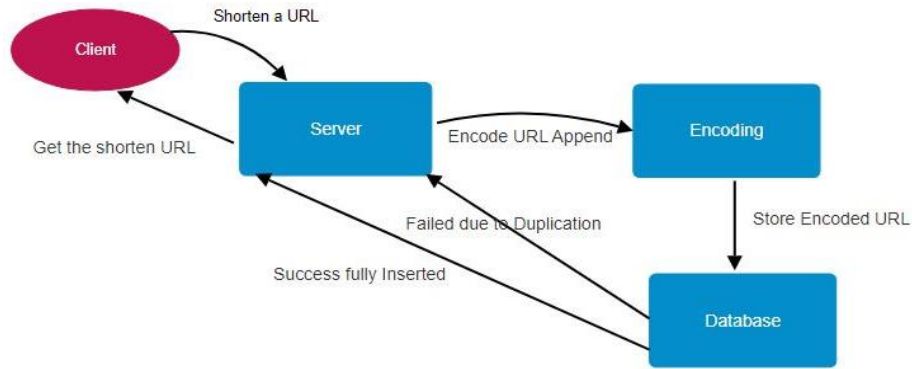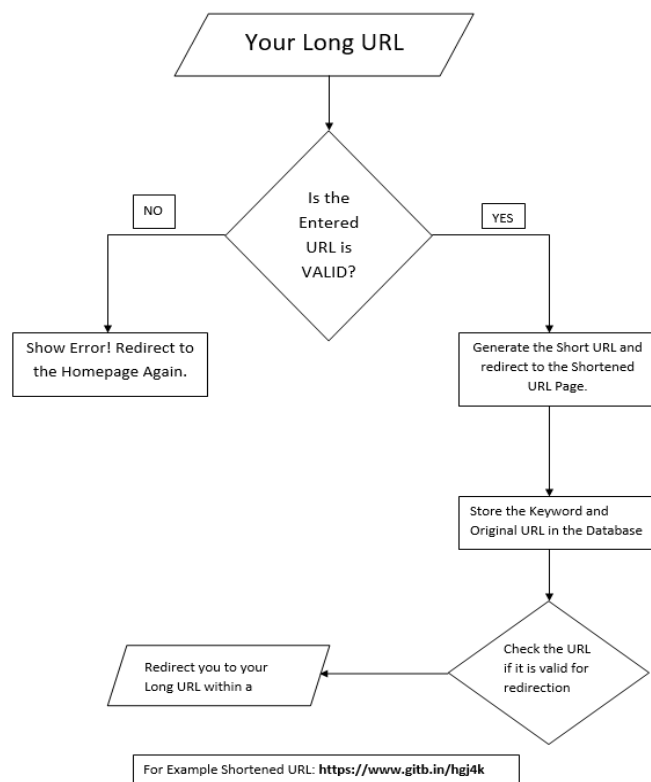
Figure 2. Communication with different stages

Figure 3. Workflow

2.2. Test cases

For critical discussion, test cases have been shown within the paper. By applying the proposed algorithm, many weblinks have got shortened link and without any error as shown in Table 1. The new shortened link is performing well within the designed statistics.

Table 1. Shortened link of test cases

Original link	Shorten link
"http://bit.ly/accenture-campus-hiring"	"http://gitb.in/kpzOTv"
"https://www.javatpoint.com/splunk"	"http://gitb.in/rsgRFu"
"https://www.javatpoint.com/computer-graphics-flood-fill-algorithm"	"http://gitb.in/qweOTo"
"https://www.sarkariresult.com/"	"http://gitb.in/iopGBd"
"https://www.flipkart.com/"	"http://gitb.in/vefJKg"

3. RESULTS AND DISCUSSION

Some major findings: the Table 2 has been designed by getting some good gaps and crisp outcome from many valued publications. Unsolved problems for the user's perspective listed here: i) tracking and analytics feature has been accommodated by introducing the shorten keywords within the specified link and location into the website; ii) improve readability: easy to read the content for a particular website; iii) space saving: URL's keywords are getting less space into the memory now; iv) easy sharing also possible for effective communication; v) redirect management for good accebility; and vi) customization is also benefitted from the proposal [19]. A glimpse of the technology implemented showing the shortener identifier. Also, the IP address for the origin identifier remains unchanged, shown by Figure 4.

Table 2. Major finding

| S. No.	Research methods	Major findings
[20]	*Analyzed code of popular URL shorteners. *Performed black-box testing. *Evaluated security vulnerabilities.	*Most shorteners used encryption for shortened URLs, but some had weaknesses. *Some shorteners were vulnerable to phishing attacks.
[21]	*Literature review of existing threats and defenses. *Proposed a taxonomy of URL shortener vulnerabilities. *Evaluated existing countermeasures.	*Identified various threats like phishing, malware, and tracking. *Proposed defense mechanisms like stricter domain validation and user education.
[22]	*Analyzed privacy policies of popular URL shorteners. *Performed privacy testing experiments. *Compared data collection practices.	*Many shorteners collected user data beyond URL shortening needs. *Data collection practices varied significantly across services.
[23]	*Proposed cryptographic techniques for privacy-preserving shortening. *Implemented and evaluated a prototype system. *Analysed performance and security trade-offs.	*Cryptographic techniques can enable anonymized URL shortening. *Trade-offs exist between privacy and performance.
[24]	*Analyzed real-world traffic through URL shorteners. *Identified malicious redirects and tracking practices. *Evaluated the impact on user privacy and security.	*Malicious actors exploited shorteners for phishing and malware distribution. *Shorteners often used opaque tracking mechanisms.
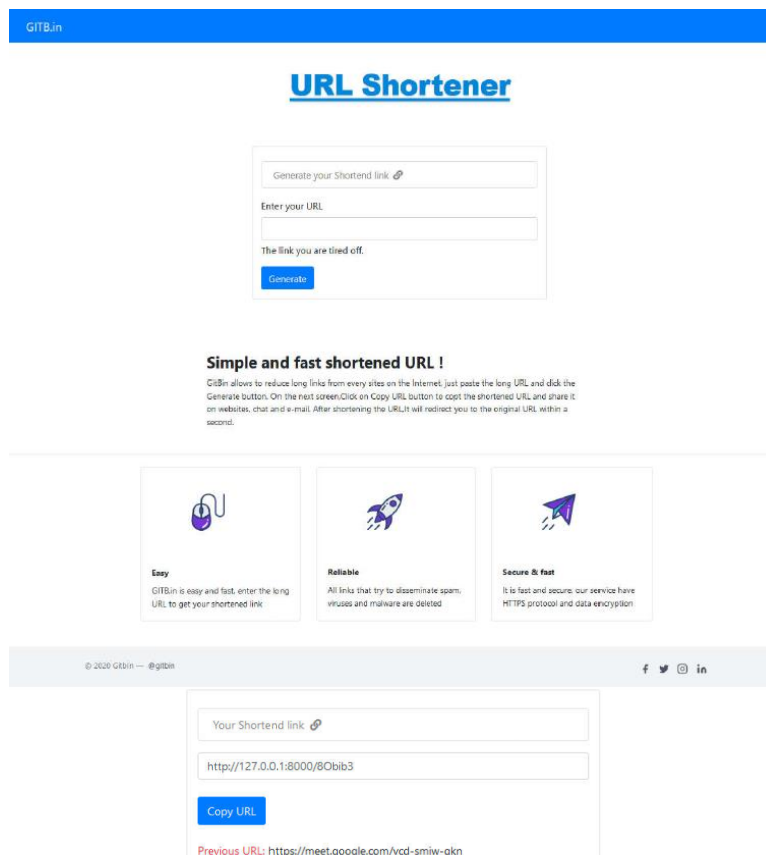
Figure 4. Outcome by proposed methodology

3.1. Several URL shortening services

All the URL shortening services listed in the Table 3 follows structure: "https://<USS_name>/<random_path>". Where, this is a URL shortening service called "USS_name>" "<random_path>" is the random character produced by the service [25]. Table 4 is able to show some advancement into existing methodology by introducing some services that has been specified.

Table 3. Several URL shortening services

URL shortening services	URL shortening
"bit.ly"	"https://bit.ly/Ufy0c2"
"tinyurl.com"	"https://tinyurl.com/8y5tsdx"
"goo.gl"	"https://goo.gl/J13Lb"
"mcafe.ee"	"https://mcaf.ee/93ojk"

Table 4. Optimized services

Feature	Proposed application	Other services (specify which ones)
Platform	Python-Django	(Examples: Bitly, TinyURL, Rebrandly)
Customizable URLs	Yes	Yes/No
Analytics	Yes (details on provided analytics)	Yes/No (details on what they offer)
Security measures	(Describe your security features)	(Compare theirs to yours)
Data storage and privacy	(Explain how data is stored and privacy ensured)	(Compare theirs to yours)
Availability and uptime	(Describe your uptime guarantee)	(Compare theirs to yours)
Cost	Free/premium (details on each tier)	Free/premium (details on theirs)

4. CONCLUSION

In software development every software or and application has future scope or developer has the future for the software which are release once software is commercially active. Scope is defined features and functions of a product. Scope involves getting information required to start the project, including the features to start a project, including the features the product needs to meet its requirement. The future scope of the gitb.in is to add the premium service in the project which consist of some paid membership at minimum cost (at most benefit likely to be adhere in future). Premium include features like self-structure shorten links or in simple words it can be customizable shorten URL can be created, a greater number of URL can be shorten using the website application. The website consists many updates which will be reflected as per future requirement and response of the website in beta testing. In addition to security, availability, and confidentiality consideration the proposed URL shortening structure has the following features: By putting the original host name in the shorten URL, users will be more likely to trust the service, and when the shorten URL is attacked, the user can still access the original site and its original web page. In previous services, the shortening process was unknown and random. The available URL shortening services replace the entire URL with a new short URL which bears no relationship to the original URL. Using our proposed service, we replace the entire URL with some characters from the original one.

REFERENCES

[1]   L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "Web caching and zipf-like distributions: evidence and implications," in *Proceedings-IEEE INFOCOM*, 1999, vol. 1, pp. 126–134, doi: 10.1109/INFCOM.1999.749260.

[2]   M. Cha, H. Kwak, P. Rodriguez, Y. Y. Ahnt, and S. Moon, "I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system," in *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, 2007, pp. 1–14, doi: 10.1145/1298306.1298309.

[3]   M. Cha, A. Mislove, and K. P. Gummadi, "A measurement-driven analysis of information propagation in the Flickr social network," in *WWW'09-Proceedings of the 18th International World Wide Web Conference*, 2009, pp. 721–730, doi: 10.1145/1526709.1526806.

[4]   H. Chun, Y. Y. Ahn, H. Kwak, S. Moon, Y. H. Eom, and H. Jeong, "Comparison of online social relations in terms of volume vs. interaction: a case study of cyworld," in *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, 2008, pp. 57–69, doi: 10.1145/1452520.1452528.

[5]   M. E. Crovella and A. Bestavros, "Self-similarity in world wide web traffic evidence and possible causes," *Performance Evaluation Review*, vol. 24, no. 1, pp. 160–169, 1996, doi: 10.1145/233008.233038.

[6]   C. Grier, K. Thomas, V. Paxson, and M. Zhang, "@Spam: the underground on 140 characters or less," in *Proceedings of the ACM Conference on Computer and Communications Security*, 2010, pp. 27–37, doi: 10.1145/1866307.1866311.

[7]   Shantanu, B. Janet, and R. J. A. Kumar, "Malicious URL detection: a comparative study," in *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, Mar. 2021, pp. 1147–1151, doi: 10.1109/ICAIS50930.2021.9396014.

[8] V. Geetha and P. V. Kallapur, "Web security: research challenges and open issues," *Advances in Computer, Communication, Control and Automation*, pp. 397–404, 2012, doi: 10.1007/978-3-642-25541-0_51.

[9] R. Kumar, J. Novak, and A. Tomkins, "Structure andevolution of online social networks," in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2006, pp. 611–617.

[10] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a social network or a news media?," in *Proceedings of the 19th International Conference on World Wide Web, WWW '10*, 2010, pp. 591–600, doi: 10.1145/1772690.1772751.

[11] K. Lerman and R. Ghosh, "Information contagion: an empirical study of the spread of news on Digg and Twitter social networks," in *ICWSM 2010-Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, 2010, pp. 90–97, doi: 10.1609/icwsm.v4i1.14021.

[12] J. Leskovec, L. A. Adamic, and B. A. Huberman, "The dynamics of viral marketing," *ACM Transactions on the Web*, vol. 1, no. 1, 2007, doi: 10.1145/1232722.1232727.

[13] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, pp. 29–42, 2007.

[14] S. Yaqub, S. Gochhait, H. A. H. Khalid, S. N. Bukhari, A. Yaqub and M. Abubakr, "WhatsApp Chat Analysis: Unveiling Insights through Data Processing and Visualization Techniques," *2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS), Manama, Bahrain,* 2024, pp. 862-865, doi: 10.1109/ICETSIS61505.2024.10459604.

[15] S. Gochhait and M. Paliwal, "Green Internet of Things (GIoT) Based Smart Security Surveillance Cobot," *2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India*, 2023, pp. 1517-1521, doi: 10.1109/ICIRCA57980.2023.10220598.

[16] N. Bansal and A. K. Singh, "Development of a new task scheduling in cloud computing," *International Journal of System Assurance Engineering and Management*, vol. 14, no. 6, pp. 2267–2275, Dec. 2023, doi: 10.1007/s13198-023-02068-y.

[17] G. Kumar, N. Bansal, M. K. Sharma, and S. Chawla, "Actual effects of cloud computing technology," in *2023 International Conference on Disruptive Technologies (ICDT)*, May 2023, pp. 424–427, doi: 10.1109/ICDT57929.2023.10150630.

[18] M. S. Pandey and S. Agarwal, "Computer assisted automated detection of knee osteoarthritis using X-ray images," *Journal International Conference on Electrical, Electronics, Engineering Trends, Communication, Optimization and Sciences (EEECOS)*, vol. 1, pp. 1–6, 2015.

[19] M. S. Pandey, S. Soam, and S. Tripathi, "A framework for mathematical methods in medical image processing," *Imaging*, 2017.

[20] N.-A. L.-Khac and M. T. Kechadi, "Security threats of URL shortening: a user's perspective," *Journal of Advances in Computer Networks*, vol. 3, no. 3, pp. 213–219, 2015, doi: 10.7763/JACN.2015.V3.169.

[21] T. Bakhshi, "Securing wireless software defined networks: appraising threats, defenses and research challenges," in *2018 International Conference on Advancements in Computational Sciences (ICACS)*, Feb. 2018, pp. 1–6, doi: 10.1109/ICACS.2018.8333279.

[22] C. Pilton, S. Faily, and J. Henriksen-Bulmer, "Evaluating privacy-determining user privacy expectations on the web," *Computers and Security*, vol. 105, Jun. 2021, doi: 10.1016/j.cose.2021.102241.

[23] D. Demmler, "Towards practical privacy-preserving protocols," *it-Information Technology*, vol. 64, no. 1–2, pp. 49–53, Apr. 2022, doi: 10.1515/itit-2022-0005.

[24] N. A. ALfouzan and N. C, "A systematic approach for malware URL recognition," in *2022 2nd International Conference on Computing and Information Technology (ICCIT)*, Jan. 2022, pp. 325–329, doi: 10.1109/ICCIT52419.2022.9711614.

[25] N. Bansal, A. Jain, M. K. Sharma, and G. Kumar, "A novel matrix for analyzing cloud services in top MNCs," in *2023 7th International Conference on Trends in Electronics and Informatics (ICOEI)*, Apr. 2023, pp. 704–708, doi: 10.1109/ICOEI56765.2023.10125916.

BIOGRAPHIES OF AUTHORS

Saikat Gochhait ⓘ 🆂 🆂🅲 ⊘ teaches at Symbiosis Institute of Digital and Telecom Management, Constituent of Symbiosis International Deemed University, Pune, India and Neurosciences Research Institute-Samara State Medical University, Russia. He is Ph.D. and Post-Doctoral Fellow from the UEx, Spain and National Dong Hwa University, Taiwan. He was Awarded DITA and MOFA Fellowship in 2017 and 2018. His research publication more than 60+ with foreign authors indexed in Scopus, ABS, ABDC, and Web of Science. He is a senior IEEE member. He was awarded the Doctoral Honoris Causa by ITUEM, Mexico in 2023. He can be contacted at email: saikat.gochhait@sidtm.edu.in.

Yogesh Singh Rathore ⓘ 🆂 🆂🅲 ⊘ resident of Noida, Uttar Pradesh. He is a graduate of Kurukshetra University, Kurukshetra where he earned a Bachelor's Degree in Science with Computer Science as a vocational subject. As his enthusiasm grew in computer applications, he met post-graduation in computer applications from Gurukul Kangri Vishwavidyalya, Hardwar, where he earned his first Master's Degree in computer applications. After a short span of teaching computer science, he indulges more in computer technology and completed his Master's in Technology from Kurukshetra University, Kurukshetra thereafter he earned his Ph.D. from Mewar University, Rajasthan After that, he became a full-fledged educator and chooses this field as his profession. He can be contacted at email: yogesh.rathore@sharda.ac.in.

Irina Leonova ⓘ 📇 SC ⊙ Faculty of Lobachevsky State University of Nizhni Novgorod, Social Sciences Department, Russia. She has published many papers with indexed by Scopus and WoS. She can be contacted at email: irina.leonova@unn.ru.

Mahima Shanker Pandey ⓘ 📇 SC ⊙ has received his Ph.D. from Dr. A.P.J.A.K.T.U Lucknow. He is working as an Assistant Professor in the Department of Computer Science and Applications in Sharda University, Greater Noida, India. He has published many papers in various international journals and reputed conferences. He has also organized many workshops. He can be contacted at email: mahimashanker@gmail.com.

Bal Krishna Saraswat ⓘ 📇 SC ⊙ is a dynamic professional with a multifaceted background in academia, sales, and marketing. With nearly 10 years of experience in academia and 5 years in sales and marketing operations, Saraswat has demonstrated a keen ability to excel in diverse fields. His journey began with a solid educational foundation, culminating in a Ph.D. in Computer Science and Engineering from Dr. APJ Abdul Kalam Technical University, where he focused on formal verification of recovery techniques in distributed systems using Event-B. Prior to his doctoral studies, he earned a Master of Technology in Computer Science and Engineering and a Master of Business Administration in Marketing with Finance. He has completed numerous workshops, short-term courses, seminars, and training programs, further enhancing his expertise in areas such as entrepreneurship, software testing, and artificial intelligence. He can be contacted at email: bal.saraswat@sharda.ac.in.

Santosh Kumar Maurya ⓘ 📇 SC ⊙ Assistant Professor at Department of Computer Science and Engineering, IIMT University, Meerut. He has done B.Tech. in 2000 and M.Tech. in 2013 in Computer Science and Engineering. His research interest is cloud computing, network security, information security and AI. He can be contacted at an email: santoshcse200@gmail.com.

Hare Ram Singh ⓘ 📇 SC ⊙ Associate Professor in GNIOT Engineering College, Greater Noida, District-G.B Nagar, U.P, (approved by AICTE, New Delhi and affiliated to AKTU, Lucknow). He has more than 22 years teaching and research experience in the domain of computer science and engineering. His interest research areas are network security, cyber security, data mining, IoT, machine learning. He contributed many research papers into the reputed journals, international conferences, reputed book chapter along with patent. He has also served as session chair in many conferences. He has also been reviewer for the reputed journals. He can be contacted at email: hrsingh.2000.2000@gmail.com.

Nidhi Bansal ⓘ 📇 SC ⊙ born in December 1990, Meerut, India. She has done Ph.D. in Computer Science and Engineering from AKTU (UPTU) Lucknow. With total 10 years of teaching experience, currently she is working as an Assistant Professor in Sharda University, Greater Noida, Uttar Pradesh, India. Her current research interests include scheduling algorithms, cloud computing technologies, IoT concepts, machine learning algorithms and its applications. She can be contacted at an email: nidhi18jul@mail.com.