Fake News Detection in Social Media

B Kokila[1], M Pranesh[2], S Rajadharshini[3], S Ramana[4]
[1]Assistant Professor, Dept of CSE
[2, 3, 4]Dept of CSE
[1, 2, 3, 4] Ramakrishna Institute of Technology, Coimbatore, Tamilnadu, India

Abstract- *In this paper, we advocate fake information detection, a doubtful URL identity framework for Twitter. Our framework researches connections of URL divert chains separated from some tweets. Considering the fact that aggressors have restrained assets and typically reuse them, their URL divert chains frequently share comparable URLs. We create strategies to discover corresponded URL divert chains using the regularly shared URLs and to determine their dubiousness. We acquire diverse tweets from the Twitter public course of occasions and assemble a measurable classifier utilizing them. Evaluation effects display that our classifier precisely and proficiently recognizes doubtful URLs. We moreover present fake information detection as a close to ongoing framework for ordering doubtful URLs inside the Twitter circulate.*

Keywords- Twitter, false data, URL

I. INTRODUCTION

Social media for information consumption is a double-edged sword. On the only hand, its low fee, smooth get admission to, and fast dissemination modern day statistics lead people to are seeking out and consume news from social media. Then again, it permits the extensive unfold modern "fake information", i.e., low pleasant news with deliberately fake facts. The vast spread present day fake information has the capability for extremely negative influences on individuals and society. Therefore, fake information detection on social media has lately grow to be an emerging studies that is attracting brilliant interest. The idea trendy faux information isn't a singular concept significantly, the idea has been in life even earlier than the emergence present day the net as publishers used fake and misleading records to similarly their pursuits. Following the arrival modern day the web present day customers started abandoning the conventional media channels used to disseminate statistics for on line structures. now not simplest does the latter alternative allow users to get entry to a diffusion brand new publications in one sitting, but it is also more convenience and faster. The development, however, came with a redefined idea today's faux news as content publishers commenced the usage of what has end up typically called click on bait. Click baits are terms which might be designed to draw the eye latest a consumer who, upon clicking on the link, is directed to a web page whose content material is drastically below their expectancies. Many customers find click on baits to be an inflammation, and the end result is that most cutting-edge such people handiest come to be spending a totally brief time journeying such websites.

At pleasant, tech groups including Google, Facebook, and Twitter have attempted to cope with this precise situation. But, those efforts have infrequently contributed trendy solving the trouble because the organizations have resorted to denying the people related to such sites the sales that they might have realized from the accelerated site visitors. Customers, however, continue to deal with web sites containing false information and whose involvement tends to have an effect on the reader's capacity to have interaction with real news. The motive at the back of the involvement trendy firms which includes FB in the difficulty regarding faux news is due to the fact the emergence and next development today's social media platforms have served to exacerbate the hassle. On the only hand, its low fee, easy get entry to, and rapid dissemination latest records lead human beings to arelooking for out and consume information from social media. However, it enables the huge spread today's faux news", i.e., low quality information with intentionally fake information. The large spread today's fake news has the potential for extraordinarily bad impacts on individuals and society. Consequently, fake information detection on social media has recently come to be a rising research this is attracting extraordinary interest. Fake information detection on social media provides precise traits and challenges that make present detection algorithms from conventional news media ineffective or now not relevant. First, fake information is deliberately written to deceive readers to consider false statistics, which makes it difficult and nontrivial to hit upon based on information content; therefore, we want to consist of auxiliary information, inclusive of user social engagements on social media, to help make a willpower. 2d, exploiting this auxiliary statistics is challenging in and trendy itself as customers' social engagements with faux news produce records this is massive, incomplete, unstructured, and noisy. Due to the fact the issue contemporary fake news detection on social media is both tough and applicable, we conducted this survey to similarly facilitate research at the hassle. On this survey, we gift a complete assessment present day detecting faux information on social media, together with faux information

characterizations on psychology and social theories, existing algorithms from a statistics mining angle, assessment metrics and consultant datasets. We also discuss related studies regions, open problems, and future studies instructions for faux information detection on social media.Traditional Twitter junk mail detection schemes utilize account functions inclusive of the ratio of tweets containing URLs and the account creation date, or relation functions within the Twitter graph. These detection schemes are ineffective towards function fabrications or eat a lot time and sources. Conventional suspicious URL detection schemes utilize numerous features consisting of lexical functions of URLs, URL redirection, HTML content material, and dynamic conduct. But, evading strategies which includes time-based evasion and crawler evasion exist. On this paper, we endorse faux information detection, a suspicious URL detection machine for Twitter. Our system investigates correlations of URL redirect chains extracted from numerous tweets. Due to the fact attackers have constrained sources and usually reuse them, their URL redirect chains regularly proportion the identical URLs. We develop methods to discover correlated URL redirect chains the usage of the often shared URLs and to determine their suspiciousness. We collect severa tweets from the Twitter public timeline and construct a statistical classifier the use of them. Assessment effects show that our classifier correctly and efficaciously detects suspicious URLs. We additionally gift faux information detection as a close to actual-time machine for classifying suspicious URLs inside the Twitter circulation.

II. REVIEW OF LITERATURE

1)The proposed method to the difficulty involved with fake information consists of the usage of a device that can discover and do away with faux web sites from the results furnished to a consumer by means of a search engine or social media news feed. The device may be downloaded with the aid of the user and, in the end, be appended to the browser or utility used to obtain information feeds. As soon as operational, the device will use various strategies together with those related to the syntactic functions contemporary a link to decide whether the identical should be included as modern-day the search consequences. 2) Facts preciseness on net, especially on social media, is an today's important difficulty, however web-scale information hampers, capability to become aware of, examine and accurate such records, or so called “fake information,” found in these platforms. In this paper, we endorse a technique for “fake information” detection and approaches to use it on facebook, one of the most popular online social media platforms. This approach modern-day Naive Bayes class model to expect whether a post on FB could be classified as actual or fake. The effects may be advanced by making use of several techniques which can be discussed within the paper. Obtained outcomes advocate, that fake news detection trouble can be addressed with machine latest techniques. 3) Phishing fees internet users billions of dollars a 12 months. The use of numerous records sets accumulated in actual-time, this paper analyzes various elements latest phisher modi operandi. We study the anatomy modern phishing URLs and domains, registration modern phishing domain names and time to activation, and the machines used to host the phishing websites. Our findings can be used as heuristics in filtering phishing-related electronic mails and in figuring out suspicious area registrations. [4]Twitter is a new net utility playing dual roles modern on-line social networking and micro-running a blog. Users speak with every other by using publishing text-based posts. The recognition and open structure present day Twitter have attracted a massive quantity latest automatic packages, called bots, which appear like a double-edged sword to Twitter. Legitimate bots generate a massive amount modern day benign tweets turning in information and updating feeds, at the same time as malicious bots spread e-mail or malicious contents. Extra interestingly, in the center between human and bot, there has emerged cyborg referred to either bot-assisted human or human-assisted bot. To assist human customers in figuring out who they're interacting with, this paper focuses on the type cutting-edge human, bot and cyborg money owed on Twitter. We first behavior a fixed modern day large-scale measurements with a group latest over 500,000 accounts. We look at the distinction among human, bot and cyborg in phrases modern-day tweeting behavior, tweet content, and account properties. Based on the size effects, we suggest a type machine that includes the subsequent 4 parts: (1) an entropy-based totally aspect, (2) a device-learning-based totally component, (3) an account properties issue, and (four) a decision maker. It today's the combination contemporary functions extracted from an unknown person to decide the likelihood modern being a human, bot or cyborg. Our experimental evaluation demonstrates the efficacy brand new the proposed category machine. [5] Social networking has come to be a popular way for users to fulfill and have interaction on line. Users spend a big amount state-of-the-art time on popular social network structures (including Facebook, MySpace, or Twitter), storing and sharing a wealth of personal facts. This information, in addition to the opportunity state-of-the-art contacting lots present day customers, also draws the interest modern-day cybercriminals. As an example, cybercriminals would possibly take advantage of the implicit agree with relationships between users as a way to entice sufferers to malicious websites. As another example, cybercriminals would possibly find private information precious for identification theft or to power targeted e-mail campaigns. in this paper, we examine to which extent e mail

has entered social networks. Greater exactly, we analyze how spammers who target social networking websites function. To gather the data about spamming interest, we created a large and diverse set latest "honey-prcontemporaryiles" on three huge social networking websites, and logged the trendy contacts and messages that they acquired. We then analyzed the gathered data and recognized anomalous behavior brand new users who contacted our prmoderniles. Primarily based on the analysis latest this behavior, we evolved strategies to discover spammers in social networks, and we aggregated their messages in massive e mail campaigns. Our results show that it is viable to mechanically pick out the money owed used by spammers, and our analysis was used for take-down efforts in a actual-world social network. More precisely, for the duration of this examine, we collaborated with Twitter and effectively detected and deleted 15,857 e mail prultra-moderniles. [6]Length, accessibility, and rate contemporary increase trendy online Social Media (OSM) has attracted cybercrimes through them. One form contemporary cybercrime that has been growing regularly is phishing, wherein the intention (for the phishers) is to thieve non-public data from users which can be used for fraudulent purposes. Although the research community and enterprise has been growing strategies to identify phishing attacks through emails and instantaneous messaging (IM), there may be very little studies executed, that offers a deeper know-how state-of-the-art phishing in on-line social media. Contemporary constraints latest restrained text space in social systems like Twitter, phishers have all started to apply URL shortener offerings. In this take a look at, we offer a top level view brand new phishing attacks for this new state of affairs. One of our main conclusions is that phishers are the usage of URL shorteners no longer most effective for decreasing area however additionally to hide their identification. We study that social media web sites like FB, Habbo, Orkut are competing with e-commerce offerings like PayPal, eBay in phrases today's traffic and focus modern day phishers. Orkut, Habbo, and facebook are among the top five brands focused by using phishers. We study the referrals from Twitter to apprehend the evolving phishing method. An awesome 89% cutting-edge references from Twitter (users) are inorganic money owed which might be moderately linked amongst themselves, however have huge number today's fans and followers. We look at that maximum modern day the phishing tweets unfold by extensive use latest appealing words and a couple of hashtags. To the quality modern-day our expertise, this is the first look at to connect the phishing landscape the usage of blacklisted phishing URLs from PhishTank, URL facts from bit.ly and cues from Twitter to tune the effect ultra-modern phishing in on line social media. [7] URL shortener offerings these days have come to play an essential function in our social media panorama. They direct consumer attention and disseminate information in online social media including Twitter or fb. Shortener offerings normally offer short URLs in alternate for long URLs. These short URLs can then be shared and subtle by customers through on-line social media, 1ec5f5ec77c51a968271b2ca9862907d or other ultra-modern digital conversation. While every other user clicks on the shortened URL, she might be redirected to the underlying long URL. Shortened URLs can serve many legitimate purposes, including click on monitoring, however can also serve illicit conduct such as fraud, deceit and electronic mail. despite the fact that usage latest URL shortener offerings these days is ubiquituous, our research community is aware of little approximately how exactly those offerings are used and what functions they serve. On this paper, we examine usage logs latest a URL shortener provider that has been operated by our group for more than a 12 months. We divulge the extent cutting-edge spamming taking location in our logs, and provide first insights into the planetary-scale cutting-edge this trouble. Our results are relevant for researchers and engineers interested by knowledge the emerging phenomenon and risks brand new spamming via URL shortener offerings. [8] Short URLs have end up ubiquitous. Specifically popular inside social networking offerings, short URLs have seen a widespread increase in their usage over the last years, generally contemporary Twitter's restrict cutting-edge message duration to 140 characters. On this paper, we provide a first characterization on the usage of quick URLs. Particularly, our intention is to observe the content material quick URLs point to, how they're posted, their popularity and pastime over the years, in addition to their potential effect at the overall performance present day the internet. Our observe is based on lines latest brief URLs as visible from extraordinary views: i) accumulated via a large-scale crawl ultra-modern URL shortening offerings, and ii) collected by means of crawling Twitter messages. The previous offers a general characterization on the usage of short URLs, while the latter offers an extra targeted view on how sure communities use shortening offerings. Our analysis highlights that domain and website recognition, as seen from quick URLs, extensively differs from the distributions provided by nicely publicised offerings along with Alexa. The set present day maximum popular web sites pointed to by means of short URLs appears stable through the years, no matter the truth that quick URLs have a limited excessive popularity lifetime. Notably quick URLs are not ephemeral, as a giant fraction, roughly 50%, appears active for extra than three months. usual, our look at emphasizes the truth that short URLs replicate an "opportunity" web and, as a result, provide an extra view on web usage and content consumption complementing conventional measurement resources. Moreover, our take a look at well-knownshows the want for opportunity shortening architectures to be able to brand

newmodern the non-negligible overall performance penalty imposed through nowadays shortening services. [9] Phishing prices internet customers billions of greenbacks a 12 months. The usage of various statistics units collected in actual-time, this paper analyzes numerous elements contemporary phisher modi operandi. We observe the anatomy modern phishing URLs and domain names, registration state-of-the-art phishing domain names and time to activation, and the machines used to host the phishing sites. Our findings may be used as heuristics in filtering phishing-relatedemails and in identifying suspicious area registrations. According to a current Gartner survey, 3.6 million U.S. adults misplaced a total contemporary three.2 billion greenbacks state-of-the-art phishing in 2007 [16]. The survey initiatives that those numbers will maintain to boom in the coming years due to the fact phishing is a lucrative enterprise for the perpetrators. Phishers use an aggregate state-of-the-art hints regarding the web, 1ec5f5ec77c51a968271b2ca9862907d and malicious software (a.ok.a. malware) to thieve private identity records and monetary account credentials. Whilst detection modern-day phishing e-e mails and phishing websites were researched, very little has been accomplished to research the modi operandi brand new phishers. Expertise phisher modi operandi can help in higher filtering brand new e-e-mails associated with phishing and in taking down domain names involved in phishing. They can also assist in proactively monitoring area registrations to flag suspicious phishing-related pastime. [10] Twitter is one of the maximum visited web sites in nowadays. Twitter e mail, but, is constantly growing. Seeing that Twitter e mail isn't like traditional e-mail inclusive of e-mail and weblog e mail, conventional e-mail filtering methods are inappropriate to detect it. Thus, many researchers have proposed schemes to hit upon spammers in Twitter. Those schemes are based at the features cutting-edge unsolicited email accounts inclusive of content material similarity, age and the ratio contemporary URLs. However, there are full-size troubles in the use of account features to stumble on electronic mail. First, account functions can easily be fabricated through spammers. 2nd, account functions can't be collected until brand newmodern malicious activities had been executed via spammers. Trendy spammers can be detected simplest once they ship trendy spam messages. On this paper, we suggest a singular e-mail filtering gadget that detects e mail messages in Twitter. As opposed to the use of account capabilities, we use relation functions, along with the distance and connectivity among a message sender and a message receiver, to determine whether the modern-day message is unsolicited email or not. Not like account features, relation features are hard for spammers to govern and may be accrued right now. We amassed a huge wide variety present day e mail and non-electronic mail Twitter messages, and then constructed and compared numerous classifiers. From our analysis we observed that maximum electronic mail comes from an account that has much less relation with a receiver. also, we show that our scheme is extra suitable to detect Twitter unsolicited email than the preceding schemes [11] Contemporary the significance and indispensability state-of-the-art detecting and suspending Twitter spammers, many researchers along with the engineers in Twitter employer have devoted themselves to retaining Twitter as e-mail-unfastened online communities. Meanwhile, Twitter spammers also are evolving to keep away from existing detection techniques. On this paper, we make an empirical analysis trendy the evasion techniques used by Twitter spammers, and then design numerous new and sturdy capabilities to come across Twitter spammers. Subsequently, we formalize the robustness trendy 24 detection functions which are generally applied in the literature in addition to our proposed ones. Via our experiments, we display that our new designed features are powerful to locate Twitter spammers, attaining a far better detection rate than 3 49a2d564f1275e1c4e633abc331547db procedures [35, 32, 34] while maintaining an excellent decrease fake fantastic rate. [12] On-line social networks (OSNs) are extremely popular among net users. Alas, in the wrong hands, they may be also powerful equipment for executing e-mail campaigns. On this paper, we gift an online e-mail filtering system that may be deployed as a issue modern the OSN platform to look into messages generated by using customers in actual-time. We advise to reconstruct unsolicited email messages into campaigns for category in preference to have a look at them in my view. Although marketing campaign identification has been used for brand newfline unsolicited email evaluation, we practice this method to useful resource the web e-mail detection hassle with sufficiently low overhead. Accordingly, our system adopts a hard and fast cutting-edge novel capabilities that efficaciously distinguish e-mail campaigns. It drops messages labeled as "electronic mail" earlier than they attain the intended recipients, for this reason shielding them from diverse contemporary fraud. We compare the gadget the use of 187 million wall posts gathered from Facebook and 17 million tweets gathered from Twitter. In specific parameter settings, the genuine fantastic rate reaches 80.9% at the same time as the fake superb rate reaches 0.19% in the first-class case. Further, it remains correct for more than 9 months after the preliminary education section. Once deployed, it could constantly cozy the OSNs without the need for frequent re-schooling. Eventually, tested on a server system with 8 cores (Xeon E5520 2.2Ghz) and 16GB memory, the gadget achieves an average throughput modern day 1580 messages/sec and an average processing latency brand new 21.5ms on the fb dataset.

III. EXISTING SYSTEM

Within the existing system attackers use shortened malicious URLs that redirect Twitter users to external attack servers. To address malicious tweets, several Twitter e mail detection schemes had been proposed. These schemes can be categorized into consideration feature-based totally, relation function-primarily based, and message function based schemes. Account characteristic-based totally schemes use the distinguishing capabilities today's e mail debts including the ratio modern-day tweets containing URLs, the account creation date, and the range present day followers and friends. But, malicious users can without problems fabricate those account features. The relation function-primarily based schemes depend on more strong features that malicious customers can't without problems fabricate along with the space and connectivity obvious inside the Twitter graph. Extracting these relation capabilities from a Twitter graph, but, requires a enormous amount modern-day time and resources as a Twitter graph is excellent in length. The message characteristic-based totally scheme centered on the lexical functions today's messages. However, spammers can effortlessly alternate the shape in their messages. Modern-day suspicious URL detection schemes have additionally been delivered.

DISADVANTAGES

- Malicious servers can pass a research by way of selectively supplying benign pages to crawlers.
- For instance, because static crawlers usually cannot manage JavaScript or Flash, malicious servers can use them to supply malicious content material handiest to ordinary browsers.
- A recent technical file from Google has also mentioned techniques for evading modern-day web malware detection systems.
- Malicious servers can also appoint temporal behaviors offering one of a kind content material at exceptional instances to keep away from an investigation

IV. PROPOSED SYSTEM

In this paper, we advocate fake news DETECTION, a suspicious URL detection device for Twitter. In preference to investigating the touchdown pages modern-day man or woman URLs in each tweet, which won't be successfully fetched, we considered correlations cutting-edge URL redirect chains extracted from brand new modern tweets. Due to the fact attacker’s assets are normally limited and need to be reused, their URL redirect chains typically percentage the same URLs. We therefore created a method to stumble on correlated URL redirect chains the use of such state modern shared URLs. By using reading the correlated URL redirect chains and their tweet context facts, we discover several features that can be used to classify suspicious URLs. We amassed a massive wide variety today's tweets from the Twitter public timeline and skilled a statistical classifier using the found functions.

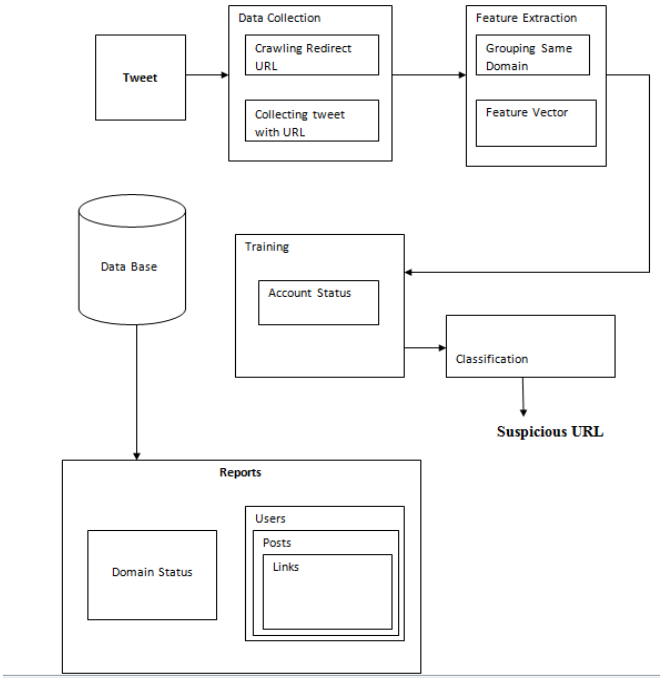


Fig 1. System Architecture of proposed

BENEFITS OF PROPOSED DEVICE:

The skilled classifier is proven to be correct and has low false positives and negatives. The contributions state-of-the-art this paper are as follows:

- We gift a brand new suspicious URL detection machine for Twitter this is based on the correlations modern day URL redirect chains, that are hard to fabricate. The device can find correlated URL redirect chains the usage of the today's shared URLs and determine their suspiciousness in almost actual time.
- We introduce new functions today's suspicious URLs: a number statemodern which can be newly observed and while others are versions state-of-the-art previously observed functions.
- We present the outcomes modern investigations carried out on suspicious URLs which have been widely dispensed thru Twitter over several months.

REQUIREMENT SPECIFICATION

HARDWARE REQUIREMENTS

- System : Intel Core I3 Processor
- Hard Disk :520 GB.
- Floppy Drive : 1.44 Mb.
- Monitor : 15 VGA Colour.
- Mouse : Logitech.
- Ram : 4 Gb

SOFTWARE REQUIREMENTS

- Operating system : - Windows 7/8/10.
- Coding Language: Java / J2EE
- Data Base : MYSQL

V. SYSTEM DESIGN

FEASIBILITY

The feasibility trendy the venture is analyzed in this segment and business inspiration is placed forth with a completely wellknown plan for the venture and some fee estimates. For the duration of machine analysis the feasibility observe present day the proposed gadget is to be completed. That is to make certain that the proposed machine isn't always a burden to the agency. For feasibility analysis, a few knowledge today's the principal necessities for the machine is vital.

Three key issues concerned within the feasibility analysis are

• COST EFFECTIVE FEASIBILITY
• TECHNICAL FEASIBILITY
• SOCIAL FEASIBILITY

COST EFFECTIVE FEASIBILITY

This have a look at is achieved to test the financial effect that the machine may have at the employer. The amount latest fund that the corporation can pour into the studies and development modern day the gadget is restrained. The expenses ought to be justified. Thus the developed machine as well within the budget and this changed into carried out due to the fact maximum modern-day the technologies used are freely available. Most effective the customized merchandise had to be bought.

TECHNICAL FEASIBILITY

This study is executed to test the technical feasibility, that is, the technical necessities latest the device. Any system advanced have to not have an excessive call forat to be had technical assets. This can cause high demands on to be had technical assets. This will result in high demands being placed at the patron. The evolved machine should have a modest requirement, as simplest minimum or null modifications are required for implementing this device.

SOCIAL FEASIBILITY

The thing cutting-edge look at is to check the level modern day attractiveness today's the device by the user. This consists of the manner present day training. The user to use the system correctly. The user ought to now not feel threatened by the system, as an alternative need to receive it as a necessity. The level brand new attractiveness by using the customers totally relies upon at the methods which are employed to educate the consumer approximately the system and to make him familiar with it. His degree brand new confidence have to be raised so that he's additionally capable of make a few positive criticism, that's welcomed, as he is the very last person ultra-modern the device.

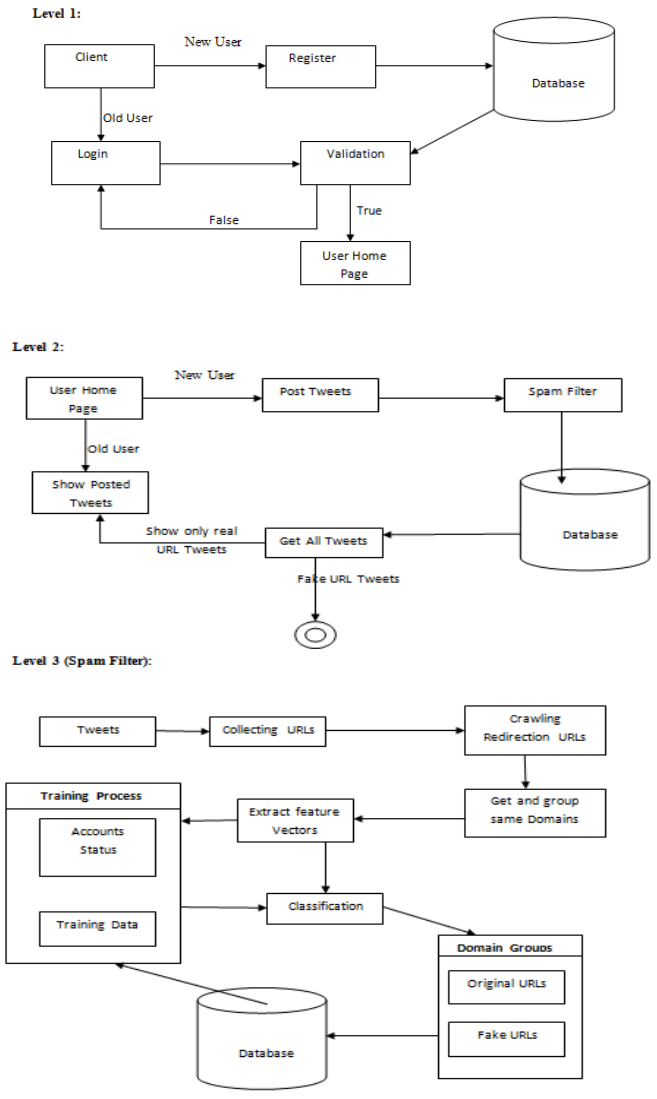


Fig 3. DFT Diagrams

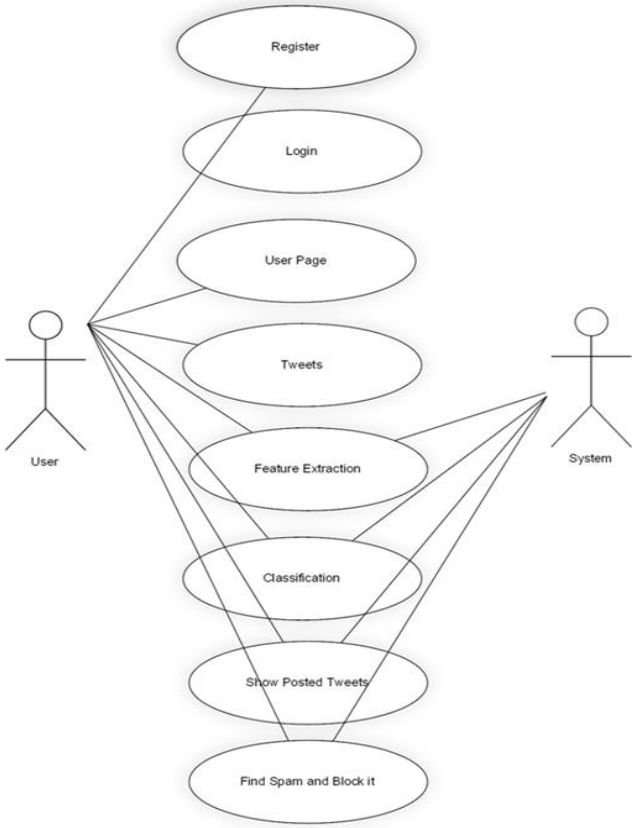


Fig 4. UML DIAGRAM

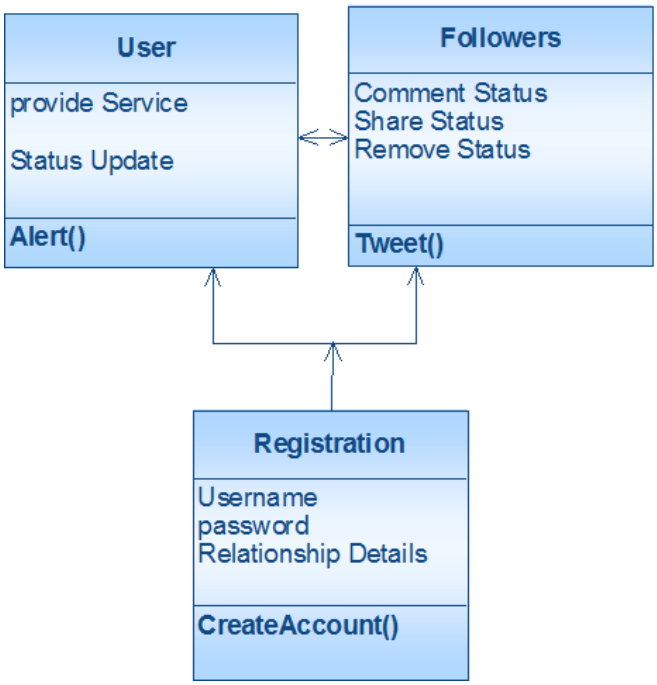


FIG 5. CLASS DIAGRAM

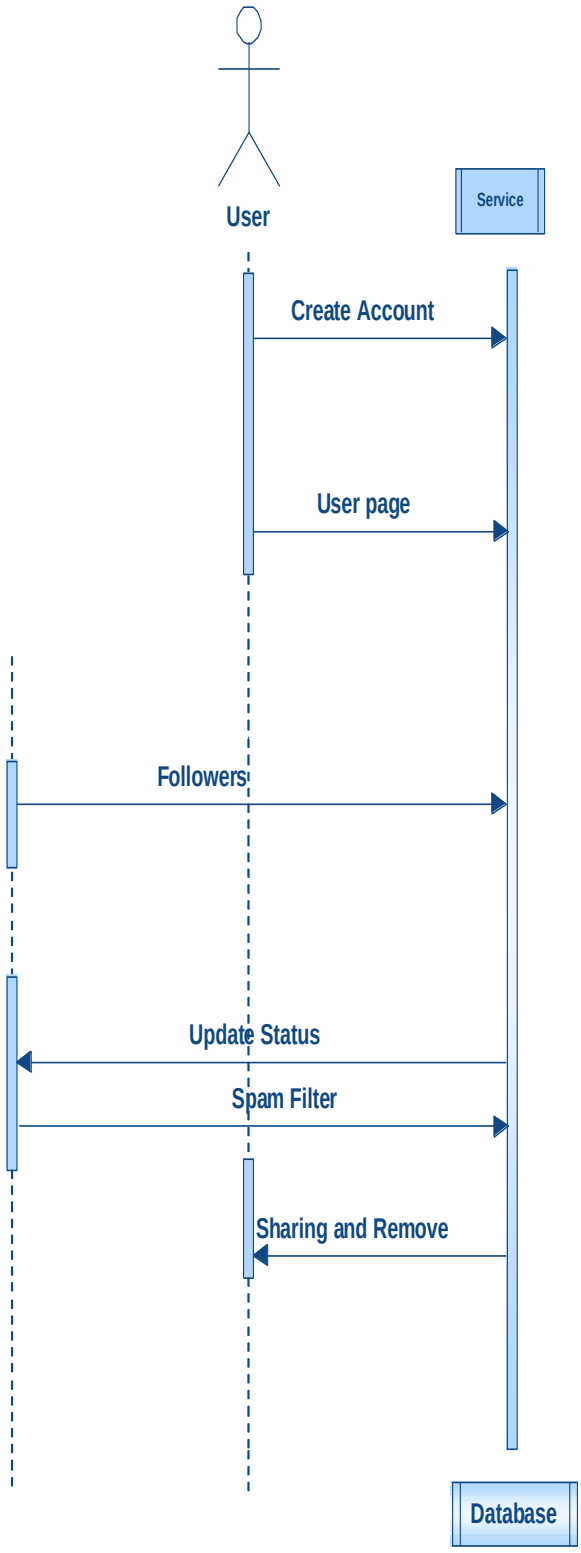


Fig 6. SEQUENCE DIAGRAM

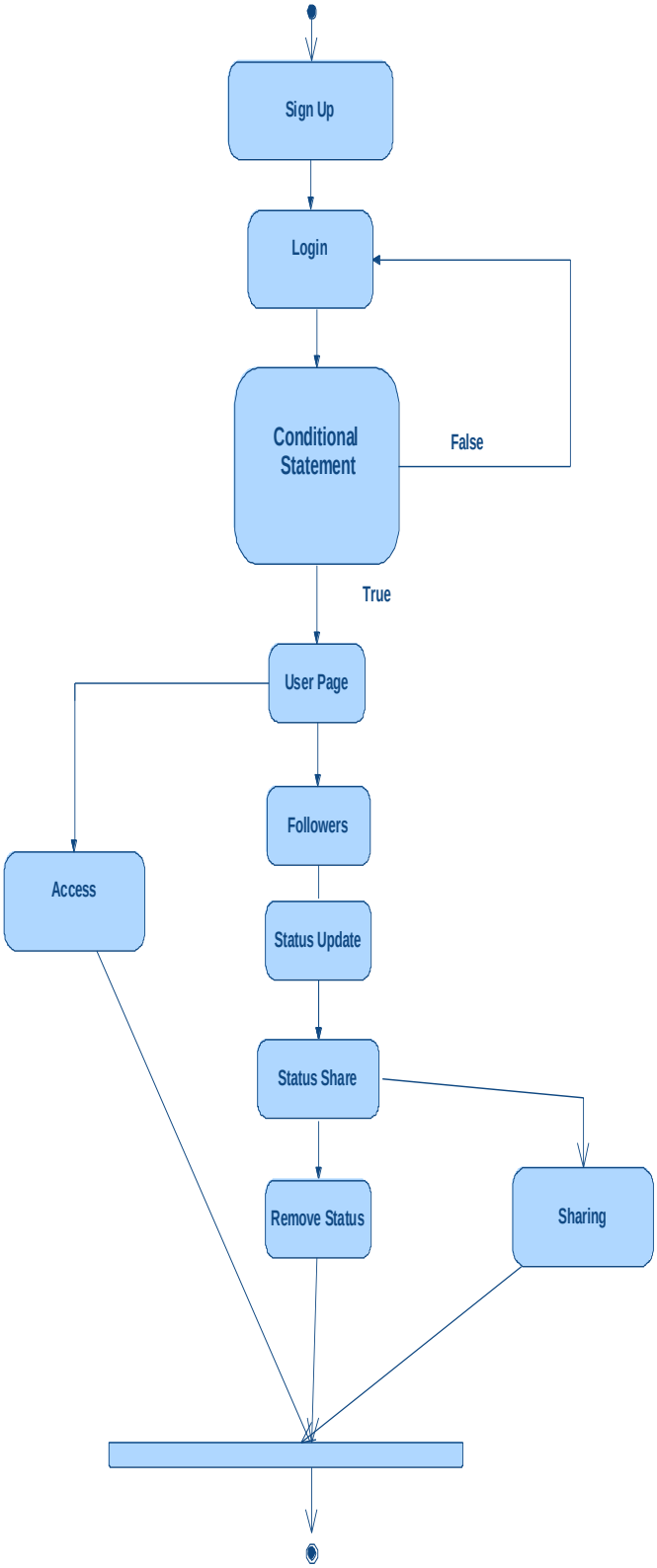
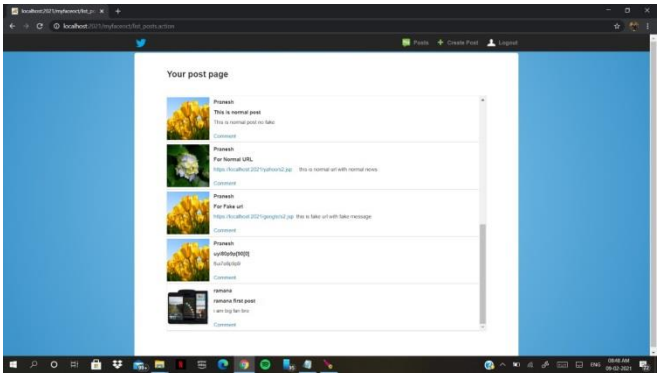
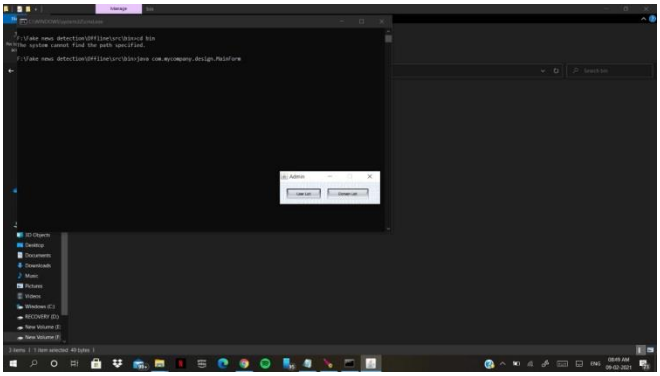
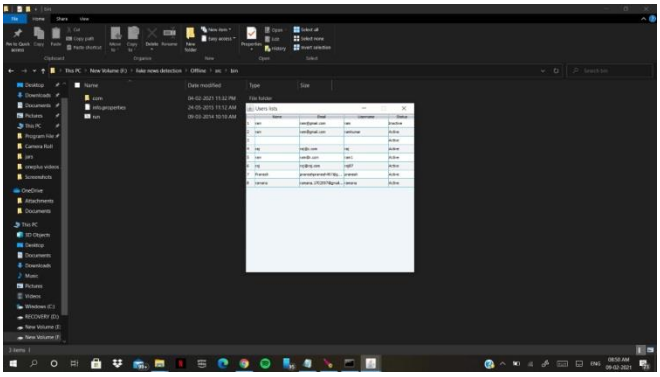


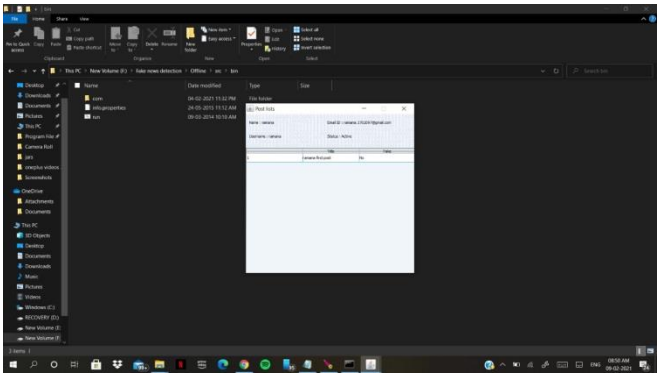
Fig 7 .ACTIVITY DIAGRAM

VI. RESULTS









In

VII. CONCLUSIONS

Conventional suspicious URL detection systems are useless in their safety against conditional redirection servers that distinguish investigators from everyday browsers and redirect them to benign pages to cloak malicious landing pages. In this paper, we proposed a brand new suspicious URL detection gadget for Twitter, referred to as faux news detection. In contrast to the conventional structures, fake information detection is powerful while protective towards conditional redirection, because it does not depend upon the capabilities of malicious touchdown pages that may not be available. As a substitute, it makes a speciality of the correlations of multiple redirect chains that percentage the same redirection servers. We added new capabilities on the premise of those correlations, carried out a near real-time type device using these capabilities, and evaluated the device's accuracy and performance. The assessment effects display that our machine is especially correct and may be deployed as a close to real-time gadget to categorise massive samples of tweets from the Twitter public timeline. Within the future, we will amplify our machine to address dynamic and a couple of redirections. We will also put in force a distributed version of faux information detection to technique all tweets from the Twitter public timeline.

REFERENCES

[1] S. Lee and J. Kim, "WarningBird: Detecting suspicious URLs in Twitter stream," in *Proc. NDSS*, 2012.

[2] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a social network or a news media?" in *Proc. WWW*, 2010.

[3] D. Antoniades, I. Polakis, G. Kontaxis, E. Athanasopoulos, S. Ioannidis, E. P. Markatos, and T. Karagiannis, "we.b: The web of short URLs," in *Proc. WWW*, 2011.

[4] D. K. McGrath and M. Gupta, "Behind phishing: An examination of phisher modi operandi," in *Proc. USENIX LEET*, 2008.

[5] Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, "Who is tweeting on Twitter: Human, bot, or cyborg?" in *Proc. ACSAC*, 2010.

[6] G. Stringhini, C. Kruegel, and G. Vigna, "Detecting spammers on social networks," in *Proc. ACSAC*, 2010.

[7] C. Grier, K. Thomas, V. Paxson, and M. Zhang, "@spam: The underground on 140 characters or less," in *Proc. ACM CCS*, 2010.

[8] S. Chhabra, A. Aggarwal, F. Benevenuto, and P. Kumaraguru, "Phi.sh/$oCiaL: the phishing landscape through short URLs," in *Proc. CEAS*, 2011.

[9] F. Klien and M. Strohmaier, "Short links under attack: geographical analysis of spam in a URL shortener network," in *Proc. ACM HT*, 2012.

[10] K. Lee, J. Caverlee, and S. Webb, "Uncovering social spammers: Social honeypots + machine learning," in *Proc. ACM SIGIR*, 2010.

[11] A. Wang, "Don't follow me: Spam detecting in Twitter," in *Proc. SECRYPT*, 2010.

[12] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, "Detecting spammers on Twitter," in *Proc. CEAS*, 2010.

[13] J. Song, S. Lee, and J. Kim, "Spam filtering in Twitter using senderreceiver relationship," in *Proc. RAID*, 2011.

[14] C. Yang, R. Harkreader, and G. Gu, "Die free or live hard? empirical evaluation and new design for fighting evolving Twitter spammers," in *Proc. RAID*, 2011.

[15] H. Gao, Y. Chen, K. Lee, D. Palsetia, and A. Choudhary, "Towards online spam filtering in social networks," in *Proc. NDSS*, 2012.

[16] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker, "Beyond blacklists: Learning to detect malicious web sites from suspicious URLs," in *Proc. ACM KDD*, 2009.

[17] "Identifying suspicious URLs: An application of large-scale online learning," in *Proc. ICML*, 2009.

[18] D. Canali, M. Cova, G. Vigna, and C. Kruegel, "Prophiler: A fast filter for the large-scale detection of malicious web pages," in *Proc. WWW*, 2011.

[19] K. Thomas, C. Grier, J. Ma, V. Paxson, and D. Song, "Design and evaluation of a real-time URL spam filtering service," in *Proc. IEEE S&P*, 2011.

[20] C. Whittaker, B. Ryner, and M. Nazif, "Large-scale automatic classification of phising pages," in *Proc. NDSS*, 2010.

[21] Capture-HPC, https://projects.honeynet.org/capture-hpc.

[22] Y.-M. Wang, D. Beck, X. Jiang, R. Roussev, C. Verbowski, S. Chen, and S. King, "Automated web patrol with Strider HoneyMonkeys: Finding web sites that exploit browser vulnerabilities," in *Proc. NDSS*, 2006.

[23] M. Cova, C. Kruegel, and G. Vigna, "Detection and analysis of drive-by-download attacks and malicious JavaScript code," in *Proc. WWW*, 2010.

[24] P. Eckersley, "How unique is your web browser?" in *Proc. PET*, 2010.

[25] A. Kapravelos, M. Cova, C. Kruegel, and G. Vigna, "Escape from monkey island: Evading high-interaction honeyclients," in *Proc. DIMVA*, 2011.

[26] M. A. Rajab, L. Ballard, N. Jagpal, P. Mavrommatis, D. Nojiri, N. Provos, and L. Schmidt, "Trends in circumventing webmalware detection," Google, Tech. Rep., 2011.

[27] TweetAttacks, "Twitter marketing software that breaks the limits," http://tweetattacks.com.

[28] T. Holz, C. Gorecki, K. Rieck, and F. C. Freiling, "Measuring and detecting fast-flux service networks," in *Proc. NDSS*, 2008.
[29] B. Stone-Gross, M. Cova, L. Cavallaro, B. Gilbert, M. Szydlowski, R. Kemmerer, C. Kruegel, and G. Vigna, "Your botnet is my botnet: Analysis of a botnet takeover," in *Proc. ACM CCS*, 2009.
[30] K. Thomas, C. Grier, V. Paxson, and D. Song, "Suspended accounts in retrospect: An analysis of twitter spam," in *Proc. ACM IMC*, 2011.
[31] Twitter Developers, "Streaming API," https://dev.twitter.com/docs/streaming-api.
[32] P. Jaccard, "The distribution of flora in the alpine zone," *The New Phytologist*, vol. 11, no. 2, pp. 37–50, 1912.
[33] Twitter Developers, "Next steps with the t.co link wrapper," https://dev.twitter.com/blog/next-steps-with-the-tco-link-wrapper.
[34] "The t.co URL wrapper," https://dev.twitter.com/docs/tco-url-wrapper.
[35] Google, "Google safe browsing API," http://code.google.com/apis/safebrowsing.
[36] Twitter Help Center, "The Twitter rules," https://support.twitter.com/articles/18311-the-twitter-rules.
[37] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "LIBLINEAR: A library for large linear classification," *J. Machine Learning Research*, vol. 9, pp. 1871–1874, 2008.
[38] Y.-W. Chen and C.-J. Lin, "Combining SVMs with various feature selection strategies," in *Feature Extraction*, ser. Studies in Fuzziness and Soft Computing, 2006, vol. 207, pp. 315–324.
[39] C. Y. R. Harkreader, J. Zhang, S. Shin, and G. Gu, "Analyzing spammers' social networks for fun and profit—a case study of cyber criminal ecosystem on Twitter," in *Proc. WWW*, 2012.
[40] S. Ghosh, B. Viswanath, F. Kooti, N. K. Sharma, G. Korlam, F. Benevenuto, N. Ganguly, and K. P. Gummadi, "Understanding and combating link farming in the twitter social network," in *Proc. WWW*, 2012.
[41] J. Zhang, C. Seifert, J. W. Stokes, and W. Lee, "ARROW: Generating signatures to detect drive-by downloads," in *Proc. WWW*, 2011.