Defamation 2.0: New Threats in Digital Media Era - An Overview on Forensics Approaches in the Social Network Ecosystem

Cristina Nastasi and Sebastiano Battiato
Dipartimento di Matematica ed Informatica, Università degli Studi di Catania, Italy

Keywords: Digital Forensics, Social Media, Defamation, Criminal Procedure.

Abstract: Recently, social networks have become the largest and fastest growing websites on the Internet. These platforms contain sensitive and personal data of hundreds of millions of people, and are integrated also into millions of other websites so it is more and more important to focus on security and privacy issues. In this work, we expose the defamation issue in the social network context and apply some known methods to recover data of person who offends reputation of others over 250 different social media frameworks. The datasets, that it is possible to exploit, can contain various profile information (user data, photos, etc.) and associated meta-data (internal timestamps and unique identifiers). These data are significant in the field of digital forensics to be properly used as evidences in front of Court.

1 INTRODUCTION

Every day, millions of people connect to the Internet and most of them use “Social Networks" to work, to keep in touch with friends or simply for fun. This tendency to use these "communication platforms" has made social networks the undisputed masters of media communication on the web in recent years. Users can share information, take care of their interpersonal relationships or create new ones, can advertise their business or even set up real marketing campaigns. There are different types of social networks like professional or entertainment and specific categories: animals, sports, music, etc. Unfortunately, it is not always good to expose your personal data on social media. A news published on the web, a post on a social network, an inappropriate comment on a chat of a Facebook or a "Whatsapp" group are able to easily reach an unspecified number of people and can, however, be quite dangerous whenever the subject of the diffused message has a disparaging and defamatory nature towards its recipient. In recent years the issue of defamation through the use of social networks has been the subject of extensive debates because it is one of the criminal offenses that are most commonly used. Thanks to anonymity, the web induces the most impudent (called haters or keyboard lions) to offenses and insults of all kinds. The number of cases of insult and defamation on these social networks are increasing and it is necessary to be ready to react in the appropriate forms. Defamation or offense that occurs on social media is punishable and their certified acquisition can become evidence in criminal or civil court proceedings.

In this paper we present on overview on existing social network platforms, brief notes on the defamation crime on the net and show the procedure to recover and freeze useful data to be used as evidences at Court. Finally we expose the obtained result and concludes the paper with the explanation of our future goal in this field.

2 SOCIAL MEDIA

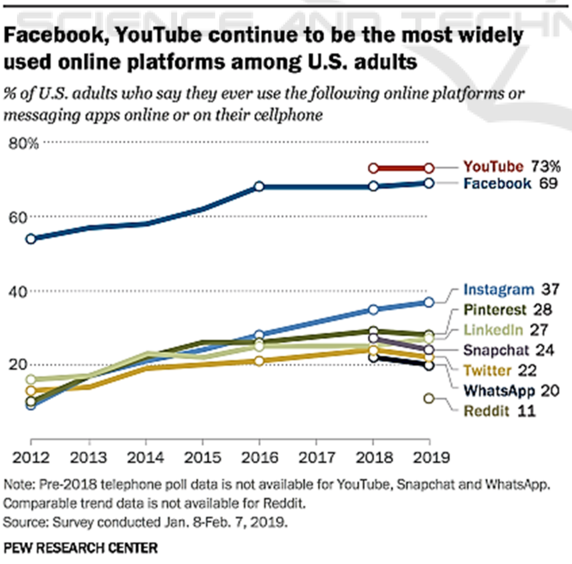
Social networks, born in the late nineties, allow users to create an appropriate user profile, to organize a list of people to keep in touch with, to publish their own stream of updates and access that one of others. A social network (Boyd & Ellison, 2007) is a service offered through the Internet, typically usable in a completely free way through the web or by specific applications for mobile devices, whose purpose is to facilitate the management of social relationships by allowing communication and sharing of digital content. Most social networks have common

Nastasi, C. and Battiato, S.
Defamation 2.0: New Threats in Digital Media Era - An Overview on Forensics Approaches in the Social Network Ecosystem.
DOI: 10.5220/0010463601210127
In *Proceedings of the International Conference on Image Processing and Vision Engineering (IMPROVE 2021)*, pages 121-127
ISBN: 978-989-758-511-1

characteristics that can be identified in three main elements:

- the creation of a personal profile (public or semi-public);
- the creation of a list of friends;
- the exploration of one's own network and that one of friends.

Once the registration phase is completed, in which we are asked to provide an e-mail address and a password, we move on to the creation and management of a personal user profile through a series of questions (on the city of birth, on the place where we live, on the school we attend or have attended, on the job, on personal interests, hobbies and more). This page contains general information about the user with images or photos, videos and a short self-description. Then we move on to the creation of a list of friends. The list of "contacts" is usually expanded with the help of machine learning algorithms through reference to the answers we gave in the user profile compilation phase, suggest friends we met in real life and "potential new friends", selecting from the registered users, people who have characteristics corresponding to our indications. Another typical social networks feature is the ability to explore the profiles of friends who are part of our friends list and those who are part of our friends' friends (even if they have not made friends directly with us); you can visit the personal pages of users (friends), observe their favourite activities, musical tastes, etc. and of course interact directly with people we don't know.


Facebook, YouTube continue to be the most widely used online platforms among U.S. adults

% of U.S. adults who say they ever use the following online platforms or messaging apps online or on their cellphone

Note: Pre-2018 telephone poll data is not available for YouTube, Snapchat and WhatsApp. Comparable trend data is not available for Reddit.
Source: Survey conducted Jan. 8-Feb. 7, 2019.

PEW RESEARCH CENTER

Figure 1: The results of the survey conducted from the 8th of January 2019 to the 7th of February 2019 by Pew Reserarch Center. It is showed the percentage of U.S. adults who say they use certain online platforms or apps online or on they cellphone.

A new Pew Research Center survey (Smith & Anderson, 2018) of U.S. adults finds that the social media landscape in early 2018 is defined by a mix of old trends, but also new emerging apps. Facebook andYouTube dominate this landscape, while at the same time, younger (especially between 18 and 24) use a variety of other platforms frequently. Moreover, there are substantial differences in social media use by age and there is a substantial amount of overlap between users of the various sites.

Social networks have become the largest websites on the Internet. This web site, such as Facebook or LinkedIn, contain sensitive and personal data of hundreds of millions of people, and are integrated also into millions of other websites. Research has acknowledged the importance of these websites and recently, a number of publications have focused on security issues. In particular, a number of empirical studies on online social networks (L. Bilge, Balzarotti, & Kirda, 2009) (Gao, et al., 2010) (Jagatic, Johnson, Jakobsson, & Menczer, 2007) (M. Huber, 2011) (Wondracek, Holz, Kirda, & Kruegel, 2010) highlight challenges to the security and privacy of social network users and their data.

2.1 How Many Are the Social Media Sites and Apps?

The world of social media continues to maintain a great power of attraction, in the last year the number of users has grown again by 17%. According to the site (La Stampa, 2018), each person who frequents social networks is registered on an average of seven sites. The existing social networks are not limited only to Twitter, Facebook, LinkedIn and Blog. But how many social networks are there in the world? Wikipedia (Wikipedia, 2020) lists 206 of them, while (Social Media List, 2020) has registered 250 ones. In alphabetical order they go from Academia.edu, a site for teachers and researchers that helps to make their work known by sharing scientific publications, up to Greek Zoo.gr frequented by those who want to play online. Some have a few thousand members and are dedicated to individual passions, such as books or cinema, or to communities of people who share particular situations. The social network founded by Mark Zuckerberg remains the most popular and today has more than 2 billion and 100 million users.

2.2 Defamation Crime in Italian Law

In Italian law, defamation is the crime provided for by art. 595 of the Criminal Code, which reads: "Anyone who, apart from the cases indicated in the previous article, by communicating with more people, offends the reputation of others" and "is punished with imprisonment of up to one year or a fine of up to euro 1,032.00 ". Law 48/2008 has introduced into our legal system a series of new offenses generically classified as computer crimes, but has not added anything for the possibility of configuring the defamation crime through computer networks or telematics. In the typical crime of defamation, the principal legal asset is the reputation and its structural elements are: the offense to the reputation of others that is an injury to personal, moral, social, professional qualities, etc. of an individual; communication with several people, where the expression "several people" must certainly be understood as "at least two people"; the absence of the offended person. It is not difficult to argue, however, that the offense referred to in art. 595 of the Criminal Code, is sufficiently generic to also include all those offensive behaviors that are carried out through computer networks and modern communication techniques.

3 INVESTIGATE ON SOCIAL MEDIA CRIME

Forensic analysis of social networks is the method used by investigators to identify and prosecute dangerous subjects present in a social network service.

The acquisition and collection of evidence of defamation is a key element, but it is necessary to certify and verify the integrity and authenticity of the collected evidence. The authentication of the page or profile that has carried out the defamation, insult or slander can be followed by a Notary equipped or by a forensic computer expert who acquires the pages with the messages defamatory or abusive. The digital appraisal aimed at documenting the defamation and the offense or injury occurred on the Internet through computerized evidence that can be extended through OSINT investigations and searches also to the acquisition of data relating to the owners or users of the profiles, groups or pages on where defamatory messages are published.

3.1 OSINT & SOCMINT

The Open Source INTelligence (OSINT) is the activity of collecting information by consulting publicly accessible sources. OSINT sources are distinguished from other forms of intelligence because they must be legally accessible to the public without violating any copyright or privacy laws. Indeed, OSINT includes all sources of information accessible to the public. This information is available online or offline, following some examples:

- Access to the Internet, which includes forums, blogs, social networking sites, video sharing sites, wikis, Whois records of registered domain names, metadata and digital files, geolocation data, IP addresses, people's search engines and everything that can be found online.
- Traditional mass media (TV, radio, journals, book).
- Specialized journals, academic publications, dissertations, conference proceedings, company profiles, annual reports, company news, employee profiles and resumes.

There are organizations specializing in OSINT services. Some of them are based on government services others are private companies that offer their services to various entities such as government agencies and commercial companies on a subscription basis; among the best known: government bodies, international organizations, military agencies, but also companies whose power is information.

The term of Social Media Intelligence (SOCMINT) indicates a set of techniques and technologies that allow private or public agencies to monitor social media platforms. SOCMINT's activities concern the monitoring of content, such as messages or photos posted, and any other kind of data produced during an activity session on social media. Such information, whether private or public, involves interactions between people, between people and groups or between different groups. The methods used to analyze the data produced through social networks are different: they may also include the manual correction of content, public or private, or of entire pages; o reviewing the results of some research or some questions; o modification of activities or content posted by the user; or scraping, which translated means scraping and which consists of extracting the content of a web page and duplicating it in a way accessible to those involved in social media intelligence. Clearly, SOCMINT's activity includes a series of procedures to collect, store, and analyze the data produced on social media, data that are subsequently translated into analyzes and trends. The term Social Media Intelligence is sometimes replaced by the equivalent Open Source Intelligence (OSINT), although there is a substantial difference between the two activities': while the OSINT analyzes only public data, such as articles, sites and blogs, SOCMINT analyzes both those public and private ones, i.e. messages and chats.

3.2 Identify Profile, Page or Group with Defamatory Content

In order to perform a correct analysis on a profile, page or group, it is necessary to identify the ID code that uniquely identifies it. The profile name can in fact be changed by the owner, as well as the address that appears in the browser's URL bar. To locate the ID code of the profile or page from which the defamation comes, you can use a site such as Find My FB ID (FindMyFbID, 2020), by pasting the profile or page address in the text field and pressing the "Find numeric ID" button. Once you have entered the address of the profile or page where the defamation is present, you will get a number to copy or print, to "freeze" the unique identifier that will allow you to find the profile or page even in the event of a name change or URL and to ask the Judicial Authority for any log files or defamatory content. If it is not possible to use online sites that identify the ID, it is advisable to save the page or profile on which the defamation was detected. Within the page code, you will find two items containing the ID codes searched: "pageID"; (for Facebook pages) and "profile id" (for profiles).

3.3 Find the Unique Reference of the Defamatory Post or Comment

Once the User ID of the owner of the profile from which the defamation occurred or the Page ID of the page containing the defamatory text has been established, it is also necessary to "freeze" the post or comment itself, including the date, to then use it as IT proof of the defamation and allow the IT forensic consultants who will be hired to carry out an IT expertise. The address or URL that identifies the post itself will be of the following type:

www.facebook.com/profile.name/posts/10213357451991856.

To identify a specific comment, by clicking on the date and time under the comment itself, after the "Like" link, the post will be opened in a new page with the comment highlighted, a URL of the following type:

www.facebook.com/profile.name/posts/10213357451991856?comment id=10213357955884453

The first code, highlighted in bold, is the ID code of the post, while the second one is the "comment id", that is the unique identifier of the defamatory comment.

3.4 “Freeze" a Digital Proof of Defamation

It is always important to make a certified copy of a profile or page containing defamatory messages. However, it is possible (also to protect oneself in the event of cancellation) to begin the "crystallization" phase using some precautions, such as the free FAW, Forensic Acquisition of Websites (FAW, 2020), or Legal Eye (Legal_eye, 2020) software that allows for the forensic acquisition of web pages or social profiles network with some guarantees on the originality of the acquired data. There are also web services that allow you to download an authentic copy of pages or posts as long as they are public and not private, such as Perma.cc or Archive.is that permits to create a copy of an Internet page on a third server, carried out by a third party, a strategic activity in particular in the event that the defamatory messages are modified or removed.

4 OVERVIEW ON FORENSIC ANALYSIS IN SOCIAL MEDIA ECOSYSTEM

In the recent years, Social Media Applications received attention from many forensic researchers, because of their exponential growth, due to their ease of use and efficiency reaching out to people, allow the development of many malicious activities and serious cybercrime (Mohtasebi & Dehghantanha, 2011).

In 2012 Al Mutawa et al. (Mutawa, Baggili, & Marrington, 2012) focus their attention on mobile device analyzing forensic artifacts of several Social Media apps on various mobile platforms: MySpace, Twitter and Facebook each on Blackberry phone, iPhone (iOS) and Android. In 2013, M. Baca et al (Baca, Cosic, & Cosic, 2013) conduct an analysis of Facebook artifacts in internet and were able to find significant evidence traces related to Facebook activity. Other research based on the analysis of

WhatsApp, Viber and Skype artifacts was carried out (Mahajan, Dahiya, & Sanghvi, 2013) (Thakur, 2013) (Al-Saleh, I., & Forihat., 2013). In 2015 Walnycky et al. (Walnycky, Baggili, Marrington, Moore, & Breitinger, 2015) conduct a network and device forensic analysis of twenty android social messaging apps to explore digital evidence strictly limited to messaging service only. A forensic analysis of three social media apps (Facebook, Viber and Skype) in windows 10 was carried out by Majeed et al. (Majeed, Zia, Imran, & Saleem, 2015). They explored and examined the potential locations of storage finding interesting artifacts for all three applications in plain text. In 2017 Yusoff et al. (Yusoff, Dehghantanha, & Mahmod, 2017) conduct an investigation and analysis of social media and instant messaging services focus on residual remnants of forensics value in FireFox OS. They examined three social media services (Facebook, Twitter and Google+) as well as three instant messaging services (Telegram, OpenWapp and Line).

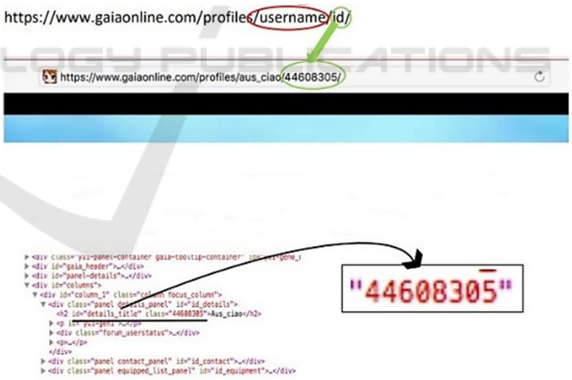
A very interesting research focused on authorship attribution for Social Media Forensics was conduct by Rocha et al. (Rocha, et al., 2016). Their research is based on the fact that all authors possess peculiarities of habit that influence the form and content of their written works. These characteristics can often be quantified and measured using machine learning algorithms. Rocha et al. provide a comprehensive review of the methods of authorship attribution that can be applied to the problem of social media forensics. Further, they examine emerging supervised learning based methods that are effective for small sample sizes, and provide step-by-step explanations for several scalable approaches as instructional case studies for newcomers to the field.

A work on Forensics of Social Network Relationship based on Big Data was carried out in 2020 by Junjing et al. (Junjing, Yan, & Jinqiang, 2020). They expound the forensics mode of social network relationship and the forensics process of mobile phones, and puts forward the forensics method of social network relationship based on Wechat platform, analyses the instance data set, obtains the social network diagram, and intuitively and clearly shows the relationship and intimacy between multiple members. The possibility to extract information about images uploaded on social platform has been exploited in (Giudice, Paratore, Moltisanti, & Battiato, 2017) / (Moltisanti, Paratore, Battiato, & Saravo, 2015).

Nowadays images are the main way by which people and companies share news and opinions through social platforms. Fake images (e.g., deep fake) and mislead images are under severe discussion by technicians in this last period. The field that studies how and why images are used to convey opinions (i.e., sentiment) (Ortis, Farinella, & Battiato, 2020) is named visual sentiment analysis. Related to this field there are several possible tasks, which would benefit from forensics analysis. Moreover, the outcomes of such analyses could be used as evidences (e.g., popularity dynamics, etc.).

4.1 Analysis Results

Our analysis has been conducted on the 250 social networks listed by (Social Media List, 2020). We have examined the different social networks in order to find the way to identify useful information that can be used in forensic investigations, like User Id, Post Id, Comment Id and any other to retrieve univocally to the defamation author. For each Social Network website, different type of approaches has been adopted to trace these identifiers: an inspection and analysis of the URL of the page or profile of interest, and a "Inspect Element" to examine the source code of the profile or post page; the inspection element analysis, mainly in HTML code, aims to find a script containing the string with the useful information for the investigation (for example an ID Users or others relevant ID) .



```
https://www.gaiaonline.com/profiles/username/id/

https://www.gaiaonline.com/profiles/aus_ciao/44608305/
```

```
▶ <div class="yui-panel-container gaia-tooltip-container" id="yui-gen0_(
▶ <div id="gaia_header">…</div>
▶ <div id="panel-details">…</div>
▼ <div id="columns">
  ▼ <div id="column_1" class="column focus_column">
    ▼ <div class="panel details_panel" id="id_details">
        <h2 id="details_title" class="44608305">Aus_ciao</h2>
      ▶ <p id="yui-gen1">…</p>
      ▶ <div class="forum_userstatus">…</div>
      ▶ <p>…</p>
      </div>
    ▶ <div class="panel contact_panel" id="id_contact">…</div>
    ▶ <div class="panel equipped_list_panel" id="id_equipment">…</div>
```

Figure 2: An example of URL analysis and Code Inspection applied on a Social Network Website.

Figure 2 shows an example of URL analysis and Code Inspection applied on a Social Network website.

Our analysis has shown that is not always possible to use both these approaches for every social network examined, or they are not always able to discover interesting information to be used in front of a court in the event of any cases of defamation.

In some cases the forensic study has not been possible to conduct; for example when it is not able to extract useful evidences or when the social site has characteristics that do not allow the defamation crime.

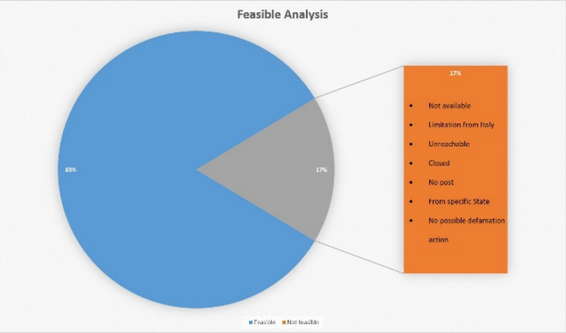


Figure 3: Our results show that in the 83% of the cases is possible apply a forensic analysis of the Social Network Platform obtaining useful evidences.

Figure 3 shows our analysis results conduct on 250 Social Network Platform: in the 83% of the cases it was possible to pull out useful information applying the forensics standard approaches said before. In the remaining 17% of the cases it was not possible conduct a forensic analysis because of different reasons (closed web site, unreachable app or website or because there are not post on the website to analyze).

In the Social Media Forensics the main information to detect defamation author are basically obtained from URL analysis or Code Inspection. Focus our attention on the inspectable Social Network Websites, in Figure 4 we show how many Social Media is possible to investigate with a URL analysis approach, how many with a Code Inspection approach and how many with both of them.

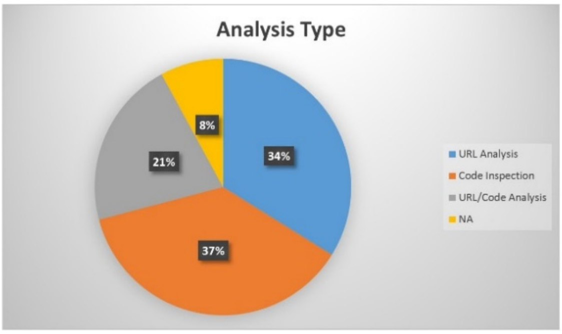


Figure 4: Types of Forensic approaches applicable on Social Media Websites.

In the 8% of the analyzed cases it is not possible to proceed with these techniques but it is necessary to contact the social network provider to recover valued information.

Figure 5 shows the number of Social Network where it was possible pull out some evidences such as IDUser, ID Post and ID Comment.

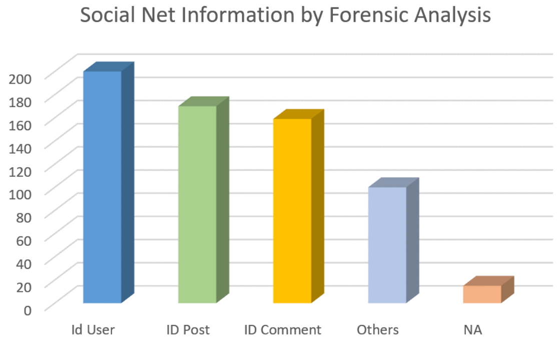


Figure 5: Typology of mainly Information that is possible extract with a forensic Social Network Analysis.

Moreover it is possible to find also other numerous information that characterize social network as shown in figure 6.

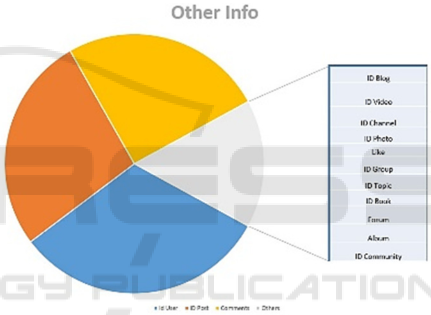


Figure 6: Other typology of information that is possible extract with a forensics Social Network Analysis.

5 CONCLUSIONS

Social networks are changing the way forensics examinations are done. In this paper we achieve an overview on how is possible to conduct a digital forensics investigation on 250 different Social Network Platforms to search and identify potential evidences. Our research highlights that in the most of the cases, also without the collaboration of the social network service, a forensic investigation is able to identify relevant potential information that can be presented as proofs in front of a court to pursue, in particular, the defamation crime. Our evaluation demonstrates that, in many cases, it is reasonably possible to extract forensic useful information of a given social networking account, of a given post or comment.

Moreover it has been shown how the main forensic approaches can be applied, specifically on this particular type of investigation, and how is possible to collect the evidences.

Future work in this domain could be related to an implementation of an automatic tool able to provide a digital evidence collection by means of a forensic investigation on social networking activities.

REFERENCES

Al-Saleh, I., M., & Forihat., Y. A. (2013). Skype forensics in android devices. *International Journal of Computer Applications*, (pp. 38-44).

Baca, M., Cosic, J., & Cosic, Z. (2013). Forensic analysis of social networks (case study). *Information Technology Interfaces (ITI), Proceedings of the ITI 2013 35th International Conference on* (pp. 219-223). IEEE.

Boyd, D. m., & Ellison, N. B. (2007). Social Network Sites: Definition, History and Scholarship. *Journal of Computer-Mediated Communication*, (pp. 210-230).

FAW. (2020). https://it.fawproject.com/.

FindMyFbID. (2020). https://findmyfbid.in/.

Gao, H., Hu, J., Wilson, C., Li, Z., Chen, Y., & Zhao, B. (2010). Detecting and characterizing social spam campaigns. *Proceedings of the 10th annual conference on Internet measurement* (pp. 35-47). ACM.

Giudice, Paratore, Moltisanti, & Battiato. (2017). A classification engine for image ballistics of social data. *International Conference on Image Analysis and Processing*, (pp. 625-636).

Jagatic, T., Johnson, N., Jakobsson, M., & Menczer, F. (2007). Social phishing. *Communications of the ACM*, (pp. 94-100).

Junjing, T., Yan, B., & Jinqiang, M. (2020). Research on Forensics of Social Network Relationship Based on Big Data. *Journal of Physics: Conference Series 1584 012022. DMCIT 2020*. doi:10.1088/1742-6596/1584/1/012022

L. Bilge, T. S., Balzarotti, D., & Kirda, E. (2009). Allyour contacts are belong to us: automated identity theft attacks on social networks. *Proceedings of the 18th international conference on World wide web* (pp. 551-560). ACM.

La Stampa. (2018). Retrieved from https://www.lastampa.it/cultura/2018/02/03/news/quanti-social-network-esistono-1.33975738

Legal_eye. (2020). https://www.legaleye.cloud/public.

M. Huber, M. M. (2011). Friend-in-the-middle attacks: Exploiting social networking sites for spam. *Internet Computing.*

Mahajan, A., Dahiya, M. S., & Sanghvi, H. P. (2013). Forensic analysis of instant messenger applications on android devices. arXiv preprint arXiv:1304.4915.

Majeed, A., Zia, H., Imran, R., & Saleem, S. (2015, December). Forensic analysis of three social media apps in windows 10. *2015 12th International Conference on High-capacity Optical Networks and Enabling/Emerging Technologies (HONET)*, (pp. 75-79). Islamabad (Pakistan). doi:10.1109/HONET.2015.7395419

Mohtasebi, S., & Dehghantanha, A. (2011). Defusing the Hazards of Social Network Services. *Int. J. Digit. Inf. Wirel. Commun.*, 504-516.

Moltisanti, M., Paratore, A., Battiato, S., & Saravo, L. (2015). Image manipulation on facebook for forensics evidence. *International Conference on Image Analysis and Processing*, (pp. 506-517).

Mutawa, A., Baggili, & Marrington. (2012). Forensic analysis of social networking applications on mobile devices. *DIgital Investigation, 9*, S24-S33.

Ortis, A., Farinella, G. M., & Battiato, S. (2020). Survey on visual sentiment analysis. *IET Image Processing*, (pp. 14(8), 1440-1456).

Rocha, A., Scheirer, W. J., Forstall, C. W., Cavalcante, T., Theophilo, A., Shen, B., Stamatatos, E. (2016). Authorship Attribution for Social Media Forensics. *IEEE Transactions On Information Forensics And Security.*

Smith, A., & Anderson, M. (2018, March 1). Social Media Use in 2018. Pew Research Center. Retrieved from www.pewresearch.org

Social Media List. (2020). Retrieved from https://socialmedialist.org/social-media-apps.html

Thakur, N. S. (2013). Forensic analysis of WhatsApp on Android smartphones. University of New Orleans Theses and Dissertations.

Walnycky, D., Baggili, I., Marrington, A., Moore, J., & Breitinger, F. (2015). Network and device forensic analysis of Android social-messaging applications. *Digital Investigation Impact Factor: 0.99.*

Wikipedia. (2020). Retrieved from https://en.wikipedia.org/wiki/JPEG.

Wikipedia. (2020). Retrieved from https://en.wikipedia.org/wiki/List of social networking websites.

Wondracek, G., Holz, T., Kirda, E., & Kruegel. (2010). A Practical Attack to De-Anonymize Social Network Users. *Proceedings of the IEEE Symposium on Security and Privacy.*

Yusoff, M. N., Dehghantanha, A., & Mahmod, R. (2017). Forensic Investigation of Social Media and Instant Messaging Services in Firefox OS: Facebook, Twitter, Google+, Telegram, OpenWapp and Line as Case Studies. *Contemporary Digital Forensic Investigations Of Cloud And Mobile, Chapter 4*, 41-62.