INTENSIF, Vol.7 No.2 August 2023
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

Exploring Alternative Approaches for TwitterForensics: Utilizing Social Network Analysis to Identify Key Actors and Potential Suspects

Received:
29 October 2022
Accepted:
17 July 2023
Published:
1 August 2023

[1*] Irwan Sembiring, [2]Ade Iriani, [3]Suharyadi
[1-3]Faculty of Information Technology,
Universitas Kristen Satya Wacana
E-mail: [1]irwan@uksw.edu, [2]ade.iriani@uksw.edu,
[3]suharyadi@uksw.edu
*Corresponding Author

Abstract—SNA (Social Network Analysis) is a modeling method for users which is symbolized by points (nodes) and interactions between users are represented by lines (edges). This method is needed to see patterns of social interaction in the network starting with finding out who the key actors are. The novelty of this study lies in the expansion of the analysis of other suspects, not only key actors identified during this time. This method performs a narrowed network mapping by examining only nodes connected to key actors. Secondary key actors no longer use centrality but use weight indicators at the edges. A case study using the hashtag "Manchester United" on the social media platform Twitter was conducted in the study. The results of the Social Network Analysis (SNA) revealed that @david_ornstein accounts are key actors with centrality of 2298 degrees. Another approach found @hadrien_grenier, @footballforall, @theutdjournal accounts had a particularly high intensity of interaction with key actors. The intensity of communication between secondary actors and key actors is close to or above the weighted value of 50. The results of this analysis can be used to suspect other potential suspects who have strong ties to key actors by looking.

Keywords— SNA, Twitterforensics; Secondary Key Actor; Key Actor

Corresponding Author:

Irwan Sembiring,
Information System,
Universitas Kristen Satya Wacana,
Email: irwan@uksw.edu
ID Orcid: http://orcid.org/0000-0002-6625-7533

INTENSIF, Vol.7 No.2 August 2023
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

I. INTRODUCTION

Social media sites like Twitter in contemporary society provide innovative concepts and expertise across a wide range of sectors. This development has both advantages and disadvantages. According to the statistics, out of the 13,169 tweets that were gathered, 5,561 contained hate speech, of which 3,575 were directed at specific people and 1,986 were directed at a group [1]. According to Patrosiber Data Indonesia, there were up to 4228 complaints of online libel between October 2020 and November 2021[2]. In order to gather data and provide an answer to 5WH, a forensic digital investigation is unavoidably required in this scenario [3][4].

A standard framework has been used to carry out the digital forensic procedure. Collection, examination, analysis, and reporting make up the whole investigative framework. The Systematic Digital Forensic Investigation Model (SRDFIM), which concentrates on cybercrime and cyber fraud investigations, is one of the frameworks that are frequently utilized in conducting digital forensic operations [5]. The Integrated Digital Forensics Process Model (IDFPM), which suggests a four-step model of Preparation, Incident, Digital Forensics investigation, and Presentation, is another framework [6]. A more recent system called D4I makes the semi-automated examination and investigation of cyberattacks [7]. These existing frameworks, when implemented on social media platforms, provide broad and superficial analysis. For instance, it fails to identify the communication patterns between important actors and other participants. These patterns are crucial for extending the analysis and gaining clarity on the suspect's role. For the findings of the research to characterize user profiles (identifying) and communication patterns between actors, forensic analysis on social media calls for a distinct methodology.

Currently, the SNA (Social Network investigation) approach is used to do forensic investigations on social media [8]. The purpose of this study is to obtain other suspects besides the main suspect, this suspect profile can be seen using the SNA method. SNA modeling is necessary to observe patterns of social interaction within a network, which starts with identifying the key actors [9]. A similar study related to digital forensics was conducted using the Twitter social media platform, employing a combination of PCAP (Packet capture) analysis and network forensics. However, the communication patterns between nodes could not be determined [10]. The contribution of this research lies in the expansion of the analysis of other actors with SNA, not only the factors that have been found so far but also by looking at other actors who have the potential to be used as other suspects based on additional values of weight and edges.

INTENSIF, Vol.7 No.2 August 2023
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

II. RESEARCH METHOD

This study contributes by adopting a modified digital forensic framework and incorporating analysis using SNA, as seen in figure 1.

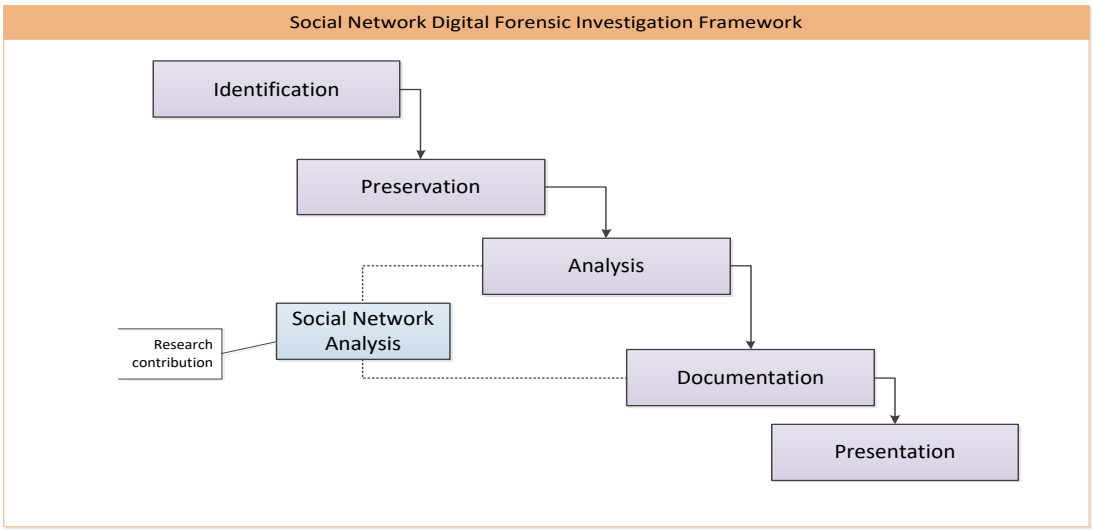

Fig 1. Research Stages

2.1 Identification

Manchester United is a term that was used in the data collecting of network interaction and information distribution on Twitter. With the help of the TwitterStreaming Importer plugin for Gephi applications, data is collected using data crawling techniques. There are 7677 nodes in the data that has been gathered. Id, label, timestamp, created_at, description, profile, actual name, and location make up the identifying field for investigative requirements. Finding prospective sources of information about the incident, identifying them, and then classifying and logging them is their aim. The information included inside must be acquired later in order to preserve the source's integrity [11]. Digital forensics places a lot of importance on this procedure. Priority is given to sterilizing the evidence. Digital forensics begins when electronic evidence is gathered at the crime scene. When evidence is first retrieved from a crime scene, mistakes can result in the loss of crucial details about the crime that is being investigated or even the rejection of electronic evidence for use as evidence in court.

2.2 Preservation

Both in terms of form and substance, digital evidence must be kept in a clean environment. To be sure there are no alterations, keep this in mind. Because even little alterations to digital evidence might alter an investigation's conclusions. Digital evidence can easily be damaged, lost, changed, or deleted in an accident since it is by its very nature transient (volatile). Data isolation, data security, and data maintenance are performed at this stage. To protect data from the possibility of data degradation, the imaging or cloning procedure is crucial at first. The process

INTENSIF, Vol.7 No.2 August 2023
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

of gathering, examining, and disclosing digital data in order to get information or data of evidential value is known as digital forensics (DF) [12][13]. It is impossible to undervalue the role that DF plays in contemporary criminal investigations. Keeping up with technological advancements and inventions in numerous criminal cases presents this sector with everyday challenges. [14] [15]. The Digital Forensics Workflow Model (DFWM) [16] describes workflow stages as digital forensic procedures that include data sizing and planning, identification, handling, preservation, and collection. Digital forensic models form the basis for digital investigation. This model guides researchers with the steps and procedures to be taken during research [17][18].

Digital forensics generally performs five steps of work, namely (1) identification which includes what evidence exists and where it is stored in what format, (2) preservation aims to maintain the originality of evidence, (3) analysis with the aim of investigating the reconstruction of data fragments and drawing conclusions. in accordance with the evidence found, (4) Documentation, namely documenting all evidence including appropriate cases and making chain of custody, (5) Presentation is the process of making reports and conclusions to be submitted to law enforcement [19]. One of the most important processes at the digital forensics stage is data integrity in the preservation section. MD5 and SHA-1 message-digest algorithms as one-way cryptographic hash functions are used in integrity validation [20][21]. Four non-linear functions in a 512-bit block in the MD5 Algorithm as equation 1. The primary MD5 method uses a 128-bit state that is split into four 32-bit words called A, B, C, and D. These are set to a set of preset constants upon initialization. The primary algorithm then modifies the state using each of the 512-bit message blocks in turn. Four comparable steps, known as rounds, make up the processing of a message block; each round is made up of 16 related operations based on the non-linear function F, modular addition, and left rotation. There are four potential functions, and each round uses a new one, $\wedge \oplus \wedge \vee \neg$ denote the XOR, AND, OR and NOT operations respectively.

$$\begin{gathered} F(B,C,D) = (B \wedge C) \vee (\neg B \wedge D) \\ G(B,C,D) = (B \wedge D) \vee (C \wedge \neg D) \\ H(B,C,D) = B \oplus C \oplus D \\ I(B,C,D) = C \oplus (B \vee \neg D \end{gathered} \qquad (1)$$

Another alternative that is often used as a second comparison is the method of data validation by hashing SHA-1 as shown in equation 2 by adding bit "1" to the message, then adding k bit "0", where k is the minimum number 0 to the length of the message. congruent with 448 (mode 512).

$$\begin{gathered} f = (B \wedge C) \vee (\neg B \wedge D) \\ f = B \oplus C \oplus D \\ f = (B \wedge C) \vee (B \wedge D) \vee (C \wedge D) \\ f = B \oplus C \oplus D \end{gathered} \qquad (2)$$

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

The non-profit organization ACPO (Association of Chief Police Officers) has created regulations for law enforcement in England, Wales, and Northern Ireland. Since its founding in 1948, Acpo has served as a platform for police personnel to exchange ideas, plan strategic operational responses, and, in some situations, such as terrorist attacks and public emergencies, advise governments. ACPO organizes cooperative law enforcement, extensive investigations, cross-border law enforcement, and national police operations. Law enforcement and all those who assist in the investigation process of cybersecurity incidents and crimes utilize materials released by ACPO as guidance [22].

1. No action taken by law enforcement organizations, people working for them, or law enforcement officials can replace information that can later be accounted for in court.
2. In situations when access to original data is required, the staff who need it must be competent and able to support its needs with evidence demonstrating its importance and significance.
3. Digital evidence processing requires the creation and maintenance of audit trails or other records of the entire process. To analyze such methods and produce the same results, there must be an impartial third party.
4. It is entirely the responsibility of the team working on the investigation to ensure that these rules and guidelines are followed.

Next, the imaging results were tested using the MD5 and SHA-1 message-digest algorithms as one-way cryptographic hash functions [2,7]. For data validation and integrity, all preservation stages are documented in the chain of custody. To ensure that the imaging findings align with the physical evidence, forensic imaging methods are now employed to create duplicates of electronic evidence, such as imaging physical hard disks. The hashing operations applied to both can ascertain the degree of similarity between them.

2.3 Analysis

SNA is modeling users represented by nodes and interactions between users represented by edges, this analysis is necessary because it brings new opportunities to understand individuals or communities about their social interaction patterns [23][24]. SNA can be used to study network patterns of organizations, ideas, and people connected in various ways in the environment [25] [26].

1. Degree centrality counts the number of connections or interactions a node has. To calculate the value of the degree of centrality (CD), i = node, N = total node, representation of a node, X = neighbor matrix, Xij = connectedness value like equation (3) [27].

INTENSIF, Vol.7 No.2 August 2023
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

$$CD(i) = \sum_{\substack{j=1 \\ i \neq j}}^{N} Xij \qquad (3)$$

2. Closeness centrality (CC) calculates the average distance between a node and all other nodes on the network, N= Total Node, dij = The shortest number of paths connecting the ni and nj nodes. This measure describes the proximity of this node to other nodes [14] as in equation (4).

$$cc(i) = \frac{N-1}{\sum_{J=1}^{N} dij} \qquad (4)$$

3. Betweenness centrality (BC) calculates how often a node is passed by another node to go to a particular node in the network, gjk(i) = number of shortest paths from node j to node k passing through node I, gjk = number of shortest paths between two nodes in the network. This value serves to determine the role of the actor who is the bridge that connects interactions in the network. To calculate the value of degree centrality such as equation (5) [27].

$$Cb(i) = \sum_{J=1}^{N} \sum_{k=1}^{j-1} \frac{gjk(i)}{gjk} \qquad (5)$$

At this point, the image file is thoroughly examined with the aim of obtaining digital data consistent with the investigation. To achieve this, the forensic analysis must receive a complete picture of the facts of the case from the investigator, ensuring that what the forensic analysis is looking for and ultimately finding is the same (matching).) As expected by the investigator for the progress of his investigation. After getting an overview of the facts of the case, the forensic analysis searches the image files for the required files or data. In this phase in addition to the general SNA approach calculates the degree of centrality, closeness centrality, and centrality between centralities. The contribution in this study added a weighted value to the edges used to measure the intensity of interaction between nodes. The Weight parameter is a value on the edges that is used to measure the intensity of interaction between nodes. The value of this weight parameter will later be used to determine the secondary key actor where A, B on the left = Origin Nodes and B, C on the right = Destination Nodes as equation (6).

$$A \rightarrow B, B \rightarrow C \qquad (6)$$

Equation 4, when applied to the data in the Gephi application, will result in a weight value of 1 at origin node A of goal B, and result in a weight value of 1 at origin node B of destination C. After the previous evaluation process produces the required file or digital data, the data is thoroughly and deeply examined to establish the nature of the crime and the relationship between the offender and the relationship between the offender and the criminal offender. The results of

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

the examination of digital data are referred to as digital evidence as follows and are subject to scientific and legal accountability in court.

2.4 Documentation

Documentation is the process of sequencing the data obtained and the findings of the analysis so that they can be accounted for or, if necessary, recreated for the conclusion of evidence. The steps involved in creating a report on the results of the process of viewing through and analyzing digital evidence before entering data into a technical report. In this process, a record of all visible data must be created. It helps in creating a crime scene or crime scene and reviewing it. In this case it involves proper crime scene documentation along with photographing, sketching, and mapping the crime scene. In this study, due to data collection on social media, the time and profiles of accounts that interacted were documented.

2.5 Presentation

In this last step, the process summarizes and explains the conclusions in detail. This stage is to explain who the key and secondary actors are according to the results of the analysis carried out. The determination of the report format can be adjusted to the needs of law enforcement. This stage also explains the importance to which digital evidence will be examined, authenticated, and compared with original evidence. This is important because the judge will be presented with evidence and information that has been tested for authentication. Testing of evidence from forensic digital analysis compared with original evidence will be tested using hashing techniques. What has been analyzed must be ensured not to change the integrity of the original evidence. This form of testing is carried out in digital forensics using MD5 and SHA-1 message-digest algorithms as one-way cryptographic hash functions used in data integrity validation.

III. RESULT AND DISCUSSION

Social Network Analysis (SNA) in research is used to analyze communication patterns among nodes. SNA analysis helps to obtain more comprehensive information regarding the relationships between nodes. This is essential in determining the potential involvement of other suspects besides the main suspect, which was previously only observed based on key actors. Research related to the topic of Twitter forensic as a database in several analytical purposes including, Berube et al [28] using machine learning techniques, especially natural language processing and modeling the topic of Latent Dirichlet Allocation (LDA) at the Manchester arena bombing 2017. The base was collected from Twitter in the first 24 hours after the attack. These findings make it possible to track different types of social reactions over time and to identify sub-events that have a significant impact on public perception. Priyanka and Satheesh stated that approximately 6,000 tweets are sent every second on Twitter. Forensic analysis of the Twitter app is crucial for crime

INTENSIF, Vol.7 No.2 August 2023
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

investigators because it can contain a rich collection of evidence artifacts. The physical acquisition of an Android device could uncover forensic artifacts stored in the Twitter app's database, but only recent tweets and messages. This paper introduces a new methodology for forensically extracting Twitter cloud data using existing access tokens on Android devices. The token allows investigators to gain authenticated access to Twitter's cloud servers and further access to the rest of the data using Twitter's APIs. Umrani et al [29] conducted a network traffic analysis of Twitter, a popular social media application that uses encryption to protect information transmitted over the network. We concentrate on the Android platform to generate Twittertraffic, analyze fixed patterns, and extract artifacts based on various user actions. The firewall is also employed to examine the adaptability of Twitter's concealed design and explore alternative connectivity options.

3.1 Identification and Preservation

The identification process commences with the collection of data by querying keywords related to "Manchester United" using the Twitter streaming importer plugin within the Gephi application. The data is obtained by selecting the user's network to apply network logic. The collected data consists of user interactions, which are categorized as nodes. Subsequently, a cleaning process is conducted to eliminate attributes that are not relevant for this analysis. The data that has been collected is 7677 nodes. The identification fields for investigation need consist of *id, label, timestamp, created_at, description, profile, real name, and location* as shown in figure 2.

Id	created_at	description	followers_count	real_name	location	closnesscentrality	betweenesscentrality
@david_ornstein	Sat Jun 13 19:09:04 ICT 2009	Football Correspondent, @TheAthleti	878011	David Ornstein	UK	0.366323	0.350327
@purelyfootball	Fri Oct 25 04:46:43 ICT 2013	PurelyFootballÂ®ï¸Premier League f	564689	PurelyFootball	England, United King	0.342648	0.284139
Bringing you the Late	Match Reports	Stats	54	28	Manchester United f	10	0.184542
@footballfunnnys	Mon Dec 03 01:47:35 ICT 2012	Bringing you the best stats, facts, pict	823934	FootballFunnys	United Kingdom	0.295945	0.109578
@cr7_inside	Mon Nov 12 00:18:56 ICT 2018	ðŸ¨â€ðŸ» ADM : @souheilK07	63580	CR7 inside âž	MANCHESTER IS RED	0.236573	0.102818
@skysportspl	Tue Jul 24 16:18:08 ICT 2012	The official account for the Sky Sports	6511682	Sky Sports Premier League	London	0.311257	0.101523
@football__tweet	Fri Jan 13 03:31:01 ICT 2012	Your home of football discussion.	616491	Football Tweet âš½ï¸€	team@thesocialsup	0.301677	0.094929
@hadrien_grenier	Mon Apr 30 00:30:22 ICT 2012	ðŸ²² Journaliste @le10sport | ðŸ¨â€	93359	Hadrien Grenier	Paris, France	0.285291	0.046882
@goal	Fri Mar 27 01:14:06 ICT 2009	Sharing footballâ€™s passion â	2973774	Goal	Everywhere	0.294113	0.043657
@skysportsnews	Tue Jan 26 17:22:29 ICT 2010	The official Twitter account for Sky Sp	9280574	Sky Sports News	England	0.290214	0.038853
@leaguetotal	Sun Aug 11 23:21:15 ICT 2019	Compte FranÃ§ais pour les fans de Pr	17152	Total Premier League ðŸ‡¬ðŸ‡		0.239693	0.033516
@devilsofunited	Mon Oct 28 23:51:24 ICT 2013	Manchester United Football Club is ou	141412	Devils of United ðŸ˜	18+	0.267842	0.030157
@elou_pnl	Sat Jun 23 04:41:53 ICT 2018	Quâ€™une vie est heureuse ðŸ¨quai	3069	El Hadj (ðŸ‡²ðŸ‡·ðŸ‡¸ðŸ‡³)	Mauritania	0.266235	0.029429
@thelifeofmino	Sun Apr 23 06:21:36 ICT 2017		105	Certified Lover Boy		0.266235	0.029429
@caleb_mufc	Wed Nov 27 04:47:18 ICT 2013	ð?½ð?‡±ð?¼ð?† ð?‡±ð?„ð?‘€	9486	ð?Šð?¢ð?ð?¦ð?£âš¡	ðŸŒŠ	0.280439	0.029362
@ericlaurie	Sat Mar 06 23:47:50 ICT 2010	Performance Analyst & Academy Coac	37844	Eric Laurie		0.287028	0.027783
@kwasigazy	Mon Oct 22 16:14:00 ICT 2018	A Red Devil , Sarkodie	8710	Rashford	Accra, Ghana	0.23977	0.026739
@sansah_alfred	Thu Jul 29 18:25:34 ICT 2021	My uniqueness and how feeling I'm	89	Alfred Sansah		0.337673	0.022906
@teamcronaldo	Wed Nov 13 20:46:30 ICT 2013	Everything you need to know about @	226257	TCR.	Paris, France	0.315464	0.022707
@brfootball	Tue Jul 02 00:51:17 ICT 2013	Get the Free B/R App â¬‡ï¸	3478037	B/R Football		0.2886	0.020777
@manutd	Fri Apr 20 22:17:43 ICT 2012	The home of Manchester United. Sho	27191590	Manchester United	Old Trafford, Manch	0.270268	0.019256

Fig 2. Data Nodes Query Word *Manchester United*

After the data collection and identification process has been carried out, then the evidence is carried out by imaging as part of preservation. The imaging results were tested for integrity using

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

the MD5 checksum and SHA1 Hash methods. The file imaging in this case is validated as shown in Figure 3. The results of these two methods (MD5 and SHA1) show a match between the *computed hash* and the *report hash.* The FTK imager 3.1.2.0 application is used in the imaging and testing stages of evidence. The results of this test prove that the imaging results are exactly.

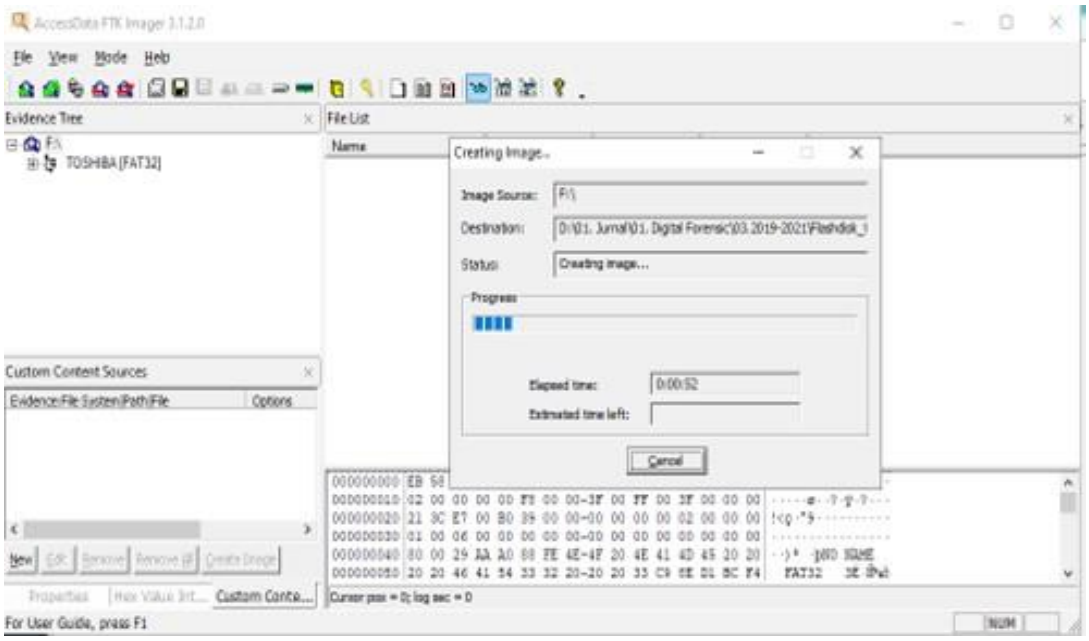

Fig 3. The Process of Evidence Acquisition.

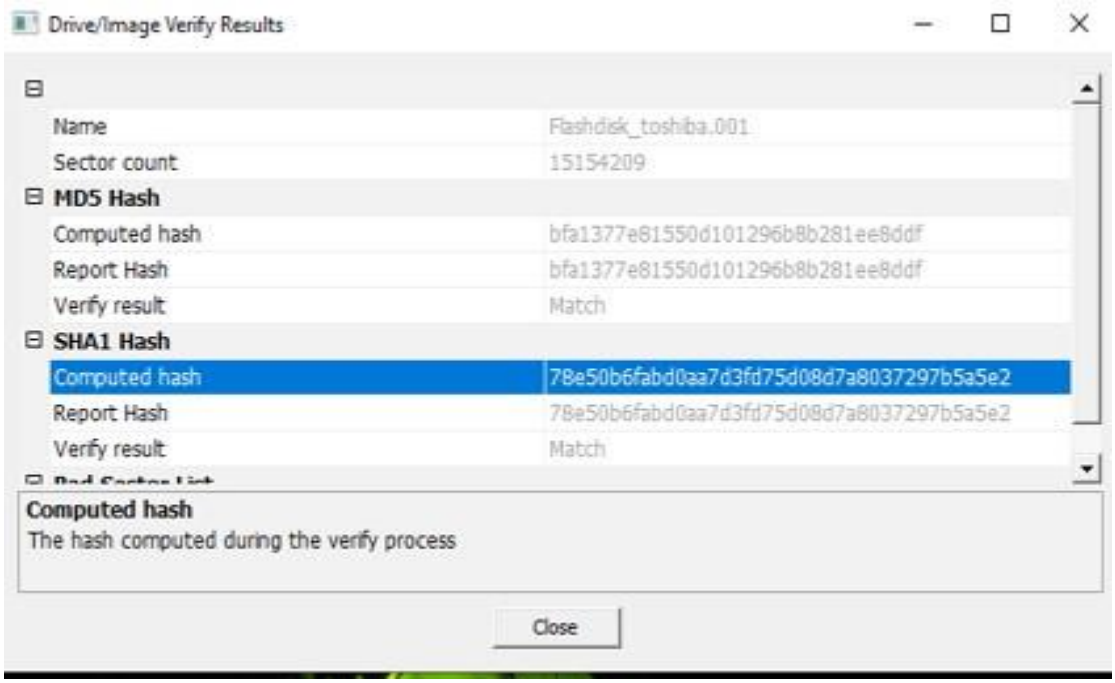

Fig 4. Integrity Checksum with MD5 and SHA.

In forensic science, the originality of evidence is the main key in the investigation. Digital evidence must be ensured that it is not contaminated during the search and seizure process. Proof is done by checking the integrity of files or documents using hashing with MD5 and SHA1

INTENSIF, Vol.7 No.2 August 2023
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

methods as shown in Figure 4. If this test fails or does not match, it cannot be continued in the next process.3.2 Forensic Twitter Analysis with SNA. Digital forensic analysis using SNA consists of determining the key actor and secondary actor. This assessment makes the investigation process more focused on certain actors.

3.2 Key Actor

The range of nodes from source to target needs to be known, including the type of communication, whether to mention or retweet. Figure 5 shows the edges that show the *source*, *target type of twitter: mention*, *retweet*, *quote* there are 15243 edges in this study.

Source	Target	Type	Kind	Id	timeset	Weight
@david_ornstein	@theathleticuk	Directed	Mention	2	<[2021-09-13T07:17:21.691Z, 202	1533
@hadrien_grenier	@david_ornstein	Directed	Mention	4502	<[2021-09-13T08:01:54.353Z, 202	100
@hadrien_grenier	@theathleticuk	Directed	Mention	4503	<[2021-09-13T08:01:54.353Z, 202	100
@footballlforall	@david_ornstein	Directed	Mention	11	<[2021-09-13T07:17:26.031Z, 202	86
@brfootball	@manutd	Directed	Mention	609	<[2021-09-13T07:22:58.259Z, 202	83
@theutdjournal	@david_ornstein	Directed	Mention	12998	<[2021-09-13T09:35:33.936Z, 202	47
@unitedsupdate	@david_ornstein	Directed	Mention	101	<[2021-09-13T07:18:15.041Z, 202	29
@overthebarfb	@daniireadd	Directed	Mention	10884	<[2021-09-13T09:13:40.264Z, 202	28
@teamcronaldo	@david_ornstein	Directed	Mention	7433	<[2021-09-13T08:32:49.186Z, 202	26
@talksport	@mrjamieohara	Directed	Mention	10536	<[2021-09-13T09:10:04.761Z, 202	24
@afcstuff	@arsenalacademy	Directed	Mention	249	<[2021-09-13T07:19:37.083Z, 202	23
@afcstuff	@charliepatino10	Directed	Mention	250	<[2021-09-13T07:19:37.083Z, 202	23
@munitedfr	@richjolly	Directed	Mention	559	<[2021-09-13T07:22:36.801Z, 202	22
@_jaffy	@utdnathan77	Directed	Quote	1427	<[2021-09-13T07:29:54.676Z, 202	22
@unitedredscom	@david_ornstein	Directed	Mention	43	<[2021-09-13T07:17:44.133Z, 202	19
@footyaccums	@david_ornstein	Directed	Mention	1147	<[2021-09-13T07:27:04.656Z, 202	19
@fermaldonado_11	@mediotiempo	Directed	Mention	1493	<[2021-09-13T07:30:39.415Z, 202	18
@fermaldonado_11	@mediotiempo	Directed	Retweet	1494	<[2021-09-13T07:30:39.415Z, 202	18
@psghub	@david_ornstein	Directed	Mention	5635	<[2021-09-13T08:12:38.812Z, 202	16
@psghub	@theathleticuk	Directed	Mention	5636	<[2021-09-13T08:12:38.812Z, 202	16
@nigerianewsdesk	@todayng	Directed	Mention	7520	<[2021-09-13T08:33:37.405Z, 202	15

Fig 5. Data edges query word *manchester united*

Determination of *key actors* is done to determine the level of influence of actors in the network. This is important to determine which actors play the most important role in communication patterns. The following is a sequence of key candidate candidates that have been sorted based on the value of degree centrality, closeness centrality and betweenness centrality degree on the topic of Manchester *United* as shown in figure 6.

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

Id	Degree Centrality	Closeness Centrality	Betweenness Centrality
@david_ornstein	2298	1	0.000028
@manutd		1	0.000009
@brfootball			0.000005
@hadrien_grenier		1	0.000003
@teamcronaldo			0.000003
@footballdaily		1	0.000002
@kwasigazy		1	0.000002
@iam_wilsons		1	0.000001
@munitedfr		1	0.000001
@talksport		1	0.000001
@purelyfootball	1922		
@footballfunnnys	993		
@cr7_inside	876		
@football__tweet	762		
@skysportspl	660		
@goal	368		
@skysportsnews	329		
@devilsofunited	278		
@leaguetotal	278		
@_iaffy		1	
@mufc_malaysia		1	

Fig 6. Assessment of DC, CC, and BC. values

SNA analysis results determine account @ david_ornstein is a key actor. This actor is included in the top 10 actors on every measurement taken. Measurements have been carried out with a combination of 2298 degree centrality connections, the distance between nodes (closeness centrality) is 1 and the frequency of a node being passed when communicating (betweenness centrality) is 0.000028. Visualization of information dissemination network interaction data using the keyword manchester united on the Twitter social networking site was carried out using Gephi software version 0.9.1. The data is visualized into a sociogram where the points in the image are called 'nodes' or vertices representing an individual who are connected by lines called 'vertices'. Two connected nodes are indicated by the presence of a line connecting them. The thicker the line, the greater the number of interactions that occur between the two nodes. Network visualization is done using the *Fruchterman Reingold algorithm* with an area of 10,000, a Gravity attribute of 10.0, a speed attribute of 100.0 and color visualization grouped by degree value. Modeling communication data as a network facilitates the discovery of patterns of relationships in complex networks so that visualization and measurement of relationships between actors can be carried out. The following is a visualization of the Manchester United Twitter thread interaction network as shown in Figure 7.

INTENSIF, Vol.7 No.2 August 2023
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

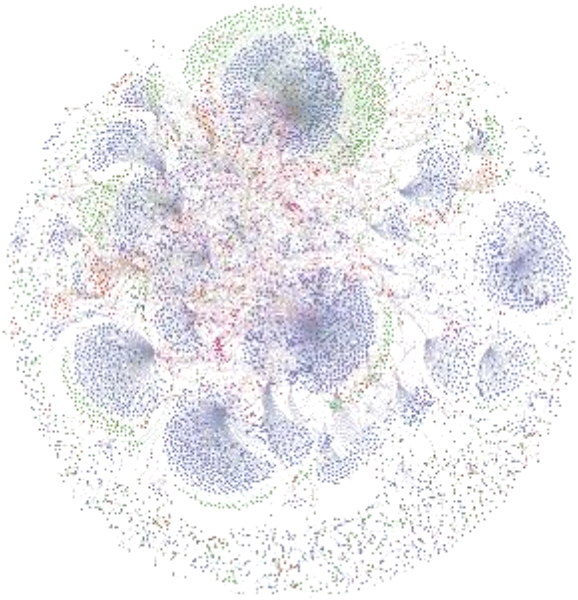

Fig 7. Visualization of The Manchester United Network

3.3 Secondary Key Actor

The secondary key actor refers to other actors outside of the key actor. These actors have the potential to become additional suspects based on the evidence of weight values. A high weight value indicates a high intensity of communication with the key actor. In general, legal cases, these actors can also be considered as intellectual actors. The complete dataset is initially filtered to retain only the nodes connected to the key actor. In the resulting visualization, the key actor is represented by the red node positioned at the center, and the thickness of the lines represents the weight value. The line thickness increases with higher weight values. Figure 8 is a data visualization depicting the network of nodes connected to the key actor.

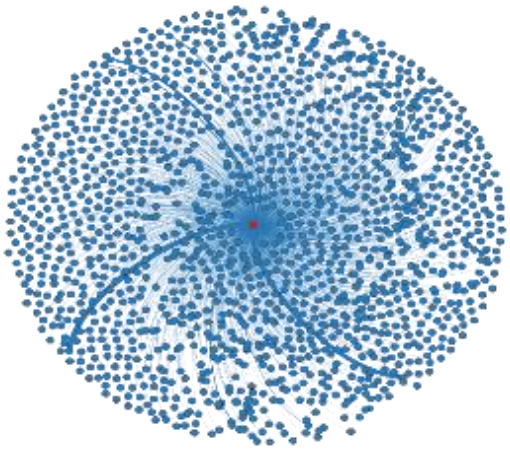

Fig 8. Visualization of Network Interaction Patterns With Key Actors.

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

As the focus of this research is to find the secondary key actor, there is an indicator that can be used to determine the intensity of the interaction between the key actor and the nodes connected to it so that the secondary key actor can be found. The indicator in question is the weight value on the edges, along with the top 10 weight values for each relationship between actors.

Source	Target	Weight
@hadrien_grenier	@david_ornstein	100.0
@footballlforall	@david_ornstein	86.0
@theutdjournal	@david_ornstein	47.0
@unitedsupdate	@david_ornstein	29.0
@teamcronaldo	@david_ornstein	26.0
@unitedredscom	@david_ornstein	19.0
@footyaccums	@david_ornstein	19.0
@psghub	@david_ornstein	16.0
@mufcscoop	@david_ornstein	11.0
@unitedpeoplestv	@david_ornstein	10.0
@football_taik	@david_ornstein	10.0

Fig 9. Relationship Between Nodes with Weight.

Figure 9 shows the three accounts that are considered as secondary key actors, namely users @hadrien_grenier, @footballforall, @theutdjournal, with the weight value being close to or above 50.

IV. CONCLUSION

Twitter forensics with SNA analysis has proven to be able to be carried out to expand the information on the results of the investigation in determining who is the main actor (keyctor) and the actor has the strongest relationship with the main actor (secondary actor). The case in the manchester united keyword query that becomes the key is the account id @david_ornstein and the secondary actor is the id @hadrien_grenier. This analysis method can be used as a source of information to determine the communication pattern of a Twitter thread. This research can be continued by collecting more complete fields related to coordinates and developing investigation methods on *anonymous accounts* with text mining.

REFERENCES

[1] M. O. Ibrohim and I. Budi, "Multi-label Hate Speech and Abusive Language Detection in Indonesian Twitter," 2019. doi: 10.18653/v1/w19-3506.

[2] "Statistik laporan masyarakat," 2021. https://patrolisiber.id/ (accessed Jun. 20, 2022).

[3] W. Anwar, I. S. Bajwa, M. A. Choudhary, and S. Ramzan, "An empirical study on forensic analysis of Urdu text using LDA-based authorship attribution," *IEEE Access*, vol. 7, pp. 3224–3234, 2019, doi: 10.1109/ACCESS.2018.2885011.

[4] X. Du, N. A. Le-Khac, and M. Scanlon, "Evaluation of digital forensic process models with

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

respect to digital forensics as a service," in *European Conference on Information Warfare and Security, ECCWS*, 2017, pp. 573–581.

[5] A. Agarwal, M. Gupta, S. Gupta, and S. C. Gupta, "Systematic digital forensic investigation model," *Int. J. Comput. Sci. Secur.*, vol. 5, no. 1, pp. 118–131, 2011, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.227.8647&rep=rep1&type=pdf

[6] M. D. Kohn, M. M. Eloff, and J. H. P. Eloff, "Integrated digital forensic process model," *Comput. Secur.*, vol. 38, 2013, doi: 10.1016/j.cose.2013.05.001.

[7] A. Dimitriadis, N. Ivezic, B. Kulvatunyou, and I. Mavridis, "D4I - Digital forensics framework for reviewing and investigating cyber attacks," *Array*, vol. 5, p. 100015, Mar. 2020, doi: 10.1016/J.ARRAY.2019.100015.

[8] A. Aslam, S. M. Maher, L. Kanwal, and M. A. Shah, "An aspect of internet of things security: Analysis of digital fingerprinting of generic Twittersessions by using forensic tool," *ICAC 2019 - 2019 25th IEEE Int. Conf. Autom. Comput.*, no. September, pp. 1–5, 2019, doi: 10.23919/IConAC.2019.8895172.

[9] Y. Wang, H. Sun, Y. Zhao, W. Zhou, and S. Zhu, "A Heterogeneous Graph Embedding Framework for Location-Based Social Network Analysis in Smart Cities," *IEEE Trans. Ind. Informatics*, vol. 16, no. 4, pp. 2747–2755, 2020, doi: 10.1109/TII.2019.2953973.

[10] I. Sembiring, Suharyadi, A. Iriani, J. V. B. Ginting, and J. A. Ginting, "A Novel Approach to Network Forensic Analysis: Combining Packet Capture Data and Social Network Analysis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 3, pp. 466–472, 2023, doi: 10.14569/IJACSA.2023.0140353.

[11] G. Bissias, B. N. Levine, M. Liberatore, and S. Prusty, "Forensic Identification of Anonymous Sources in OneSwarm," *IEEE Trans. Dependable Secur. Comput.*, vol. 14, no. 6, pp. 620–632, 2017, doi: 10.1109/TDSC.2015.2497706.

[12] P. Lewulis, "Digital forensic standards and digital evidence in Polish criminal proceedings. An updated definition of digital evidence in forensic science," *Int. J. Electron. Secur. Digit. Forensics*, vol. 13, no. 4, 2021, doi: 10.1504/IJESDF.2021.116024.

[13] D. Cozzolino and L. Verdoliva, "Noiseprint: A CNN-Based Camera Model Fingerprint," *IEEE Trans. Inf. Forensics Secur.*, vol. 15, no. 1, pp. 144–159, 2020, doi: 10.1109/TIFS.2019.2916364.

[14] P. Reedy, "Interpol review of digital evidence 2016 - 2019," *Forensic Sci. Int. Synerg.*, vol. 2, pp. 489–520, 2020, doi: 10.1016/j.fsisyn.2020.01.015.

[15] V. R. Kebande, P. P. Mudau, R. A. Ikuesan, H. S. Venter, and K.-K. R. Choo, "Holistic digital forensic readiness framework for IoT-enabled organizations," *Forensic Sci. Int. Reports*, vol. 2, p. 100117, Dec. 2020, doi: 10.1016/j.fsir.2020.100117.

[16] G. Horsman and N. Sunde, "Unboxing the digital forensic investigation process," *Sci. Justice*, vol. 62, no. 2, pp. 171–180, Mar. 2022, doi: 10.1016/J.SCIJUS.2022.01.002.

[17] D. Mothi, H. Janicke, and I. Wagner, "A novel principle to validate digital forensic models," *Forensic Sci. Int. Digit. Investig.*, vol. 33, p. 200904, Jun. 2020, doi: 10.1016/J.FSIDI.2020.200904.

[18] Q. Li, G. Sovernigo, and X. Lin, "BlackFeather: A framework for background noise forensics," *Forensic Sci. Int. Digit. Investig.*, vol. 42, p. 301396, Jul. 2022, doi: 10.1016/j.fsidi.2022.301396.

[19] N. M. Karie, V. R. Kebande, and H. S. Venter, "Diverging deep learning cognitive computing techniques into cyber forensics," *Forensic Sci. Int. Synerg.*, vol. 1, pp. 61–67, Jan. 2019, doi: 10.1016/J.FSISYN.2019.03.006.

[20] A. Mohammed Ali and A. Kadhim Farhan, "A novel improvement with an effective expansion to enhance the MD5 hash function for verification of a secure E-Document," *IEEE Access*, vol. 8, pp. 80290–80304, 2020, doi: 10.1109/ACCESS.2020.2989050.

[21] S. Long, "A Comparative Analysis of the Application of Hashing Encryption Algorithms for MD5, SHA-1, and SHA-512," in *Journal of Physics: Conference Series*, 2019, vol. 1314, no. 1. doi: 10.1088/1742-6596/1314/1/012210.

[22] DAC Janet Williams QPM, "Revised Good Practice Guide for Digital Evidence_Vers 5_Oct 2011_Website," 2012, [Online]. Available: https://www.npcc.police.uk/documents/crime/2014/Revised Good Practice Guide for Digital Evidence_Vers 5_Oct 2011_Website.pdf

[23] S. Sen Zhang, X. Liang, Y. D. Wei, and X. Zhang, "On Structural Features, User Social Behavior, and Kinship Discrimination in Communication Social Networks," *IEEE Trans. Comput. Soc. Syst.*, vol. 7, no. 2, pp. 425–436, 2020, doi: 10.1109/TCSS.2019.2962231.

[24] A. Matakos, C. Aslay, E. Galbrun, and A. Gionis, "Maximizing the Diversity of Exposure in a Social Network," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 9, pp. 4357–4370, 2022, doi: 10.1109/TKDE.2020.3038711.

[25] M. Mirtaheri, S. Abu-El-Haija, F. Morstatter, G. Ver Steeg, and A. Galstyan, "Identifying and Analyzing Cryptocurrency Manipulations in Social Media," *IEEE Trans. Comput. Soc. Syst.*, vol. 8, no. 3, pp. 607–617, 2021, doi: 10.1109/TCSS.2021.3059286.

[26] D. Vimalajeewa, S. Balasubramaniam, B. O'Brien, C. Kulatunga, and D. P. Berry, "Leveraging Social Network Analysis for Characterizing Cohesion of Human-Managed Animals," *IEEE Trans. Comput. Soc. Syst.*, vol. 6, no. 2, pp. 323–337, 2019, doi: 10.1109/TCSS.2019.2902456.

[27] A. A. Al-Shargabi and A. Selmi, "Social Network Analysis and Visualization of Arabic

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v7i2.18894

Tweets During the COVID-19 Pandemic," *IEEE Access*, vol. 9, pp. 90616–90630, 2021, doi: 10.1109/access.2021.3091537.

[28] M. Bérubé, T. U. Tang, F. Fortin, S. Ozalp, M. L. Williams, and P. Burnap, "Social media forensics applied to assessment of post–critical incident social reaction: The case of the 2017 Manchester Arena terrorist attack," *Forensic Sci. Int.*, vol. 313, 2020, doi: 10.1016/j.forsciint.2020.110364.

[29] A. Umrani, Y. Javed, and M. Iftikhar, "Network Forensic Analysis of TwitterApplication on Android OS," *Proc. - 2022 Int. Conf. Front. Inf. Technol. FIT 2022*, pp. 249–254, 2022, doi: 10.1109/FIT57066.2022.00053.