INTERNATIONAL E-JOURNAL OF CRIMINAL SCIENCES
Supported by DMS International Research Centre

Can metadata be used to measure the anonymity of *Twitter* users? Results of a Confirmatory Factor Analysis

Zoraida Esteve, Asier Moneva and Fernando Miró-Llinares [1]
Research Centre for the Study and Prevention of Crime
Miguel Hernandez University of Elche

Abstract

Anonymity is one of the elements traditionally associated with criminal and antisocial behaviour. Anonymity depends on several factors, such as natural surveillance or the visibility created by the physical or digital environment. Certain digital environments, such as social networks, exhibit characteristics that facilitate or limit the degree of anonymity of their users. Social networks are places in cyberspace where people interact with each other and with the environment, where they increasingly carry out their daily activities and where they also commit crimes. This paper attempts to test the hypothesis that certain elements of the social network environment define the anonymity of their users. To this end, an empirical process for quantifying anonymity is proposed, which can be applied transversally to all places in cyberspace that permit user accounts. Subsequently, a data set of 162 users has been obtained from the social network Twitter which also collects the metadata associated to their accounts. To test this hypothesis, a Confirmatory Factor Analysis (CFA) has been conducted to determine whether the data obtained fit the model based on a theoretical concept proposed by the researchers. The results show a moderate fit for the model, suggesting that some metadata (i.e., geopositioning) do not contribute to defining the latent variable anonymity. We suggest the proposed model needs to be reconsidered and applied to a larger sample to improve its fit. Finally, the applicability of the proposed methodology for measuring anonymity and future lines of research are discussed.

Keywords: Anonymity, cyberspace, metadata, Twitter, CFA.

[1] Avda. de la Universidad, s/n. Edificio Hélike, 03201 Elche, Alicante (Spain) z.esteve@crimina.es

INTERNATIONAL E-JOURNAL OF CRIMINAL SCIENCES
Supported by OMS International Research Centre

1. Introducción

Information and Communication Technologies (ICT) have revolutionized the way we relate to each other. First the growth of the Internet and later the growth of ICT as a whole, have been fundamental in creating an environment which is distinct from physical space and which allows content sharing and instant interaction with other members of the virtual community. Thus, cyberspace, understood as a virtual space which resembles a geographic space without being the same (Miró-Llinares & Johnson, 2018), has altered the meaning of distance and time, which are now compressed to the point of disappearance (Miró-Llinares, 2011, 2012). Along with these intrinsic characteristics, cyberspace has been configured by other extrinsic characteristics such as its transnationality, neutrality, decentralization, universality and anonymity, attributes that have driven its popularization, but that have also imposed obstacles to the prevention and prosecution of the crimes that occur within it (Miró-Llinares, 2011). Thus, for example, the absence of barriers, which in the physical space represent the borders that configure the different states, makes it enormously difficult for the justice system to act outside its boundaries, lost in a labyrinth of infinite and complex legal rules and procedures. The neutral and non-centralized nature of the Internet reduces or practically eliminates restrictions on accessing websites or disseminating information, making it difficult to control the flow of content and, therefore, also the behaviour of its users. Finally, anonymity, the driving force behind the popularisation of the Internet, inhibits social controls (Armstrong & Forde, 2003; Finn, 2004; McGrath & Casey, 2002) and provides criminals with an attractive environment (McGrath & Casey, 2002) by increasing the sense of impunity for crime and dissociating their anonymous online actions from their behaviour in physical space (Suler, 2004).

In this sense, scientific literature has extensively described the role of anonymity in the genesis of criminal behaviour, both in physical space (Haney, Banks, & Zimbardo, 1973; Lelkes, Krosnick, Marx, Judd, & Park, 2012; Rogers & Ketchen, 1979), and in cyberspace (Baggili & Rogers, 2009; Hinduja, 2008; Ševčíková & Šmahel, 2009). Thus, research such as that of Armstrong and Forde (2003) on paedophilia or that of Finn (2004) on harassment have shown that online environments can favour a false sense of intimacy, as well as uninhibited behaviour that increases risk-taking and antisocial behaviour. It seems, therefore, that it is definitely the virtual space which, given its specific configuration favouring anonymous behaviour, provides the appropriate conditions for criminal events to occur. Thus, by focusing on this new place we have called cyberspace, we find ourselves required to ask whether it makes sense to study crime from a theoretical perspective, under the same postulates as we do in physical space. Hence, when we examine the consistency of the so-called crime theories applied to this "no place" (Miró-Llinares & Johnson, 2018) we find that, if the minimum elements present in the criminal event are the same, that is, a potential offender, a suitable target, and the absence of a

capable guardian, whose convergence occurs in a (digital) place and at a specific moment (Cohen & Felson, 1979; Miró Llinares, 2011), it seems logical to think that the fundamental premises of crime theories can be adapted to help explain the commission of crimes in cyberspace.

In any case, we know that people's daily activities are moving progressively into cyberspace and that, as in physical space, criminal opportunities are not randomly distributed in cyberspace, but are concentrated in certain places where the risk of a crime occurring is greater (Miró-Llinares & Johnson, 2018). In this way, users who buy on certain pages and do not check their security before making payment become appropriate targets for cyber-scammers who are aware of such vulnerability (Pratt, Holtfreter, & Reisig, 2010). Similarly, users who use email to exchange messages and files with others are more likely to receive spam (Yeargain, Settoon, & McKay, 2004) or suffer a malware infection (Hoar, 2005). It is true that in all these crimes and in others such as hate speech, the convergence between the aggressor and the victim takes place in a different way from traditional offences. However, convergence is still necessary, as it is essential that there is a place where the hate message is expressed and where another user receives it (Miró-Llinares, Moneva, & Esteve, 2018). Thus, just as with certain crimes in physical space, such as theft, which tend to be concentrated in the places where businesses are located (Wikström, 1995), in cyberspace both crime and the perception of insecurity are also distributed according to the characteristics of cyberplaces (Castro-Toledo & Miró-Llinares, 2018; Miró-Llinares & Johnson, 2018; Miró-Llinares et al, 2018) and the interaction between users, such as forums, chats and, mainly, social networks.

The popularization and universality of social networks means that virtual convergence among users is sometimes more frequent than in physical space, which in turn increases criminal opportunities. For example, on the social network Twitter, many users publish real personal information such as multimedia content, location, routines, etc. that put them in a position of vulnerability with regards to certain cybercriminals. On the contrary, other users interact anonymously, hiding their identity through a pseudonym or false names (Peddinti, Ross, & Cappos, 2014), which allows them to express opinions and publish sensitive information without fear of being identified (Peddinti, Ross, & Cappos, 2017b). Although not all Twitter users hide behind a veil of anonymity to commit criminal or deviant behaviour, the work of Peddinti and colleagues (2017b; 2017a) shows that there is a relationship between the anonymity of users and the pornographic, homophobic and Islamophobic content published from their accounts. Both the anonymity provided by this social network and the ease with which it is accessed, which allows users with different personal characteristics to identify with certain ideologies (Perry & Olsson, 2009), have favoured Twitter becoming a platform where some users emit radical and hateful messages that remain fixed over time, thereby reaching a massive audience and thus increasing their harmfulness (Miró-Llinares et al., 2018). However, scientific research

has not yet shown what the specific influence of anonymity is on the commission of criminal behaviour on Twitter, mainly due to the methodological difficulties involved in quantifying this condition.

2. A step-by-step proposal to measure the anonymity of online users

There is some consensus on the existence of a relationship between anonymity on the Internet and criminal behaviour, despite the fact that few empirical studies have delved into its influence at the individual level and despite the fact that some who have done so tangentially have not extracted conclusive results (Bautista-Ortuño, 2017). Perhaps one of the reasons for this derives from the fallacious understanding of cybercrime as a single event when, in reality, there are multiple criminal modalities of very different nature and in which anonymity can play a very different role (Miró-Llinares, 2015). In addition, cyberspace is not univocal either, so it seems unrealistic to try to measure all the factors that configure "anonymity on the Internet", since they vary according to the configuration of each digital environment. In this sense, the few exceptions that have attempted to define and measure anonymity in cyberspace rightly try to circumscribe it to certain places. Peddinti et al. have done so on Twitter via an approach which is more similar to analysis of deviance than crime (Peddinti, Korolova, Bursztein, & Sampemane, 2014; Peddinti, Ross, & Cappos, 2017b, 2017a, 2014) and others have tried to replicate their methodology in isolation (Xue, Yang, Ross, & Qian, 2017). According to these investigations, anonymous Twitter users are more uninhibited, interact more, follow more accounts and are more willing to display their activity to the general public. The scale of anonymity developed by Peddinti and colleagues is, however, debatable as a method for classifying Twitter users according to their anonymity, especially with a view to analysing their relevance in relation to criminal or antisocial activity. The authors divide Twitter users into four categories according to their degree of anonymity:

- *Anonymous*: The user has not provided a first and last name or a URL in their profile.
- *Partially anonymous*: The account contains a first name or last name, but not both in their profile.
- *Identifiable*: The user has indicated their name and surname in their profile.
- *Unclassifiable*: The user does not indicate name or surname, but they do have a URL in their profile (p.85).

Although this is an interesting first approximation, identifying anonymity exclusively with the wording of a name or surname, or with the presence of a URL in the profile, does not seem to be an adequate strategy to measure the degree of anonymity with which a user acts at a specific time. It is even less so if we bear in mind that Twitter is characterized

by intense visualization of contents that are published from very heterogeneous profiles. However, given the different structural and communicative nature of the different environments that make up the Internet, we do think it is opportune to confine this analysis to a specific place in cyberspace.

This paper proposes a methodology to measure the degree of anonymity of online users by following a sequential and systematic five-phase process (Figure 1): (1) select a specific cyber place; (2) identify the relevant metadata from a user's profile; (3) collect additional relevant information associated with the profile; (4) operationalize the identified anonymity variables; and (5) conduct a Confirmatory Factor Analysis (CFA) to model the anonymity of users.

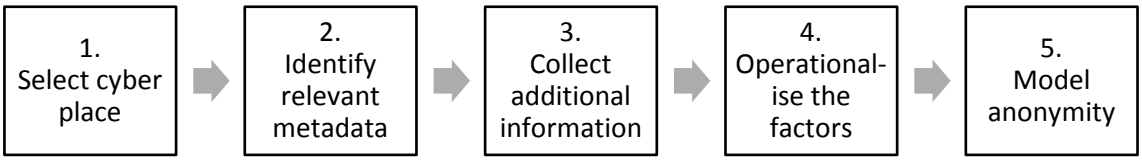

Figure 1. Systematic five-step sequential process to measure online users' anonymity

2.1. Selection of a specific cyber place

Since each cyber place has a different configuration that conditions the degree of exposure of its users, the first step to measure exposure is to select a specific digital environment that will be subject to the analysis. According to Miró-Llinares and Johnson (2018), cyber places can be classified according to (1) the type of contact they allow, (2) the natural surveillance they enable and the self-protection measures they make available to their users, and (3) the type of activity carried out there. All cyber places are configured to a certain extent by the way in which these three elements are combined. Therefore, it is essential to pay attention to these characteristics and how they relate to the potential anonymity of the users who visit them.

Regarding the former, most cyber places have store-and-forward information transmission channels, although some of them, such as the social network Twitter, also include streaming contact systems such as Periscope. The fundamental implication in relation to anonymity is that while the first modality implies greater control for the user over the information that they wish to transmit and the way to do it, since they can dedicate an indeterminate time to plan its publication, the second requires greater capacity for improvisation and, therefore, risks are assumed. Although the essential form of communication on Twitter is the tweet (i.e., any message published on Twitter that can

contain photos, videos, links and text), it is possible to communicate privately through direct messages. The second characteristic that defines cyber places is their natural surveillance, defined by the volume of information traffic generated by their users and the degree of publicity of the content they include. In this sense, Twitter is a social network with a significant daily influx of users and which offers users a series of self-protection tools that allow filtering of certain messages or blocking annoying content. In addition, the configuration of social networks in general, and Twitter in particular, allows the user's privacy to be adjusted. With respect to the third element, the configuration of cyber places determines to a great extent the routines of its users and, therefore, the type of activity carried out there. While some users use Twitter for professional purposes - providing their real identity, establishing a network of professional contacts, and paying special attention to the content they publish from their profile (for example, academic) - others do not need to provide personal information, as they use the social network for leisure.

In addition, cyber places can be analysed at the macro or micro level. Once you have selected the macro-place where you want to measure the anonymity of your users, the next step is to identify the micro-place that contains the relevant information related to them. In the macro cyber place "Twitter" this information can be found in the micro cyber place "user account" (Miró-Llinares et al., 2018).

2.2. Identification of the relevant metadata associated with the user profile

In addition to the personal information collected, these accounts generate additional information linked to each user that reflects their interaction and, in turn, defines their public exposure. For example, a Twitter user's account contains information related to both visibility and anonymity (Miró-Llinares et al., 2018). Regarding the second factor, each Twitter account contains at least the following information: (1) if it is a verified account, (2) the user's name, (3) if it adds a short biography, (4) if it is a geopositioned account, (5) an external link, (6) if it includes a location, (7) if it has a profile photo, (8) if it has a cover photo, and (9) if it has other photos. The first of these elements has definite importance for determining the degree of anonymity of a user, since if Twitter grants a user the blue badge, it means they have passed a rigorous process of identity verification. Given that the rest of the information can be falsified to a greater or lesser extent, a subsequent categorization is necessary to reflect this and to establish a range depending on the greater or lesser anonymity provided by each element.

2.3. Collect information from the profile

Although some of the anonymity-related factors are only accessible through the Application Programming Interface (API) that Twitter only makes available to users with developer permissions, others are publicly accessible and can therefore be collected

manually. In relation to the API, access to this information is increasingly restricted, as recent problems related to the filtering of users' personal information from the social network Facebook has affected privacy and has called into question the ability of these companies to ensure the protection of user data. Nevertheless, provided that the applicant meets a series of requirements and, in some cases, passes an evaluation regarding the justification it has provided for access to such information, a significant number of APIs can be accessed that in turn facilitate access to information stored on social network servers such as Twitter, Facebook, or YouTube, among others. In the case of manual collection processes, it is necessary to adopt a strategy for the systematic observation of the cyber places described in phase 1 that gather the information identified in phase 2. However, when accessing each of the user profiles to record the appropriate information, difficulties may arise related to the privacy settings that each user has set up for their profile, or the account may have been sanctioned by Twitter and closed accordingly. In any case, the identification of the relevant metadata in order to capture the degree of exposure of an online user is not only a characteristic of social networks but, rather, it is transversal to all those cyber places configured to store user profiles.

2.4. Operationalization of anonymity factors

Once the relevant variables have been selected, the next step is to proceed with their operationalization. The approach to operationalising these variables depends on the analysis technique that is used afterwards, so there is no single criterion to carry out this process. However, it is preferable to opt for a quantitative categorization whenever possible, as these data are more flexible to treatment and allow the variables to be re-operationalised in a qualitative format if necessary. On the other hand, a qualitative operationalisation of the variables will not allow the inverse step to be carried out with such a high level of detail. However, it is true that it is sometimes complex to quantify certain characteristics and a qualitative operation must be used, in which case it is essential that the resulting variable is ordinal. The values taken by the variables must describe a range that goes through an anonymity-exposure scale in order to be able to interpret the results of the analyses in one sense or another; that is, it is necessary to know whether a category corresponds to a greater or lesser degree of anonymity in order to determine the direction of the resulting relationship. For example, it would be possible to quantify the number of photos that a user has posted from their account or it would be possible to discretise this information to know if the user has ever posted a photo or not. While from the first method it is always possible to obtain the coding for the second, knowing whether photos have ever been published from a profile does not allow us to know to what extent this is so. Other metadata such as name or location do not allow quantification, so they should be used as a nominal dichotomous scale. In this case, the

absence of the attribute is considered an increase in the anonymity of the user and its presence the opposite.

2.5. Modelling anonymity

The analytical strategy adopted to measure anonymity depends on how the data have been recorded and categorized. In this pilot study, an approach to the anonymity model is proposed through the application of a Confirmatory Factor Analysis (CFA). CFA is a type of factorial analysis that is used to model a latent variable whose value is unknown with a set of manifest variables whose value is known. A fundamental requirement for the application of CFA is that the relationships established between the manifest and hypothetical variables within the proposed model are guided by a solid theoretical approach (Schreiber, Nora, Stage, Barlow, & King, 2006). In this case, the variables based on metadata identified in phase 2 have been included in a model to adjust the latent variable that reflects the degree of anonymity-exposure of each user and that has been defined theoretically from the previous variables. In order to execute this statistical technique, the free software R is used, loading the functionalities offered by the 'lavaan' package (Rosseel, 2012).

3. The pilot study

The purpose of this study is to illustrate an application of the proposal to measure the anonymity of online users by executing a CFA on a sample of Twitter users. The aim is to determine whether these data fit the anonymity model previously hypothesized by the researchers.

3.1. Sample

After acquiring developer permissions on the Twitter platform, it is possible to access the Streaming API to make a formal request for information about the online activity of users of this social network. After the terrorist attack on London Bridge in June 2017, a request was made to the Twitter server to obtain a sample of the tweets issued by users in order to study the prevalence of violent communication and hate speech in this environment of digital interaction. In order to delimit the query, a filtering criterion by language and keywords was established. Regarding the first criterion, only those messages published in English were requested. Regarding the second criterion, three keywords were defined based on the hashtags that at the time of the request were global trending topics on Twitter: #LondonBridge, which sought to identify those messages referring to the terrorist event from a neutral position; #PrayForLondon, which sought to filter messages of solidarity or support; and #StopIslam, which sought to collect

expressions of negative or discriminatory content. This query returned a sample of 200,882 records in the unstructured JSON format containing information associated with Twitter users and the tweets they had published after the attack (Miró-Llinares et al., 2018). After a process of classifying the messages with a criterion of inter-rater agreement trained in the Taxonomy of hate and violent communication on the Internet (Miró-Llinares, 2016), an additional dichotomous attribute was assigned to each record indicating whether the message could be included within the categories defined in the taxonomy or not.

To carry out the proposed pilot study, 200 users were selected who had published a message in the sample described; 100 of whom had published at least one hate message. Subsequently, each of the selected profiles was accessed to check whether the user account is still active. Given the impossibility of collecting some essential data to conduct the CFA model proposed here (i.e., suspended or closed accounts), 38 of these users were excluded from the sample. The final number of users included in the sample for the systematic observation of their profile is 162. Since this study is going to apply a CFA on a single sample and taking into account the consensus established in the specialized literature that suggests the incorporation of 10 records for each of the estimated parameters is acceptable, the sample size is adequate (for example, Schreiber et al., 2006).

3.2. Manifest variables

After excluding the variable that determines whether a user account is verified, the 8 remaining variables identified in phase 2 were selected and the data collected following a process of systematic observation of the 162 identified profiles. Subsequently, the variables were quantitatively categorised by describing an anonymity-exposure range, as indicated in phase 3. Table 1 below summarizes the characteristics of the manifest variables included in the model.

INTERNATIONAL E-JOURNAL OF CRIMINAL SCIENCES
Supported by OMS International Research Centre

Table 1. Operationalisation of the manifest variables for the CFA

Variable	Categorisation
Name	0: the user name is fictitious; 1: the user name can be real
Biography	0: the user does not have a biography; 1: the user's biography does not provide relevant information about the user's identity; 2: the user's biography provides information that may be relevant to identify the user.
Geopositioning	0: the user has not activated the geopositioning of their tweets; 1: the user has activated the geopositioning of their tweets.
External URL	0: the user does not include a URL in their profile; 1: the website to which the profile URL redirects does not provide relevant information about their identity; 2: the website to which the profile URL redirects may be relevant to identify them.
Location	0: the user does not include a location in their profile; 1: the user includes a location in their profile that is fictitious; 2: the user includes a location in their profile that may be real
Profile photo	0: user does not include a profile picture; 1: user includes a profile picture that does not show a person; 2: user includes a profile picture of a person
Cover photo	0: the user does not include a cover photo; 1: the user includes a cover photo that does not show a person; 2: the user includes a cover photo that shows a person.
Other photos	0: the user has not posted messages with images; 1: the user has posted messages with images that do not show people; 2: the user has posted messages with images that show a person.

For the CFA model, we conducted 1 regression for each of the manifest variables, giving a total of 8. Taking into account the nature of the variables included in the model, the use of a robust weighted least squares estimator (WLSMV; Rosseel, 2012) has been selected.

4. Results

Since our model assumes the existence of relationships between the variables that define it, the standardized covariance matrix of the variables has been extracted to observe the differences between the expected and observed correlations in the model (Table 2). Values > 0.1 are indicators of relationships that can be conflicting when adjusting the model and that will be reflected in the error measurement. In general, these residual

correlations show acceptable values, with some exceptions (for example, name, geopositioning, location).

Table 2. Covariance matrix of variables for CFA

Manifest variable	1	2	3	4	5	6	7	8
1. Name								
2. Biography	-0.12							
3. Geopositioning	0.12	-0.13						
4. External URL	-0.01	0.10	-0.03					
5. Location	0.07	0.07	0.16	0.14				
6. Profile photo	0.16	0.01	0.09	-0.09	-0.09			
7. Cover photo	-0.08	-0.03	-0.07	-0.10	-0.17	0.03		
8. Other photos	-0.12	-0.09	-0.08	0.00	-0.06	-0.05	0.09	

Note. The variables have been standardized in accordance with the following parameters: M = 0.00; SD = 1.00 (N = 162).

Below are the main indicators of the fit of the specified model based on the main indices that, according to Schreiber and colleagues (2006; see also Kline, 2011), are fundamental in evaluating unique analyses such as the one presented in this study: (1) the Tucker-Lewis index (TLI), which should be ≥ 0.96 for categorical data for a good fit; (2) the comparative fit index (CFI), which should be ≥ 0.95; and (3) the root mean square error of approximation (RMSEA), which should be < 0.60. While the first two serve to compare the model, the latter is a measure of error based on the residual correlations reflected in Table X. In our model the indices described take the following values: TLI = 0.98; CFI = 0.99; RMSEA = 0.05. In its robust version, the values are: TLI = 0.98; CFI = 0.99; RMSEA = 0.05. In its robust version, the values are: TLI = 0.98; CFI = 0.99; RMSEA = 0.05: TLI robust = 0.95; CFI robust = 0.96; RMSEA robust = 0.07. These results show a moderate degree of fit between the model and the observed data, since, although normal indices give good results, their robustness does not exceed the cut criteria in two of the three cases.

As expected, all variables included in the model have factor loads that are significantly related to the latent anonymity construct. Most standardized coefficients (i.e., Betas) are above the threshold of 0.60, although it is true that they range from 0.26 regarding geopositioning to 0.74 with regards to other photos (Table 3). Figure 2 shows the proposed CFA model for the anonymity latent variable with standardized parameters for each manifest variable. Although the coefficients associated with the geopositioning variable suggest that discarding it could improve the fit of the model, no post hoc

modifications have been made, since the hypothesized model derives from a sufficiently justified theoretical approach.

Table 1. *Regression coefficients for the variables in the CFA*

Manifest Variable	B	SE	Z	Beta	sig.
Name	1.00	0.09	4.95	0.43	***
Biography	1.58	0.07	10.05	0.68	***
Geopositioning	0.60	0.11	2.44	0.26	*
External URL	1.58	0.06	11.41	0.68	***
Location	1.53	0.07	8.99	0.66	***
Profile photo	1.42	0.06	9.63	0.62	***
Cover photo	1.60	0.07	9.96	0.69	***
Other photos	1.72	0.06	12.88	0.74	***

Note. * p value < 0.05; ** p value < 0.01; *** p value <0.001

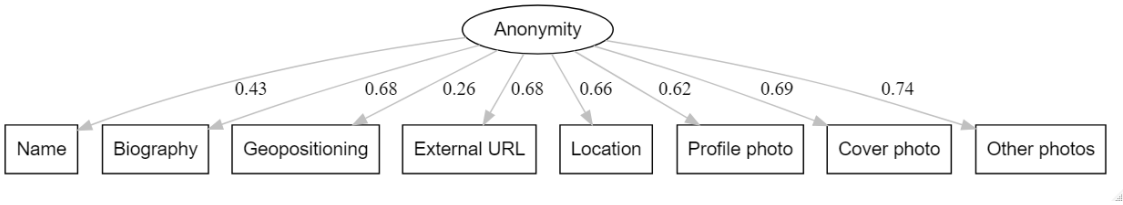

Figure 1. CFA Model for latent variable anonymity. Robust Chi-squared = 0.015. Degrees of freedom = 20.

5. Discussion and conclusions

In the present study a CFA has been performed to test the hypothesis that the metadata of Twitter users' accounts (i.e., manifest variables) define their anonymity (i.e., latent variable) measured as a theoretical construct. Since the proposed model is based on both a conceptualization exercise and a review of the literature on the operationalization of anonymity on Twitter (Miró-Llinares et al., 2018; Peddinti, Korolova, et al., 2014; Peddinti, Ross, et al., 2014; Peddinti et al., 2017a, 2017b), a CFA has been chosen rather than an Exploratory Factor Analysis (EFA). Unlike the CFA, the EFA pursues the creation of new latent variables starting from groups of factors whose relationship is unknown from a theoretical point of view. It can be affirmed that while the CFA is guided

by theory (Schreiber et al., 2006), the EFA is guided by pragmatics (Tabachnick, Fidell, & Ullman, 2019), and while at the time of executing a CFA the variables of the model are previously defined, the EFA serves precisely to define the number of variables that define a construct (Williams, Onsman, & Brown, 2010). For the same reason, even when the results do not show a good fit for the data, when a CFA is performed it is not necessary to include or eliminate factors to improve the fit of the model. In the present case, the CFA results show that the proposed manifest variables, modelled as a single linear combination of factors, show a moderate fit; that is, the proposed model can be said to have limited capacity to explain the latent variable anonymity. In summary, the data partially support the proposed hypothesis and show that it is possible to develop more rigorous anonymity measurement models than those currently in use.

Beyond the results, another contribution of this paper is the proposal to quantify anonymity. In contrast to other attempts to measure user anonymity on Twitter (for example, Peddinti et al., 2017b), the proposed process is not only more exhaustive, as it makes use of metadata associated with users' online accounts, which allows for the incorporation of a wide variety of elements that characterise such a digital environment, but also transcends the social network Twitter, as its sequential design is devised for application in any digital environment configured to record and store user accounts. Thus, for example, this proposal can be extrapolated to email user accounts, forums, or web applications. The versatility of this methodology makes it possible to explore the influence of anonymity, as well as other constructs, on criminal behaviour and deviant behaviour in cyberspace.

Regarding the method of data collection, researchers have used the open data policy of Twitter to make a request for information that has allowed access to certain information that otherwise would not be accessible and that has subsequently been supplemented with additional information collected manually. It should be noted that other social networks do not allow access to their data for research purposes or restrict it altogether. The same obstacle may arise in the case of other cyber places that also involve user accounts. Therefore, in order to follow the process of quantification of anonymity proposed in this paper, it will be necessary to rely on manual data collection methods that will only be valid to the extent that they are also systematic and rigorous. These methods pose a number of problems: (1) some factors that hypothetically allow latent variables to be defined from theoretical approaches will not be accessible; (2) the costly sample collection process will greatly limit the ability to obtain large amounts of information that would allow more robust models to be developed; and (3) such a process is difficult for other researchers to replicate, limiting the ability to contrast results or extrapolate them to other similar contexts for comparative analysis.

Future research on the use of CFA to measure the anonymity of online users should work on improving the process of obtaining and coding the variables included in the

model to improve their fit. To this end, it is necessary to identify new variables related to the anonymity of the accounts and to propose new ways of operationalizing the existing ones in order to achieve greater completeness and precision in their modelling. Secondly, it is necessary to carry out studies with a larger sample to give greater robustness to the results obtained, so it is advisable to use the APIs as the main way of obtaining data. Finally, it would be interesting to see to what extent the new construct of anonymity obtained after modelling acts as a predictor of criminal or deviant behaviour in cyberspace (for example, hate speech, spam, fraud), as its relationship with cybercrime is often taken for granted, but its empirical study in cyberspace has been neglected.

References

Armstrong, H. L., & Forde, P. J. (2003). Internet anonymity practices in computer crime. *Information Management & Computer Security*, *11*(5), 209–215. https://doi.org/10.1108/09685220310500117

Bautista-Ortuño, R. (2017). ¿Eres un ciberhater? Predictores de la comunicación violenta y el discurso del odio en Internet. *International E-Journal of Criminal Sciences*, *11*, 1–28.

Castro-Toledo, F. J., & Miró-Llinares, F. (2018). ¿Nos parecen más inseguros los ciberlugares después de un ciberataque? *International E-Journal of Criminal Sciences*, *12*, 1–25.

Finn, J. (2004). A Survey of Online Harassment at a University Campus. *Journal of Interpersonal Violence*, *19*(4), 468–483. https://doi.org/10.1177/0886260503262083

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed). New York: Guilford Press.

McGrath, M. G., & Casey, E. (2002). Forensic psychiatry and the internet: Practical perspectives on sexual predators and obsessional harassers in cyberspace. *The Journal of the American Academy of Psychiatry and the Law*, *30*(1), 81–94.

Miró-Llinares, F. (2015). Cibercrimen y vida diaria en el mundo 2.0. In F. Miró-Llinares, J. R. Agustina-Sanllehí, J. E. Medina-Sarmiento, & L. Summers (Eds.), *Crimen, oportunidad y vida diaria* (Dykinson, pp. 415–456). Madrid.

Miró-Llinares, F. (2016). Taxonomía de la comunicación violenta y el discurso del odio en Internet. *IDP: Revista de Internet, Derecho y Política*, *22*, 82–107.

Miró-Llinares, F., & Johnson, S. D. (2018). Cybercrime and Place: Applying Environmental Criminology to Crimes in Cyberspace. In G. J. N. Bruinsma & S. D. Johnson (Eds.), *The Oxford Handbook of Environmental Criminology* (pp. 883–906). https://doi.org/10.1093/oxfordhb/9780190279707.013.39

Miró-Llinares, F., Moneva, A., & Esteve, M. (2018). Hate is in the air! But where? Introducing an algorithm to detect hate speech in digital microenvironments. *Crime Science*, *7*(15), 1–12. https://doi.org/10.1186/s40163-018-0089-1

Peddinti, S. T., Korolova, A., Bursztein, E., & Sampemane, G. (2014). Cloak and Swagger: Understanding Data Sensitivity through the Lens of User Anonymity. *2014 IEEE Symposium on Security and Privacy*, 493–508. https://doi.org/10.1109/SP.2014.38

Peddinti, S. T., Ross, K. W., & Cappos, J. (2014). 'On the internet, nobody knows you're a dog': A twitter case study of anonymity in social networks. *Proceedings of the Second Edition of the ACM Conference on Online Social Networks - COSN '14*, 83–94. https://doi.org/10.1145/2660460.2660467

Peddinti, S. T., Ross, K. W., & Cappos, J. (2017a). Mining Anonymity: Identifying Sensitive Accounts on Twitter. *ArXiv:1702.00164 [Cs]*. Retrieved from http://arxiv.org/abs/1702.00164

Peddinti, S. T., Ross, K. W., & Cappos, J. (2017b). User Anonymity on Twitter. *IEEE Security & Privacy*, *15*(3), 84–87. https://doi.org/10.1109/MSP.2017.74

Rosseel, Y. (2012). Lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2). https://doi.org/10.18637/jss.v048.i02

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *The Journal of Educational Research*, *99*(6), 323–338. https://doi.org/10.3200/JOER.99.6.323-338

Suler, J. (2004). The Online Disinhibition Effect. *CyberPsychology & Behavior*, *7*(3), 321–326. https://doi.org/10.1089/1094931041291295

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2019). *Using multivariate statistics* (Seventh edition). Boston: Pearson.

Williams, B., Onsman, A., & Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine*, *8*(3). https://doi.org/10.33151/ajp.8.3.93

Xue, M., Yang, L., Ross, K. W., & Qian, H. (2017). Characterizing user behaviors in location-based find-and-flirt services: Anonymity and demographics: A WeChat Case Study. *Peer-to-Peer Networking and Applications*, *10*(2), 357–367. https://doi.org/10.1007/s12083-016-0444-5

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 740773.

This research has been funded by the Ministry of Science, Innovation and Universities under FPU Grant Reference FPU16/01671.

This research has been funded by the Spanish National Cybersecurity Institute (INCIBE) under "*Ayudas para la excelencia de los equipos de investigación avanzada en ciberseguridad*" Grant Reference INCIBEI-2015-27349.