A REVIEW OF THE INTERSECTION TECHNIQUES ON HUMINT AND OSINT

Alcides Macêdo, Laerte Peotta and Flávio Gomes

Department of Electric Engineering, University of Brasília, Brasília/DF

ABSTRACT

The traditional and doctrinal concepts of cybersecurity involve the concepts of physical, logical and social perspectives, prevailing the consensus that the system will be secure if these three perspectives have appropriate levels of compliance. However, the difficulties in conceptualizing “security” for human and social interfaces have compelled industry and academic research to treat cybersecurity with a focus on the physical and logical perspectives. Within the scope of the social perspective, the threats are shaped in the interactions between information and technology what shifts the central point of the discussions to the informational society with the interdependencies of human relations based on the use of social networks that result in the growing volume of information supply. This research aims to broaden the underlying debate about malicious user behaviour by proposing a framework that allows classifying this type of user according to predefined profiles. The objective of this research is to develop a framework based on OSINT (Open Source Intelligence) and HUMINT (Human Intelligence) concepts that can help cybersecurity professionals to select potential collaborators with objective criteria that make it possible to measure the reliability of the information obtained. Besides, the research aims to evaluate how the approach based on HUMINT (Human Intelligence) techniques and applied in a virtual environment can enhance cybersecurity.

KEYWORDS

Open Source Intelligence, Human Intelligence, Human Sources, Information Gathering, Social Engineering

1. INTRODUCTION

The traditional and doctrinal concepts of cybersecurity involve, in a simplified way, the physical, logical and social perspectives, prevailing the consensus that the system will be safe if these three perspectives have appropriate levels of compliance. Faced with the difficulties in conceptualizing “security” for human and social interfaces, industry and academic research have compelled the treatment of cybersecurity to a focus on the physical and logical perspectives [1]. However, from the mid-2000s onwards, the field of research on online crime expanded from theoretical discussions about the nature of cybercrime in a virtual environment to approaches related to physical spaces where tangible damage to property and physical well-being and people's emotions are made. In addition to the expansion of research, the scope of academic investigation has also expanded from isolated approaches in the early 2000s to large-scale population surveys, including population samples of young people and new criminal and non-criminal methods that have been used to understand the nature of cybercrime [2] [3] [4] [5] [6].

According to the most researched approaches, at the level of the physical and logical perspectives, effective security depends on the implementation of a cybernetic strategy based on
understanding the hacker mindset, penetration tests by the red team, broad and deep defense, and blue team actions [7].

Within the scope of the social perspective, the threats are shaped in the interactions between information and technology, that shift the central point of the discussions to the informational society, with the interdependencies of human relations based on the use of social networks that result in the growing volume of information supply.

From the perspective of the physical and logical perspectives, there are several frameworks for proactive cyber defense, with only a few approaching the social layer tangentially, through treatment centered on the user's “security culture”.

By affecting all fields of human activity, the massive supply of information also poses challenges to cybersecurity, particularly to the collection of quality information. Faced with the possibility of generating information by the user himself through social networks, the interaction between information and technology results in new possibilities for collecting and processing information on the web. This works as a subsidy for the implementation of information security policies, from corporate environments in the private sector to law enforcement or the national security.

The collection of information as part of the public security activity fundamentally covers the collection and processing of information to subsidize the managerial process in the administration of public security bodies or in the conduction of actions to repress crime. This happens, mainly, in the cybernetic space where series of services are offered also and in their environment transit confidential information which are subject to threats from cybercrimes in the first level, electronic industrial espionage in the second level and cyber war in the third level [8].

This research intends to broaden the underlying debate about malicious user behaviour by proposing a framework that allows the classification of this type of user according to predefined profiles, so that security teams can assess the possible approach for a collaboration proposal to search for information, in order to facilitate the handling of the volume of information created in social networks that permeates all fields of human activity.

Considering the potential of data collection in the huge amount of material available on the web through forums and message boards, review and opinion sites, social bookmarks, media sharing, blogs, microblogs and social networks, as well as the limitations regarding their implementation, the following research question is presented: How can the approach based on HUMINT (Human Intelligence) techniques combined with OSINT (Open Source Intelligence) techniques be applied in a virtual environment to enhance cybersecurity?

The objective of this research is to develop a framework based on concepts that might help cybersecurity professionals to select potential collaborators with objective criteria that make it possible to measure the reliability of the information obtained.

The effectiveness of using human sources to achieve the objectives of law enforcement agencies and national security was demonstrated in the UK Government Office report, that accounted for more than 30 terrorist threats, in addition to 3500 arrests and 500 weapons during the year 2018 as a result of information collected from human sources [9]. Despite these results, obtaining information from the use and management of informants is still a poorly researched area [10] [11], and even the ability to obtain information through interaction with human beings composes the subculture of police activity [12].
This research is justified by the fact that public security is one of the main Brazilian problems nowadays, and although the analysis of information collected in web environments to support government actions is a recurring theme in the area, there are no studies in Brazil on the performance of police agents in collecting information from the human relationships that happen in the internet. In this sense, the empirical research proposal will be able to discuss the practices of the police subculture and continue in search of new meanings for the collection of information through human interactions in the internet.

The empirical diagnosis as a basis for theoretical discussion on the information collection through human interactions, as well as on its legal, ethical and methodological implications, can offer answers related to the use of new tools, mainly those mediated by technology, in order to improve the effectiveness of public security actions through the collection of information, discussion forums, websites or social networks [13], so that online user profiles can be created in a framework intended for cybersecurity professionals.

With the advent of the internet, traditional surveillance, dependent on human capabilities to analyze and capture data without the intermediation of technology, lost importance in intelligence activity as a result of the end of the Cold War and of the advances in technology. This disseminated in bureaucratic and operational structures the belief that the disciplines grouped in the field of techint could account for most of the information collection, underestimating the collection of information from human sources – discipline of HUMINT (Human Intelligence) [14] – collection of information from human relations that happen in the internet. In this sense, the empirical research proposal will be able to discuss the practices of the police subculture and continue in search of new meanings for the collection of information through human interactions in the internet.

1.2. Problem Description

According to Steele [15], the theoretical and organizational framework of intelligence activity can be represented by Figure 01, which positions open-source information as the basis for knowledge production, according to Allen Dulles' understanding of the preponderance of information collection from "normal, open and transparent methods" [16] [17], which, after being processed, form the basis for the collection disciplines (HUMINT (Human Intelligence), SIGINT (Signal Intelligence), IMINT (Image Intelligence), MASINT (Mass Intelligence)). Thus, the processed open information (OSINT) is the basis for the analysis of information and is founded on the collection in public sources.

![Figure 1. Framework of Intelligence Activities [15]](image)
With the increasing use of the internet and the information dynamics formation that occur in the virtual environment based on the user’s behaviour, the OSINT approach can be combined with HUMINT techniques to work beyond the mere provider of subsidies for the other collection disciplines, consolidating a new aspect of cybernetic intelligence destined to select eventual collaborators.

Faced with the myriad of open sources of information and their importance to support criminal investigation activities or the production of strategic information, this research seeks evidence on the applicability of the combined approach of OSINT and HUMINT techniques to select potential collaborators from OSINT and HUMINT techniques as a subsidy for actions that might increase cybersecurity measures.

Considering the contributions of data collection on the huge amount of material made available in the web through forums and message boards, review and opinion sites, social bookmarks, media sharing, blogs, microblogs and social networks, as well as the limitations regarding its implementation, the following research question is presented: Is it possible to combine OSINT and HUMINT techniques to select potential collaborators in an online environment?

Thus, the scope of this research is limited to the selection of the best practices in order to collect information for the creation of online user profiles in a framework intended for cybersecurity professionals’ use.

The main contribution of this study is to broaden discussions on (1) the applicability of OSINT and HUMINT techniques, together, with cybersecurity professionals, what may lead to the proposition of a framework of actions aimed at the collecting of information so that user profiles can be created, as well as: (2) gather greater knowledge about intelligence activity in the context of the web, highlighting the peculiarities of information collection from open sources, discipline of OSINT (Open Source Intelligence), and recruitment of human sources, discipline of HUMINT (Human Intelligence); (03) identify good practices for combining the collection of information from open sources, an OSINT (Open Source Intelligence) discipline, with the management of human sources, a HUMINT (Human Intelligence) discipline and (4) propose a framework based on good practices for gathering information from open sources.

2. LITERAL REVIEW AND BACKGROUND WORK

The collection and processing of information consists of steps that are grouped as “structured information flows” in the single-sources collection and all-sources analysis dimensions, which relate, respectively, the operational means and the analytical process used in collecting and processing information [15].

The work of police intelligence, as a rule, begins with the collection of information that can be obtained in an easier and safer way and advances to more difficult and risky methods, in the single-sources collection dimension. Thus, actions begin with the collection of information available from open sources and social media (OSINT and SOCMINT, respectively, Open-Source Intelligence and Social Media Intelligence) and evolve to more complex and expensive methodologies, passing through HUMINT (Human Intelligence) to the network of seismic sensors that detect atomic weapons tests deep underground. The next step takes place in the all-sources analysis dimension, in which the information collected is gathered and interpreted through analytical processes that differentiate the product of the intelligence activity from the information collected by its explanatory and/or predictive capacity.

However, watertight methods structured as single-sources or all-sources need to be resized to accommodate changes in interactions between information and technology, in order to develop
theoretical and applied solutions that increase the accessibility of the growing supply of information that results from the diversity of these interactions and that may be related to malicious activity.

2.1. Human Intelligence (HUMINT) in an Online Environment

The framework that results from this research proposes the expansion of the collect activity beyond the capture of information in a virtual environment, typical of OSINT and SOCMINT. However, the fusion of HUMINT, OSINT and SOCMINT in a virtual environment depends on actions inserted in contexts of restricted groups of individuals with dynamics of limited sharing, differently from the content that reaches broader and more diverse audiences [14].

Many of these individuals interact in the network "camouflaged" by pirated and/or fake accounts that aim to harm the image of other participants, through "trolling" or disturbing the digital environment with offensive posts, instigating confrontations between participants, in a way that preference is given to participants in which they may have repercussions, in addition to obtaining and/or deleting information from participants for the commission of cybercrimes and also crimes in the “physical world”, as pointed out by Glenny [8].

In their research on terrorism, Briggs and Strugnell [17] maintain that the use of the internet by extremist groups made it possible to break down existing barriers in the physical world for certain groups of people, particularly for the participation of women in jihadist movements, as the physical relationships between men and women can be seen in the Arab world as unacceptable outside the family circle, making it difficult for women to participate in these groups.

The same authors also claim that anonymity and freedom of action can make the expression of certain thoughts in public acceptable, with the advantage of reaching many users interested in the topic [17]. For Glenny [8] “the internet is a theory of the big bubble – we solve a problem that affects it, but another, apparently intractable, surfaces elsewhere”. In this way, the internet generates a huge amount of data and information that has little or no value, another amount with little or no interpretation and a small portion is dangerous for its falsehood [14].

For Gioe [18], technological innovation that increases the importance of collecting data from OSINT and SOCMINT (Social Media Intelligence) will not diminish the importance of the HUMINT discipline, but will give it a new dimension through new capabilities and new challenges: (1) the massive supply of data will facilitate the effective evaluation of the collection operations based on human sources; (2) the dissemination of the use of biometric functions can help in the identification of people; (3) the widespread use of social media can increase the vulnerability of collection operations under SOCMINT. The same author argues that the solutions are not easy and do not involve only the intelligence community to prospect more innovative and creative solutions allowed by the same technologies.

In this virtual scenario characterized by an abundant supply of data, Stottlemyre [19] suggests that the main change in the activity of collecting information stems from the imbrication between OSINT and HUMINT through the so-called crowdsourcing intelligence, a new specialized discipline of collection based on human interactions in social media.

However, the collection of information in a virtual environment would depend, in addition to OSINT and SOCMINT skills, on the acquisition process of the human source, which is structured in four interdependent phases, that begin with (1) identification of the potential source, (2) collaboration development, (3) recruitment and (4) dismissal [19] [20]. These phases configure a process that begins with actions to confirm whether the eventual human source has access to the
requested information, actions that are related to the knowledge of biographical and professional data, as well as motivations and behavioural characteristics for the development of strategies to establish links and mutual trust, in order to convince the human source to provide information, echoing virtual relationships in the real world.

When selecting a unit as a research object with the proposal to analyze it in depth [21], the choice of the study methodology to develop a proposed framework based on human intelligence (HUMINT) and open source intelligence (OSINT) to information collection in a virtual environment has an exploratory, descriptive and qualitative character in relation to the target group composed of cybersecurity professionals who perform the collection of information in internet environments.

According to references in the consulted literature, the basic premises of the study are related to the process of acquiring information through a human source in a virtual environment, involving the collaborator as an information supplier and the controller in the position of the cybersecurity professional who is in charge of establishing the management of interpersonal relationships related to:

a) Controller personality: the literature consulted lists the existence of personality profiles characterized by confidence, extroversion, emotional stability and good judgment of people [1] [22] [23] [24] [25], as well as the ability to confront informants [26];

b) Rapport: ability to establish interpersonal relationships based on empathy and respect, ability to listen and understand [27] [28] [29] [30];

c) Informant's motivation: the precise identification of the elements that drive the informant is pointed out by several authors as fundamental for the collaboration relationship [31] [32]. Studies conducted in the US and UK have identified a range of motivations, including leniency in the criminal justice system, financial reward, revenge or removal of criminal competitors, and even moral or interpersonal motivations [33] [34];

d) Establishing cooperation: this aspect emphasizes that an informant is not an interviewee with access to information, but an active participant in the collection and processing of information and must have the centrality of anticipation [35] [36] [37] [38] [39] [40];

e) Obtaining information: based on information elicitation techniques, including interview methods aimed at providing information and the way it can be improved [41] [30] [42];

f) Detection of false information: involves deliberate action to detect and curb the transmission of false information [43];

g) Social engineering: among hackers and cybersecurity professionals, the practice of accessing and controlling confidential information is called “social engineering” [44] [45], and it involves a wide spectrum of devices to compromise security systems of information [46]. From the perspective of perpetrators of online crimes, social engineering is casually and fluidly applied by people with “attributes” related to the social context, to the frailties of human nature, to the complexities of social networks, to the role of social conventions and the to limitations of human processing and reasoning. These attributes are organized into four categories corresponding to different stages of social engineering action [48] [49] [50]. The recent study by Kevin Steinmetz et al [51] demonstrates that many “social engineers” apply their scams after rigorous planning that considers all circumstances that may influence the potential victim's susceptibility, presentation and chances of success. Thus, planning involves the following steps:
1) Research: consists of gathering as much information as possible in order to help the “social engineer” set up the coup based on pretexts, false identity or setting up a scenario;

2) Assessment: assessment of the resulting skills and previous experiences or technical training related to the victim;

3) Timing: consists of evaluating the duration of the blow to build “pretexts” strong enough to withstand the duration of the blow;

4) Proximity: establishing proximity with the victim consists of “building a relationship” to give the impression of being an integrated participant in the victim's network or a kind person. To achieve this objective, the following techniques can be used:
   • Rapport: approach strategy that avoids an incisive, aggressive or intimidating approach, with preference for subtle and friendly presentations.
   • Networking integration: adoption of measures to simulate participation in the victim's social environment.

5) Activation: stage after the establishment of proximity, when the perpetrator makes maneuvers to obtain the expected attitude, related to the provision of critical information or favorable attitudes to the “social engineer”. The main techniques are asking for help or offering incentives;

6) Concealment: measures adopted by the perpetrator to avoid the emergence of suspicions if the activation stage results in failure, thus avoiding detection maintaining the integrity of the ruse for later reuse.

The social engineering techniques summarized above can be used by cybersecurity professionals in some situations. However, given the importance of managing human sources for cybersecurity actions, it is necessary to discuss and implement techniques to manage these sources effectively and professionally.

Operations for collecting information derived from human sources have their center of gravity in the “controlling agents” who are responsible for collecting information and for the human sources themselves [47]. From the perspective of human sources, in general, the sensitivity pyramid proposed by Herman can be used to designate the number of sources and the informational value in the area of HUMINT [11].

3. METHODOLOGY

The references found indicate the prevalence of research interest in the development of tools to explore the potential of data supply in internet environments. The existence of more studies in this area would support the elaboration of strategies capable of integrating two or more collection disciplines in order to increase the production of subsidies for cybersecurity actions.

From the preliminary phase, it was observed that data collection for the production of informational subsidies destined to cybersecurity actions occurs through direct research in online search engines, social media sites and restricted access databases. The quality of information, which has a great impact on cybersecurity in collecting information, with little group integration during the all-sources collection phase.

In order to understand the informational dynamics that occur in the domain of information for cybersecurity, the case study is the most suitable for the development of understanding little-known organizational phenomena and for the development of new approaches without the requirement of representativeness or measurement of statistical frequencies, in addition to providing the analytical generalization of the researched phenomena [21].
It is, therefore, an empirical task to investigate a contemporary phenomenon inserted in its real context without defined limits between the phenomenon and the context [21]. Nevertheless, it is necessary to adopt a research strategy that has as its object the detailed analysis of a phenomenon [48], which relate the case study methodology with the need to “illuminate a decision or a set of decisions” through the study of the motivation decision-making process, the way it is implemented and what are its results [49].

The research of a phenomenon in its natural environment through the observation and the application of structured interviews constitutes the case study as a qualitative methodology [50], focusing on the subjective nature of the studied object to understand the dynamics of group-target, in order to provide descriptions, test theories or generate theories and models [51] [52]. When selecting a unit as a research object with the proposal to analyze it in depth [53], the choice of study methodology to develop research on HUMINT (Human Intelligence) techniques combined with OSINT (Open Source Intelligence) techniques applied in a virtual environment to intensify cybersecurity, has an exploratory, descriptive and qualitative character in relation to the target group composed of cybersecurity professionals who perform the collection of information in internet environments.

The basic assumptions of the case study are:

a) Elaboration of a study in the field of information security related to the use of a human intelligence approach (HUMINT) combined with open source techniques (OSINT) for data collection in a virtual environment;

b) Improvement in information production procedures to subsidize the selection of collaborators.

After carrying out a bibliographical research, the collection activity was explored through semi-structured interviews to obtain contributions in order to understand the informational dynamics of human relationships that involve the procedures for the selection of collaborators. Depending on the exploratory nature of the research, a structured questionnaire was made to explore the basic topics with the possibility for the respondent to make comments in a very open manner, in order to deepen the theme according to the respondent's perspective. 26 questions related to the selected good practices were presented, addressing the following topics:

   a) Definition of OSINT;
   b) Definition of HUMINT;
   c) Opportunities arising from the use of OSINT;
   d) Opportunities arising from the use of HUMINT;
   e) Difficulties resulting from the use of OSINT;
   f) Difficulties resulting from the use of HUMINT.

3.1. Evaluations of Results

The good practices selected in the theoretical framework and initiated in the case study, as well as the verified factual context, allowed the confirmation of the good practices that follow presented in the final form of each practice within the previously identified categories. Questionnaires for validation of the framework were submitted to 40 specialists at the end of 2021 and the beginning of 2022 who working in data collection in a virtual environment. The questionnaire was designed with statements to be answered using a Likert scale with six alternatives with equivalence between 1 and 6, with 1 corresponding to “completely disagree” and 6 “completely agree”.

In the end, the expert assessment confirmed good practice. However, some showed a high degree of disagreement. Based on the participants' responses, the final form of each good practice is
presented, emphasizing that the threshold of each good practice is not very clear, including that some may permeate the context of another.

4. CONCLUSIONS

The advances that began in the Bletchley Park laboratories and reached the sophisticated surveillance networks structured in satellite equipment, passing through the ambitious ECHELON system to the TEMPORA and PRISM Projects, resulted in the proposition of the term Dataveillance, which consists of the “systematic use of data for investigating or monitoring the actions or communications of one or more people” [54], as opposed to the traditional surveillance approach, which relies on human capabilities not mediated by technology to analyze and capture data.

The possibilities of systematic data monitoring resulting from the growth of social media represent a promising proposal to improve the effectiveness of public safety actions through the configuration of the discipline or Social Media Intelligence (SOCMINT). In addition to providing horizon scanning and strategic alerting capabilities, SOCMINT can make decisive contributions to prevent cybercrime by identifying potential collaborators or perpetrators of cyberthreats.

In this sense, the proposition of a framework aimed at good practices for recruiting collaborators in a virtual environment can increase the study and development of measures to enlarge cybernetic defenses aimed at the social perspective and recognize human behaviour as a critical infrastructure that can be exploited and must also be defended.

REFERENCES


