Cyber Personalities as a Target Audience

Miika Sartonen¹, Petteri Simola², Jussi Timonen¹ and Lauri Lovén³,
¹Finnish National Defence University, Finland
²Finnish Defence Research Agency, Finland
³Center for Ubiquitous Computing, University of Oulu, Finland

Abstract: Target audience analysis (TAA) is an essential part of any psychological operation. In order to convey a change in behaviour, the overall population is systematically segmented into target audiences (TAs) according to their expected responsiveness to different types of influence and messages, as well as their expected ability to behave in a desired way. The cyber domain poses a challenge to traditional TAA methods. Firstly, it is vast and complex, requiring effective algorithms to filter out relevant information within a meaningful timeframe. Secondly, it is constantly changing, representing a meshwork in formation, rather than a stable collection of TAA-specific data. The third challenge is that the target audience (TA) consists not of people, but of digital representations of people, whose true identity and characteristics cannot usually be verified. To address these challenges, the authors of this article suggest that the concept of TAA has to be revised for use in the cyber domain. Instead of trying to analyse physical people through the cyber interface, the authors have conceptualized an abstract entity whose physical identity might not be known, but whose behavioural patterns can be observed in the cyber environment. These cyber personalities, some of which are more or less intelligent algorithms, construct and share their interpretation of reality as well as carefully planned narratives in the digital environment. From the viewpoint of TAA, the only relevant quality of these entities is their potential ability to contribute to the objectives of a psychological operation. As a first step, this article examines the cyber domain through a five-layer structure and looks at what TAA-relevant data is available for analysis. The authors also present ways of analysing cyber personalities and their networks, in order to conduct a TA that effectively supports psychological influence in the cyber domain. As a way of better utilizing the digital nature of cyber personalities, a concept of dynamic TAs is also introduced.

Keywords: psychological operations, PSYOPS, target audience analysis, TAA, cyber personality, dynamic target audience

1. Introduction

The concept of war is being transformed by the introduction of information war in the digital environment. The cyber domain, and the internet in particular, allows global audiences to be reached in battles of influence. Facts, often presented as more of a question of a viewpoint than an objective entity, are hard to find among the stories of aggressors, victims, justified attacks and illegal defence. Von Clausewitz’s famous statement, of war being “nothing but a continuation of politics with the admixture of other means”, finds new relevance with opportunities provided by the digital revolution. If military objectives can be achieved without use of force or without accountability, the threshold for clandestine military psychological operations is lowered.

From a military viewpoint, the internet is an ample theatre of operation for anyone skilled enough and willing to use it. The information that people and organizations give away in their daily business, both directly and indirectly, can be very valuable to anyone with the means of obtaining it. It is now possible to gather mission-relevant data almost in real time, often with the willing help of the TAs themselves. On the other hand, both individuals and organizations obtain much of their information from the internet, meaning that anyone with means of either directly or indirectly influencing that information has a global reach of operations. Thus, the fast and reciprocal nature of the internet provides real time opportunities for both data mining and message dissemination in support of psychological operations.

This article presents the idea of cyber personalities as the TA of psychological operations and seeks to answer the question of “what characteristics of cyber personalities can be analysed?” Following the introduction, section two looks at cyber personalities from the viewpoint of psychological operations, and how digital cyber personalities can be found in a five layer structure framework as presented by Sartonen, Huhtinen and Lehto (2016). Section three presents an idea of dynamic TAs for cyber personalities, followed by a conclusion and discussion in sections four and five.
The aim of military psychological operations is to change the behaviour of TAs (or in some cases to maintain it despite hostile influence attempts). A psychological operation can be described in three major steps. In step one, the objectives of the operation are defined and the TAs (including the means and the messages used to influence their behaviour) are selected. In step two, the influence messages are produced, approved and delivered to the TA. During step three the effectiveness of the ongoing operation is assessed and, if necessary, the influence attempts are altered. Thus, for a successful operation to take place, the TA must be reached in a way that allows conducting the TAA, disseminating the influence messages, and evaluating their effectiveness (FM 3-05.30, 6-1 – 6-4).

TAA is a detailed and systematic process of selecting the most viable TAs that can be reached and influenced, and whose behavioural change will effectively produce results that support the overall military operation (FM3-05-301, 5-1). In order for a set of messages to be effective, a TA typically needs to consist of a homogenous group of people with similar conditions and vulnerabilities. There are various ways of building these homogenous groups, such as looking for common demographic or geographic features. Centres of gravity (people or small groups that have large degree of power over others) or key communicators are also desirable, although typically small TAs (FM3-05-301, 5-3 – 5-4).

Once initial TAs have been selected, a more detailed scrutiny is applied, beginning with the conditions that the TA has to cope with. Conditions, according to FM3-05-301, are the life-affecting elements over which the TA has little or no control. In an ideal TA, common conditions within the group lead to similar needs. These unfulfilled or perceived needs are manifested as vulnerabilities, as the TA’s desire to fulfil, alleviate or eliminate the needs acts as motivation to change behaviour. A successful TAA provides the most effective means of satisfying the TA’s needs in a way that allows the PSYOPS objectives to be reached (FM3-05-301, 5-4 – 5-9).

From the point of view of TAA, the cyber domain is a challenging theatre of operations, as TAs in the digital environment consist of digital representations of people, whose true identity often cannot be verified. Although the internet provides many ways of identifying persons behind their aliases, the resources available may still leave the TA with a large number of unidentified portions. This may lead to ineffective results, as the unidentified persons, whose characteristics cannot be analysed, may distort the qualities of TAA, and lead to less effective influence attempts. On the other hand, if the unidentified persons are left outside of the TA pool and their number is too high, the overall effectiveness of the influence operation will suffer and the operational targets may not be achieved.

The question, then, is whether or not it is possible to conduct psychological operations solely in the cyber domain, without physically identifying all the members of a TA. In order to succeed, the above-mentioned three main steps would need to be achieved entirely within the cyber domain. Within the scope of this article, the authors concentrate on the first step, i.e. defining TAs and the most effective messages of influence.

As a possible solution, the authors suggest a concept of a cyber personality. A cyber personality is a collection of all the interconnected information in the five layers of the cyber domain, as presented by Sartonen, Huhtinen and Lehto (2016), which creates a single abstract entity. Digital pictures, audio and video files, social media posts, location data history, topographic analysis of both physical and virtual networks, and finally observed behaviour in the digital environment are all examples of characteristics that can be analysed. In case of humans, this cyber personality is a digital representation of a human being. Humans are not, however, the only inhabitants of the cyber domain. The concept of cyber personality allows for different types of influential algorithms to be acknowledged as operational factors in line with human beings. In the following sections, the five layers of cyber domain are viewed from the perspective of the information that they provide for the identification of a cyber personality.

As shown above, a cyber personality interacts with the layers of the cyber domain in various ways. Figure 1 illustrates the complex relations between the different aspects of a cyber personality and the cyber domain layers at large.
Figure 1: Cyber personality aspects in relation to the five layers of the cyber domain

2.1 Physical layer

The physical layer includes all the physical devices and networks of the cyber domain (Sartonen, Huhtinen and Lehto, 2016). For example, mobile devices with location services double as surveillance assets, and combined with the information provided by digital maps, they can provide constant information about the location of a physical identity of a cyber personality. The same applies to network topologies and switching diagrams together with the network traffic generated by the virtual identity of a cyber personality. With a long enough surveillance time, and provided that the user keeps the location services on, this information reveals geographic information, behavioural patterns, and even lifestyle choices. Where a cyber personality lives and works, how many hours, which stores are visited daily or weekly, what recreational locations the cyber personality visits etc., can all be established with information from the physical layer.

Conversely, when provided a physical trajectory of a cyber persona as well as the corresponding service or semantic level activities, it is possible to analyse whether there are any discrepancies, omissions or conflicts between the levels. These could point to possible fabrication or a fraudulent cyber personality belonging to an artificial construct. In order for the cyber personality to be convincing, the location information needs to correlate with the personality’s activities.

2.2 Syntactic layer

The syntactic layer consists of the software that operates the devices of the physical layer (Sartonen, Huhtinen and Lehto, 2016). The corresponding cyber personality aspect is a virtual identity: a local user account on a computer or device. In other words, once a cyber personality starts using a new device (computer, mobile phone), a virtual identity has been created in the syntactic layer.

A single virtual identity can provide access to multiple network identities, such as e-mail addresses or cloud-based user IDs, and can thus be the means of connecting multiple network identities to a single cyber personality. Linking a physical device, such as a computer on a campus or in a working place, to a virtual identity also provides demographic information about the physical identity of a cyber personality. The browser used by
the cyber personality is also a good source of information. It can leave traces of past browsing and other information (such as user agent and operating system) (Wang et al., 2016).

Again, conversely, supposing we have established a possible connection between the physical as well as the virtual identities of a cyber personality, we can assess the likelihood of the connection being real by comparing the information on both levels. Is the network usage pattern as expected, and does it correspond with the physical trajectory? If there are discrepancies, it is possible that the cyber personality is fraudulent, such as an automated social media bot that is not utilizing a browser and is only focusing on application programming interface (API) (Chu et al., 2012). Discrepancies can also occur if a cyber personality uses different techniques, such as encryption (Gupta et al., 2014) and TOR network (Haraty and Zantout, 2014) to avoid detection.

2.3 Semantic layer

The semantic layer consists of data and information provided by the cyber personalities (Sartonen, Huhtinen and Lehto, 2016). This information includes images, text and audio files that people use to communicate and share their views of the world. The semantic layer may in many cases be the most useful in terms of TAA, as in everyday life, personality is manifested in our behaviour and interaction with others (Mehl, Gosling, & Pennebaker, 2006). Along with Facebook, Twitter updates provide a great deal of information about personality. In their study in which they were able to identify several aspects of personality, Qiu et al. noted in 2012 that “you are what you tweet”. They also showed that some of the content that they analysed was more specifically related to gender and age than personality (Qiu, Ramsay, and Yang, 2012). Personality related information can be extracted from both text-based and online social media information (Tskhay and Rule, 2014). For example, extroverted people use positive words more often, whereas the use of positive words is negatively correlated with neuroticism (Pennebaker and King, 1999; Yarkoni, 2010). In a similar manner, criminals behind digital personas can be tracked using vast amounts of text data sets from social media (Rashid et al., 2013). Insider threats can be recognised from social media feeds by using a natural language processing system and combined with risk assessment (Symonenko et al., 2004).

2.4 Service layer

The service layer consists of public and commercial digital services, with various social media as the main focal point of psychological operations (Sartonen, Huhtinen and Lehto, 2016). In the past few years, an increasing amount of studies have focused on understanding how our actions in social media reflect a personality in the physical world. There is evidence showing that personality traits relate to the content of status updates and choice of profile pictures on Facebook (Winter et al., 2014; Wu, Chang, and Yuan, 2014), showing that people, adding a particular post or picture, also reveal a part of their personality. Ikeda et al. (2013) developed a method for demographic estimation of Twitter users. Demographics such as age, gender, area, hobby, occupation and marital status were estimated by tracking the tweet history and clustering of followers/followees (Ikeda et al., 2013). It is also possible to perform sentiment analysis on Twitter feeds. Along with this analysis, devices and other entities can be extracted from the data set (Saif et al., 2012).

Furthermore, whole social networks may be analysed to find and target fraudulent, artificial (such as social botnets) or “troll” networks. A natural social network follows the shape of a certain type of random graph. Divergences from this shape indicate non-naturalness: the greater the divergence, the higher the probability of fraud (Zhang et al., 2013). It has to be noted that, concerning social media, the semantic and service layers are often intertwined. What someone says, how it is said and to whom, what links are used, the overall behavioural pattern of the writer, etc. fall under different categories of semantic, network, demographic and other analysis types. The main difference, however, is the digital tools used for analysing different aspects of a cyber personality.

2.5 Cognitive layer

The cognitive layer is the ultimate target of psychological operations in the cyber domain. This layer consists of rational and emotional human processes that direct the information flow through all the layers of the cyber domain (Sartonen, Huhtinen and Lehto, 2016). Successful influence attempts affect this layer and produce a change in behaviour that can be observed through the behavioural residue in the other layers. Although we interact with digital representations of personalities on the internet, behind a social media account there is typically still a person (although the number of bots and other algorithms is increasing). This person exists in the
Miika Sartonen et al.

physical world and interacts with other entities both in the physical world and the cyber domain — he or she is not just bits in cyberspace. This person has a unique personality, motivations and desires that direct his or her behaviour and interaction with others. One may be overly extrovert or may want to keep to oneself (McCrae & John; 1992); similarly, one may be motivated by gaining power and dominance above anything else, or maintaining positive social relationships at the expense of one’s own needs (McClelland 1988). To understand a personality we must understand both the environment that he or she lives in, as well as the personality and motivations that mediate behaviour in that environment (Funder, 2006, 2009).

Once the characteristics of personality have been identified, this information can be used to include basic elements of influence psychology in the operation (Cialdini, 1984). For example, if we identify individuals as extroverts we can use influence methods that exploit the principles of liking, reciprocation, social proof and scarcity (Alkış & Temizel, 2015; Uebelacker & Quiel, 2014). Similarly agreeable individuals, in other words individuals who are amicable, compassionate and trust others (McCrae & John, 1992) are the most vulnerable to influencing attempts such as phishing (Parrish 2009). They are also the most vulnerable to influence methods using the principles of authority, reciprocity and liking (Alkış & Temizel, 2015). On the other hand, individuals who are identified as highly conscientious tend to follow safety and security regulations (Parrish et. al 2009; Darwish et. al., 2012) and as such, are less likely to react to influence attempts that try to make them break these regulations. However, it has been suggested that conscientious individuals may be vulnerable to influence methods using principles of reciprocity, authority, commitment and consistency (Alkış & Temizel, 2015; Uebelacker & Quiel, 2014).

Cultural aspects, identified through demographic and geographic data in addition to behavioural information, can also indicate successful means of influence. As an example, Orji (2016) demonstrated that collectivists and individualists vary with respect to their responsiveness to influence. For individuals living in a collective culture authority, reciprocity, consensus and liking-based influence seems to be somewhat more effective than on individuals living in individualist cultures (Orji, 2016) For TAA, the data available on each above-mentioned level must be identified. The interactions in each layer leave traces that can be used for data analysis. Furthermore, there is ample context information available in each layer, which is illustrated in further detail in Figure 2.

Figure 2: Context and interaction data available on the five levels of the cyber domain
3. Cyber personalities in dynamic target audiences

In addition to changing how digital TAs are comprehended, successful psychological operations in the cyber domain may need more than just the initial composition of the TA to be adjusted. The rhizomatic nature of this environment means that the information flows and the integrity of data cannot be controlled. A message, once sent, may or may not end up being read by the intended TA in the form it was written. In addition, other messages may suddenly gain wide audiences and approval, possibly colliding with existing themes and creating something entirely new. This leads to sudden, abrupt changes in how matters are viewed and thus may lead to friendly parts of TA becoming hostile and vice versa (Sartonen, Huhtinen and Lehto, 2016).

Reacting (preferably faster than adversaries) to these changes in global digital moods requires changing the messages and their recipients. In other words changing the TAs. This cannot, however, be done without a feedback mechanism, i.e. without knowing how to change messages and which members to include or exclude in a TA. For this purpose alone, the authors of this article argue that the input – feedback loop has to be both constant and very fast. Constancy requires supervision, which in turn means that authority for making decisions about the composition of a TA and thus the targets of a psychological operation has to be very low hierarchically.

Combining fast or even real time feedback from the digital personalities’ observed behaviour with the possibilities provided by the sheer amount of information mined from extensive data requires re-thinking the concept of a TA. Traditionally, as in FM3-05-301, the TA is basically an instrumental tool, an object of certain handcrafted influence messages and remaining the same once approved by a proper authority. The cyber domain, however, allows for temporal changes in the composition of TAs.

The main benefit of creating and moderating dynamic TAs, i.e. audiences whose composition would be allowed to change constantly, would be to reach susceptible audiences more effectively. Such TAs would be labelled by varying features, which would not need to be pre-defined. Such labels could include the type of conditions or vulnerabilities that a group shares, or behavioural patterns (such as observed willingness to support the PSYOPS campaign’s objectives without apparent external reason), or of demographic similarities, etc. A single cyber personality could belong to multiple TAs, provided that the requirements of belonging to a group apply.

Typically, any psychological operation has to begin with pre-defined TAs as the base of operations. There are, however, two main ways of applying dynamic TAs to these initial groups. One would be to allow new dynamic TAs to be created outside the initial groups, and to let it be populated by new cyber personalities in addition to those already inside an initial TA. Another would be allowing also the initial TAs to become dynamic in nature (and possibly non-existent in the long run), but this would mean greater risk of losing the scope of the operation.

The above-mentioned concept means that many such TAs would be temporal in nature, utilized as long as the requirement of supporting the PSYOPS objectives is met. The concept would also require the influence channels linked to the TA to be flexible, allowing for messages to be sent by the means that are currently the most effective. The usability and effectiveness of different channels would be a constantly changing variable, updated either through the same feedback loop as the TA, or by external analysis.

4. Conclusion

Many digital tools are capable of providing very specific information on different subjects of interest. It is, however, the meaningful combination of the information from all the five layers of the cyber domain that makes it possible to conduct effective TAA. As noted above, this combination allows gathering data about digital TAs’ geographic information, demographics, personality, behaviour and conditions, all of which are requirements for TAA. Thus it can be concluded that at least in theory, it is possible to perform TAA exclusively in the cyber domain.

Attaching cyber personalities to dynamic TAs, as suggested in this article, would allow the reality of the rhizomatic cyber domain to be accepted as an inseparable part of operations. In order words, all the messages in the cyber domain that shape the information environment would be accepted as an environmental factor, rather than a possible disruption of ongoing operations. The concept also includes the mind-set of influencing, instead of controlling, the information environment. This would mean that instead of seeking controlled and evaluated cause – effect changes, a more holistic approach would be applied, with only the end result in mind. For a western military thinker, such an approach may sound haphazard in its uncontrollability, but the authors of this article argue that the cyber environment has a level of chaos built in as a feature.
5. Discussion

This article joins the debate on the internet as a platform of influence operations. The authors argue that although psychology, which forms the backbone of psychological operations, advances in a relatively careful manner, changes in the digital environment are rapid and revolutionary. Stepping from the industrial into the information age may need an entirely new approach instead of trying to adjust current processes to work in a new environment.

One critical limitation of the operations described in this article, is that the usability of TAs consisting entirely of cyber personalities depends on their representation of the overall target group in the real world. In other words, the critical question is whether the cyber TAAs have the capability of conveying the desired behavioral change in the overall population. This probability depends on many factors, such as the percentage of internet users in the target population and cultural differences in the use and importance of the digital environment. These factors may be difficult to estimate reliably in advance. Further studies are thus needed to assess the real world usefulness of the theoretical approach presented here.

To verify the theoretical approach presented in this article the authors suggest three methods. The first would simply test the ability of this approach to deliver the desired behavioural changes in a target audience. However, as with any other psychological operation’s evaluation, the problem with this approach is that the causalities behind observed behavioural changes are very difficult to prove.

The second approach would be to compare the results from two influence operations, performed on similar audiences but with different TAA-methods. The first method would be the more traditional model of seeking out human identities through the cyber interface, while the other would suffice with the cyber personalities only. Comparing the results and the overall practicality (such as time consumed) of these methods would show whether or not human identification is necessary for satisfactory results. To conduct several tests with various audiences with different percentages of internet access would yield additional information on the applicability of different methods. In order to have reliable results, the properties of the target audiences would have to be known, which may present the greatest challenges of this approach. To make the test in laboratory conditions may lack the ‘real world’ complexity of target audiences, while tests with real audiences seldom have clear cause and effect correlations.

The third method would be to access data from an already conducted target audience analysis, and perform the TAA with cyber personality approach on the same target audience. The challenge with this approach is that the conditions of the target audience may have changed and thus the results may not be comparable.

Disseminating and analysing a cyber personality through the five layers in the way presented in this article may also include the benefit of recognizing bots and other algorithms more easily. One suggestion for further studies would be to look at what features artificial cyber personalities lack in comparison to humans in the five cyber layers and whether these differences are platform-specific. The lack of human qualities would thus enable the identification of artificial entities.

References


