IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 3, 2020, accepted December 28, 2020, date of publication December 30, 2020,
date of current version January 15, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.3048269

Semantically Modeling Cyber Influence Campaigns (CICs): Ontology Model and Case Studies

NATHAN JOHNSON[1], (Graduate Student Member, IEEE), BENJAMIN TURNBULL[2], (Member, IEEE), THOMAS MAHER[2], AND MARTIN REISSLEIN[1], (Fellow, IEEE)

[1]School of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ 85287-5706, USA
[2]School of Engineering and Information Technology, University of New South Wales, Campbell, ACT 2601, Australia

Corresponding author: Martin Reisslein (reisslein@asu.edu)

ABSTRACT This paper presents a novel ontological model of Cyber Influence Campaigns (CICs). The model accepts both physical and cyber based actions. The model represents the mechanics, linkages, and structure of tailored data to concurrently analyse both the physical and cyber realms. Influence modeling and ontological based analysis of social media has to date mainly focused on the use of ontologies to categorise or cluster the results of text based feature extraction. Whilst this is highly important for detection of misinformation that has the potential to influence a network, these methods do not provide a mechanism for mapping events across realms in order to quantify the influence in meaningful ways. By developing a novel semantic model and unique classes that leverage the graph nature of the ontological representation, our ontological model provides causal linkages and a framework which is applicable for analysis and deeper insights into CICs. This study also builds two tailored datasets for our ontological model from raw Twitter data as the IEEE DataPort Cyber Influence Campaign Ontology dataset (DOI 10.21227/70kc-yx38) and details how to analyze various CIC scenarios.

INDEX TERMS Cyber influence, conflict, influence campaign, ontology, semantic model, social media.

I. INTRODUCTION

A. MOTIVATION

Daily life is now a hybrid of social media, social networking, digital communications, as well as physical communications and interactions. Facebook, Twitter, and Instagram boast over two billion monthly active users [1], and as such, their ability to directly and indirectly connect the world's population has never been easier or more far reaching. Online content delivery, and the algorithms that govern it, have changed both the method and the speed at which our world communicates, consumes, decides, and progresses. There is now sufficient data that has been voluntarily shared on social media by users, that their beliefs and behavioural responses can be predicted and manipulated [2]. An important open question remains: are the underlying principles and mechanics of influence between the physical and digital (cyber) realms fully defined and understood? In order to progress the understanding of influence flow between the physical and cyber realms, there is a requirement to not only contextualise but attribute actions to reactions across both realms. Building a model that can combine both physical events with cyber events helps understand the mechanics of the influence exchange and becomes the basis of tools that take a holistic approach to influence analysis.

Social Networking Services (SNS) remain largely unregulated by government [3], and have been weaponized to become *Cyber Influence* [4]. Social influence is defined by psychologists as "the change in one's beliefs, behaviour or attitudes due to external pressures that may be real or imagined" [5]. Cyber influence, is social influence via digital means, which research shows is commensurate and in some cases even more powerful than physical influence [6]. Cyber influence has already been employed to mobilise oppressed populations, win elections, fight wars, and undermine drug cartels [4].

The associate editor coordinating the review of this manuscript and approving it for publication was Xiao Liu.

B. CONTRIBUTIONS

We develop the concept of a Cyber Influence Campaign (CIC) and define a CIC as ''actions employed by an agent in order to change the attitudes of their desired audience''. We generate a novel semantic CIC model that spans both the cyber and physical realms and explores the mechanics and underlying principles of cyber influence.

This paper makes the following key contributions:

- Definition of the classes of objects and relationships within a CIC, encapsulated into a semantic model and data structure.
- Development of a novel ontology capable of capturing both the physical and cyber elements of a CIC.
- Two physical datasets derived from real life events and the corresponding cyber datasets extracted from raw Twitter data made publicly available as the IEEE DataPort Cyber Influence Campaign Ontology dataset (DOI 10.21227/70kc-yx38).
- The dataset development code and search queries used in the research analysis.
- Two case studies with exemplary analysis to confirm real life (physical realm) observations.

This paper is structured as follows. Section II outlines the background of cyber influence and explores the related fields of study that build the fundamentals of cyber influence. In Section II, existing studies and their limitations are outlined, defining the scope for our research. Section III provides an overview of a number of real CICs examples and their role in conflict. Section IV reviews the concept of semantic modeling as well as the technology and recent examples of ontologies. Section V introduces the novel ontology model that we have developed to represent the influence of both physical and cyber events. Section V covers the novel CIC model, the data pipeline, and its application to the influence model. Section V contains the majority of our contribution as well as key insights into the challenges involved. Section VI contains the analysis of two different CIC datasets as case studies, and confirms the consistencies with real life observations. Section VII concludes this study and outlines future research directions.

II. BACKGROUND AND RELATED WORK

The vast data associated with SNS makes isolating influence difficult, however, there is evidence of not only state-sponsored influence, but individuals determined on shaping the nature of SNS and the populations that engage with SNS [7]. Extensive research has already surveyed and defined the hierarchical schema of SNS [8] which have been used in various research studies, for example, text mining research to explore the trending or popular actions [9]. Studies have also defined areas for potential future research [10], while others have applied semantic analysis of SNS to track and assess the influence of content shared across these platforms [11]. Whilst some of this research has touched on the physical realm associated with cyber influence, most of this prior research is limited to only the cyber realm.

The role of cyber influence in physical conflict is highly important, and is commensurate to that of conflict propaganda [12] first popularized in WWII by Joseph Goebbles [7]. Whilst Goebbles was limited by the media available at the time, his propaganda still had significant influence on the population and final outcomes of the conflict. Propaganda and the role of influence in conflict has evolved continuously from the ''hearts and minds'' campaign from the Vietnam War [13] to modern information warfare [14]. With social media fast becoming the sole source of information for individuals, it is very likely that social media will be the future front line of conflict. Shaping what a user is exposed to, consumes, and believes is as effective as any kinetic effect could be [12]. The race to build tools that absorb and analyse big data will result in a competitive advantage [15].

The fundamentals of cyber influence are mainly derived from the following four areas of related research. The first area of related research is computer platforms used to propagate influence. The second area is the analysis of social networking and their place within modern society. The third area is the theory of influence and how influence is modeled. The fourth area is the role of influence within conflict. These four related areas are discussed in the following four subsections.

A. COMPUTER PROPAGATED INFLUENCE

The founder of the concept of persuasive computers was Fogg [16] who theorized that computers have the ability to persuade individuals and coined the term ''Computer as Persuasive Technology (CAPT)'' [17], [18]. Further work, termed Mass Interpersonal Persuasion (MIP) [19], applied CAPT at scale. When understanding the outcomes and effects of influence, captology only considered compliance of the agent as the result. Xie *et al.* [20] expanded on the outcomes of social influence by including **obedience** and **conformity**. These terms were drawn from the work of Cialdini in 2004 [21]. In light of [21], Xie *et al.* recognized three effects of cyber influence: *compliance*, *conformity*, and *obedience*. Compliance refers to a particular type of response known as acquiescence, i.e., a request. The request may be explicit or implicit, but in all cases the target recognises being urged by the source to respond in a desired way [21]. Conformity is the act of changing one's behaviour to match the responses of others [21]–[23].

B. ANALYSIS OF SOCIAL MEDIA

The second area that relates to cyber influence is the analysis of online communities, social media, social networking, and their role within our modern society. Over 4.5 billion people are now estimated to be online [24]–[26]. Facebook reports a monthly active user group of over 2.45 billion and over 1.62 billion daily active users [27]. Combined with smart phones, an individual is easily identified within cyber meet spaces, which allows for categorizing individuals into age,

gender, and ethnic groups for the unique and targeted delivery of content and marketing.

The power and implications of social media are well established and have become a popular research field, resulting in a spectrum of studies from the psychological, socio-economic, academic, and industrial fields [28]–[32]. When analyzing Twitter specifically, the size and scale of the tweet stream is problematic; hence, the development of techniques and methods to categorise tweets is important [33]–[40]. Many studies assist commercial endeavours for businesses interested in sentiment assessment or optimizing product exposure and promotion [40]–[42]. Subsequently, the ability to fabricate highly technical results by Social-media Data Providers (SDP) can impact corporation models of new and developing businesses. Hence, [43], [44] investigated methods of verifying correctness and completeness of their data and the results generated from social media analysis. Also, [45] uses an ontological approach to build trust and transparency into social media data. These related studies show the depth of research already provided and we leverage the knowledge gained by them in building our cross-realm model.

C. INFLUENCE MODELS AND MODELING

The third related area of research draws upon influence modeling. The theory of quantitative models of influence began with the ''two step'' flow model [46], [47]. The ''two step'' model represented a person with an established reputation within the community (celebrity), who would then pass on their influence to their network. Watts and Dodds critically reviewed the ''two step'' model and quantified the effect of so-called *influentials* [48]. This research negated the theory that specific individuals have significantly more influence over others within a network. Whilst these individuals do exist, they were considered only one factor, with the state of the network as a whole being a larger factor. This research was completed before the existence of modern SNS, but is applicable to any form of communication that can be represented as a network. The Watts and Dobbs model manipulated the thresholds to activate neighbouring cells to initiate information cascades. By changing the nature of the network, Watts and Dobbs [48] were able to demonstrate the ability to create information cascades regardless of who within the network was activated first.

The *Sick, Injured, Recovering* (SIR) model for pathogen modeling was theorized by Coleman in 1957 [49]. The SIR model differs from the influence flow model [48], in that the SIR model has no memory. Rather, the SIR model treats each interaction as ''pure'', as opposed to observations over time. This relates to social media because, each interaction and information exchanged is typically accepted or considered pure. As such, consumers of information and content from these networks are in a highly vulnerable position. Recently, the AAS study [50] investigated the cat-and-mouse type game of detecting and countering detection of fake news on social media. The AAS study outlined the importance of intervention measures to protect the public, such as education and personal fact checking, in addition to platform structural changes to prevent exposure to such material. The AAS study specifically outlines that ''There are no comprehensive data-collection system to provide a dynamic understanding of how pervasive systems of fake news provision are evolving''.

As mentioned in Section II-B, commercial applications are popular, which has resulted in the emergence of the field of influence maximization [51]. The study [51] improves computational efficiency by using cloud computing and specifically designed algorithms. Other influence modeling examines security threats in Social Gaming Networks (SGN) [52], using influence modeling to identify why certain players are targeted by scams or cyber attacks. Influence modeling has also been employed to improve the detection of subtle and long-running radicalization of individuals [53]. These examples show that there are many forms of influence modeling and methods of defining a campaign. For the purpose of our CIC model, we define influence as an outcome of an action, contributing to or resolving conflict.

D. CYBER INFLUENCE IN CONFLICT

The power of cyber influence is increasingly being recognized by governments and militaries around the world. In 2016, NATO published the study [54] which focused on Influence Cyber Operations (ICO) as a subset of Influence Operations. The NATO study [54] identifies operations that are conducted in the logical layer of cyberspace with ICOs targeting attitudes, behaviours, and decisions, and specifically, ''hacking minds by shaping the environment in which political debate takes place''. A key prediction from [54] is the increased employment of ICOs due to the promise of ''victory through non-kinetic means to erode the adversary's willpower, confuse him, constrain his decision making and undermine public support'' with little to no attribution.

Singer and Brooking [4] draw a parallel between cyber influence and Clausewitz's concept of war being ''an extension of policy by other means''. War is used to enforce one nations narrative or policy on another. Cyber influence does the same without the physical violence or destruction. Cyber influence can be achieved through communication directly or indirectly with the population itself, thus bypassing or reinforcing diplomatic channels. The US Army Cyber School recognized that adversaries are weaponizing social media to attack the American social and political environment [55]. The study [55] highlights the malicious nature of ''foreign governments employing a combination of state-sponsored media and personas who support their positions on social media and disrupt free discourse'' and America's requirement to advance their cyber and information operations to counter this threat. In order to do so, [55] proposed the development of the *1st Troll Battalion* to conduct both offensive and defensive ''trolling'' operations.

Assessing a state's cyber power by measuring cyber influence was investigated by [56]. The study [56] found that the ''logic relies on the assumption that the same skills that allow actors to be successful at social media operations

also enable them to be successful at offensive and defensive cyber operations''. The study concluded that cyber operations require orders of magnitude of greater skill and technology compared to cyber influence. The study [56] also identifies that there is very little cross-over of skill between the two disciplines, cyber operations skill sets being technology driven, whilst cyber influence skill sets being psychology based. The framework [57] posed four research questions: 1) How do groups use social media to recruit and shape the ideology of potential followers? 2) How do elites and world leaders use social media? 3) How do technology advances influence the strategic interactions of actors in highly dynamic settings? 4) Does the reduced entry cost of communication increase partisan and ethic polarization, as well as erode the trust in mainstream media?

There remains significant research to be conducted that investigates the interplay between both the physical and cyber realms in cyber influence. In order to identify and quantify information operations or propaganda, there is a requirement to build organized and flexible data structures and datasets capable of representing influence flow across realms. There is also a pressing need to develop logical and flexible analytical tools that leverage these next-generation data structures to identify influentials, regardless of their genesis, as well as misinformation and automated activities.

III. CYBER INFLUENCE CAMPAIGN EXAMPLES

The previous section discussed research fields that contributed to the concept of CICs, the role of CICs within conflict, and the requirement to evolve cyber influence concepts within the technical, policy, and academic fields. This section explores real examples of CICs to demonstrate their scale, outcomes, and time frames. This section shows that CICs can transfer influence in both the physical and cyber realms with a spectrum of methods and techniques.

A. SCALE

The concept of scale of a CIC is target dependent, which could be as small as a single agent, or as large as an entire state. Chicago Gangs often use low-level or individual-orientated CICs daily. Commonly refereed to as ''Cyber Banging'' [58], these individual CICs are typically initiated by single agents to support gang violence [59]. The CIC techniques include tagging oneself in rival gang territory or posting inflammatory comments on rival gang members' posts [60]–[62]. More complicated techniques involve gang members increasing status by promoting ones persona to their digital audience [61].

At the other end of scale, there is state-on-state conflict. The use of international CICs is now becoming commonplace [4]. A recent example was the February 2019 India vs. Pakistan conflict which was started by a terrorist attack against an Indian convoy [63]. After an Indian retaliatory strike on a terrorist camp [64], a small CIC quickly became a large CIC that leveraged popular celebrities to increases it's impact [65]. The #IndiaStrikesBack and #BalakotAirstrike networks were prominent and quickly became politicized and led the narrative of the conflict as well as the upcoming government elections [66].

B. OUTCOMES

The desired outcomes or influence effects of a CIC will determine the target or targets, the techniques to be used, and the required scale. Outcomes achievable with a CIC are also on a spectrum from personal or local, all the way to political and international. Small-scale CICs typically target individuals with personal or commercial outcomes. Larger CICs can have far greater and longer lasting outcomes. For example, the Al Hayat Media Center is an Islamic State of Iraq and Syria (ISIS) media branch [67]. Al Hayat are funded to operate CICs for various outcomes, such as recruiting to ensure the survival of the group [68], re-branding the group as a legitimate government alternative [67], influencing potential candidates, and inciting violence using coordinated tactics [69]–[72]. Al Hayat were also the first group to use a large-scale CIC concurrently with an application called 'Dawn of Glad Tidings' [70]. Whilst eventually shut down, 'The Dawn of Glad Tidings' served to reinforce ideals and opinions, creating what is commonly known as an ''Echo Chamber'', restricting nuance and only allowing strict ideological messaging [73].

There are also acute examples of large-scale CICs that can be hijacked or repurposed for other outcomes. In late March 2014, Russian forces were lawfully invited into the Crimean Peninsula to help settle a social unrest [74]. A legitimate request on the surface, however, it was the result of a long-running large-scale CIC. It began with a domestic unrest due the Ukraine President ceasing discussion on a EU trade agreement. Domestic protests were initiated by individuals using the #euromaidan network. The hashtag gained popularity as a single point of coordination and voice of the people. At the same time, Russia saw this as an opportunity to reclaim Crimea. Russia used what is now known as the Dulles Doctrine [75] to dominate narratives within social media. Russia defines operating within social media as an evolutionary Information Warfare, ''a permanently operating front through the entire territory of an enemy state'', which can asymmetrically lower an adversary's combat potential [76]. Russia used state-level resources to push pro-Russian messaging on the #euromaidan network and influence support for Russian intervention in Crimea.

C. TIME FRAMES

CIC time frames are closely linked to the used techniques and the desired outcomes. Intuitively, there is a linear relationship between the scale of a CIC versus the time frame and investment. Whilst individual small-scale CICs can be launched almost instantly from a single agent account, large-scale influence requires CICs to closely coordinate a critical mass of accounts. In the #euromaidan example, the hijacking was possible because the accounts posting to the network looked legitimate. Russia's Internet Research Agency (IRA) [77]

or ''Troll Factory'' [78] accounts looked legitimate because hundreds of bloggers were paid to build false identities. They then pushed pro-Russian messaging, praised Putin, and denounced opposition in forums, social networks, and comments boards; thus, achieving a coordinated effect [79]–[81].

For purposes of demonstrating our model and ontology we selected two of the reviewed CIC examples for case studies: The #euromaidan campaign is a particularly interesting case due to the corruption of the network as well as the CIC evolving into a state-on-state conflict. The second case study is India vs. Pakistan and the Balakotstrike. This CIC will be valuable due to the highly correlated physical events. The #AlleyesonISIS campaign would also be very interesting given the significant influence and intimidation achieved, however, much of the graphic content posted has been scrubbed from Twitter and thus makes detailed analysis infeasible.

IV. SEMANTIC MODELING, RDF/RDFS, AND ONTOLOGIES

A. SEMANTIC MODELING

Ontologies have been well researched with some modern examples found in [82]–[85]. At its core, an ontology is a graph, using graph theory to collate and organise information. Essentially, an ontology is a number of definitions, relationships, and inference rules. For example, heterogeneous data provided by various devices and sources can all be integrated and applied with commonality and uniformity [86]. An advantage of semantic modeling is the ability to link established ontologies. This means that terminologies of objects with their inherent properties for common concepts have to be defined only once and remain the same in various ontologies. This reduces replication and maintains consistency once an ontology is stable, but also means they can be leveraged by other models. For example, [87] generates separate ontologies within an evaluation model to categorise tweets and detect spam.

Cyberthreat researchers employ semantic modeling to categorise large and unstructured datasets collected from cyber attacks. This approach allows to ''provide a flexible framework for representing and structuring the large variety of data with which security analysts are confronted'', the framework can then be used for implementing cyber security analytic tools [88]. One of the key benefits of semantic modeling is that a single query will result in all the information about a particular instance or object, thus improving search and time efficiencies within large datasets.

Semantic modeling and information structures are applicable beyond computer science, cyber security, and engineering. Ontologies are able to logically and conceptually map information, making them versatile and valuable to numerous research fields [89]–[92]. Masolo *et al.* [93] proposed a formal framework to examine the relationship between (scientific) models and empirical observations. The study [93] uses an ontological approach to address the problem of observational conclusions and the potential for inconsistencies that underline the knowledge gained from the observations.

B. RDF/RDFS

The Resource Descriptive Framework (RDF) is a standard for data interchange on the web [94]. The Web Ontology Language (OWL) is built using RDF Schema (RDFS) which extend link data via Unique Resource Identifiers (URI) resulting in only one instance of data being allowed to exist. RDF/RDFS allows for linking even if the underlying data schemata are different.

C. ONTOLOGIES

This section outlines a selection of related existing ontologies. The advantage of exploring these existing ontologies is that they can either be leveraged by our ontology or tailored to support our requirements. Generally, as we explore these ontologies, it is important to remember the question, ''why should we use an ontology?'' The simple answer is because an ontology is well suited to artificial intelligence (AI) applications. As we will discuss in both this section and in Section VI, the ontology is one method to enable AI to bring meaning to an environment. An ontology achieves this by building causal links of ''related'' data to enable the discovery of new information. As shown in [8], an ontology can ''identify the order of relationship among the entities'' which can then be processed by an AI algorithm. This same principle is employed in [95], [96]. Hence, the development of an ontology is a building block of AI research. The focus of this study is on developing and demonstrating a functioning practical ontology for CIC modeling. Future research on AI algorithms can then build on the ontology developed in this study to discover and infer new information and make better, more accurate decisions about CICs based on the developed CIC ontology model.

1) GOOD ONTOLOGIES

A good ontology as defined by the World Wide Web Consortium (W3C), means that it is well documented, differentiable, used by independent data providers, and possibly supported by existing tools [97]. Many of the ontologies used in research and academia, or published in the public domain, use good ontologies as a baseline. A well-known ontology is Friend of A Friend (FOAF) [98], which represents relational networks of online social media and was one of the first ontologies to highlight the potential of semantic modeling. In [98], an individual person can be linked to others using the `foaf:knows` relationship as well as online artifacts, such as documents and URLs, building an understanding of social media. The Socially Interconnected Online Communities (SIOC) ontology [99] is commonly used to represent communities. The Dublin Core (DC) ontology [100] is a lightweight generalist ontology used to describe metadata. Many of the following ontologies extend these good ontologies for a specific purpose or requirement.

2) CONSENT ONTOLOGY

In response to the introduction of personal data laws in Turkey in 2016, researchers at Ege University developed an extension of the FOAF ontology to track the consent of a person to process personal medical data. The semantic solution [101] allowed Turkey to comply with international laws but also to manage this data. The extended ontology [101] imported the FOAF ontology, leveraging the `Person` class. Secondly, due to the legal age requirements, FOAF was further extended with additional classes, such as '`foaf:HasMinAge` which is Boolean and either above or below 18 years old. To allow for consent to be granted by a parent or legal guardian, additional classes are imported, namely `foaf:MotherOf foaf:FatherOf foaf:RepresentativeOf`. This consent ontology now tracks if consent is provided (another Boolean class of :permission or :prohibition) and who provided that for legal history. The extended ontology [101] demonstrates the ability to import established ontologies and extend them for other purposes.

3) FOAF ACADEMIC

Kalemi and Martiri [102] developed an extension to the existing FOAF vocabulary to include professional achievements and bring people closer to others with similar interests, topics, and research. Kalemi and Martiri [102] focused on the academic community, extending FOAF to cover academic-specific terms and relationships. An example of this is the 'afoaf:university' class. A main class of the ontology, narrowing down the academic community to a geographic location. FOAF academic also defines axioms which allow for richer information, but also assurance of the information. For example, Rule 1: If person A and person B are at University C, they are colleagues. Rule 2: If person X and person Y work at a department D, then they are in the same department and Rule 1 is inferred. This ontology shows the power of axioms and ability to enrich information in meaningful ways.

4) OSN EXTENSION TO FOAF

El Kassiri and Belouadha [103] extended FOAF to address the evolution of Online Social Networks (OSNs) through a Unified Semantic Model (USM). The USM leverages three good ontologies, FOAF, Semantically Interlinked Online Communities (SIOC), and Simple Knowledge Organization System (SKOS). USM extends FOAF by using membership, association, and organization to imply ideals and potential persuasions. The study [103] demonstrates that unique extensions (including classes not traditionally associated with FOAF) can provide specific deep insights.

5) SNS ANALYSIS

Nie *et al.* [104] used text based analysis to identify bursty hot events within Twitter. They clustered key words using a domain ontology. Leveraging the graph structure of an ontology, enabled measurements of the distance between words through the graph. Hence, key words could be used in different contexts, but their syntactic and semantic definitions remained the same. Fang *et al.* [105] proposed a unified ontological model for cross-media events, allowing combinations of SNS platform data for SNS analysis. Dhiman and Toshniwal [106] used an ontological model for specific event detection, focusing again on text analysis and then the generation of graphs around related textual content to form relationships and linkages [107].

Whilst important research and applicable to SNS analysis, specifically to the automated detection of malicious events, these studies differ significantly from our work. Both our model and interests are positioned at a higher level of abstraction. Our ontology takes into account both physical and cyber events, enabling our ontology to determine the causal influence between physical and cyber events, as well as to quantify the influence of a physical versus cyber action.

6) INFLUENCETRACKER ONTOLOGY

To the best of our knowledge, the closest study to ours is the InfluenceTracker ontology developed in 2014 by Razis and Anagnostopoulos [108] to specifically represent the influence of Twitter accounts on each other. A very specific instance of cyber influence, the study limits influence strictly to the cyber domain and only uses the Twitter platform. Important for future studies are the metrics developed to quantify influence of one account over another. For example, the Followers to Following ratio (FtF ratio) and the Tweets Creation Rate (TCR) provides quantifiable and measurable metrics to determine if one account influences the network more than another. The InfluenceTracker leverages the FOAF ontology for representing an agent and uses a similar hierarchy of classes.

V. PROPOSED CYBER INFLUENCE CAMPAIGN (CIC) MODEL

Figure 1 depicts the proposed cyclical CIC model, showing the flow of influence from action, to network to agent, through the cognitive filter and back to action in a cycle. The green boxes are classes, the black links are predicates, and the blue boxes are states of the cognitive process. This model is then integrated into an ontological representation using the Terse Triple Language (TTL) and abbreviated to `cicmod`. Whilst researchers have been able to observe that nefarious injection of content can steer the climate and discourse of an issue [7], [109], [110], their outcomes are based on data analysis and pattern recognition. Our model and ontology formalises the underlying relationships, identifying the foundational causal links between the physical and cyber realms in terms of influence flow. By creating this framework, we can understand how an influence campaign starts with a cyber action, flows through agents and networks, and results in real-world physical actions. The model observes actions taken by the agents, applied to networks of agents who then take further action. The cycle stops only when all agents within a network take no

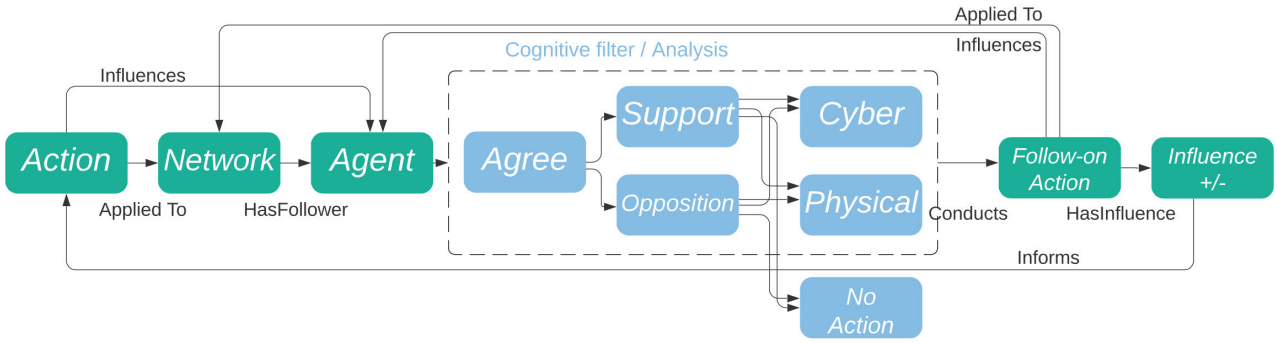


FIGURE 1. Proposed CIC model with flow of influence, from an action to network to agent, through the analysis (cognitive filter), back to action with influence being the outcome that informs the action. This is a continuous cycle in both the physical and cyber realms until all agents in a network take no action.

TABLE 1. Novel classes defined for CIC model as well as predicates defined to connect the classes.

Classes: `cicmod:Action`, `cicmod:Network`, `cicmod:Realm`, `cicmod:Analysis`, `cicmod:Influence`
Predicates: `cicmod:Initiates`, `cicmod:Informs`, `cicmod:AppliedTo`, `cicmod:Influences`, `cicmod:HasFollower`, `cicmod:IsFollowing`, `cicmod:HasAgent`, `cicmod:BelongsTo`, `cicmod:Conducts`, `cicmod:ConductedBy`, `cicmod:HasInfluence`

action. Section V-A explains the design decisions behind the classes and level of abstraction. We have developed the novel classes and predicates to connect the classes in Table 1. The agent's cognitive filter is represented by the `cicmod:Analysis` class, a process that captures disposition, motivation, and realm in boolean sub classes. The `cicmod:AppliedTo` and `cicmod:Influences` predicates are shown linking back to the next network and agent, respectively, in the cycle. Finally, the `cicmod:Influence` class is the result of observing any action taken by an agent. The `cicmod:Influence` class also provides feedback to the CIC via the `cicmod:Informs` predicate in either positive or negative states.

Our ontology allows for abstract concepts, such as Influence and Follows, to be represented simply with predicates. Whereas, these abstract relationships between objects would be very difficult to capture with statistical models or conventional database structures. Moreover, our ontology allows for objects to be related across domains. A predicate can relate a physical event, such as a protest, to an object in another domain, such as an online message or CIC. As a result, through our CIC ontological model, we can translate abstract cross-domain concepts into a format that can be machine interpreted. Thus, our CIC ontological model allows for the application of logic rules to discover information within large and potentially unrelated data.

A. KEY MODEL COMPONENTS

The following sections describe each class in detail.

1) ACTION

The action class is designed for any type of action that could have an influence effect. Any online publication or communication, e.g., post, tweet, vlog, blog, or opinion, is considered a cyber action or an action taken in the cyber realm. By exclusion, all other activities not taken in cyber space are deemed to be physical actions taken in the physical realm, e.g., talking, voting, and violence. Actions are applied to both the physical and cyber networks. This means that an agent applies an action to a network, and another agent consumes the action by being connected to the network. In the physical realm this would be attending a lecture, presentation, or address. In the cyber realm, this is subscribing, following, liking, or searching for any point of reference of the agent's action. Whilst the agent conducts an action, the influence is the result of the action. Hence, the action connects influence to agents.

The `cicmod:HasInfluence` predicate represents the resultant influence of the action in either the positive or negative state. Many of the platforms have their own metrics for this already, such as upvote, downvote, like, dislike, thumbs up, and favourite. These metrics align with one of the three established influence categories, see Sec. II-A. The predicate `cicmod:Influences` is the action's effect on the agent and assumes the same metrics as `cicmod:HasInfluence` (like, dislike, and voting). This design allows both simple and complex actions to be represented, from traditional support campaigns to false flag campaigns.

2) NETWORK

The network class represents the first degree contacts of an agent or thing. An agent or thing can have multiple networks, in both the cyber and physical realms. For example, an agent may have multiple Twitter, Instagram, and YouTube accounts, (#musicfestival, #lollapalooza) representing multiple cyber networks, also referred to as cyber agent networks. In the physical realm, an agent may have multiple networks, such as, friends, colleagues, and family. In either realm, the networks can be accessed by other agents within the

network. For example, when #election2016 was hijacked and used against a running candidate. Or, a town hall meeting can be used to voice the opinion of anyone that attends. As such, the metrics for networks must be native to the network itself. Data properties, such as **Followers** and **Following**, belong with the network class. The following-to-followers ratio is an established influence metric [108], however, our model determines scale by considering the `HasFollowers` metadata.

3) AGENT

The ontology employs the `foaf:Agent` [98] with its established definitions. Our cicmod ontology establishes additional predicates and linkages, but the definition of the agent class remains the same.

4) REALM

Two possible representations of realm were considered. The first being an individual object class realm. For an object class to stand alone, it must be contextual and defined to make sense. That is, an action, is contextual, a cyber, is not. Therefore, the second representation of realm was employed, this being a sub class of an object class. That is, an action within the cyber realm being a cyber action. Sub classes for each object class were developed, such as an agent cyber network, `cicmod:AgentCyberNetwork`, and an initial physical action, `cicmod:InitialPhysicalAction`.

5) ANALYSIS

An analysis class was originally developed to represent a cognitive filter function as shown in the center box of Figure 1. The cognitive filter is an individual's analytical process. This analytical process is highly complex and beyond the scope of this study. We abstract the process into three yes or no questions: 1) Does the agent **agree** with the message or content of the action? 2) Is the agent **motivated** to take further action? 3) If motivated, in which **realm** will the agent take action? The output of the analysis determines the follow-on action, and influence can be determined by observing the action which is explained further next.

6) INFLUENCE

Quantifying influence is highly complicated (and specific to every use case) for a number of reasons: 1) An agent's disposition with respect to a CIC cannot be assumed, i.e., does the agent already support or oppose the theme of the CIC? 2) How did an action achieve influence? However, the influence can be relatively easily determined by observing the follow-on action of an agent. Our model quantifies influence by observing the state of a follow-on action, i.e., retweets indicate support, while downvotes, dislikes, and thumbs down indicate opposition. No prior information is required for this assessment of influence as positive or negative; rather, this assessment can be based on native metadata.

B. MODEL FLOW

The cicmod ontology is designed to be cyclical, see Figure 1: Actions are applied to networks of agents that in turn take more actions. These iterate forever until all agents in a network do not take any further action. The following is an example of one cycle of the model:

- *:Initiate* A group decides to begin a Cyber Influence Campaign and engages an agent.
- *:Action* The engaged agent posts 'Elect John for President' to their Twitter account.
- *:Network* The agent's followers on Twitter are delivered the post.
- *:Agent* An individual is part of the engaged agent's network and consumes the post.
- *Analysis* The individual makes a decision whether or not to act.
- *:Action* The individual re-posts the initial post with a 'thumbs up'.
- *:Influence* Positive influence is inferred due to the 'thumbs up' associated with the repost.
- *:Informs* The positive action taken informs the group that the campaign is working as desired.

In this example, we can observe influence flowing from agent to network to agent to action. Therefore, we can observe the model representing the action, network, agent, and influence. The influence is captured by observing the nature of the action taken. That is, if the follow-on action is supportive, then the influence was positive.

C. REFINEMENT OF THE MODEL

We tested our cicmod ontology with small datasets from TrackMyHashtag [111], which are discussed further in Section V-D2. To ensure that our model continued to reflect reality, we made the following refinements:

- 1. The analysis class was not required. The decision process does not change the outcome, nor does it provide any additional insight into the influence assessment of the action. Hence, the cognitive filter (`cicmod:Analysis`) was removed.
- 2. Cross platform indicators were removed. There is no additional value in knowing which platform the action is taken on, as for this study's purposes, all action have the same potential influence.
- 3. Initially, influence was assessed at the agent. As mentioned, only an action has influence, hence, the influence must connect the action to the agent, not the agent to the agent.
- 4. `cicmod:Following` and `cicmod:Followers` were changed to data properties of the network. As this allows for the networks to have scale and for an agent to have multiple networks in different realms.

D. DATA PIPELINE

In order to test the ontology using real-world data, we needed a real-world dataset from a campaign. This was a complicated

process, as there are a number of steps required to take raw data from a social media platform and turn it into triples (i.e., semantic objects) for a functioning ontology. This was achieved through the following steps:

1) CAMPAIGN SELECTION

We identified that the campaign needed to have two key elements. First, the campaign needed a strong physical timeline of actions and events that were easily distinguishable and consistent in reporting. Second, the cyber activity had to be of significant scale, i.e., above the noise floor. We began this process by considering a number of well-known CICs. We discovered that a suitable campaign should be bipartisan, as this reduced complexity. Also, the involvement of a military resulted in reporting being somewhat consistent and readily available. Thus, we selected two campaigns, namely the euromaidan protests during the Crimea crisis of 2013/2014 and the Balakot Airstrike during the Indian-Pakistan hostilities in 2019.

2) DATA COLLECTION

In order to achieve a complete understanding of social influence propagation through the network we require full-take or ''fire hose'' Twitter datasets. Without loss of generality, we focused on the Twitter platform as it is simplistic and consistent. Also, the obtained example data showed that Twitter metadata contained network detail, hashtags, and user generated content. Text logs from any of the other SNS platform, e.g., Instagram, WeChat, and Facebook, would be equally suitable for our cicmod ontology model. To test the ontology, small trial datasets were used. These are manageable percentages of the full-take Twitter stream and were easily obtainable. Trial Twitter data came from two different sources. 1) TrackMyHashtag [111] which only provided 100 tweet samples, and 2) Spritzer style Twitter logs from the Internet Archive [112]. These were suitable for testing and also helped confirm the nature of each criterion, e.g., hashtag, date range, and agents. From testing with the [112] logs we confirmed the suitability of the two campaigns. Then, we employed the third-party website TweetBinder [113] to access the developer.twitter API [114] and to provide the key hashtags over the date ranges. The number of tweets were confirmed by cross-referencing the quotes provided by Twitter academic support staff and [113]. The #euromaidan campaign resulted in over a million tweets and more than 3 million tweets with first-tier related hashtags. Similar numbers were achieved for the Balakotstrike campaign.

3) PHYSICAL EVENTS

The following timeline details the physical events from the #euromaidan campaign which we have been translated into triples for our ontology. We built the timeline using numerous conventional media sources on the conflict [74], [115], [116]. However, when building this timeline, a decision to define an event as either an initial physical action or a physical reaction had to be made. Unfortunately, the definition of initial physical action versus physical reaction can be individually interpreted and potentially introduce inconsistencies. Therefore, to ensure consistency, we interpreted only the first physical event as an initial physical action; all subsequent physical events are interpreted as physical reactions. Therefore, the timeline for the #euromaidan campaign is as follows and a similar timeline was built for the Balakotstrike campaign.

- November 21, 2013: Cessation of EU agreement discussions by President Viktor Yanukovych
- November 21–23, 2013: Small demonstrations in Kiev in response to failed EU association agreement.
- November 30, 2013: Ukraine special police, Berkut, beat unarmed peaceful protesters.
- December 01, 2013: Ukraine anti-government protesters have smashed their way into Kiev's city hall.
- December 13, 2013: Parliament passes restrictive anti-protest laws as clashes turn deadly.
- December 16, 2013: Protesters begin storming regional government offices in western Ukraine.
- December 28, 2013: Prime Minister Mykola Azarov resigns.
- February 14, 2014: 234 protesters arrested since December are released.
- February 18, 2014: Clashes erupt, with reasons unclear: 18 dead.
- February 21, 2014: Crimean parliament members called for an extraordinary meeting.
- February 22, 2014: Vote to remove President Yanukovych and Putin holds meeting to regain the Crimean peninsula.
- February 24, 2014: Parliament votes to ban Russian as the second official language.
- February 26, 2014: Large scale clashes during opposing rallies in Simferopol.
- February 27, 2014: Undeclared Russian troops enter Crimean parliament and Russia commences military training exercise in vicinity of Crimea peninsula.
- March 1, 2014: Aksyonov declared head of police, immediately requests support from Russia to maintain order.
- March 16, 2014: Public vote held to align with Russia.
- March 17, 2014: The EU and US impose travel bans and asset freezes on several officials from Russia and Ukraine over the Crimea referendum.
- March 18, 2014: President Putin signs a bill to absorb Crimea into the Russian Federation.

4) DATA INGEST

In order to translate the data from the raw collection into triples, a unique data translation script was designed, written, and tested. The following points detail key design elements.

- 1. Inspecting the raw data. By inspecting the data before progressing, we ensured that the fields and meta data contained enough detail to populate the ontology. Moreover, the date range of actions covered the course of our specific campaign. We decided to extend the date

range by 10 percent before and after the expected campaign dates to ensure that we caught the initial and final actions. Reference [114] uses the Java Script Orientation Notation (JSON) format to output the raw data. This was advantageous as the JSON dictionary format allowed for simple inspection and the JSON toolset within Python allowed for easy manipulation and processing.

- 2. The code was designed to loop through the JSON file building agents, cyber agent networks, actions, and hashtag networks as triples from each tweet which was contained within a JSON dictionary. Moreover, the code used ''mention'' and ''retweet'' information to build additional networks and agents as they were referenced.
- 4. Privacy issues. Whilst the publication of tweets is public, the agent's user name and alias are not important to our research. Hence, the script anonymized agent names and aliases.
- 5. Additional information. Our real dataset also included artifacts of the actions and networks, such as favourite, location, and language, which were not present in the [112] JSONs. These are all highly valuable search criteria for the ontology and needed to be included. Whilst not in our initial test data, including these elements made our ontology richer and more valuable.

5) TRIPLE STORE

Once the ontology and dataset triples had been built, the file contained in excess of 20 million triples. Therefore, careful consideration of a triple store was required, as many stores cannot handle datasets of this size. We initially had used Protege [117] and WebVOWL [118] to build and view the ontology; however, these were not capable of handling the large dataset. We selected the application Stardog which has been stable and user friendly and included a GUI, the Stardog Studio.

VI. ANALYSIS OF CASE STUDY CICs

The cicmod ontology is applicable to all conceivable CICs, as the object relationships and causal connections remain the same. The graph-based structure of cicmod allows for unique connections to be made though ontological reasoning or inference rules. This section presents the specific analysis of the two selected case study datasets to showcase the evolution from intuitive results through to a deep analysis of behaviors across both the physical and cyber realms of a CIC. Each query has been specifically designed to demonstrate various possible types of analysis. We have analyzed the two selected CICs, #euromaidan and Balakotstrike, to compare and contrast key metrics, such as the size of the campaign, influence actions achieved, and key artifacts of the engaged networks.

It is also important to reinforce the point that our CIC ontological modeling and feature extraction has been tailored to the CIC use case. Our deep understanding of these specific case study CICs and their corresponding physical events provided insights that enabled us to extract suitable features for our analysis. We then used the extracted features for the database and model. Without this deep understanding, there is a potential for misinterpretation which could result in false positives or a model failure.

6) SPARQL

The SPARQL Protocol and RDF Query Language is the semantic query language used with data stored in an RDF dataset. Hence, to access the novel information generated as part of the ontology, specifically designed SPARQL queries must be written. Therefore, a unique query is written for each element of our analysis in order to extract the detail from our ontology. Each query is published with the dataset and hosted together for ease of reference and use as the IEEE DataPort Cyber Influence Campaign Ontology dataset (DOI 10.21227/70kc-yx38).

A. ACTIONS PER DAY

The number of cyber actions (tweets) per day is an elementary quantitative metric for the comparison of campaigns and gives an initial appreciation for the overall volume of the campaign. The number of actions per day does not directly indicate influence; however, provides some initial insights into the scale and behaviour of the CIC and potential time periods that require further analysis. The #euromaidan and Balakotstrike campaigns cover a period of 131 days and 46 days, respectively. The differences in timeline do not impact the numbers of actions per day, which still reflect the relative volume of interest in the issue over time. The SPARQL query first searches all actions, physical or cyber, and then sorts the actions by day and counts the number of actions per day. Figures 2 and 3 show the results of these queries with the physical events represented as vertical lines. This is because each day has a maximum of one event per day in both campaigns, except for February 24th 2014, when there were four physical events attributed to the #euromaidan campaign.

From Figures 2 and 3, we can identify clusters of physical events that correspond to increases in actions per day. This is an intuitive result that demonstrates that our data is accurate and our model reflects reality. Of note, in Figure 2, there are some offsets, as the physical events that were reported on the 13th and 16th of January 2014 did not have an immediate SNS response; the SNS response began to increase on the 19th of January 2014, potentially due to details of the physical events being released. The drop in tweet activity on the Balakotstrike campaign in Figure 3 on the 12th of April 2019 likely indicates the calming influence of the independent inspection of the airstrike location on the 10th of April. In Figure 2, the actions per day for the #euromaidan campaign peaked at around 80,000 tweets per day at the height of the hostilities between protesters and the Ukraine government. Whilst a shorter campaign, the Balakotstrike SNS activity spiked to almost 230,000 tweets per day in Figure 3 on the 26th of February 2019, the day of the retaliatory Indian strike against Pakistan.

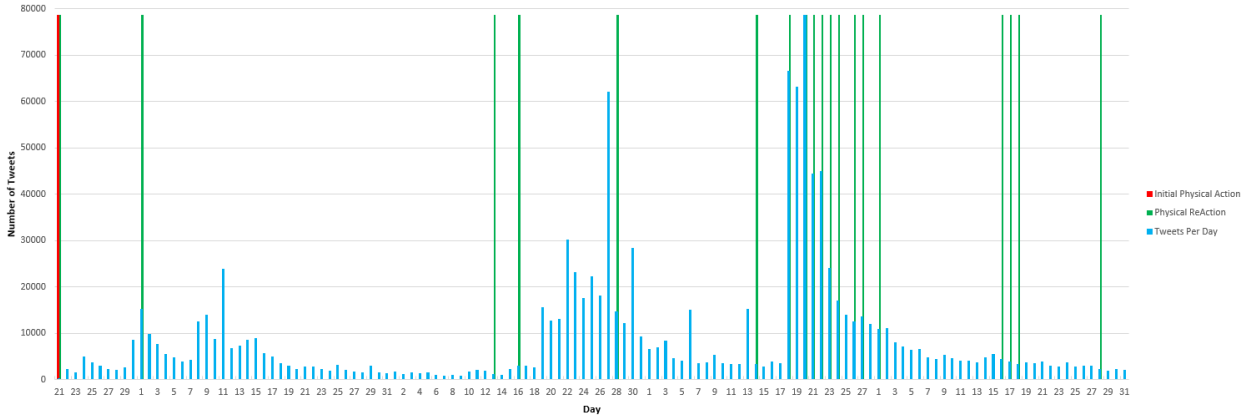


FIGURE 2. Total number of cyber actions (tweets) per day of #euromaidan campaign. The numbers on the x-axis are days, beginning with November 21, 2013 and running through March 31, 2014.

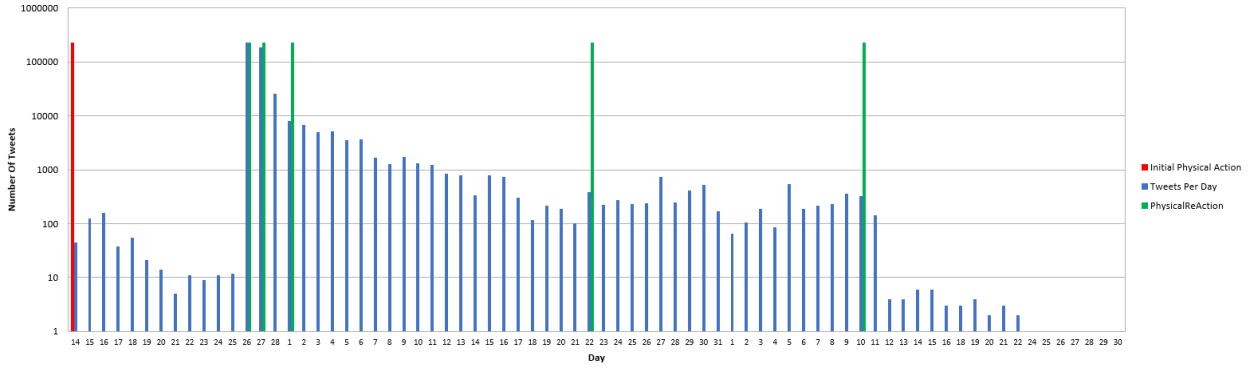


FIGURE 3. Total number of cyber actions (tweets) per day of Balakotstrike campaign. The numbers on the x-axis are days, beginning with February 14, 2019 and running through April 30, 2019.

B. NETWORKS PER DAY

The number of networks per day is a similar quantitative metric as the number of actions per day, however, the number of networks per day metric considers the volume of networks being engaged. The number of networks per day metric represents the diversity of how these actions are applied to the SNS platform. The cyber agent networks are the only networks that individual agents can post to and control. The predicate used in this situation was `cicmod:ControlledBy`. An agent can also ''mention'' another agent in an action. By mentioning another agent, a relationship represented by the predicate `cicmod:Mentions`, connects another agent's network to the action. The hashtag networks connected to the action are captured with the `cicmod:appliedTo` predicate, as a hashtag network is not controlled by an agent. This means that a hashtag network can be manipulated by any agent or narrative. Our raw data did not contain the follower or following metric for the hashtag networks; future research may consider hashtag networks with the follower and following metric.

For both campaigns, our first observation is that the numbers of networks per day in Figures 4 and 5 correlate closely with the number of actions per day in Figures 2 and 3. This is logical as the number of actions taken by agents is expected to be similar to the number of unique networks, because most agents will first post to their own network. Figure 4 shows limited hashtag employment compared to Figure 5. Potentially due to the limited public awareness of hashtags, only a small number of hashtags were used throughout the #euromaidan campaign.

Generally, the smaller the number of agents, the more limited the distribution of information, which curtails the dilution and manipulation of the information and details. For organizing events, a single source of truth is preferable for an organizer if the priority is to coordinate demonstrations; however, the limited distribution restricts the exposure of the campaign. We observe relatively low numbers of cyber agent networks in the early phase of the #euromaidan campaign in Figure 4, which helps maintain the consistency of information. These low numbers of cyber agent networks may also contribute to the very low numbers of hashtag networks in the left part of Figure 4.

Generally, an action needs to be taken in the cyber realm in order to enable the subsequent ''Mention'' as a cyber reaction. That is, mention networks are predominately retweets, which we have corroborated in additional data analysis that is not

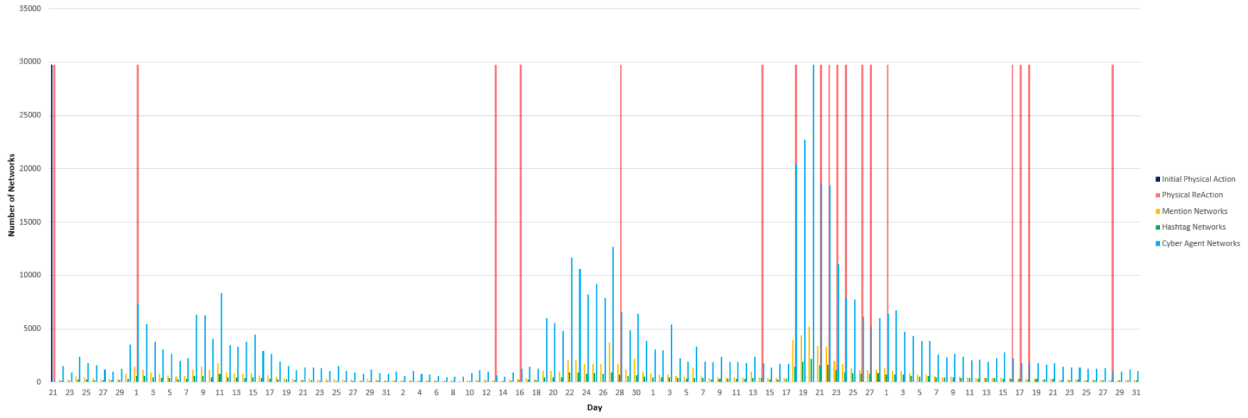


FIGURE 4. Total numbers of cyber agent networks, mention networks, and hashtag networks by day of #euromaidan campaign.

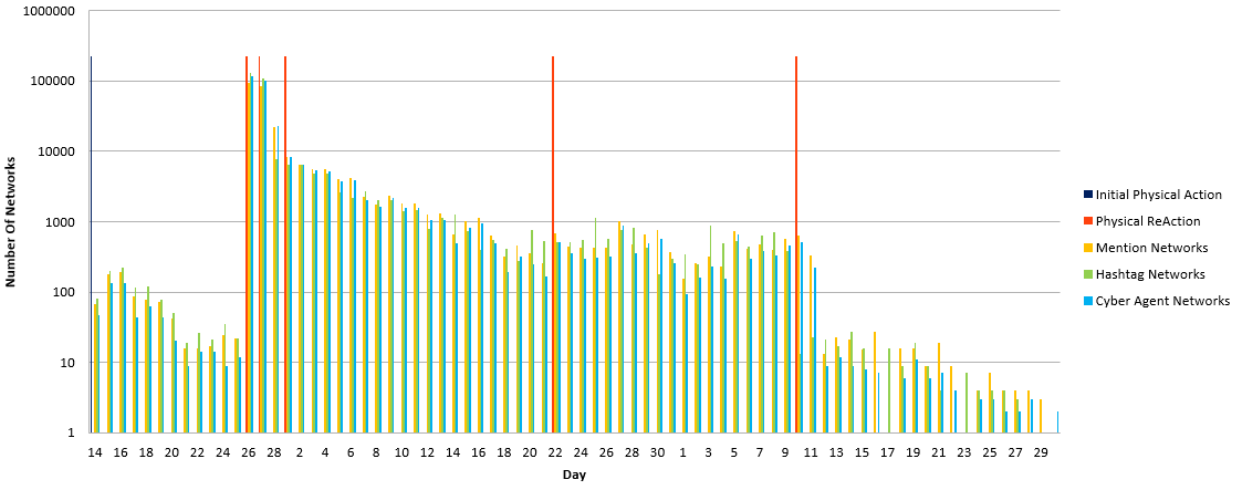


FIGURE 5. Total numbers of cyber agent networks, mention networks, and hashtags networks by day of Balakotstrike campaign.

included here to keep the analysis presentation concise. Our additional data analysis has suggested that the CyberReActions are connected to all three types of networks, giving them the most exposure.

C. POTENTIAL INFLUENCE

Quantifying potential influence is possible by using the `cicmod:AppliedTo` predicate and accumulating the followers (agents) of the total networks that an action influences (`cicmod:Influences`) on a given day d. We define a quantitative potential influence metric $\Pi(d)$ which represents the number of end nodes of our graph. Formally, we denote $a(n, d)$ for the number of actions on a given network n on a given day d and denote $f(n, d)$ for the number of followers of a given network n on a given day d. Furthermore, we define $\tau \in \mathcal{N} = \{\mathtt{a}, \mathtt{h}, \mathtt{m}\}$ as an indicator variable for the network type, which can take on values from the set $\mathcal{N}$ of network types, specifically cyber agent (`a`) networks, hashtag (`h`) networks, and mention (`m`) networks in the context of Twitter. We define $N_\tau(d)$ to denote the number of networks of type τ on a given day d. We then define the potential influence score $\Pi(d)$ on a given day d as:

$$\Pi(d) = \sum_{\tau \in \mathcal{N}} \sum_{n=1}^{N_\tau(d)} a(n, d) f(n, d). \tag{1}$$

The potential influence metric $\Pi(d)$ is akin to assessing the magnitude of a campaign by summing the numbers of network followers (agents) which are influenced by campaign actions. Thus, the $\Pi(d)$ metric allows for a fair comparison of two CICs. The values for Π quickly become massive, reaching orders of 10^9; these numbers reflect the potential end nodes, not the actual agents engaged.

Figure 6 displays the potential influence Π by day for both campaigns on a logarithmic scale. Using this quantitative Π metric we can see that the potential influence of the retaliatory strike from India had a large influence on the Twitter population and by extension the world, reaching Π values above 10^{10}, which are higher than for any of the #euromaidan events. However, the #euromaidan campaign had a sustained

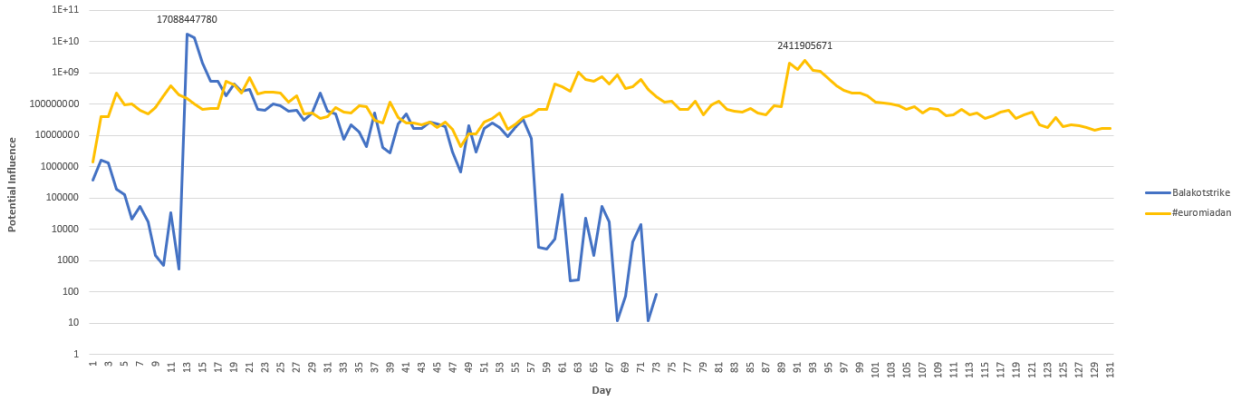


FIGURE 6. Potential influence $\Pi(d)$ of the #euromaidan and Balakotstrike campaigns as a function of the day d, whereby days are numbered starting from the beginning of each campaign.

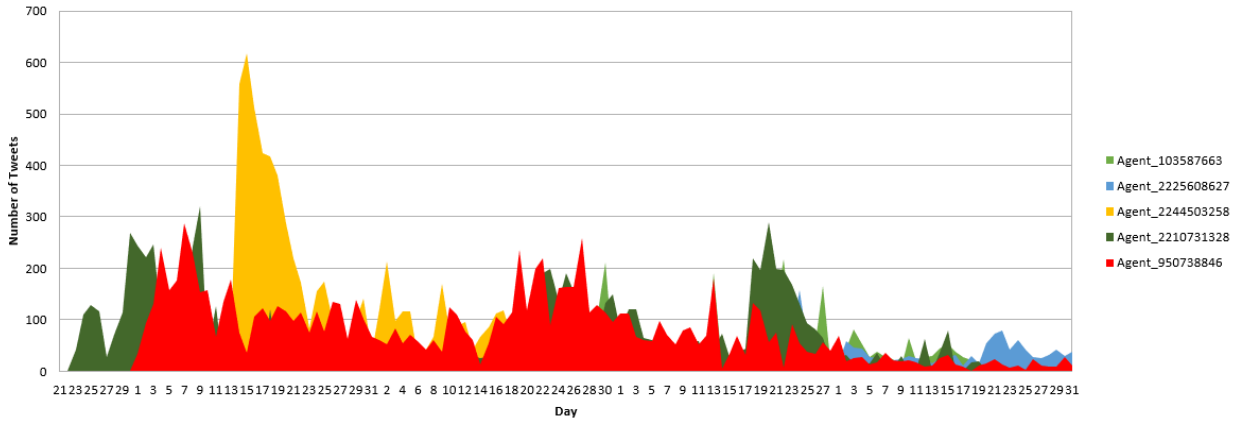


FIGURE 7. Number of tweets per day of top 5 agents within #euromaidan campaign.

influence over time, while the Balakotstrike campaign covered a much shorter time period. This comparison based on the potential influence metric Π as defined in Eqn. (1) provides a unique and novel ability to measure and compare one physical or cyber event against another with a consistent metric.

D. ACTIVE AGENTS OVER TIME

Having established the potential influence of a campaign, the following example has been selected to show the ability to focus on specific details of the dataset. The agent behaviour over time showcases the flexibility of the ontology. This analysis gives a quantitative appreciation of which agents were most active and when. The ability to not only identify key agents within CICs, but to also confirm human or automated behaviours is highly advantageous for the operational analysis of CICs. We conducted this analysis with two sequential queries: the first query discovering the most active agents over time; the second query grouping agent actions over time.

In Figure 7 for the #euromaidan campaign, Agent-950738846 in the data set is observed as part of an initial intense activity along with Agent-2210731328. Their activity peaks at over 300 tweets per day in early January; potentially organizing or reporting on the euromaidan demonstrations. However, their activity is quickly surpassed in mid January by Agent-2244503258, who peaks at over 600 tweets in one day, but then rather suddenly ceases all activity by mid February. This discontinuation of activity warrants further investigation, as it may provide insights into potential automated or state-sponsored activity.

Similarly, in Figure 8, Agent-728486277683777536 peaks at just under 600 tweets per day early in the Balakotairstrike campaign, but then ceases any action. This dynamic suggests that this individual or account was only interested in the initial physical action and not in the subsequent physical events that happened in response.

The actions of Agent-2244503258, who tweeted over 600 times in a day as shown in Figure 7, are shown in Figure 9 per hour, over an eight day period. From Figure 9 we can observe that the activity of Agent-2244503258 maintains periodicity with normal patterns of life for a human agent. This means, sleep patterns are maintained at night as well as

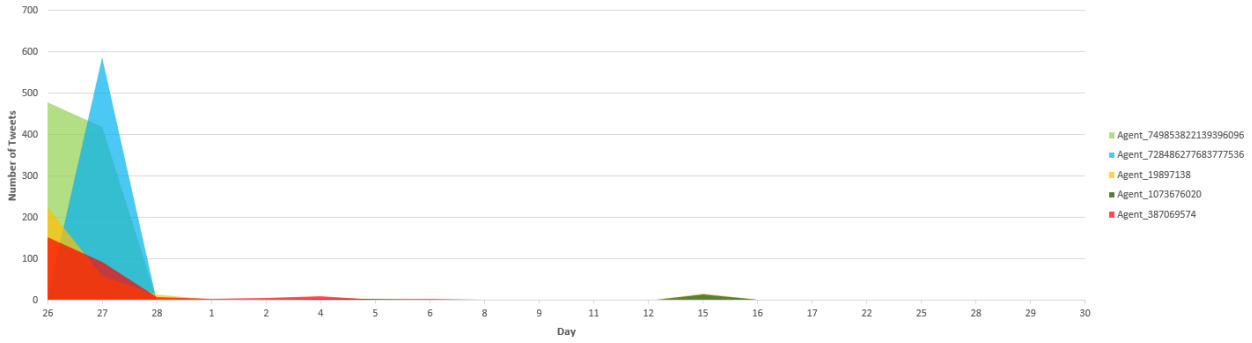


FIGURE 8. Number of tweets per day of top 5 agents within Balakotstrike campaign.

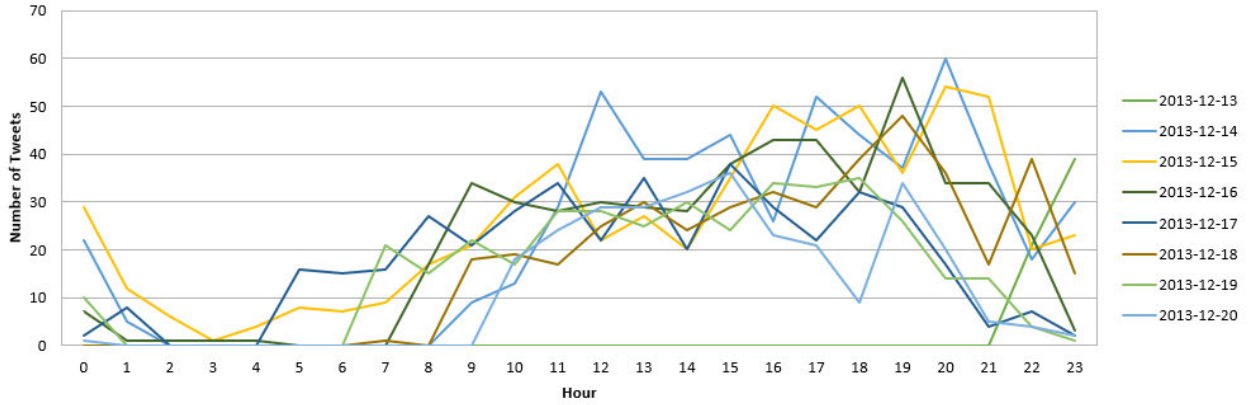


FIGURE 9. Agent-2244503258 actions as a function of hour of the day for eight day period.

peak periods of activity around early evening each day. Other fields within the dataset can also be leveraged to provide evidence of a human or automated agent. The device used for all actions by Agent-2244503258 was a desktop based interface of VK.com and each action came from the same location within the Ukraine. The combination of a single interface device and location supports the theory of a human agent using a desktop interface to publish a high number of tweets over a sustained period.

E. OUTCOME

The quantitative analysis using the CIC ontology has validated the mechanics of our semantic model and shown that the linkages and relationships of objects within the dataset reflect real life. The modeling and ontological representation is novel and provides the basis for future work in the field of cyber influence and the investigation of CICs. The application of our ontology to other research will allow for the integration of physical and cyber events in feature extraction and other machine learning (ML) techniques used in social media analysis. Having established an influence flow semantic model as the basis of our ontology, it is now possible to track and identify influence across realms through leveraging established ontologies.

F. LIMITATIONS AND FUTURE RESEARCH

The presented case study analyses only represent samples of the types of analysis possible with our cicmod ontological model. With the flexibility of the SPARQL query language and graph based cicmod ontology, key insights can be gained into the behaviours and nature of CICs and cyber influence in general. The use of the location and language fields within the dataset is highly versatile for operational and thematic analysis in conflict. The intent of this research was to provide a novel and flexible ontology to progress the field of cyber influence.

We acknowledge that to the best of our knowledge, a theoretical analysis of failure or error bounds of the introduced CIC ontology model is intractable. From an empirical research perspective, two independent CICs have been assessed with the developed CIC ontology in this article. Future research should explore additional CICs in order to determine if there are any scenarios or types of CICs that do not fit the introduced CIC model or cause it to fail. In order to support future research, the ontology, datasets, code for the data pipeline, and SPARQL queries have been hosted as the IEEE DataPort Cyber Influence Campaign Ontology dataset (DOI 10.21227/70kc-yx38).

ML can be employed to recognise and define indicators of activity that may lead to physical events. For example,

determining the preconditions that result in physical demonstrations or potentially a change of leadership within a state. With the ability to compare physical events against each other, we can also use the SNS activity to suggest when activity reaches a threshold to cross domain into the physical realm.

This study has focused on developing and evaluating an ontology model for analyzing cyber influence campaigns in conflicts conducted in social media networks. Social media networks can also give indications of emerging cyber security threats [119]–[123]. One interesting future work direction is to adapt our ontology model to uncover the sources and agents behind emerging cyber threats. Moreover, social media can be used to spread misinformation to wide audiences. In future research, our model could be adapted to identify the sources of potential misinformation.

VII. CONCLUSION

We have introduced a Cyber Influence Campaign (CIC) ontology model called cicmod, our design methodology, and the case study analysis of two real-life CICs. We explained the rationale behind the cicmod ontology model and the translation into a functioning ontology. We have demonstrated and provided the unique code required for a data pipeline, as well as a tailored dataset in ontological format ready for use with a triple store or existing ontology. As a result of this research, we can conclude that our cicmod ontology model functions as desired and accurately reflects the reality of CICs in conflict. Our specific analysis demonstrated that the cicmod ontology quantifies potential influence and that it can scale from localized individual actions through to global state-funded campaigns. This work provides the mechanics of CICs and confirms that the level of abstraction is appropriate to provide detail without complexity. The cicmod ontology can become the basis for understanding and further modeling of cyber influence. Combined with potential future research, cicmod can form the basis for powerful tools in the analysis of SNS and cyber security in the future.

REFERENCES

[1] A. B. Dhiraj. (Mar. 2019). *The 20 Top Most Used Social Networking Sites And Apps in The World, 2019*. [Online]. Available: https://ceoworld.biz/2019/03/03/the-20-top-most-used-social-networking-sites-and-apps-in-the-world-2019/

[2] C. Cadwalladr. (2018). *I Created Steve Bannon's Psychological Warfare Tool: Meet the Data War Whistleblower*. [Online]. Available: http://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopherwylie-faceook-nix-bannon-trump

[3] R. Tench and B. Jones, "Social media: The wild west of CSR communications," *Social Responsibility J.*, vol. 11, no. 2, pp. 290–305, Jun. 2015.

[4] P. Singer and E. Brooking, *LikeWar: The Weaponization of Social Media*. Rancho Cucamonga, CA, USA: Houghton Mifflin Harcourt, 2018.

[5] R. E. Guadagno and R. B. Cialdini, "Preference for consistency and social influence: A review of current research findings," *Social Influence*, vol. 5, no. 3, pp. 152–163, Jul. 2010.

[6] M. Chen, S. Mao, and Y. Liu, "Big data: A survey," *Mobile Netw. Appl.*, vol. 19, no. 2, pp. 171–209, Apr. 2014.

[7] S. Zannettou, T. Caulfield, W. Setzer, M. Sirivianos, G. Stringhini, and J. Blackburn, "Who let the trolls out: Towards understanding state-sponsored trolls," in *Proc. 10th ACM Conf. Web Sci.*, 2019, pp. 353–362.

[8] G. Razis, I. Anagnostopoulos, and S. Zeadally, "Modeling influence with semantics in social networks: A survey," *ACM Comput. Surv.*, vol. 53, no. 1, pp. 1–38, May 2020.

[9] A. Karami, M. Lundy, F. Webb, and Y. K. Dwivedi, "Twitter and research: A systematic literature review through text mining," *IEEE Access*, vol. 8, pp. 67698–67717, 2020.

[10] B. Abu-Salih, B. Bremie, P. Wongthongtham, K. Duan, T. IssaKit, K. Y. Chan, M. Alhabashneh, T. Albtoush, S. Alqahtani, A. Alqahtani, M. Alahmari, N. Alshareef, and A. Albahlal, "Social credibility incorporating semantic analysis and machine learning: A survey of the state-of-the-art and future research directions," in *Proc. Workshops Int. Conf. Adv. Inform. Netw. Appl.*, 2019, pp. 887–896.

[11] S. Bayrakdar, I. Yucedag, M. Simsek, and I. A. Dogru, "Semantic analysis on social networks: A survey," *Int. J. Commun. Syst.*, vol. 33, no. 11, pp. e4424.1–e4424.30, 2020.

[12] R. Stengel, *Information Wars: How we Lost the Global Battle Against Disinformation and What We Can Do About it*. New York, NY, USA: Grove/Atlantic, , 2019.

[13] A. Miller. (2019). *To Win the Hearts and Minds: The Combined Action Program During the Vietnam War*. Accessed: Jul. 21, 2020. [Online]. Available: https://ir.vanderbilt.edu/handle/1803/9474

[14] E. X. Schaner, "What is military information power?" *Mar. Corps Gazette*, vol. 104, no. 4, pp. 17–19, Apr. 2020.

[15] V. Rauta, "Towards a typology of non-state actors in 'hybrid warfare': Proxy, auxiliary, surrogate and affiliated forces," *Cambridge Rev. Int. Affairs*, vol. 33, no. 6, pp. 868–887, Oct. 2020.

[16] B. Fogg, "Persuasive computers: Perspectives and research directions," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, 1998, pp. 225–232.

[17] H. Oinas-Kukkonen and M. Harjumaa, "A systematic framework for designing and evaluating persuasive systems," in *Proc. Int. Conf. Persuasive Technol.*, 2008, pp. 164–176.

[18] A. Sara and H. Mostafa, "Exploring persuasive systems using comparative study between actual technologies," in *Proc. Int. Conf. Big Data Smart Digit. Environ.* Cham, Switzerland: Springer, 2018, pp. 369–379.

[19] B. J. Fogg, "Mass interpersonal persuasion: An early view of a new phenomenon," in *Proc. Int. Conf. Persuasive Techn.*, 2008, pp. 23–34.

[20] Y. Xie, M. Chen, H. Lai, W. Zhang, Z. Zhao, and C. M. Anwar, "Neural basis of two kinds of social influence: Obedience and conformity," *Frontiers Hum. Neurosci.*, vol. 10, pp. 51.1–51.8, Feb. 2016.

[21] R. B. Cialdini and N. J. Goldstein, "Social influence: Compliance and conformity," *Annu. Rev. Psychol.*, vol. 55, no. 1, pp. 591–621, Feb. 2004.

[22] M. Deutsch and H. B. Gerard, "A study of normative and informational social influences upon individual judgment.," *J. Abnormal Social Psychol.*, vol. 51, no. 3, pp. 629–636, 1955.

[23] K.-L. Thomson and R. von Solms, "Information security obedience: A definition," *Comput. Secur.*, vol. 24, no. 1, pp. 69–75, Feb. 2005.

[24] S. Kemp. (2019). *Digital Trends 2020: Every Single Stat You Need to Know About the Internet*. [Online]. Available: https://thenextweb.com/growth-quarters/2020/01/30/digital-trends-2020-every-single-stat-you-need-to-know-about-the-internet/

[25] *Cisco Annual Internet Report (2018–2023) White Paper*, Cisco, San Jose, CA, USA, Mar. 2020. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

[26] S. Edosomwan, S. K. Prakasan, D. Kouame, J. Watson, and T. Seymour, "The history of social media and its impact on business," *J. Appl. Manage. Entrepreneurship*, vol. 16, no. 3, pp. 79–91, 2011.

[27] *Facebook Reports Third Quarter 2019 Results*, Facebook Inc., Menlo Park, CA, USA, 2019.

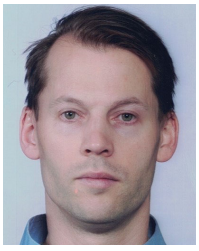
[28] M. Altuwairiqi, E. Arden-Close, N. Jiang, G. Powell, and R. Ali, "Problematic attachment to social media: The psychological states vs usage styles," in *Proc. 13th Int. Conf. Res. Challenges Inf. Sci. (RCIS)*, May 2019, pp. 1–6.

[29] Y. Shibuya and H. Tanaka, "Using social media to detect socio-economic disaster recovery," *IEEE Intell. Syst.*, vol. 34, no. 3, pp. 29–37, May 2019.

[30] X. Xu, B. Shen, X. Yin, M. R. Khosravi, H. Wu, L. Qi, and S. Wan, "Edge server quantification and placement for offloading social media services in industrial cognitive IoV," *IEEE Trans. Ind. Informat.*, early access, Apr. 16, 2020, doi: 10.1109/TII.2020.2987994.

[31] Z. Ning, Y. Liu, J. Zhang, and X. Wang, "Rising star forecasting based on social network analysis," *IEEE Access*, vol. 5, pp. 24229–24238, 2017.
[32] A. Maistrelli, "Palestinian transnational advocacy network: A comparative analysis of the three most powerful network actors and their discursive practices on Twitter," Ph.D. dissertation, Dept. Media Commun., Bournemouth Univ., Poole, England, 2019.
[33] A. Belhadi, Y. Djenouri, J. C. Lin, and A. Cano, "A data-driven approach for Twitter hashtag recommendation," *IEEE Access*, vol. 8, pp. 79182–79191, 2020.
[34] R. Harakawa, S. Takimura, T. Ogawa, M. Haseyama, and M. Iwahashi, "Consensus clustering of tweet networks via semantic and sentiment similarity estimation," *IEEE Access*, vol. 7, pp. 116207–116217, 2019.
[35] Y. He, C. Wang, and C. Jiang, "Mining coherent topics with pre-learned interest knowledge in Twitter," *IEEE Access*, vol. 5, pp. 10515–10525, 2017.
[36] R. R. de Mendonça, D. F. de Brito, F. de Franco Rosa, J. C. dos Reis, and R. Bonacin, "A framework for detecting intentions of criminal acts in social media: A case study on Twitter," *Information*, vol. 11, no. 3, pp. 154.1–154.40, Mar. 2020.
[37] A. Rosyiq, A. R. Hayah, A. N. Hidayanto, M. Naisuty, A. Suhanto, and N. F. Avuning Budi, "Information extraction from Twitter using DBpedia ontology: Indonesia tourism places," in *Proc. Int. Conf. Informat., Multimedia, Cyber Inf. Syst. (ICIMCIS)*, Oct. 2019, pp. 91–96.
[38] S. Sharma and A. Jain, "Cyber social media analytics and issues: A pragmatic approach for Twitter sentiment analysis," in *Advances in Computer Communication and Computational Sciences*. Singapore: Springer, 2019, pp. 473–484.
[39] D. Yu, D. Xu, D. Wang, and Z. Ni, "Hierarchical topic modeling of Twitter data for online analytical processing," *IEEE Access*, vol. 7, pp. 12373–12385, 2019.
[40] A. Arora, S. Bansal, C. Kandpal, R. Aswani, and Y. Dwivedi, "Measuring social media influencer index—Insights from Facebook, Twitter and Instagram," *J. Retailing Consum. Services*, vol. 49, pp. 86–101, Jul. 2019.
[41] P. Penas, R. del Hoyo, J. Vea-Murguia, C. Gonzalez, and S. Mayo, "Collective knowledge ontology user profiling for Twitter—Automatic user profiling," in *Proc. IEEE/WIC/ACM Int. Joint Conf. Intell. (WI) Intell. Agent Technol. (IAT)*, Nov. 2013, pp. 439–444.
[42] É. Deparis, M.-H. Abel, and J. Mattioli, "Modeling a social collaborative platform with standard ontologies," in *Proc. 7th Int. Conf. Signal Image Technol. Internet-Based Syst.*, Nov. 2011, pp. 167–173.
[43] Y. Zou, X. Yao, Z. Chen, and M. Zhao, "Verifiable keyword-based semantic similarity search on social data outsourcing," *IEEE Access*, vol. 7, pp. 5616–5625, 2019.
[44] E. Sediyono, Suhartono, and C. Nivak, "Measuring the performance of ontological based information retrieval from a social media," in *Proc. Eur. Model. Symp.*, Oct. 2014, pp. 354–359.
[45] O.-J. Lee, H. L. Nguyen, J. E. Jung, T.-W. Um, and H.-W. Lee, "Towards ontological approach on trust-aware ambient services," *IEEE Access*, vol. 5, pp. 1589–1599, 2017.
[46] P. F. Lazarsfeld and P. K. Merton, "Friendship as a social process: A substantive and methodological analysis," *Freedom Control Mod. Soc.*, vol. 18, no. 1, pp. 18–66, 1954.
[47] G. Weimann, "The Influentals: Back to the concept of opinion leaders?" *Public Opinion Quart.*, vol. 55, no. 2, pp. 267–279, Summer 1991.
[48] D. J. Watts and P. S. Dodds, "Influentials, networks, and public opinion formation," *J. Consum. Res.*, vol. 34, no. 4, pp. 441–458, Dec. 2007.
[49] J. Coleman, E. Katz, and H. Menzel, "The diffusion of an innovation among physicians," *Sociometry*, vol. 20, no. 4, pp. 253–270, 1957.
[50] D. M. J. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S. A. Sloman, C. R. Sunstein, E. A. Thorson, D. J. Watts, and J. L. Zittrain, "The science of fake news," *Science*, vol. 359, no. 6380, pp. 1094–1096, 2018.
[51] S. Chen, X. Yin, Q. Cao, Q. Li, and H. Long, "Targeted influence maximization based on cloud computing over big data in social networks," *IEEE Access*, vol. 8, pp. 45512–45522, 2020.
[52] A. Alturki, N. Alshwihi, and A. Algarni, "Factors influencing players' susceptibility to social engineering in social gaming networks," *IEEE Access*, vol. 8, pp. 97383–97391, 2020.
[53] U. Kursuncu, M. Gaur, C. Castillo, A. Alambo, K. Thirunarayan, V. Shalin, D. Achilov, I. B. Arpinar, and A. Sheth, "Modeling islamist extremist communications on social media using contextual dimensions: Religion, ideology, and hate," *ACM Hum.-Comput. Interact.*, vol. 3, no. CSCW, pp. 151.1–151.22, Nov. 2019.
[54] P. Brangetto and M. A. Veenendaal, "Influence cyber operations: The use of cyberattacks in support of influence operations," in *Proc. 8th Int. Conf. Cyber. Conflict*, May 2016, pp. 113–126.
[55] S. W. Hart and M. C. Klink, "1st troll battalion: Influencing military and strategic operations through cyber-personas," in *Proc. Int. Conf. Cyber Conflict*, Nov. 2017, pp. 97–104.
[56] D. Herrick, "The social side of 'cyber power'? Social media and cyber operations," in *Proc. 8th Int. Conf. Cyber Conflict (CyCon)*, May 2016, pp. 99–111.
[57] T. Zeitzoff, "How social media is changing conflict," *J. Conflict Resolution*, vol. 61, no. 9, pp. 1970–1991, Oct. 2017.
[58] D. U. Patton, J. Lane, P. Leonard, J. Macbeth, and J. R. S. Lee, "Gang violence on the digital street: Case study of a south side chicago gang member's Twitter communication," *New Media Soc.*, vol. 19, no. 7, pp. 1000–1018, Jul. 2017.
[59] E. Anderson, *Code Streets—Decency, Violence, Moral Life Inner City*. New York, NY, USA: W. W. Norton & Company, 1999.
[60] S. H. Decker and D. C. Pyrooz, "Gangs: Another form of organized crime?" in *Oxford Handbook of Organized Crime*, L. Paoli, Ed. London, U.K.: Oxford Univ. Press, 2013, pp. 1–9.
[61] M. L. Storrod and J. A. Densley, "'Going viral'and 'Going country': The expressive and instrumental activities of street gangs on social media," *J. Youth Stud.*, vol. 20, no. 6, pp. 677–696, Jul. 2017.
[62] A. V. Papachristos, D. M. Hureau, and A. A. Braga, "The corner and the crew: The influence of geography and social networks on gang violence," *Amer. Sociol. Rev.*, vol. 78, no. 3, pp. 417–447, Jun. 2013.
[63] L. Maheshwari. (Apr. 2019), *India's February 2019 Strike Pakistani Territory: A Jus ad Bellum Analysis*. [Online]. Available: https://www.lawfareblog.com
[64] S. A. Philip. (Feb. 2020). *Inside Story of Attack on Balakot—From IAF Officer Who Planned and Executed it*. [Online]. Available: https://theprint.in/
[65] (Jun. 2019). *IAF Pilots in Balakot Air Strikes Say Op Was Over in 90 Seconds Families Didn't Even Know*. [Online]. Available: https://theprint.in/
[66] V. Kaura, "India's Pakistan policy: From 2016 'surgical strike' to 2019 Balakot airstrike," *Round Table*, vol. 109, no. 3, pp. 277–287, May 2020.
[67] B. Durr, "ISIS: The use of social media," Ph.D. dissertation, Dept. Cybersecur., Utica College, New York, NY, USA, 2016.
[68] K. Anderson. (2016). *Cubs of the Caliphate: The Systematic Recruitment, Training, and Use of Children in the Islamic State*. Accessed: May 3, 2018. [Online]. Available: www.ict.org.il/UserFiles/ICT-Cubs-of-the-Caliphate-Anderson.pdf
[69] S. H. B. O. Alkaff and R. Mahzam, "Islamic state after the fall of Mosul and Raqqa," *Counter Terrorist Trends Analyses*, vol. 10, no. 1, pp. 57–61, 2018.
[70] I. Awan, "Cyber-extremism: ISIS and the power of social media," *Society*, vol. 54, no. 2, pp. 138–149, Apr. 2017.
[71] A. Perešin, "Fatal attraction: Western muslimas and ISIS," *Perspect. Terrorism*, vol. 9, no. 3, pp. 21–38, Jun. 2015.
[72] A. Speckhard, *Bride of ISIS*. Advances Press LLC, 2015. Accessed: Dec. 29, 2020. [Online]. Available: https://www.advancespress.com
[73] P. Barberá, J. T. Jost, J. Nagler, J. A. Tucker, and R. Bonneau, "Tweeting from left to right: Is online political communication more than an echo chamber?" *Psychol. Sci.*, vol. 26, no. 10, pp. 1531–1542, Oct. 2015.
[74] G. C. Tracker. (Jan. 2014). *Ukraine Crisis: Timeline*. [Online]. Available: https://www.bbc.com/news/world-middle-east-26248275
[75] N. MacFarquhar, "Inside the Russian troll factory: Zombies and a breakneck pace," *New York Times*, Feb. 18, 2018. Accessed: Dec. 31, 2020. [Online]. Available: https://www.nytimes.com/2018/02/18/world/europe/russia-troll-factory.html
[76] M. Galeotti. *The 'Gerasimov Doctrine' and Russian Non-Linear War*. Accessed: Dec. 29, 2020. [Online]. Available: https://inmoscowsshadows.wordpress.com/2014/07/06/the-gerasimov-doctrine-and-russian-non-linear-war
[77] I. J. Strudwicke and W. J. Grant, "#JunkScience: Investigating pseudoscience disinformation in the Russian Internet Research Agency tweets," *Public Understand. Sci.*, vol. 29, no. 5, pp. 459–472, 2020.
[78] U. A. Mejias and N. E. Vokuev, "Disinformation and the media: The case of Russia and Ukraine," *Media, Culture Soc.*, vol. 39, no. 7, pp. 1027–1042, Oct. 2017.
[79] A. Garmazhapova, "Where do trolls live. and who feeds them, Novaya Gazeta," Tech. Rep., 2013.

[80] P. R. Gregory, "Putin's new weapon in the Ukraine propaganda war: Internet trolls, Forbe," *Forbes*, Dec. 9, 2014. [Online]. Available: https://www.forbes.com/sites/paulroderickgregory/2014/12/09/putins-new-weapon-in-the-ukraine-propaganda-war-internet-trolls
[81] S. Shane, "The fake Americans Russia created to influence the election," *New York Times*, Sep. 7, 2017. Accessed: Dec. 31, 2020. [Online]. Available: https://www.nytimes.com/2017/09/07/us/politics/russia-facebook-twitter-election.html
[82] M. El Asikri, J. Laassiri, S.-D. Krit, and H. Chaib, "Contribution to ontologies building using the semantic Web and Web mining," in *Proc. Int. Conf. Eng. MIS (ICEMIS)*, Sep. 2016, pp. 1–5.
[83] C. Bonacchi and M. Krzyzanska, "Digital heritage research re-theorised: Ontologies and epistemologies in a world of big data," *Int. J. Heritage Stud.*, vol. 25, no. 12, pp. 1235–1247, Dec. 2019.
[84] G. Kuang and Y. Du, "Building semantic business Web services based on ontology," in *Proc. Int. Conf. New Trends Inf. Sci. Service Sci.*, 2010, pp. 634–637.
[85] S. N. Han, G. M. Lee, and N. Crespi, "Towards automated service composition using policy ontology in building automation system," in *Proc. IEEE 9th Int. Conf. Services Comput.*, Jun. 2012, pp. 685–686.
[86] S. Bischof, A. Karapantelakis, C. Nechifor, A. P. Sheth, A. Mileo, and P. Barnaghi. (2014). *Semantic Modelling of Smart City Data*. [Online]. Available: https://corescholar.libraries.wright.edu/knoesis/572
[87] B. Halawi, A. Mourad, H. Otrok, and E. Damiani, "Few are as good as many: An ontology-based tweet spam detection approach," *IEEE Access*, vol. 6, pp. 63890–63904, 2018.
[88] S. Bromander, A. Jøsang, and M. Eian, "Semantic cyberthreat modelling," in *Proc. STIDS*, 2016, pp. 74–78.
[89] J. Breuker, A. Valente, and R. Winkels, "Legal ontologies in knowledge engineering and information management," *Artif. Intell. Law*, vol. 12, no. 4, pp. 241–277, Dec. 2004.
[90] R. Hoekstra, J. Breuker, M. Di Bello, and A. Boer, "The LKIF core ontology of basic legal concepts," in *Proc. Workshop Legal Ontol. Artif. Intell. Techn.*, Jan. 2007, pp. 43–63.
[91] J. Campbell, "Ex machina: Technological disruption and the future of artificial intelligence in legal writing," Sturm College Law, Univ. Denver, Denver, CO, USA, Legal Stud. Res. Paper 20-4, Feb. 2020.
[92] B. Brodaric, T. Hahmann, and M. Gruninger, "Water features and their parts," *Appl. Ontology*, vol. 14, no. 1, pp. 1–42, Feb. 2019.
[93] C. Masolo, A. Botti Benevides, and D. Porello, "The interplay between models and observations," *Appl. Ontol.*, vol. 13, no. 1, pp. 41–71, Feb. 2018.
[94] *Resource Description Framework (RDF)*. Accessed: Dec. 29, 2020. [Online]. Available: https://www.w3.org/RDF/
[95] V. Bindu and C. Thomas, "Knowledge base representation of emails using ontology for spam filtering," in *Advances in Artificial Intelligence and Data Engineering*, N. N. Chiplunkar and T. Fukao, Eds. Singapore: Springer, 2021, pp. 723–735.
[96] K. Angele, D. Fensel, E. Huaman, E. Kärle, O. Panasiuk, U. Şimşek, I. Toma, and A. Wahler, "Semantic Web empowered E-tourism," in *Handbook of e-Tourism*, Z. Xiang, M. Fuchs, U. Gretzel, and W. Höpken, Eds. Cham, Switzerland: Springer, 2020, pp. 1–46, doi: 10.1007/978-3-030-05324-6_22-1.
[97] *Good Ontologies*. Accessed: Dec. 29, 2020. [Online]. Available: https://www.w3.org/wiki/Good_Ontologies
[98] *FOAF Vocabulary Specification 0.99, Namespace Document—Paddington Edition*. Accessed: Dec. 29, 2020. [Online]. Available: http://xmlns.com/foaf/spec/
[99] *Semantically Interlinked Online Communities (SIOC) Project*. Accessed: Dec. 29, 2020. [Online]. Available: http://sioc-project.org/
[100] S. L. Weibel and T. Koch, "The Dublin Core Metadata Initiative: Mission, current activities, and future directions," *D-Lib Mag.*, vol. 6, no. 12, Dec. 2000. [Online]. Available: http://www.dlib.org/dlib/december00/weibel/12weibel.html
[101] E. Olca and O. Can, "Extending FOAF and relationship ontologies with consent ontology," in *Proc. 3rd Int. Conf. Comput. Sci. Eng. (UBMK)*, Sep. 2018, pp. 542–546.
[102] E. Kalemi and E. Martiri, "FOAF-academic ontology: A vocabulary for the academic community," in *Proc. 3rd Int. Conf. Intell. Netw. Collaborative Syst.*, Nov. 2011, pp. 440–445.
[103] A. E. Kassiri and F.-Z. Belouadha, "A FOAF ontology extension to meet online social networks presentation and analysis," in *Proc. IEEE Int. Conf. Power, Control, Signals Instrum. Eng. (ICPCSI)*, Sep. 2017, pp. 3056–3061.
[104] X. Nie, W. Zhang, Y. Zhang, and D. Yu, "Method to predict bursty hot events on Twitter based on user relationship network," *IEEE Access*, vol. 8, pp. 44031–44040, 2020.
[105] M. Fang, Y. Li, Y. Hu, S. Mao, and P. Shi, "A unified semantic model for cross-media events analysis in online social networks," *IEEE Access*, vol. 7, pp. 32166–32182, 2019.
[106] A. Dhiman and D. Toshniwal, "An approximate model for event detection from Twitter data," *IEEE Access*, vol. 8, pp. 122168–122184, 2020.
[107] A. Agarwal and D. Toshniwal, "Face off: Travel habits, road conditions and traffic city characteristics bared using Twitter," *IEEE Access*, vol. 7, pp. 66536–66552, 2019.
[108] G. Razis and I. Anagnostopoulos, "Semantifying Twitter: The influence tracker ontology," in *Proc. 9th Int. Workshop Semantic Social Media Adaptation Personalization*, Nov. 2014, pp. 98–103.
[109] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The rise of social bots," *Commun. ACM*, vol. 59, no. 7, pp. 96–104, Jun. 2016.
[110] S. Kumar, J. Cheng, J. Leskovec, and V. S. Subrahmanian, "An army of me: Sockpuppets in online discussion communities," in *Proc. 26th Int. Conf. World Wide Web*, Apr. 2017, pp. 857–866.
[111] *Twitter Hashtag Tracking*. Accessed: Dec. 29, 2020. [Online]. Available: https://www.trackmyhashtag.com/
[112] *Twitter Stream Archive Team: The Twitter Stream Grab*. Accessed: Dec. 29, 2020. [Online]. Available: https://archive.org/details/twitterstream?and[]=year%3A%222014%22
[113] *TweetBinder (TB)*. Accessed: Dec. 29, 2020. [Online]. Available: https://www.tweetbinder.com/
[114] *Twitter Developers*. Accessed: Dec. 29, 2020. [Online]. Available: https://developer.twitter.com/en
[115] *Encyclopedia Britannica*. (Aug. 2014). *Ukraine Crisis*. [Online]. Available: https://www.britannica.com/topic/Ukraine-crisis
[116] M. Fisher. (Sep. 2014). *Everything You Need to Know About the Ukraine Crisis*. [Online]. Available: https://www.vox.com/2014/9/3/18088560/ukraine-everything-you-need-to-know
[117] *Protege: A Free, Open-Source Ontology Editor and Framework for Building Intelligent Systems*. Accessed: Dec. 29, 2020. [Online]. Available: https://protege.stanford.edu/
[118] *WebVOWL*. Accessed: Dec. 29, 2020. [Online]. Available: http://vowl.visualdataweb.org/webvowl/old/
[119] A. Bose, V. Behzadan, C. Aguirre, and W. H. Hsu, "A novel approach for detection and ranking of trendy and emerging cyber threat events in Twitter streams," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining*, Aug. 2019, pp. 871–878.
[120] M. C. K. Michel and M. C. King, "Cyber influence of human behavior: Personal and national security, privacy, and fraud awareness to prevent harm," in *Proc. IEEE Int. Symp. Technol. Soc. (ISTAS)*, Nov. 2019, pp. 1–7.
[121] R. Riesco and V. A. Villagrá, "Leveraging cyber threat intelligence for a dynamic risk framework," *Int. J. Inf. Secur.*, vol. 18, no. 6, pp. 715–739, Dec. 2019.
[122] K. Simran, P. Balakrishna, R. Vinayakumar, and K. Soman, "Deep learning approach for enhanced cyber threat indicators in Twitter stream," in *Proc. Int. Symp. Secur. Comput. Commun.*, vol. 2019, pp. 135–145.
[123] R. Syed, "Cybersecurity vulnerability management: A conceptual ontology and cyber intelligence alert system," *Inf. Manage.*, vol. 57, no. 6, Sep. 2020, Art. no. 103334.



NATHAN JOHNSON (Graduate Student Member, IEEE) has been a member of the Royal Australian Army, since 2001. He received the bachelor's degree in physics and economics from the University of New South Wales, Australia, in 2003, and the master's degree in military electrical system engineering from Cranfield University, U.K., in 2013. He has extensive experience in military RF and electrical systems engineering as well as project management. His Ph.D. research focuses on influence modeling of social media using ontological approaches and its application to conflicts. In 2017, he received the Chief of Army's Scholarship, the General Sir John Monash Foundation Scholarship, the American Australian Association Scholarship, and the President of Arizona State University Scholarship to complete his Ph.D. studies in cyber and electronic warfare at Arizona State University.

BENJAMIN TURNBULL (Member, IEEE) is currently a Senior Lecturer with the University of New South Wales at the Australian Defence Force Academy, Canberra. He was working in digital forensics, network security, and simulation for 17 years. His previous work as a defense research scientist saw him develop and deploy new technologies to multiple clients globally. His research interests include the intersection of cybersecurity, simulation, scenario-based learning, and the security of heterogeneous devices and future networks. He is a Certified Information Systems Security Professional (CISSP).

THOMAS MAHER received the bachelor's degree in computing and cybersecurity (BCCS) from the University of New South Wales, in 2018. He is currently studying a BCCS-Honours. He has been a member of the Australian Army, since 2016. In 2019, he received the Chief of Army's Honours Scholarship. His research interest includes the weaponization of social media, specifically, the promulgation of misinformation on online social media platforms.

MARTIN REISSLEIN (Fellow, IEEE) received the Ph.D. degree in systems engineering from the University of Pennsylvania, Philadelphia, in 1998. He is currently a Professor with the School of Electrical, Computer, and Energy Engineering, Arizona State University (ASU), Tempe. He received the IEEE Communications Society Best Tutorial Paper Award in 2008, the Friedrich Wilhelm Bessel Research Award from the Alexander von Humboldt Foundation in 2015, and the DRESDEN Senior Fellowship in 2016 and 2019. He serves as an Associate Editor for the IEEE TRANSACTIONS ON MOBILE COMPUTING, the IEEE TRANSACTIONS ON EDUCATION, IEEE ACCESS, and *Computer Networks*. He is also an Associate Editor-in-Chief of the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, and a Co-Editor-in-Chief of *Optical Switching and Networking*. He has chaired the Steering Committee of the IEEE TRANSACTIONS ON MULTIMEDIA from 2017 to 2019.