Automating Social Interactions: A Network Perspective of the Role of Bots on Social Media

Abdullah Alrhmoun

Supervisor: János Kertész

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Network Science

Department of Network and Data Science
Central European University
Vienna, Austria

2023

Abdullah Alrhmoun
Automating Social Interactions: A Network Perspective of the Role of Bots on Social Media

Researcher Declaration

I, Abdullah Alrhmoun, certify that I am the author of the work **Automating Social Interactions: A Network Perspective of the Role of Bots on Social Media**. I certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, the contributions of others. The thesis contains no material accepted for any other degrees in any other institutions. The copyright of this work rests with its author. Quotation from it is permitted, provided that full acknowledgement is made. This work may not be reproduced without my prior written consent.

Statement of inclusion of joint work

I confirm that Chapter 3 is based on the paper titled, **"Emergent Local Structures in an Ecosystem of Social Bots and Humans on Twitter,"** which is accepted for publication in EPJ Data Science. Following my idea, both Dr. Kertész and I designed the social bot experiment on Twitter to investigate the effectiveness of various bot strategies from a network standpoint. I wrote the code, executed the experiment and conducted the data analysis. Both contributors were actively involved in interpreting the observations and writing the paper. Dr. Kertész expresses his agreement to this statement by providing his signatures below.

I hereby verify that Chapter 5 is based on the paper titled **"Automating Terror: The Role and Impact of Telegram Bots in the Islamic State's Online Ecosystem,"** published in Terrorism and Political Violence [1]. This work was a collaborative effort with Dr. Charlie Winter and Dr. János Kertész. The idea of exploring the role of bots within the ISIS Telegram ecosystem originated from me. Both Dr. Kertész and I established the methodologies employed, while I conducted the data collection and analysis. Dr. Winter provided expertise on terrorist networks. All authors participated in the paper's writing. Dr. Winter and Dr. Kertész express their agreement to this statement by providing their signatures below.

Endorsing statement of joint work:

PhD Candidate

Date: 07 December 2023

Supervisor: Dr. János Kertész

Date: 23 August, 2023

Collaborator: Dr. Charlie Winter

Date: 22 August 2023

Abstract

This thesis comprises three projects examining the emerged network structures due to bot activities on Twitter and Telegram. Specifically, we delve into these structures' origins, impacts, and roles in a broader network context. Our analysis uses data-driven, experimental, and network-based concepts to identify varying network structures across contexts and user groups, such as core-periphery structures, local structures (i.e., motifs), community formations, and rich-club organization.

The first project is an experimental study on Twitter where we deployed six bots in pairs of two, each pair assigned different strategies: a trend-targeting strategy (TTS), a keywords-targeting strategy (KTS), and a user-targeting strategy (UTS). We then assessed interaction patterns, including targeting users, message dissemination, relationship propagation, and engagement. While TTS was the most effective in obtaining human feedback, it displayed the least diverse local structure patterns. In contrast, UTS was the least effective but activated a broader spectrum of complex, local structures. Furthermore, content-related strategies (TTS and KTS) had a significant overlap in terms of local structures activated. Notably, the KTS shows promise in bridging the benefits of content-focused and user-focused approaches by targeting content that resonates with particular users. This strategy has shown the ability to create engaging patterns while effectively disseminating content, which is vital to success on social media platforms.

The second project is concerned with the network structures of three extremist groups on Telegram: the Islamic State of Iraq and Syria (ISIS), far-right groups (FR), and pro-Russian actors tied to the conflict in Ukraine (PR). We expected different authority structures: ISIS lacked a centralized authority, the pro-Russian actors displayed a pronounced central authority, and the far-right group combined decentralized and centralized elements. Network metrics-based analysis supported the expectations that the three extremist groups follow different organizational principles and platform usage purposes, which results in different structures. Our application of the 'rich-club' detection method disclosed variations in the nodes' roles and positions within these networks. Bots are present in the rich-club of the ISIS and FR networks, while the PR network's rich-club is exclusively composed of human users. Given their automatic nature, bots could increase the pace of information spread within the network, but in a very centralized network, there is no need for such an augmentation.

The third project identified two primary communities of bots and channels/groups associated with ISIS on Telegram. These basic bots, notwithstanding their simplicity, remain pivotal in sustaining the online presence of the Islamic State, especially in the light of Telegram's intensive countermeasures. Furthermore, the core of both communities mainly consists of bots, with their peripheries comprising a mixture of channels and groups. A func-

tional explanation is that the core-periphery structures have emerged because of constant activity from the core in maintaining content distribution efforts and chat moderation, which was conducted in the periphery.

إلى أمّي وأبي: قُطبَي عَالَمي وما بَينَهما

إلى زَوجَتي وطِفلَتي: حَبة القلبِ وريحَانته

إلى أخوَتي وأخواتِي: بكم تُختَصر الحياة

إلى الثَورَة السُوريّة وشهدائِها: أنتُمِ الذَاكرة التي تأبَى النسيان والخَيال الذي يصنع المستقبل والهُويّة التي تَعلو على كلِّ شيء.

عبدُ اللّه

"The battle for your reality begins in the fields of digital interaction"

Douglas Rushkoff

Acknowledgements

I am very thankful to the amazing people who have helped bring this thesis to fruition. First and foremost, I want to thank my advisor, János Kertész, for his guidance, support, patience, and belief in me. János is a great scientist and a kind person.

I also want to thank my colleagues Teodoro Criscione, Maria Gabriela Juncosa Calahorrano, Tamer Khraisha, and Felipe Vaca for their friendship and encouragement during my PhD program, along with the entire Department of Network and Data Science at CEU for the positive and supportive learning and research environment they have created.

I'm grateful to my close friends who have helped and supported me in many ways over the years. I want to thank Matthew King for his interesting ideas and helpful writing tips; Ashraf Alkilany for his coding support; and Tam Preston for her design help. I also want to give special thanks to Amy Braman, who has carefully edited and reviewed this thesis, and supported me during this journey in many ways.

I also extend my gratitude towards the Accelnet-Multinet program and its coordinators, specifically Ann McCranie and Tara Holbrook. Their diligent efforts made my visit to the United States feasible, a truly appreciable feat considering the struggles I navigate through due to the unfair realities of our current world toward those of Syrian descent.

Furthermore, I am especially grateful to Filippo Menczer and Alessandro Flammini, two brilliant scientists at Indiana University whose insights and comments significantly influenced the fourth chapter of my thesis.

Most of all, I am deeply thankful to my wife, Salwa, who has been my rock, my friend, and my partner in every way. Her support, understanding, encouragement, and love have given me strength throughout this journey, especially during the past year while she was pregnant with our baby. She inspires me, and I dedicate this thesis to her with love.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

> *"Robots will neither be common nor very good in 2014, but they will be in existence."*
>
> Isaac Asimov

Tracing back to the early days of the internet, and even before, in networked systems like Usenet, social bots have been an integral part of the digital world [2]. Simultaneously, artificial intelligence research communities have been striving to create intelligent bots capable of realistically communicating with humans. An early, notable example is ELIZA, a chatbot developed in the 1960s [3]. The prominence of social bots has grown substantially since then, with the emergence of social media platforms. At the same time, data growth and AI technology advancements have enabled them to become more sophisticated and convincing [4, 5]. Today, these bots serve a variety of purposes throughout social media platforms, such as aiding in content distribution [6], boosting user popularity [7], and performing a wide array of other tasks [8, 9].

Over time, internet bots have evolved from basic tools to crucial components in an ever-changing ecosystem that increasingly depends on automated accounts. However, bots operating within social media platforms such as Twitter, Facebook, and Reddit are designed to appear as authentic members of the online community. They mimic the behavior of ordinary users by engaging with human users through activities like posting, liking, sharing, and commenting [10, 11, 12, 13, 14]. Furthermore, social bots interact with one another in a coordinated manner to influence popularity metrics and generate a sense of consensus or support for specific ideas [15, 16, 17].

The internet provides social bots with a vast playground to assume various roles, such as political activism, gaming competition, and digital assistance. Within this diverse digital space, the interactions between social bots and humans in online social networks (OSNs) have emerged as a pivotal component of the expanding field of human-machine interaction [18]. Social bots exist in online environments alongside humans, where they engage, compete, cooperate, and collaborate. Their relationships with humans impact various aspects of social media platforms, including community formation and the spreading of information.

The negative public perception of social bots is attributed mainly to their association with the 2016 US presidential election [19, 4]. Although bots can be harmful, they can also be beneficial, and this dual nature represents only a part of the social bots narrative. A broader view reveals that the negative aspects of bots originate not from inherent shortcomings but from a flawed social media landscape that promotes polarization, radicalization, and the decay of truth [20, 21, 22].

While social Bots are often linked to election manipulation, dissemination of inaccurate information ("fake news"), and other malicious activities [23, 24, 25], they display polarized behavior as a manifestation of human polarization and are exploited by malicious entities [26, 27]. However, their contribution to these issues could be substantially limited by adequate governance measures [28]. When developed responsibly, bots can positively serve humanity as investigative scientific tools and aid in tasks such as emergency management and locating as well as the scheduling of vaccine appointments. [29, 30].

Although bots can perform harmful actions, their negative assessment was premature and possibly unfair after the 2016 US elections. It goes without saying that the harm caused by malicious bots (or, rather, malicious people employing bots) should be minimized as much as possible, however, the abundance of bots in online communication is a fact of life, leading to an ecosystem in which humans and bots will mingle in both positive and negative interactions. Understanding how this ecosystem functions is not only a challenge for science but also an inevitable element of regulating the online world for the benefit of humanity.

While the impact, role, and identification of bots in social media have been of focal interest in the literature [25, 31, 32, 33, 34, 35, 11, 36], the way they interact with humans and the processes that make social bots effective are still unclear. This thesis proposes that the network structures formed by or co-evolved with the activity of these bots play an important role in their effectiveness. Social bots build communities, form core-periphery structures, trigger both simple and complex local structures (motifs) via human interactions, and participate in the rich-club organization of networks. This thesis is comprised of multiple projects to investigate bot-human ecosystems both for harmless and harmful bots from these points of view.

The first project was an experiment on Twitter by deploying social bots, pre-programmed with different strategies, to interact with humans. The fundamental premise of this experiment was to illustrate that social bots and humans together form an ecosystem within OSNs. Notably, the strategies employed determine the emergent properties of this ecosystem, which is characterized by fundamental building blocks (i.e., local structures) that illustrate patterns of interactions, such as propagation of content and bridging relationships.

The second project examined humans and bots within the channels/groups of three extremist organizations on Telegram: ISIS, far-right groups (FR), and pro-Russia actors related to the war in Ukraine (PR). This research suggests that the network of social bots and humans organizes itself in different ways depending on its history and goals as well as on interventions by the service provider. A decentralized structure emerges if the external pressures strong and the network's main role is the overall maintenance of the community of similarly thinking people while a more centralized structure is formed, which can enhance influence, consolidate control, and reduce inter-group competition costs. The roles of bots reflect these differences in their network positions and activities.

The third project examined bot-only interactions within ISIS channels/groups on Telegram. Telegram's API documentation provides comprehensive details about the bots' functionalities. Users employ these bots for diverse tasks, including content dissemination, group management, and service provision. ISIS embraced Telegram early on, and the project's objective was to understand the role and impact of bots within this terrorist organization. Two communities of Islamic State supporters were identified with one community focusing on facilitating discussion and exchange, and the other one augmenting content distribution efforts and the bots were employed according to the role of these communities.

My research is based on two kinds of data. First, I collected and analyzed data from Telegram, the increasingly popular online messenger service with Russian roots, which is a preferred communication space for extremist groups and which allows or even encourages the use of bots. Second, I developed bots and published them on Twitter to generate and analyze data about the bot-human interaction network.

This thesis is structured in the following manner:

> Chapter 2: I provide a comprehensive overview of the various attempts to define social bots, recounting their history, tracing their origins back to Usenet and Internet Relay Chat (IRC), and charting their progression from basic chatbots to sophisticated social bots on modern platforms such as Twitter and Reddit. I further explore the multifaceted roles of social bots in social media, discussing their positive and negative implications and the influence of human intentions and inputs on their functionality. I highlight the role of bots in social media, particularly Telegram and Twitter, as the thesis focuses on them. In addition, given the significance of political bots in the historical development and impact of social bots, I dedicate a section to examining their emergence. Moreover, I present a literature review on the use of bots for the propagation of terrorism and extremism, as well as their deployment in conflict situations. At the same time, I introduce the extremist and terrorist networks where such bots operate. Finally, I highlight the contributions of network science in facilitating a deeper understanding of the roles and impacts of social bots in various contexts.
>
> Chapter 3: I present the findings of an experiment conducted on Twitter, which aimed to investigate the local structures (motifs) present in networks corresponding to various bot strategies. I outline the experimental design, the setting of the social bots, and the strategies employed to facilitate interaction with human users and among the bots themselves. Furthermore, I present the findings of the emergent structural and colored motifs, where the latter distinguishes between bots and humans and the former does not. Finally, I summarize their implications for the strategies and the broader social media context.
>
> Chapter 4: This chapter explores the networks of three extremist groups on Telegram: ISIS, far-right communities, and pro-Russia actors in Ukraine, from two angles: 1) statistical analyses highlighting distinct usage patterns and strategies and 2) the underlying network structure, which reveals their information operations on the platform. Additionally, the chapter highlights the significance of bots within these networks.
>
> Chapter 5: I examine the roles of bots in the Islamic State's Telegram network. It covers the data collection process, detection of bot communities, their activities, and the core-periphery structure of the network. The conclusion interprets the emerging structures in ISIS's online ecosystem in relation to the real-world shifts of ISIS as a

political movement in recent years.

Chapter 6: In this chapter, I summarize the key results and contributions in chapters 3, 4, and 5. Moreover, I suggest potential avenues for future research and conclude the thesis.

Chapter 2

Literature Review

> *"We shape our tools and thereafter our tools shape us"*
>
> John M. Culkin

With technology constantly improving and changing our online and offline lives, the widespread use of social bots offers both possibilities and difficulties. As these bots become more prevalent, the consequences of using them, the networks that facilitate their activities, the technologies that help them function, and the ethical issues related to their use in the digital world still need to be fully understood.

Social bots are automated programs, sometimes equipped with artificial intelligence, designed to interact with people on social networks [4, 37, 5]. These bots can be found on popular social media platforms like Facebook [38], Instagram [39], Twitter [40], Reddit [41], and Telegram [1], as well as in online communities like Wikipedia and Discord [42, 43].

The origins of social bots can be linked to their initial roles in content posting and removal in two discussion or chat systems: Usenet and Internet Relay Chat (IRC) [44, 45, 46]. In the IRC ecosystem, bots played dual roles as helpful assistants and harmful adversaries, reflecting the ongoing complexity of their nature in the digital world [47, 48]. Similarly, on platforms like Reddit, bots perform supportive and contentious functions [41].

Multi-user domains (MUDs) were once described as "The Africa of bots." [2] In these virtual worlds, social bots experienced significant growth, underlining their increasing role in shaping interactions and communities within online games. Here, they served as virtual residents [49]. In a notable example, a simulated virtual city reported that over half of its inhabitants were bots [2].

Social bots can have both positive and negative effects on such platforms. On one hand, they can help create engaging gaming experiences by controlling non-player characters (NPCs) that interact with human players and move stories forward [50]. On the other hand, they can negatively impact gameplay by cheating, affecting game balance, or enabling real-world trading and market manipulation [51, 52].

Recently, a type of social bot known as 'political bots' emerged with the intent to impact online political debates, designed to influence public opinion, spread false information, and polarize online communities [53]. For example, a study examining Russian Twitter activity between early 2014 and late 2015 found that more than half of its political tweets came from bots [54]. Moreover, political bots were extensively deployed during the 2016 US presidential election to spread misinformation and influence voters, resulting in a negative perception of bots in the eyes of the public [19].

Although social bots are perceived mainly as harmful, not all are designed for malign purposes. Many positive examples exhibit how they can be used for good [55, 8]. During the COVID-19 pandemic, for instance, The Washington Post built a chatbot to help citizens make better choices about retirement [56]. The New York Times launched a Slackbot to more directly connect readers with their newsroom during the 2016 US presidential election, which in turn encouraged audience involvement and community growth [57]. Even political campaigns are incorporating the use of bots. The Biden campaign, for instance, created a Facebook Messenger chatbot designed to motivate people to vote [58].

Wikipedia bots, while positive in intent, have sometimes resulted in unexpected 'negative effects.' While these bots were designed to create, improve, and maintain the encyclopedia's vast content, they also exhibit interactions that, at times, counteract each other's edits. This leads to a phenomenon called "sterile fights" where they undo each other's changes in lengthy, unproductive cycles [42, 59].

In this chapter, I will define social bots, explore their positive and negative effects, examine their origins, and discuss their roles in social media platforms, with a focus on Twitter and Telegram. Moreover, I will highlight the significant contributions of network science to understanding social bots and the ecosystems they create in interactions with humans[1].

2.1 Defining Social Bots

Social bots are software apps designed to operate within online social networks (OSNs) like Twitter and Facebook [60]. Derived from the word "robot," the term "bot" relates to automated software programs executing specific online tasks. These bots operate as computer programs written by a human developer, who usually maintains control, albeit sometimes only partially.

Programmed with diverse objectives, social bots engage users by liking or commenting on posts, initiating conversations, or suggesting content [60]. Even though they are software apps, they frequently pretend to be humans on social media platforms, complicating the differentiation between genuine users and automated accounts [62].

As bots interact with users in online environments, they adapt and evolve [63]. However, uncontrolled and unlimited interactions between bots and humans can be dangerous. Case in point: in 2016, Microsoft introduced Tay, a Twitter bot emulating a 19-year-old girl [64]. Despite Tay's capacity to post nearly 100,000 tweets in a single day, it faced suspension due to the dissemination of controversial and offensive content [65].

Social bots exhibit context-sensitive behaviors, continuously learning from their environments, which raises an anthropological question. The interactions between humans and machines are dynamic and situational, not static. To comprehend the social and cultural roles of bots, one must explore their potential for influence and agency. Winner (1980) suggests that tools can hold authority, exert control, and even discriminate and influence [66].

[1] For an overview of social bots, I recommend the following books [60, 23, 61].

Suchman (2006) agrees, emphasizing that developers/designers can not anticipate every outcome when crafting objects [67].

Though bots' impacts are typically intentional or pre-programmed, the mere act of information dissemination can inadvertently induce change [60]. This unpredictability stems from bots functioning within complex ecosystems interlaced with diverse processes, forces, and types of users [40]. Adding to this complexity is the potential human emotional bond with bots. For example, Weizenbaum's creation of ELIZA showcased the degree to which individuals attribute human traits to machines [60].

2.1.1 The dual nature of social bots

The design of social bots determines their effects. Some bots are used for harmful purposes, such as spreading false information, manipulating public opinion, or participating in other negative activities. Others play positive roles, distributing helpful information and connecting people with similar interests.

There are political bots that make use of algorithms to manipulate the spread of information in order to achieve their political objectives[60]. These bots promote certain messages while suppressing opposing views. They also monitor OSNs to gather data on public opinion and trends, which can be used to target individuals with content that can influence their voting choices or instill fear about a particular policy.

Conversely, social bots have the potential to positively impact society, promoting transparency and accountability. For instance, a Twitter bot launched to track Wikipedia edits from official government IPs worldwide to identify instances where government offices might be involved in public relations or propaganda activities [68]. Such transparency bots utilize social media platforms to engage in digital activism. They aim to unveil wrongdoings, promote transparent governance, and bolster political accountability [69]. Building on this foundation of fostering transparency, news bots, another type of social bot, come into play.

These news bots exhibit various forms and degrees of societal involvement [70]. Often, they are deployed to create and disseminate news content, offering a more cost-effective and efficient alternative to traditional newsmaking methods. Their potential to advance transparency lies in the timely and widespread distribution of information. However, it is essential to acknowledge that news bots, like all bots, inherently reflect the opinions and views of their developers. This means that the content they produce might carry the biases or viewpoints of those who programmed them, underscoring the importance of ensuring their ethical and unbiased design[60].

The platform Twitch serves as a prime example of where bots and human users coexist and collaborate positively within a community. In various Twitch communities, bots take on diverse roles that can be organized into four main categories [71]: [i] Moderation Function: Bots assist in moderating chat rooms by enforcing community rules, filtering out inappropriate content, and maintaining a positive user atmosphere. [ii] Entertainment Function: Bots add an extra layer of entertainment by involving users in interactive games, performances, or other engaging activities that complement the main content. [iii] Information Function: Bots deliver relevant information to users on topics such as gaming statistics, news updates, or other essential details. [iv] Social Function: Bots foster social interactions within the community, greeting newcomers, assisting members in forming connections, and suggesting related content or links to boost engagement.

2.2 The Role of Social Bots on Social Media

Social bots are widely perceived in a negative light. Many academic and public opinions agree that they disrupt the information environment of OSNs with false or misleading information, distract and confuse users, influence political election outcomes [72, 19, 4], spread financial disinformation [73], distort narratives associated with extreme events such as mass shootings [74], or promote terrorist activities [1]. However, it is essential to remember that some bot creators' malicious intentions have led to their negative impact on digital spaces [75].

Despite the negative reputation of social bots, their prevalence is on the rise. A recent study found that bots make up between 9 to 15% of Twitter accounts [76] (though this was a hefty dispute between Twitter and Elon Musk in their legal battle before the acquisition [77]). Similar estimates have suggested that bots account for up to 15% of Instagram accounts and that 10% of Facebook's accounts are fake [78]. While many of these bots are engaged in malicious activities, not all are. Some perform beneficial services, such as sending vaccine appointment reminders or providing critical information in emergency situations [30, 8].

One of the notable consequences of social bots on OSNs is the artificial inflation of popularity for celebrities and politicians [79]. Numerous companies offer services that sell fake online followers for a nominal fee, allowing individuals to boost their perceived influence and reach[60]. This practice can lead to a skewed perception of an individual's popularity and potentially distort public opinion or even manipulate the outcome of political elections [79, 80].

Moreover, there is a growing trend in the employment of social bots for discussions on health-related topics, which constitutes a dangerous domain for social bot interventions [81]. In a particular study, it was discovered that social bots, suspected of being involved in conversations regarding COVID-19, accounted for approximately 9.27% of the discourse on Twitter [82]. Furthermore, these social bots demonstrated greater negativity than their human counterparts when discussing COVID-19 in the United States. The study also revealed that social bots successfully incited bot-to-human anger transmission during exchanges concerning COVID-19 within the United States.

2.2.1 Twitter bots

A Twitter bot is a special type of bot designed to operate within the Twitter ecosystem. It is crafted to generate and express ideas, interacting with users through human-like language and other communicative methods. Prior to Elon Musk's acquisition, Twitter's API permitted extensive automation, enabling bots to send and receive tweets almost as seamlessly as human users.

The inception of bots on Twitter predates its official 2009 launch. In fact, the start of Twitter bots can be traced back to 2006, when Jack Dorsey, one of Twitter's co-founders, sent the platform's inaugural tweet [12]. Intriguingly, this message originated from a script on Dorsey's computer, bypassing conventional web interfaces or mobile applications.

Over time, the relevance and ubiquity of social bots on Twitter have surged. They've been employed for a diverse range of tasks, from curating content and disseminating news to providing customer service and propagating political agendas. Some have even been used to disseminate misinformation. Hence, these automated entities have interwoven themselves into the fabric of Twitter, perpetually shaping user interactions and the platform's dynamics.

Researchers have proposed multiple classification systems for Twitter bots. For instance, the Bot Summit 0.2 approach, originating from an informal assembly of bot developers, divided Twitter bots into two categories: independent and dependent [83]. Independent bots generate content autonomously, whereas dependent bots harness Twitter data to formulate their messages. Contrarily, Veale and Cook's methodology categorizes Twitter bots into three: feed, watcher, and interactor bots [84]. Feed bots predominantly post tweets, watcher bots identify and relay specific events, and interactor bots engage with users, adapting based on these interactions.

One noteworthy subtype of Twitter bots is the political bot. These bots have the capability to shape public discourse on major political matters, from elections and referendums to policy discussions [85, 86, 31, 87, 88, 89, 90, 91, 92]. Such bots have been used by politicians, political parties, and governments to amplify their influence on public sentiment. Post the 2016 US election and the UK's Brexit vote, the term "bot" has been tinged with negativity, frequently associated with deceptive practices and manipulation.

Bots on Twitter exhibit hyper-social behavior, generating a substantial proportion of retweets. To elaborate, 1.45% of bots were responsible for producing 4.5% of retweets and tended to retweet content originating from their opinion group [93]. Essentially, according to a recent study, the primary function of bots on the Twitter platform is to enhance the spread of viewpoints by retweeting and potentially delivering false or misleading information to accounts with many followers. On the other hand, these bots rarely act as the primary source of such misinformation [94]. For instance, although social bots contribute up to 15.4% of the content in climate change-related conversations on Twitter, their capacity to spark significant discussions seems limited, as shown by their minimal interaction with human users [95].

2.2.2 Telegram bots

Telegram has garnered attention for its appeal to extremist groups, partly because of its features, including the widespread use of bots. The platform has found favor among various groups spanning diverse regional and ideological backgrounds. Notable examples include ISIS [96], Al-Qaeda [97], and far-right communities [98]. While ISIS has been particularly recognized for its early adoption of Telegram, using it for tasks like recruitment [99], attack coordination [100], identity creation, and community support [101, 100], other groups have followed suit.

During the conflict between Russia and Ukraine, both sides have leveraged the capabilities of Telegram, with bots playing a pivotal role in disseminating and gathering information [102]. The Russians have employed a systematic approach, flooding Telegram with disinformation bots. They craft fake "war correspondents" and push narratives through Kremlin-friendly channels that give the appearance of impartial reporting. However, instead of impartiality, it has been a source of disinformation and propaganda. Conversely, the Ukrainians utilize Telegram bots more proactively and defensively. These bots serve as lines for civilians to relay critical information to the authorities. Civilians can report on-the-ground intelligence, such as movements of Russian troops and armored vehicles.

This system proved effective when a tip received through this method enabled Ukraine's Security Service to launch a successful attack on Russian vehicles outside Kyiv [103]. Similarly, Ukrainian authorities introduced a Telegram bot to document and report Russian war crimes in Ukraine [104]. Another bot, which Russia demands its removal, scrapes the platform for evidence of Russian service members being captured or killed in Ukraine [105].

2.3 The Emergence of Political Bots

The first instances of political bots were likely "cancelbots," which were used to enforce specific policies or ideologies by removing content deemed inappropriate or non-compliant in IRC [106]. The employment of "cancelbots" marked the beginning of bots influencing online discussions and user behavior [107]. Political bots existed as early as 1980 in systems like Usenet, with the appearance of Serdar Argic on Usenet. This bot was designed to search for the term "Turkey" and post denials of the Armenian Genocide, aiming to create doubt about Turkey's involvement.

Decades later, during the 2016 US presidential election, bots' use in political contexts significantly increased. Bots were crucial in spreading the unfounded "Pizzagate" conspiracy theory, which falsely linked Hillary Clinton to a pedophilia ring in a Washington DC pizza parlor [61].

Political bots have many forms, such as disinformation bots and harassment bots[60]. Disinformation bots were seen in cases like the Russian government's denial of its involvement in the assassination of opposition leader Boris Nemtsov in 2015 [108]. Harassment bots came to prominence during the 2013 Turkish pro-democracy Gezi Park protests, where independent journalists covering the demonstrations against President Erdogan faced aggressive harassment campaigns, illustrating how bots are used to target and intimidate individuals [109].

Political bots were also utilized in democratic countries to affect the democratic processes. Although they played a minor role, bots strategically contributed to the Twitter discussion, mainly supporting the "Leave" campaign during the UK referendum on EU membership. Hashtags related to the "Leave" stance significantly dominated the conversation, as different viewpoints employed varying levels of automation [110].

In recent years, political bots have become easier to access and more affordable, leading to increased use in various political situations[60]. Open-source codes for bot creation are readily available online. For instance, in 2016, over 4,000 repositories for deploying Twitter bots were hosted on GitHub, a popular code-sharing platform [111].

The political bots phenomenon is where technology and power cross. Langdon Winner contends that different technological structures, like computer systems (including algorithms and bots), power plants, and highways, can represent various types of power and authority [66]. The design and setup of these technologies can effectively solve problems in a specific community, influencing both social and political dynamics. These human-made systems often require or are highly compatible with certain political relationships. Being inherently political technologies, they can either reinforce existing power structures or create new ones. Winner ultimately suggests that technology is not neutral; it can significantly affect social connections and the distribution of political power.

2.3.1 Terrorist bots

Scholars from an array of disciplines have long tracked how online terrorist activities have developed in synergy with technological advancements and shifts in both the physical and information security environments. Collectively, this work has demonstrated that much of the time, developmental trends are intuitive and borne of an iterative process of bottom-up innovation [112, 113]. However, occasionally – and often when the shifts they result in are most impactful – there is evidence of top-down influence playing a role as well [114].

For decades, violent Islamist extremists have set out to adopt and exploit new technologies to facilitate their operations (both military/terroristic and recruitment-focused). As Torres-Soriano has pointed out, the late 1990s and early 2000s saw prominent organizations like the Global Islamic Media Front (GIMF) and al-Qa'ida transitioning from static websites to closed forums [115]. In the first half of the 2010s, after more than a decade of use, these forums were supplanted by "conventional" social media platforms like Twitter and Facebook [116]. Following effective and systematic targeted disruption from these mainstream tools, most jihadist outreach online is now confined to the partially encrypted broadcasting and chat platform Telegram, which, as of late, has become increasingly inhospitable to militancy-related activism and remains a more open and functional space than sites like Twitter [117].

Despite its more optimal security and functionality, Telegram does not possess a full monopoly on jihadist outreach, as numerous researchers have pointed out in recent years. Other similarly oriented and secured apps like WhatsApp, Element, and Hoop have also emerged as preferred platforms for sensitive communication between adherents of extremist Islamism [118]. Moreover, since 2018, in particular, static websites have become increasingly important spaces once again, especially in the context of propaganda archiving and distribution. And, as Winter, Sayed and Alrhmoun have observed, for more conventional state-based groups like Hamas, Hizbullah and the Afghan Taliban, Twitter has never been more important [119]. That being said, Telegram still remains a hegemonic presence for jihadist outreach activities online, activities that typically take one of two overlapping forms: propaganda production and distribution and group identity formation. As Wagemakers observes, jihadist organizations have long invested a significant amount of their time and energy in propaganda work [120, 121]. From the influence network of Abu Jandal al-Azdi, one of al-Qa'ida's (A.Q.) most important Internet ideologues in the late 2010s, to al-Shabab's sprawling covert media apparatus, online spaces have long been replete with examples of jihadist organizational outreach.

It is important to note that while the Islamic State is arguably the most prominent example of jihadist propagandizing to date, it is by no means the only militant Islamist group to have seen the value in resource-intensive strategic outreach. Hamas, Hizbullah, and the Afghan Taliban all preside over similarly sized (if not bigger) media networks, which, as Khatib notes, they each deploy – like the Islamic State – to shape the narrative landscape and attract followers, legitimize their actions, and intimidate their opponents [122, 123]. Moreover, alongside their official, organizational outreach infrastructures – whether in the context of Sunni or Shi'i militancy – a vast array of supporter-run media outlets and agencies operate, peddling their respective ideological lines and amplifying their messages – often, incidentally, with assistance from automated bots [124].

In the more tangible, operationally direct sphere of identity formation (of which active recruitment can be a downstream consequence), jihadist outreach is characterized by a combination of hierarchical design and organic, volunteer-led activism. The role of social media platforms – Twitter in particular – in social absorption and formal enlistment has been pivotal over the last decade, as indicated in numerous case study-led assessments focusing on foreign fighter networks in Syria (including several network analyses) [125, 126, 127]. These studies have shown that on-/offline recruitment networks, while largely organic, are often carefully groomed spaces populated by official operatives and unofficial advocates, with the latter serving as connectors or beacons that directly elicit engagement from curious onlookers, drawing them in before furnishing them with the information they need to physically sign up – i.e., who to talk to, where to fly to, how to evade being apprehended, and so on [128, 129]. While these online encounters are important, a measure of face-to-face interaction is usually also required to facilitate the process of joining any militant movement, something that accounts for the continuing prevalence of what Conway describes as real-world, social network-based recruitment patterns [130].

Recognizing the new centrality of Telegram to extremist networking online, particularly in the last four years, scholars have begun to turn away from Twitter to focus their attention on how these same community behaviors and activities pan out on other platforms like Telegram, prominent among them Clifford and Powell and, separately, Amarasingam, Maher, and Winter [131, 132, 117]. However, to date, there have been no quantitative or network analyses of extremist networks on this platform, presumably because there are significantly more technical obstacles when it comes to collecting Telegram data versus Twitter data. This leaves us with a partial understanding of how and why Telegram continues to be a preferred arena for supporters of groups like the Islamic State today – a gap in the knowledge that this chapter seeks to go some of the way towards remedying.

Moreover, while there have been many exploratory and speculative studies looking at the role of bots in terrorist networks, these have so far refrained from addressing the specific challenges they present on Telegram. Among the first to substantively trace the practical impact of bots in the context of terrorism was Berger, whose early analysis of the role of the Islamic State's "Dawn of the Glad Tidings" app on Twitter was something of a trailblazer [133]. Berger showed that, through this app, the Islamic State was able to post 40,000 synchronized tweets in a single day when Mosul fell to its forces in 2014. While not technically driven by bot activity, this effort was an early precursor to the group's broader experimentation with artificially augmented social media activities [134, 135]. Other analyses from the likes of Bondy in 2017 and Sultan in 2019 have been more forward-looking, speculating as to how bots might be deployed to support malign activities at both a state and non-state level over the course of the next decade. While useful, these and studies like them are predominantly hypothetical, drawing on anecdotal observations or small datasets. On that basis, in this chapter, we hope to both add to and elevate the existing knowledge base regarding terrorist usage(s) of bots.

2.4 Network Science Contributions to Social Bots Research

Network science methods provide valuable insights into social bots' organization, spread, and influence in OSNs. For example, network science techniques help to (I) monitor how information spreads through networks and the role bots play in this process, (II) find communities within networks where bots operate, and (III) assess the importance of bots in networks by employing methods such as centrality and hierarchy, among other tasks.

For example, a study employed network analysis to examine the spread of mis- and disinformation on Twitter in the events of the US presidential election in 2016 [136]. By studying the structure of diffusion networks, consisting of nodes (Twitter accounts) and links (retweets, quoted tweets, replies, and mentions), the researchers observed patterns in how fake news spreads on Twitter by analyzing 400 thousand claims. The study revealed that social bots played a significant role in spreading these claims, as bot-generated content comprised many retweets. Additionally, bots were found to target influential users by mentions and replies.

Social bots' impact is due to their ability to hold strategic positions within social networks, allowing their messages to spread more broadly [19]. This is especially important on platforms like Twitter, where harmful bots are common. It has been suggested that the most effective method to identify and remove harmful bots from the OSNs efficiently is usually a mix of connections topology with the review of content [137].

Two other approaches are used to explore social bots in networks: community detection and centrality measures. For example, a study looked into bots' effects on political discussion networks, using a community detection algorithm to outline the network structure and identify separate communities based on retweet relationships. The researchers also determined centrality measures to locate influential users within each identified community [138]. A study examined social bot activities and their interactions with humans on Twitter during the 2018 U.S. midterm elections [27]. Using network analysis, researchers noticed varying strategies between conservative and liberal bots. Conservative bots occupied a more central location in the social network. They balanced their interactions with humans and right-wing bots, while liberal bots mainly concentrated on engaging with human users.

The research paper "The Strength of Weak Bots" by Marijn A. Keijzer, inspired by the classical work of "The Strength of Weak Ties," [139] extensively employs network science principles to explore the impact of bots within social networks [140]. Initially, the study utilized ring and lattice network structures to simulate belief dissemination, mirroring the high clustering seen in OSNs. The influence of network clustering on belief spread was examined, revealing that reduced clustering accelerated the adoption of bot beliefs. The research also used spatial random graphs to explore the dynamics of directed networks [141], similar to OSNs like Twitter.

Furthermore, the study introduced a model to measure the bots' connectivity and activity within the network, uncovering paradoxically that the weakly connected and moderately active bots are more effective in spreading beliefs in the network at large. Homophily, the tendency of individuals to interact primarily with those sharing similar beliefs [142], was integrated into the model, reflecting its significant role in human interactions and its potential amplification by online platform algorithms.

Bots play a pivotal role within extremist networks, spreading content, managing communities, and facilitating recruitment. Evidence highlights their involvement in communicating basic information to potential recruits and influencing terror attack narratives on platforms like Twitter [143, 144]. Despite continuous efforts by platforms such as Telegram to curb their activities, bots remain integral components of extremist groups' operations [1].

The application of network science has significantly enriched our understanding of these extremist networks, both online and offline. Distinct properties, such as centrality and hierarchy, have been identified within these networks, influencing how information flows among other processes [145, 146, 147]. Utilizing such properties, experts can anticipate and potentially diminish the effects of terror activities [148].

As extremist networks evolve, their structures often shift from hierarchical designs to more fragmented ones made up of radical and moderate factions [149]. This transition may emerge organically or result from observing and adapting successful strategies from groups like ISIS [150]. Consequently, many extremist entities have swiftly adopted similar methodologies in both online and offline domains [151, 152, 153]. Notably, far-right groups on Twitter seem less cohesive compared to their ISIS counterparts [154].

While many real-world extremist networks exhibit hierarchical designs [148, 155], there is a pressing need for further research to determine whether their online presence mirrors this structure. With bots becoming increasingly integrated into these networks, be they organic or artificial, centralized or decentralized, it is essential to investigate the nature and structure of these online networks, and the roles bots fulfill within various network structures must be defined.

Network science provides tools to assess such hierarchical designs. For example, it is possible that extremist networks, be it in Telegram or any other platform, and like other large networks, show a scale-free degree distribution, indicating hierarchy in importance (degree centrality) [156, 157]. Moreover, the rich-club phenomenon, which is common in large-scale networks ranging from transport to scientific collaborations, offers another perspective of such hierarchical organization [158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169].

2.5 Summary

The interdisciplinary nature of social bot research highlights the complexity of the topic. While philosophical ethics, for instance, delves into the ethical aspects of using bots in political and social contexts [170, 171, 172, 173, 174, 175], social sciences investigates bots' actions and interactions with humans [176, 177, 178, 179, 180, 181] and the history of technology documents their origin, role, and advances [182, 183]. However, network science, machine learning, and similar data-driven approaches focus on issues such as identifying social bots on platforms like Twitter, mapping out their networks or the networks they are part of, and measuring their impact on OSNs. This thesis potentially belongs to the latter category.

Chapter 3

Emergent Local Structures in an Ecosystem of Social Bots and Humans on Twitter

> *"When you light a candle, you also cast a shadow."*
>
> Ursula K. Le Guin

The dynamics on social media platforms are deeply influenced by how users, both human and bots, interact within these spaces. Notably, the strategies employed—whether innate to humans or programmed into social bots—determine outcomes on these platforms. At the heart of this discussion is a fundamental premise: social bots and humans together form an ecosystem within social media. This ecosystem is characterized by foundational building blocks or local structures, which collectively result in the emergent properties that define the flow of interactions in online social networks (OSNs). In this chapter[1], we examine the local structures (motifs) of interactions between humans and bots, as well as the effects of different bot strategies on the networks' emergent properties. We aim to provide insight into whether the content or the users are more influential in shaping network structures, which may inform the design and deployment of more effective and beneficial social bots. This way we also hope to contribute to the ongoing debates about the nature of social networks.

For our study, we deployed six social bots on Twitter for a five-month period and analyzed the evolution of the networks they participated in using a network science framework. To understand these networks, we identified and examined the network motifs present within them. Network motifs, also known as "motifs" or "local structures," are statistically significant mesoscale structures that exist within larger graphs or structures [184]. Similar to social cliques, motifs are considered essential components for the higher-order organization of complex real-world networks [185]. They have been widely used to describe the dynamics of networks in natural systems, such as biological and ecological systems, and man-made networks, including power grids and social networks [186, 187, 188].

Network motifs have been extensively studied in OSNs like Twitter and Yahoo, they have

[1] The results of this chapter are summarized in a paper accepted for publication in EPJ Data Science.

been mainly used to map emotional expressions during emergency events [189], explain knowledge-sharing and question-answer patterns [190], and characterize opinion formation processes [191]. It is essential to note that network motifs are platform-specific, and their types and roles vary significantly from one OSN to another [192]. To date, there has been no research into network motifs in the context of human-bot interactions.

This work has potential to contribute towards improving the effectiveness of bots in carrying out useful tasks, as well as combating malicious bot activities. It also offers insights into how social bots can coexist with humans by adapting to the structures of existing networks. Understanding the patterns of human-bot interactions in online social networks is crucial for this progress.

The chapter is structured as follows: Section 1 describes the experimental setup, including the design of the social bots, their strategies for interacting with other users, and the content they posted. Section 2 provides an overview of the bots' networks and their evolution. Section 3 analyzes and interprets the motifs in those networks. Finally, in Section 4, we discuss the implications of our findings and the limitations of our experiment.

3.1 Experimental Setup

For this experiment, we created and released six bots on Twitter and monitored their interactions with humans over five months. These bots published content using a language model [37], and had the same interaction capabilities as a human on an OSN. Nevertheless, all of them were given specific "rules of engagement" and each pair of bots was assigned a different strategy that directed them to act in a distinct manner.

This experiment was composed of five main components: the environment, humans, bots, a mediator and a data collector. (1) The environment is the medium in which the experiment is conducted; in this case, Twitter. The platform's policies and regulations on safety, privacy, authenticity and more define Twitter's operating environment. (2) Humans are persons with a Twitter account who occasionally interact with the experiment. (3) Bots are automated agents: computer programs created for the experiment that communicate and engage in the chosen environment. (4) The mediator is a program that controls the bots' communications and engagements with other bots and humans. (5) A data collection utility securely stores the data into a server for further analysis. This section focuses on the bots and mediator of this experiment, detailing their design methodology and specifications.

3.1.1 Building social bots

Perception, decision, and action framework

In order to create a social bot on Twitter, two steps must be taken: 1) writing a script with instructions for the bot's behavior and 2) creating an application through the developer portal for the Twitter account [12]. This experiment employed a three-step logic framework of perception, decision, and action (see 3.1). Perception involves collecting and processing data, such as identifying trending hashtags, scanning user timelines for recent tweets, reading tweet text, and recording retweets and other engagements. In the decision stage, the bot chooses its next move based on insights from the perception stage. For instance, this decision is made by randomly selecting one or two most retweeted or liked tweets in one of the top five trending hashtags. Finally, in the action stage, the bot performs an action or series of actions. For example, if the bot decides to respond to a user's tweet, it will request text from a content generator API, which will then be used as a reply. For more details on the implementation of the framework, the bots' workflow and the data collection process check the supporting information (see A).

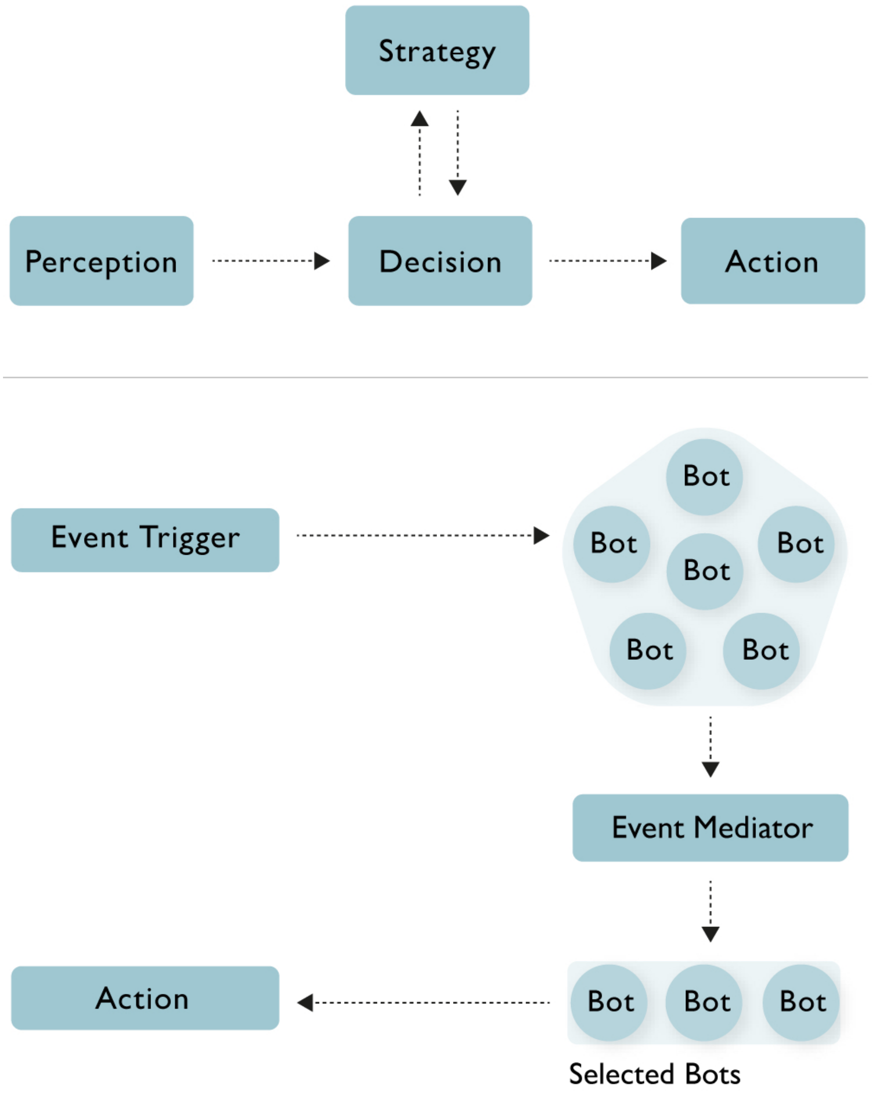


Figure 3.1: Diagrams of the perception, decision, and action framework and the workflow of the mediator. (Top) An illustrative diagram of the perception, decision, and action framework that dictates the workflow of each bot during the cycle of each interaction they make. (Bottom) The mediator API is built to listen to the events from Twitter and select the interaction type, the bots and the time of the interaction, and then put the response in a queue.

Profile and activity settings of social bots

Giving bots a human personality in their public operating environment requires at least names and photos [193]. We chose neutral names for the bots' profiles, such as "Philip Nolan" from Edward Everett Hale's novel *Man Without a Country*. This name could be either real or fictional, thus reducing the chances of recognizing that the accounts belong to bots as well as avoiding declaring that the content is not generated by a human. Additionally, we used fictional profile pictures to avoid infringing on other people's privacy by using real humans' photos [194]. To avoid any bias, we did not make any information about location or gender (except possibly the name) public on the bots' accounts. We also did not disclose whether these accounts were run by humans or bots.

The programming of the bots determines their "personality" and the way they interact with the environment. The bots were designed to take advantage of all the features that the environment provides; they could do anything on Twitter - tweeting, retweeting, following, unfollowing, replying, mentioning and liking - that a human user can. The bots were not told to follow each other, but during the process they began to follow one another.

3.1.2 The bots' workflow

The workflow scheme for each bot was composed of five parts: an environment perception API, a content generating API, a tweeting API, an interaction API, and a data collection API. The purpose of each component and how it interacts with the other components are outlined below (see Fig 3.2).

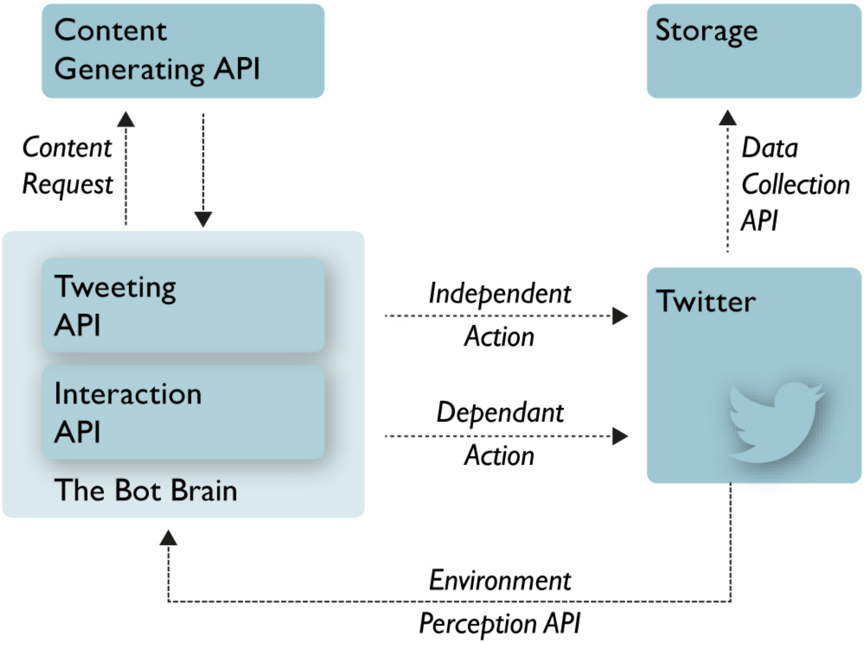


Figure 3.2: The bots' workflow diagram shows the interconnected APIs that facilitated bot functioning. The bot brain consists of both the tweeting API and the interaction API. The content generating API serves the brain by content. The environment perception API senses the environment and sends signals to the brain for analysis and decision. Finally, the data collection API collects the data in real-time. Two types of actions occur in this workflow: (1) independent action, which is tweeting, and (2) dependent action, which is retweeting, replying, liking, and following other users.

Environment perception API

The environment perception API acted as a sensor, gathering the necessary data from the bots' environment to enable it to make decisions. This data included the text of tweets and other information (trending list, engagement metrics, retweets, etc.) that was specific to the strategies and their interpretation. In the *Perception, Decision, and Action* Framework, this API was the technical implementation of the perception step as illustrated in the main manuscript. It was activated by a timer and when one of the bots wanted to interact or when a triggering event in the environment occurred, it sent the collected data to the mediator.

Tweeting API and interaction API

The tweeting API and the interaction API together formed the basis of the bot's brain. Two types of action were possible in this workflow: 1) independent action, which was tweeting, and 2) dependent action, which included retweeting, replying, liking, and following other users. The tweeting API used Twitter's API to allow bots to tweet according to a predetermined strategy or program; this was an independent action that was not influenced by external stimuli. The interaction API used Twitter API to retweet, like, reply, follow and mention users based on the bot's program; this was a dependent action as it relied on signals from the environment perception API. The "independent" action followed the strategy but still took into account information from the relevant external data - e.g., trends and users' timelines.

Content generating API

The content generation API was employed in two scenarios of the workflow: when the tweeting API requested content to tweet, and when the interaction API asked for content to reply to or mention a user in a quote. To create this API, we used DialoGPT, a state-of-the-art, transformer model [195]. We chose this model because it is conversational and trained on a Reddit dataset; this made its content outputs closely aligned to general social media language patterns and within Twitter's character limits. Of the three available versions of DialoGPT, we used the largest one (762m parameters) to maximize the bots' ability to respond naturally. The content generation API was only used for tweeting and replying in English.

Data collection API

This API uses webhooks to connect to Twitter, gathering all tweets and interactions—whether made by a human or a bot—in real-time. It does this by creating a web app for each bot, registering a webhook URL to make POST requests, and subscribing to an account. The collected data is then stored on a server for further analysis.

3.1.3 Common rules of the bot's engagement

The six bots were given three rules to abide by when interacting with other accounts and humans. Firstly, they could only respond to external prompts such as a reply to their tweet or an alert about new trending content. Secondly, the bots were instructed to not only reactively interact with other accounts but also post content proactively in order to look normal and interesting. Lastly, a mediator was responsible for deciding when the bots could interact. This was done for two reasons: first, without a mediator, the bots might interact with each other and others simultaneously, potentially violating Twitter's spam policy [196]; second, the mediator helped to balance out the bot-bot interactions among and within strategies.

3.1.4 Overview of the interaction strategies

The bots in our experiment were each assigned a strategy to interact with users in their network. Each strategy was designed to ensure that the bot's behavior was consistent and could help it gain the trust of human users. By varying the strategies of the bots, we were able to replicate the diversity of human users who interact in a variety of ways online. More importantly, each strategy was used as an independent variable to measure its effect on mesoscopic network structures, our dependent variable. The strategies reflected different ways in which OSN users interact with other users; for example, some people are interested in trends and novelties while others are more interested in influencers and their published content.

We created three strategies and gave each one to two bots. The strategies are designed to observe and react differently to changes in the environment and to other bots' and humans' actions, leading to "intelligent" bot behaviors [197].

The bots with the Trends-Targeting Strategy (TTS) kept an eye on the top five hashtags from the global trends list on Twitter [198]. They would then post about these hashtags or interact with other users who were using them by liking or retweeting their content, or by following them. The idea behind this strategy is that engaging with a trending hashtag increases a bot's visibility and, as a result, its chances of becoming embedded in the network. Trends are like digital public gatherings, and in such gatherings, there is a greater chance of meeting new people. For example, after a football game, a trending hashtag will draw in people who are interested in sports to discuss, converse and maybe fight.

The second strategy, the Keywords-Targeting Strategy (KTS), was also content-focused but more passive. The bots using this strategy did not depend on current or upcoming trends and conversations, but rather on a consistent interest in certain topics. They would search other users' tweets for the presence of one or more of these topics and interact with them if found. They would also post tweets about their topics of interest. The idea behind this strategy was to find users with similar interests to interact with. KTS is based on the reasonable assumption that most people have a limited range of interests when browsing the internet and are less likely to engage randomly with any topic.

Twitter's advertising campaign guidelines suggest using a minimum of 25 keywords [199]. For this experiment, we opted to use 50 English keywords that were selected from the University of Vermont Complex Systems Center's research on universal positivity bias and happiness [200]. This study provided us with a list of non-polarizing keywords that fit with the purpose of our study. These included words such as *music, jokes, forests, family, cake, kitten, success, holidays, and beach*; a full list of chosen keywords can be found in the accompanying supporting information (see A).

The User-Targeting Strategy (UTS) was also employed in this experiment, which involved bots following influential Twitter accounts. This was based on the idea that by doing so, the bot's exposure to their followers would increase, thus boosting its own popularity within the network. The influential accounts chosen were diverse and politically balanced, and their levels of popularity were comparable; some of these included Barack Obama, Donald Trump, Justin Bieber, Cristiano, Bill Gates and CNN. The influential accounts were identified by the number of followers they had regardless of why they had them. A full list of users is available in the supporting information section (see A).

Our decision to employ these three strategies was based on two key considerations. Firstly, each strategy operates with a different level of focus. The TTS is more dynamic, continuously adapting to current trends and thus encompassing a broad array of rapidly changing, largely unrelated topics. Conversely, KTS narrows its focus to a defined set of 50 keywords, yielding a more limited range of topics. UTS is the most targeted of all, specifically centering on a select group of users. Secondly, the strategies can be classified into two categories based on their orientation: content-centric (TTS and KTS) and user-centric (UTS). Our objective is to draw meaningful comparisons between networks that arise from a focus on content and those shaped by users' interests.

3.1.5 The mediator workflow

The mediator's role was to oversee and coordinate the bots' activities and interactions. This API was designed as a program within which the bots interacted autonomously, but they were aware of when to interact after receiving a signal to do so. The workflow followed a four-step process (see Fig 3.1): (1) being notified of a triggering event, such as a reply, a mention or a new trend; (2) randomly selecting one or two bot(s) to act and randomly choosing the type of action; (3) if the chosen response involves text, requesting a response text from the content generation API; and (4) scheduling the bot's response or the responses of multiple bots, ensuring there is an interval between responses if more than one bot was instructed to respond. The mediator only governs the interaction API, which is the first part of the bots' logic. The tweeting API, which is the other part of their logic, works independently from the interaction API and according to each bot's strategy.

Network Representation and Data

To track the bots' activities, we created network maps of their interactions. This was done through snapshot representation, which allowed us to analyze the network as its structure evolves [201]. Snapshot representation creates a discrete-time sequence of networks, as shown in Eq. (3.1). We aggregated the interactions after one day, built the resulting network, and repeated this process for each subsequent day:

$$\mathcal{N} = \{N(1), N(2), ..., N(t_{max})\}, \tag{3.1}$$

where $N(t_{max})$ is the number of networks.

The experiment ran for five months (from July, 2020 to November, 2020), with unavoidable interruptions to data collection due to external constraints. The interruptions occurred due to two primary causes. First, Twitter occasionally restricted the tweeting and interaction APIs used by the bots, resulting in restriction periods ranging from 24 hours to two weeks during which the bots could not take any action. Second, there might have been technical issues in the data collection API which prevented full data collection but did not affect the bots activity.

Despite these challenges, our analysis was designed to ensure the validity and integrity of our results. Specifically, we filled any gaps in data with information from the preceding period based on our assumption that the bots' activity would likely remain unchanged until new data were received. This approach was employed as a means to maintain the continuity of the experiment.

As a result, we collected 75 days of data on bots using KTS, 95 days of data on bots using TTS, and 90 days of data on bots using the UTS, with significant overlap between the observation periods. The two most significant periods of inactivity were 15 days for the bots using KTS and 9 days for those using UTS. These were both due to suspensions from writing to the Twitter API. However, these suspensions were removed automatically.

We acknowledge that this may raise some concerns about the completeness of our data. However, it's important to note that these interruptions mirror real-world behaviors of human users on social media platforms who often experience their own periods of inactivity or suspension, so a pause in tweeting or interacting does not significantly compromise the integrity of our results. To further validate this approach and our results, we ran our analysis on both the complete data set and a sample with no data collection interruptions. Importantly, we found identical results in both cases, which adds a layer of confidence to our findings.

The data shows five types of interactions: *like, follow, retweet, reply, and mention*, and four settings of interactions: bot-bot interactions, bot-human interactions, human-bot interactions, and human-human interactions (see Fig. (3.3)). However, we excluded 26 interactions executed by a human and targeted at another human from our analysis for two reasons. First, they are not relevant to this work's research questions. Second, explaining how bots facilitated these human-human interactions requires special handling and different data collection methodology that is not centered around bots.

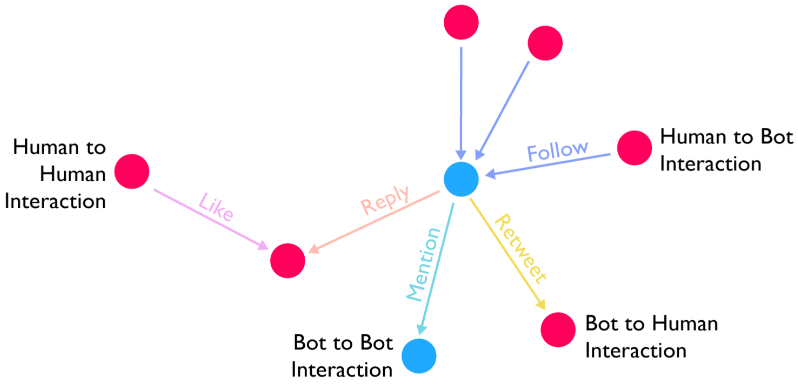


Figure 3.3: This diagram illustrates the interactions of humans and bots within the networks. Bots, represented in blue, and humans, depicted in red, engage in four potential types of interaction: bot to bot (blue-blue), bot to human (blue-red), human to bot (red-blue), and human to human (red-red). These interactions, denoted by links between respective nodes, can manifest as follows, likes, mentions, retweets, or replies - each distinguished by a unique color. The graphic provides an overview of the potential dynamics within this ecosystem, clarifying the participants in each interaction and the various forms these interactions can take.

We also created bot-human networks for each strategy and a bot-human network that tracked the activity of all bots (see Fig. (3.4)). Each node in the network represented either a bot or a human, while links between nodes indicated one of the five types of interactions, such as "likes" or "follows". Our network map treated all forms of interaction equally, regardless of quantity or type. This is a simplified version of a more intricate reality. The quality of a link can vary depending on factors such as whether one user is following another user or simply retweeting their content, or if they are following a node with one hundred followers or one hundred thousand followers.

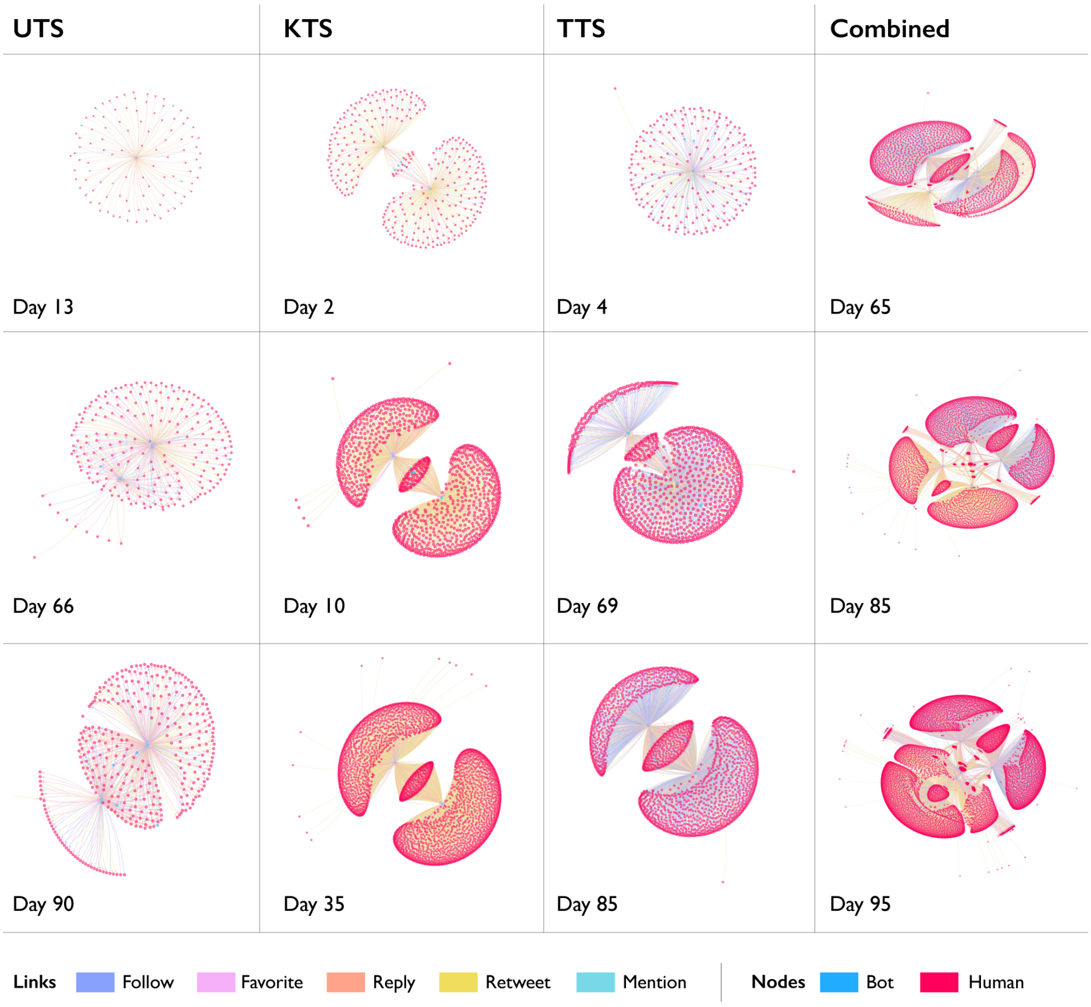


Figure 3.4: Snapshots of the interactions between humans and bots. Each graph represents the state of connections on a specific day for: UTS, KTS, TTS, and the strategies combined. The state of connections on a specific day is represented by each graph and denotes a red node for humans and a blue node for bots. The link between them is one of the five types of interactions.

Altogether, the bots tweeted around 21000 times and were involved in around 43000 interactions. Fig. (3.5) demonstrate that the TTS had the most incoming and outgoing interactions. The rewards, which are incoming interactions from humans to bots, are generally much higher with the TTS than with the other two strategies, except that the "incoming likes" are much higher in KTS. The UTS received the least rewards in all incoming interactions.

While bot-bot interactions played a relatively minor role in our analysis, we recognize their potential importance in future metaverse-like environments. However, our study focuses on analyzing interactions between bots and humans, in both directions. We found that "likes" and "replies" were the most common types of interactions between humans and bots, while "mentions" and "retweets" were less common. Other interactions fell somewhere in the middle. Our findings suggest that bot-bot interactions appear to be rather

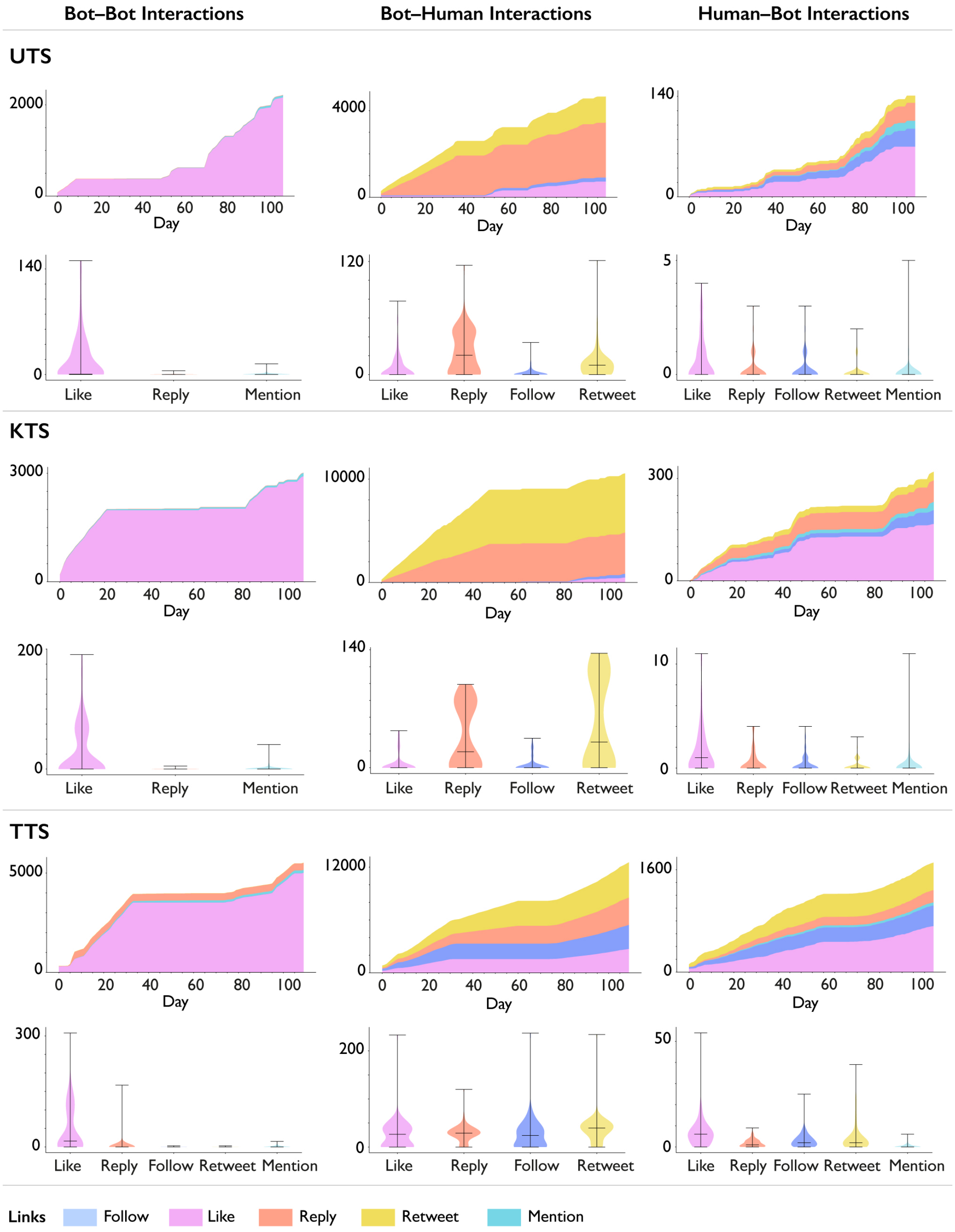


Figure 3.5: Summary graphs to illustrate the interactions of the UTS, KTS, and TTS. The stacked line graphs allow for a comparison of the cumulative performance of three ways of interactions: bot-bot, bot-human, and human-bot. The graphs have distinct colors to represent each of the five different types of interactions. Additionally, the violin plots display the daily maximum, median, and minimum number of interactions for each of the three interaction types: bot-bot, bot-human, and human-bot.

mechanistic and limited in diversity, interactions involving humans and bots are more varied and meaningful, highlighting the potential of online platforms to facilitate rich and engaging exchanges. This implies that although bots participate fully in the ecosystem, much of the complexity comes from the human participants (see Fig. (3.5)).

3.2 Analysis of Local Structures

3.2.1 Detecting local structures

By examining local structures, referred to as motifs, in the networks, we were able to gain a better understanding of bot-human interactions' patterns. This experiment focused on three-node motifs, which are that involve three nodes (either human or bot) connected by a type of link (interaction). In directed networks, such as those found on online social networks, there are only 13 possible types of three-node connected motifs, which can be seen in the image below (see Fig. (3.6)).

We chose to analyze three-node motifs for three main reasons. Firstly, they are much more prevalent in social networks than four- or five-node motifs. Secondly, the functions of three-node motifs are more meaningful and straightforward in social networks; they represent the most basic level of group connection. Lastly, they require less computing power to identify than larger motifs. The 13 motifs identified in a network each have a unique shape and purpose, which can be used by individual nodes or users to increase their influence. Our experiment demonstrated that each motif has a specific function, and analyzing their prevalence in different networks is essential for understanding the strategies of bots.

We used the subgraph enumeration algorithm (ESU) to identify network motifs [202], ensuring that these motifs met the graph isomorphism condition. Upon detecting all motifs, we classified them into one of 13 pre-defined motif types.

To distinguish between significant motifs (those reflecting a pattern) and random occurrences, we employed the configuration model. This statistical model generates random graphs by reshuffling the edges of the original network while preserving the same degree distribution. The random networks, therefore, maintain the degree sequence of the original networks but have randomized structures. For more detailed information on the configuration model, please refer to the (see A).

We applied the ESU algorithm and the isomorphism condition to these randomized networks, enumerating and categorizing the motifs found within them. By comparing the frequencies of each motif type in the original networks and the randomized networks, we could identify which motifs in the original networks occurred with a significantly higher frequency.

Significant motifs are those that occur more frequently than would be expected by chance in the randomized networks. A motif was considered "frequent" if it appeared in the original network significantly more often than in the majority of randomized networks.

Motifs that did not meet this frequency criterion were deemed likely to be random occurrences, and thus less informative for understanding bot strategy performance. Following the construction of the networks, we carried out motif detection, classification, and analysis at regular intervals based on snapshots captured nearly daily.

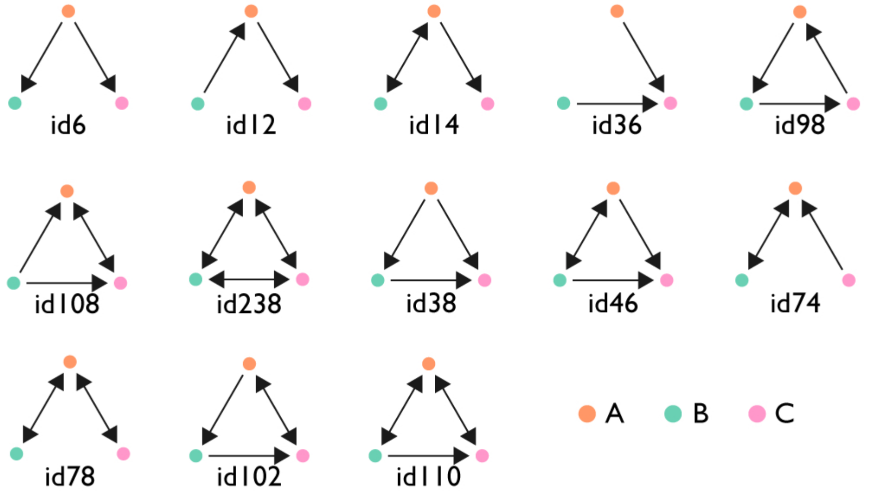


Figure 3.6: Detailed illustration and interpretation of all 13 possible three-node connected motifs in directed networks, highlighting six significant motifs and their contextual relevance in OSNs. 1. Motif id6 (“Diverging” or “Fan-out” motif): Characterized by a source node (A) diverging to two other target nodes (B and C), this motif can indicate broadcasting or information spreading scenarios in OSNs, with a single user disseminating content to two recipients. 2. Motif id12 (“Chain” motif): Displays a sequential pattern of connectivity, forming a directed “chain” or “path” ($B- > A- > C$). This motif is indicative of sequential information exchange, typically observed in instances like information cascades. 3. Motif id14: Comprises a chain-like structure between two nodes (A and B), with node A also connecting directly to a third node (C). 4. Motif id36 (“Converging” or “Fan-in” motif): Inverts the diverging motif structure, featuring two source nodes (A and B) converging into a single target node (C), representing shared support or common targets within a network. 5. Motif id98 (“Circular” motif): Constitutes a cyclical, unidirectional pattern where each node is linked to one other node ($A- > B$, $B- > C$, and $C- > A$), enabling a path returning to the origin node. 6. Motif id108 (“Feed-forward Loop” variant): Diverges from the traditional “Feed-forward Loop,” with one node (B) reaching out to two other nodes (A and C), with nodes A and C sharing a reciprocal connection. 7. Motif id238 (“Fully connected motif”): This motif signifies a network condition where all nodes have mutual connections. 8. Motif id38 (“Feed-forward Loop”): Presents a motif where node A connects to nodes B and C, while node B also connects to node C, illustrating instances of information cascading via a direct and intermediary path. 9. Motif id46 (“Feed-forward Loop” variant): Contains two mutually connected nodes (A and B), both independently connecting to a third node (C), indicative of either independent information sharing or non-reciprocal friendship situations within a network. 10. Motif id74: Mimics Motif id14’s configuration, but with nodes A and B sharing mutual connection, and node C also connecting to the first node (A). 11. Motif id78 (“Star” or “Hub” motif): Demonstrates a node A maintaining mutual connections with nodes B and C, but without a direct link between nodes B and C. 12. Motif id102 (“Feed-forward Loop” variant): Incorporates an additional mutual connection between nodes A and C, compared to the typical “Feed-forward Loo” motif that forms a directed chain between nodes A, B and C. 13. Motif id110 (“Feed-forward Loop” variant): This motif exhibits nodes A and B with a mutual connection, a directed edge from node B to node C, and a mutual connection between nodes A and C.

We detect motifs when at least one of the nodes is a bot. The motifs defined above without the specification of the node types are called structural motifs. We have two kinds of nodes, humans and bots, which can be represented by different colors, leading to colored motifs.

3.2.2 Emerged local structures

Our analysis of the dataset revealed the presence of six out of the thirteen structural motif types: id6, id12, id36, id74, id46, and id38. Each of these motifs represents a specific pattern of interaction between bots and/or humans on social media platforms. Motifs id6, id12, and id36 have simpler relations between two nodes out of the three. For example, motif id6 can occur when one bot retweets the posts of two humans, whereas motif id36 results from two bots following the same human. In contrast, motif id74 is more complex; for instance, it can occur when two bots interact with each other and a human interacts with one of them. Motif id74 is complex because it involves more complicated interactions between nodes (i.e., have a reciprocal relation between two nodes out of the three). Motifs id46 and id38 have a closed-triangle structure with interactions flowing between three nodes.

Fig. (3.7) illustrates the relative importance of motif types in the networks of UTS, KTS, and TTS. It is evident that different strategies produced distinct local structures with varying levels of significance. The UTS, KTS and TTS all shared two simple motifs, identified as id12, id36 and one complex motif identified as id74. Complex motifs, however, were not shared across the board. A motif is considered simple if it cannot close the triangle or has only one reciprocal link; a complex motif has a closed triangle and more than one reciprocal link. In summary, the UTS activated six types of motifs, the KTS activated four, and the TTS activated three.

Fig. (3.7) also demonstrates the daily frequency of the significant motifs during the observation period. Most of the motif types had a steady but mild growth rate. However, id12 and id6 had a much more significant increase in growth compared to the other motif types, with an increase of an order of magnitude more than any other type. This suggests that these two motifs were dominant in the networks being studied.

The data also show a substantial similarity between TTS and KTS in terms of activated local structures. They both activated three motifs: id74, id12, and id36. This might be due to them being content-oriented. However, KTS and TTS were distinguished by the absence of the id6 motif in TTS. Motif id6 refers to a dissemination pattern where one user amplifies the posts of two different users. This pattern may manifest itself in a variety of ways, including combinations of likes and retweets to two other users, be they humans or bots. Even though TTS yielded higher rewards in terms of followers and replies, KTS demonstrated better results in terms of "likes" and in propagating the id6 motif.

In TTS, we noted the emergence of motif id74. This motif, characterized by two nodes interacting with a third node and only one node from the pair receiving engagement from the targeted individual, suggests a certain level of engagement focused on content in social media. However, the fact that TTS and KTS activated only this one complex motif, despite its significance, indicates a relatively weaker form of engagement compared to UTS.

The absence of motif id6 in TTS suggests that engagement, such as mentioning another user, doesn't automatically lead to wider content spreading like retweeting. This implies that TTS might engage actively but not necessarily disseminate content widely. On the contrary, KTS activates both id74 and id6 motifs, indicating an effective blend of engagement

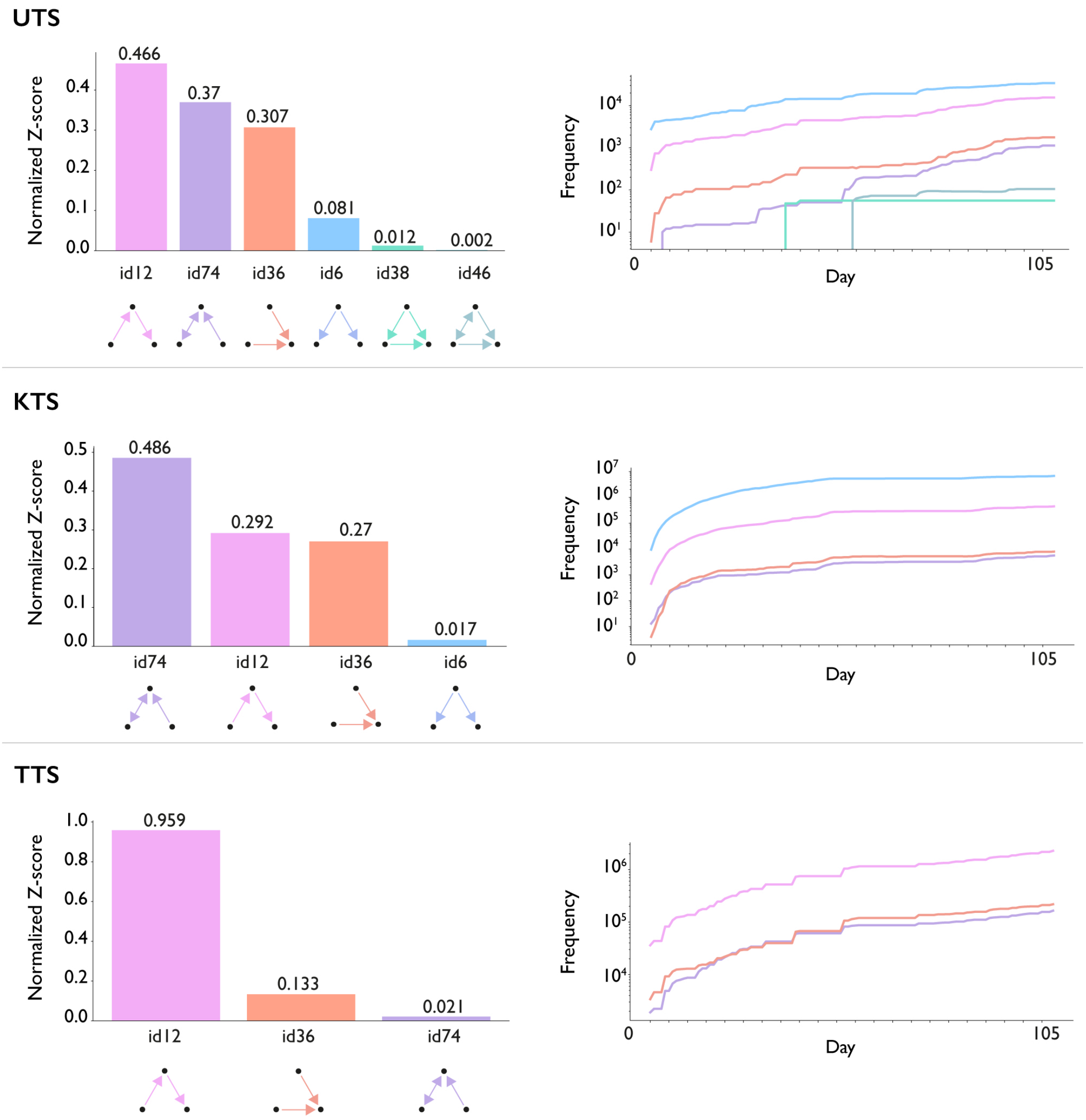


Figure 3.7: Emerged local structures in UTS, TTS and KTS. (Left) Bar plot illustrates the average normalized z-score only for significant motifs. UTS showed six network motifs, KTS showed 4 motifs, and TTS showed 3 motifs. (Right) Line graph shows the changes in frequency of significant motifs over time.

and content spreading. Thus, KTS not only involves direct interactions but also amplifies content, providing a balanced approach to social media interaction.

Both the UTS and KTS also shared the common characteristic of frequently activating the id6 motif. Motif id6 is a spreading pattern, and the results show that both a content-based strategy and user-based strategy are able to spread (amplify) content, when the focus is high. The difference between KTS and UTS lies in the former focusing on a broad array of topics instead of concentrating on individual users. Such observations may also suggest that focusing on targeted content also implies a limited number of users' interests.

UTS uniquely exhibited more complex motifs such as id38 and id46, pointing to the intricacies of user-focused engagement. The manifestation of these motifs in the UTS implies that a tighter focus, as embodied in this strategy, tends to amplify the level of engagement. This not only reiterates the correlation between the preciseness of an interaction strategy and the pattern of engagement within these networks but also emphasizes the importance of direct and bidirectional engagement in building interactive relationships.

In summary, the two primary patterns identified in our study, id6 and id36, demonstrate the most common behaviors exhibited by bots in social media: spreading, where a bot retweets multiple tweets, and targeting, where multiple bots follow a single human. These strategies are frequently used for marketing, propaganda, and misinformation purposes [15]. We found that all the strategies employed in our study demonstrated these patterns, except for TTS, which did not activate id6. However, all strategies activated id74, which indicates engagement due to the existence of a reciprocal link between two out of the three nodes. In KTS, motif id74 was the most significant, suggesting that a keyword-based approach is effective in increasing engagement between users.

Building upon our structural motif analysis, we incorporated a color-based methodology to more precisely identify the unique interaction patterns arising in the bot-human network. Inspired by prior research [203], we labeled nodes and edges with specific colors: nodes were designated either as "bots" or "humans," while edges were colored to represent one of five types of interactions – retweet, like, follow, mention, or reply. This approach allows for a nuanced representation of network relationships, as shown in Figure (3.8).

The color-based analysis unveiled a higher number of unique patterns in each strategy: TTS, KTS, and UTS yielded 479, 388, and 245 unique colored patterns, respectively. Intriguingly, across all strategies, the colored variations of motif id74 emerged as a foundational structure. Nevertheless, its distribution varied significantly: while it constituted 85% of all patterns in TTS and 78% in KTS, UTS exhibited a relatively diverse pattern distribution with id74 making up 59% of all patterns (see Figure (3.8) for more details).

Despite its frequent occurrence, motif id74 is far from trivial. The id74 structure demonstrates a situation where two nodes (A and B) maintain a reciprocal relationship, typically a "friendship" on the platform, with a third node (C) only connecting with one of them (A). This structure carries the potential for future engagement between node C and B, either directly or indirectly through node A.

In TTS, for instance, the most common id74 pattern consisted of a bot and human in a mutual follow relationship, with another human liking the bot's content. For UTS, the id74 pattern most often showed a bot and another bot engaged in a reciprocal relationship, with a human liking the first bot's content. In KTS, the prevailing id74 pattern showcased a bot and a human in a mutual follow relationship, with another human liking the bot's

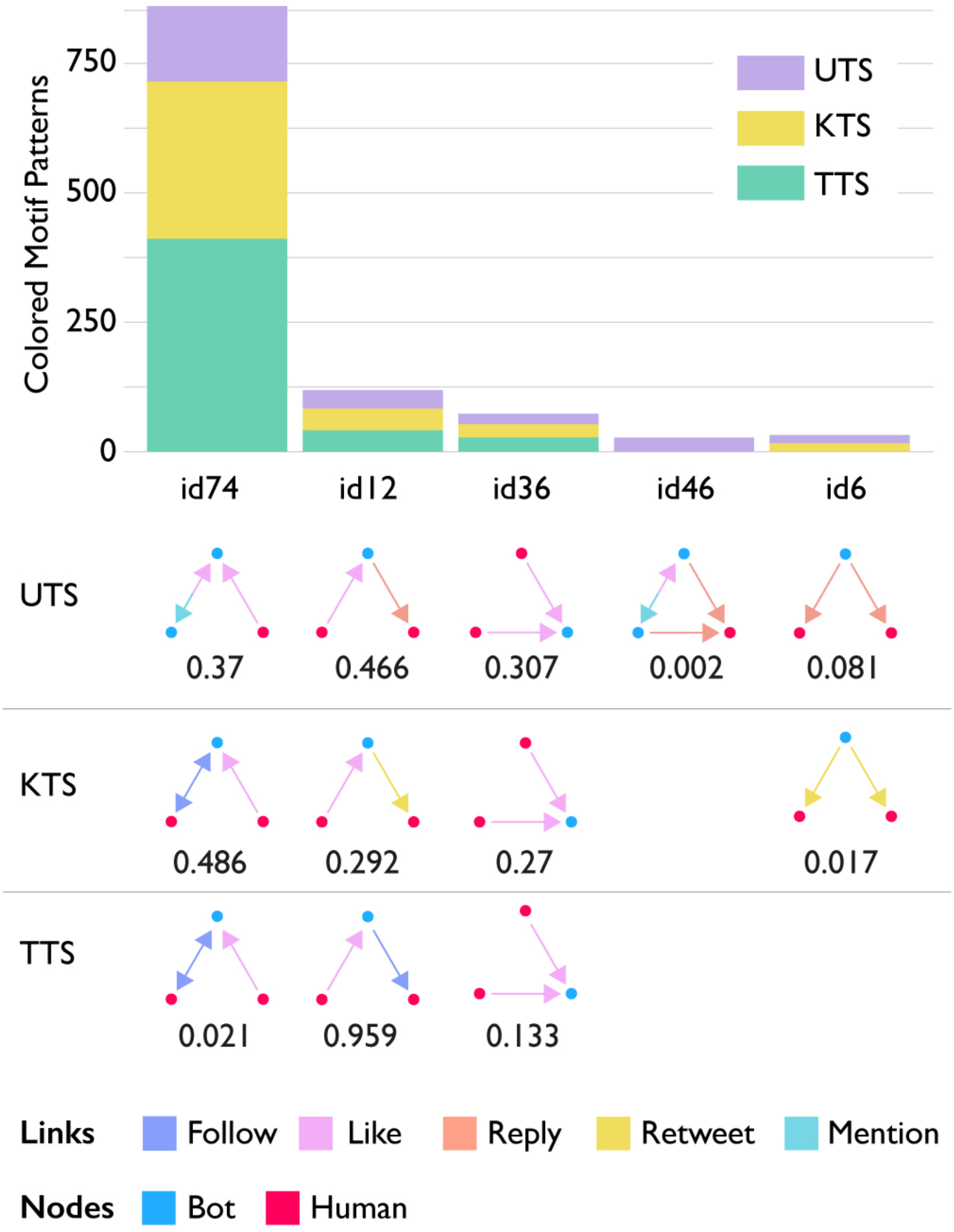


Figure 3.8: Stacked bar plot illustrating the distribution of colored motifs across strategies. The top section showcases the prominence of Motif id74 across all strategies, constituting 59% (146 instances) of patterns in UTS, 85% (411 instances) in TTS, and 78% (304 instances) in KTS. In comparison, Motif id12 takes up approximately 10% (42 instances) in KTS, 8.5% (41 instances) in TTS, and 14.6% (36 instances) in UTS. Motif id36 is represented with 6.7% (26 instances) in KTS, 5.6% (27 instances) in TTS, and 8% (20 instances) in UTS. Moreover, id46 and id6 make up 11% (27 instances) and 6.5% (16 instances) of UTS respectively, while id6 constitutes 4% (16 instances) in KTS. The bottom part of the figure depicts the most prevalent pattern for each distinct motif type within each strategic approach, underlining the unique interaction dynamics nurtured by each.

content.

These motifs, dynamic in nature, often evolve over time, transforming into different types. For example, node C, in motif id74 case, may later establish a direct relationship with node B, where node A acts as a bridge between B and C, which then leads to closing the triangle. Or, node C could propagate information to B through A. In this context of information cascades, such indirect relationships can be as impactful as direct ones.

A feature shared across these strategies is the motif id36, which typically displays two humans liking a bot's content. This consistency suggests that regardless of the implemented strategy, bot-generated content has the capability to attract and engage human users.

Moreover, the analysis exposes a fundamental aspect that transcends these strategies: human-initiated interactions with bots are crucial for activating significant network motifs. For instance, the common patterns of motif id74 in TTS and id12 in KTS both originate with a human liking a bot's content, leading to subsequent interactions. Similarly, across all strategies, motif id36 shows two humans like a bot content. Thus, even though bots can influence the dynamics of these interactions, human engagement with bots appears to be the initial spark that ignites these patterns.

This ecosystem of humans and bots exhibits two key properties: it is self-organized, meaning that we did not instruct the bots to follow or retweet the same humans, yet they did so in a pattern that emerged organically. Additionally, the ecosystem is evolving in two ways: firstly, it is gaining slow but steady traction as more users engage with the bots; secondly, the interactions are progressing into more complex local structures, where interactions can have diverse and nuanced meanings.

3.3 Discussion

In this study, we explored the impact of social bot strategies on the local structures that emerged within the networks of humans and bots on Twitter. By creating six bots and observing the changes in network structures, we identified three-node motifs that reflected patterns such as targeting users, spreading messages, propagating relationships, and engagement among users.

Our findings indicate that the most rewarded strategy (TTS) had the least diverse local structure patterns, while the least rewarded strategy (UTS) had the largest set of local structures, activating up to six different types of motifs. This suggests that the choice of bot strategies should be informed by the specific goal for which the bots are intended.

Furthermore, we found that content-related strategies (TTS and KTS) had a significant overlap in terms of local structures activated, while the user-oriented strategy was the only one to activate more complex motifs. Each of the three strategies could activate motif id74, a local structure consisting of two reciprocally linked nodes and an additional node interacting with one of these. This structure, serving as a foundational component for all strategies, bears potential to drive information cascades. Notably, the KTS shows promise in bridging the benefits of both approaches, by targeting content that resonates with particular users. This strategy has shown the ability to create engaging patterns while effectively disseminating content, key to achieving success on social media platforms.

It is worth noting that there is a debate about what defines a social network: is it the content or the users (i.e., their networks)? [204] Our findings suggest that the activation of more diverse motifs in UTS indicates the potential success of a user-focused social network, which prioritizes building connections between users. KTS' success highlights the importance of understanding the role of both content and user interactions in shaping the structure and dynamics of online social networks. While typically experiments require a large number of bots to influence individual opinions and demonstrate the impact of content on people, we conducted our experiment on a small scale and used non-sensitive material to maintain neutrality and technical simplicity.

It is also important to note that the visualization of these motifs do not reflect the temporal character of the links. While the existence of these motifs suggests that certain patterns of interaction between bots and humans may be more likely to occur in sequence, further study is needed to uncover the temporal-causal aspects of the motifs [205].

This study concentrates on uncovering local interaction patterns that emerge within networks comprising both bots and humans. While we intentionally omitted certain human-to-human interactions from our data analysis, we recognize their potential importance and the unique interaction dynamics they could reveal. The role of bots in shaping these omitted interactions presents a compelling avenue for future research, particularly with regard to how bots might alter these interpersonal dynamics and potentially influence the relationships established among human users.

We must also note that our research is contextualized within a specific timeline, preceding the managerial changes introduced by Elon Musk's acquisition of Twitter in 2022. As a result, the direct generalizability of our specific results to the evolving Twitter landscape might be limited. However, the value of our study extends beyond these findings and is also strongly associated with the methodologies and analytical approaches we employed. The use of network motifs to understand local interactions proves instrumental in comprehending user dynamics on digital platforms. While these motifs are platform-specific, the methodology itself is adaptable, providing a blueprint for researchers studying other social media platforms.

Chapter 4

Structural Differences in Extremist Networks on Telegram

> "The real question is not whether machines think but whether men do. The mystery which surrounds a thinking machine already surrounds a thinking man"
>
> B. F. Skinner

The growth of social media platforms and the increased impact of digital life on all individuals have resulted in a new public domain that can be used for good and bad. On the bad side, the internet is widely recognized as a crucial factor in radicalization, attack planning, terrorist recruitment, and propaganda spreading [206, 207, 208, 209, 210].

For example, Twitter took action against digital terrorism in August 2015 by deactivating over 300,000 accounts linked to ISIS at the request of the US Cyber Command, highlighting the widespread use of social media by terrorists [100]. Influenced by the digital era, terrorism has evolved significantly, with this transformation being marked by the generation of a vast amount of data. Researchers have used and analyzed this data to understand terrorist behaviors, objectives, and networks [211, 212]. Terrorism's digital transformation has been driven by three technological advancements: the creation of the internet, the rise of social media platforms like Twitter and Facebook, and the development of instant messaging applications such as Telegram and Viber.

Facebook and Twitter were initially started to help users to connect and grow their online social networks. Over time, they incorporated instant messaging functionalities, such as audio and video calls and file transfers. In contrast, Telegram and Viber began as instant messaging apps. However, they later integrated features typical of online social networks, such as forming groups where users can connect and disseminate content.

Social media platforms and instant messaging apps leverage the power of network effects, enabling individuals to interact seamlessly, irrespective of geographical or other constraints. These tools and platforms have facilitated decentralized terrorism, allowing individuals from any location to become radicalized and engage in or carry out extremist activity. Recent events, such as mass shootings in the United States (e.g., the Buffalo, NY shooting) broadcasted live on platforms like Twitch, demonstrate this phenomenon [213], as do lone-wolf attacks orchestrated online and aimed at the Western world [214, 215].

Initially introduced as a peer-to-peer messaging app, Telegram has now evolved into a multifaceted platform serving a wide range of purposes. It has grown to include social media features like content sharing and user interaction, as well as public channel and group creation [216]. By June 2022, the app had attracted 700 million users [217]. Extremist groups prefer Telegram's messaging platform due to its high-level encryption, somewhat anonymous features, and looser content restrictions compared to other services [96].

This work focuses on three extremist groups: supporters of the Islamic State of Iraq and Syria (ISIS), pro-Russia actors involved in the Ukraine conflict (PR), and far-right communities (FR). I compare how these groups utilize the platform, their ultimate goals, and the emerging network structures that enable their activities. Collectively, we expect that each group uses Telegram for various purposes, such as reporting events, spreading disinformation, building communities, and coordinating protests and violence.

Extremist and terrorist networks often evolve to a decentralized state, enhancing their resilience and adapting to real-world conditions. However, the specific emergent properties of these networks vary [218, 219, 220, 221, 222, 223]. For instance, ISIS has been a primary target of Telegram's disruptive efforts. As a result, we anticipate that the ISIS network on Telegram is fragmented, self-organized, and consists of highly engaged members. It's unlikely that a central authority exists on the platform to guide or dictate activities. Conversely, the PR network, which is less exposed to control and measures from the platform, has a centralized structure. Moreover, as this network is likely supported or even controlled by the Russian state, the centralized form is much more appropriate for conveying messages of a central power. FR seems to take an intermediate position. Its overlapping yet distinct ideologies and operations across different geographical regions point toward decentralization. However, the competing nature of FR groups highlights a tendency to evolve into a centralized structure.

Disruptions on the platform influence the stability of ISIS's presence on Telegram, adding another layer of complexity to the dynamics. While ISIS primarily focuses on online continuity and survival (i.e., community building and support), FR and PR emphasize propaganda and war reporting (or misreporting). The evident advantage FR and PR enjoy is partially due to the impact of aggressive suspensions of channels and groups linked to ISIS networks, considering that Telegram's treatment of these groups varies and can influence their networks.

This study explores the network structures of these groups in Telegram, what these structures could facilitate, and the consequences of different structural arrangements/emergence. I analyzed the activity within these networks, the roles of different types of users, such as bots, and the potential organizational principles emerging from the activities of these groups.

This chapter unfolds in the following manner: Section one describes Telegram and its features, outlines the data collection process, and provides a representation of Telegram networks. Section two explains the methods and unveils the findings of the study. In the final section, I discuss these findings' implications and suggest future research directions.

4.1 Telegram: Data and Network

Telegram is an instant messaging app where users can send messages, voice notes, and multimedia files such as images, audio, and video and make voice and video calls. Telegram messages can be encrypted and self-destruct. Telegram was launched in 2013 by brothers Nikolai and Pavel Durov, who previously founded the Russian social network VKontakte (VK) [224]. There are two types of communities on Telegram: channels and groups. These types enable admins to distribute content, open conversations, and chat between users. Channels serve as a one-to-many communication tool, where admins share content for subscribers to read and, in some cases, interact by using likes and emojis but not replies or comments. Groups, called chats, function as many-to-many communication platforms, enabling members to share content, respond to posts, and interact with others. A recent update permits channels to connect to discussion groups, allowing users to comment on posts and engage with the content. However, users cannot directly post within the channel (see Fig 4.1). Telegram also features mega-groups (or super-groups), officially classified as channels by Telegram but operating similarly to groups, supporting user interaction, posting, and replies. Both channels and groups can be public, private, or semi-private, meaning a link to a channel can be shared to join, but searching for it is not possible. I regarded semi-private as public in our analysis.

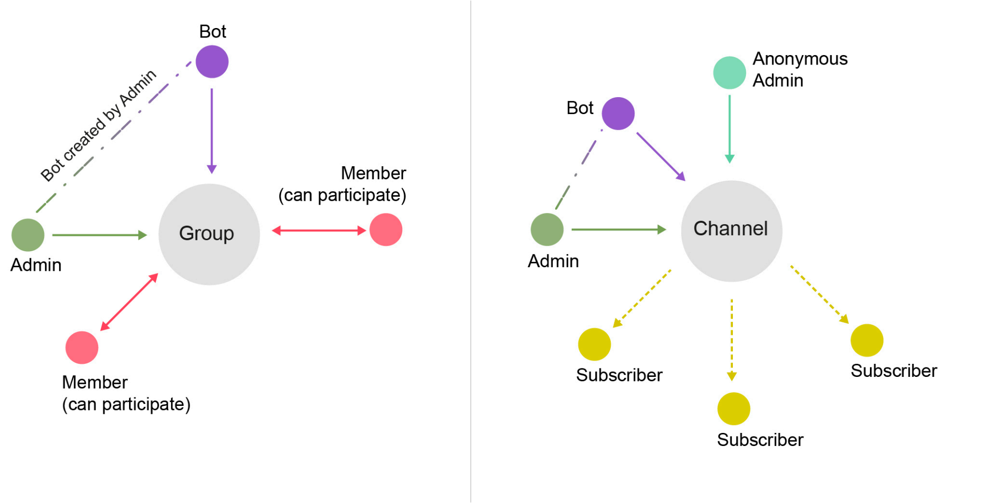


Figure 4.1: Understanding the difference between channels and groups in Telegram. Telegram offers two mediums for content sharing and community building: channels and groups. The diagram illustrates how channels operate under the management of admins, which can be bots, allowing users to subscribe and view the content. Meanwhile, groups enable users to join, share content, and engage with one another under the oversight of admins, who could also be bots. Both channels and groups can be managed by "anonymous admins" with unknown identities. Extremists often exploit these features to disseminate information and establish communities.

People use Telegram for various purposes, such as communicating with friends and family via private messages, subscribing to channels to access and view content, or engaging in groups with other individuals. The platform provides an "anonymous admin" feature, which enables admins to share content without revealing their identity. Generally, users register with their phone numbers but often make them private. In most cases, they must have at least a given name that is different from their actual name. They can also create usernames (i.e., IDs) that allow others to search for them. Telegram also permits the discovery of users, groups, and channels nearby. Furthermore, anyone can create bots in Telegram, which can be added as admins to channels and groups. These bots can, for example, receive private messages, moderate groups, and post content.

The "anonymous admin" function was developed following the arrest of revolution channel admins by authorities in Belarus [225]. This feature enables admins to share material without disclosing their identity, as the content is credited to the channel's name. While this is considered a way for Telegram to aid protesters in oppressive countries, it may also be exploited by criminals and terrorists.

In this study, I categorize users into three types: human, anonymous, and bot. Anonymous accounts do not have a username, first name, or last name. In contrast, human accounts have at least one of these identifiers. However, these identifiers might not represent authentic names, maintaining a level of anonymity.

Many groups and channels employ bots to assist with moderation tasks. Bots are not independent entities, but must be linked to a pre-existing account. However, the link between the bot and its owner is only known if mentioned. To create a bot, users can access BotFather (https://t.me/BotFather), an official Telegram account that manages bot registration. The capabilities of bots are limited only by the expertise of the developer and the available application programming interface (API) functionalities, which can enable functions such as content sharing and group administration.

4.1.1 Data collection process

The data collection process covered Telegram activities from September 2017 to April 2023 and followed four steps:

i Channel/Group identification: Initially, a few Telegram channels/groups were identified by keyword searches. These search terms associated with the extremist group under investigation were utilized to locate a number of primary channels, referred to as "seed channels." These seed channels served as the foundation for discovering additional channels/groups that were either related to or mentioned by the seed channels. The procedure was carried out until an adequate number of channels were identified.

ii Channel/Group relevancy verification: A manual examination of each channel was conducted to ascertain its affiliation with one of the three extremist groups. For ISIS, criteria included the dissemination of propaganda and both official and unofficial support for its cause. For FR, criteria encompassed the distribution of white supremacist content, anti-Jewish sentiment, and Nazi sympathizers. Lastly, for PR, criteria involved promoting and supporting Ukraine's annexation and Ukraine's de-Nazification sentiments. This step involved a qualitative assessment, which is subjective in some cases where the affiliation is unclear.

iii Data collection and iterative channel/group identification: The Telegram official API was leveraged to collect messages and their authors from the selected channels/groups. This step operated concurrently with the previous steps, allowing for an iterative process of data collection, identification of new channels/groups, filtering of irrelevant channels/groups, and finally, data collection from relevant channels/groups. The collected data was subsequently cleaned and organized in preparation for analysis.

iv Bot identification: The difference between bot and non-bot accounts was facilitated by Telegram's requirement that all bot usernames terminate with the word "bot" (e.g., `tetris_bot` or 'TetrisBot') [226]. This enabled the filtering of user data to include information related to messages posted by bots and identify all posts containing a URL referring to a bot.

4.1.2 Networks representation

Telegram is a complex, multi-layered network comprised of users with diverse connection types. These types include direct peer-to-peer communication, public interactions (for instance, between two members of a group), user-to-channel interactions (such as an admin publishing content within a channel), and user-to-group interactions (like a group member posting content in the group). Nodes within this network represent various entities, including users, channels, and groups.

I decided to reduce the network's complexity by employing a bipartite representation. A bipartite network consists of two distinct sets of nodes that interconnect, with no links occurring within each set. In this case, the two sets of nodes are users (including bots) and a combined set of channels and groups. Network links are based on user membership roles like channel admins and bots or participants in groups (see Fig 4.2).

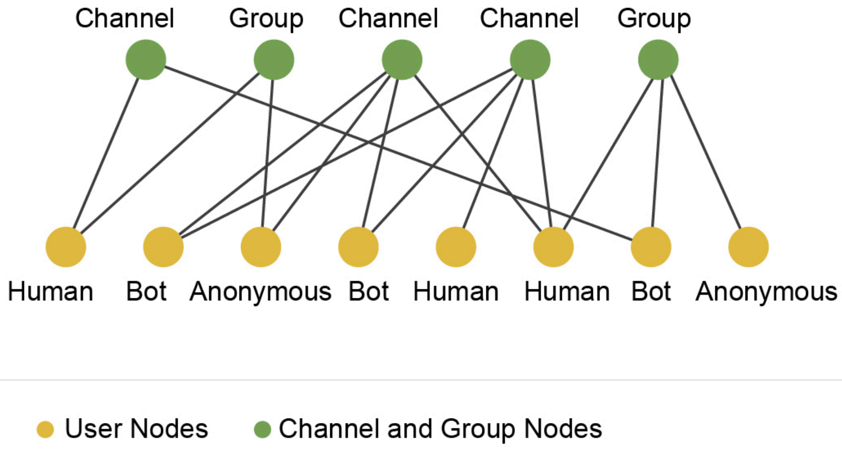


Figure 4.2: Bipartite network representation of user nodes set and group/channel nodes set.

As illustrated in Fig 4.3, the degree distributions exhibit distinct characteristics. For the ISIS network, the degree distribution of user nodes adheres to a power-law behavior, which indicates the presence of hub nodes, representing users, that are interconnected to numerous channels and groups. Nevertheless, the majority of the users are associated with only a limited number of channels. Furthermore, a similar power-law behavior is observed when examining the degree distribution of channels and groups within the ISIS network. While channels with a significantly high degree (indicating a large user base) exist, the majority maintain a moderate to low degree. These observations underscore the scale-free nature of the ISIS network, which is often a sign of organic network evolution.

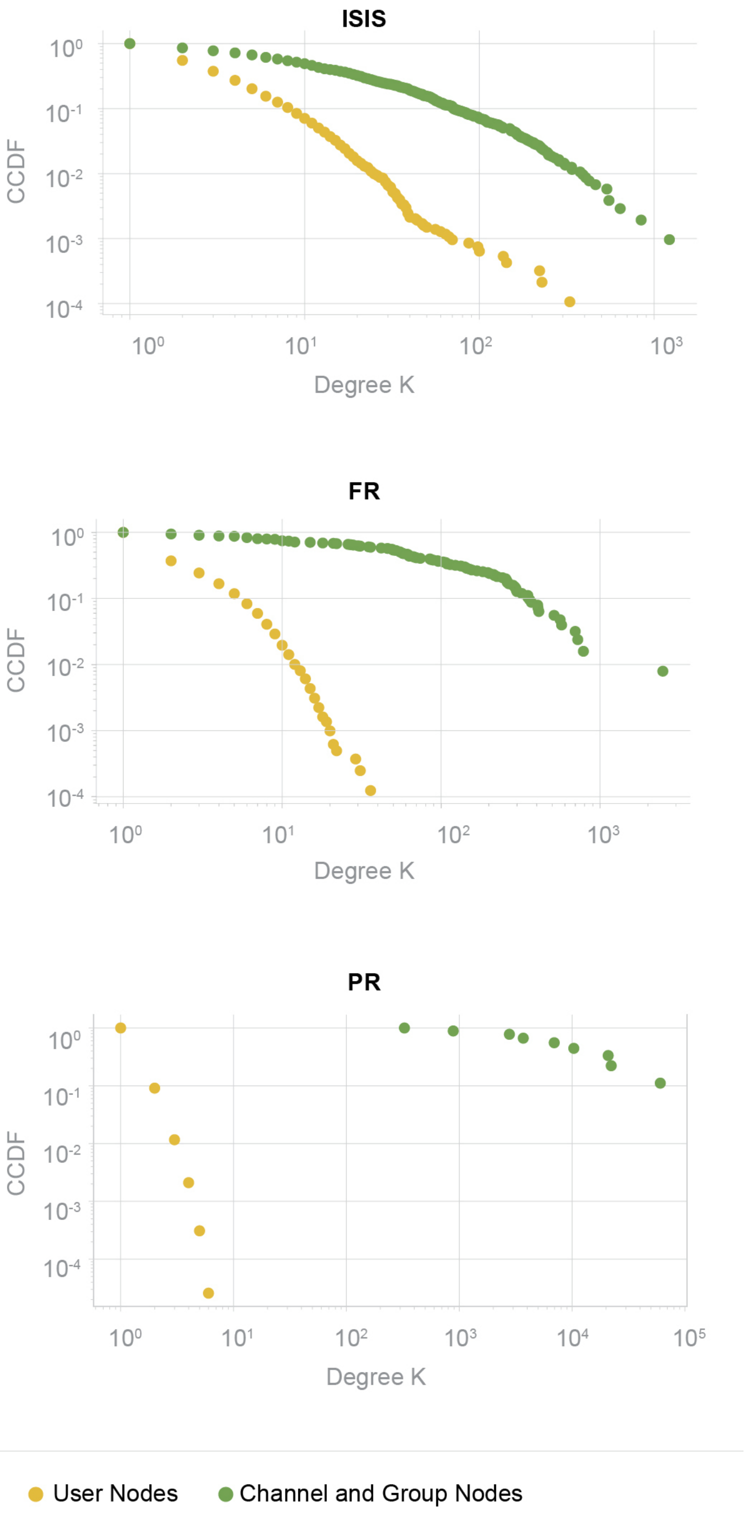


Figure 4.3: Degree distributions of user and channel/group nodes in ISIS, FR and PR bipartite networks.

The degree distribution of channels and groups within the FR network suggests a fat-tailed distribution, aligning with the broader definitions of scale-free systems. Notably, channels and groups manifesting as hubs, characterized by a high user count, coexist with the majority of channels that maintain moderate to minimal user connections. Contrarily, the user nodes within this network follow an exponential degree distribution. This stands in contrast to scale-free topologies, where hub nodes dominate. In exponential networks, as opposed to scale-free networks, the count of nodes decreases more steeply as their degree increases.

For the PR network, user and channel nodes (noting the absence of group entities) present an indistinct distribution pattern attributed to the limited number of channels (n=9), which could indicate inorganic (i.e., engineered) network activities.

I conducted our rich-club analysis on the user projection of the network, focusing on one set of nodes, which include humans, bots, and anonymous users. In reality, this set forms a heterogeneous or multi-layered network due to the presence of multiple node types. In this projection, nodes connect based on shared connections in the original bipartite network. Two nodes in this projection link if they have one or more common channels or groups.

4.2 Findings

4.2.1 Statistical observations

The analysis of each group's activity during the observation period highlighted variances in content sharing, community building (e.g., channel and group creation), and bot involvement. The FR network exhibited the highest messaging activity, with 3,296,205 messages. It was succeeded by the PR network with 2,041,037 messages and the ISIS network with 1,082,260 messages.

Fig. 4.4 illustrates daily messaging activity and other metrics. The FR network published more content, with an average of 2,535 daily messages, surpassing both ISIS and the PR community, which averaged 995 and 956 messages daily, respectively. Interestingly, the PR network showed the highest average of 312 daily active users, followed by FR (236) and ISIS (109). A similar pattern appeared for daily active human accounts, with PR leading and ISIS lagging. ISIS had a notable presence of active bots and anonymous users daily. Moreover, ISIS led with 34 daily active channels, with FR at 22 and PR at only 3. Both ISIS and FR shared similar counts of daily active groups, whereas PR had none.

The PR community stood out regarding new account creation with daily additions of new human and anonymous users. These users joined a few number of channels, indicating targeted activity. Conversely, ISIS saw the highest influx of daily new bot accounts, anonymous users, and channels/groups. These numbers, collectively, suggest divergent growth approaches among the groups.

Fig. 4.5 shows the activity as a function of time. FR demonstrated steady content sharing, with only PR showing a peak in activity before and at the onset of the Russia-Ukraine war. The data suggests that PR's channel inactivity lasted for a significant period, yet there was a notable surge in user additions prior to the war. This may be an organic surge due to the severity of the event, but the concentration of users on a few specific channels likely indicates engineered activity. ISIS maintained consistency in channel creation and activity levels of "active" channels relative to the other groups.

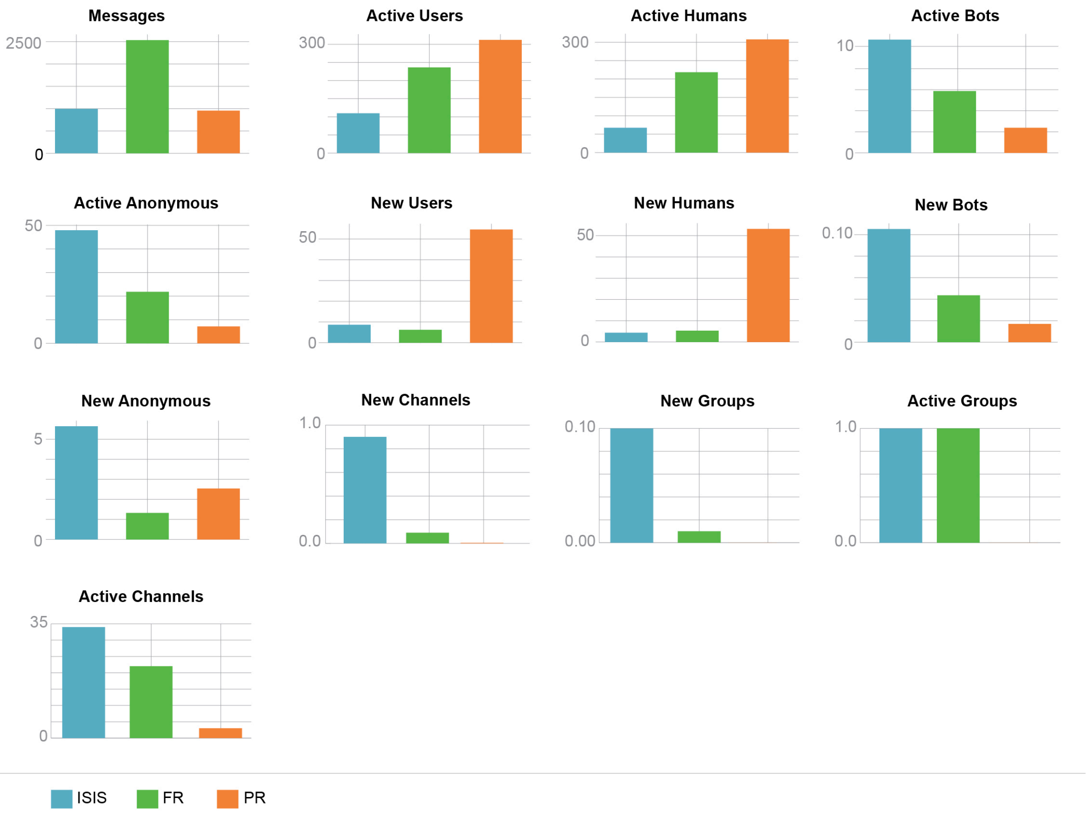


Figure 4.4: Messaging activity statistics on Telegram for ISIS, FR, and PR channels/groups. The graphs contrast various 'daily average' metrics (evident from the bar graph titles.)

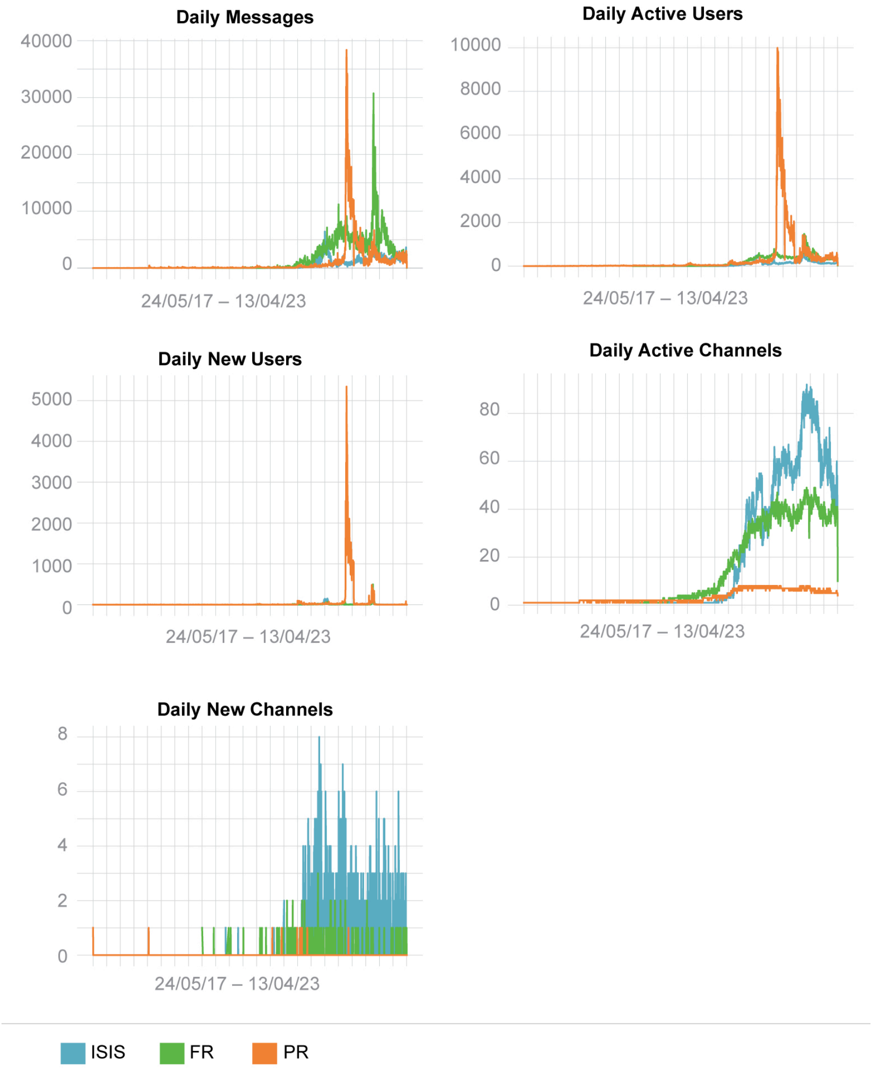


Figure 4.5: Line graphs showing messaging activity trends over time for ISIS, FR, and PR channels/groups on Telegram. The x-axis represents the date range, and the y-axis denotes daily metrics that include messages, active users, new users, active channels, and new channels.

The previous statistical observations highlight the emergence of specific patterns. ISIS frequently establishes new channels, likely as a response to ongoing channel removals by the platform. This group's reliance on bots and its sizable anonymous user base suggests that security is a primary concern in their operations. Despite having fewer channels, the PR network witnessed a surge in user membership, particularly in channels that had been dormant for extended periods. This sudden activity, especially noticeable just before and during the early stages of the conflict, hints at a potentially centralized, orchestrated operation. In contrast, FR operates under less immediate pressure, neither from the platform nor external crises like PR's war. This relaxed stance affords FR the luxury to strategize for the long term, emphasizing content dissemination, community building, and ideological focus.

4.2.2 Rich-club property

The rich-club (RC) property is a feature of complex networks where well-connected nodes ("rich" nodes) have dense connections, forming a close-knit group. This property plays a crucial role in improving the flow of information in networks, as these highly connected hubs help share and spread information. The rich-clubness is a well-defined quantity in network science (see B for more details).

Moreover, inspired by a study on the rich-club organization of the human brain [162], the nodes in the network are grouped into two categories: rich-club (RC) nodes and non-rich-club (NRC) nodes. RC nodes are nodes with a degree greater than k that contribute to the observed rich-club effect. The effect is where the real network's $\rho(k)$ exceeds that of the randomized networks. All other nodes are considered NRC nodes. The RC phenomenon can lead to a hierarchy in the network, where RC nodes have more influence over the network.

The RC property is an organizing principle that has emerged across many social networks [158], and the social messaging platform Telegram is no exception. A normalized RC coefficient with respect to the degree k shows the prevalence of an 'oligarchy' – a group of highly interconnected individuals engaging in reciprocal communication (see B for more details on RC). In the observed networks, as shown in Fig 4.6, the RC property is evident: the normalized coefficient $\rho(k)$ surpasses 1, indicating a denser connectivity among RC nodes than what we would expect from a null model.

RC organizations within these networks have substantial implications for their dynamics. For instance, any content shared within a channel that encompasses a substantial proportion of RC members can rapidly propagate through the entire network, given the multiple channel memberships of these individuals. Furthermore, the members of this rich-club can potentially display a greater degree of influence within the network. Their coordinated actions may lead to significant influence within the broader community dynamics.

However, as illustrated by Fig 4.7, RC nodes constitute a minor fraction of the three networks: 0.34% of ISIS, 0.38% of FR, and 0.21% of PR. In detail, in the PR network, all 17 nodes forming the rich-club are human, with no representation from bots or anonymous users. Contrarily, in the ISIS network, there are four bots (≈ 0.125), 11 humans (≈ 0.34), and 17 anonymous users (≈ 0.53). The FR network consists of 2 bots (≈ 0.06), 26 humans (≈ 0.83), and three anonymous users (≈ 0.09). These numbers underscore a more prominent role of bots within the ISIS and FR networks.

Fig 4.8 presents a comparative analysis of the "Observed Activity Duration" (OAD) of RC and NRC nodes across the networks, grouped by user types. OAD stands for the duration

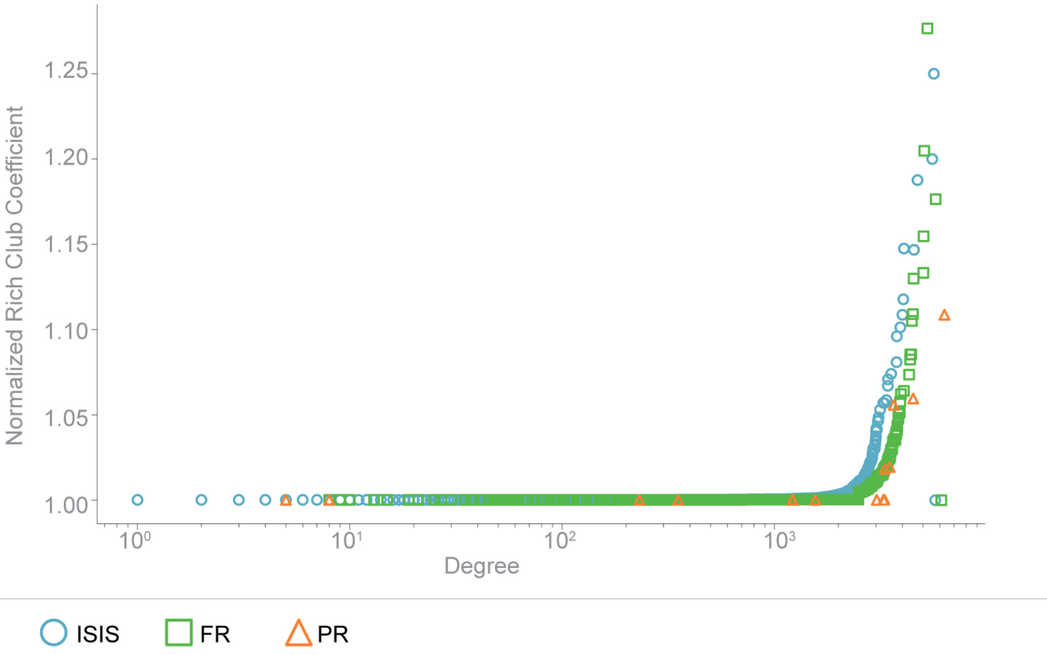


Figure 4.6: RC organization illustrated across ISIS, FR, and PR networks. The plot displays the normalized RC coefficient $\rho(k)$ as a function of degree k. Defined as $\rho(k) = \frac{\phi(k)}{\phi_{rand}(k)}$, where $\phi(k)$ represents the density of subgraphs G_k which only include nodes with a degree greater than k. In contrast, $\phi_{rand}(k)$ represents this value for a maximally randomized version of the network, maintaining the same degree distribution.

between a user's first and last messages. Within the ISIS and FR networks, RC humans and anonymous users display relatively more extended OAD, while RC and NRC bots exhibit potential for prolonged OAD. The PR network contains humans exclusively within its RC nodes, generally with high OAD. Despite NRC nodes demonstrating the capability to persist for substantial OAD, RC nodes are typically observed to have a longer OAD.

This observation illustrates that these central users either adopted the platform early and remained active or have effectively circumvented the platform's regulatory systems. These results also suggest that, in Darwinian terms, the fittest survive the longest, and those who survive the longest join the rich-club. It is also important to note that the rich-club can evolve; a current RC node may become inactive, and a new one may emerge once a certain threshold is met.

Interestingly, Fig 4.9 depicts the networks in which RC nodes within ISIS and FR exhibit limited participation. The total count of channels featuring at least one RC node is 19 out of 585 in the ISIS network and 20 out of 100 in the FR network, yielding respective reach ratios of 0.03 and 0.2. However, the PR network shows a higher level of participation, with 4 out of 7 channels involving at least one RC node, resulting in a reach ratio of 0.6.

These numbers suggest that information dissemination could be faster within the PR network since RC nodes cover 0.6 of the entire network. Conversely, this process might be slower within ISIS due to its decentralized structure and only moderate within FR due to its mixed characteristics. RC nodes could fulfill roles beyond 'information spreading,' such as establishing a 'hierarchy' that exercises authority over the network.

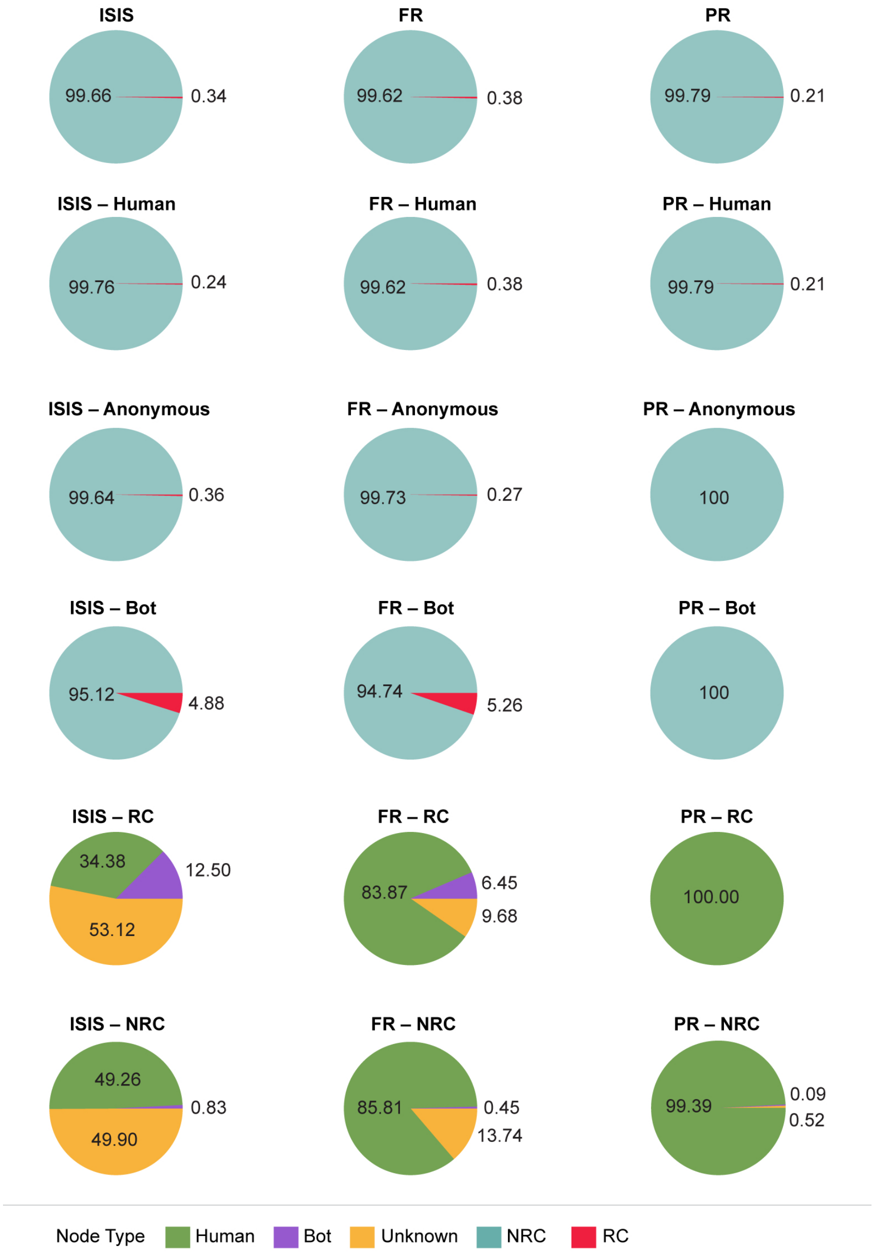


Figure 4.7: The plots display the percentage of RC and NRC nodes in each network per user group and the participation of each group in RC and NRC nodes.

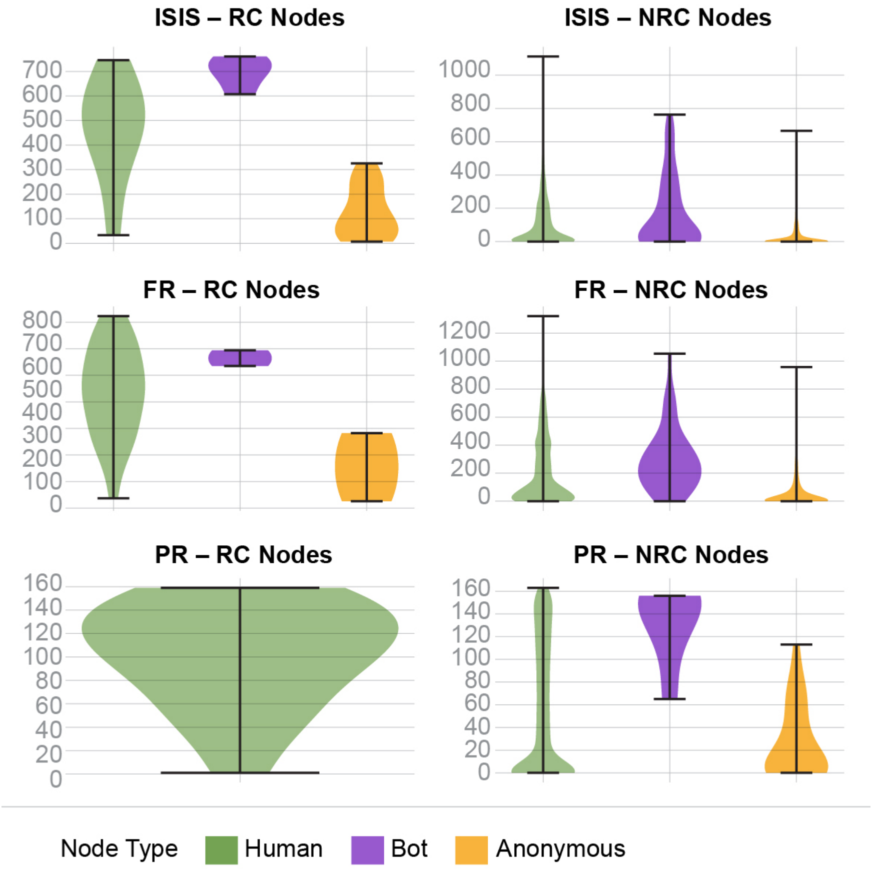


Figure 4.8: Comparative OAD analysis of RC and NRC nodes across user types.

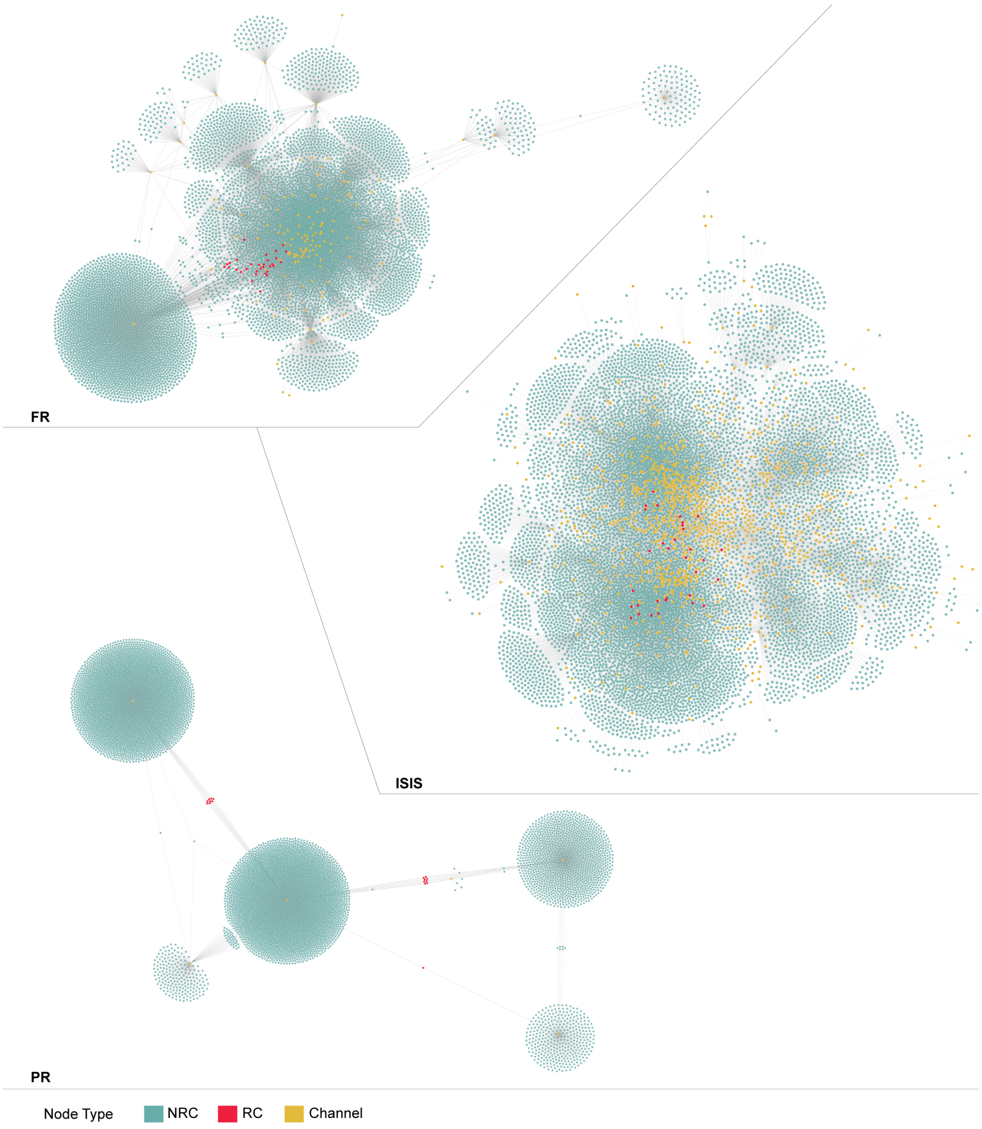


Figure 4.9: Network representation of ISIS, FR and PR channels, rich-club (RC) and Non-rich-club (NRC) nodes.

4.3 Discussion

The three groups' behavioral patterns and network structures suggest distinct strategies for their platform usage. ISIS and PR appear to utilize the platform in a crisis mode. ISIS adapts to survive in an increasingly restrictive environment by adopting a scale-out strategy: they create more channels, employ bots to handle parts of their network management, and their network evolved to function organically, marked by distributed local authorities. In contrast, PR uses the platform to win the information war by implementing a scale-up strategy, evident from the more extensive user base in its channels and a network structure characterizing highly connected nodes with central authority over the network. FR does not operate under any emergency pressure and instead focuses on a long-term plan involving higher content sharing and user recruitment.

The RC property within the networks significantly impacts the dynamics of their information spread. For example, the PR network's RC organization could make it better at message propagation, potentially increasing communication efficiency and disinformation spread. Furthermore, the RC organization might provide a certain degree of resilience to the network. Given their typically extended observed activity, the RC nodes could form the network's backbone, allowing it to resist disruptions and retain functionality over time.

A noteworthy point from our analysis is the role of bots in these networks. Bots are present in the RC organizations of the ISIS and FR networks, While the PR network's RC is exclusively composed of human users. Given their automatic nature, bots could increase the pace of information spread within the network, and this impact could be further magnified if these bots are RC members.

The insights derived from our RC analysis could guide potential mitigation tactics concerning information operations and strategic communication. Considering the crucial role of RC members in disseminating information, focusing interventions on these nodes could effectively control misinformation spread.

NRC nodes still hold significance for network functioning. These nodes play a covert role in the network periphery, where RC nodes lack influence. While RC lacks global authority over ISIS and, somehow, FR networks, NRC nodes maintain local authority within the networks, making this structure more organic and thus robust.

Chapter 5

The Role and Impact of Telegram Bots in the Islamic State's Online Ecosystem

> *"The Imperial need for control is so desperate because it is so unnatural. Tyranny requires constant effort – it breaks, it leaks. Authority is brittle, oppression is the mask of fear"*
>
> Karis Nemik - Andor

To understand and ultimately mitigate the threat from terrorism today, which is increasingly reliant on the massive, multivariate usage of internet-based technologies, researchers, policymakers, and practitioners require new approaches that rely on complexity science [227]. Over the past decade, network theory has contributed tremendously to our understanding of how and why violent extremist communities form and thrive [228, 229, 230, 231, 232]. Among other things, network science has helped to demonstrate that the topological structure of these illicit networks has common features with that of other complex systems and social phenomena [156], even though extremist communities are often treated as a social aberration at a policy level [233]. At the same time, features specific to terrorist networks have been revealed through network science, like the counter-intuitive role women play in making them robust [234], and the relationship between structural network characteristics and the severity of the attacks carried out by the actor in question [218].[1]

In recent years in particular, several valuable efforts have been made to develop and/or apply quantitative tools to identify and analyze complex structures present within violent extremist networks in online spaces. Among these tools are community detection algorithms [235, 236], which can be used to find topologically related clusters reflecting similar interests and activities in groups of nodes in social ecosystems. Using these algorithms along with a range of other approaches, several scholars have attempted to unearth and understand community structures within terrorist networks, tracing and anticipating the dynamics that shape them and locating common characteristics among their members [237].

Today, one of the most important but under-researched aspects of terrorist networking online is the use of bots, something that we explore in this chapter using network science meth-

[1] The results of this chapter are summarized in a paper published in Terrorism and Political Violence [1].

ods. Specifically, we investigate the role of bots as they appear in the context of Islamic State supporter communities on Telegram. An instant messaging app and social media platform, Telegram is favored by the Islamic State [132, 100] (as well as many other violent extremist movements) on account of its broad array of functionalities – which includes anything from content hosting and broadcasting to peer-to-peer chats – and branding focus on user privacy/user sovereignty [224]. Although most of Telegram's offerings are currently unencrypted, users can choose to implement encryption technologies when using its services. Besides its peer-to-peer messaging functionality, Telegram has channels and groups. Channels are one-way broadcasting lists, where only admins can send messages and users can subscribe (and unsubscribe). They can be of two types: public and private. Anyone can join a public channel, but users need an invitation link, sometimes shared publicly, to join private channels. In groups, which can also be public or private, anyone can send a message and interact with other members of the group, whether they are an admin or not.

Bots on social media, including but not limited to Telegram, are best understood as automated accounts that execute specific tasks such as publishing, sharing, and resharing content [4]. Anyone can build and deploy a bot: all that is required is a sustainable and open Application Programming Interface (API) access. In recent years, bot networks have been used increasingly across both mainstream and lesser known platforms (including Telegram) to promote products and services [238], spread misinformation [136] and low-credibility content [6], probe algorithmic and political bias [239], manipulate elections and public opinions [19], and mitigate the challenges posed by violent extremism [240]. As this study demonstrates, in the context of the Islamic State and its supporters' activities on Telegram, bots generally perform one of three key functions: publishing content, moderating discussions, and acting as gatekeepers. In this capacity, they play a central, lubricating role in amplifying the movement's ideology and cultivating its community of sympathizers, automating administrative tasks like blocking users that violate group policies, and permitting new members to join.

Below, drawing on 1,215,850 data points that were collected from Telegram between February 1 and September 30, 2021 via an ingest program, we study the activities of the Islamic State's "terrorist bots" within the community dynamics amidst which they operate. We map out their interaction network and, in addition to presenting a schema of their activities and impacts, we sequentially apply community detection algorithms to the data with a view to determining the extent to which they operate in a structured or unstructured manner. These methodologically distinct approaches parse the network's underlying structure by dividing its nodes into communities. All of the applied methods show that the structure is by no means randomly distributed but, rather, made up of clusters (or modules) of closely interconnected nodes. Our analysis of the network's modular structure and clustered activities implies the existence of a hybrid system of functional groupings that have been proactively, and collectively, developed to augment the Islamic State's presence in Telegram channels and groups, as well as a spontaneous process of unorganized supporter-generated community formation. Based on these findings – which speak to the flexibility, ease, and effectiveness with which bots can be deployed to further the interests of bad actors – we contend that the allowances that Telegram makes for bot development are a central factor driving the Islamic State's years-long preference for it over other platforms that are demonstrably more secure. This is in spite of the (valid and widely implemented) assertion by Telegram in its FAQs that "we do block terrorist (e.g. ISIS-related) bots and channels" [241].

The chapter proceeds as follows. First, we set out the data collection and analytical methodology, explaining how we collected the data on the bot network and how we processed and interrogated it. After that, we describe the overarching characteristics of the network, touching on what functions these bots performed, how frequently they were active and for how long, and in what language they operated. We then present the findings of the community detection analysis itself; in it, we explore the clustered and utility-driven topology of the network. We conclude by weighing up the implications of this study and suggesting further avenues for research.

5.1 Methodology

This section gives an overview of both the data collection process and the analytical methodologies that were used to interrogate said data.

5.1.1 Collecting the data

In order to better understand the role and impacts of bots deployed in support of the Islamic State on Telegram, we collected an original dataset from a manually selected community of 3,940 public groups and channels that were considered to be either controlled by or supportive of the Islamic State. The data collection period was February to September 2021. We selected these groups and channels for analysis because they were explicitly and proactively aligned with the Islamic State. A group or channel was deemed to be "explicitly aligned" if, on a sustained basis, its users' focus was on either news or ideological matters relating to the Islamic State. Other overt indicators of pro-Islamic State orientation that were taken into account include: re-posting or sharing of official media, the creation of unofficial pro-Islamic State media, or overt declarations of support by group administrators for the movement's mission, goals, activities, and operations [100]. Once selected for inclusion, each group or channel was tagged according to the topic it prioritized. In total, we applied 16 tags to the dataset, meaning that the groups and channels generally revolved around one of 16 issues or spheres of activity. These are: general commentary; news; Afghanistan; anti-Shi'a; al-Hol/Roj; Kashmir/India; information security; links sharing; media campaigns; medical advice; nashids; content distribution; networking; theology; tactics, techniques, and procedures (TTPs); and women's affairs. The largest category by far was 'general commentary,' which describes groups and channels (but mainly the former) in which anything from conflict news to theology is discussed and both unofficial and official propaganda materials are shared. It is important to state that this is not a complete sample of all Islamic State channels and groups on Telegram, however, it does serve as an extensive representation of the pro-Islamic State community on the platform.

5.1.2 Cleaning the data

Once we had selected this sample, we accessed and archived all the publicly available messages shared by these groups and channels – programmatically stripping away all metadata besides the unique (and randomized) identification numbers associated with accounts that published on them. Next, we requested user objects from Telegram's API using the methods "users.getFullUser" and "users.getUsers." This provided us with the usernames of all the active accounts in the network that had opted to make their user details listed on Telegram's public directory when signing up to the platform. In total, we requested the usernames of 16,682 accounts through Telegram's API, recognizing that the vast majority of them would not be publicly listed – and, indeed, only 258 were. Determining which of these 258 publicly listed accounts were bots and which were not was simple, because Telegram mandates that all bots have usernames that end with the word "bot" (such as `tetris_bot` or 'TetrisBot') [226]. This meant we could filter the user data to include only information related to or messages posted by bots as well as identifying all posts that contained a URL referring to bot. In total, we identified 106 accounts in this network that were bots. Having ascertained this, we discarded all other user data and collected all names and usernames directly associated with these bots, including those of the groups and channels in which they had been most active. We also collected all dates and times of posts and relevant data about content type (i.e., was it a JPEG, MP3, MP4, or PDF). While some of these bots are Telegram natives, the majority of them were compiled by self-appointed Islamic State khuruq al-buwatt ("bot creators"), who advertise their bot development services as part of their support for the broader extremist community. (Interestingly, these self-described khuruq typically label bots as hudhud, which is a reference to the Islamic scripture; the hudhud was a bird which served under the Prophet Solomon and supported him by collecting information about his enemies) [242].

5.1.3 Constructing the network

In order to determine how, and to what end, these bots were deployed by Islamic State supporters during the data collection period, we built a network map capturing their actions and interactions in the ecosystem. The nodes of this bipartite network, which is visualized in Figure 5.1, are bots (blue) and channels and groups (green and yellow respectively). Although this network is bipartite, the platform architecture is multilevel, due to the ability to link channels to groups, where each new post from a channel can automatically be forwarded to a connected discussion group by a so-called "discuss link" [243]. The edges are directed links representing messages posted by bots in said groups/channels or mentions of these bots in said groups/channels. Although the edges represent two types of linkage, we opted to treat them as one for the sake of simplicity (given the unwieldy nature of network maps that are both bipartite and multilayer). We placed a link between a given bot and group or channel if the bot was found to have posted one or more messages in the group or channel in question. When this happens, it indicates that the bot is a group member or channel admin. Bots usually post far more than one message, but, for the purposes of this aspect of the study, we opted to ignore the weight of the link and count only the existence of the connection, not its strength.

5.1.4 Identifying communities

In network science, communities refer to parts of a connected network within which nodes are linked to each other more than they are to the rest of the network. Such communities are generally understood to have a functional role and their efficient identification is a major scientific challenge [244]. Communities can be disjunct – in which case finding them means to identify a partition of the network – or they can be overlapping, leading to a so-called cover of the network [245]. Considering the multitude of tasks carried out by bots on Telegram, it is a natural question to ask whether, in this network of bots, channels and groups, we can identify whether the implementation of these tasks is organized, or whether its topology

To reiterate, we define the word "community" as it is customarily understood in network science—i.e., as a topological entity based on the connectedness of its constituent nodes.

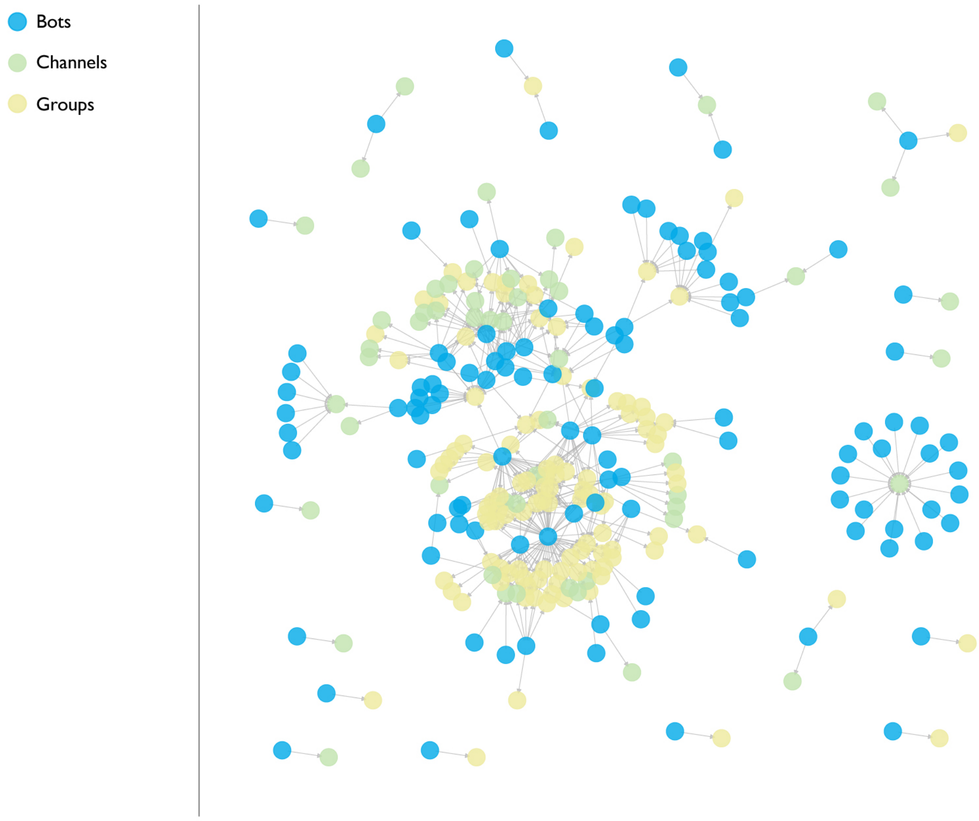


Figure 5.1: The visualization of the bots-groups-channels network. The blue nodes are bots, the green nodes are channels, and the yellow nodes are groups. The network consists of 241 nodes and 346 links. 106 of the nodes are bots, 98 are groups, and 37 are channels.

is decoupled from the respective functions of the bots. To study this problem, we used six different community detection methods to parse the network. We present and compare the outcomes of this analysis in the Appendix (C). While the methods yield somewhat different communities, common features emerge indicating robust topological structures. To better illustrate these features, we use the partition resulting from the Girvan-Newman (GN) algorithm [246] (see Figure 5.4 below), which represents them in the clearest way and leads to the most meaningful community structure. (The GN algorithm is known to be a powerful method, though only when restricted to relatively small networks, which makes it ideal for the present context.)

5.2 Findings

The 106 bots we identified via Telegram's API were either active in or mentioned on 98 distinct groups and 37 distinct channels. Their level of activity was significant: collectively, they posted some 39,211 messages. Most of these were published in groups (33,487 in groups and 5,724 in channels).

The vast majority of the content the bots published was text (36,715 posts of the 39,211 we collected). After that was image files (1,748), PDFs (315), video clips (185), audio files (90), and emojis and links (158). From a linguistic perspective, the bots mainly posted in Arabic (29,605 messages in total). Besides that, there were 4,194 posts in English, 2,123 in Urdu, 922 in Farsi, 147 in Soko, and 45 in Bahasa-Indonesia, and a smattering in other languages. Figure 5.2 shows the percentage of the type of content and the top languages.

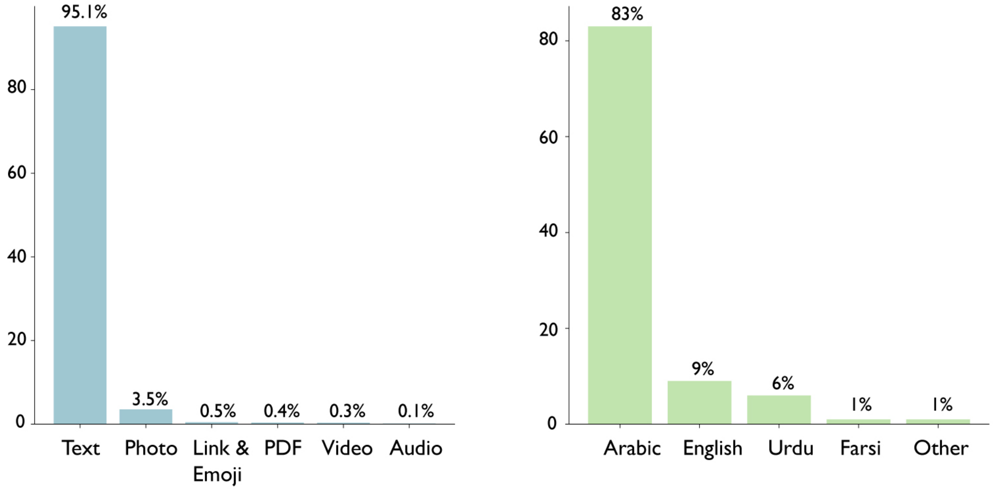


Figure 5.2: The percentage of content types published by the bots and the language distribution of the content. Left: The percentage of content types published by the bots. Most of the content is text; the rest comprises of image files, PDF files, video clips, audio files, and RAR files and emojis. Right: The language distribution of the content. Most of the text content is Arabic, then English, Urdu, Farsi.

The bots' respective lifespan was highly variable. By the end of the data collection period, none of the non-Telegram native bots were active anymore, with some having been censored by Telegram and others having been shut down by their administrators. The longest bot lifespan was 213 days, and the shortest was just one day. On average, they were active for around 18 days. Naively, one would expect a linear dependence between the bots' lifespan and their cumulative activity. However, Figure 5.3 below shows that there are a lot of fluctuations in the activity, resulting in a correlation coefficient $R = 0.65$. Moreover, the median and mode of the lifespans of bots are both about one day, which suggests that the lifespan of individual bots is typically short but, due to some bots having longer periods of activity, the system as a whole is surprisingly robust. Given what we know about Telegram's disruption policy when it comes to the Islamic State, this suggests that longer-life bots engaged in acts that were less explicitly or overtly supportive of the movement (i.e., activities other than media distribution).

Notably, there were several instances in the data of the removal-recreation cycle that characterizes how pro-Islamic State communities respond to Telegram's moderation efforts. For example, the `@UrduNashir_22bot`was removed and recreated half a dozen times during the data collection period, each time coming back with a different but still recognizable username (e.g., `@UrduNashir_24bot`, `@UrduNashir_27bot` and `@Urdu_nashir28bot`). This means that, generally, the removal of individual bots did not explicitly cause sustained disruption to the network, which was able to function in spite of Telegram's censors.

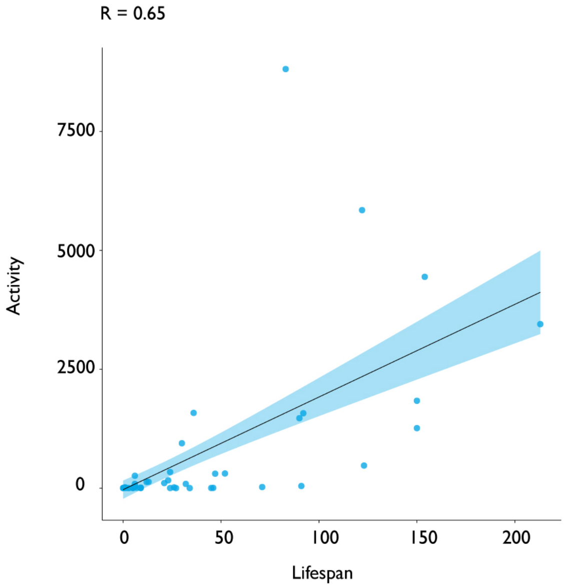


Figure 5.3: Pearson correlation coefficient R shows a positive relationship between the lifespan of the bots and their activity.

On average, the bots posted 176 messages each day (140 text, 13 images, 12 PDFs, four video clips, and two audio files). For reasons that are not immediately clear, they were most active on 2 August 2021, when they collectively posted 1,431 messages. By contrast, on a not insignificant number of days, they were almost completely inactive, posting just one message per day during such periods.

From the perspective of the range of their activities, some bots were specialized and others were multifunctional. Most bots were associated with 'general commentary' channels and groups; of the 86 that operated in that context, there were 26 media amplification bots, 24 links sharing bots, 19 media activism bots, and 17 news-posting bots.

5.2.1 Analyzing the communities

The bots-groups-channels network in question, as visualized above in Figure 5.1, consists of several components: small ones, consisting of two or three nodes; one medium size one, the shape of a star; and one large connected component containing the majority of the nodes. Below, we restrict our analysis to this largest component, which dominates the network. It is not clear *a priori* what governing principles could be behind this community structure. One possibility could be that partition would lead to its splitting into clusters of bots and clusters of groups/channels. This is because, at base, it is a bipartite and nebulous network (i.e., there are no direct connections among the groups and channels or the bots). Another possibility could be that the bots would get positioned randomly in the matrix of channels and groups, irrespective of their functions. On applying community detection algorithms to the data, however, we found that the communities we detected are in fact diverse and different but highly structured, containing an array of groups, channels and bots. Indeed, once parsed by the GN algorithm, we see two well separated communities that make up the structure of the network (see Figure 5.4).

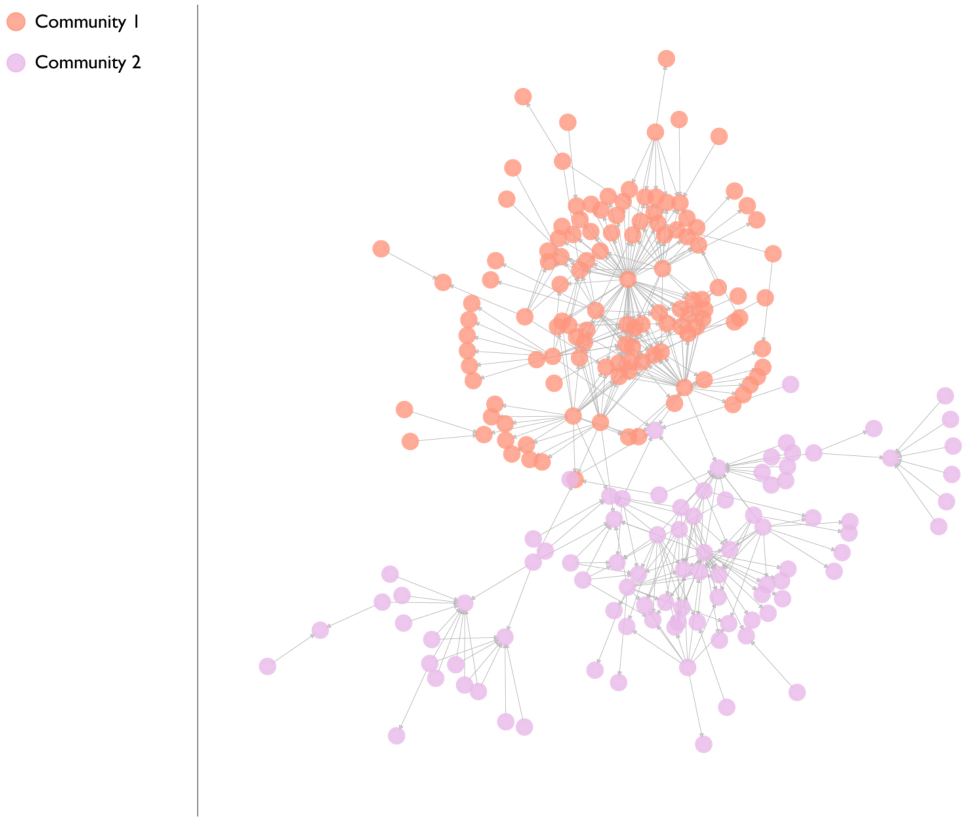


Figure 5.4: Visualization of the communities in the largest connected component of the bots-channels-groups network, as produced by the GN algorithm. It was partitioned into two parts with communities of similar sizes 93 (orange) and 125 (pink).

One hundred and twenty five nodes belong to Community 1, marked in orange, and 93 belong to Community 2, marked in pink (see Figure 5.4). Interestingly, while they are similarly sized, the composition of these two communities is rather different: 44 nodes in Community 2 are groups/channels and 49 are bots; on the other hand, 27 nodes in Community 1 are bots, and 98 are groups/channels.

Role analysis

Across the two communities, we identified two types of bots, with three core functions emerge: content distribution, basic group administration such as blocking spammers and deleting messages, and gatekeeping (i.e., allowing users to join and sharing links). Admin bots mostly perform administrative functions, including discourse moderation, link sharing, and gate keeping. Content bots engage directly with groups and channels, sharing content and exchanging information with group members. This latter grouping is clustered into many different clusters, with different bots connected with different channels or groups that publish different content types.

We manually tagged each bot as either being content-focused or admin-focused. Some bots can be both, but we tagged them by the majority of their activities. Community 1 has 20 admin bots and seven content bots and Community 2 has 42 content bots and seven admin bots (see Figure 5.5). Notably, the number of channels/groups in Community 1 is twice as big (n=98) as the number of channels/groups in Community 2 (n=44). By reviewing the content that the bots published in each community, we were able to prove the validity of the overall partitioning process. That is to say, the two communities identified by the algorithm differed significantly when it comes to the amount and type of content they each posted. Community 1 published some 30,251 items, with Community 2 publishing just 6,284 items, even though it is only moderately smaller in scale. Notably, the content shared by Community 2 is significantly more diverse. Specifically, it shared 4,021 text-based posts, 1,562 images, 315 PDF files, 168 video clips, 158 links and emojis, and 60 audio files. In contrast, Community 1 overwhelmingly published text-based items; in total, it shared 30,073 text-based posts and just 135 images, 29 audio files, and 14 video clips. (See Figure 5.6 below). This suggests that Community 1 is focused more on facilitating/moderating chatter between group members than it is sharing content (official or otherwise), something that is broadly left to Community 2.

The two communities also differ linguistically. Figure 5.6 shows that the language of the text corpus in each community differs significantly. The bots in Community 1 overwhelmingly published Arabic messages (26,533), and significantly less in English (2,831). In contrast, the bots in Community 2 published messages in Urdu (1,781), in Arabic (1,328), in Farsi (823), and in English (486).

Each and every bot, group, and/or channel in these two communities serves a specific purpose for the broader ecosystem. In both communities, we found that their operations were trifurcated in a way that spoke to both parallel missions and targeted audiences: general commentary on the Islamic State, often grounded in distribution of and discussion around content produced by its official media apparatus; media materials that have been translated and/or engineered to respond to/support the interests of local pro-Islamic State communities in specific locations or contexts (like ISKP outreach networks in Afghanistan and IDP camps in Syria and Iraq); and efforts geared towards facilitating the activities of the Islamic State and its supporters online (like information security advice and links to specific technical services).

Importantly, while they are structurally different, the two communities we identified do not exist in two entirely separate spaces. They engage with the same audience segments, but, due to their functionality and nature, have evolved to serve different but complementary

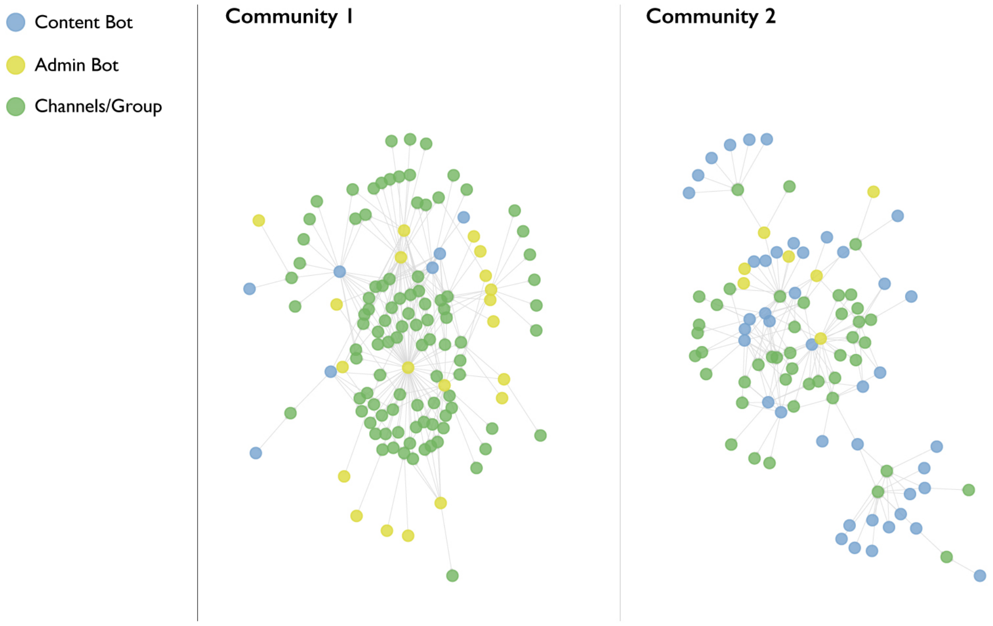


Figure 5.5: Content bots versus admin bots. The graph shows that Community 1 has more admin bots than content bots and the admins are connected to more channels than in Community 2, which has significantly more content bots than admin bots.

purposes. That being said, their overarching audience and orientation is the same – aiding and abetting people who share and support Daesh's mission either online or offline, locally or internationally. The efficient and clearly resilient recreation cycle that characterises the Islamic State's presence on Telegram today, not to mention the persistent presence of bots within it, speaks to their widespread utility and impact.

Core-periphery analysis

Demonstrably, the network is not organized by single person or controlling authority but formed incrementally, at least partly in a self-organized manner, by loosely connected groups of Islamic State supporters. This corresponds to the broader decentralized character of the Islamic State's influence activities, a network quality that has been cultivated because it makes exceptionally difficult to localize and break down the ecosystem in a sustained and effective manner.

To identify the mesoscale structure of both communities and the position/role of the bots and channels/groups in this structure, we used Rombach's Algorithm to analyze the core-periphery (C-P) structure of both networks [247]. The algorithm is a modified version of the continuous Borgatti-Everett algorithm [248], an improvement to detect multiple cores in the core-periphery structure in networks (See C for the mathematical details of the algorithm). Simply put, the structure comprises a few core nodes and many periphery nodes, where the core is composed of densely interconnected nodes, and the periphery comprises sparsely interconnected nodes.

The communities exhibit C-P structure as a result of a partly spontaneous process (see Figure 5.7). The coreness of both communities is very close (the coreness of Community 1 and Community 2 is ≈ 0.30 and 0.29 respectively). Community 1's core nodes are made up of

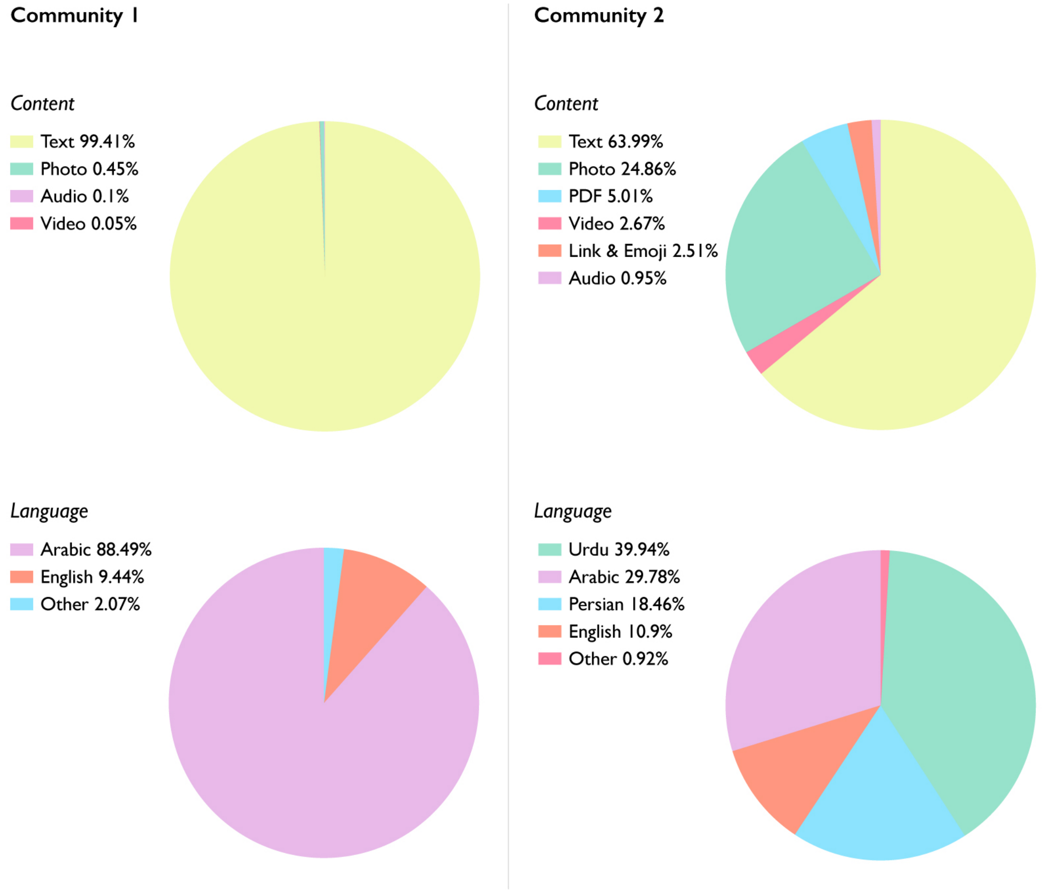


Figure 5.6: Type and language of content shared by Communities 1 and 2. It shows that Arabic text dominates Community 1 while Community 2 has more diverse content types and mixed languages.

many admin bots in the core. Indeed, the node with highest coreness is an administrator bot. This reflects our observation that Community 1 has the task of assuring information flow. Community 2, by contrast, has only one admin bot at the center of its core; the rest of its core nodes are content bots, which implies, as per our initial observation, that this community's main role is content distribution.

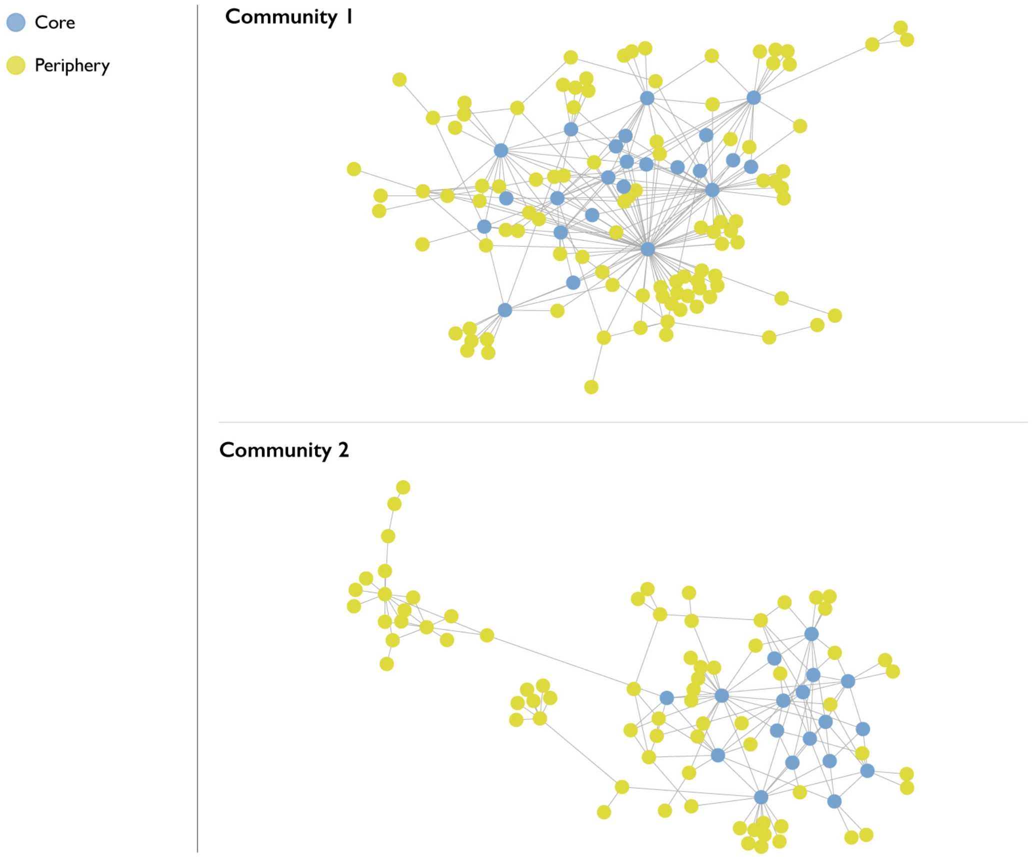


Figure 5.7: Network visualization that indicates the core and periphery nodes in both communities.

The different core-periphery structures and, as part of that, the proportion of admin bots over content bots, speaks to differences in how bots operate when they are focused on information flow over content distribution. The former involves keeping the network available (and 'clean') to facilitate the latter. These bots are focused on making the content supply chain efficient; they do not produce or provide the goods, but they help to keep the distribution running (it is about control). The latter is all about production and provision, sharing content with as many members of the community as possible (it is about participation). Inasmuch as that is the case, and although this splitting of responsibilities is not clean-cut, the two communities share a common objective while having distinct roles.

In any case, the core of both communities is mainly comprised of bots, with their peripheries made up of a mixture of channels and groups. Based on that structure, one could argue that the core-periphery structures emerged because of constant activity from the core in maintaining content distribution efforts and chat moderation, which was being conducted in the periphery, assuming that the core–core relations are not as important as the core–periphery relations.

The network has two main attributes: i) core-like nodes comprising admin bots that control and moderate the activities of groups and channels; and ii) facilitating general commentaries and exchanges between Islamic State supporters, with less emphasis on content distribution. These attributes together imply that its principal impact (and, consequently, utility) corresponds to its ability to maintain the activities of the broader ecosystem of supporters.

The contrasting structure of each network suggests that different bots serve different purposes and that, in doing so, their roles and applications have iteratively evolved alongside the communities they serve. In other words, the way these bots (or groups of bots) operate and/or are applied has helped to shape these networks in a manner that best serves their specific purposes. In community 1, one core can be seen dominating the network, which shows a level of discipline; the other community, on the other hand, has multiple cores, which speaks to a measure of flexibility and scope for equal engagement.

The evolution of these two communities, which has left one highly centralised and the other more multipolar, has a rough corollary in the Islamic State's own evolution as a political movement in recent years. Its shift from territorial protostate to covert insurgency in its core territories of Iraq and Syria, in tandem with sustained pressure from the online platforms it once favoured, forced the munasirin community to transform from a model of organic but acute centralisation to one characterized by multipolarity. That is to say, in its early days, both the Islamic State and its community of supporters prioritized discipline and control online. However, with its decline and ultimate territorial collapse, discipline and control were no longer feasible and were instead swapped for an approach that was more in line with the characterisation the second community we detected. Importantly, these bots must be understood as being part of a bigger picture, not the picture itself; they alone do not constitute the ecosystem. Rather, they help to construct and curate its constituent networks (thus communities) and – by pushing the limits of scalability of human-made networks – they serve to augment and strengthen its operations. Inasmuch as this is the case, the bot network is a mirror of the larger human community within which it operates.

We can be confident, given the conditions within which it exists, that both the bot network and the human community have developed evolutionarily, responding opportunistically to the stresses, strains, and obstacles they have faced over the years. Due, however, to a lack of historical data, it is not possible to empirically trace the scale of that evolution – at least, not as part of this study. That being said, the communities in this network are not completely disconnected. On the contrary, they are very much interlinked. However, each community's respective nodes are more connected with each other than they are with the nodes of the other community. On that basis, we believe these two groupings have evolved out of one original network entity.

5.3 Conclusion

On Telegram, anyone can create and manage a bot. This means that users can deploy bots to do most of the activities that human users can: post content, receive messages, manage groups, provide services, and even accept payments. Unlike human users, though, bots can be programmed to be active, and immediately responsive, around the clock. This makes them ideal for tasks that are otherwise labor-intensive or tedious. Indeed, they are in many ways better-placed to manage groups and share content because they can do so at large scale and with immediacy.

Bots can be used for good and bad; they can be truly harmful when used by violent extremists. Indeed, as we have shown above, for the Islamic State – similar to many other violent extremist groups – bots are being used to lubricate and augment influence activities, including facilitating content amplification and community cultivation efforts. They are standing in for official Islamic State operatives and advocates, connecting people with the movement based on common behaviors, shared interests, and/or ideological proximity while minimizing risk for the broader organization.

As we have shown above, the bots we identified during the data collection period are mainly clustered around two communities of Islamic State supporters, with one group seemingly focusing on facilitating discussion and exchange, and another augmenting content distribution efforts. Crucially, as things stand, they are not 'intelligent.' Indeed, we did not identify any that use language models or intelligent systems to interact or generate content, even though such a capability is technically possible. Instead, these bots deploy basic scripts to execute specific – and simple –tasks. Notwithstanding their simplicity, though, they are clearly and fundamentally embedded within the Islamic State's online ecosystem. Together, they directly facilitate its survival, amplifying and minimizing risk when it comes to the spread of content and, more importantly, keeping the consumers of this content within reach of it –something that has become increasingly difficult to do in recent years due to Telegram's highly aggressive and well-targeted disruption efforts.

Understanding how these bot communities operate and, more to the point, on what basis they are structured is critical if their efforts are to be meaningfully and sustainability undermined any time soon. On that basis, future research would do well to build on these findings, perhaps studying temporal aspects of the network with a view to tracing its evolution and robustness. Another option for further study would be to address one of the major limitations of this study: the fact that that, due to ethical considerations, content posted by human users is missing from the analysis, and only content posted by the bots as out-of-context conversations is explored. Analyzing how bot content overlaps and interacts with human user content is a logical and necessary next step for exploration.

Chapter 6

Conclusion

"The future is already here – it's just not very evenly distributed."

William Gibson

This thesis highlights that social bots are integral to online social networks (OSNs), coexisting with humans to form a dynamic ecosystem. Bots play roles from moderating communities on platforms such as Reddit and Telegram to contributing to Wikipedia articles, interacting personally on Facebook Messenger, and shaping Twitter trends and user engagement.

My research reveals network structures that enables bots' roles and impacts. Social bots activate local structures on Twitter, form communities within platforms like Telegram, occupy network cores and are part of the rich-club organization of networks.

Chapter 2 identifies the emergence of local structures, known as 3-node motifs, in networks that feature interactions between bots and humans. These structures reflect user targeting, message and relationship propagation, and user engagement patterns in three different strategies: Trend-Targeting (TTS), Keyword-Targeting (KTS), and User-Targeting (UTS). The findings show that these strategies produce different local structures with varying significance levels. Additionally, we discovered that content-related strategies (TTS and KTS) activate similar local structures, while a user-oriented strategy triggers simple and complex motifs.

This study draws attention to a local structure wherein two nodes (A and B) uphold a relationship, often a 'friendship' on the platform, while a third node (C) connects solely to one of them (A). This structure sets the stage for potential future engagement between nodes C and B, directly or indirectly through node A. This inherently dynamic local structure can evolve and morph into different structures. For instance, node C might eventually form a direct relationship with node B, with node A as a bridge, forming a triangle. Alternatively, node C could relay information to B through A. Such indirect relationships can wield as much influence within an information cascade as direct ones.

Chapter 3 reveals that extremist networks on Telegram employ diverse strategies on the platform, reflecting their goals and the environment's dynamics. While all exhibit a specific network structure, known as the rich-club property, the roles of rich-club nodes vary among these networks. One network might centralize authority, while another may have distributed local authorities. Bots in these networks have dual roles: as members of a rich-club with assumed global authority over the network, and as participants in independent local authority managing parts of the network.

Chapter 4 illustrates that an ISIS network of bots and channels on Telegram can be categorized into two distinct communities. Within these communities, we identified two types of bots, each with three core functionalities: content distribution, basic group administration—including spam blocking and message deletion—and gatekeeping, such as controlling user access and sharing links. Both communities primarily consist of bots at their core, with their periphery populated by channels and groups. This arrangement suggests that the core-periphery dynamic emerged from the core's continuous activity in content distribution and chat moderation, which is then reflected in the periphery. It is presumed that the core and periphery relationships hold more significance than intra-core relationships.

While this thesis focused on the emerged structures in networks with bots and humans or exclusively bots, it omitted human-human interactions. By recognizing the importance of such interactions, future studies should investigate the influence of bots on human-human dynamics. Lastly, our representation of these networks does not incorporate the temporal traits of links. However, exploring the temporality of these networks can help unravel the dynamics driving these network formations.

Appendix A

Supplementary Material of Chapter 3

A.1 Action Model

As detailed in Table A.1, each bot, irrespective of the strategy employed (TTS, KTS, or UTS), uses one of two primary action probability models. In Model 1, the actions (retweet, like, follow, reply, and mention) are treated uniformly and assigned a probability value of 0.2. Conversely, Model 2 discriminates these actions based on the relative frequency of their occurrence among users. Specifically, actions 'retweet' and 'like' receive an augmented probability of 0.3, whereas 'reply' and 'mention' are attributed a reduced probability of 0.1. The 'follow' action is assigned a probability of 0.2.

The purpose of this diverse probability allocation is to enable a thorough examination of the influence of the action model on the bot's performance. Furthermore, it aspires to explore the conjecture that the effectiveness of a bot may significantly vary depending on the overall strategy or the specific action settings - a topic for future investigation.

Bots adhere to these probability models, supplemented by a set of constraints that modulate their activity rate, ensuring a human-like engagement pattern. These constraints restrict the bots to a maximum of two actions within an interval of one to two hours, and limit their self-initiated tweets to a range of 100 to 150 per day.

Moreover, TTS bots, beyond the basic probability rule, scan the worldwide trending list and engage with one or two of the top tweets within the selected trends. In contrast, UTS bots focus on specific users, scanning their timelines over the past 7 days and interacting with the top-ranking tweet from this period (for a comprehensive understanding of these bots' design, refer to the provided code in the following GitHub repository).

A.2 Measuring the Significance of Motifs

Many statistical measures have been developed to measure the significance of motifs. In this analysis, we used Z-score and normalized Z-score [249]. We chose the later method of calculation to compare between strategies, as we treat them as different classes of networks [184].

For a given motif M, a Z-score can be calculated as:

		TTS		KTS		UTS	
Strategy		Bot 1	Bot 2	Bot 3	Bot 4	Bot 5	Bot 6
Action probabilities	Retweet	0.2	0.3	0.2	0.3	0.2	0.3
	Like	0.2	0.3	0.2	0.3	0.2	0.3
	Follow	0.2	0.2	0.2	0.2	0.2	0.2
	Reply	0.2	0.1	0.2	0.1	0.2	0.1
	Mention	0.2	0.1	0.2	0.1	0.2	0.1
Number of actions in each interaction		1-2	1-2	1-2	1-2	1-2	1-2
Daily tweets range		100-150	100-150	100-150	100-150	100-150	100-150
Interaction time interval		5-8 minutes	5-8 minutes	5-8 minutes	5-8 minutes	5-8 minutes	5-8 minutes
Action time interval		1-2 hours	1-2 hours	1-2 hours	1-2 hours	1-2 hours	1-2 hours
Top tweets range		1-2	1-2	1-2	1-2	2-3	2-3
Trend's range		1-5	1-5	-	-	-	-
Last days		-	-	-	-	7	7
Top last day's tweet		-	-	-	-	1	1

Table A.1: Comparison of action probability models and operational parameters for TTS, KTS, and UTS bots. Note: An interaction from bots can consist of two actions. For example, a bot might tweet and then, after an interval of 5 to 8 minutes, like the same tweet. These two actions together are considered a single interaction.

$$Z_M = \frac{Nreal_M - < Nrand_M >}{\sigma(Nrand_M)} \tag{A.1}$$

In this equation, $Nreal_M$ is the number of motifs in the empirical network, and $< Nrand_M >$ and $\sigma(Nrand_M)$ are the mean and standard deviation of the number of motifs in the randomized network, respectively. The significance is calculated by measuring its frequency compared to the average frequency of the same motif in a 1000 random networks (configuration model). The motif is over-represented if the Z-score Z_M has a high positive value, and under-represented if it is negative with high absolute value [250].

We then used the vector of Z-scores to calculate the normalized Z-score:

$$NZ_i = \frac{Z_i}{(\sum Z_i^2)^{\frac{1}{2}}} \tag{A.2}$$

The normalized Z-score measures on the relative importance of the motifs by creating a common scale between 0 and 1 and the degree of significance increases as the score tends towards 1. We used the relative Z-score due to the fact that any non-zero Z-scores show high significance.

A.3 Configuration Model

The configuration model of network theory is designed to generate random networks that maintain a predefined degree sequence [251]. The essence of this model is that each node in the resulting network has a predetermined degree k_i, while the network itself is wired randomly, resulting in the most random network with a predefined degree sequence. When this procedure is repeatedly applied to the same degree sequence usually generated from a degree distribution p_k, it generates an ensemble of different networks.

A.4 Code and Data

A.4.1 GitHub repository

This GitHub repository contains the bots design.

A.4.2 Data availability

The datasets generated and analysed during the current study are available in the Harvard Dataverse repository, `https://doi.org/10.7910/DVN/ZSI5MH`.

Appendix B

Supplementary Material of Chapter 4

B.1 Rich-Club Detection

Rich-club (RC) detection is a method in network science aimed at identifying a subset of high-degree nodes that demonstrate a denser interconnectedness than expected from their node degrees alone. The behavior of the rich-club coefficient as a function of the degree k is a probe for the topological correlations that yields important information about the organizing principle of complex networks. The RC detection process has been well-documented in previous works [158, 252] and involves the following stages:

1. For a given degree k, discard all nodes with a degree $\leq k$.

2. Compute the rich-club coefficient $\phi(k)$ using the density of the induced subgraph formed by the retained nodes. The density is given by the ratio of existing edges between the nodes to their potential total. This relationship is expressed as:

$$\phi(k) = \frac{2E_{>}k}{N_{>k}(N_{>k} - 1)} \tag{B.1}$$

   Here, $E_{>}k$ is the count of edges between the nodes with a degree exceeding k.

3. Construct 100 random networks that maintain the initial network's degree distribution through the configuration model. For each of these networks, compute the subgraph density (or average rich-club coefficient, $\phi_{rand}(k)$) for nodes with a degree greater than k.

4. Compute the normalized rich-club coefficient, $\rho(k)$, by dividing $\phi(k)$ by the average coefficient from the random networks, $\phi_{rand}(k)$. A $\rho(k)$ exceeding 1 indicates the rich-club phenomenon relative to the null model:

$$\rho(k) = \frac{\phi(k)}{\phi_{rand}(k)} \tag{B.2}$$

5. Repeat steps 1-4 for k values from the lowest to the highest degree observed in the network.

6. A rich-club is suggested if $\rho(k)$ remains above one over a range of degree values. This indicates that the observed $\phi(k)$ surpasses typical in randomized networks, implying denser interconnectivity among high-degree nodes.

B.2 Network Data

Table B.1 presents the data used for rich-club detection. Due to computational constraints in constructing and analyzing the complete PR network, we used a segment version from the last five months, from 1/11/2022 to 13/04/2023.

	ISIS	FR	PR	PR (segment)
Number of human accounts	4615	6917	113524	8026
Number of anonymous accounts	4680	1106	2869	42
Number of bots	83	39	12	7
Number of channels	1010	124	9	7
Number of groups	27	2	0	0
Total nodes	9378	8062	-	8075
Total edges	2,870,396	5,221,621	-	10,842,172
Average degree, $\langle k \rangle$	612	1295	-	2685
Maximum degree, k_{max}	7514	6523	-	6511
Degree thresholds	1823	1521	-	16

Table B.1: Network data summary for each group. The average degree $\langle k \rangle$ represents the typical degree within a network, while k_{max} is the maximum degree. The degree threshold, pivotal for rich-club analysis, spans from the minimum to the maximum degree held by nodes in each network. Note: "-" denotes unavailable data for that particular metric.

Appendix C

Supplementary Material of Chapter 5

C.1 Identifying the Communities

In network science, communities refer to parts of a connected network within which nodes are linked to each other more than to the rest of the network. Such communities are supposed to have functional role and their efficient identification is a major scientific challenge[244, 253]. Communities can be disjunct – then finding them means to identify a partition of the network – or they can be overlapping, leading to a so-called cover of the network [245]. Having collected the data and built the initial network, and after we identified the largest connected component (LCC), we used four models to identify communities within the network to which it pertained. The models we used are: Girvan-Newman algorithm (GN); asynchronous label propagation algorithm (ALP); stochastic block model (bSBM); and nested stochastic block model (nSBM). Each of these is described in more detail below.

1. **Girvan-Newman (GN)** [246]: GN operates by finding community boundaries and breaking a given network into smaller groups by iteratively eliminating the edges with the highest scores for betweenness centrality. (Betweenness centrality measures the importance of an edge in connecting different parts of a network [254]). Edges with high betweenness centrality usually join communities, so by locating these edges one-by-one we can work towards ascertaining the topological structure of a network. The GN algorithm has four steps: i) calculate the betweenness centrality score for all edges in the network; ii) remove the edge with the highest betweenness centrality score; iii) recalculate the edge betweenness centrality score for every remaining edge; and iv) repeat this procedure iteratively. The community structure is found by the partition resulting from stopping the procedure where a global objective function (modularity) reaches its maximum.

2. **Asynchronous label propagation (ALP)** [255, 256]: ALP is an iterative community detection algorithm developed to be time-efficient and computationally less expensive in identifying communities in networks. It requires no prior information, such as the number and size of the communities in question, and it does not optimize any objective function, relying only on the overarching network structure (more specifically on the immediate neighbors of each node). Put simply, ALP operates based on the idea that nodes belong to communities that their proximal neighbors belong to. It, too, has four steps: i) attach a label to every node in a network; ii) let the label

propagate throughout the network and, as the labels propagate, track the formation of densely connected nodes that form consensus around unique labels; iii) perform this process iteratively (for each step, each node updates its labels based on the labels of its immediate neighbors); and iv) cease the updating process when each node has the label that appears most frequently among its neighbors.

3. **Stochastic block models (SBMs)** [257]: SBMs were originally invented for providing generative tools to construct networks with prescribed properties, but they can be used for inferring communities in networks. In a SBM of a network, nodes are assigned into groups called 'blocks' and general assumptions are made about the probabilities of links within and between the blocks. In the basic version denoted by bSBM, the partition of the nodes into groups and the matrix of these probabilities as parameters are considered as parameters, where the element p_{rs} of the matrix specifies the independent probability of having an edge between a node of group r and a node of group s [258]. Using Bayesian inference, we avoid the problem of overfitting. To overcome underfitting, we also used a nested version of the SBM (nSBM), which was introduced to discover the network's hierarchical structure (if it exists at all) [259]. The approach consists of first inferring an SBM from the network data, and since this result can be represented as a multigraph with the groups as nodes and the new edges given by the corresponding edges of the nodes inside each group, another SBM can be inferred based on the results. A third SBM can be inferred from the second multigraph, and by continuing this process, it is possible to obtain a nested hierarchy.

C.2 Comparing the Methods

Adjusted Rand index ($\boldsymbol{ARI}$): $\boldsymbol{ARI}$ is used to compare the performance of the different community detection algorithms [260]. $\boldsymbol{ARI}$ is used to measure the similarity of partitioning of two different community detection algorithms. It works by calculating the fractions of nodes that are similarly classified by two algorithms irrespective of the permutations of the nodes in a community. It is defined as follows [261]:

$$\boldsymbol{ARI} = \frac{s+d}{n(n-1)/2} \tag{C.1}$$

Where s, which stands for 'similar,' is the number of pairs of nodes which are in the same communities in both algorithms to be compared, while d is the number of pairs of nodes, which are in the same community for one algorithm but different communities as determined by the other algorithm. Thus, $s + d$ is the total number of pairs of nodes placed in their corresponding groups in one algorithm and their corresponding groups in the other algorithm. The nominator $n(n-1)/2$ is the number of pairs of n nodes. The higher the value of the $\boldsymbol{ARI}$, the more similar the community structures as calculated by the compared algorithms.

C.3 Analyzing the Communities

When we applied the models to the data, they identified different topological structures (see figure C.1). However, they all agreed to cut the network into mixed groups of bots and channels/groups. Figure C.2 shows that SBMs were very similar ($\boldsymbol{ARI} \approx 0.91$), due to that they share same process; bSBM and nSBM inferred three communities. ALP and the other three models were different; GN was very similar with the bSBM and nSBM ($\boldsymbol{ARI} \approx 0.87 \approx 0.92$ respectively). Only ALP was different from the rest due to the fact it looks for a higher resolution. On other words, GN and the SBMs were able to extract the biggest possible cohesive communities while ALP was able to provide zoomed-in and smaller communities.

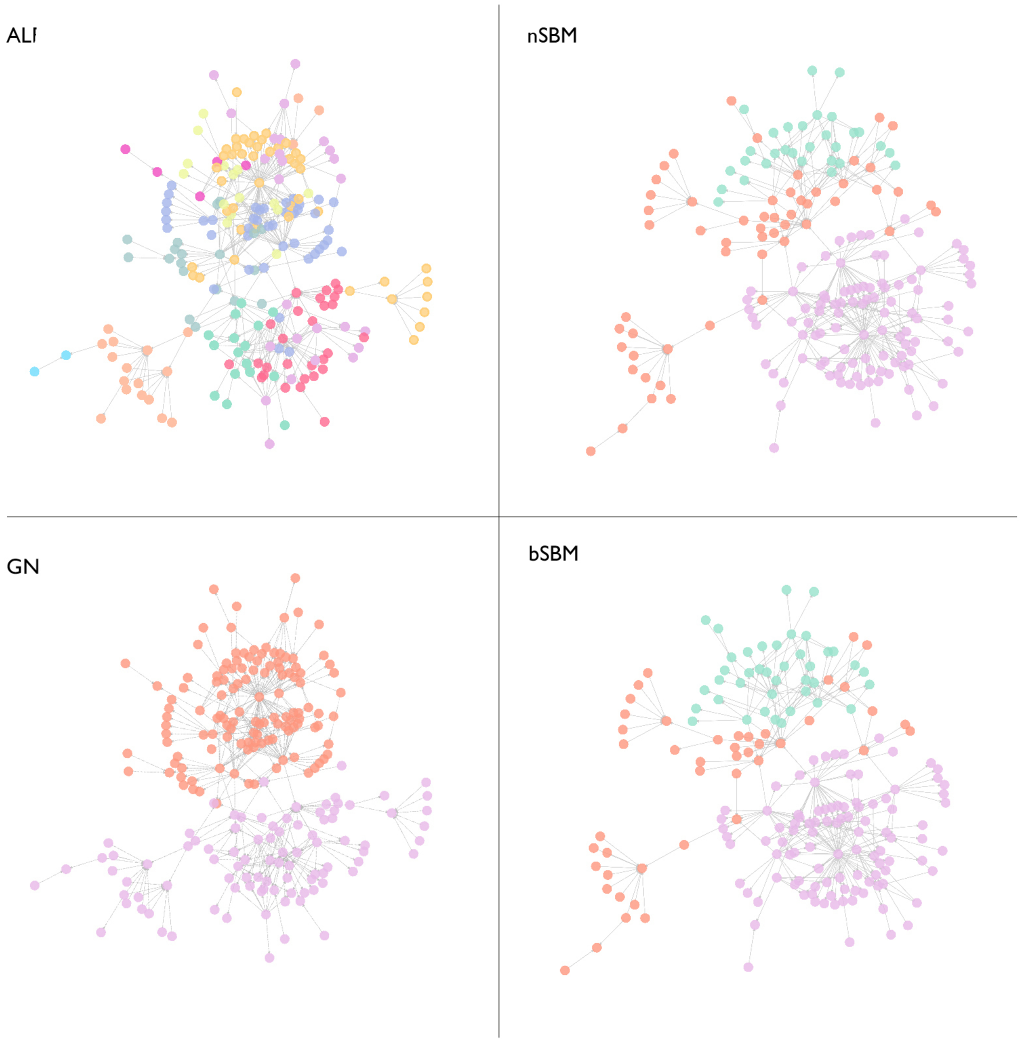
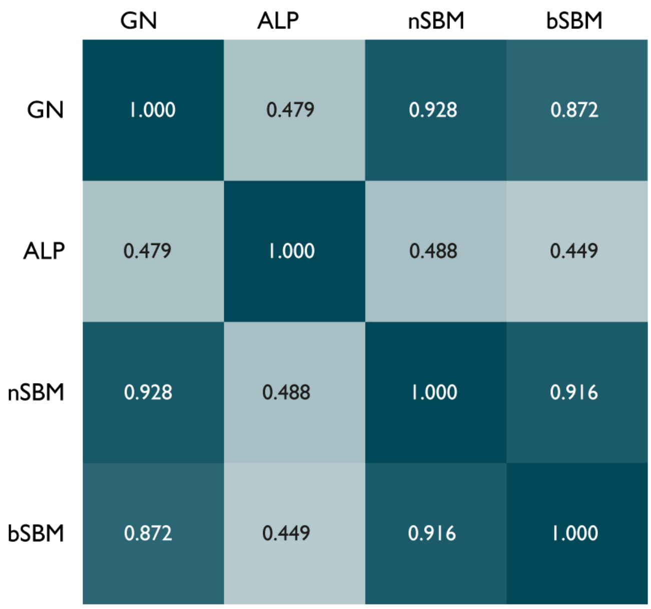


Figure C.1: Network visualization of the communities produced by the four models.

C.4 Detecting the Core-Periphery Structure

The core-periphery (C-P) structure in networks is a mesoscale structure consists of two sets of nodes: core and periphery. The core is a group of highly interconnected nodes and the periphery is a group of sparsely interconnected nodes [262]. A core node is well connected to other core nodes as well as it is well connected to peripheral nodes. Scientists introduced many models to detect the C-P structures. To compute the coreness, we used Rombach's method [247], which is a modified version of the continuous Borgatti-Everett



	GN	ALP	nSBM	bSBM
GN	1.000	0.479	0.928	0.872
ALP	0.479	1.000	0.488	0.449
nSBM	0.928	0.488	1.000	0.916
bSBM	0.872	0.449	0.916	1.000

Figure C.2: Similarities between four different models as measured by ARI. ARI is used to measure similarity in the results of the respective models (GN, ALP, bSBM, and nSBM).

method [248]. Rombach defined the coreness of a graph as follows:

$$R_\gamma \ = \sum_{i,j} A_{ij} C_{ij} \tag{C.2}$$

where γ is a vector that parametrizes the coreness and it is = (α, β). α characterizes the sharpness of the transition from core to periphery, and β is the ratio of nodes belonging to the core. A_{ij} is the adjacency matrix of the network, and C_{ij} is calculated as follow

$$C_{ij} \ = C_i C_j \tag{C.3}$$

where $C_i \ = \ 1$ if i is in the core and 0 otherwise, i.e., C_{ij} indicates if both i and j are in the core. In a monopartite network the C_{ij} matrix can be arranged such that it forms a diagonal block of 1-s. In the bipartite network we deal with the corresponding block in the C_{ij} matrix is part of an off diagonal part (see figure C.3).

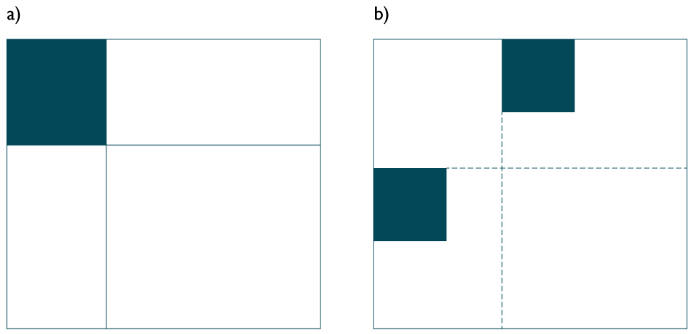


Figure C.3: The structure of the reference matrix C_{ij} for a) a monopartite network; b) a bipartite network. In b) the dashed line is for separation of the node labels in the two interconnected sets.

Appendix D

Some Important Network Science Concepts

Assortativity (degree assortativity) The tendency of nodes in a network to connect to other nodes with similar degree. It is quantified by the assortativity coefficient r, measured by the correlation. Positive r means assortativity, negative one means dis-assortativity.

Community or module A subgraph in a connected network within which nodes are linked to each other more than to the rest of the network. Community detection is an algorithm to identify the communities in (large) networks.

Core-periphery structure is the phenomenon in networks that the nodes can be divided in a densely connected and a loosely connected part.

Degree The number k_i of edges node i has in the network. Average degree: $\overline{k} = (1/N)\sum_{i=1}^{N} k_i$ in a network of N nodes.

Network Motif A significantly over-represented set of topologically equivalent subgraphs of interconnected nodes that exist within larger graphs, where over-representation is with respect to a null model, usually the uncorrelated random model with fixed degree sequence (configuration model). Colored motifs are network motifs where nodes have some (discrete) properties. In our case a node can be a human or a bot.

Rich club phenomenon is the tendency of high degree nodes to connect to each other (irrespective of what happens to low degree nodes). It is measured by the rich club coefficient.

Subgraph In a graph (or network) G(V,N) with a set of nodes $1, 2, ..., i, j, ... \in V$ and a set of edges E, where $e_{ij} \in E$ such that nodes $i, j \in V$ a subgraph is a graph $G'(V', E')$ such that $V' \subset V$ and $E' \subset E$ such that $e'_{i',j'} \in E'$ if $i', j' \in V'$.

Appendix E

Ethical Declaration

The research project presented in chapter 3 received an ethical compliance certification from the Ethical Research Committee of my home institution (Central European University). The certification states that the '..research project titled "The Ecosystem of Humans and Bots in Online Social Networks" led by Abdullah Alrhmoun (PhD candidate, Network Science) had been reviewed for ethical research issues by the appropriate bodies, as defined by the Ethical Research Policy of Central European University, available at: https://documents.ceu.edu/documents/p-1012-1v220.'

The research project presented in chapter 4 was approved by the institutional ethics committee at Indiana University Bloomington (protocol # 14273), where I was a visiting scholar when the project was started. According to the approved protocol, i) I only collected data from publicly available Telegram channels/groups and some "private" Telegram channels/groups that act like public ones because their links are openly accessible online. ii) I did not interact with the subjects on these channels/groups.

Bibliography

[1] Abdullah Alrhmoun, Charlie Winter, and János Kertész. "Automating Terror: The Role and Impact of Telegram Bots in the Islamic State's Online Ecosystem". In: *Terrorism and Political Violence* (2023). DOI: `10.1080/09546553.2023.2169141`.

[2] Andrew Leonard. *Bots Are Hot!* Apr. 1996. URL: `https://www.wired.com/1996/04/netbots/`.

[3] Joseph Weizenbaum. "ELIZA—a computer program for the study of natural language communication between man and machine". In: *Communications of the ACM* 9.1 (1966), pp. 36–45.

[4] Emilio Ferrara et al. "The rise of social bots". In: *Communications of the ACM* 59.7 (2016), pp. 96–104.

[5] Terrence Adams. "AI-powered social bots". In: *arXiv preprint arXiv:1706.05143* (2017).

[6] Chengcheng Shao et al. "The spread of low-credibility content by social bots". In: *Nature communications* 9.1 (2018), pp. 1–9.

[7] Tobias R Keller and Ulrike Klinger. "Social bots in election campaigns: Theoretical, empirical, and methodological implications". In: *Political Communication* 36.1 (2019), pp. 171–189.

[8] Lennart Hofeditz et al. "Meaningful Use of Social Bots? Possible Applications in Crisis Communication during Disasters." In: *ECIS*. 2019.

[9] Florian Brachten et al. "Threat or opportunity?-examining social bots in social media crisis communication". In: *arXiv preprint arXiv:1810.09159* (2018).

[10] Stefan Stieglitz et al. "Do social bots (still) act different to humans?–Comparing metrics of social bots with those of humans". In: *Social Computing and Social Media. Human Behavior: 9th International Conference, SCSM 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part I 9*. Springer. 2017, pp. 379–395.

[11] David A Broniatowski et al. "Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate". In: *American journal of public health* 108.10 (2018), pp. 1378–1384.

[12] Tony Veale and Mike Cook. *Twitterbots: Making Machines that make meaning*. MIT Press, 2018.

[13] Guido Caldarelli et al. "The role of bot squads in the political propaganda on Twitter". In: *Communications Physics* 3.1 (2020), p. 81.

[14] Yazan Boshmaf et al. "The socialbot network: when bots socialize for fame and money". In: *Proceedings of the 27th annual computer security applications conference*. 2011, pp. 93–102.

[15] Tuja Khaund et al. "Social bots and their coordination during online campaigns: A survey". In: *IEEE Transactions on Computational Social Systems* 9.2 (2021), pp. 530–545.

[16] Emilio Ferrara. "Disinformation and social bot operations in the run up to the 2017 French presidential election". In: *arXiv preprint arXiv:1707.00086* (2017).

[17] Tuja Khaund et al. "Analyzing social bots and their coordination during natural disasters". In: *Social, Cultural, and Behavioral Modeling: 11th International Conference, SBP-BRiMS 2018, Washington, DC, USA, July 10-13, 2018, Proceedings 11*. Springer. 2018, pp. 207–212.

[18] Iyad Rahwan et al. "Machine behaviour". In: *Machine Learning and the City: Applications in Architecture and Urban Design* (2022), pp. 143–166.

[19] Alessandro Bessi and Emilio Ferrara. "Social bots distort the 2016 US Presidential election online discussion". In: *First monday* 21.11-7 (2016).

[20] Abeer Aldayel and Walid Magdy. "Characterizing the role of bots' in polarized stance on social media". In: *Social Network Analysis and Mining* 12.1 (2022), p. 30.

[21] Michael D Rich et al. *Truth decay: An initial exploration of the diminishing role of facts and analysis in American public life*. Rand Corporation, 2018.

[22] Matthew Benigni and Kathleen M Carley. "From tweets to intelligence: Understanding the islamic jihad supporting community on twitter". In: *Social, Cultural, and Behavioral Modeling: 9th International Conference, SBP-BRiMS 2016, Washington, DC, USA, June 28-July 1, 2016, Proceedings 9*. Springer. 2016, pp. 346–355.

[23] Sinan Aral. *The hype machine: how social media disrupts our elections, our economy, and our health–and how we must adapt*. Currency, 2021.

[24] Marina Azzimonti and Marcos Fernandes. "Social media networks, fake news, and polarization". In: *European Journal of Political Economy* 76 (2023), p. 102256.

[25] Emilio Ferrara. "Bots, elections, and social media: a brief overview". In: *Disinformation, Misinformation, and Fake News in Social Media: Emerging Research Challenges and Opportunities* (2020), pp. 95–114.

[26] Harry Yaojun Yan et al. "Asymmetrical perceptions of partisan political bots". In: *New Media & Society* 23.10 (2021), pp. 3016–3037.

[27] Luca Luceri et al. "Red bots do it better: Comparative analysis of social bot partisan behavior". In: *Companion proceedings of the 2019 world wide web conference*. 2019, pp. 1007–1012.

[28] Stefano Pedrazzi and Franziska Oehmer. "Communication rights for social bots?: options for the governance of automated computer-generated online identities". In: *Journal of Information Policy* 10 (2020), pp. 549–581.

[29] Wen Chen et al. "Neutral bots probe political bias on social media". In: *Nature Communications* 12.1 (2021), pp. 1–10.

[30] Kurt Wagner. *The Surprising Usefulness of Vaccine Bots*. Apr. 2021. URL: https://www.bloomberg.com/news/newsletters/2021-04-08/the-surprising-usefulness-of-vaccine-bots.

[31] Johan Fernquist, Lisa Kaati, and Ralph Schroeder. "Political bots and the Swedish general election". In: *2018 ieee international conference on intelligence and security informatics (isi)*. IEEE. 2018, pp. 124–129.

[32] Martin N Ndlela. "Social media algorithms, bots and elections in Africa". In: *Social media and elections in Africa, Volume 1: Theoretical perspectives and election campaigns* (2020), pp. 13–37.

[33] Javier Pastor-Galindo et al. "Twitter social bots: The 2019 Spanish general election data". In: *Data in brief* 32 (2020), p. 106047.

[34] Luca Luceri et al. "Evolution of bot and human behavior during elections". In: *First Monday* (2019).

[35] Joshua Uyheng and Kathleen M Carley. "Bots and online hate during the COVID-19 pandemic: case studies in the United States and the Philippines". In: *Journal of computational social science* 3 (2020), pp. 445–468.

[36] Menghan Zhang et al. "Social Bots' Involvement in the COVID-19 Vaccine Discussions on Twitter". In: *International Journal of Environmental Research and Public Health* 19.3 (2022), p. 1651.

[37] Dennis Assenmacher et al. "Demystifying social bots: On the intelligence of automated social media actors". In: *Social Media+ Society* 6.3 (2020), p. 2056305120939264.

[38] Florian Daniel, Cinzia Cappiello, and Boualem Benatallah. "Bots acting like humans: Understanding and preventing harm". In: *IEEE Internet Computing* 23.2 (2019), pp. 40–49.

[39] Koosha Zarei, Reza Farahbakhsh, and Noel Crespi. "Typification of impersonated accounts on instagram". In: *2019 IEEE 38th International Performance Computing and Communications Conference (IPCCC)*. IEEE. 2019, pp. 1–6.

[40] Norah Abokhodair, Daisy Yoo, and David W McDonald. "Dissecting a social botnet: Growth, content and influence in Twitter". In: *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 2015, pp. 839–851.

[41] Adrienne L Massanari. "Contested play: The culture and politics of reddit bots". In: *Socialbots and their friends*. Routledge, 2016, pp. 126–143.

[42] Milena Tsvetkova et al. "Even good bots fight: The case of Wikipedia". In: *PloS one* 12.2 (2017), e0171774.

[43] Charles Kiene and Benjamin Mako Hill. "Who uses bots? A statistical analysis of bot usage in moderation teams". In: *Extended abstracts of the 2020 CHI conference on human factors in computing systems*. 2020, pp. 1–8.

[44] Michael L Mauldin. "Chatterbots, tinymuds, and the turing test: Entering the loebner prize competition". In: *AAAI*. Vol. 94. 1994, pp. 16–21.

[45] John C Paolillo. ""Conversational" codeswitching on usenet and internet relay chat". In: *Language@ Internet* 8.3 (2011).

[46] Guillaume Latzko-Toth. "The socialization of early Internet bots: IRC and the ecology of human-robot interactions online". In: *Socialbots and Their Friends*. Routledge, 2016, pp. 63–84.

[47] John Canavan. "The evolution of malicious IRC bots". In: *Virus bulletin conference*. 2005, pp. 104–114.

[48] Nicole M Radziwill and Morgan C Benton. "Evaluating quality of chatbots and intelligent conversational agents". In: *arXiv preprint arXiv:1704.04579* (2017).

[49] Daniel Lehtovirta and Tobias Jönsson. "Agents in Computer Games". In: *First Blekinge Institute of Technology Student Work-shop on Agent Programming* (2000), p. 73.

[50] Muhammad Fikri Hasani and Yogi Udjaja. "Immersive experience with non-player characters dynamic dialogue". In: *2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI)*. Vol. 1. IEEE. 2021, pp. 418–421.

[51] Alex Mcconnell. *How Bots are Ruining Online Gaming for Players and Publishers*. Aug. 2021. URL: https://netacea.com/blog/bots-are-ruining-online-gaming-for-players/.

[52] Paul H. Müller. *Winning against bots: How gaming apps can fight back*. Apr. 2019. URL: https://venturebeat.com/games/winning-against-bots-how-gaming-apps-can-fight-back/.

[53] Samantha Bradshaw and Philip N Howard. “The global disinformation order: 2019 global inventory of organised social media manipulation”. In: (2019).

[54] Denis Stukal et al. “Detecting bots on Russian political Twitter”. In: *Big data* 5.4 (2017), pp. 310–324.

[55] Kevin M Blasiak, Marten Risius, and Sabine Matook. ““Social Bots for Peace”: A Dual-Process Perspective to Counter Online Extremist Messaging”. In: Association for Information Systems. 2021.

[56] Michelle Singletary and Youjin Shin. *Trying to plan for retirement during the pandemic? We made a bot to help you*. May 2020. URL: https://www.washingtonpost.com/graphics/2020/business/retirement-planning-bot/.

[57] Jeremy Gayed. *How Building a Slack Bot Helped Us Send News Notifications*. Apr. 2019. URL: https://open.nytimes.com/how-building-a-slack-bot-helped-us-send-news-notifications-f28c681a5b3b.

[58] Shruti Dhapola. *How a chatbot helped Joe Biden become US President*. Jan. 2021. URL: https://indianexpress.com/article/technology/tech-news-technology/amplify-ai-chatbot-how-it-helped-joe-biden-become-us-president-7157578/.

[59] R Stuart Geiger. “The lives of bots”. In: *arXiv preprint arXiv:1810.09590* (2018).

[60] Nick Monaco and Samuel Woolley. *Bots*. John Wiley & Sons, 2022.

[61] Philip N Howard. *Lie machines: How to save democracy from troll armies, deceitful robots, junk news operations, and political operatives*. Yale University Press, 2020.

[62] Eiman Alothali et al. “Detecting social bots on twitter: a literature review”. In: *2018 International conference on innovations in information technology (IIT)*. IEEE. 2018, pp. 175–180.

[63] Joseph Seering et al. “It takes a village: integrating an adaptive chatbot into an online gaming community”. In: *Proceedings of the 2020 chi conference on human factors in computing systems*. 2020, pp. 1–13.

[64] Marty J Wolf, K Miller, and Frances S Grodzinsky. “Why we should have seen that coming: comments on Microsoft’s tay" experiment," and wider implications”. In: *Acm Sigcas Computers and Society* 47.3 (2017), pp. 54–64.

[65] Hope Reese. “Why Microsoft’s “Tay” AI bot went wrong”". In: *Tech Republic* 24 (2016).

[66] Langdon Winner. “Do artifacts have politics?” In: *Daedalus* (1980), pp. 121–136.

[67] Lucy Suchman and Lucy A Suchman. *Human-machine reconfigurations: Plans and situated actions*. Cambridge university press, 2007.

[68] Joseph Cox. *These Bots Tweet When Government Officials Edit Wikipedia*. July 2014. URL: https://www.vice.com/en/article/pgaka8/these-bots-tweet-when-government-officials-edit-wikipedia.

[69] Heather Ford, Elizabeth Dubois, and Cornelius Puschmann. “Automation, algorithms, and politics| Keeping Ottawa honest—one tweet at a time? Politicians, journalists, Wikipedians and their Twitter bots”. In: *International Journal of Communication* 10 (2016), p. 24.

[70] Tetyana Lokot and Nicholas Diakopoulos. “News Bots: Automating news and information dissemination on Twitter”. In: *Digital Journalism* 4.6 (2016), pp. 682–699.

[71] Joseph Seering et al. “The social roles of bots: evaluating impact of bots on discussions in online communities”. In: *Proceedings of the ACM on Human-Computer Interaction* 2.CSCW (2018), pp. 1–29.

[72] Filippo Menczer and Thomas Hills. *Information Overload Helps Fake News Spread, and Social Media Knows It*. Dec. 2020. URL: https://www.scientificamerican.com/article/information-overload-helps-fake-news-spread-and-social-media-knows-it/.

[73] Serena Tardelli et al. "Characterizing social bots spreading financial disinformation". In: *Social Computing and Social Media. Design, Ethics, User Behavior, and Social Network Analysis: 12th International Conference, SCSM 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part I 22*. Springer. 2020, pp. 376–392.

[74] Ross Schuchard et al. "Bots fired: examining social bot evidence in online mass shooting conversations". In: *Palgrave Communications* 5.1 (2019), pp. 1–12.

[75] Samuel C Woolley and Douglas R Guilbeault. "Computational propaganda in the United States of America: Manufacturing consensus online". In: *Computational Propaganda Research Project* 22 (2017).

[76] Onur Varol et al. "Online human-bot interactions: Detection, estimation, and characterization". In: *Proceedings of the international AAAI conference on web and social media*. Vol. 11. 1. 2017, pp. 280–289.

[77] Clare Duffy and Brian Fung. *Elon Musk commissioned this bot analysis in his fight with Twitter. Now it shows what he could face if he takes over the platform*. Oct. 2022. URL: https://edition.cnn.com/2022/10/10/tech/elon-musk-twitter-bot-analysis-cyabra/index.html.

[78] Andreas Hepp. "Artificial companions, social bots and work bots: Communicative robots as research objects of media and communication studies". In: *Media, Culture & Society* 42.7-8 (2020), pp. 1410–1426.

[79] Nicholas Confessore et al. "The follower factory". In: *The New York Times* 27 (2018).

[80] Katrina Lee. "Your Honor, on Social Media: The Judicial Ethics of Bots and Bubbles". In: *Nev. LJ* 19 (2018), p. 789.

[81] Jon-Patrick Allem and Emilio Ferrara. "Could social bots pose a threat to public health?" In: *American journal of public health* 108.8 (2018), p. 1005.

[82] Wen Shi et al. "Social bots' sentiment engagement in health emergencies: A topic-based analysis of the COVID-19 pandemic discussions on Twitter". In: *International Journal of Environmental Research and Public Health* 17.22 (2020), p. 8701.

[83] *Bot summit*. URL: https://tinysubversions.com/2013/11/bot-summit/.

[84] Will Schenk. *Bot design patterns*. Nov. 2014. URL: https://willschenk.com/articles/2014/bot-design-patterns/.

[85] Samuel C Woolley and Philip N Howard. "Social media, revolution, and the rise of the political bot". In: *Routledge handbook of media, conflict and security*. Routledge, 2016, pp. 302–312.

[86] Michelle Forelle et al. "Political bots and the manipulation of public opinion in Venezuela". In: *arXiv preprint arXiv:1507.07109* (2015).

[87] Fabian Schäfer, Stefan Evert, and Philipp Heinrich. "Japan's 2014 general election: Political bots, right-wing internet activism, and prime minister Shinzō Abe's hidden nationalist agenda". In: *Big data* 5.4 (2017), pp. 294–309.

[88] Fenwick McKelvey and Elizabeth Dubois. "Computational propaganda in Canada: The use of political bots". In: (2017).

[89] Sippo Rossi et al. "Detecting political bots on Twitter during the 2019 Finnish parliamentary election". In: (2020).

[90] Florian Brachten et al. “Strategies and Influence of Social Bots in a 2017 German state election-A case study on Twitter”. In: *arXiv preprint arXiv:1710.07562* (2017).

[91] Philip N Howard et al. “Junk news and bots during the US election: What were Michigan voters sharing over Twitter”. In: *CompProp, OII, Data Memo* 1 (2017).

[92] Dan Arnaudo. “Computational propaganda in Brazil: Social bots during elections”. In: (2017).

[93] Xiaoyi Yuan, Ross J Schuchard, and Andrew T Crooks. “Examining emergent communities and social bots within the polarized online vaccination debate in Twitter”. In: *Social media+ society* 5.3 (2019), p. 2056305119865465.

[94] Matthew R DeVerna et al. “Identification and characterization of misinformation superspreaders on social media”. In: *arXiv preprint arXiv:2207.09524* (2022).

[95] Chang-Feng Chen et al. “Social bots’ role in climate change discussion on Twitter: Measuring standpoints, topics, and interaction strategies”. In: *Advances in Climate Change Research* 12.6 (2021), pp. 913–923.

[96] Ahmet S Yayla and Anne Speckhard. “Telegram: The mighty application that ISIS loves”. In: *International Center for the Study of Violent Extremism* 9 (2017).

[97] Jacob Zenn. “Debates and Controversies over the Legitimacy of “Internet Sources” in Scholarship on Jihadism: The Online Dimension in the Persistence of the “Al-Qaeda Narrative” in Boko Haram Studies”. In: *Studies in Conflict & Terrorism* (2021), pp. 1–25.

[98] Richard Rogers. “Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media”. In: *European Journal of Communication* 35.3 (2020), pp. 213–229.

[99] JM Berger. “Tailored online interventions: The islamic state’s recruitment strategy”. In: *CTC Sentinel* 8.10 (2015), pp. 19–23.

[100] Ahmad Shehabat, Teodor Mitew, and Yahia Alzoubi. “Encrypted jihad: Investigating the role of Telegram App in lone wolf attacks in the West”. In: *Journal of strategic security* 10.3 (2017), pp. 27–53.

[101] Nico Prucha. “Is and the jihadist information highway–projecting influence and religious identity via telegram”. In: *Perspectives on Terrorism* 10.6 (2016), pp. 48–58.

[102] Vera Bergengruen. *How Telegram Became the Digital Battlefield in the Russia-Ukraine War*. Mar. 2022. URL: `https://time.com/6158437/telegram-russia-ukraine-information-war/`.

[103] Yaroslav Druziuk. *A Citizen-like chatbot allows Ukrainians to report to the government when they spot Russian troops — here’s how it works*. Apr. 2022. URL: `https://www.businessinsider.com/ukraine-military-e-enemy-telegram-app-2022-4?r=US&IR=T`.

[104] Lexi Lonas. *Ukraine launches Telegram bot to collect evidence of war crimes*. Mar. 2022. URL: `https://thehill.com/policy/international/597449-ukraine-launches-telegram-bot-to-collect-evidence-of-war-crimes/`.

[105] Kyiv Independent news desk. *Russia demands Telegram remove bots that search the platform for evidence of Russian servicemen captured or killed in Ukraine*. Mar. 2022. URL: `https://kyivindependent.com/russia-demands-telegram-remove-bots-that-search-the-platform-for-evidence-of-russian-servicemen-captured-or-killed-in-ukraine/`.

[106] Douglas E Cowan. “Contested spaces: Movement, countermovement, and e-space propaganda”. In: *Religion Online*. Routledge, 2013, pp. 255–272.

[107] Christopher Lueg and Danyel Fisher. *From Usenet to CoWebs: interacting with social information spaces*. Springer Science & Business Media, 2003.

[108] Vicky Baker. “Battle of the bots”. In: *Index on Censorship* 44.2 (2015), pp. 127–129.

[109] Erkan Saka. “Social media in Turkey as a space for political battles: AKTrolls and other politically motivated trolling”. In: *Middle East Critique* 27.2 (2018), pp. 161–177.

[110] Philip N Howard and Bence Kollanyi. “Bots,# StrongerIn, and# Brexit: computational propaganda during the UK-EU referendum”. In: *arXiv preprint arXiv:1606.06356* (2016).

[111] Bence Kollanyi. “Automation, algorithms, and politics| Where do bots come from? An analysis of bot codes shared on GitHub”. In: *International Journal of Communication* 10 (2016), p. 20.

[112] Gabriel Weimann. *Terrorism in cyberspace: The next generation*. Columbia University Press, 2015.

[113] Maura Conway, Lee Jarvis, and Orla Lehane. *Terrorists’ use of the Internet: Assessment and response*. Vol. 136. Ios Press, 2017.

[114] Bruce Hoffman. “Inside Terrorism Columbia University Press”. In: *New York NY* (1998).

[115] Manuel Ricardo Torres-Soriano. “The dynamics of the creation, evolution, and disappearance of terrorist Internet forums”. In: *International Journal of Conflict and Violence (IJCV)* 7.1 (2013), pp. 164–178.

[116] Andrew Glazzard. *ISIS: The state of terror*. 2015.

[117] Amarnath Amarasingam, Shiraz Maher, and Charlie Winter. “How Telegram disruption impacts jihadist platform migration”. In: *Centre for Research and Evidence on Security Threats* (2021).

[118] Aaron Brantly. “Banning encryption to stop terrorists: A worse than futile exercise”. In: *CTC Sentinel* 10.7 (2017), pp. 29–33.

[119] Charlie Winter, Abdullah Alrhmoun, and Abdul Sayed. *The Taliban’s vast propaganda machine has a new target*. Aug. 2021. URL: https://www.wired.co.uk/article/taliban-propaganda-news-afghanistan.

[120] Joas Wagemakers. “Al-Qa ‘ida’s editor: Abu Jandal al-Azdi’s online jihadi activism”. In: *Politics, Religion & Ideology* 12.4 (2011), pp. 355–369.

[121] Christopher Anzalone. “Continuity and Change: The evolution and resilience of Al-Shabab’s media insurgency, 2006–2016”. In: *Hate Speech International* 9 (2016).

[122] Lina Khatib. *Image politics in the Middle East: The role of the visual in political struggle*. Bloomsbury Publishing, 2012.

[123] Charlie Winter and Abdullah Alrhmoun. *Mapping The Extremist Narrative Landscape In Afghanistan*. 2021. URL: https://public-assets.extrac.io/reports/ExTrac_Afghanistan_201120.pdf.

[124] Telegram. *ISIS Watch*. URL: https://t.me/ISISwatch.

[125] JM Berger. “How terrorists recruit online (and how to stop it)”. In: *Brookings Institution* 9 (2015).

[126] Jytte Klausen. “Tweeting the Jihad: Social media networks of Western foreign fighters in Syria and Iraq”. In: *Studies in Conflict & Terrorism* 38.1 (2015), pp. 1–22.

[127] Jonathon M Berger and Jonathon Morgan. "The ISIS Twitter Census: Defining and describing the population of ISIS supporters on Twitter". In: *Brookings Institution* (2015).

[128] Bruce Hoffman. "Using the Web as a Weapon: The Internet as a Tool for Violent Radicalization and Homegrown Terrorism". In: *Statement provided to the US House Committee on Homeland Security (November 6, 2007)* (2009).

[129] Bruce Hoffman. "The Myth of Grass-Roots Terrorism-Why Osama bin Laden Still Matters". In: *Foreign Aff.* 87 (2008), p. 133.

[130] Maura Conway. "From al-Zarqawi to al-Awlaki: The emergence and development of an online radical milieu". In: *CTX: Combating Terrorism Exchange* 2.4 (2012), pp. 12–22.

[131] Bennett Clifford. "Migration moments: Extremist adoption of text-based instant messaging applications". In: *Retrieved from London* (2020).

[132] Bennett Clifford and Helen Powell. "Encrypted extremism: Inside the English-speaking Islamic state ecosystem on telegram". In: *The George Washington University Program on Extremism* (2019), pp. 27–53.

[133] J. Berger. *How ISIS Games Twitter.* June 2014. URL: https://www.theatlantic.com/international/archive/2014/06/isis-iraq-twitter-social-media-strategy/372856/.

[134] Oz Sultan. "Tackling Disinformation, Online Terrorism, and Cyber Risks into the 2020s". In: *The Cyber Defense Review* 4 (2019), pp. 43–60.

[135] Matthew Bondy. *Bad Bots: The Weaponization of Social Media*. The Project on International Peace and Security. The College of William and Mary. Apr. 2017. URL: https://www.wm.edu/offices/global-research/research-labs/pips/_documents/pips/2016-2017/Bondy.Matthew.pdf.

[136] Chengcheng Shao et al. "The spread of fake news by social bots". In: *arXiv preprint arXiv:1707.07592* 96 (2017), p. 104.

[137] Guido Caldarelli. "A perspective on complexity and networks science". In: *Journal of Physics: Complexity* 1.2 (2020), p. 021001.

[138] Loni Hagen et al. "Rise of the machines? Examining the influence of social bots on a political discussion network". In: *Social Science Computer Review* 40.2 (2022), pp. 264–287.

[139] Mark S Granovetter. "The strength of weak ties". In: *American journal of sociology* 78.6 (1973), pp. 1360–1380.

[140] Marijn A Keijzer and Michael Mäs. "The strength of weak bots". In: *Online Social Networks and Media* 21 (2021), p. 100106.

[141] Ling Heng Wong, Philippa Pattison, and Garry Robins. "A spatial model for social networks". In: *Physica A: Statistical Mechanics and its Applications* 360.1 (2006), pp. 99–120.

[142] Miller McPherson, Lynn Smith-Lovin, and James M Cook. "Birds of a feather: Homophily in social networks". In: *Annual review of sociology* 27.1 (2001), pp. 415–444.

[143] Matthew Bondy. "Bad Bots". In: *The Project on International Peace and Security, Institute for the Theory and Practice of International Relations, College of William and Mary* (2017), pp. 2016–2017.

[144] Lisa Grobelscheg, Ema Kusen, and Mark Strembeck. "Automated Narratives: On the Influence of Bots in Narratives during the 2020 Vienna Terror Attack". In: *Proceedings of the 7th International Conference on Complexity, Future Information Systems and Risk*. Proceedings of the 7th International Conference on Complexity,

Future Information Systems and Risk. 2022, pp. 15–25. DOI: 10.5220/0011034000003197. URL: http://eprints.cs.univie.ac.at/7564/.

[145] Justin Magouirk, Scott Atran, and Marc Sageman. "Connecting terrorist networks". In: *Studies in Conflict & Terrorism* 31.1 (2008), pp. 1–16.

[146] Roy Lindelauf, Peter Borm, and Herbert Hamers. *Understanding terrorist network topologies and their resilience against disruption*. Springer, 2011.

[147] Sean F Everton. *Disrupting dark networks*. 34. Cambridge University Press, 2012.

[148] Philip Vos Fellman. "The complexity of terrorist networks". In: *International journal of networking and virtual organisations* 8.1-2 (2011), pp. 4–14.

[149] Marc Sageman. *Understanding terror networks*. University of Pennsylvania press, 2004.

[150] James Barnett, Shiraz Maher, and Charlie Winter. "Literature review: Innovation, creativity and the interplay between Far-Right and Islamist extremism". In: *International Centre for the Study of Radicalisation* (2021).

[151] C Mathias. "The enemy of my enemy is my friend: What Neo-Nazis like about ISIS". In: *Huffington Post* (2017).

[152] Julia Ebner. *Mein Kampf Meets Jihad: How Neo-Nazis Are Copying the ISIS Terror Playbook*. 2017.

[153] Travis Morris. *Dark ideas: How neo-Nazi and violent jihadi ideologues shaped modern terrorism*. Lexington Books, 2016.

[154] John M Berger. "Nazis vs. ISIS on Twitter: A comparative study of white nationalist and ISIS online social media networks". In: *Center for Cyber & Homeland Security (GWU)* (2016).

[155] Chris Dishman. "The leaderless nexus: When crime and terror converge". In: *Studies in Conflict & Terrorism* 28.3 (2005), pp. 237–252.

[156] Albert-László Barabási. "Network science". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 371.1987 (2013), p. 20120375.

[157] Guido Caldarelli. *Scale-free networks: complex webs in nature and technology*. Oxford Finance, 2007.

[158] Vittoria Colizza et al. "Detecting rich-club ordering in complex networks". In: *Nature physics* 2.2 (2006), pp. 110–115.

[159] Yihong Hu and Daoli Zhu. "Empirical analysis of the worldwide maritime transportation network". In: *Physica A: Statistical Mechanics and its Applications* 388.10 (2009), pp. 2061–2071.

[160] José J Ramasco, Vittoria Colizza, and Pietro Panzarasa. "Using the weighted rich-club coefficient to explore traffic organization in mobility networks". In: *Complex Sciences: First International Conference, Complex 2009, Shanghai, China, February 23-25, 2009. Revised Papers, Part 1 1*. Springer. 2009, pp. 680–692.

[161] Michael Szell and Roberta Sinatra. "Research funding goes to rich clubs". In: *Proceedings of the National Academy of Sciences* 112.48 (2015), pp. 14749–14750.

[162] Martijn P Van Den Heuvel and Olaf Sporns. "Rich-club organization of the human connectome". In: *Journal of Neuroscience* 31.44 (2011), pp. 15775–15786.

[163] Martijn P Van Den Heuvel et al. "High-cost, high-capacity backbone for global brain communication". In: *Proceedings of the National Academy of Sciences* 109.28 (2012), pp. 11372–11377.

[164] Zhuqing Jiao et al. "Rich club characteristics of dynamic brain functional networks in resting state". In: *Multimedia Tools and Applications* 79 (2020), pp. 15075–15093.

[165] Dae-Jin Kim and Byoung-Kyong Min. "Rich-club in the brain's macrostructure: Insights from graph theoretical analysis". In: *Computational and Structural Biotechnology Journal* 18 (2020), pp. 1761–1773.

[166] Shi Zhou and Raúl J Mondragón. "The rich-club phenomenon in the Internet topology". In: *IEEE communications letters* 8.3 (2004), pp. 180–182.

[167] Christopher Ansell, Renata Bichir, and Shi Zhou. "Who says networks, says oligarchy? Oligarchies as "rich club" networks". In: *Connections* 36.1 (2016), pp. 20–32.

[168] Ye Wei et al. "The rich-club phenomenon of China's population flow network during the country's spring festival". In: *Applied Geography* 96 (2018), pp. 77–85.

[169] Luis M Vaquero and Manuel Cebrian. "The rich club phenomenon in the classroom". In: *Scientific reports* 3.1 (2013), pp. 1–8.

[170] Carolina Alves de Lima Salge and Nicholas Berente. "Is that social bot behaving unethically?" In: *Communications of the ACM* 60.9 (2017), pp. 29–31.

[171] Miranda Mowbray. "Ethics for bots". In: *Cognitive, Emotive and Ethical Aspects of Decision Making and Human Action* 1 (2002), pp. 24–28.

[172] Miles C Coleman. "Bots, social capital, and the need for civility". In: *Journal of Media Ethics* 33.3 (2018), pp. 120–132.

[173] Tathagata Chakraborti and Subbarao Kambhampati. "(When) can ai bots lie?" In: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. 2019, pp. 53–59.

[174] Jonas Haeg. "The ethics of political bots: should we allow them for personal use?" In: *Journal of Practical Ethics* 5.2 (2017).

[175] David J Gunkel. "The other question: Socialbots and the question of ethics". In: *Socialbots and Their Friends*. Routledge, 2016, pp. 246–264.

[176] Erhardt Graeff. "What we should do before the social bots take over: Online privacy protection and the political economy of our near future". In: (2013).

[177] Minha Lee et al. "Bots mind the social-technical gap". In: *15th European Conference on Computer-Supported Cooperative Work (ECSCW 2017)*. European Society for Socially Embedded Technologies (EUSSET). 2017, pp. 35–54.

[178] Evelien Heyselaar and Tibor Bosse. "Using theory of mind to assess users' sense of agency in social chatbots". In: *Chatbot Research and Design: Third International Workshop, CONVERSATIONS 2019, Amsterdam, The Netherlands, November 19–20, 2019, Revised Selected Papers 3*. Springer. 2020, pp. 158–169.

[179] Maria Bakardjieva. "Rationalizing sociality: an unfinished script for socialbots". In: *The Information Society* 31.3 (2015), pp. 244–256.

[180] Grant Bollmer and Chris Rodley. "Speculations on the sociality of socialbots". In: *Socialbots and their friends*. Routledge, 2016, pp. 163–179.

[181] Florian Muhle. "Embodied conversational agents as social actors? Sociological considerations on the change of human-machine relations in online environments". In: *Socialbots and Their Friends*. Routledge, 2016, pp. 102–125.

[182] Eleni Adamopoulou and Lefteris Moussiades. "Chatbots: History, technology, and applications". In: *Machine Learning with Applications* 2 (2020), p. 100006.

[183] Emilio Ferrara. "The history of digital spam". In: *Communications of the ACM* 62.8 (2019), pp. 82–91.

[184] Ron Milo et al. "Network motifs: simple building blocks of complex networks". In: *Science* 298.5594 (2002), pp. 824–827.

[185] Austin R Benson, David F Gleich, and Jure Leskovec. "Higher-order organization of complex networks". In: *Science* 353.6295 (2016), pp. 163–166.

[186] Paul Schultz, Jobst Heitzig, and Jürgen Kurths. "Detours around basin stability in power networks". In: *New Journal of Physics* 16.12 (2014), p. 125001.

[187] D. Braines et al. "he Role of Motifs in Understanding Behavior in Social and Engineered Networks". In: *SPIE* 10653 (2018).

[188] Xu Hong-lin et al. "Social Network Analysis Based on Network Motifs". In: *J. Appl. Math.* (2014), p. ID 874708.

[189] Ema Kušen and Mark Strembeck. "An analysis of emotion-exchange motifs in multiplex networks during emergency events". In: *Applied Network Science* 4.1 (2019), pp. 1–33.

[190] Lada A Adamic et al. "Knowledge sharing and yahoo answers: everyone knows something". In: *Proceedings of the 17th international conference on World Wide Web*. 2008, pp. 665–674.

[191] Mauro Coletto et al. "Automatic controversy detection in social media: a content-independent motif-based approach". In: *Online Social Networks and Media* 3 (2017), pp. 22–31.

[192] Alexandru Topirceanu, Alexandra Duma, and Mihai Udrescu. "Uncovering the fingerprint of online social networks using a network motif based approach". In: *Computer Communications* 73 (2016), pp. 167–175.

[193] David M Beskow and Kathleen M Carley. "Its all in a name: detecting and labeling bots by their name". In: *Computational and Mathematical Organization Theory* 25.1 (2019), pp. 24–35.

[194] Kashmir Hill and Jeremy White. *Designed to Deceive: Do These People Look Real to You?* Accessed: March 22, 2021. Nov. 2020. URL: https://www.nytimes.com/interactive/2020/11/21/science/artificial-intelligence-fake-people-faces.html.

[195] Yizhe Zhang et al. "Dialogpt: Large-scale generative pre-training for conversational response generation". In: *arXiv preprint arXiv:1911.00536* (2019).

[196] Twitter. *Twitter's platform manipulation and Spam Policy*. Apr. 2022. URL: https://help.twitter.com/en/rules-and-policies/platform-manipulation.

[197] Philip N Howard, Samuel Woolley, and Ryan Calo. "Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration". In: *Journal of information technology & politics* 15.2 (2018), pp. 81–93.

[198] Twitter. *Trends Overview | docs | twitter developer platform*. URL: https://developer.twitter.com/en/docs/twitter-api/v1/trends/trends-for-location/overview.

[199] *Keyword targeting*. URL: https://business.twitter.com/en/help/campaign-setup/campaign-targeting/keyword-targeting.html.

[200] Peter Sheridan Dodds et al. "Human language reveals a universal positivity bias". In: *Proceedings of the national academy of sciences* 112.8 (2015), pp. 2389–2394.

[201] Renaud Lambiotte and Naoki Masuda. *A guide to temporal networks*. World Scientific, 2016.

[202] Sebastian Wernicke. "Efficient detection of network motifs". In: *IEEE/ACM transactions on computational biology and bioinformatics* 3.4 (2006), pp. 347–359.

[203] Christoph Adami et al. "Information content of colored motifs in complex networks". In: *Artificial Life* 17.4 (2011), pp. 375–390.

[204] d boyd. *The significance of social software*. na, 2007.

[205] Lauri Kovanen et al. "Temporal motifs reveal homophily, gender-specific patterns, and group talk in call sequences". In: *PNAS* 110.45 (2013), pp. 18070–18075.

[206] Sarah Beadle. "How does the Internet facilitate radicalization". In: *London, England: War Studies Department, King's College* (2017).

[207] Özen Odag, Anne Leiser, and Klaus Boehnke. "Reviewing the role of the Internet in radicalization processes". In: *Journal for deradicalization* 21 (2019), pp. 261–300.

[208] Daniel Koehler. "The radical online: Individual radicalization processes and the role of the Internet". In: *Journal for Deradicalization* 1 (2014), pp. 116–134.

[209] Paul Gill et al. "Terrorist use of the Internet by the numbers: Quantifying behaviors, patterns, and processes". In: *Criminology & Public Policy* 16.1 (2017), pp. 99–117.

[210] Alexander Tsesis. "Social media accountability for terrorist propaganda". In: *Fordham L. Rev.* 86 (2017), p. 605.

[211] Gary LaFree and Laura Dugan. "Introducing the global terrorism database". In: *Terrorism and political violence* 19.2 (2007), pp. 181–204.

[212] Bart Schuurman. "Research on terrorism, 2007–2016: A review of data, methods, and authorship". In: *Terrorism and Political Violence* 32.5 (2020), pp. 1011–1026.

[213] Alex Hern and Dan Milmo. *Online hate under scrutiny after Buffalo shooter streamed massacre on Twitch*. May 2022. URL: https://www.theguardian.com/us-news/2022/may/16/buffalo-shooter-twitch-streamed-online-hate.

[214] Sarah Teich. "Trends and developments in lone wolf terrorism in the western world: An analysis of terrorist attacks and attempted attacks by Islamic extremists". In: *International Institute for Counter-Terrorism* 19 (2013).

[215] Stanislaus Riyanta. "Shortcut to Terrorism: Self-Radicalization and Lone-wolf Terror acts: a case study of indonesia". In: *Journal of Terrorism Studies* 4.1 (2022), p. 2.

[216] Massimo La Morgia et al. "Uncovering the Dark Side of Telegram: Fakes, Clones, Scams, and Conspiracy Movements". In: *arXiv preprint arXiv:2111.13530* (2021).

[217] The Telegram Team. *700 million users and telegram premium*. June 2022. URL: https://telegram.org/blog/700-million-and-premium.

[218] Matteo Gregori and Ugo Merlone. "Comparing operational terrorist networks". In: *Trends in Organized Crime* 23 (2020), pp. 263–288.

[219] Jonathan Matusitz. "Similarities between terrorist networks in antiquity and present-day cyberterrorist networks". In: *Trends in Organized Crime* 11 (2008), pp. 183–199.

[220] Ofra Klein and Jasper Muis. "Online discontent: Comparing Western European far-right groups on Facebook". In: *European societies* 21.4 (2019), pp. 540–562.

[221] Nasrullah Memon et al. "Understanding the structure of terrorist networks". In: *International Journal of Business Intelligence and Data Mining* 2.4 (2007), pp. 401–425.

[222] Alexandre Bovet and Peter Grindrod. "Organization and evolution of the UK far-right network on Telegram". In: *Applied Network Science* 7.1 (2022), pp. 1–27.

[223] Scott Helfstein and Dominick Wright. "Covert or convenient? Evolution of terror attack networks". In: *Journal of Conflict Resolution* 55.5 (2011), pp. 785–813.

[224] DARREN LOUCAIDES. *How Telegram Became the Anti-Facebook*. Feb. 2022. URL: https://www.wired.com/story/how-telegram-became-anti-facebook/.

[225] Mariëlle Wijermars and Tetyana Lokot. "Is Telegram a "harbinger of freedom"? The performance, practices, and perception of platforms as political actors in authoritarian states". In: *Post-Soviet Affairs* 38.1-2 (2022), pp. 125–145.

[226] Telegram. *Bots: An introduction for developers*. URL: https://core.telegram.org/bots.

[227] Neil F. Johnson. "New terrorism reveals new physics". In: *APS News Back Page* (), November 2016.

[228] Paul Staniland. "Organizing insurgency: Networks, resources, and rebellion in South Asia". In: *International Security* 37.1 (2012), pp. 142–177.

[229] Paul Stephen Staniland. "Explaining cohesion, fragmentation, and control in insurgent groups". PhD thesis. Massachusetts Institute of Technology, 2010.

[230] Cynthia Stohl and Michael Stohl. "Networks of terror: Theoretical assumptions and pragmatic consequences". In: *Communication Theory* 17.2 (2007), pp. 93–124.

[231] Jonathan David Farley. "Breaking Al Qaeda cells: A mathematical analysis of counterterrorism operations (A guide for risk assessment and decision making)". In: *Studies in Conflict & Terrorism* 26.6 (2003), pp. 399–411.

[232] Stuart Koschade. "A social network analysis of Jemaah Islamiyah: The applications to counterterrorism and intelligence". In: *Studies in Conflict & Terrorism* 29.6 (2006), pp. 559–575.

[233] Marc Sageman. *Understanding terror networks*. University of Pennsylvania press, 2011.

[234] Pedro Manrique et al. "Women's connectivity in extreme networks". In: *Science advances* 2.6 (2016), e1501742.

[235] Yahui Tian and Yulia R Gel. "Fusing data depth with complex networks: Community detection with prior information". In: *Computational Statistics & Data Analysis* 139 (2019), pp. 99–116.

[236] Gian Maria Campedelli, Iain Cruickshank, and Kathleen M Carley. "A complex networks approach to find latent clusters of terrorist groups". In: *Applied Network Science* 4.1 (2019), pp. 1–22.

[237] Matthew C Benigni, Kenneth Joseph, and Kathleen M Carley. "Online extremism and the communities that sustain it: Detecting the ISIS supporting community on Twitter". In: *PloS one* 12.12 (2017), e0181405.

[238] Felix Brünker et al. "'The Tireless Selling-Machine'-Commercial Deployment of Social Bots during Black Friday Season on Twitter." In: *Wirtschaftsinformatik (Zentrale Tracks)*. 2020, pp. 1522–1527.

[239] Wen Chen et al. "Neutral bots reveal political bias on social media". In: *arXiv preprint arXiv:2005.08141* (2020).

[240] Samuel Woolley and Mark Kumleben. "Social Bots for Peace: Combating Automated Control with Automated Civic Engagement?" In: *Social media as a space for peace education*. Springer, 2020, pp. 203–223.

[241] Telegram. *Telegram FAQ*. URL: https://telegram.org/faq?setln=en.

[242] Ibrahim B. Syed. *Birds in the Quran: The Hoopoe*. 2021. URL: https://aboutislam.net/muslim-issues/science-muslim-issues/birds-quran-hoopoe/.

[243] Telegram. *Focused Privacy, Discussion Groups, Seamless Web Bots and More*. URL: https://telegram.org/blog/privacy-discussions-web-bots#broadcasts-meet-group-chats.

[244] Santo Fortunato and Darko Hric. "Community detection in networks: A user guide". In: *Physics reports* 659 (2016), pp. 1–44.

[245] Andrea Lancichinetti, Santo Fortunato, and János Kertész. "Detecting the overlapping and hierarchical community structure in complex networks". In: *New journal of physics* 11.3 (2009), p. 033015.

[246] M. Girvan and M. E. J. Newman. "Community structure in social and biological networks". In: *Proc. Natl. Acad. Sci. USA* 99 (2002), 7821–7826.

[247] M Puck Rombach et al. "Core-periphery structure in networks". In: *SIAM Journal on Applied mathematics* 74.1 (2014), pp. 167–190.

[248] Stephen P Borgatti and Martin G Everett. "Models of core/periphery structures". In: *Social networks* 21.4 (2000), pp. 375–395.

[249] Feng Xia et al. "A survey of measures for network motifs". In: *IEEE Access* 7 (2019), pp. 106576–106587.

[250] Giovanni Ciriello and Concettina Guerra. "A review on models and algorithms for motif discovery in protein–protein interaction networks". In: *Briefings in Functional Genomics and Proteomics* 7.2 (2008), pp. 147–156.

[251] Albert-László Barabási. "Network science book". In: *Network Science* 625 (2014).

[252] Alessandro Muscoloni and Carlo Vittorio Cannistraci. "Rich-clubness test: how to determine whether a complex network has or doesn't have a rich-club?" In: *arXiv preprint arXiv:1704.03526* (2017).

[253] Patrick Doreian, Vladimir Batagelj, and Anuska Ferligoj, eds. *Advances in Network Clustering and Blockmodeling*. Statistics for Social Sciences. Wiley, 2020.

[254] Mark Newman. *Networks*. Oxford university press, 2018.

[255] Usha Nandini Raghavan, Réka Albert, and Soundar Kumara. "Near linear time algorithm to detect community structures in large-scale networks". In: *Physical review E* 76.3 (2007), p. 036106.

[256] Gergely Tibély and János Kertész. "On the equivalence of the label propagation method of community detection and a Potts model approach". In: *Physica A: Statistical Mechanics and its Applications* 387.19-20 (2008), pp. 4982–4984.

[257] Tiago P. Pexoto. "Bayesian stochastic blockmodeling". In: *Advances in Network Clustering and Blockmodeling*. Ed. by Patrick Doreian, Vladimir Batagelj, and Anuska Ferligoj. Wiley, 2020, pp. 289–332.

[258] Tiago P Peixoto. "Nonparametric Bayesian inference of the microcanonical stochastic block model". In: *Physical Review E* 95.1 (2017), p. 012317.

[259] Clement Lee and Darren J Wilkinson. "A review of stochastic block models and extensions for graph clustering". In: *Applied Network Science* 4.1 (2019), pp. 1–50.

[260] Lawrence Hubert and Phipps Arabie. "Comparing partitions". In: *Journal of classification* 2.1 (1985), pp. 193–218.

[261] Alexander J Gates and Yong-Yeol Ahn. "The impact of random models on clustering similarity". In: *arXiv preprint arXiv:1701.06508* (2017).

[262] Andrew Elliott et al. "Core–periphery structure in directed networks". In: *Proceedings of the Royal Society A* 476.2241 (2020), p. 20190783.