STUDENT SESSION

SOCIAL BOTS IN THE AGE OF AI - DETECTION, BEHAVIOR AND CHALLENGES

Anđela Pavlović[1]*,
[0009-0004-0548-8778]

Marko Šarac[2]
[0000-0001-8241-2778]

[1]Student,
Singidunum University,
Belgrade, Serbia

[2]Singidunum University,
Belgrade, Serbia

Abstract:

Social media platforms are increasingly filled with content, including conversations, posts, and user interactions. This content is generated not only by human users but also by automated accounts, such as bots. While both participate in the online environment, they differ in their characteristics, as humans express multiple identities shaped by context, whereas bots typically operate within predefined roles aligned with specific purposes.

This paper analyzes the characteristics of bots on social media, examines their role in online communication, and explores methods for their detection based on linguistic and behavioral patterns. In addition, it addresses challenges associated with sophisticated bots, including both positive and negative examples of their use in social media environments.

Keywords:

Artificial Intelligence, Social Media, Social Bots, Bot Detection.

INTRODUCTION

In recent years, the term “bot” has become increasingly common in both online and offline discussions. Bots are widely present on digital platforms, and many users have interacted with them, often without realizing it. Research has shown that bots are frequently used for activities such as spreading toxic content, amplifying hate speech, and influencing public opinion, including artificially boosting the perceived popularity of political actors during elections. These practices have been linked to increased social tensions and, in some cases, have contributed to the escalation of protests and hate speech. [1]

According to authors [1], a bot is a metaphor for inauthentic online user. Predominantly, definitions of bots have been associated with negative aspects. However, bots can also be utilized beneficially, providing notifications and entertainment. A bot is a computer algorithm. It can be defined as an automated account, controlled by software and can be used in multiple ways. It can create and distribute content, collect data and manage relationships.

Correspondence:
Anđela Pavlović

e-mail:
andjela.pavlovic.21@singimail.rs

DOI: 10.15308/Sinteza-2026-565-569

That means that it can make posts, share them across networks, gather information from interactions, and form or break social connections. A bot is not just a passive tool on the internet, it actively interacts with the environment, influencing the flow of information and social relationships.

Studies show that during U.S. elections, bot activity spikes above 40%. Often, during political events, such as elections, bots become more active, in attempts to shape public opinion.

To distinguish bots from humans, a study was conducted using NetMapper software, counting how many times specific linguistic or behavioral features appear in tweets. It has been shown that bots use a more abusive tone and tweet more frequently than humans, while humans use more first-person pronouns and positive statements. Humans are more likely to tweet more personal content. [1]

Examples of typical bot tweets are: "New Smartphone A is coming out! Check out the specs now! #SmartphoneA #TechNews", "Team B won again! Can't believe the score! #TeamB #Win www.link.com", and "New Movie C is trending! Everyone is talking about it! #MovieC www.link.com". Correspondingly, a human would say something along the lines: "I'm thinking of upgrading my Smartphone A to Smartphone AB. Has anyone tried it yet?", "I'm so excited about Team B's win tonight! Their goalie was amazing!" and "Just watched Movie C. The soundtrack is on repeat!". Using these examples, we can analyze the cues present in linguistic and behavioral patterns. Bots use more generic, direct language to announce information. Humans tend to generate content that is personal, interactive and contextual. First and second person pronouns are predominantly used by humans, while third person pronouns can be used by bots too to generalize statements for wider audiences. When examining emotional cues, bots rely on neutral or negative statements that exaggerate claims, while humans display positive sentiment or excitement within personal contexts. Bots rely on links and hashtags to disseminate information and humans rely on images and emojis to enrich their content and engage others. Cue counts are essential for measuring linguistic, emotional, and behavioral features in social media content. By counting cues, researchers can compare behavior, detect patterns, and identify automated content. Cue counts serve as a foundation for machine learning and AI detection algorithms. The 'Bot Cue Counts' column uses a binary format, where each position refers to an example. A value of 1 represents that the cue is present in that sentence, while 0 means it is not. For example, a sequence of 1/0/0 means that the cue is present in the first sentence and absent in the other two.

Depending on their mission, bots adopt a specific identity and use words associated with that. In contrast, humans have multiple identities [1]. This means that a person may present different identities related to their family, work, and hobbies. The primary purpose of a bot, however, is to complete its assigned task, whether that involves informing users about the weather or spreading misinformation. The language used by a bot aligns with that task, although the person uses more complex language.

Table 1. Cue Counts for Bot vs Human Tweets

Cue	Bot Tweet Examples	Bot Cue Counts	Human Tweet Examples	Human Cue Counts
Semantic First-Person Pronouns (FP)	"New Smartphone A is coming out!" / "Team B won again!" / "New Movie C is trending!"	0 / 0 / 0	"I'm thinking of upgrading my Smartphone A…" / "I'm so excited about Team B win tonight!" / "Just watched Movie C…"	1 / 1 / 0
Semantic Second-Person Pronouns (SP)	(None)	0 / 0 / 0	(None)	1/ 0 / 0
Semantic Third-Person Pronouns (TP)	"the score" / "the score" / "Everyone"	1 / 1 / 1	"Their goalie" / "The soundtrack"	1 / 1
Emotion: Positive Sentiment	(None)	0 / 0 / 0	"excited" / "amazing" / "on repeat"	2 / 1 / 1
Emotion: Negative Sentiment	"Can't believe"	1 / 0 / 0	(None)	0 / 0 / 0
Metadata: Hashtags	#SmartphoneA #TechNews/ #TeamB #Win / #MovieC	2 / 2 / 1	(None)	0 / 0 / 0
Metadata: URLs	www.link.com/ www.link.com	0 / 1 / 1	(None)	0 / 0 / 0

To summarize, a bot is a computer algorithm. It is neither inherently good nor bad. Its impact depends on how it is used, as the actions it performs can be classified as either beneficial or harmful. On social media, there are different types of bots: general, bridging, political, chat and activist. These can serve multiple purposes, including data collection, performance optimization, message amplification, user support, and crisis management, as well as opinion manipulation, political influence, posting offensive content, or sparking activism. Bots often adopt specific identities; for example, they can present themselves as "conservative" and "American" to engage in political discussions, while those portraying themselves as "man" or "son" focus on family-related topics. Bots have a limited dictionary and they often repeat similar expressions. However, detecting these bots has become increasingly difficult and poses a significant challenge for cybersecurity specialists today. ChatGPT helps with the generation of human-like content, and contributes to the development of more sophisticated botnets that are AI powered. [1]

2. CHALLENGES IN SOCIAL BOT DETECTION

Coordinated Inauthentic Behavior (CIB) is a strategy that consists of coordinated operations that can use bots to manipulate discourse on social media. In that case, bots would work together on multiple platforms at once. In 2021, Meta identified a network of fake social media accounts engaged in coordinated anti-vaccination campaigns against COVID-19 vaccinations. Meta had to remove those accounts. CIB tactics can have a massive impact on public discourse that can disturb decision-making processes. This strategy can include bots that are fully automated, but can also include half-automated accounts on social media and accounts that hide the identity of a person running them. [2]

Bots typically operate in groups called botnets. These networks share a common goal, which may include creating and spreading fake information or manipulating ratings and elections. One detection method of botnets involves analyzing information shared among bots within a botnet. The evolution of botnets is characterized by high adaptability, in which advances in detection techniques lead to the development of more sophisticated methods to avoid detection, requiring continuous innovation in detection strategies. With the rise of ChatGPT and technologies with a similar purpose, which can be used as a language model or image creation, bot detection is becoming increasingly difficult. In one study, people had to manually review more than 4,000 accounts and categorize them as spam bots, authentic accounts or unclassifiable. Human accuracy in this task was lower than 0.5. [3]

Algorithms for bot detection use machine-learning methods that collect data from those profiles based on their content and patterns. Sometimes, detection models disagree whether an account is a bot, because models are trained on different datasets that are taught to trigger on different signals. Some social media bots have accounts that are so detailed and mimic human actions and characteristics so well that they are hard to notice. Details from those accounts could be stolen from a real person or generated by deep neural networks. [4]

Other data that is analyzed when deciding if an account is a bot its metadata, specifically account creation date, follower count, following count and their ratio. For instance, bot accounts are often created recently with a high number of followers and following. The content they create is repetitive, with the same language patterns and buzzwords. These methods can be avoided, since bots made nowadays are far more humanlike with their activity on social media. To detect bots that are progressively sophisticated, newer detection models are trained to identify anomalies, biases and patterns in the AI-generated text. Bot detection does not violate users' rights or compromise their personal information. One of the main concerns in social bot detection is the risk of falsely identifying a real user as a bot. This can result in account suspension, content removal and denying access to information. Detection methods should comply with data protection regulations. [5]

There are a few ways to detect a social bot. When analyzing a profile, factors such as the ratio of followers and following and the rate of the content publication, you can decide if a bot is behind that profile. If a profile publishes content often, at consistent intervals, during the whole 24-hour cycle, it raises suspicion. Next step is looking into the content itself and if it is too generic and not personalized, it may suggest automated behavior. If a profile is repeating the same phrases and using third-person pronouns frequently, it's another indication of a bot action. If the language used seems too perfect and consistent, it may point to AI language models, such as ChatGPT, to polish the language. Often, bot profiles are connected to other bot profiles, making a botnet structure.

Social bots today can simulate human-like behavior, which creates a challenge for current detection methods. To illustrate this, consider a hypothetical social media account. Its posting intervals may appear realistic, for example, during working hours or in the evening after work. Content can be generated using AI models, especially language models. Network factor can also be bypassed if the bot interacts with real users in a genuine, natural way. Language models like ChatGPT, Claude and others can generate text that is contextual, logically structured and varied. Because of this, distinguishing between a human and a bot account becomes more difficult. As a result, detection methods have to be more complex and have the ability to detect AI-driven bots. These methods should include frameworks that are hybrid and adaptive, combining multiple techniques into a single detection system.

Social bots also impact cryptocurrency markets and financial scams. They have reshaped public perception, in a way that they created fake demand by spreading false information and messages about the state of market trends. In the future, AI-generated content is rapidly evolving, becoming harder to distinguish from real, human-generated. This evolution is followed by the development of policies and guidelines that address about ethical, legal and societal dimensions. [5]

An example of platforms mislabeling users as bots can be seen on Facebook, Instagram and Twitter. Often, this resulted in suspensions, being locked out of the account, and having to provide Facebook with private information to prove that they are not a bot. This process is often complicated and users may have to wait for a long time for a response or a review of their appeal. These platforms, along with TikTok, made policies regarding bot usage. They are not strictly prohibited, but their actions and usage are monitored. On Twitter, now X, bots are not allowed to be used for spamming, deceptive behavior, aggressive following and unfollowing, but they can be used for informational or entertainment purposes. On Facebook, especially Messenger, bots have a responsive time and cannot be used for deceptive behavior. On Instagram, they are not allowed to impersonate others, do anything against the law, create accounts or collect information without the service's permission. Lastly, on TikTok, bots can't be used to make fake reviews, comments or increase likes or shares. Bot detection datasets should include diverse data, collected from a wide range of user demographics, different cultures, languages, genders... This helps algorithms become more fair and avoid biases. [6]

A new type of bot is known as a sleeper bot, developed using AI. Earlier, before the age of AI, bots could be easily identified, but this bot can mimic human behavior so effectively, and users can engage in conversations and interactions with them, without realizing that an algorithm is behind the persona. It's called a sleeper bot, because during the Cold War, there were 'sleeper agents', agents from an opposing side, portaying a patriot, on a mission to steal information. They are mostly used in political objectives, because they can influence and shape public opinion. [7]

Nowadays, bot creation is largely automated. It starts with data harvesting. Bots are trained on data from various sources, even leaked databases and public records. The next step is profile creation, where bots build human-like profiles. They include detailed information, interests, and a realistic online history. Communication is trained on models such as ChatGPT. To present themselves as real users, they replicate human behavior, including login patterns and response timing. Typically, large numbers of bots are automatically made. Bots continuously learn, even after their creation phase is done, so they become more resilient to cybersecurity attacks. Different sectors may face issues and challenges related to bot activity. For example, the financial sector can face an issue where a bot can pose as a real human and request funds or initiate account changes, which can lead to financial losses. The healthcare sector might struggle with bots that ask for medical records or prescriptions. Bots can even make phishing emails and send them to companies to request funds. Travel agencies and airlines may experience manipulation with interfering with services and inflating prices. [8]

Bots are not only used for malicious purposes. Chatbots can be highly efficient computer programs designed to simulate human conversation in text or voice. There are three types. They can be rule-based, operating and following strict guidelines, AI-driven, where they learn from users and improve over time and can handle difficult conversations or they can be hybrid. They are often utilized in customer support, serving as a brand voice and supporting marketing campaigns. They eliminate the time zone between brands and users, offering instant answers, which results in increasing customer satisfaction. This can positively reflect on a brand, trust and loyalty. Despite all the positive aspects, challenges remain with ethical issues, transparency and privacy concerns, since these bots collect a large amount of personal data. Bots used for marketing and social media can result in higher conversion rates and can gather data from customers that would support business scaling. [9]

3. CONCLUSION

This paper explores the role of social bots in the new age, especially where AI tools and technologies have become incorporated in daily life. There are many definitions of social bots and their growing sophistication as AI-driven agents. Their impact depends on how, why and when they are employed, as well as the purpose they serve. On the one hand, social bots can be useful, for example, customer service, content distribution and crisis communication. On the other hand, the term 'bot' is associated with misuse and is presented as a serious challenge to the public.

It has been shown that social bots operate across various platforms and industries. From influencing election outcomes, amplifying political narratives, to generating fake reviews and inflating follower counts in marketing, bots have reshaped the flow of information in the digital space. As the author states [10], AI technologies have had a huge impact on the dynamics of misinformation, enabling the creation of convincing false narratives. AI relies on multi-layered neural networks and machine learning. Natural language processing (NLP) systems work alongside these technologies, providing them with the function to understand and generate human language. Platforms have introduced guidelines to monitor and limit deceptive bot behavior, but with the rise of sophisticated bots, many bots still manage to go undetected.

Chatbots [11] do not rely only on machine learning and natural language processing. Other technologies that are as equally important include sentiment analysis, which detects the tone and emotion of a user's message, in order to provide a more appropriate response, as well as contextual awareness and dialogue management, which contribute to more relevant conversations. Bots can be deployed in crisis management. AI chatbots are trained to pull information from trustworthy sources. They automatically update their responses in real time, and provide people with more accurate guidelines during natural disasters, pandemics or evacuations. During such situations, human agents are unable to respond to a large number of queries, and that's where AI bots become useful. In companies, AI bots were used for brand management. For example, airlines used chatbots to handle refunds and rebookings.

Overall, the development of social bots and AI technologies can be seen in everyday, digital spaces. While bots can be used for the good, such as communication management and efficiency, supporting users in crisis situations, enhancing customer experience, helping business scale and manage inquiries, misuse is still a major concern. The increasing ability of bots to mimic human behavior, makes them difficult to detect. This raises important questions about trust, privacy and authenticity of online information. The evolution of AI should be followed by improved detection methods, clearer regulations and guidelines, and a stronger emphasis on the responsible use of technology to minimize potential risks while maximizing its benefits.

REFERENCES

[1] L. H. X. Ng and K. M. Carley, "What is a Social Media Bot? A Global Comparison of Bot and Human Characteristics," 2025, arXiv. doi: 10.48550/ARXIV.2501.00855.

[2] M. Murero, "Coordinated inauthentic behavior: An innovative manipulation tactic to amplify COVID-19 anti-vaccine communication outreach via social media," Front. Sociol., vol. 8, Mar. 2023, doi: 10.3389/fsoc.2023.1141416.

[3] S. Lopez-Joya, J. A. Diaz-Garcia, M. D. Ruiz, and M. J. Martin-Bautista, "Bot Detection in Twitter: An Overview," Lecture Notes in Computer Science. Springer Nature Switzerland, pp. 131–144, 2023. doi: 10.1007/978-3-031-42935-4_11.

[4] J. Giroux, G. Ariyarathne, A. C. Nwala, and C. Fanelli, "Unmasking social bots: how confident are we?," EPJ Data Sci., vol. 14, no. 1, Mar. 2025, doi: 10.1140/epjds/s13688-025-00536-y.

[5] E. Ferrara, "Social bot detection in the age of ChatGPT: Challenges and opportunities," FM, Jun. 2023, doi: 10.5210/fm.v28i6.13185.

[6] L. H. X. Ng, E. Pan, M. M. Yoder, and K. M. Carley, "FATe of Bots: Ethical Considerations of Social Bot Detection," 2026, arXiv. doi: 10.48550/ARXIV.2602.05200.

[7] J. Doshi et al., "Sleeper Social Bots: a new generation of AI disinformation bots are already a political threat," 2024, arXiv. doi: 10.48550/ARXIV.2408.12603.

[8] J. P. Castro, "The Future of Social Engineering: How Automated Bots are Reshaping Cybersecurity," Apr. 18, 2025. https:// researchgate.net/publication/390920504_The_Future_of_Social_Engineering_How_Automated_Bots_are_Reshaping_Cybersecurity

[9] Jothy K P, "A Study on the Role of AI and Chatbots in Social Media Marketing: Enhancing Customer Engagement and Experience," IJLTEMAS, vol. 14, no. 6, pp. 884–889, Jul. 2025, doi: 10.51583/ijltemas.2025.140600097.

[10] R. Cazzamatta and A. Sarısakaloğlu, "AI-Generated Misinformation: A Case Study on Emerging Trends in Fact-Checking Practices Across Brazil, Germany, and the United Kingdom," Emerging Media, vol. 3, no. 2, pp. 214–251, Jun. 2025, doi: 10.1177/27523543251344971.

[11] M. Cate and M. Cate, "AI-Powered Chatbots and Virtual Assistants for Crisis Communication," Jul. 01, 2023. https:// researchgate.net/publication/390741250_AI-Powered_Chatbots_and_Virtual_Assistants_for_Crisis_Communication