Article

The Role of Social Media in the Amplification of Disinformation and Hate Speech against Afghan Refugees in Iran: Challenges and Mitigation Strategies

Shahriar Khonsari *

Article

The Role of Social Media in the Amplification of Disinformation and Hate Speech against Afghan Refugees in Iran: Challenges and Mitigation Strategies

Shahriar Khonsari

Freelance Researcher; khonsari2000@gmail.com

Abstract: In Iran, fake news has been refined to weapon due to the popularity of new media and the growing social media-based assault on Afghan refugees. The recent geopolitical tensions and migrations, of course, have brought with them an increased number of false information reports and with them the rapid spread of negative stereotypes and the fueling of anti-Afghan sentiment. The article examines disinformation dispersion and its exacerbation in the digital space. It discusses in specifics ways in which misleading and hate speech against Afghan refugees pass through Iranian social media, in a thorough study of the psychological factors behind these narratives, that favors or fights such disinformation. The study screens the virtual screening areas that distribution of toxicity and disinformation by and engagement with social media. The disinformation technologies as well as the socio-political and historical contexts they derive from. Based on a detailed overview of the cyber-world, this study exposes the myriad challenges of suppressing false narratives and hate speech, focusing particularly on Afghan refugees in Iran. Besides, it points out to methods for both disallowing it to be distributed and curing the disease itself, basically, by redesigning digital agendas to present only the truth. The paper is aimed at unveiling the subversive forms of disinformation, armory with the data concerning digital tools and platforms, as well as input by external societies into the tapestry of the web that will construct a more inclusive digital discourse through critical engagement. Setting apart the genuine and not so wrong investigative procedures show great progress in society through these demanded efforts. Besides, the study impinges on more general issues such as media ethics and fact-checking, in what may be called the development of reliable information systems. This paper impinges on disinformation and humanizing toxic information, a vital aspect that currently has been managed only to a limited extend through the top-down disapproving or regulatory method and it creates strong ecosystem.

Keywords: disinformation; hate speech; Afghan refugees; social media; iranian digital environment; algorithms; mitigation strategies

1. Introduction

"The proliferation of social media has dramatically changed the way information is consumed, shared, and produced. Not only are the ways of personal interactions different, but it changes public discourse and societal attitudes. This increasing presence of social media tends to act like a double-edged sword in many countries of the world, as is the case with Iran. While these platforms allow space for the voices of ideas and expression of marginalized people, they are also fertile grounds from which the spread of disinformation and hate speech consecutively spreads against Afghan refugees. As Aldamen mentions, social media platforms have become a breeding ground for xenophobic

rhetoric and false narratives on refugees that may deepen divisions of society and worsen existing tensions. Most of these platforms are designed to maximize user engagement through algorithmic structures that favor sensationalist and divisive content, therefore building an ecology in which misinformation can more easily spread.

The case is particularly worrying in Iran: Afghan refugees, who already face discrimination and exclusion linked to both historical and contemporary socio-political dynamics, are increasingly victimized by hate speech amplified through social media. According to a study conducted by Zahra and Gholamreza (2021), many of those algorithms serve to amplify the falsity about Afghan refugees, furthering their negative stereotyping and adding hostilities. This amplification of such destructive content, as Bakir and McStay (2018) say, can manipulate users' emotional responses and psychological biases, further fueling engagement in xenophobic content.

Resorting to these narratives is illustrative of several psychological factors, particularly in situations where economic or social instability bites. Current studies argue that fear, anxiety, and the need to scapegoat, especially in times of uncertainty, provide fertile ground for users toward believing and sharing disinformation on minority groups. In Iran, these dynamics overlap with powerful discourses of 'the Other;' Afghan refugees and migrants are viewed as outsiders, and that generates feelings of exclusion and hostility.

This research discusses how these dynamics come together to perpetuate disinformation and hate speech against Afghan refugees in Iran, focusing on the roles driven by psychological factors. It also delves into the demonstration of how those narratives would resonate with Iranian social media users, and what kinds of strategies might help reduce their impact. It will explore measures to limit the proliferation of disinformation and promote more integrated online discussion, informed by the experiences of disinformation campaigns across the world and efforts to defeat them. (Wardle & Derakhshan, 2017).

2. Methodology

Since many atatistics are not available for researchers in Iran, his paper uses qualitative methods to get a better result.

2.1. Content Analysis

2.1.1. Data Sources

Social media platforms (Instagram, Facebook), news sites, and relevant online forums will be analyzed.

The enhancement of this content will be closely followed, especially around the topics that are emotionally charged and divisive.

2.1.3. Analysis Techniques

Natural Language Processing can help the process of keyword and trend detection.

2.2. Surveys and Interviews

2.2.1. Participants

The surveys and interviews will consist of journalists, media professionals and technology experts, among others. Their opinions and experiences will be collected.

2.2.2. Key Topics

Questions concerning the connection which can lead to the mental health of a person who is misinformed and recommendations for further developments will be asked in surveys. Interviews will be on the topic of mathematical algorithms and media ethics in the spread of disinformation.

2.2.3. Data Collection

Qualitative data will be acquired through face to face interviews.

2.3. Case Studies

2.3.1. Case Selection

There will be a selection of specific cases of false information,

2.3.2. In-depth Analysis

In every instance, elements such as decision ecology, programming, and Facebook's real-time distribution algorithms will be dissected by applying these principles to discourses of disinformation in domestic and in a global media environment.

2.3.3. Comparative Approach

In terms of the other country cases, these will be pitted against local cases of prevention (for example, on hate speech and disinformation on refugees), with the purpose of clarifying the interrelation issues which impact the situation differently.

3. Disinformation and Hate Speech's Mechanisms

The ways in which disinformation and hate speech are spread across social media are not only technological, but are multifaceted, involving psychologically and societally multilingual. An analysis of the characteristics of each of them illustrates how they are merging to carve out the context of meaning that leads to the creation of harmful promotion materials directed at groups like Afghan refugees.

3.1. Role of Algorithms

Algorithms are indeed central to shaping individuals' online experiences, especially on platforms such as Twitter, Facebook, and Instagram.

These algorithms advance the posts that attract an enormous amount of users' engagement. In the most cases, these should be the ones that people will respond to with strong emotions, like disgust, fear, or anger (Scout, 2018). This interest in the involvement of the public in the preparation of the content has opened doors to the most diverse range of disinformation which pops up in the society, particularly through the affected community, that is not visible.

Prevalent sensational untruths are the type that most easily get noticed on the weaker ones, which are the more predominant among these. Hence, as per the study revealed by Zeng et al. (2021), emotional triggering dis-information spread like wildfire compared to viewpoints backed by solid data and got disseminated all the more quickly since it quickly trifled with the users of attention and was very engaging, and that is what algorithms are most attracted to.

One of the groups that were regularly attacked by these stories were Afghan refugees in Iran. Incidences like that together with hate speech often go viral on social media because they enter the sites of alienation masked by them and eventually they are able to amplify social biases which they are associated with. To demonstrate, Essalhi-Rakrak and Pinedo-González (2023) point out that the algorithmic promotion of this content serves not only to extend the reach of fake news but also to further polarize communities. Their working paper on Twitter's hate speech against refugees looks into how the chosen algorithm might unintentionally build up anti-immigrant sentiments thus leading to creating a hate speech zone where hatred is not censured.

Iran is specifically problematic when the spread of such divisive content is of paramount implication, as it is essentially the result of a long-existing societal bias against Afghan refugees. According to Pasquetto, Swire-Thompson, and Amazeen's (2020) research, which algorithm-driven disinformation is the utilization of service psychological vulnerabilities and the existence of

prejudices already in existence, thus, leading to deepening of the teenager situation. Consequently, the websites that insulate from the tangled web of engagement criterionishness, prolonged-periodicly and, throughout all the existing ones of the sort come to revitalize the stories through a different and portray narrow-mindedness as the taboo, at an even higher level, stop date, a more extreme phase.

3.2. Psychological Factors

Cognitive biases, notably confirmation bias, have a central part in defining how people consume and process information. Postulated by Pennycook and Rand (2019), people are predisposed to believe and share the data that fits into their current beliefs, even if it is false. This inclination is particularly visible in the case of Afghan refugees in Iran who are the victims of several neck-deep prejudiced stereotypes and historic mildew. Thus, the spread of disinformation is facilitated. But are these factor and hate speech actually responsible for the total selling of the financial behaviour?

The dishonesty and hate-mongering are deeply rooted in human psychology and, to them, the first essential thing to understanding is exactly what causes such content to resonate with users and spread at such a large scale. Disinformation and disinformation over refugees in Iran also arise

from the fears and anxiety of the people. Remarkable are the researches on social media hate speech that are based on the economic fears of a particular society, such as the fear of loss, the fear of economic competition, the fear of national security, and the fear of cultural differences (Forson, et al., 2024). Conditions in the cases of socio-economic challenges, whether due to bad economy or social disturbances, the threats are much more vivid, resulting in the people to get influenced with disinformation that is going against the minority groups such as Afghan refugees causing problems to the country.

"Otherization" of the process of assimilation further worsens the situation of misinformation. According to Guillén-Nieto et al. (2023), the creation of refugees as "the Other" who are the ones that are a threat to the social, economic, and cultural integrity of the host society is a very important reason for the spreading of hate speech. This dehumanization of the problem allows misinformation to spread. The individuals circulating misinformation tend to think of those people as pointless of having an attitude of gratitude or shielding. The narrative is very familiar, for instance, in case of Afghan refugees who are pointed at as a primary cause of cultural as well as economic downfall. Therefore, it becomes easier to dehumanize them through disinformation campaigns since these are the presumed tools that make them prone to hate speech and, eventually, violence (Weitzel, 2022).

White supremacy and xenophobia are the two underlying psychological drivers behind the perpetuation of false information and hate speech across social media platforms. Bringing these areas of anarchy is not only about the eagle and the tiller, for example, the improved vigilance framework or, instead, a hard-hitting deconstruction of the hidden social disaggregation that forces defamatory and offensive attitudes about Afghan refugees.

4. Mitigation Strategies

An effective disinformation and hate speech relieving strategy, specifically in cases of social media targeting vulnerable groups such as refugees from Afghanistan in Iran, necessitates the adoption of a fully-integrated and diversified solution. Central tactics revolve around building on digital tools, educating people about media through media literacy, transparency in algorithmic method, and pushing an agenda that supports the wider participation of civic society and tech companies with the government

4.1. Digital Tools and Platforms for Countering Hate Speech

The most significant of the methods to tackle the issues of disinformation and hate speech is in the use of digital tools such as fact-checking mechanisms and AI-based content moderation systems. According to González-Baquero et. al. (2022), machine learning models can be trained on large databases to recognize patterns in hate speech so they can immediately flag or remove offending content. These tools will use NLP techniques and AI to find words full of hate, xenophobic narratives,

and fake news targeting Afghan refugees. AI-based moderation tools alleviating the burden on systems like Facebook, Twitter, and YouTube are showing promising results, but issues regarding the adequacy of these systems to non-Western countries, such as Iran, where the language, cultural nuance, and political implications are much more subtle, still persist.

One of the crucial functions of fact-checking organizations is to debunk the false claims about the refugees. These independent fact-checkers collaborate with social media platforms to check the validity of the claims and ensure that the information is accurate. Typically, they address the lies that are targeting the Afghan refugees. Such platforms as organizations cooperating with the major social media networks, like Facebook, are responsible of trying to limit misinformation. Although the situation is not perfect as government supervision is not present in these areas, such the case of Iran, where regulations regarding fact-checking organizations are weaker. Without a well-functioning regulatory body, the application of countermeasures is weakened which it should be accountable to the platform and the government alike.

4.2. Promoting Digital Literacy

Not only recognizing but reacting to the disinformation and hate statements on the Internet by the users are the primary tools of the fight. In this case, Lewandowsky et al. (2017) posits that digital technology skills should be taught as the new literacy needed in the 21st Century. Programs that teach users how to critically assess information they find online can actually uplift them out of their gullibility to manipulation.

Actually this would be the missed opportunity for digital literacy if it only focused on detecting fake news. The root of the main psychological mechanisms such as confirmation bias, and emotional manipulation that lead disinformation should be sought for. Educational campaigns are essential for specific regional and cultural contexts such as Iran that might introduce intolerant and harmful narratives about Afghan refugees, and empower users to identify them, resist the inclination to engage with them, or not pass it on.

4.3. Collaboration between Governments, Civil Society, and Tech Companies

The initiation of a more extensive digital debate by involving all those who are affected is also a must to achieve a more inclusive digital world. Social media companies must follow the instructions of the legislation and work together with the government in order to adopt more fair and transparent content moderation practices. It is the job of the government to create laws that are in line with the implementation of these non-disinformation and hate speech rules, as well as to keep it clear that platforms that are negligent in their use will be terminated.

Besides spotting and calling out the issues derived from online disinformation, civil society organizations can do more actions such as defending migrants' rights and take both states and tech companies to account. According to the findings of Milojević and Reddan (2020) which the involvement of civil society is crucial in the monitoring and the abolishment of hate speech and disinformation against the marginalized communities and to keep the policies inclusive and human right-oriented.

Tech companies should also commit to full disclosure of their content moderation strategies, which will include how the algorithms work and the steps they take to stop hate speech. Collaboration is needed among these sides not only for finding the right solutions that will encounter the problems of misinformation but also in the power regions of the world like Iran where the Afghan refugees are often made as scapegoat.

5. Case Study: Legal Insecurity and Citizenship Issues

Afghan refugees are the objects of disinformation campaigns on a regular base, mainly over internet sites. These campaigns lead to the marginalization of Afghan refugees by the use of different actions such as creating lies and promoting hate speech, which in turn results in the growth of the hostile and violent reactions towards them (Aldamen, 2023). Moreover, social media algorithms tend

to zero in on the most sensational materials, which deepen the divisions rather than repairing them, and reinforcing the negative stereotype of Afghan expatriates (Essalhi-Rakrak & Pinedo-González, 2023).

The Afghan refugees living in Iran are part of the most vulnerable collaterals of undocumented migration who are disposed to the host state's rapid rule-making changes. Many of them are incapable of getting necessary documents and thus, can't take advantage of such vital services as health, education, and protection (Essalhi-Rakrak & Pinedo-González, 2023). Those that have refugee status are also at risk of being forcibly removed because of the government's attempts to balance their budgets and maintain their sovereignty (Zahra & Gholamreza, 2021). Iranian laws do not recognize Afghan children born in Iran as citizens, and this is why they are often stateless. Consequently, they are deprived of rights and opportunities (Zahra & Gholamreza, 2021). Their stay backstage society hinder them from participation in Iranian society at a longer vista (Pasquetto et al., 2020). The possibility of deportation, the state of poverty, and their isolation from the rest of society also contribute to mental health problems among Afghan refugees and their vulnerability. Many suffer from trauma due to the violence they fled in Afghanistan, the discrimination they face only adds to their suffering while they are in Iran. However, such services, especially mental health ones, are limited, leaving refugees to deal with their own mental illnesses, such as anxiety, depression, and PTSD (post-traumatic stress disorder) without the help of someone else (Pennycook & Rand, 2019).

Afghan refugees are often targets of socially biased and xenophobic Iranian society as economic struggles and refugee invasion are increased, the situation becomes worse. Usually, media and social platforms portray refugees as characters of ill will or as the main source of the economic problem. It has cultivated an aggressive environment in which hate speech and disinformation aimed at Afghan refugees (Aldamen, 2023) are prevalent. Cultural exclusion furthermore ostracizes Afghan refugees as there are many instances when they are treated as The Others which in turn obstructs their preceding as they are Iranian(Guillén-Nieto et al., 2023).

Lately, this negative sentiment against the exiles has turned into a structured form ever more notably on such social media like Twitter where those sentences such as #اخراج_افغانی_مطالبه_ملی (Expel-Afghan-National-Demand) trend frequently. Nevertheless, these feelings often originate from wrong data about crime rates and economic competition, since the findings of the research illustrate that the major part of the cases that are connected with Afghan refugees impact the people from their community and other citizens of Iran are not always involved in the cases. There are tweets on the social media platform that have been released in X in 2024, the tweets that are racist to the legal and illegal immigrants with high verbal violence, they even tease and insult, in fact, most part of Iranian experts and officials prefer the solution of the problem through the return of illegal immigrants and the legal, as well as they only disagree on the model but in the cyberspace, the issue seems to go in a different way. "We say throw out the Afghans, they say that they are foreigners, throw out the Iranians, that means the cucumber has an IQ 10 higher than this", "Shi'a Salafist Toronto in choosing Afghanistan lessens the citizen's Iran! It is such an existential thought that one can infuse with society, "15m immigrants come into a country and people of that country don't vote for them yet they clot at the border thinking they have a right to live and suck the life out of our people and our country without our permission, is a sort of transgression both in terms of intellect and law", "Referring to Afghans, by reason of them being illegal citizens, don't call them illegal citizens because there is no difference between legal and illegal, but why?

Let's say Afghans are already illegal aliens (refugees arriving from abroad) what are we going to say to the legal ones that are inside our country as well? "First of all, there is no such thing as anti-immigrant(!) foreigners in Iran! It's all on the surface and easy to discover being a political talking point!It is nothing else but the truth! Changing the fabric and the population of Iran is a favor done by the statesmen.

But what has really gained ground in the virtual space as well as some media in the last few months is the phenomenon which is in a way related to Afghanophobia, but at the same time it is more than just Afghanophobia. Some data analysis and monitoring firms do studies on the web and give feedback accordingly. "Lifeweb", a data collection and analysis firm, has shared a report on this

doi:10.20944/preprints202410.0102.v1

virtual wave to the farhikhtegan newspaper. According to the authors of the aforementioned report, "The Afghanistan-centric posts are made about 10,000 times a day on the social network X (Twitter). This is not only one of the most discussed subjects in the last few months but also among X Farsi users in the last six months and hence we can say that it is the presence of Afghan immigrants in Iran ." The human aspect of the refugees who have been taken in by Iranian people and the link to "Afghanophobia" are displayed by the `Status of Achievement` presentation of the latest incoming information of the report.

In response to the fact that a negative approach to the issue of Afghans was just about 20% of the content in April 2023, it has been the case that, in the last month, more than 50% of the mentioned themes, which are presented daily in this X overflow through Telegram and Instagram and engage the wide audience, are related to the" Afghanistan-phobia" idea. This action was initiated on the first day of 1401. The news of the stopping of Afghans in the environment near Chitgar Lake, the argument over the Hirmand border, the Iran-Afghanistan border seeds, and the Afghans swimming in Naqsh-e Jahan Square, which had a larger negative impression in the virtual space were among these.

October 2023, was the period of even more discussions on these two issues, where one of the main topics was the Afghans in the country and that of the usage of the hashtag "#Afghan-Eject-National-Demand". Even before the initiation of the al-Aqsa storm operation, this trend was top on Twitter. Also, in October 2023, some interethnic confrontations occurred between Iranians and Afghans in both, Twitter and the real world. The case of a police officer in Bushehr, who was shot dead by an Afghan refugee, the detention of an Afghan crew which had been robbing houses in Alborz, the Iranians forcible seizure of Afghans in Shahryar and Qazvin are only a few instances to mention.

The virtual movement is explicitly arranged. For sure, a rosen that this story is properly managed is the hash-tag # اخراج_افغانی_مطالبه_ملی.

Besides a few, most of the group have started their Twitter accounts in 2021, 2022, and 2023. Almost 100 accounts are opposing the plan of the political action and the role of the organizers in the cyberspace. There are no private data on these pages except hate speech. There are also claims of holding training sessions for certain twitter users on this matter. Meanwhile, some of the Iranian Human rights workers in the immigration sector have been victims during this time. The only way out for this faction is the mass expulsion of through the route of Pakistan. Besides, the political inclination of this group is somehow affiliated with the government. But some people that are the pure ones won't say

In some of their tweets, the users are calling Afghans parasites and calling Afghanistan Angalestan(country of the parasites).

The gossip launched says that Afghans are inhuman. So the Iranians who would accept the oppression of the Rwandans became neutral on restrictions concerning foreigners and on segregation restricting them to informal, bad, and low-wage jobs that the indigenous citizens don't want to do are common in sectors such as construction and agriculture. This is another important reason they are still poor and insecure (Guilén-Nieto et al., 2023). In addition, discrimination from the society also narrows their chances, as many employers are hesitant to offer fair salaries to Afghan people (Forson et al., 2024).

Even if Afghan children go to under-resourced educational institutions, they can be prone to being bullied or may drop out, because of these hardships (Essalhi-Rakrak & Pinedo-González, 2023).

However, if the Iranian society allows Afghan to be treated as they are not humans, the access to healthcare gets uneven. Most of Afghan refugees do not enjoy the luxury of health insurance making them to be not covered and so these services mainly become either inaccessible or too expensive for them (Zahra & Gholamreza, 2021). Due to these viral misinformation, even those who access healthcare services will be discriminated, resulting in their inability to access healthcare facilities (Bakir & McStay, 2018).

6. Conclusion

The dissemination of disinformation and hate speech against Afghan refugees on Iranian social media has become a dangerous issue due to the biases of the computer program and the psychological aspects. It is innocuous that Social media platforms, with their engagement-focused algorithms, have inadvertently contributed to the spread of harmful rhetoric and thus exacerbated societal divisions. It is necessary for this that we adopt a strategy that includes fact-checking, algorithmic reform, and digital literacy initiatives in order to cope with it. Policymakers and platforms that acknowledge the definitive causes of disinformation can produce direct scenarios consequently through the provision of inclusive and less biased digital environments. This in turn also helps to support constructive public discourse and reduces harm to vulnerable populations like Afghan refugees.

References

Aldamen, Y. (2023). Xenophobia and hate speech towards refugees on social media: Reinforcing causes, negative effects, defense and response mechanisms. Societies, 13(4), 83. https://doi.org/10.3390/soc13040083

Bakir, V., & McStay, A. (2018). Fake news and the economy of emotions: Problems, causes, solutions. Digital Journalism, 6(2), 154-175. https://doi.org/10.1080/21670811.2017.1345645

Essalhi-Rakrak, A., & Pinedo-González, R. (2023). #EspañaInvadida. Disinformation and hate speech towards refugees on Twitter: A challenge for critical thinking. Profesional de la Información, 32(3), e320310. https://doi.org/10.3145/epi.2023.may.e320310

Forson, C., Healy, G., Öztürk, M. B., & Tatli, A. (2024). Research handbook on inequalities and work. Edward Elgar Publishing.

González-Baquero, W., Amores, J. J., & Arcila-Calderón, C. (2022). From belief to hate: An analysis of the beliefs supporting hate speech towards migrants on Twitter. Tripodos, 53, 39-51. https://doi.org/10.51698/tripodos.2022.53.03

Graves, L. (2018). Boundaries not drawn: Mapping the institutional roles of journalism in the fact-checking movement. Journalism Studies, 19(5), 613-631. https://doi.org/10.1080/1461670X.2016.1196602

Guillén-Nieto, V., Pais, A. D., & Stein, D. (2023). From fear to hate: Legal-linguistic perspectives on migration. De Gruyter Mouton.

Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "post-truth" era. Journal of Applied Research in Memory and Cognition, 6(4), 353-369. https://doi.org/10.1016/j.jarmac.2017.07.008

Milojević, A., & Redden, J. (2020). Social media, algorithmic harms, and the amplification of hate speech in digital spaces. Journal of Media Ethics, 35(3), 177-189. https://doi.org/10.1080/23736992.2020.1784468

Pasquetto, I., Swire-Thompson, B., & Amazeen, M. (2020). Misinformation and disinformation: Socio-psychological vulnerabilities in the digital age. Journal of Media Ethics, 35(3), 177-189.

Pennycook, G., & Rand, D. G. (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. Proceedings of the National Academy of Sciences, 116(7), 2521-2526. https://doi.org/10.1073/pnas.1806781116

Tufekci, Z. (2018). Twitter and tear gas: The power and fragility of networked protest. Yale University Press.

Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policy making. Council of Europe.

Zahra, N., & Gholamreza, M. (2021). Afghan refugees in Iran: Societal integration and discrimination. Migration Studies Journal, 12(3), 45-67.

Zeng, J., Liu, H., & Wang, Y. (2021). The spread of disinformation on social media: A socio-political challenge. International Journal of Information Management, 58, 102-116.