YouTube Algorithm and Disinformation Issues

Jong-bin Lee*

ABSTRACT

With the proliferation of smartphones and the advancement of internet speeds, the usage of social media has experienced a significant and rapid surge. In this evolving landscape, the predominant platform for video content is none other than YouTube. Despite operating a robust recommendation algorithm that leverages collaborative and content-based filtering technologies, YouTube grapples with the formidable challenge of generating misinformation due to its vast and diverse user base. Indeed, the scale of fake news, when examined statistically, is undeniably substantial. This misinformation has given rise to a myriad of societal issues, prominently manifesting in the form of the echo-chamber effect and the filter bubble effect. Furthermore, these effects have played a pivotal role in various instances of political exploitation, adding an extra layer of complexity to the challenges at hand. The algorithm's inherent efficiency, coupled with widespread user engagement, has inadvertently facilitated the production and dissemination of misleading information, necessitating a comprehensive approach to address the multifaceted challenges posed by misinformation on the platform.

Key Words : YouTube Algorithm, Social Problems, Recommendation Algorithm, Echo-Chamber Effects, Filter Bubble Efftets, Political Abuse

I. Introduction

With the proliferation of smartphones and the improvement of internet speeds, the usage of social media continues to show an upward trajectory. Among numerous video platforms, YouTube currently holds the largest user base. This has led to a trend where people acquire necessary information not through direct web searches but rather by consuming video content created by others. From a technological YouTube standpoint, is recognized second-largest search engine globally, following Google. Notably, YouTube employs a powerful recommendation algorithm to ensure users stay within the platform, tailoring content suggestions based on users' diverse viewing histories. As a result, user engagement on YouTube continues to rise. In addition to its consistent user retention strategies, YouTube

introduced YouTube Shorts in 2021. Unlike traditional videos on the platform, YouTube Shorts allows users to create and share videos with a maximum duration of 1 minute. These shorts feature automatic loop playback and a seamless transition to the next video with a simple scroll. Interestingly, even YouTube Shorts are subject to the platform's recommendation algorithm, further enhancing user interaction.

The contemporary news landscape has transitioned predominantly to video media, with leading newspapers such as Chosun Ilbo and JoongAng Ilbo establishing channels on YouTube. However, it is crucial to note that channels delivering news content on YouTube may not always be affiliated with certified institutions or media outlets. Channels covering a diverse range of topics, including information dissemination, can be operated by

[•] First Author: Yonsei University Department of Applied Information Engineering & FTI Co. Ltd, Belle.lee@Yonsei.ac.kr, 정회원 논문번호: 202401-006-E-RN, Received December 31, 2023; Revised March 15, 2024; Accepted March 16, 2024

individuals. As mentioned earlier, YouTube employs a personalized algorithm that includes both news outlets and user-generated content in recommendation lists. YouTube utilizes visual cues like thumbnails and video cover images to encourage user clicks. In contrast to formal stances adopted by news agencies, individual channels may use more sensational language, given platform's flexibility. the Consequently, users naturally encounter videos produced by individuals, and with the introduction of YouTube Shorts, even brief, captivating videos have the potential to capture users' attention. Moreover, individual-produced videos, lacking objectivity and accuracy, may be perceived as factual by the YouTube user base, influencing public perception.

Historically, inaccuracies in information have been a source of societal confusion, and the advent of video media, coupled with increased accessibility to information, has further heightened these concerns. In the specific context of this paper, our primary aim is to delve into the workings of the current YouTube algorithm and shed light on the various societal issues stem specifically from the recommendations. Also, this paper seeks to provide an in-depth exploration of the YouTube algorithm's approximate workings and, more importantly, to unravel the challenges and problems associated with misinformation caused by this algorithm. Specially this method faces challenges such as the "cold start problem" for new users or items and the "long tail problem" of emphasizing popular items, despite its effectiveness in providing recommendations based on extensive user and item data

By focusing on the implications of YouTube's recommendation algorithms, we aim to contribute to a nuanced understanding of the specific issues arising from information inaccuracies within this digital platform.

II. Current Technology and Problems with YouTube Algorithm

2.1 Circumstance of YouTube Korea

Before delving into the issues arising from the YouTube algorithm, it is essential to examine the impact of YouTube in South Korea. According to Playboard, a specialized YouTube statistics analysis firm, as of the end of 2020, South Korea had 97,934 revenue-generating YouTube channels (channels with over 1,000 subscribers), resulting in a ratio of one channel for every 529 individuals. In comparison, the United States, considered the birthplace of YouTube, had a ratio of one channel per 666 individuals. Additionally, according to DataReportal, an internet data analysis agency in Singapore, South Korea recorded an overwhelming usage rate of 39.9% on YouTube, as revealed in the 2022 Monthly Social Media App Usage report[Fig 1]. This far surpassed other platforms, with TikTok at 16.9%, KakaoTalk at 11%, Facebook at 7.8%, and Daum at 7.7%. Notably, TikTok experienced the highest growth at 22%, while YouTube demonstrated a 5% increase in the annual growth rate. Despite the relatively modest growth rate, the substantial user base of YouTube remains evident, consistently maintaining its dominance in the market[1].

These statistics reflect the global trend, indicating continuous growth of video platforms in South Korea. Furthermore, Mobile Index, an app analysis service based on Data Management Platform (DMP), reported that as of September 2022, the monthly active users (MAU) of YouTube in South Korea surpassed 43.19 million, accounting for 83% of the total population of 51.78 million. The average monthly viewing time on YouTube in South Korea was reported as 30 hours and 34 minutes, significantly exceeding the global average of 23 hours and 24 minutes. Moreover, on average, YouTube users in South Korea accessed the platform for 16.9 days per month. The frequency of



Fig. 1. Time spent using social media apps in South Korea

usage was observed to increase with younger age groups, with teenagers accessing the platform for 20 days, individuals in their 20s for 19.1 days, those in their 30s for 16.7 days, 40s for 16.1 days, 50s for 16.3 days, and those aged 60 and above for 15.8 days. This indicates a higher frequency of YouTube access among the younger demographic^[1,2].

2.2 YouTube Algorithm

YouTube has not publicly disclosed the specific algorithm it employs when recommending videos to users. However, it is widely known to utilize machine learning methods based on collaborative filtering and content-based filtering technologies.

Collaborative filtering stands out as a robust recommendation system that capitalizes on user behavior records or preferences personalized suggestions. By scrutinizing interactions among users, collaborative filtering excels at predicting or recommending items favored by users with similar tastes. Despite its effectiveness in providing recommendations based on extensive user and item data, this method faces challenges such as the "cold start problem" for new users or items and the "long tail problem" of emphasizing popular items. To address these challenges, YouTube has integrated content-based filtering [Table 11 recommendation system. Content-based filtering videos based recommends on the intrinsic characteristics and content of the videos themselves,

Table 1. Method of content-based filtering

1. Profile Creation	Collecting information on user preferences and characteristics to generate user profiles, including keywords, genres, and rated items.
2. Item Feature Analysis	Analyzing the characteristics of each item, extracting keywords, genres, descriptions, etc.
3. Item Profile	Generating profiles for each
Creation	item based on its features.
4. User and Item	Comparing user profiles with
Profile	item profiles to predict items
Comparison	of interest to the user.
5. Recommendation Generation	Adding items that best reflect user interests to the recommendation list.

irrespective of user preferences or behavior. This technique analyzes the features of each video, compares them to user profiles, and selects relevant items for recommendation. Although content-based filtering effectively tackles the cold start problem and offers personalized recommendations, its reliance on item characteristics alone can limit its ability to capture diverse perspectives and individual preferences^[3]. In response to these limitations, YouTube has developed hybrid recommendation systems that seamlessly blend collaborative filtering and content-based filtering. This hybrid approach aims to leverage the strengths of both methods, providing a more comprehensive and accurate recommendation system. By combining collaborative filtering's ability to capture user preferences based on past interactions with content-based filtering's focus on video characteristics, YouTube enhances the overall recommendation accuracy. This integration ensures a more nuanced understanding of user preferences and effectively addresses the challenges posed by the cold start problem and the long tail problem. The continuous evolution of YouTube's recommendation system reflects its commitment to delivering a personalized and engaging user experience. The synergistic combination of collaborative content-based filtering techniques contributes to YouTube's ability to cater to the diverse preferences of its vast user base^[3,4].

The results obtained from collaborative filtering and content-based filtering undergo a nuanced ranking process before being presented to users. This crucial step involves the consideration of various factors to ensure a curated and personalized list of recommended videos. The key factors taken into account during the ranking process include^[5]:

- Personalized Interest: Calculating personalized interest by analyzing user preferences, behavior, and past interactions.
- · Current Popularity: Prioritizing videos that are currently popular among a large number of users.
- Diversity: Ensuring diversity in the recommendations by including videos from various genres and topics in the ranking.
- · Playback Time: Considering the playback time of

- videos to recommend content that aligns with user time preferences.
- User Behavior: Adapting the recommendation list based on the user's previous video views, clicks, and liked videos.
- Recommendation Algorithm Weights: Assigning weights to the results of collaborative filtering and content-based filtering, determining the significance of each video in the recommendation list.

This meticulous ranking process, integral to YouTube's dynamic recommendation system, is designed to continually enhance user experience. It not only reflects the platform's commitment to delivering personalized content but also underscores its ongoing efforts to refine and improve the recommendation algorithms over time. By considering a diverse array of factors, YouTube aims to present users with a thoughtfully curated list of videos that align with their individual interests and preferences, thereby fostering a more engaging and tailored user experience.

2.3 definition of disinformation

The terminology used to characterize news containing inaccurate information encompasses a range of terms, and the absence of clear-cut criteria contributes to the interchangeable use of these terms. Commonly encountered expressions include "fake news," "rumor." "misinformation," "disinformation," collectively falling under the umbrella of false information. However, these concepts are often used interchangeably and lack strict delineation. Among these terms, "fake news" stands out as the most frequently employed, signifying instances where media reporting deviates from factual accuracy. "Misinformation" denotes unverified information that may possess some basis in truth, while "rumor" closely aligns with misinformation, predominantly spreading through informal channels like word of mouth. Terms like "misinformation" or "disinformation" involve the deliberate distortion of facts with the intent to deceive^[6].

Despite the distinctions, the definitions of these

concepts remain fluid, contributing to potential confusion. In this paper, false information is defined as any information lacking a clear source and not verified as true. Even if the content itself may hold factual accuracy, it is categorized as false information if sufficient reliability is not established. It is crucial to note that false information, regardless of the specific term used, has the potential to instigate various societal issues. These may include public confusion, erosion of trust in information sources, and the manipulation of public opinion. False information can contribute to the spread of baseless beliefs, exacerbating social divisions and hindering informed decision-making processes^[7].

Building upon this foundation, the paper aims to explore the interconnected societal issues arising from false information, particularly as related to the YouTube algorithm.

2.4 South Korea's Seriousness of Fake News

Overall, the viewership of YouTube news among Koreans is 1.5 times higher than the global average. However, a significant majority of the Korean population, accounting for 89%, perceives fake news as a serious issue. It is evident that there is a general aversion to fake news among the citizens. Nevertheless, according to a global survey conducted by Ipsos, encompassing 25 countries to investigate perceptions of "fake news," 85% of respondents in Korea reported having fallen victim to fake news^[8]. This statistic underscores the undeniable impact of fake news, emphasizing the importance of not underestimating its influence. Notably, YouTube emerged as the primary platform for disseminating false information, with 22% of internet users admitting to distributing misinformation on the platform^[9]. Additionally, 34% reported having watched or received videos on YouTube that were perceived as fake news. This signifies that one in three YouTube users has encountered or shared content classified as fake news. Notably, this trend was particularly pronounced among individuals in their 20s and the elderly, representing both ends of the age spectrum^[10].



Fig. 2. the route through which false information is circulated the most



Fig. 3. Fake News Access Experience



Fig. 4. Recognition towards fake news

Furthermore, in a survey targeting 1,200 adults aged 20-59, both male and female, the most crucial factor identified in fake news was its perceived political intent. This suggests a significant likelihood that the production of fake news is driven by political motives. Additionally, a study conducted by Embrain Trend Monitor among men and women aged 19-59[Fig. 4] found that over 85% of respondents expressed concern about the excessive prevalence of fake news, attributing it to the deepening societal divisions^[8,11]. Despite the continuous increase in YouTube viewing time in South Korea, the production of fake news persists, and the public recognizes it as a pressing issue in society.

2.4.1 News/Politics on YouTube and YouTube Algorithm

The operation of the YouTube algorithm is not determined by individual choices but is rather essential and automatically executed during the use of YouTube. The provision of tailored content in this manner has significant advantages from a commercial perspective, as it can extend user engagement on the platform. However, it raises concerns about the

limitation of the types and content of the shown material within a certain framework. Currently, there over 500 channels dedicated to "News/Politics"[4] category on YouTube. Among them, channels are operated not only by media organizations but also by individuals, including politicians. Although media channels generally have higher viewership than individual channels, as of December 2023, channels operated by individuals are not significantly lagging behind. Moreover, the sponsorship feature on YouTube, known as Super Chat, is popular enough to involve substantial monetary transactions[Fig. 5]. According to prior research, videos produced by media outlets and those created by individuals are not consistently classified under the same category.

Additionally, there was prior research aimed at verifying recommendation tendencies through AI recommendation algorithms. In this experiment, two accounts were created, with Account 1 exclusively viewing news from media outlets and Account 2 exclusively consuming User-Generated Content (UGC). Three keywords were searched, and four videos were selected for each keyword. After watching these videos, the initial recommended videos at the time of account creation and the subsequent recommendations after the experiment were compared on the YouTube homepage. The experiment involved selecting four videos for each keyword, and the test was conducted by searching for keywords in the top search bar to find and watch the videos. The chi-square test results for the experiment indicated an association between YouTube algorithm recommendations and the type of content viewed (news or UGC) based on the accounts^[3].

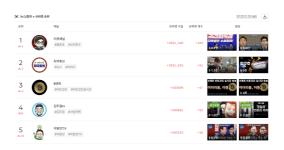


Fig. 5. YouTube Korea News/Political SuperChat Ranking

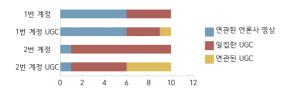


Fig. 6. Experiment Outcome of YouTube Algorithm

Specifically as [Fig 6], if a user exclusively watches videos created by media outlets, the recommendation ratio between media outlet-produced and user-generated content (UGC) was 6:4. On the other hand, if the user only consumed videos created by individuals, this ratio shifted to 9:1[3]. Although videos unrelated to the searched keywords also appeared, the overarching conclusion is that when users predominantly watch personally created videos with a focus on the 'News/Politics' category, YouTube assumes an interest in channels operated by individuals.

2.5 Echo-Chamber Effects

In tandem with the surge of fake news on YouTube, the emergence and fortification of echo-chamber effects among the platform's user base pose a substantial concern. An echo chamber denotes a scenario in which individuals encounter information that aligns with and reinforces their pre-existing beliefs, creating a self-perpetuating loop of confirmation bias.

YouTube's recommendation algorithm, designed to amplify user engagement, inadvertently contributes to the formation of echo chambers. The algorithm, propelled by collaborative filtering and content-based filtering, frequently suggests videos akin to a user's past engagements. While this personalized strategy seeks to cater to individual preferences, it unintentionally restricts the range of perspectives users are exposed to^[6].

Users may discover themselves confined within a content bubble that mirrors their current worldview, curtailing exposure to diverse opinions and alternative viewpoints. This echo-chamber effect can deepen the entrenchment of existing beliefs, fostering polarization, and impeding open discourse. The algorithm's emphasis on user engagement metrics,

such as watch time and clicks, may prioritize content that evokes strong reactions or aligns with pre-existing beliefs, thereby amplifying the echo-chamber phenomenon.

Research underscores that individuals entrenched in echo chambers are more likely to have their existing beliefs reinforced while being less exposed to information challenging their perspectives. This holds broader societal implications, including heightened polarization, diminished understanding between distinct ideological groups, and a potential erosion of shared realities^[12].

Recognizing and remedying the echo-chamber effects within YouTube's recommendation system becomes pivotal for cultivating a more diverse and inclusive information environment. Striking a delicate balance between personalized content recommendations and ensuring exposure to a spectrum of perspectives is indispensable for nurturing a robust, healthy, and well-informed public discourse on the platform.

2.6 Filter Bubbling

While echo-chamber effects on YouTube contribute to reinforcing existing beliefs by presenting users with content that aligns with their preferences, another noteworthy phenomenon is the emergence of filter bubble effects. The filter bubble, coined by internet activist Eli Pariser, refers to the personalized information ecosystems that result from algorithmic curation, where users are selectively exposed to content that matches their past online behavior, preferences, and demographics.

In contrast to echo chambers, filter bubbles are characterized by a more individualized and personalized information experience. The YouTube recommendation algorithm, in its pursuit of enhancing user engagement, tailors content suggestions based on an individual's watch history, likes, and clicks. This personalized curation can create a unique information bubble around each user, where their exposure to diverse perspectives is selectively limited^[3].

The filter bubble effect is driven by algorithms prioritizing relevance and user engagement metrics. As a result, users may find themselves in a curated

digital space where their pre-existing preferences are continually reinforced, and dissenting or diverse viewpoints are underrepresented. Unlike echo chambers, filter bubbles focus on the individual's online journey, crafting a digital environment that aligns with their preferences.

The implications of the filter bubble extend beyond reinforcing existing beliefs. Users may unintentionally shielded from information challenges their worldview, hindering comprehensive understanding of various issues. While the echo chamber tends to foster polarization by reinforcing group beliefs, the filter bubble accentuates an individual's isolation from diverse content, potentially leading to a more personalized yet narrow information landscape^[12].

Addressing the filter bubble effect necessitates a critical examination of the algorithms that govern content recommendation on platforms like YouTube. Striking a balance between personalization and exposing users to a variety of perspectives is essential for fostering a more open and informed digital space. Acknowledging the dual challenges posed by both echo chambers and filter bubbles is crucial for designing recommendation systems that prioritize diversity, transparency, and a more comprehensive exchange of ideas.

2.7 Political Abuse

The confluence of filter bubble and echo chamber effects in online platforms, particularly in the political domain, has given rise to a concerning phenomenon known as political abuse. This phenomenon manifests when individuals, groups, or entities exploit the algorithmic mechanisms of platforms like YouTube to manipulate or weaponize political content for their gain. Understanding the intertwined dynamics of filter bubbles and echo chambers within the political landscape is crucial for comprehending the potential consequences of political abuse. The combination of filter bubble and echo chamber effects in the political landscape sets the stage for a multifaceted and complex phenomenon - political abuse^[13]. This phenomenon involves the strategic exploitation of algorithmic tendencies to curate personalized content and foster ideological echo chambers. Recognizing the interplay between these effects is essential for devising strategies to mitigate political abuse, preserve information diversity, and uphold the integrity of democratic discourse in the digital age.

2.7.1 Filter Bubble and Political Abuse

The filter bubble effect, driven by personalized content curation, can be strategically employed for political manipulation. Bad actors can capitalize on the algorithm's tendency to prioritize relevance and engagement, tailoring political content to cater specifically to users' existing beliefs. In this context, political abuse involves the deliberate creation and dissemination of content that reinforces a particular political narrative, isolating users within their personalized information bubbles.

By leveraging the filter bubble, political abusers can ensure that users are consistently exposed to content that aligns with a specific ideological agenda. This strategic manipulation of information flow can amplify confirmation biases, limit exposure to alternative viewpoints, and foster an environment where individuals are more susceptible to propaganda or misleading political narratives^[14].

2.7.2 Echo Chamber and Political Abuse

Echo chambers, characterized by the reinforcement of existing beliefs within homogeneous communities, offer a fertile ground for political abuse. In a politically charged echo chamber, misinformation or politically biased content can spread rapidly and be amplified by like-minded individuals. Political abuse within echo chambers involves the intentional dissemination of content to exploit the emotional resonance of a particular political narrative within a closely-knit community.

The echo chamber effect facilitates the creation of ideological silos where dissenting opinions are marginalized or excluded. Political abusers can manipulate these closed environments to disseminate propaganda, sow discord, or even engage in targeted disinformation campaigns. By capitalizing on the echo chamber's tendency to strengthen group identity, political abuse can deepen societal polarization and

erode trust in democratic processes^[14].

III. Conclusions

In conclusion, while YouTube's video recommendation algorithm undoubtedly enhances user experience, it poses inherent challenges, particularly in the realms of societal and political biases. The inadvertent propagation of misinformation remains a pressing concern, given the platform's vast content landscape. Addressing this issue requires multi-faceted approach, involving not only technological advancements in algorithm refinement but also legal frameworks to combat misinformation. To improve the YouTube algorithm, targeted adjustments within the news and politics domain could provide nuanced a more content recommendation, fostering a balanced informational environment. Periodic algorithm resets might serve as a preventive measure, ensuring the system doesn't inadvertently contribute to the formation of echo chambers and filter bubbles over time. Moreover, policy interventions focused on strengthening legislation around misinformation can act as a robust line of defense. By holding content creators accountable for the accuracy of their information, legal measures can mitigate the spread of false narratives and uphold the integrity of online information spaces.

While this discussion has introduced echo-chamber effect and filter bubble effect, it's essential to acknowledge the ongoing exploration needed to fully understand and address these complex societal phenomena. This comprehensive approach, combining technological innovation, periodic evaluations, and legal reinforcement, is crucial to navigate the intricate landscape of challenges posed by the YouTube algorithm. In essence, fostering a more responsible and accountable digital environment necessitates collaboration between technology developers, policymakers, and users. By collectively addressing the shortcomings in the algorithm, society can move towards a more informed, diverse, and resilient digital discourse on platforms like YouTube.

References

- [1] https://www.insightkorea.co.kr/news/articleView.html?idxno=112386, Dec. 30. 2023.
- [2] https://www.kpf.or.kr/front/research/selfDetail. do?seq=595768&link_g_homepage=F, Dec. 30. 2023.
- [3] S. Kim, S. Kim, and J. Ahn. "A Study on the bias and regulation of AI algorithm," *Korean Telecommun. Policy Rev.*, vol. 29, no. 2, pp. 111-144, 2022. (https://doi.org/10.37793/ITPR.29.2.4)
- [4] I. Kim and J. M. Kim, "YouTube algorithm and confirmation bias," *The Korean Assoc. Comput. Edu. Conf.*, vol. 25, no. 1(A), pp. 71-74, 2021.
- [5] P. Covington, J. Adams, and E. Sargin, "Deep neural networks for YouTube recommendations," in *Proc. 10th ACM Conf. RecSys* '16, Assoc. Comput. Machinery, pp. 191-198, New York, NY, USA, 2016. (https://doi.org/10.1145/2959100.2959190)
- [6] J. Hwang, "The effect of YouTube's personalization algorithm on perceptions of disinformation experiences," Korean Telecommun. Policy Rev., vol. 29, no. 2, pp. 177-213, 2022. (https://doi.org/10.37793/ITPR.29.2.6)
- [7] S. O. Yun, "A discussion on the concept and scope of fake news," *J. Media Law, Ethics* and Policy Res., vol. 17, no. 1, pp. 51-84, 2018. (https://doi.org/10.26542/JML.2018.4.17.1.51)
- [8] http://www.mhdata.or.kr/bbs/board.php?bo_tabl e=koreadata&wr id=88, Dec. 30. 2023.
- [9] S. Yang, "Reality and challenges of traditional journalism in the era of YouTube journalism," *J. Soc. Sci.*, vol. 31, no. 1, pp. 245-262, 2020. (https://doi.org/10.16881/jss.2020.01.31.1.245)
- [10] J. A. Yang, S. W. Oh, "Use of YouTube videos and experiences of exposure to misinformation," *Media issue of Korea Press Foundation*, vol. 4, no. 8, 2018.08.30.
- [11] J. A. Yang, "News and fake news as perceived by people," *Media issue of Korea*

- Press Foundation, vol. 5, no. 1, 2019.02.25.
- [12] S. Flaxman, S. Goel, and J. M. Rao, "Filter bubbles, echo chambers, and online news consumption," *Public Opinion Quarterly*, vol. 80, no. S1, pp. 298-320, 2016. (https://doi.org/10.1093/poq/nfw006)
- [13] M. Kim, "The effect of perception of the usefulness of YouTube algorithm recommendation on media trust on YouTube: Mediated effects of perceived harm," Confirmation Bias, and Privacy Concerns, Korean J. Commun., vol. 21, no. 4, pp. 7-42, 2022.
- [14] M.-H. Kang, "Does YouTube reinforce confirmation bias?: A study on the political uses of YouTube and its effects," *Korean J. Commun.*, vol. 20, no. 4, pp. 261-288, 2021.

Jong-bin Lee



Dec. 2005~May 2019: Director, Intel Co.

Mar. 2020~Jul. 2023: Professor, Sangmyung University

Mar. 2023~Current: Professor, Yonsei University & CEO, FTI Co.

<Research Interests> ICT technology, wireless communications, Data Analysis, Machine Learning

[ORCID:0000-0002-6144-2406]