Fediverse Migrations: A Study of User Account Portability on the Mastodon Social Network

Haris Bin Zia*
h.b.zia@qmul.ac.uk
QMUL
London, United Kingdom

Jiahui He*
jhe976@connect.hkust-gz.edu.cn
HKUST (GZ)
Guangzhou, China

Ignacio Castro
i.castro@qmul.ac.uk
QMUL
London, United Kingdom

Gareth Tyson†
gtyson@ust.hk
HKUST (GZ)
Guangzhou, China

Abstract

The advent of regulation, such as the Digital Markets Act, will foster greater interoperability across competing digital platforms. In such regulatory environments, decentralized platforms like Mastodon have pioneered the principles of social data portability. Such platforms are composed of thousands of independent servers, each of which hosts their own social community. To enable transparent interoperability, users can easily migrate their accounts from one server provider to another. In this paper, we examine 8,745 users who switch their server instances in Mastodon. We use this as a case study to examine account portability behavior more broadly. We explore the factors that affect users' decision to switch instances, as well as the impact of switching on their social media engagement and discussion topics. This leads us to build a classifier to show that switching is predictable, with an F1 score of 0.891. We argue that Mastodon serves as an early exemplar of a social media platform that advocates account interoperability and portability. We hope that this study can bring unique insights to a wider and open digital world in the future.

CCS Concepts

• **Information systems → Social networks**.

Keywords

Fediverse, Mastodon, Platform Migration, Instance Switching, Interoperability

ACM Reference Format:
Haris Bin Zia, Jiahui He, Ignacio Castro, and Gareth Tyson. 2024. Fediverse Migrations: A Study of User Account Portability on the Mastodon Social Network. In *Proceedings of the 2024 ACM Internet Measurement Conference (IMC '24), November 4–6, 2024, Madrid, Spain.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3646547.3689027

*Joint first authors
†Also with QMUL

IMC '24, November 4–6, 2024, Madrid, Spain

ACM ISBN 979-8-4007-0592-2/24/11
https://doi.org/10.1145/3646547.3689027

1 Introduction

The concept of "platform lock-in" has become a major concern for users of online social platforms. Giants, such as Facebook and Twitter/X, have created systems that, while providing a seamless and integrated experience, act as a walled garden that prevents users from easily switching to competing providers. For example, users cannot move their account to another social media platform easily (*e.g.* from Twitter/X to Threads), without incurring significant costs (*e.g.* the loss of social data and their social network). This deliberate obstruction not only stifles competition, but also limits users' autonomy and choice, effectively monopolizing their online social presence. In response to these concerns, regulators have begun to take steps. Most prominently, the European Union has taken legislative action in the form of the Digital Markets Act (DMA) [4]. The DMA implements rules to prevent large platforms from imposing unfair conditions on consumers and businesses. One of its key objectives is to promote greater interoperability between different platforms, which will allow users to seamlessly migrate their data (including their social network) from one platform to another. It is hoped that this will facilitate new market entrants, and spur greater competition.

An early adopter of this form of open, interoperable account migration is *Mastodon* [14]. This is the largest decentralized microblogging platform, which has gained prominence as a alternative to mainstream social media platforms like Twitter/X [3, 8, 9]. Mastodon avoids the pitfalls of platform lock-in by giving users the ability to seamlessly switch their accounts between servers (aka "instances") while maintaining their social graph. This paradigm promotes a diverse and pluralistic social media environment and aligns closely with the principles of the DMA (which naturally encourages competition and user sovereignty). However, despite the growing importance of data portability and user autonomy, there is still a limited understanding of the factors that influence user behavior in such open environments. By investigating these factors, we aim to uncover insights that challenge the dominance of closed social media ecosystems. As such, we argue that Mastodon is the perfect lens through which to study the process of account migration. By demonstrating that it is feasible to build a social media ecosystem that respects choice and promotes data portability, we argue that Mastodon sets a precedent for challenging the status quo of prior walled gardens.

With this in-mind, we gather a first-of-its-kind dataset from Mastodon, covering instance switching behavior. We address the following research questions:

RQ1: How common is instance switching on Mastodon, and what preferences do users exhibit regarding the instances they switch to?

RQ2: What factors influence a user's decision to switch instances, and how does instance switching impact the user's social media engagement, particularly in terms of posting frequency and content?

RQ3: Can we predict whether a particular user is going to switch instances?

Our main findings include:

- Switching is more common than expected, with 1.69% of users choosing to switch at least once. We find that users tend to switch from general purpose instances to topic-focused specific instances.
- A user's social network seems to impact their decision to switch. After switching, the user's follower count increases, and the user tends to start following more people on the new instance. The median number of followers (for all switching users) before and after the switch is 1 and 72, respectively.
- However, switching instances has only a temporary effect on user engagement. The median number of likes that users receive on their posts remains constant at 3 before and after switching, except for the first 6 days after switching, where likes grow (between 4 and 7, median). Users also tend to discuss different topics before and after switching (with a hashtag Jaccard similarity of just 0.197 before/after the switch).
- It is possible to predict if a user will switch their instance (we attain an F1 score of 0.891). The most predictive features are the number of followers and followees. Switching users have only a small number of followers compared to non-switching users, and expanding their social network seems an important reason for users to switch.

2 Background

2.1 Fediverse & Mastodon

The Fediverse is a network of independently owned, operated, moderated, and interconnected "Decentralized Web" servers (known as "instances"). The architecture of the Fediverse is such that no single entity operates the entire infrastructure. Instead, instances collaborate (aka "federate") in a peer-to-peer fashion to collectively offer various types of services (*e.g.* micro-blogging, file sharing, video streaming). This federation is performed using the W3C ActivityPub [1] protocol, which allows instances to subscribe to objects provided by each other. The largest Fediverse platform is Mastodon, an open-source micro-blogging service. Each individual Mastodon instance works much like Twitter/X, allowing users to register new accounts and share **posts** with their followers — equivalent to a tweet. Mastodon instances can also **federate**, whereby users registered on one instance can follow users registered on another instance. This results in the instance **subscribing** to posts performed on the remote instance, such that they can be pushed across and presented to local users.

2.2 Instance Switching

Instance switching in Mastodon refers to the ability of a user to move their account from one instance to another while retaining their identity, followers, followees, and other associated data (as much as possible). Switching instances involves five steps: (*i*) The user creates an account on the new instance to which they intend to move. (*ii*) The user downloads their data from their original instance (in CSV format). (*iii*) The user then uploads this CSV into their account on the new instance. (*iv*) In the new account, the user sets the old account as an alias of the new account, by inputting their old account handle. (*v*) From the old account, the user initiates the one-click switching process, which initiates the transfer of data from the old instance to the new instance.

Upon completion of these steps, the user's old account is redirected to the new account. Thus, interactions directed to the old account are automatically redirected to the new server. It is important to note that while the user's data (including followers and followees) are transferred, the user's old posts do not migrate with the user.

3 Data Collection

To underpin our work, we gather a large Mastodon dataset. Our data collection process involves the following steps:

Discovering Instances. It is first necessary to obtain a list of instances to gather data from. We crawl the full list of instances from the `instances.social`, a comprehensive index of Fediverse instances. This yields a list of 17,646 unique instances. We notice that this list contains non-Mastodon instances (*e.g.* instances belonging to other Fediverse platforms) in addition to Mastodon instances. Thus, we filter out the non-Mastodon instances. This leaves us with 10,904 Mastodon instances.

Identifying Switching Users. We then crawl the profiles of all local users from these 10,904 instances using their directory endpoint using the instance API. We are only able to collect user profiles from 6,899 instances as the rest of the instances were offline. This yields 516,505 unique users. Following this, we fetch the profile metadata of these users using the lookup endpoint and filter those with a "moved" key in the metadata. This field only exists for users who switch instances by following the steps outlined in §2.2. This key further provides information about the user's new instance: the new instance name, username on the new instance, the creation date of the new account, number of followers, followees, and posts. We identify a total of 8,745 users who switch instance.

Posts, and Followees. We next collect the posts of each switching user on both their old and new instances using the statuses endpoint of the respective instances. In total, we collect 1,838,645 posts from old instances and 1,995,698 posts from new instances. Furthermore, we collect the followees of each switching user on both their old and new instances using the following endpoints of the respective instances. Our collection covers 936,876 followee relationships from the old instances and 1,486,081 relationships from the new instances.

Non-Switching Users. Finally, as a baseline, we randomly sample an equal number (8,745) of non-switching users (*i.e.* users without the "moved" key in the metadata) from our original set of 516,505 users. We then crawl all the posts posted by them on their respective instances. This covers 1,906,937 posts from the non-switched users.

Ethical Considerations. The dataset includes both user and post information and therefore might have privacy implications. We exclusively collect publicly available data using official APIs, following well-established ethical procedures for social data [22]. We have obtained a waiver from the ethics committee at the authors' institution.

4 Switching Patterns

We start by exploring the frequency of instance switching, and users' preferences for the instances they transition to and from (RQ1).

4.1 Switching Frequency

In total, we identify 8,745 users who switch their instances. While this constitutes a small percentage (1.69%) of the overall users (516,505), it still represents a considerable portion. This suggests that DMA-enforced account portability could be an appreciated function. To gain a deeper understanding of these instance switches, we examine the frequency of such switches over time. Figure 1b illustrates a time series of the number of switches per month. A clear spike is observable: 63.2% of all switches occur in November and December 2022. The frequency of switches peaks on November 18, 2022, with daily switches surpassing 600. This timeframe closely aligns with the acquisition of Twitter by Elon Musk, after which a substantial number of Twitter users migrated to Mastodon [8]. To validate this further, we investigate the time it takes for switching users to switch instances.

Figure 1c displays the Cumulative Distribution Functions (CDFs) of the time users take to switch instances (calculated as the number of days between the creation date of the new and old account). Overall, users take a median of 171 days to switch. However, we find a noticeable difference in the time taken by users who switch between November and December 2022, compared to those who switch outside of this timeframe. For instance, the median switching time is 39 days for users who switch between November and December 2022, and 253 days for users who switch outside of this period. This shorter switching time between November and December 2022 mirrors the surge in the number of account creations, driven by the migration of users from Twitter to Mastodon [8]. During this period, anecdotally, it was common for new users to join popular or general-purpose instances, before later switching to more topic-focused instances as they familiarize themselves with Mastodon. Therefore, we next proceed to investigate the instance preferences of switching users.

4.2 Instance Preferences

The 8,745 users in our dataset switch from 519 unique instances to 1,936 unique new instances. Interestingly, while the source (old) instances belong solely to Mastodon (since our data crawling is limited to Mastodon instances), the target (new) instances include other Fediverse platform instances. These include Misskey [11] and GoToSocial [7], which are also ActivityPub-supported microblogging platforms. This highlights the interconnected nature of the decentralized social media landscape and the fluidity with which users navigate between different platforms to find communities that best suit their needs.

To explore this, Figure 1a shows a chord plot of the 100 most frequent old-to-new instance pairs. Each line in the chord plot represents the switching path of one user, while the width of the instance indicates the total number of users switching in and out. In line with our hypothesis that new users initially join popular instances, we see that users mainly switch from large (with respect to the number of registered users) general-purpose instances to more topic-focused instances. For example, 36.9% of all switching users switch from `mastodon.social` and `mastodon.online` (flagship Mastodon instances operated by Mastodon gGmbH). The largest proportion (8.3% of these users) join `hachyderm.io` (a Mastodon server primarily comprised of the tech industry).

This is likely because most users do not explore Mastodon extensively upon their first introduction to the platform. This naturally leads users to join large well-known instances. As users gain experience, the federated nature of Mastodon enables them to discover more engaging communities, prompting them to reconsider their initial choices and switch to instances that better suit their preferences. That said, there are also some users who prefer larger communities, thus choosing to switch to popular instances. For example, 6.6% of switching users switch ***to*** `mastodon.social` from smaller instances. However, a large proportion of these users switch from other existing sizable communities *e.g.* 8.2% from `mstdn.social`, and 7.0% from `mastodon.online`. This trend suggests a preference for large instances, potentially driven by factors beyond community size. For instance, users may switch due to limitations or constraints imposed by their previous instance; *e.g.* `mstdn.social` only allows 500 characters in posts.

5 Switching Reasons & Impact

Next, we explore the factors influencing user decisions to switch instances and the subsequent impact on their social media engagement (RQ2).

5.1 Reasons to Switch

There are at least two possible (non-exclusive) reasons for switching across instances: (*i*) Users might decide to switch for ideological reasons *e.g.* they disagree with instance moderation policies; or (*ii*) Users might decide to switch to an instance that is more aligned with their interests and to expand their social network *e.g.* find more relevant users to follow. As the first is difficult to quantify, we investigate the latter. For this, we compare the social network (*i.e.* followers and followees) of all users before and after each switch.

Followers. First, we analyze the users' follower count before and after the switch. Recall that users take their followers with them when they switch instances (see §2.2). Therefore, when calculating the number of followers after the switch, we subtract the number of followers users have on the old instance from the number of followers on the new instance.

Figure 1d plots the CDFs of the number of followers before and after the switch. When comparing the number of followers of all

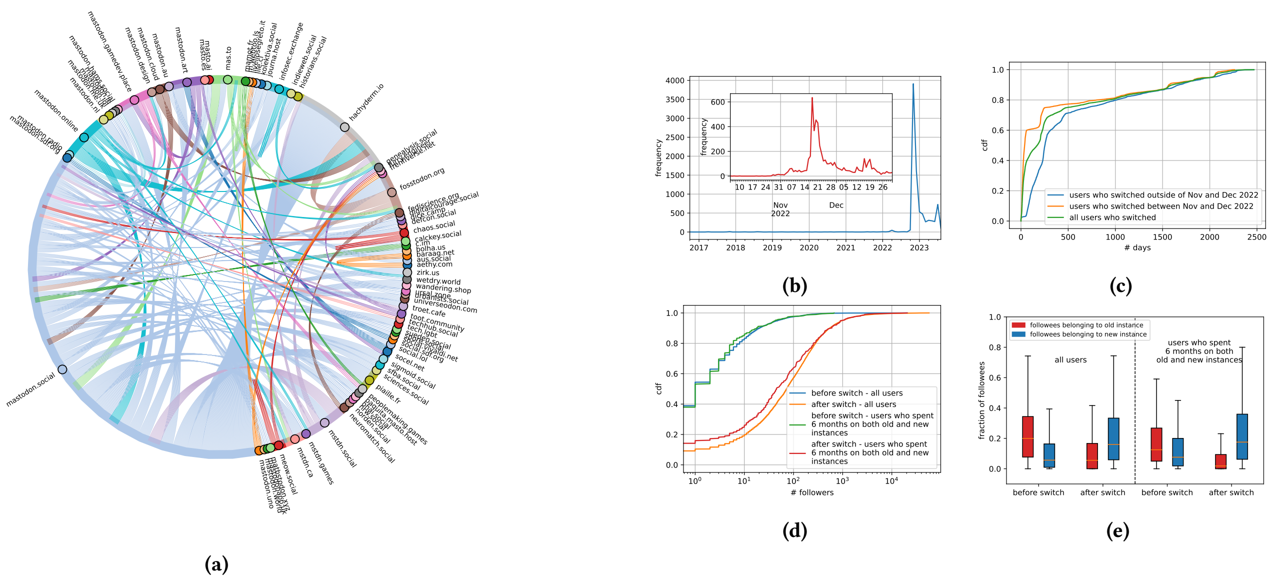


Figure 1: (a) Chord plot of top 100 most frequent old-to-new instance pairs. (b) Frequency of instance switching over time. (c) CDF of time (in number of days) it takes for switching users to switch instances. (d) CDF of the number of followers before and after the switch. (e) Distribution of fraction of followees (*i*) belonging to old instance and (*ii*) belonging to new instance before and after the switch.

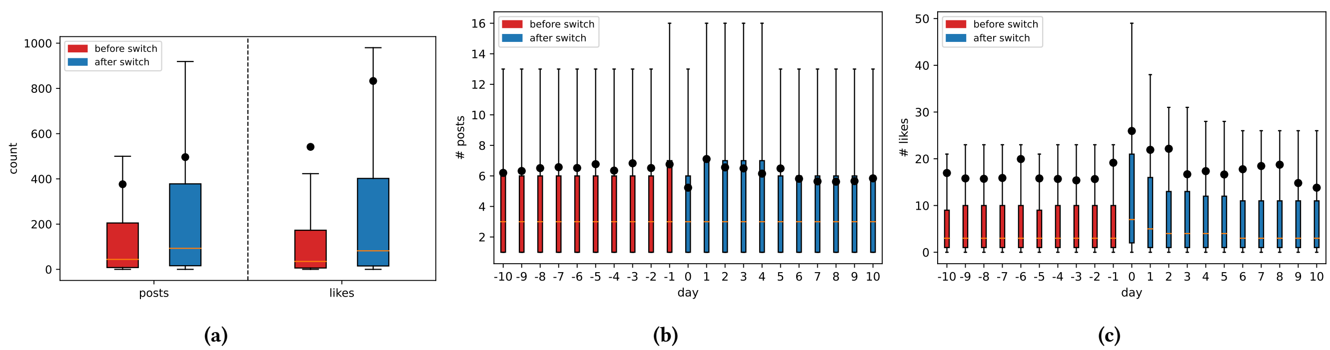


Figure 2: Distribution of the number of posts posted by each switching user and the number of likes it received before and after the switch.

8,745 users before (blue line) and after (orange line) the switch, we notice a large increase in the follower count after the switch. For example, 80% of all users just had ≤ 8 followers before the switch, whereas the same proportion of users had ≤ 280 followers after the switch. The median number of followers (for all switching users) before and after the switch is 1 and 72, respectively.

However, one limitation of the above analysis is that users may spend more time on their new instance, thereby skewing the results. For example, the median time users spend on the old instance is 171 days whereas that on the new instance is 258 days. To overcome this, we create a subset of switching users that spend approximately the same amount of time on their old and new instance. Specifically, we extract users who spend 6 months ± 1 month on both the old and new instance. This covers 4.5% of all switching users. Figure 1d presents equivalent results for this subset. Even when controlling for the duration, we observe the same increasing trend for the subset of switching users who spend the same time on both the old and new instance. The median number of followers for this subset before and after the switch is 1 and 45, respectively. This indicates that gaining followers could be the reason driving users to switch instances.

Interestingly, we also find that 361 out of all 8,745 switching users lose followers after the switch. There could be multiple possible explanations for this, most notably that the switching process is still ongoing (it could take days to complete), or that their old instance went offline during the switching process.

Followees. Next, we strive to understand if users switch instances because they find more relevant (like-minded) users on the new instance. For this, we use the followee count as a metric — we argue that following a greater number of users on an instance is a good proxy for measuring how much interest a user has in that instance's content. We therefore analyze the followee graph of each switching user and compare the fraction of each user's followees that belong to the old instance *vs.* belong to the new instance (before and after switching).

Figure 1e displays these distributions. Confirming our hypothesis, we observe a contrasting trend when comparing distributions for all switching users. Specifically, the fraction of followees belonging to the old instance decreases, while the fraction of followees belonging to the new instance increases after the switch. Before switching, the median percentage of each user's followees belonging to the old instance is 24.1%. However, this decreases to 5.5% after switching. In contrast, the median percentage of each user's followees belonging to the new instance before switching is 5.6%, which increases to 15.9% after the switch. Importantly, this trend is also present for the subset of switching users who spend the same amount of time on both the old and new instance, suggesting that finding relevant followees may influence a user's decision to switch instances.

5.2 Impact of Switching

We conjecture that switching instances impacts users' social media engagement *e.g.* users may post more frequently or post about different topics after switching. To investigate this, we inspect the posts of switching users.

Posts and Likes. We first analyze the engagement of switching users by comparing the number of posts they posted and the number of likes they receive before and after the switch. Figure 2a plots the distribution of the number of posts and likes before and after the switch.

We see an overall increase in both engagement metrics after switching. Before switching, users posted a median of 44 posts and receive a median of 35 likes on their posts, whereas after switching, the median number of posts increased to 93 and the likes received jump to 82. However, due to the disproportionate lifespans of accounts before and after the switch (as discussed in §5.1), these overall engagement metrics may be biased. Thus, to provide a more balanced comparison, we plot the distribution of the number of posts posted by each user and the likes they received 10 days before and after the switch in Figure 2b and Figure 2c, respectively.

We observe that the posting frequency remains constant with a median of 3 posts per day per user; yet, interestingly, the likes that users receive on their posts increases sharply as soon as they switch. For example, the median number of likes users receive on the last day before switching is 3, which jumps to 7 on the day of switching. We note that the number of likes remains high in the first few days of switching before going down to the pre-switch level. This suggests that users experience a temporary surge in engagement immediately after switching instances, possibly due to a change in content strategy. Thus, we proceed to analyze the post content in the following section.

Content Analysis. To understand the topical aspects of the posts posted by switching users, we briefly analyze the hashtags used in their posts. We report the top 50 hashtags (by frequency) used by switching users 10 days before and after switching in Appendix A. While politics, particularly the Russia-Ukraine war [23, 24], remains a popular topic among users both before and after the switch, we observe a diversification of topics with hashtags like #history, #movies, #writing, and #books becoming more prominent after the switch. Furthermore, the Jaccard similarity coefficient between the complete set of hashtags used by switching users 10 days before and after switching is just 0.197. This suggests that users tend to focus on community-specific topics after switching instances. Unsurprisingly, 'introduction' is the most widely used hashtag after switching, indicating users' efforts to introduce themselves to their new community.

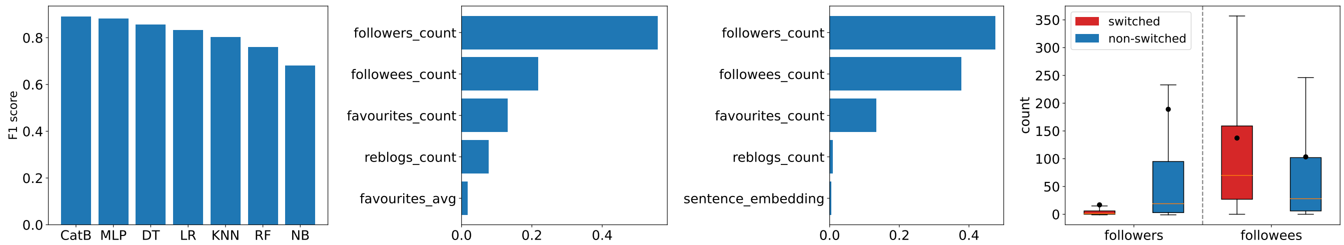
6 Switching Prediction

Finally, we aim to quantify factors predicting whether a Mastodon user will switch instances (RQ3). We build classifiers to distinguish between switched and non-switched users (see §3) and analyze feature importance to better understand the reasons behind instance switching. We argue that this is also a useful tool for administrators who might wish to identify users who are at risk of leaving their instance. Note that for switched users, only pre-switching data is used in the prediction.

Methodology. We extract 24 features for each user, including user statistics (*e.g.* number of followers and posts) and text information (*e.g.* average length of posts, number of URLs, number of hate words, sentence embedding). All features and their descriptions are in Appendix B (Table 1).

Next, we train several machine learning models with the sklearn library [12] and use GridSearchCV to perform 5-fold cross-validation to hyper-parameter tuning. The task is to discriminate between switched vs. non-switched users using a users' Mastodon timeline. Note, this only includes the users' Mastodon timeline, *prior* to switching. The models we train are: (*i*) Logistic Regression (LR) [16], (*ii*) Decision tree (DT) [17], (*iii*) Random Forest (RF) [18], (*iv*) K-nearest Neighbors (KNN) [19], (*v*) Multi-layer Perceptron (MLP) [20], (*vi*) Naive Bayes (NB) [21], and (*vii*) CatBoost (CatB) [13]. Figure 3a shows the F1 scores of the seven models on the task of predicting whether users switched or not. We predict the likelihood of user switching with F1 scores of up to 0.891 for CatBoost, our best-performing model.

Results. We use our classifiers to understand the features that correlate with a users' decisions to switch. To do this, we extract the feature importance for the three best performing models: CatB and DT (for MLP, there is no built-in feature importance function in sklearn [12]). We show the results in Figure 3b and 3c. The most important feature is *followers_count*, and *followees_count* also plays an important role in predicting user switching. To explore this further, Figure 3d plots the distribution of followers and followees for switched and non-switched users. Non-switched users have, on average, 172.1 more followers than switched users, but followees are on average 33.9 fewer than switched users. This is perhaps intuitive: users with a small number of followers may switch instances to expand their social network. Indeed, our result shows that these users gain more followers after switching (see Section 5.1). Conversely, switched users tend to have more followees, suggesting a desire to expand their social network and interact with users from different instances, and thus have a potential tendency to switch.



(a) (b) (c) (d)

Figure 3: (a) F1 scores for switching prediction. (b) CatBoost top 5 normalized feature importance. (c) Decision tree top 5 normalized feature importance. (d) Boxplot of followers and followees for switched and non-switched users.

7 Related Work

There have been a small number of studies on user migration within online social networks. Gerhart *et al.* [6] analyzed user migration from traditional social networks to anonymous social networks, arguing that social norms are the driving force behind migration. Otala *et al.* [10] studied the migration of Twitter users to Parler, noting that although Parler is not widely used, it has a significant impact on political polarization. Fiesler *et al.* [5] measured the migration activity from Fandom, tracking migrating users and their reasons for migration. Their study showed that policy and value-based aspects are key determinants of migration. A common aspect of these studies is that the users examined were constrained by the inability to carry their social network (*i.e.* followers and followees) to the new platform.

A small set of prior efforts have studied migration to decentralized social media. He *et al.* [8] and Jeong *et al.* [9] examined the migration patterns of Twitter users to Mastodon. However, in these cases, migrating users did not carry their social network with them. To the best of our knowledge, this is the first study that analyzes how interoperability allows users to switch *across* servers within a decentralized social network like Mastodon, while retaining their social network.

8 Conclusion

In this paper, we have explored user switching between Mastodon instances and made several key observations. We find that switching is more common than expected, with 1.69% of users switching, often from general-purpose to topic-specific instances. This decision is influenced by a user's social network, leading to increases in followers and followees post-switch. However, the impact on engagement is short-lived, with only a temporary rise in the number of likes after the switch. We have also demonstrated that predicting user switching is feasible, with follower and followee counts being the most predictive features.

We argue that Mastodon serves as an early exemplar of a social platform that champions the interoperability and portability of accounts. This is just the first step in a wider research initiative. We posit that more social networks (largely due to regulation) will soon begin to roll out similar features. For instance, Bluesky already offers the ability to port Twitter social data to it [2]. Consequently, in the future, we would like to extend our analysis to other social media platforms. We are also keen to explore methods by which this migration can be made more straightforward for users.

Acknowledgments

This work is supported in part by the EPSRC grants AP4L (EP/W032473/1), DSNmod (REPHRAIN EP/V011189/1), Fediobservatory, Guangzhou Science and Technology Bureau (2024A03J0684), Guangzhou Municipal Key Laboratory on Future Networked Systems (024A03J0623), and Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (2023B1212010007).

References

[1] ActivityPub. 2018. https://www.w3.org/TR/activitypub/.
[2] Leonhard Balduf, Saidu Sokoto, Onur Ascigil, Gareth Tyson, Björn Scheuermann, Maciej Korczyński, Ignacio Castro, and Michał Król. 2024. Looking AT the Blue Skies of Bluesky. *arXiv preprint arXiv:2408.12449* (2024).
[3] Lucio La Cava, Luca Maria Aiello, and Andrea Tagarelli. 2023. Drivers of social influence in the Twitter migration to Mastodon. *Scientific Reports* 13, 1 (2023), 21626.
[4] European Commission. 2022. https://digital-markets-act.ec.europa.eu/index_en.
[5] Casey Fiesler and Brianna Dym. 2020. Moving across lands: Online platform migration in fandom communities. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (2020), 1–25.
[6] Natalie Gerhart and Mehrdad Koohikamali. 2019. Social network migration and anonymity expectations: What anonymous social network apps offer. *Computers in Human Behavior* 95 (2019), 101–113.
[7] GoToSocial. 2024. https://gotosocial.org/.
[8] Jiahui He, Haris Bin Zia, Ignacio Castro, Aravindh Raman, Nishanth Sastry, and Gareth Tyson. 2023. Flocking to mastodon: Tracking the great twitter migration. In *Proceedings of the 2023 ACM on Internet Measurement Conference*. 111–123.
[9] Ujun Jeong, Paras Sheth, Anique Tahir, Faisal Alatawi, H Russell Bernard, and Huan Liu. 2023. Exploring platform migration patterns between twitter and mastodon: A user behavior study. *arXiv preprint arXiv:2305.09196* (2023).
[10] Jacqueline M. Otala, Gillian Kurtic, Isabella Grasso, Yu Liu, Jeanna Matthews, and Golshan Madraki. 2021. Political polarization and platform migration: a study of Parler and Twitter usage by United States of America Congress Members. In *Companion Proceedings of the Web Conference 2021*. 224–231.
[11] Misskey. 2014. https://misskey-hub.net.
[12] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.
[13] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. 2018. CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems* 31 (2018).
[14] Aravindh Raman, Sagar Joglekar, Emiliano De Cristofaro, Nishanth Sastry, and Gareth Tyson. 2019. Challenges in the decentralised web: The mastodon case. In *Proceedings of the internet measurement conference*. 217–229.
[15] Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084* (2019).
[16] scikit-learn. 2024. https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html.

[17] scikit-learn. 2024. https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html.
[18] scikit-learn. 2024. https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html.
[19] scikit-learn. 2024. https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html.
[20] scikit-learn. 2024. https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html.
[21] scikit-learn. 2024. https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.GaussianNB.html.
[22] Leanne Townsend and Claire Wallace. 2017. The ethics of using social media data in research: A new framework. (2017), 189–207.
[23] Yiming Zhu, Ehsan-Ul Haq, Gareth Tyson, Lik-Hang Lee, Yuyang Wang, and Pan Hui. 2023. A Study of Partisan News Sharing in the Russian invasion of Ukraine. *AAAI ICWSM* (2023).
[24] Haris Bin Zia, Ehsan Ul Haq, Ignacio Castro, Pan Hui, and Gareth Tyson. 2023. An Analysis of Twitter Discourse on the War Between Russia and Ukraine. *arXiv preprint arXiv:2306.11390* (2023).

A Hashtags Before *vs.* After Switching

Before Switching. ukraine (2547), aiart (2540), digitalart (2486), stablediffusion2 (2463), ExperimentalExtra (1210), russia (909), news (777), christmas (755), holidays (721), mastodon (697), animals (584), Algemeen (450), twitter (385), skynews (360), airlinenews (360), art (344), photography (312), santa (305), fediverse (301), cats (294), Ukrainianconflict (294), food (288), introduction (283), reuters (266), sailing (261), sail (259), snow (258), sailnews (255), music (251), WARINUKRAINE (233), photo (232), twittermigration (224), politics (217), ai (207), tvandfilm (202), anime (201), videogames (183), elonmusk (183), caturday (174), cute (170), catsofmastodon (170), nasa (156), mastoart (152), gaming (152), ElongatedMuskrat (149), tech (147), artemis (141), space (135), auspol (132)

After Switching. introduction (1351), Ukraine (1341), aiart (984), digitalArt (919), mastodon (914), stablediffusion2 (909), Russia (756), news (551), art (525), photography (519), experimentalextra (453), music (413), auspol (411), twitter (408), airline (378), skynews (378), airlinenews (378), fediverse (363), worldcup (291), warinukraine (290), cats (279), twittermigration (268), reuters (257), anime (255), MastoArt (250), animals (247), gaming (246), furryart (241), books (231), catsofmastodon (229), videogames (217), history (205), caturday (205), photo (191), mastoart (190), startrek (188), gamedev (182), ai (180), histodons (169), linux (166), retrogaming (160), furry (156), Ukrainianconflict (154), movies (149), writing (147), ttrpg (147), illustration (146), food (144), abdl (142)

B Features for Switching Prediction

Table 1 presents the full list of features used for the switching prediction task.

Feature	Description	Mean		Std Dev		Median	
		Switched	Non-Switched	Switched	Non-Switched	Switched	Non-Switched
followers_count	Number of followers of user	16.93	189.02	1.0	19.0	255.46	1414.72
followees_count	Number of users that user follow	137.22	103.33	70.0	28.0	233.37	266.91
posts_count	Number of posts of user	423.43	434.79	60.0	25.0	1929.5	3548.34
post_length_avg	Average posts length of user	160.57	142.05	147.51	125.02	88.76	106.01
hatewords_count	Number of hate words from hatebase.org	14.42	14.99	1.0	0.0	86.79	119.24
hatewords_avg	Average number of hate words from hatebase.org	0.05	0.05	0.01	0.0	0.1	0.11
urls_count	Number of URLs in posts of user	72.06	83.89	4.0	1.0	568.84	1261.03
urls_avg	Average number of URLs in posts of user	0.21	0.21	0.1	0.05	0.3	0.35
hashtags_count	Number of hashtags in posts of user	107.67	175.19	10.0	3.0	654.67	3292.97
hashtags_avg	Average number of hashtags in posts of user	0.72	0.73	0.25	0.13	1.41	1.63
mentions_count	Number of mentions in posts of user	127.73	109.94	13.0	4.0	566.28	778.25
mentions_avg	Average number of mentions in posts of user	0.52	0.41	0.5	0.33	0.45	0.46
emojis_count	Number of emojis in posts of user	5.09	5.8	0.0	0.0	95.24	109.37
emojis_avg	Average number of emojis in posts of user	0.02	0.02	0.0	0.0	0.11	0.15
replies_count	Number of posts replied by user	120.23	92.23	12.0	3.0	919.81	630.9
replies_avg	Average number of posts replied by user	0.45	0.36	0.36	0.25	0.81	0.65
reblogs_count	Number of reblogs by user	191.39	198.15	7.0	2.0	2210.64	1743.2
reblogs_avg	Average number of reblogs by user	0.92	0.9	0.17	0.1	6.89	5.41
favourites_count	Number of favourites received by users	541.54	539.74	35.0	11.0	5485.66	4081.87
favourites_avg	Average number of favourites received by users	2.39	2.38	1.0	0.79	9.67	10.15
sensitive_count	Numebr of sensitive posts of user	23.36	14.84	0.0	0.0	335.64	197.73
sensitive_avg	Average numebr of sensitive posts of user	0.05	0.05	0.0	0.0	0.12	0.15
language_index	The most frequently used languages by users	-	-	-	-	-	-
sentence_embedding	Aggregation of all posts embedding by the user (with the paraphrase-multilingual-mpnet-base-v2 model in Sentence-Transformer [15])	-	-	-	-	-	-

Table 1: Summary of all extracted features used for model training.