Chapter 5

Study of username changing behavior

Questions on how, why and who are these user accounts that frequently change their username on Twitter are, so far, unanswered. We believe that answering how users create usernames over time can aid in finding a user’s account on other social networks. Username creation methods to create usernames over time can replicate and re-occur while creating usernames on her other accounts. Literature and our prior work devise profile linking methods that can link user accounts assuming usernames are created in the similar fashion within and across networks over time \[56,112\]. Answering who are these users and why do they change usernames can help us understand if the username changing behavior is a characteristic of a specific set of users. Finding reasons for username change can indicate if the intentions are benign and valid or fraudulent.

We make the first attempt to answer these questions and characterize username changing behavior on Twitter. We carefully create a dataset of 10K users, randomly sampled from 8.7 million users, and track them for a duration of 14 months every fifteen minutes. Our work with recorded past usernames of the users can help Twitter to effectively redirect user search queries rather than either serving with a dead link or different user. With an understanding of patterns and reasons of username change, Twitter can also develop tailored username suggestion algorithms for its benign users during the sign-up process and later.

5.1 Data Collection

We use the same data collected for linking user identities across OSNs, described in detail in Section 4.4. In summary, we query 10K users via Twitter API every 15 minutes. We term the faster scan of 10K users as Fifteen-minute scan. Fifteen-minute scan starts on November 22, 2013; we bookmark the scan till January 22, 2015 and use 14 months scan for our analysis.\footnote{We continue to scan 10K users after link-score 22 '15 and record any username change.}
With fifteen-minute scan we have an advantage; we could record the exact time when user changed her username, with an error limit of 15 minutes. Further, we observe that while our regular fortnight scan took only one snapshot, fifteen-minute scan took 794 snapshots of 10K users, during November 22, 2013 - November 26, 2013. Regular fortnight scan missed 712 username changes triggered by 607 users, well captured by fifteen-minute scan. Therefore, fifteen-minute scan is successful in capturing most username changes made by the tracked users. All analysis here examines 10K users that switched usernames within 14 months on Twitter.

Table 5.1: Fortnight scan tracked over 8.7 million users every fortnight and missed many username changes due to long scan stretches of four to five days. We then initiated a fifteen-minute scan with a random sample of 10,000 users out of 10% of 8.7M users who changed their usernames at least once, and tracked them every fifteen minutes for more than a year.

<table>
<thead>
<tr>
<th>Name of scan</th>
<th>Period of scan</th>
<th># users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fortnight scan</td>
<td>October 16, 2013 - November 26, 2013</td>
<td>8,767,576</td>
</tr>
<tr>
<td>Fifteen-minute scan</td>
<td>November 22, 2013 - November 22, 2015</td>
<td>10,000</td>
</tr>
</tbody>
</table>

5.1.1 Representativeness

As mentioned earlier, it is necessary that users in 10K dataset span across diverse locations and different registration years on Twitter, to avoid any bias towards a special section of users. We examine geographical locations of 10K users to understand if they span across diverse locations. We use geo-tagged tweets by the users to record their location. We map 1,849 unique latitude, longitude pairs from where 926 users (9% of 10K) have posted their tweets (see Figure 5.1(a)). We observe that users in our dataset tweet from different locations around the world and not biased to only a few locations. Therefore, our analysis and results can be generalized to Twitter population from various global locations, who change their usernames over time. We further examine if these 10K users mainly contain newly registered users and hence they prefer to change their username to adjust to their requirements on Twitter. We note the year of account creation for these users and show the distribution in Figure 5.1(b). We observe that users in our dataset registered themselves on Twitter in years ranging from 2007 to 2013. Therefore, our dataset captures users from different levels of familiarity to Twitter.

5.2 Characterization

Before analyzing the characteristics of the entities involved in username changing process – usernames and users, we estimate the frequency of the behavior. Out of 10K users, 4,198 users changed
their usernames at least once in 14 months, constituting 14,880 username changes. About 20% users changed five times or more triggering around 12,648 (85% of all) username changes (see Figure 5.2(a)). One user changed her username 113 times in 14 months which on manual inspection, turned out to be an inorganic user [20] with half completed tweets, tweets with same text, and frequent posts in short duration. We also examine the number of days after which users trigger the change (see Figure 5.2(b)). Around 20% of username changes were triggered within a day of the previous username change. We, therefore, observe a Pareto distribution with 20% users frequently changing usernames in short intervals and 80% users rarely changing after long durations (see inset figure in Figure 5.2(a)).

5.2.1 Usernames

An action of username change involves dumping an old username and creating a new one. Often, users favorite a username and repeatedly use that username. In our dataset, we find that around 35% of users reuse an old username later, while 65% never do so. For the 65% users, it is important to understand how users create their usernames over time. This information can be helpful in predicting their username creation patterns on other social networks as well, thus helping connecting multiple profiles of a single user. In order to understand username creation patterns, we filtered out reused usernames and considered unique usernames used by the users over time. We first investigate how the usernames differ from each other. To measure the similarity between two consecutive usernames used by a user, we use Longest Common Subsequence (LCS) similarity, a well-defined metric in literature.

LCS similarity estimates the sequence of characters that appear together without penalizing for insertions made. Figure 5.2(c) shows the cumulative distribution of username changes v/s LCS
matching length between the usernames (old and new) associated with the change. LCS matching length is normalized by length longer of the two, as suggested in literature [112]. For approximately 82% of username changes, new username is un-related to the old username (length $\leq 0.5$), while for around less than 10% changes have new usernames highly similar and derived from the old ones (length $\geq 0.8$). The observation indicates that majority of users select dissimilar usernames over the time within one social network, which is a complementary observation to literature which suggests that users create similar usernames across other social network sites [83,112]. Note that such a user behavior may repeat as well across social networks and hence could challenge profile linking methods that use traditional string matching algorithms to match usernames in order to find connection between two user profiles.

We now examine the kind of change that users make to their usernames i.e. do they add special characters, numbers or alphabets, at what positions they make changes, and how they vary length changes. For the analysis, we selected usernames with LCS length $\geq 0.5$. Figure 5.2 shows distribution of users and username changes among various username alteration methods. Changes to usernames are preferred either at the beginning, at the center or in the end of the username. Possibly, users add suffixes and prefixes to create a new username out of the old one. Users prefer to use characters rather than numbers or special characters every time they create a new username (100%
changes include only characters). We conclude that users exhibit certain preferences of username creation and alteration, which can be captured to modify the search with the outdated username and for profile linking across networks.

![Graphs](image)

Figure 5.3: Comparison between 4,198 users who changed their usernames with two random samples of 4,198 users who never changed their usernames. Users who change usernames demonstrate statistically significant superiority in terms of popularity and activity.

### 5.2.2 Users

We now explore the characteristics of users who opt to change their usernames. Does their popularity or activity or familiarity with network govern the frequency of username change? We answer these questions now.
Normal v/s Changing Users

On examining if users who undergo username changes are different from others who don’t, we compare the activity and popularity of both genre of users. From 90% proportion of users from 8.7 million users that does not engage in this behavior, we extract an equal number of users as those who change usernames. A comparison of in-degree, out-degree and activity of 4,198 users with characteristics of two random samples of size 4,198 users are shown in Figure 5.3.

Both random samples do not differ in their properties but differ from the users who change their usernames over time. Users who change have higher popularity and activity as compared to users who do not change their username. Statistical significance of the difference is tested using Kolmogorov - Smirnov (KS) test. We make two pairwise comparisons of distributions and report test result for one of them. For in-degree, the value of KS $D$ statistic is 0.3943 with $p$-value $< 0.0001$, for out-degree, the $D$ value is 0.2654 with $p$-value $< 0.0001$ and for activity, $D$ value is 0.4143 with $p$-value $< 0.0001$). We receive similar values for the pairwise comparison between second normal set and changing users. Higher values of $D$ with a lower $p$-value for the comparisons signify that normal and changing users differ from each other in Twitter’s activity, popularity, and following behavior.

Popularity v/s Frequency of Change

On Twitter, users tweet, reply or converse with their username. Changing usernames by a popular user may lead to confusion among her followers or may lead to loss of tweets in case someone else picks the username. For instance, the Indian Prime Minister Office’s Twitter handle ‘@PMOIndia’ was acquired by a teenager for 30 minutes during the transfer of social accounts from the earlier government [89]. In such a scenario, we speculate that users with higher number of followers avoid any username changes. We measure popularity of 4,198 users using followers (in-degree) and plot it against frequency of username change (see Figure 5.4(a)). To find correlation between the two, we remove users with too many followers (> 1 million) or too less (< 1). We observe that username change frequency is weakly yet positively correlated to the in-degree of the user (Pearson correlation: 0.1153, $p$-value $< 0.00001$, $\alpha$: 0.05). A significant positive correlation implies that higher the popularity, higher is the frequency of change, however, weak correlation does not guarantee the same.

Activity v/s Frequency of Change

An active user on Twitter, who engages herself in conversations and group chats, may change her username less frequently to avoid confusion during tagging / replying in a tweet. We conjecture

---

2In order to justify the generalizability of observations, we pick two random samples.
that active users change their usernames less frequently. We analyze 4,198 users and measure their activity with the number of created tweets. Figure 5.4(b) shows the frequency of username change with the user’s activity. To find correlation between the two, we removed users with too many tweets (> 100K) or too few (< 1). We observe a weak and positive correlation between the two (Pearson correlation: 0.1045, p-value < 0.0001, α: 0.05). A positive yet weak correlation implies that users with high activity are inclined towards frequent username changes, however, activity does not guarantee frequency of change.

![Figure 5.4](a)

![Figure 5.4](b)

Figure 5.4: Frequency of username change v/s user popularity and activity. Weak correlations imply that popularity and activity has a little impact on the choice of changing username.

**Familiarity v/s Frequency of Change**

Intuitively, users who registered long time ago are familiar with Twitter and must have chosen stable and beneficial username for themselves than users who have registered recently and are still in exploratory stage. We examine if old user accounts engage themselves in username changing behavior, or only new users change their usernames multiple times. Figure 5.5 shows the frequency of username change with the age of the account for 4,198 users. We observe negative and very weak correlation between the age of the Twitter account and the frequency with which the account changes username (Pearson correlation: -0.0942, p-value < 0.0001, α: 0.05). Negative and weak correlation implies that both older and newer accounts engage in this behavior.

**5.3 Plausible reasons**

So far it in unclear on reasons that encourage users to change their usernames. Note that, users put in efforts to create a suitable username to converse with others on the network. A sudden change
Figure 5.5: Frequency of username change v/s account age. Very weak implies that both old and new user accounts actively change their username.

to the username may direct users to a broken link or to a different user altogether who now owns the dumped username. We, therefore, explore a set of reasons for this change based on literature and observations using data analysis and talking informally to tracked users via tweets. On manual inspection, we classify listed reasons as benign and malicious.

5.3.1 Benign Reasons

Reasons for change that are motivated by user requirements and natural behavior are marked as benign. These are:

Space Gain

On Twitter, a user can converse with another user by tagging her ‘@<username>’ in 140-character tweet. Since the tweet length is limited and maximum character limit for username is 15 characters, long usernames imply short message. We speculate that users with long old usernames may change to short new usernames to allow other users (followers) to post more content than before and benefit from space gain. This reason is motivated by the introduction of shortened URLs and RT symbol in Twitter to save space in a tweet [61]. We calculate the length difference between new and old username of users and separately represent users with old usernames less than and greater than the median length (≥ 11). We observe that 75.19% of long usernames moved to short or same length new usernames, while 60.87% short usernames picked long new usernames (see Figure 5.6). In other words, most users with old usernames of length < 11 tend to add characters in their new usernames, while most users with old usernames of length ≥ 11 prefer to remove characters for their
new usernames. With this observation, we infer that creating shorter usernames is an incentive for users to change usernames.

![Figure 5.6: Length difference between new and old username v/s length of old username. Users with long usernames pick shorter new usernames (higher negative space gain), while users with short usernames pick longer new usernames.](image)

**Maintain Multiple Accounts**

On Twitter, a user is allowed to create multiple accounts, each with a different email address. On observation, we see that few changed username in order to exchange usernames among their multiple accounts (see Figure 5.7). We think that by tracing shared username’s owners over time may help link multiple accounts of a single user within Twitter.

![Figure 5.7: Shared username is exchanged among user’s multiple accounts on Twitter.](image)

---

3https://support.twitter.com/articles/20169956
**Change Username Identifiability**

Few users in our dataset changed usernames to reverse the identifiability of the usernames i.e. either to make them personal or anonymous. For instance, a user named ‘loried ligarreto’ changed her username from ‘loriedligarreto’ to ‘sienteteotravez’ (translated: ‘feel again’) implying that user possibly intended to make her username anonymous. In other instances, we observe users who previously picked less identifiable usernames, made them personal later. For example, a user named ‘rodrigo’ changed her username from ‘unosojosverdes’ (translated: ‘green eyes’) to ‘rodrigothomas_’, thereby implicating that user probably wished to associate her real identity to her username.

**Adjust to Events**

Another user told us in a tweet that she represents Sahara India FanClub. She has supported Sahara’s Pune Warriors team in IPL event with username ‘pwifanclub’ and then Sahara F1 team with username ‘ForceIndia@!’ and therefore has changed her username (see Figure 5.8(a)).

**No specific reason**

Few users responded that they changed their usernames without any specific reasons, that they got bored of the earlier one (see Figure 5.8(b)).

**5.3.2 Malicious Reasons**

Username changes that trick and misguide users violates Twitter usage policy and result from unnatural user behavior are marked as malicious. These are:

**Obscured Username Promotion**

Owing to a limited number of users in fifteen-minute scan, we use fortnight scans of 8.7M users for this analysis. To our surprise, we find that a few user profiles collaboratively picked the same
Table 5.2: Rotational use of a shared username among users belonging to a group on Twitter. Numbers in the brackets represent follower count as recorded in the scan. We observe that users of a group (partner accounts) collaboratively share and promote the username via tweets / description to evenly distribute followers among themselves.

<table>
<thead>
<tr>
<th>ID</th>
<th>Scan - I</th>
<th>Scan - II</th>
<th>Scan - III</th>
<th>Group</th>
<th>Observation Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>12xx62463x</td>
<td><strong>CollaGe_InFo</strong> (61)</td>
<td>Dictionary_ID</td>
<td>Dictionary_ID</td>
<td>Sajan</td>
<td>2013-04-01</td>
</tr>
<tr>
<td>11xx79686x</td>
<td>DaiLy_GK (292)</td>
<td><strong>CollaGe_InFo</strong> (2,372)</td>
<td>Geo_Account</td>
<td>Sajan</td>
<td>2013-10-02</td>
</tr>
<tr>
<td>95xx18222x</td>
<td>Geonewsapk9</td>
<td>DictioNary_GK (1,279)</td>
<td><strong>CollaGe_InFo</strong></td>
<td>Sajan</td>
<td>2013-10-25</td>
</tr>
<tr>
<td>19xx56472x</td>
<td>-</td>
<td>-</td>
<td><strong>CollaGe_InFo</strong></td>
<td>Sajan</td>
<td>2013-12-04</td>
</tr>
<tr>
<td>60xx2762x</td>
<td>Peshawar_sMs (1,282)</td>
<td>MoBile_TricKes</td>
<td>BBC_PAK_NEWS</td>
<td>Sajan</td>
<td>2013-04-08</td>
</tr>
<tr>
<td>11xx37099x</td>
<td>Vip_Wife (180)</td>
<td>Peshawar_sMs (4,325)</td>
<td>UBL_Cricket</td>
<td>Sajan</td>
<td>2013-10-25</td>
</tr>
<tr>
<td>28xx1645x</td>
<td>NFS002cric</td>
<td>NaKaaM_LiFe (3,880)</td>
<td><strong>Peshawar_sMs</strong> (4,044)</td>
<td>Sajan</td>
<td>2013-11-08</td>
</tr>
<tr>
<td>70xx9502x</td>
<td><strong>FuNNy_SardaR</strong> (1,406)</td>
<td>MaST_DuLHaN (4,175)</td>
<td>KaiNaT_LipS</td>
<td>Khan</td>
<td>2013-04-01</td>
</tr>
<tr>
<td>99xx9356x</td>
<td>SaIrA_JoX (1,009)</td>
<td><strong>FuNNy_SardaR</strong> (1,841)</td>
<td>MaST_DuLHaN (1,900)</td>
<td>Khan</td>
<td>2013-10-02</td>
</tr>
<tr>
<td>12xx73970x</td>
<td>-</td>
<td>-</td>
<td><strong>FuNNy_SardaR</strong></td>
<td>Khan</td>
<td>2013-12-04</td>
</tr>
</tbody>
</table>
username at different timestamps. Table 5.2 shows two such groups and the rotation of a username among the profiles, as observed in four scans. Usernames ‘Collage_InFo’, ‘Peshawar_sMs’ and ‘FuNNy_SardaR’ were used by different user IDs at different times. All these users claimed to belong to a group, either in their name or bio attribute. We term the username which is shared by multiple accounts as *shared username* and profiles who picked the shared username at different times as *partner accounts*. We observe 70 other shared usernames in our fortnight scans. We inspect the intentions for such a behavior in following ways.

We analyze tweets and description of the partner accounts mentioned in Table 5.2. We calculate the number of ‘@’ tags mentioned in their tweets and description. It was surprising to see that irrespective of the group, the partner accounts promoted a shared username by posting “Follow @<username>” in their tweets (or in description) multiple times (see Figure 5.9). Altogether for the two groups under observation, ten accounts promoted 30 other usernames. Seventeen percent (5 out of 30) usernames are promoted by more than one user. We think that by asking other Twitter users to follow a shared username and then keep exchanging the username with each other, the intention is to obscure the real identity of the user behind the free flowing shared username and distribute the followers evenly across the partner accounts.

![Example post on Twitter where one partner account promotes another in her tweets.](image)

We explain the username promotion methodology as: an account holds a shared username \( u_s \), while other partner accounts promote the username by asking users to follow the account with username \( u_s \). The account gains followers and decides to let her partner accounts gain further. She then releases her username to be picked by any of her partner accounts, and picks another (shared) username. She starts the promotion of the username \( u_s \), along with other partner accounts. Her partner account, which picked \( u_s \) then gain followers. For the accounts mentioned in Table 5.2, we observe that a username is picked by the partner accounts with fewer followers. A similar modus operandi was observed when Recorded Future[^4] analyzed Twitter accounts of a terrorist organization, Islamic State (IS). A single username was promoted by multiple ISIS-related accounts or followers either via bio or tweets, thereby tricking and gaining followers[^42]. We suspect that the accounts listed in Table 5.2 engage in similar malicious activities.

[^4]: [https://www.recordedfuture.com/](https://www.recordedfuture.com/)
Username Squatting

On Twitter, there are four pools to which a username can belong to – free-username pool, taken-username pool, suspended / deactivated-username pool and squatted-username pool. A username belongs to a free-username pool if no one else uses it. The username moves to taken-username pool if taken by an account. If the account gets deactivated or is suspended by Twitter, the username is blocked forever (for now), which thereby moves the username in suspended / deactivated-username pool. If an inactive user account keeps the username, in order to block or preserve that username, and not to allow others to use it, the username belongs to squatted-username pool.

Username squatting is against Twitter Rules. Squatted usernames on OSNs have been investigated as a challenge in literature by researchers 29,88 to investigate cases of trademark infringement. We are curious to find if users change usernames in order to squat interesting usernames or usernames that represent an organization or an entity. Method to squatting here is to create profiles that either show no activity (i.e. no tweets) or have zero followers. For our fifteen-minute scan, we observe that for around 12% of 4,198 users, at least one of their vacated usernames are blocked by inactive Twitter profiles, either created by themselves or others. Without the access to emails, used to create user accounts, we have a little information to find if users themselves created the accounts to squat the usernames or others were opportunistic to find a free username and block the username with an inactive account. We think that future research can add these observations as features to find malign / phoney users on platforms like Twitter.

5.4 Discussion

This work aims at finding how and why users change their usernames within a social network like Twitter. Based on literature which suggests that when users create new usernames on different sites, they tend to create similar ones, we speculated that when users change their username within a network, they will pick new usernames similar to the old ones 74,83,112. However, our analysis suggests the contrary. Most users created new username un-related to the old username when they changed username within a social network. We think that un-related usernames over time could be credited to the absence of cognitive load to remember a past dumped username 112. When creating username across social networks, a user needs to remember all usernames, but when creating usernames within a network, she needs to remember only the latest one. Therefore, she has the liberty to choose it to be different from others. Un-relatedness between old and new usernames may challenge people on the network to derive new username from a user’s old username and use

5https://support.twitter.com/groups/51-me/topics/205-account-settings/articles/14609-changing-your-username
6https://support.twitter.com/articles/15348-my-account-information-is-already-taken#deactivatedaccount
7https://support.twitter.com/articles/18370-username-squatting-policy
8https://support.twitter.com/articles/18311
Twitter search engine to find her.

**Reasons for change on other platforms:** We also inquired if similar reasons for change exist for other social networks that allow username change any number of times. Wikipedia is a moderated platform which allows changes to usernames. Every time a Wiki member wants to change username, she needs to request a moderator with her old username, wished new username and the reason for change. We collated 16,167 reasons from 15,288 Wiki members listed within six years i.e., from December 20, 2007 to December 20, 2013, publicly available here.\(^9\) Reasons are described as free text, so we used grounded theory and classified the reasons in categories based on Wiki policies of username creation.\(^10\) Figure 5.10 shows the categories and the distribution of reasons within each category. We observe that 22% users request a username change as their old username is not in accordance with Wikipedia’s username policy, 30% users change to gain anonymity and avoid abuse, few for unified identity, adjust to spelling errors and capitalization, rest change for no specific reason. Few examples for username change are mentioned in Table 5.3. Study of username change on two networks, Twitter and Wikipedia, show that few users are concerned about their privacy, while others want to establish their unique identity across platforms they use. Few other reasons are platform specific e.g. username promotion on Twitter to gain followers while username change on Wikipedia to adjust to platform’s policy.

![Figure 5.10: Reasons for username change mentioned by Wiki members as part of their request.](image)

<table>
<thead>
<tr>
<th>Reason Category</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy</td>
<td>&quot;For privacy, since Sstrieu has my initials and part of my full name.&quot;</td>
</tr>
<tr>
<td>Privacy and Abuse</td>
<td>&quot;For privacy. I've attracted the attention of online bullies lately and they've been trying to harass me anywhere they can find the name FazzMunkle.&quot;</td>
</tr>
<tr>
<td>Link All Accounts</td>
<td>&quot;Consistency with other logins across the whole range of places where one can log in, including some publicly accessible online profiles (e.g. Twitter). Having these consistent usernames allows my online identity to be consistent to. Plus, I keep forgetting my Wikipedia login info, having to guess quite a lot every time and having to request it be sent to me. Thanks in advance.&quot;</td>
</tr>
<tr>
<td>Use Real Name</td>
<td>&quot;Changing my account from my nickname to my real name&quot;</td>
</tr>
<tr>
<td>Violates Wiki Policy [Promotional]</td>
<td>&quot;I have received a message from one of the administrators that my username is promotional/advertising for [[Roblox]] and he said either change your username or make a new account, so I am requesting to change my name to this.&quot;</td>
</tr>
<tr>
<td>Violates Wiki Policy [Group Usage]</td>
<td>&quot;Current username represents organisation&quot;</td>
</tr>
<tr>
<td>Violates Wiki Policy [Religious Connotation]</td>
<td>&quot;My current name is apparently too ethnic for some editors, leading to inappropriate talk page speculation about my religion.&quot;</td>
</tr>
<tr>
<td>Violates Wiki Policy [Bot]</td>
<td>&quot;Didn't read username policy, not allowed to have Bot in username.&quot;</td>
</tr>
<tr>
<td>Violates Wiki Policy [ Offensive]</td>
<td>&quot;Was told username may be offensive to some, and therefore a violation of username policy.&quot;</td>
</tr>
<tr>
<td>Random Reason</td>
<td>&quot;My current name was used for something else I was using at the time of a low point in my life and I'd like to move on from it. Every time I look at my user name I remember that day and how bad I felt and I don't want to be reminded of it just from my Wikipedia user name. Please let me change it.&quot;</td>
</tr>
</tbody>
</table>

Table 5.3: Examples of few reasons for username change listed on Wikipedia.