An Army of Me: Sockpuppets in Online Discussion Communities

Srijan Kumar
University of Maryland
srijan@cs.umd.edu

Justin Cheng
Stanford University
jcccf@cs.stanford.edu

Jure Leskovec
Stanford University
jure@cs.stanford.edu

V.S. Subrahmanian
University of Maryland
vs@cs.umd.edu

ABSTRACT

In online discussion communities, users can interact and share information and opinions on a wide variety of topics. However, some users may create multiple identities, or sockpuppets, and engage in undesired behavior by deceiving others or manipulating discussions. In this work, we study sockpuppetry across nine discussion communities, and show that sockpuppets differ from ordinary users in terms of their posting behavior, linguistic traits, as well as social network structure. Sockpuppets tend to start fewer discussions, write shorter posts, use more personal pronouns such as "I", and have more clustered ego-networks. Further, pairs of sockpuppets controlled by the same individual are more likely to interact on the same discussion at the same time than pairs of ordinary users. Our analysis suggests a taxonomy of deceptive behavior in discussion communities. Pairs of sockpuppets can vary in their deceptiveness, i.e., whether they pretend to be different users, or their supportiveness, i.e., if they support arguments of other sockpuppets controlled by the same user. We apply these findings to a series of prediction tasks, notably, to identify whether a pair of accounts belongs to the same underlying user or not. Altogether, this work presents a data-driven view of deception in online discussion communities and paves the way towards the automatic detection of sockpuppets.

Keywords

Malicious users; Antisocial behavior; Multiple account use

1. INTRODUCTION

Discussions are a core mechanism through which people interact with one another and exchange information, ideas, and opinions. They take place on social networks such as Facebook, news aggregation sites such as Reddit, as well as news websites such as CNN.com. Nonetheless, the anonymity afforded by some discussion platforms has led to some users deceiving others using multiple accounts, or sockpuppets [13]. Sockpuppetry is often malicious and deceptive, and has been used to manipulate public opinion [2, 38] and vandalize content (e.g., on Wikipedia [37]).
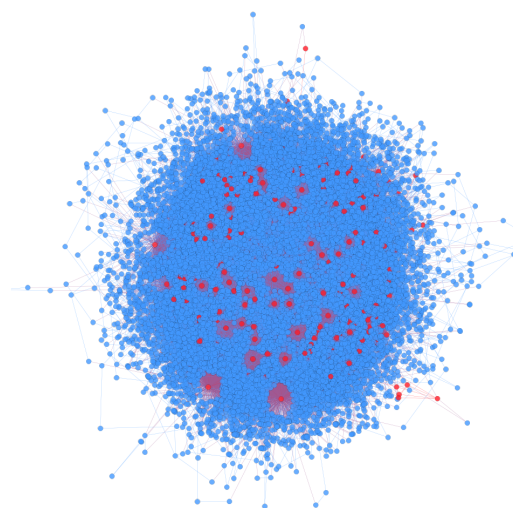
Figure 1: AVClub.com social network. Nodes represent users and edges connect users that reply to each other. Sockpuppets (red nodes) tend to interact with other sockpuppets, and are more central in the network than ordinary users (blue nodes).

Prior work on sockpuppetry and deceptive behavior has tended to focus on individual motivations [9, 17], or on identifying sockpuppets through their linguistic traits [6] (e.g., on Wikipedia [37, 40]). Further, given the difficulty of obtaining ground-truth data about sockpuppets, work has also tended to make assumptions about how sockpuppets behave, for example, assuming that they have similar usernames [26], they are only used to support one another [47], or that they write similar to each other [6]. Further, research has generally not considered how the interactions between sockpuppets controlled by the same individual could be used to accurately and automatically identify sockpuppets. As such, improved methods for identifying sockpuppets, as well as deeper analyses of how sockpuppets interact with one another may allow us to better understand, characterize, and automatically detect sockpuppetry.

The present work: Sockpuppetry in online discussion communities. In this paper, we focus on identifying, characterizing, and predicting sockpuppetry in nine different online discussion communities. We broadly define a *sockpuppet* as a user account that is controlled by an individual (or *puppetmaster*) who controls at least one other user account. By considering less easily manipulated behavioral traces such as IP addresses and user session data, we automatically identified 3,656 sockpuppets comprising 1,623

sockpuppet groups, where a group of sockpuppets is controlled by a single puppetmaster.

Studying these identified sockpuppets, we discover that sockpuppets differ from ordinary users in terms of how they write and interact with other sockpuppets. Sockpuppets have unique linguistic traits, for example, using more singular first-person pronouns (e.g., "I"), corroborating with prior work on deception [2]. They also use fewer negation words, perhaps in an attempt to appear more impartial, as well as fewer standard English parts-of-speech such as verbs and conjunctions. Suggesting that sockpuppets write worse than ordinary users on average, we find that posts are more likely to be downvoted, reported by the community, and deleted by moderators. Sockpuppets also start fewer discussions.

Examining pairs of sockpuppets controlled by the same puppetmaster, we find that sockpuppets are more likely to post at the same time and post in the same discussion than random pairs of ordinary users. As illustrated in Figure 1, by studying the network of user replies, we find that sockpuppets have a higher pagerank and higher local clustering coefficient than ordinary users, suggesting that they are more important in the network and tend to generate more communication between their neighbors. Further, we find that pairs of sockpuppets write more similarly to each other than to ordinary users, suggesting that puppetmasters tend not to have both "good" and "bad" accounts.

While prior work characterizes sockpuppetry as malicious [37, 14, 40], not all the sockpuppets we identified were malicious. In some sockpuppet groups, sockpuppets have display names significantly different from each other, but in other groups, they have more similar display names. Our findings suggest a dichotomy in how deceptive sockpuppets are – some are *pretenders*, that masquerade as separate users, while others are *non-pretenders*, that is sockpuppets that are overtly visible to other members of the community. Pretenders tend to post in the same discussions and are more likely to have their posts downvoted, reported, or deleted compared to non-pretenders. In contrast, non-pretenders tend to post in separate discussions, and write posts that are longer and more readable.

Our analyses also suggest that sockpuppets may differ in their supportiveness of each other. Pairs of sockpuppets controlled by the same puppetmaster differ in whether they agree with each other in a discussion. While sockpuppets in a pair mostly remain neutral towards each other (or are *non-supporters*), 30% of the time, one sockpuppet in a pair is used to support the other (or is a *supporter*), while 10% of the time, one sockpuppet is used to attack the other (or is a *dissenter*). Studying both deceptiveness and supportiveness, we find that supporters also tend to be pretenders, but dissenters are not more likely to be pretenders, suggesting that deceptiveness is only important when sockpuppets are trying to create an illusion of public consensus.

Finally, we show how our previous observations can be used to develop models for automatically identifying sockpuppetry. We demonstrate robust performance in differentiating pairs of sockpuppets from pairs of ordinary users (ROC AUC=0.90), as well as in the more difficult task of predicting whether an individual user account is a sockpuppet (ROC AUC=0.68). We discover that the strongest predictors of sockpuppetry relate to interactions between the two sockpuppet accounts, as well as the interactions between the sockpuppet and the community.

Altogether, our results begin to shed light on how sockpuppetry occurs in practice, and paves the way towards the development and maintenance of healthier online communities.

Community	Genre	# Users	# Sockpuppets	# Sock-groups
CNN	News	846,436	1,191	523
Breitbart	News	196,846	761	352
allkpop	Music	159,671	445	193
MLB	Sports	115,845	339	139
IGN	Games	266,976	314	142
Fox News	News	145,009	214	94
A.V. Club	Entertainment	37,332	199	90
The Hill	Politics	158,378	134	62
NPR	News	65,662	59	28

Table 1: Statistics of the nine online discussion communities.

2. DATA AND DEFINITIONS

We start by laying out the terminology that we use in the remainder of the paper. We then describe the data we used in our analysis and a robust method for automatically identifying sockpuppets.

Sockpuppetry. While in prior work sockpuppets have typically been used to refer to a false online identity that is used for the purposes of deceiving others [26, 47], we adopt a broader definition. We define a *sockpuppet* as a user account controlled by an individual who has at least one other account. In other words, if an individual controls multiple user accounts, each account is a sockpuppet. The individual who controls these sockpuppets is referred to as the *puppetmaster*. We use the term *sockpuppet group/pair* to refer to all the sockpuppets controlled by a single puppetmaster. In each sockpuppet group, one sockpuppet typically has made significantly more comments than the others – we refer to this sockpuppet as the *primary* sockpuppet, and the other sockpuppets as *secondary*. Finally, we use *ordinary user* to refer to any user account that is not a sockpuppet.

We study sockpuppets in the context of online discussion communities. In these communities, people can create accounts to comment on articles. In addition to writing or replying to *posts*, users can also vote on posts or report them for abuse. Moderators can, in turn, delete posts that do not conform to community standards. If a post is not a reply to another post, we call that post a *root post*. We define a *discussion* as all the posts that follow a given news article, and a *sub-discussion* as a root post and any replies to that post.

Data. The data consists of nine different online discussion communities that encompass a variety of topical interests – from news and politics to sports and entertainment (Table 1). Disqus, a commenting platform that hosted these discussions, provided us with a complete trace of user activity across nine communities that consisted of 2,897,847 users, 2,129,355 discussions, and 62,744,175 posts. Each user has a display name which appears next to their posts and an email address which is private. (To respect user privacy, all email addresses were stripped of the domain names and analyzed in aggregate.) Each post is also associated with an anonymized IP address of the posting user.

Identifying sockpuppets. No explicit labels of sockpuppets exist in any of the discussion communities, so to identify sockpuppets, we use multiple signals that together suggest that accounts are likely to share the same owner – the IP address of a comment, as well as the times at which comments are made, and the discussions they post in. Our approach draws on the approach adopted by Wikipedia administrators who identify sockpuppets by finding accounts that make similar edits on the same Wikipedia article, in near-similar time and from same IP address [1]. As we are primarily interested in identifying sockpuppets with high precision, the criteria we adopt is relatively conservative – relaxing these criteria may improve recall, but at the cost of more false positives.
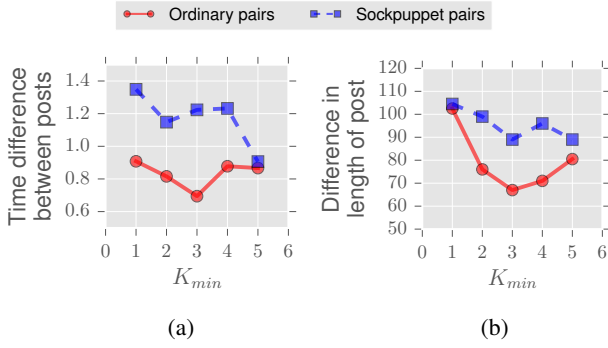
Figure 2: Varying K_{min}, the minimum number of overlapping sessions between users for them to be identified as sockpuppets. For sockpuppet pairs (blue) the time between posts, and the difference in post lengths reach a minimum value at $K_{min} = 3$.

To limit spurious detection of sockpuppets, we filter the data and remove any IP addresses used by many user accounts, as these accounts may simply be accessed from behind a country-wide proxy or intranet. We also do not consider user accounts that post from many different IP addresses, since they have high chance of sharing IP address with other accounts. Specifically, for each discussion community, we remove the top 5% most used IP addresses and the top 5% accounts that have the most IP addresses.

Then, we identify sockpuppets as user accounts that post from the same IP address in the same discussion within T minutes for at least K_{min} different discussions. Here, we set $T = 15$ minutes (larger values result in empirically similar findings). To pick an appropriate value for K_{min}, we use two metrics that prior work has found indicative of sockpuppets: the time difference between posts made by two accounts, and the difference in the length of posts [6, 40, 20, 33].

Figure 2 plots these quantities for identified pairs of sockpuppets as well as random pairs of users. We observe that regardless of the value of K_{min}, the identified sockpuppets post more closely in time and write posts more similar in length, compared to a random pair of users. Further, we find that among sockpuppet pairs, these quantities achieve their minimum at $K_{min} = 3$, which means that sockpuppets are most reliably identified at that value of K_{min}.

To summarize, we define sockpuppets as user accounts that post from the same IP address in the same discussion in close temporal proximity at least 3 times. We then define sockpuppet groups as maximal sets of accounts such that each account satisfies the above definition with at least one other account in the group. Overall, we identify a total of 1,623 sockpuppet groups, consisting of 3,656 sockpuppets from nine different online discussion communities (Table 1). As most sockpuppet groups contain two sockpuppets (Figure 3(a)), we focus our analyses on pairs of sockpuppets.

We give an example of an identified sockpuppet group below, which consists of three users: S_1 and S_2 comprise a pair of sockpuppets, while O is an ordinary user. After an unusually positive interaction between the two sockpuppets, O identifies them as being controlled by the same puppetmaster:

S_1: Possibly the best blog I've ever read major props to you.

$\hookrightarrow S_2$: Thanks. I knew Marvel fans would try to flame me, but they have nothing other than "oh that's your opinion" instead of coming up with their own argument.

$\hookrightarrow O$: Quit talking to yourself, *******. Get back on your meds if you're going to do that.
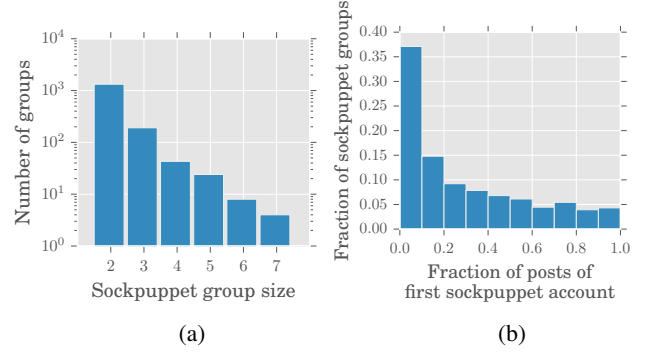
Figure 3: (a) Number of sockpuppet groups, *i.e.* sockpuppets belonging to the same puppetmaster. (b) The second sockpuppet in a group tends to be created shortly after the first.
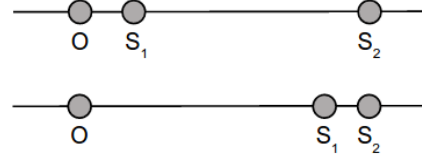
Figure 4: Two hypotheses how similarity of sockpuppet pairs and ordinary users relates to each other. Top: Under the double life hypothesis, sockpuppet S_1 is similar to an ordinary user O, while S_2 deviates. Bottom: Alternative hypothesis is that both sockpuppet accounts are highly different from ordinary users.

## 3. CHARACTERIZING SOCKPUPPETRY

Having identified sockpuppets, we now turn to characterizing their behavior. We study when sockpuppets are created and how their language and social networks differ from ordinary users across all nine discussion communities.

Sockpuppets are created early. To understand when sockpuppets are created, we examine the activity of the first sockpuppet account in each sockpuppet pair. Figure 3(b) shows the fraction of total number of posts made by the first sockpuppet before the second sockpuppet is created. The second sockpuppet tends to be created during the first 10% of the posts, with a median of 18% posts written by first sockpuppet before the second sockpuppet begins posting. In other words, sockpuppets tend to be created early in a user's lifetime, which may indicate that sockpuppet creation is premeditated and not a result of user's interactions in the community.

Matching sockpuppets with ordinary users. On average, sockpuppets write more posts than ordinary users (699 vs. 19) and participate in more discussions (141 vs. 7). To control for this disparity, in all our subsequent analyses we use propensity score matching [34] to match sockpuppets with ordinary users that have similar numbers of posts and make posts to the same set of discussions.

3.1 Do puppetmasters lead double lives?

First, we explore an important question about how the behavior of sockpuppets S_1 and S_2 controlled by the same puppetmaster relates to that of an ordinary user O. Two possible hypotheses are illustrated in Figure 4. First is the *double life* hypothesis, where the puppetmaster maintains a distinct personality for each sockpuppet – one sockpuppet, S_1, behaves like an ordinary user, while the other, S_2, behaves more maliciously. Under this hypothesis we would expect that the linguistic similarity of posts written by O and
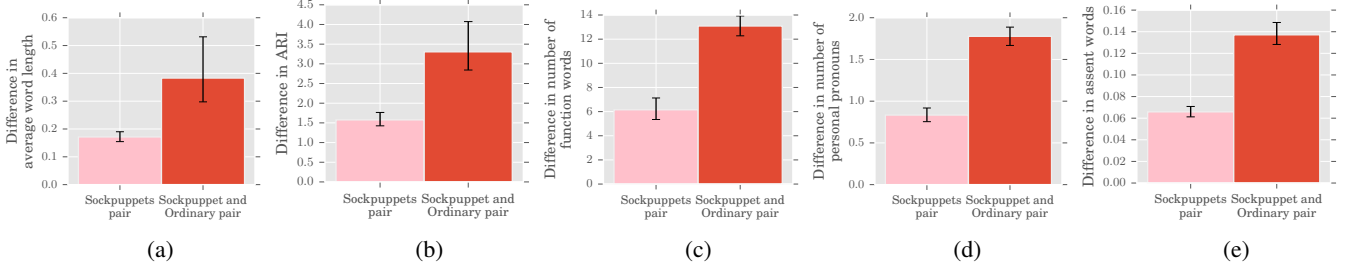
Figure 5: Difference in properties of sockpuppet pairs and that of sockpuppet-ordinary pairs. Sockpuppet pairs are more similar to each other in several linguistic attributes.

S_1 would be high, and that between both S_1 and S_2, as well as O and S_2 to be significantly lower. In the alternative hypothesis (Figure 4 (bottom)), both sockpuppets act maliciously. In this case, we might expect that the linguistic similarity between S_1 and S_2 would be low, but that between S_1 and O, and S_2 and O to be much lower.

To find out which is the case, we compare the language of pairs of sockpuppets, and that of each sockpuppet with an ordinary user. To control for user activity, we again match sockpuppets with ordinary users that have similar posting activity, and that participate in similar discussions. Specifically, for each user, we created a feature vector consisting of several linguistic features computed from that user's posts, including LIWC categories and sentiment, the average number of words in a post, the average fraction of special characters. We then compute the cosine similarity of the feature vectors.

We find that on average, the two sockpuppets are more similar to each other than either is to an ordinary user ($p < 0.001$). Figure 5 highlights that these observations hold for individual features as well – the difference between two sockpuppets' readability score (ARI), average word length, number of function words, personal pronouns and assent words are smaller than that of either sockpuppet and an ordinary user. This suggests that the double life hypothesis (Figure 4(top)) is less likely to be true than the alternate hypothesis (Figure 4(bottom)).

In other words, this experiment suggests that puppetmasters do not lead double lives, and that it is generally not the case that individual sockpuppets controlled by the same puppetmaster behave differently. Rather, sockpuppets as a whole tend to write differently from ordinary users, and sockpuppets controlled by the same puppetmaster all tend to write similarly to each other.

3.2 Linguistic Traits of Sockpuppets

Having established that different sockpuppets controlled by the same puppetmaster behave consistently, we now turn our attention to quantify their linguistic traits more precisely. Here, we focus on comparing various measures of similarity $sim(S_i, O)$ of a sockpuppet S_i ($i = \{1, 2\}$) and a matched ordinary user O. Specifically, we use LIWC word categories [31] to measure the fraction of each type of words written in all posts, and VADER [18] to measure sentiment of posts. We report the average values for sockpuppets and the corresponding p-values by performing paired t-tests for each sockpuppet and its matching ordinary user.

Do sockpuppets write differently from ordinary users? Linguistic traits have been used to identify trolls in discussions [11], vandalism on Wikipedia [32], and fake reviewers on e-commerce platforms [27]. For example, deceptive authors tend to increase their usage of function words, particles and personal pronouns [2, 5]. They use more first- and second-person singular personal pronouns (e.g., 'I', 'you'), while reducing their use of third-person singular personal pronouns (e.g., 'his', 'her'). They also tend to
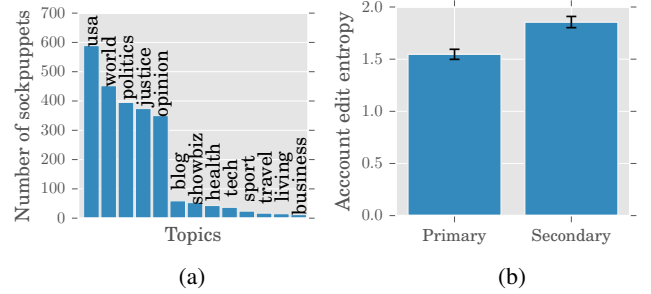
Figure 6: (a) Histogram for the most active topic for each sockpuppet account. (b) In a sockpuppet group, the secondary sockpuppets tend to be used alongside the primary sockpuppet.

oversimplify their writing style by writing shorter sentences and writing words with fewer syllables.

We make similar observations with respect to sockpuppets. First, we observe that they tend to write posts that are more self-centered – and use "I" more often than ordinary users (0.076 for sockpuppets vs 0.074 for ordinary users, $p<0.001$). Sockpuppets also use "you" more often (0.017 vs 0.015, $p<0.01$) but third-person singular personal pronouns and plural personal pronouns (i.e., 'we', 'he', 'she', and 'they') less (e.g. 0.016 vs 0.018 for 'he/she' words, $p < 0.001$), indicating that they tend to address other users in the community more directly. Similarly, we observe that sockpuppets also write shorter sentences than ordinary users (a mean of 12.4 vs. 12.9 words per sentence, $p < 0.001$). However, in contrast to prior work on deceptive writing, sockpuppets use a similar number of syllables per word (1.29 vs 1.28, $p = 0.35$).

Turning to differences in LIWC categories, we observe that sockpuppets also appear to write worse than ordinary users. They are more likely to swear (0.003 vs 0.002, $p < 0.05$) and use more punctuation (0.057 vs 0.055, $p < 0.05$), while using fewer alphabetic characters (0.769 vs 0.771, $p < 0.05$). Sockpuppets also use fewer standard English parts-of-speech, such as articles, verbs, adverbs and conjuctions (e.g. for articles, 0.062 vs 0.064, $p < 0.001$). However, while trolls wrote posts that were less readable [11], sockpuppets write posts with similar readability (automated readability index, or ARI = 11.24 vs 11.41 of ordinary users, $p=0.09$). Sockpuppets also tend to agree more in their posts (0.002 vs -0.012, $p < 0.05$), possibly to minimize conflict with others and support their other sockpuppet account. They also express less negative sentiment (0.022 vs 0.023, $p < 0.001$), though their overal sentiment, subtracting negative from positive sentiment, is similar to that of ordinary users (0.030 vs 0.028, $p = 0.43$).

3.3 Activity and Interactions

Next, we study how sockpuppets interact with the community at large, and how it responds to these sockpuppets.

Sockpuppets start fewer discussions, and post more in existing discussions. First, we note that sockpuppets start fewer discussions, but rather post more within existing discussions (65% of sockpuppets' posts are replies compared to 51% for ordinary users, $p < 0.001$). This shows that sockpuppets are mainly used to reply to other users.

Sockpuppets tend to participate in discussions with more controversial topics. Do sockpuppets create accounts to participate in certain topics? To answer this, we look at the topics of the discussions in CNN on which sockpuppets post. As shown in Figure 6(a) topics that tend to attract more controversy such as *usa, world, politics, justice* and *opinion*, also attract the majority of sockpuppets, while other topics such as health and showbiz have comparatively fewer sockpuppets. This indicates that one of the main motivations for using sockpuppets is to use them to build support for a particular position, corroborating prior work [6, 47].

Sockpuppets are treated harshly by the community. A community can provide feedback to a sockpuppet in three ways: other users can vote on or report their posts, and moderators can delete the posts. Comparing the posts made by sockpuppets with those made by ordinary users, we find that sockpuppets' posts receive a greater fraction of downvotes (0.51 vs 0.43, $p < 0.001$), are reported more often (0.05 vs 0.026, $p < 0.001$) and are also deleted more often (0.11 vs 0.08, $p < 0.001$). Moreover, sockpuppets are also blocked by the moderators more often (0.09 vs 0.07, $p < 0.001$). Overall, this suggests that sockpuppets are making undesirable comments.

Sockpuppets in a pair interact with each other more. Pairs of sockpuppets also tend to post together in more sub-discussions compared to random pairs of ordinary users (6.57 vs 0.33, $p < 0.001$). Moreover, looking at when posts are made, pair of sockpuppets also post more frequently on the same discussion within 15 minutes of each other (7.8 vs 4.28, $p < 0.001$). In other words, pairs of sockpuppets are significantly more likely to interact with one another, and post at the same time, than two ordinary users would.

Sockpuppets in a pair upvote each other more. Looking at votes, pairs of sockpuppets vote significantly more often on each other's posts than random pairs of ordinary users (9.35 vs 0.40 votes, $p < 0.001$). Among the two sockpuppets in a pair, the secondary sockpuppet votes more on primary sockpuppet's posts than vice-versa (14.2 vs 4.5 votes, $p < 0.01$). Moreover, pair of sockpuppets largely give positive votes to each other as compared to ordinary users (0.987 vs 0.952; $p < 0.05$). Altogether, sockpuppets in a pair use their votes to significantly inflate one another's 'popularity'.

Secondary sockpuppets are used in conjunction with primary sockpuppets. Puppetmasters, while controlling multiple sockpuppets, may either use multiple sockpuppets at the same time, or different sockpuppets at different times. To quantify how a puppetmaster may switch between using different sockpuppets, we compute the fraction of consecutive posts made by a particular sockpuppet, and then compute the entropy of this distribution. This way we quantify how much intertwined is the usage of both sockpuppet accounts. For instance, consider two sockpuppets controlled by the same puppetmaster. The puppetmaster first uses S_1 to write 5 posts, then uses S_2 to write 1 post, switches back to S_1 to write 4 more posts, and then finally switches back to S_2 to write 1 more post. The entropy of this post sequence for S_1 is $-\frac{5}{9}\log\frac{5}{9} - \frac{4}{9}\log\frac{4}{9}$, while that for S_2 is $-\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2}$.
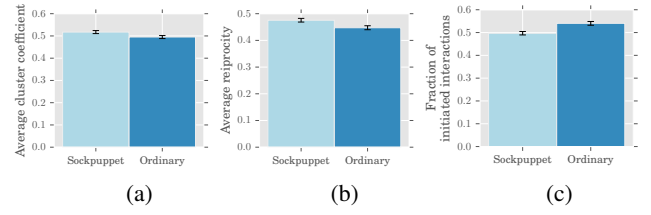
Figure 7: Comparison of egonetwork of sockpuppets and similar random users in the reply network.

Thus, a lower entropy signifies that a particular sockpuppet is not being used at the same time as the other sockpuppet, while higher entropy indicates that that sockpuppet is being used at the same time as another sockpuppet, with the puppetmaster constantly switching between the two.

Figure 6(b) shows the entropy of the primary and the secondary sockpuppets in a sockpuppet pair. We find that secondary accounts tend to have higher entropy, meaning that these sockpuppets are more likely to be used in conjunction with the primary sockpuppet, and thus may be used to support the primary account (e.g., in writing supportive replies). In contrast, primary sockpuppets have lower entropy, meaning they tend to be used more exclusively.

3.4 Reply network structure

Last, we examine the user-user interaction network of the entire discussion community. To do this, we create a *reply network*, where a node represents a user and an edge from node A to node B indicates that A replied to B's post at least once. Figure 1 shows the reply network of The AV Club discussion community, with red nodes denoting the sockpuppets and blue nodes denoting ordinary accounts. Here, we observe that the nodes denoting sockpuppets are more central in the network. In particular, we find that sockpuppets tend to have higher pagerank than ordinary users (2×10^{-4} vs 1×10^{-6}, $p < 0.001$).

To further understand the differences in how the sockpuppets interact with other users, we additionally compare the ego network of sockpuppets with that of ordinary users (Figure 7). We observe that both the number of nodes, and density of the ego networks of sockpuppets and ordinary users are similar (291.5 vs. 291.3 nodes, $p = 0.97$, and densities of 0.24 vs. 0.22, $p < 0.01$). However, the ego networks of sockpuppets are more tightly knit, as measured by the average clustering coefficient (Figure 7(a), 0.52 vs 0.49, $p < 0.001$). The nodes in a sockpuppet's ego network reply more to their neighbors, as measured by the average reciprocity (Figure 7(b), 0.48 vs 0.45, $p < 0.001$) with sockpuppets generally initiating more interactions (that is, they reply to more users than the users that reply to it, Figure 7(c), 0.51 vs 0.46, $p < 0.001$). These observations suggest that sockpuppets are highly active in their local network, and also generate more activity among the other users.

4. TYPES OF SOCKPUPPETRY

Different types of sockpuppets exist, and their characteristics suggest that they may serve different purposes. Here, in contrast to prior work which assumes that sockpuppets usually pretend to be other users [37, 14, 40], we find that sockpuppets can differ in their deceptiveness – while many sockpuppets do pretend to be different users, a significant number do not. When sockpuppets participate in the same discussions, they may also differ in their supportiveness – sockpuppets may be used to support other sockpuppets of the same puppetmaster, while others may choose not to.
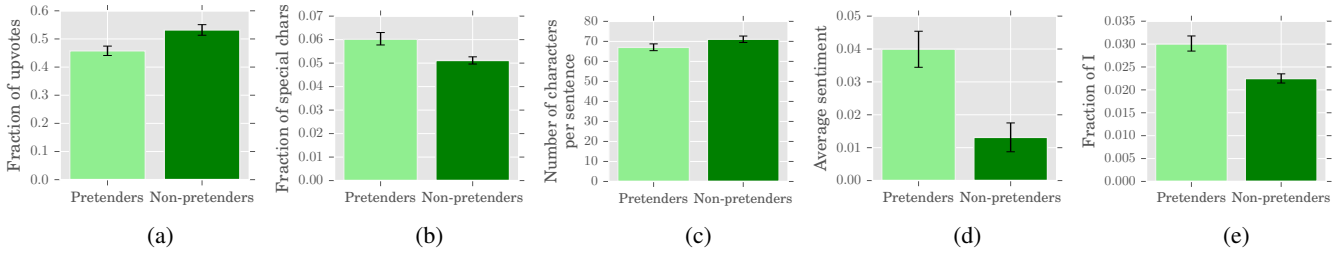
Figure 8: Differences between pretenders and non-pretenders: (a) fraction of upvotes, (b) fraction of special characters in posts, (c) number of characters per sentence, (d) average sentiment, (e) usage of first person pronoun ("I").
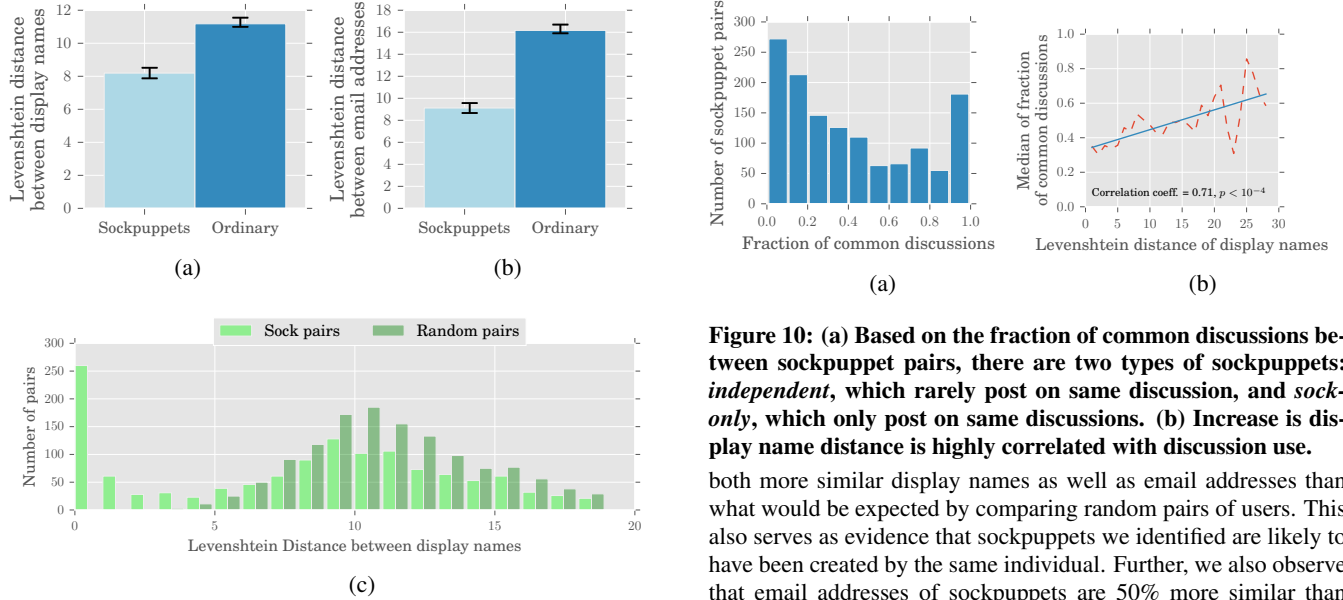

Figure 9: The (a) display names and (b) email addresses of the sockpuppet accounts are more similar to each other compared to similar random pairs. (c) Based on the distance of display names, sockpuppets can be *pretenders* (high distance) or *non-pretenders* (low distance).

4.1 Deceptiveness: Pretenders vs. non-pretenders

A pair of sockpuppets can pretend to be two separate individuals, or may simply be two user accounts an individual uses in different contexts, without any masquerading. We refer to the former group of sockpuppets as *pretenders*, and the latter group as *non-pretenders*.

One way we might quantify the deceptiveness of a sockpuppet pair is to examine the similarity of display names and email addresses (we only examine the part of the email address before the @-sign). Display names are public and show up next to user's comments, while email addresses are private and only visible to forum administrators. If a pair of sockpuppets wants to appear as two separate users, each may adopt a display name that is substantially different from the other in order to deceive community members. Puppetmaster may also adopt significantly different email addresses to avoid detection by system administrators. To quantify this difference, we measure the Levenshtein distance between two display names, as well as the corresponding email addresses.

Figures 9(a) and 9(b) compare how display names and email addresses differ between pairs of sockpuppets, and between random pairs of ordinary users. We observe that sockpuppets pairs have
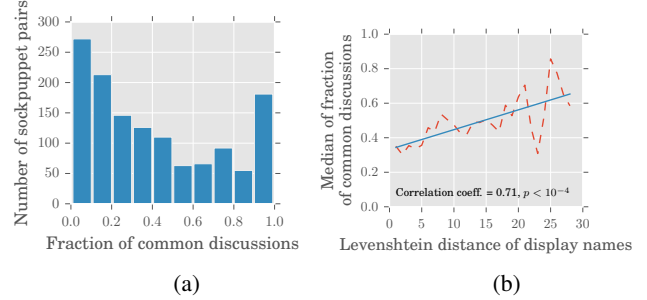
Figure 10: (a) Based on the fraction of common discussions between sockpuppet pairs, there are two types of sockpuppets: *independent*, which rarely post on same discussion, and *sock-only*, which only post on same discussions. (b) Increase is display name distance is highly correlated with discussion use.

both more similar display names as well as email addresses than what would be expected by comparing random pairs of users. This also serves as evidence that sockpuppets we identified are likely to have been created by the same individual. Further, we also observe that email addresses of sockpuppets are 50% more similar than those of ordinary accounts, while display names of sockpuppets are only 25% more similar than expected at random. This observation may be explained by the fact that sockpuppets put more effort into picking unique display names, which are public-facing, and less effort into picking unique email addresses, which are private-facing and less likely to be noticed.

But are all sockpuppets simply more likely to have more similar display names? Examining the distribution of the distances between display names in Figure 9(c), we find that the distribution for random pairs is unimodal, while for sockpuppets it is bimodal. This bimodality suggests that two types of sockpuppets pairs exist. The first type of sockpuppets has virtually identical display names (Levenshtein distance < 5), and these are what we call non-pretenders. The second type of sockpuppets has substantially different display names (Levenshtein distance ≥5), and we call them pretenders. Pretenders are likely to be created for deception and use different display names to avoid detection. Non-pretenders on the other hand have similar display names and this may implicitly signal to the community that they are controlled by the same individual, and thus may be less likely to be malicious.

Across all communities, we find 947 pretender and 403 non-pretender sockpuppet groups. We observe that pretenders tend to participate in the same discussions. For example, Figure 10(a) plots the fraction of common discussions over all sockpuppet groups. We observe bimodality here as well, which may be partially explained by the bimodality of the distribution of display name distances – Figure 10(b) shows that the likelihood of a pair of sockpuppets participating in the same discussion increases as their display names

become more different (fitted regression line shown in solid for clarity). In other words, these observations suggest that sockpuppets that participate in many common discussions have very different display names (high Levenshtein distance) and are thus pretenders, while accounts that participate in few common discussions tend to have similar display names and are thus non-pretenders.

Figure 8 additionally illustrates the other differences of pretenders and non-pretenders. We find that pretenders' posts are both more likely to be reported (0.06 fraction of all pretenders' posts are reported vs 0.03 for non-pretenders, $p < 0.001$), be deleted by moderators (0.11 vs 0.08, $p < 0.001$), and receive a smaller fraction of up-votes (0.45 vs 0.53, $p < 0.001$). Pretenders also write posts that contain more uppercase (0.07 vs 0.05, $p < 0.001$) and special characters (0.06 vs 0.05, $p < 0.001$), which suggests both shouting, as well as swearing. In contrast, non-pretenders wrote posts which were longer (35.2 words vs 38.6, $p < 0.05$) and more readable (ARI 11.15 vs 11.58, $p < 0.05$). Pretenders' posts also contained more positive sentiment (0.04 vs 0.013, $p < 0.001$) and agreement words (0.006 vs 0.005, $p < 0.001$), suggesting that they tended to be more affable.

4.2 Supporters vs. Dissenters

Prior work suggests that a primary purpose of sockpuppets is to sway public opinion by creating consensus [35]. Thus, we focus our attention on sockpuppets participating in the same discussion, and examine how they interact with each other – do sockpuppets tend to support each other?

We study two ways in which a sockpuppet pair (S_1, S_2) may interact – directly, where one sockpuppet (S_2) replies to another sockpuppet (S_1), or indirectly, where one sockpuppet (S_2) replies to a third user (O) who had replied to the first sockpuppet (S_1).

We focus on the extent to which sockpuppet S_2 agrees with sockpuppet S_1, and measure agreement as the difference between fraction of words categorized by LIWC as assenting and those categorized as either negations or dissenting [45]. We additionally adjust the sign of agreement depending on who the replying sockpuppet is replying to. For example, S_2 may write a post disagreeing with an ordinary user O. But if that ordinary user O in turn disagreed with the initial sockpuppet S_1, then we assume that the replying sockpuppet S_2 is instead in agreement with the initial sockpuppet S_1. We divide sockpuppets into three groups – *supporters*, who have a positive agreement score, *non-supporters*, who have an agreement score of zero, and *dissenters*, who have a negative agreement score.

Across all communities, we find that 60% of the sockpuppets are non-supporters, while 30% are supporters. Only 10% of sockpuppets are dissenters. Examining these discussions, we find evidence that supporters tend to support the arguments of S_1 (e.g., 'I agree, or 'so true'), and sometimes make additional arguments in their favor (e.g., 'That will cost him the election [...]').

On the other hand, dissenters tend to argue against S_1 (e.g., 'That's not what you said [...]'). We hypothesize that one reason sockpuppets may disagree with each other, despite being controlled by the same puppetmaster, may simply be to attract more attention to the argument. In some cases, we observed a dissenter making easily refutable arguments (e.g., 'Ok if your [sic] so worried about being spied Throw away all your electronics.'), which may have served to discredit the opposing view.

Altogether, these observations suggest that within discussions, sockpuppets may adopt different roles. While most sockpuppets argue for other sockpuppets controlled by the same puppetmaster, a small but significant number instead argue against other sockpuppets instead.

Nonetheless, is there a relationship between deceptiveness and supportiveness? Figure 2 shows that overall, users who support

	Pretender	Non-pretender
Supporter	0.74	0.26
Non-supporter	0.70	0.30
Dissenter	0.58	0.42

Table 2: 74% of supporters, 70% of non-supporters and 58% of dissenters are pretenders.

Feature Set	Features
Activity	Reply egonetwork clustering coefficient and reciprocity, Number of posts, proportion of reply posts, Time between posts, tenure time
Community	Whether account is blocked, fraction of upvotes, Fraction of reported and deleted posts
Post	Number of characters, syllables, words, sentences, Fraction of punctuations, uppercase characters, etc., Number of syllables per word, words per sentences, etc. Readability metrics (*e.g.* ARI), LIWC (*e.g.* swear words), Agreement, sentiment and emotion strength

Table 3: Three sets of features were used to identify sockpuppets and sockpuppet pairs.

other users in discussions are most likely to be also pretending to be other users (74% of supporters are pretenders). Interestingly, when users dissent in a discussion, they are less likely to be a pretender. This suggests that pretending is most important when a puppetmaster is trying to create an illusion of consensus.

5. DETECTING SOCKPUPPETS

Our previous analysis found that sockpuppets generally contribute worse content and engage in deceptive behavior. Thus, it would be useful to create automated tools that can help identify sockpuppets, and assist moderators in policing online communities. In this section, we consider two classification tasks, both of which relate to the prediction of sockpuppetry. First, can we distinguish sockpuppets from ordinary users? And second, can we identify pairs of sockpuppets in the communities?

Based on the observations and findings from the analyses in the previous sections, we identify three sets of features that may help in finding sockpuppets and sockpuppet pairs: activity features, community features, and post features. For each user U, we develop the following features:

Activity features: This set of features is derived from U's posting activity. Prior research has shown that activity behavior of bots, spammers, and vandals is different from that of benign users [10, 24, 25, 12, 40]. Moreover, in our analysis, we have seen that sockpuppets make more posts and they start less sub-discussions. Therefore, the activity features we consider include the number of posts, the proportion of posts that are replies, the mean time between two consecutive posts, and U's site tenure, or the number of days from U's first post. Further, we use features based on how U is situated in the reply network. Here, U's local network consists of U, the users whose posts U replied to, and the users that replied to U's posts. We then consider clustering coefficient and reciprocity of this network. In addition, for the task of identifying pairs of sockpuppets, we use number of common sub-discussions between these sockpuppets to measure how often the two comment together.

Community features: Interactions between a user and the rest of the community may also be indicative of sockpuppetry. Community feedback on an account's posts has been effective in identifying trolls and cheaters [11, 4], and we also observed that sockpuppets are treated more harshly than ordinary users. Thus, we consider the fraction of downvotes on posts U wrote, as well as the fraction that were reported or deleted, in addition to whether U was blocked.
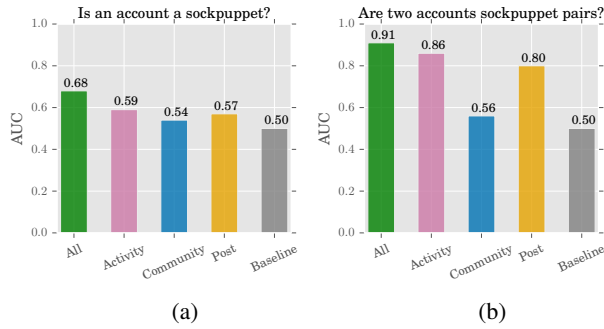
Figure 11: Classification performance to identify (a) sockpuppets from ordinary users and (b) pairs of sockpuppet accounts (bottom). Activity features have the highest performance.

Post features: Finally, we also measure the linguistic features of U's posts. These features have been very effective to identify sockpuppets [37, 6], authors of text [3, 20], deceptive writing styles [2], trolls [11], hoaxes [25], and vandalism [32]. In our analysis, we observe that sockpuppets do indeed write differently, for example, writing shorter sentences, using more swear words, and using more singular first-person pronouns. Thus, we incorporate linguistic features such as the number of characters, average word length, average number of syllables in words, number of big sentences, text readability (measured using ARI), and the different categories of LIWC features. Finally, we also consider sentiment, emotional valence, and agreement (as described previously).

5.1 Is an account a sockpuppet?

Given the posts a user has made, is it possible to distinguish a sockpuppet from an ordinary user? To control for user activity, we match sockpuppets and ordinary users on their total number of posts, as well as the discussions they post in. Matching gives a balanced dataset, where random guessing results in 50% accuracy. We perform 10-fold cross validation using a random forest classifier, then measure performance using ROC AUC.

Figure 11(a) shows results obtained individually with each feature set, as well as when considering all the features together. We observe that when using all features, the AUC is 0.68. Individually, all three feature sets perform similarly, with activity features slightly more predictive than the others (AUC=0.59).

To find the relative advantage of adding features, we perform forward feature selection. We observe that activity and post features perform close to the final AUC of 0.68, and that there is not much lift by adding the community features. This means that to identify sockpuppets, their activity and content of the post matter the most.

5.2 Are two accounts sockpuppet pairs?

Next, we turn our attention to identifying pairs of sockpuppets. Given a sockpuppet and two other users, where one is a sockpuppet, can we predict which user is the sockpuppet?

For each sockpuppet pair (S_1, S_2), we choose a matching ordinary account O for sockpuppet S_1. Again, this results in a balanced dataset and the task is to identify which of these two pairs is a sockpuppet. Features used in this experiment are the differences in the individual feature values for the two accounts each pair. We again evaluate performance using a random forest classifier.

Figure 11(b) shows the performance of the resulting classifier. We achieve a very high AUC of 0.91, suggesting that the interactions between sockpuppets are strongly predictive. Looking at the individual features, we see that activity features again perform the best, with close to 0.86 AUC. Community features perform the poorest, with an AUC of 0.56.

Overall, our results suggest that activity-based features best predict sockpuppetry. While it is possible to differentiate sockpuppets from ordinary users, it is significantly easier to find other sockpuppets in the same group once at one sockpuppet has been identified. The latter result suggests most importantly, the interactions between sockpuppets are the best way to identify them.

6. RELATED WORK

Our findings build on a rich vein of prior work in both deception and author identification.

Deception detection Sockpuppetry is situated in the broader field of deception. Deception online is aided by the virtue of anonymity [13]. It can occur as deceptive content as well as a deceptive agent [41]. The behavior of people changes when they deceive, for example, they reduce communication [48] and change the focus of their presentation [39]. When writing deceptively to hide their identity, authors tend to increase use of particles and personal pronouns, write shorter sentences, and show nervousness [2, 5, 22, 8]. Our work adds to this line of research by finding evidence of deceptive writing styles and presentation by sockpuppets – pretender sockpuppets may pretend to be different people by using different display names and they tend to write deceptively as well.

Motivations for sockpuppetry Turning to research that studies sockpuppetry specifically, one line of work has studied their motivations. Sockpuppetry is often used to avoid being banned, to create false consensus [38, 2] and support a person or a position [9], or vandalize content (*e.g.* on Wikipedia [37]). Relatedly, motivations for multiple account creation in online multiplayer games can either be benign (*e.g.* experimentation with different identities) or malicious (*e.g.* increasing in-game profit, cheating) [15, 16, 23, 9]. In our work, we find evidence for these motivations – sockpuppets in discussion communities sometimes support each other, and beyond malicious uses, some uses of sockpuppetry may be benign (*e.g.* a user may simply use different accounts to post in different topics).

Sockpuppetry and author identification Another line of work has also identified sockpuppets using textual information, link analysis and temporal information, both in online discussion forums [6, 47] and social networks [14, 26, 47, 43]. However, definitions of sockpuppets from previous research have tended to make assumptions about the usernames that sockpuppets use (*e.g.*, that they are similar [26]), their opinion towards topics (*e.g.*, they have the same opinion [6]), and their interactions (e.g., that they reply in support of each other's posts [47]). As such, these definitions tend to miss several types of sockpuppetry. In this work, we developed a robust methodology for identifying sockpuppets that makes fewer assumptions, and showed that a significant fraction of sockpuppets do use different names (i.e., the non-pretenders), and tend not to support each other in discussions (i.e., the non-supporters).

Sockpuppetry on Wikipedia has been studied extensively due to availability of manually-validated ground-truth data [37, 40, 44, 30]. However, in contrast to sockpuppetry in discussion communities, which is the focus of our work, sockpuppet editors on Wikipedia primarily edit articles, and the main purpose is not to interact with each other.

Closely related to sockpuppet detection is author identification, or the task of identifying the original author of a document [20, 21, 28, 29, 30, 33]. More recently, research has used multiple accounts of users across different social platforms to identify malicious users [42], and to identify accounts operated by the same user across different web platforms [19, 36, 46, 47]. In contrast to our work here, this line of research does not operate under the assumption of deception and thus may be less applicable to situations when authors try to obfuscate their writing [2, 5].

7. DISCUSSION AND CONCLUSION

Our findings shed light on how sockpuppets are used in practice in online discussion communities. By developing a robust methodology for identifying sockpuppets, we are able to comprehensively study their activity. Importantly, this methodology is able to identify sockpuppets that were created at significantly different times, use very different usernames or email addresses, write differently, or mostly post in different discussions.

Our work revealed differences in how sockpuppets write and behave in online communities. Sockpuppets use more singular first-person pronouns, write shorter sentences, and swear more. They participate in discussions with more controversial topics, and are especially likely to interact with other sockpuppets. These differences allowed us to build preditive models that robustly differentiate pairs of sockpuppets from ordinary users, as well as identify individual user accounts that are sockpuppets.

Nonetheless, our analysis has limitations that would be interesting to explore in future work. First, using our heuristics, we are not able to identify sockpuppets that are also throwaway accounts and are only used once before being abandoned. Next, we studied sockpuppetry in discussion forums where users are pseudonymous. It would be interesting to study the effect of using real identities (e.g., Facebook), or in completely anonymous settings (e.g., 4chan). We also primarily studied pairs of sockpuppets, but understanding how larger groups of sockpuppets function may reveal additional ways in which sockpuppets may coordinate, and may allow us to observe more pronounced effects of sockpuppetry. Further, behavior of sockpuppets in knowledge sharing platforms (e.g., StackOverflow) may be different from that in opinion expressing discussion platforms we studied – for instance, the primary purpose of sockpuppets in such platforms may primarily be to give additional 'upvotes' to their answers. Such a study would bring additional insights into sockpuppetry. Furthermore, prior work found that trolling correlates with sadism [7] – understanding the role of personality traits in sockpuppetry would also be valuable future work. Finally, even more robust methodologies for identifying sockpuppets may uncover an even wider range of behavior. For example, sockpuppets may exist beyond a single discussion community – for example, we found 14 different sockpuppet groups that were used in more than one online community. Studying these types of sockpuppets may allow us to better characterize how a sockpuppet's behavior may change in different communities.

By developing better techniques to identify sockpuppets, our work can be used to improve the quality of discussions online, where each discussant can better trust that their interactions with others are genuine. Nonetheless, while it is possible to identify sockpuppetry, one should be careful not to assume that all sockpuppets are malicious. Our work suggests that a significant number of sockpuppets do not pretend to be other users and were simply used to participate in different discussions. This observation suggests that some users find it valuable to separate their activity in different spheres of interests.

8. ACKNOWLEDGEMENT

9. REFERENCES

[1] Wikipedia sockpuppet investigation policy. https://goo.gl/89ieoY. Accessed: 2016-10-24.

[2] S. Afroz, M. Brennan, and R. Greenstadt. Detecting hoaxes, frauds, and deception in writing style online. In *Security and Privacy (SP), 2012 IEEE Symposium on*, pages 461–475. IEEE, 2012.

[3] S. Argamon and S. Levitan. Measuring the usefulness of function words for authorship attribution. In *Proceedings of the 2005 ACH/ALLC Conference*, 2005.

[4] J. Blackburn, R. Simha, N. Kourtellis, X. Zuo, M. Ripeanu, J. Skvoretz, and A. Iamnitchi. Branded with a scarlet c: cheaters in a gaming social network. In *Proceedings of the 21st international conference on World Wide Web*, pages 81–90. ACM, 2012.

[5] M. Brennan, S. Afroz, and R. Greenstadt. Adversarial stylometry: Circumventing authorship recognition to preserve privacy and anonymity. *ACM Transactions on Information and System Security (TISSEC)*, 15(3):12, 2012.

[6] Z. Bu, Z. Xia, and J. Wang. A sock puppet detection algorithm on virtual spaces. *Knowledge-Based Systems*, 37:366–377, 2013.

[7] E. E. Buckels, P. D. Trapnell, and D. L. Paulhus. Trolls just want to have fun. *Personality and individual Differences*, 67:97–102, 2014.

[8] D. B. Buller and J. K. Burgoon. Interpersonal deception theory. *Communication theory*, 6(3):203–242, 1996.

[9] A. Caspi and P. Gorsky. Online deception: Prevalence, motivation, and emotion. *CyberPsychology & Behavior*, 9(1):54–59, 2006.

[10] K.-T. Chen and L.-W. Hong. User identification based on game-play activity patterns. In *Proceedings of the 6th ACM SIGCOMM workshop on Network and system support for games*, pages 7–12. ACM, 2007.

[11] J. Cheng, C. Danescu-Niculescu-Mizil, and J. Leskovec. Antisocial behavior in online discussion communities. In *Ninth International AAAI Conference on Web and Social Media*, 2015.

[12] J. P. Dickerson, V. Kagan, and V. Subrahmanian. Using sentiment to detect bots on twitter: Are humans more opinionated than bots? In *Advances in Social Networks Analysis and Mining (ASONAM), 2014 IEEE/ACM International Conference on*, pages 620–627. IEEE, 2014.

[13] J. S. Donath et al. Identity and deception in the virtual community. *Communities in cyberspace*, 1999.

[14] K. Gani, H. Hacid, and R. Skraba. Towards multiple identity detection in social networks. In *Proceedings of the 21st International Conference on World Wide Web*, pages 503–504. ACM, 2012.

[15] R. Gilbert, V. Thadani, C. Handy, H. Andrews, T. Sguigna, A. Sasso, and S. Payne. The psychological functions of avatars and alt (s): A qualitative study. *Computers in Human Behavior*, 32:1–8, 2014.

[16] R. L. Gilbert, J. A. Foss, and N. A. Murphy. Multiple personality order: Physical and personality characteristics of the self, primary avatar and alt. In *Reinventing ourselves: Contemporary concepts of identity in virtual worlds*, pages 213–234. Springer, 2011.

[17] J. T. Hancock. Digital deception. *Oxford handbook of internet psychology*, pages 289–301, 2007.

[18] C. J. Hutto and E. Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth*

International AAAI Conference on Weblogs and Social Media, 2014.

[19] P. Jain, P. Kumaraguru, and A. Joshi. @ i seek'fb. me': Identifying users across multiple online social networks. In *Proceedings of the 22nd international conference on World Wide Web*, pages 1259–1268. ACM, 2013.

[20] F. Johansson, L. Kaati, and A. Shrestha. Detecting multiple aliases in social media. In *Proceedings of the 2013 IEEE/ACM international conference on advances in social networks analysis and mining*, pages 1004–1011. ACM, 2013.

[21] F. Johansson, L. Kaati, and A. Shrestha. Timeprints for identifying social media users with multiple aliases. *Security Informatics*, 4(1):7, 2015.

[22] P. Juola. Detecting stylistic deception. In *Proceedings of the Workshop on Computational Approaches to Deception Detection*, pages 91–96. Association for Computational Linguistics, 2012.

[23] Y. B. Kafai, D. A. Fields, and M. Cook. Your second selves: avatar designs and identity play in a teen virtual world. In *Proceedings of DIGRA*, volume 2007, 2007.

[24] S. Kumar, F. Spezzano, and V. Subrahmanian. Vews: A wikipedia vandal early warning system. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 607–616. ACM, 2015.

[25] S. Kumar, R. West, and J. Leskovec. Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes. In *Proceedings of the 25th International Conference on World Wide Web*, pages 591–602. International World Wide Web Conferences Steering Committee, 2016.

[26] D. Liu, Q. Wu, W. Han, and B. Zhou. Sockpuppet gang detection on social media sites. *Frontiers of Computer Science*, 10(1):124–135, 2016.

[27] A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance. What yelp fake review filter might be doing? In *Seventh International AAAI Conference on Weblogs and Social Media*, 2013.

[28] A. Narayanan, H. Paskov, N. Z. Gong, J. Bethencourt, E. Stefanov, E. C. R. Shin, and D. Song. On the feasibility of internet-scale author identification. In *Security and Privacy (SP), 2012 IEEE Symposium on*, pages 300–314. IEEE, 2012.

[29] J. Novak, P. Raghavan, and A. Tomkins. Anti-aliasing on the web. In *Proceedings of the 13th international conference on World Wide Web*, pages 30–39. ACM, 2004.

[30] P. P. Paul, M. Sultana, S. A. Matei, and M. Gavrilova. Editing behavior to recognize authors of crowdsourced content. In *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*, pages 1676–1681. IEEE, 2015.

[31] J. W. Pennebaker, M. E. Francis, and R. J. Booth. Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001):2001, 2001.

[32] M. Potthast, B. Stein, and R. Gerling. Automatic vandalism detection in wikipedia. In *European Conference on Information Retrieval*, pages 663–668. Springer, 2008.

[33] T. Qian and B. Liu. Identifying multiple userids of the same author. In *EMNLP*, pages 1124–1135, 2013.

[34] P. R. Rosenbaum and D. B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, pages 41–55, 1983.

[35] C. Seife. *Virtual Unreality: Just Because the Internet Told You, how Do You Know It's True?* Penguin, 2014.

[36] G. Silvestri, J. Yang, A. Bozzon, and A. Tagarelli. Linking accounts across social networks: the case of stackoverflow, github and twitter. In *International Workshop on Knowledge Discovery on the WEB*, pages 41–52, 2015.

[37] T. Solorio, R. Hasan, and M. Mizan. A case study of sockpuppet detection in wikipedia. In *Workshop on Language Analysis in Social Media (LASM) at NAACL HLT*, pages 59–68, 2013.

[38] B. Stone and M. Richtel. The hand that controls the sock puppet could get slapped. *New York Times*, 2007.

[39] C. L. Toma and J. T. Hancock. What lies beneath: The linguistic traces of deception in online dating profiles. *Journal of Communication*, 62(1):78–97, 2012.

[40] M. Tsikerdekis and S. Zeadally. Multiple account identity deception detection in social media using nonverbal behavior. *IEEE Transactions on Information Forensics and Security*, 9(8):1311–1321, 2014.

[41] M. Tsikerdekis and S. Zeadally. Online deception in social media. *Communications of the ACM*, 57(9):72–80, 2014.

[42] G. Venkatadri, O. Goga, C. Zhong, B. Viswanath, K. P. Gummadi, and N. Sastry. Strengthening weak identities through inter-domain trust transfer. In *Proceedings of the 25th International Conference on World Wide Web*, pages 1249–1259. International World Wide Web Conferences Steering Committee, 2016.

[43] B. Viswanath, A. Post, K. P. Gummadi, and A. Mislove. An analysis of social network-based sybil defenses. *ACM SIGCOMM Computer Communication Review*, 40(4):363–374, 2010.

[44] Z. Yamak, J. Saunier, and L. Vercouter. Detection of multiple identity manipulation in collaborative projects. In *Proceedings of the 25th International Conference Companion on World Wide Web*, pages 955–960. International World Wide Web Conferences Steering Committee, 2016.

[45] H. Yuan, D. O. Clifton, P. Stone, and H. H. Blumberg. Positive and negative words: Their association with leadership talent and effectiveness. *The Psychologist-Manager Journal*, 4(2):199, 2000.

[46] R. Zafarani and H. Liu. Connecting users across social media sites: a behavioral-modeling approach. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 41–49. ACM, 2013.

[47] X. Zheng, Y. M. Lai, K.-P. Chow, L. C. Hui, and S.-M. Yiu. Sockpuppet detection in online discussion forums. In *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2011 Seventh International Conference on*, pages 374–377. IEEE, 2011.

[48] L. Zhou and Y.-w. Sung. Cues to deception in online chinese groups. In *Hawaii international conference on system sciences, proceedings of the 41st annual*, pages 146–146. IEEE, 2008.