



Malicious Behavior on the Web: Characterization and Detection

Srijan Kumar (@srijankr)

Justin Cheng (@jcccf)

Jure Leskovec (@jure)

Slides are available at <http://snap.stanford.edu/www2017tutorial/>

Tutorial Outline

Malicious users

Trolling

Sockpuppets

Vandals

Misinformation

Fake reviews

Hoaxes

<http://snap.stanford.edu/www2017tutorial>



amazon



WIKIPEDIA



reddit



You Tube

Example

Why DC is better than Marvel



April 28, 2013 by [Eric_17](#)



bdiaz209
Possibly

bdiaz209 posts only on this discussion to support and defend Eric_17



Eric_17

April 28 2013, 12AM

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



Fellstrike

April 29 2013, 6PM

Quit talking to yourself, *****. Get back on your meds if you’re going to do that



Sock puppetry and fake reviews: publish and be damned



The Hand That Controls the Sock Puppet Could Get Slapped

By BRAD STONE and MATT RICHTEL JULY 16, 2007



WIKIPEDIA

Wikipedia blocks hundreds of 'scam' sock puppet accounts

🕒 2 September 2015 | Technology

🔗 Share

Wikipedia: Orangemoody

254 articles
edited/created
by 381 accounts

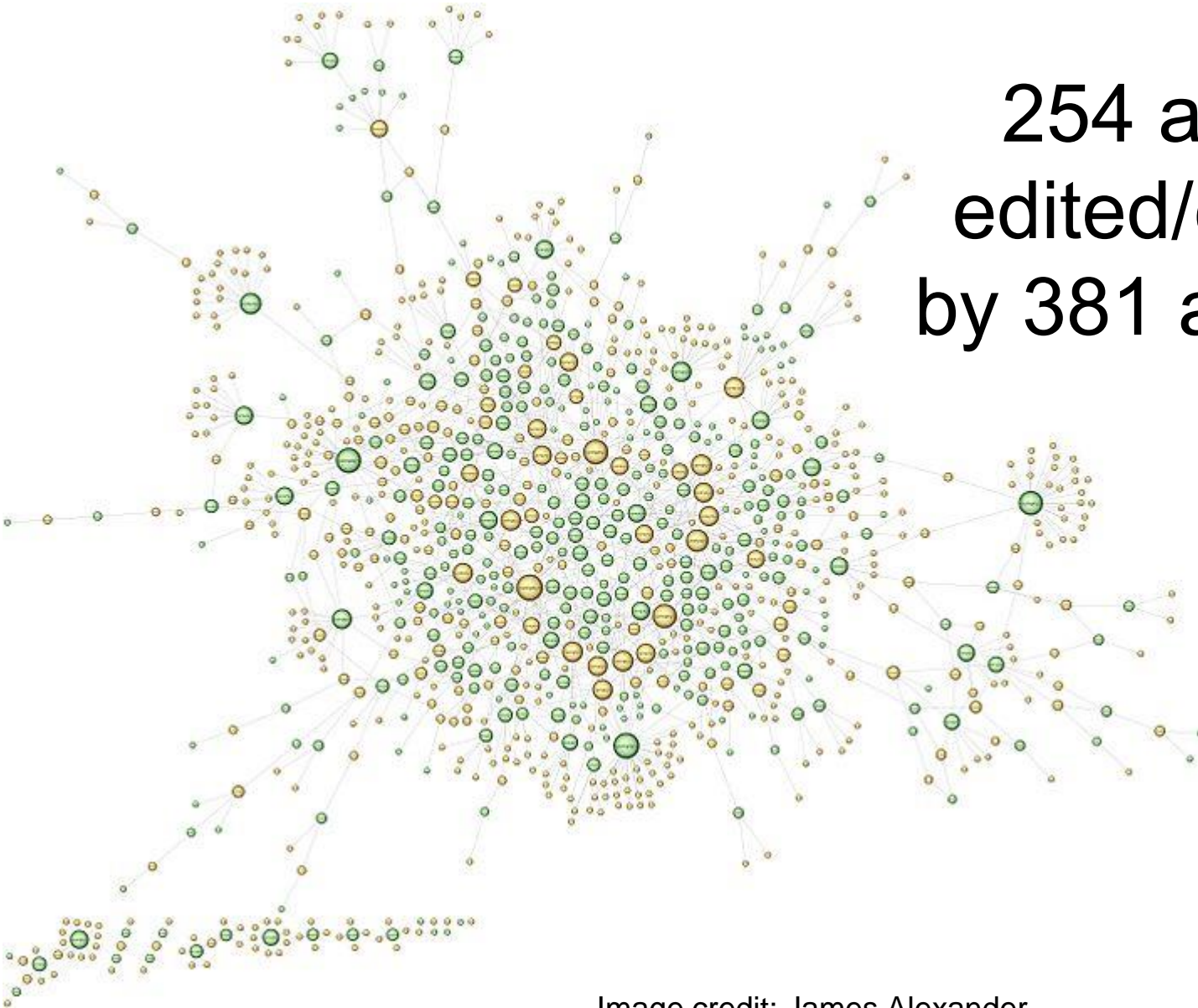


Image credit: James Alexander

Online sockpuppetry – why?



Diversify identity



Anonymize identity



Multiply identity



Privacy concern

Sockpuppets in Wikipedia

Inappropriate uses of alternative accounts

Editors must not use alternative accounts to mislead, deceive, disrupt, or undermine consensus. This includes, but is not limited to:

- **Creating an illusion of support**
- **Strawman socks**
- **Editing project space**
- **Circumventing policies or sanctions**
- **Contributing to the same page or discussion with multiple accounts**
- **Avoiding scrutiny**
- **Editing logged out to mislead**
- **Misusing a clean start.**
- **"Good hand" and "bad hand" accounts**

Legitimate uses

Valid reasons for an alternative account include:

- **Security**
- **Privacy**
- **Doppelgänger accounts**
- **Clean start under a new name**
- **Username violations**
- **Compromised accounts**
- **Humor accounts**

Wikipedia Sockpuppets



WIKIPEDIA



An editor has expressed a concern that this account may be a **sock puppet** of [Example](#) ([talk](#) · [contribs](#) · [logs](#)).

Please refer to [the sockpuppet investigation](#) of the sockpuppeteer, and editing habits or [contributions](#) of the sock puppet for evidence. [This policy subsection](#) may also be helpful.

Account information: [block log](#) – [current autoblocks](#) – [contribs](#) – [logs](#) – [abuse log](#) – [CentralAuth](#)

Suspected by volunteers,
confirmed and deleted by administrators

Sockpuppets in social media

- Similar login time
- Similar login IP address
- Similar usernames

Liu et al. (FCS 2016)

- Write similar to each other
- Similar point of view

Bu et al. (KBS 2013)

- Support one another

Zheng et al. (IIH-MSP 2011)

Sockpuppets in online discussions

Sockpuppets in online discussions



- IP address
- Posts on discussions
- Time of post

Sockpuppets are accounts that post from the same IP address in the same discussion very close in time, in at least 3 different instances.

Sockpuppets in online discussions

DISQUS



BREITBART



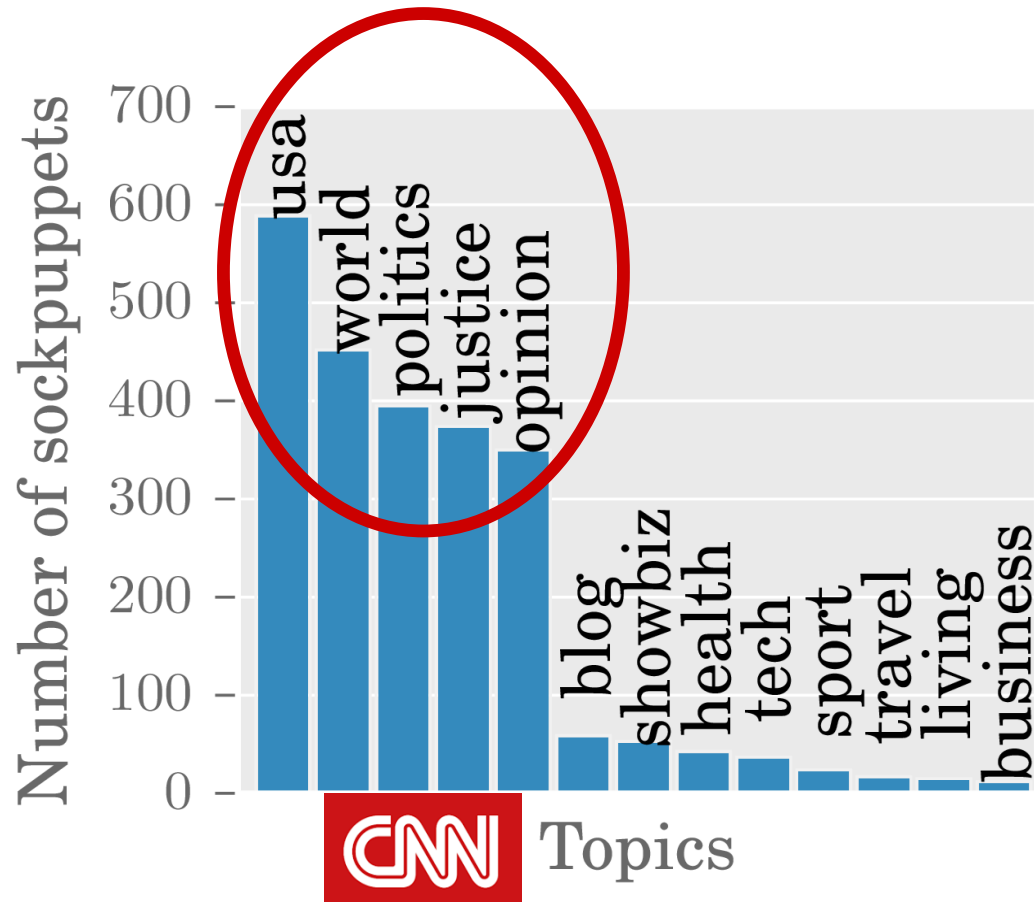
A.V. CLUB



2.9M	2.1M	62M	3,656
Users	Discussions	Posts	Sockpuppets

Characteristics of sockpuppets

Sockpuppets love controversy



One of the worst that I have seen yet...

jakey008

Feb 5 2013, 2PM

should have read the reviews first :(

ricobeans27

5 2013, 3

Couldn't agree more!!

Start fewer discussions

$p < 10^{-3}$

Agree more

$p < 10^{-3}$



Falcon-X32

I agree. You are absolutely right!

Feb 5 2013, 3PM

Write shorter sentences

$p < 10^{-3}$

Address others directly

$p < 10^{-3}$

Downvoted more

$p < 10^{-3}$

Write more self-centered posts

$p < 10^{-3}$



One of the worst that I have seen yet...



jakey008

Feb 5 2013, 2PM

should have read the reviews first :(



ricobeans27

Couldn't agree more.

Upvote each other more

$$p < 10^{-3}$$



Falcon-X32

I agree. You are absolutely right!



Feb 5 2013, 3PM



Smoothzilla

Thanks for your support!!!!

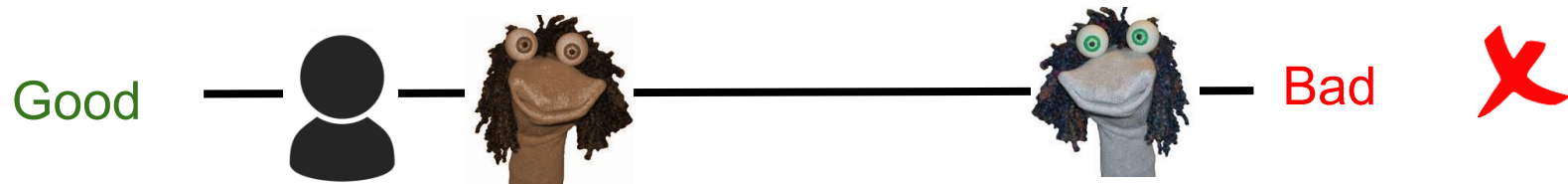
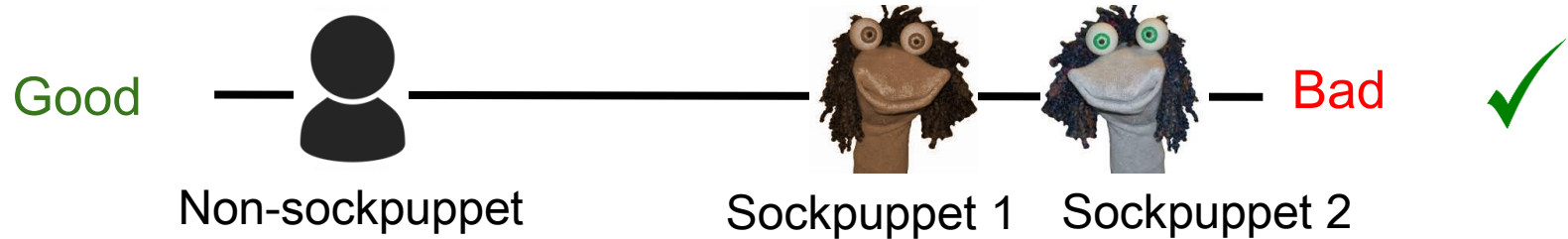


Feb 5 2013, 3PM

Interact more with each other

$$p < 10^{-3}$$

Do puppetmasters lead double lives?



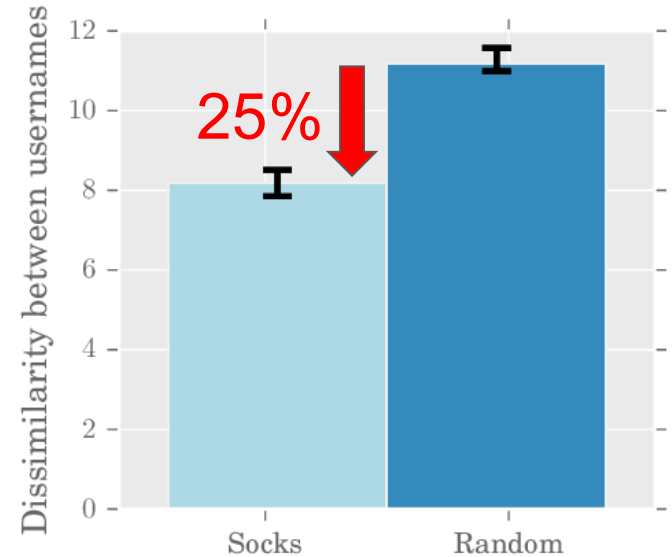
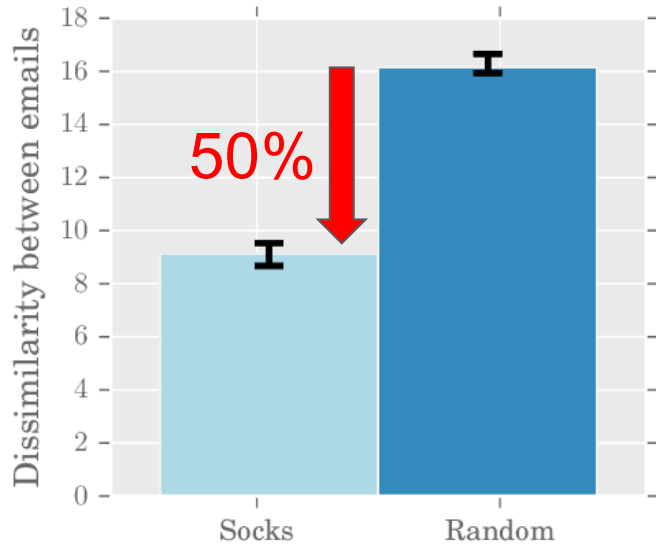
Both sockpuppets are more similar to
each other

$$p < 10^{-3}$$

“Good sock/Bad sock” not common

Types of sockpuppets

Are sockpuppets deceptive?



Sockpuppets pick **distinct usernames**
but
not as distinct email ids

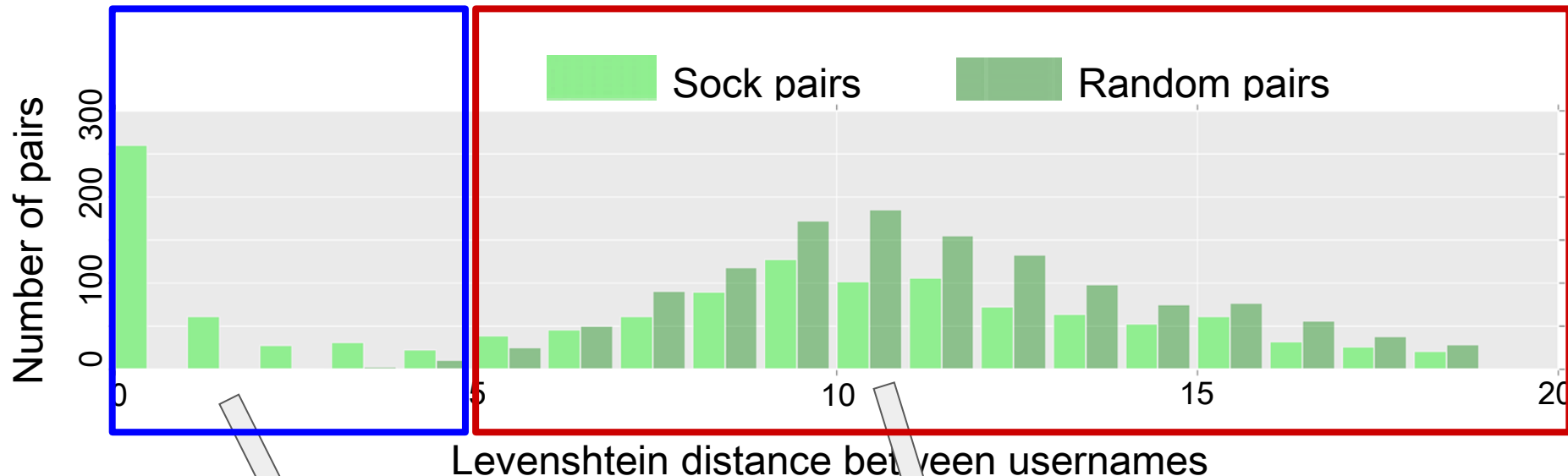
1/3

Deceptiveness

2/3

Non-Pretenders

Pretenders



srijan



srijan2



srijan



theRealBatman

One of the worst that I have seen yet...



Falcon-X32

Feb 5 2013, 2PM

best article i have read!!!



ricob

More opinionated

$p < 10^{-3}$

Feb 5 2013, 3PM

But this article doesn't make any sense



Smoothzilla

YOU ARE STUPID AND

Feb 5 2013, 3PM



Falcon-X32

Feb 5 2013, 3PM

i agree.. these morons dont know a thing



Swear more

$p < 10^{-3}$

Downvoted and
reported more

$p < 10^{-3}$



One of the worst that I have seen yet...



Falcon-X32

Feb 5 2013, 2PM

best article i have read!!!



ricobeans27

Feb 5 2013, 3PM

But this article doesn't make any sense



Falcon

Feb 5 2013, 3PM

Your opinion, my opinion



ricobeans27

Feb 5 2013, 4PM

You here to join?



Falcon-X32

Feb 5 2013, 5PM

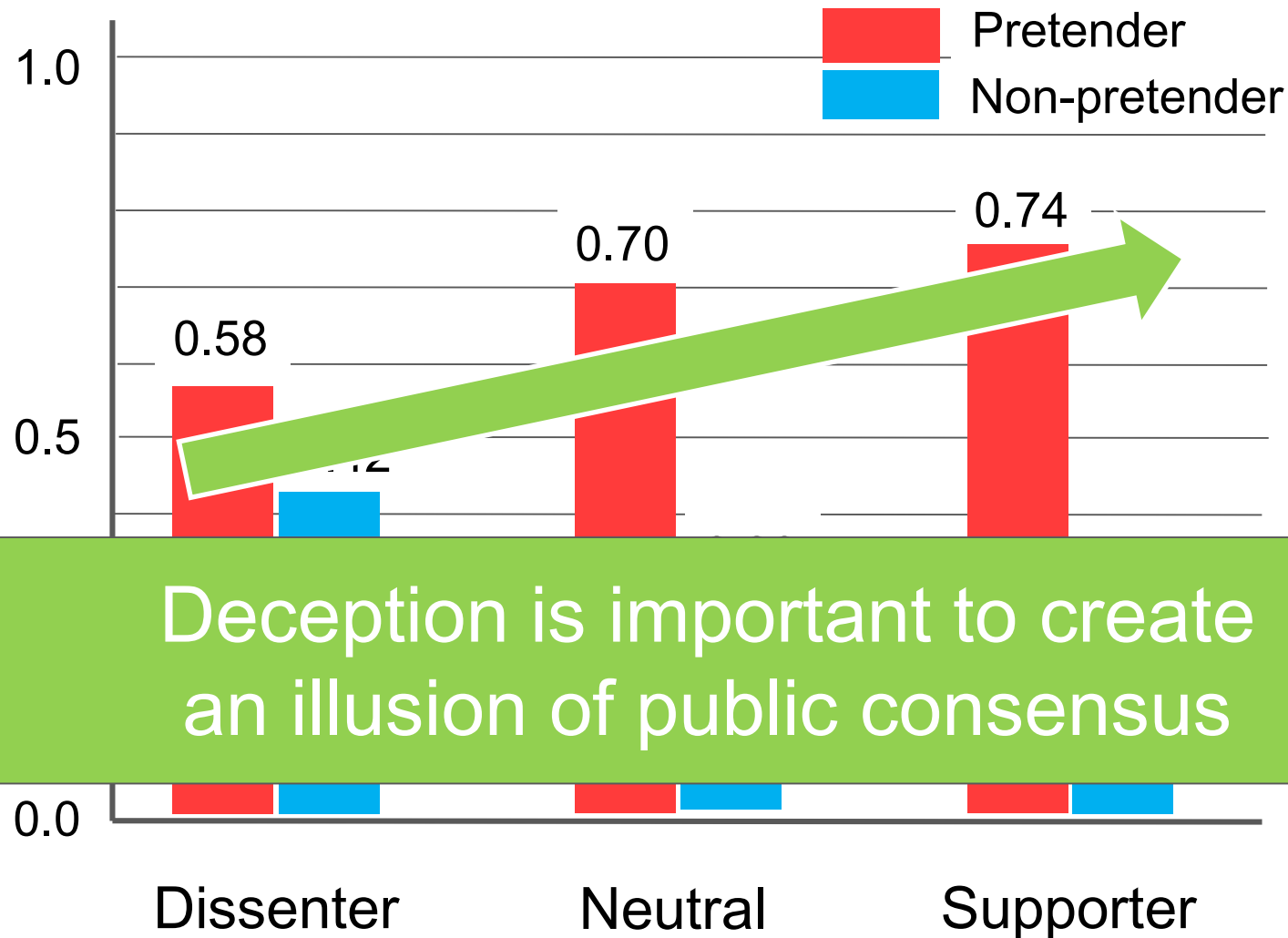
Sorry accidentally posted from my other account

Supportiveness

Does one sockpuppet support the other
or
defend from disagreeing user?

60%	30%	10%
Neutral	Support	Oppose

Supportiveness and Deceptiveness

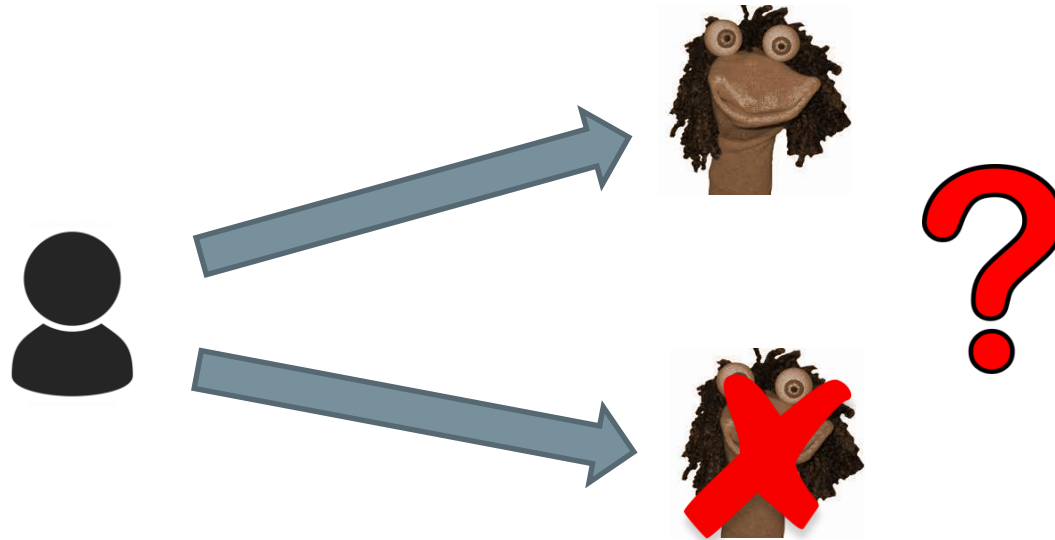


Detecting sockpuppets

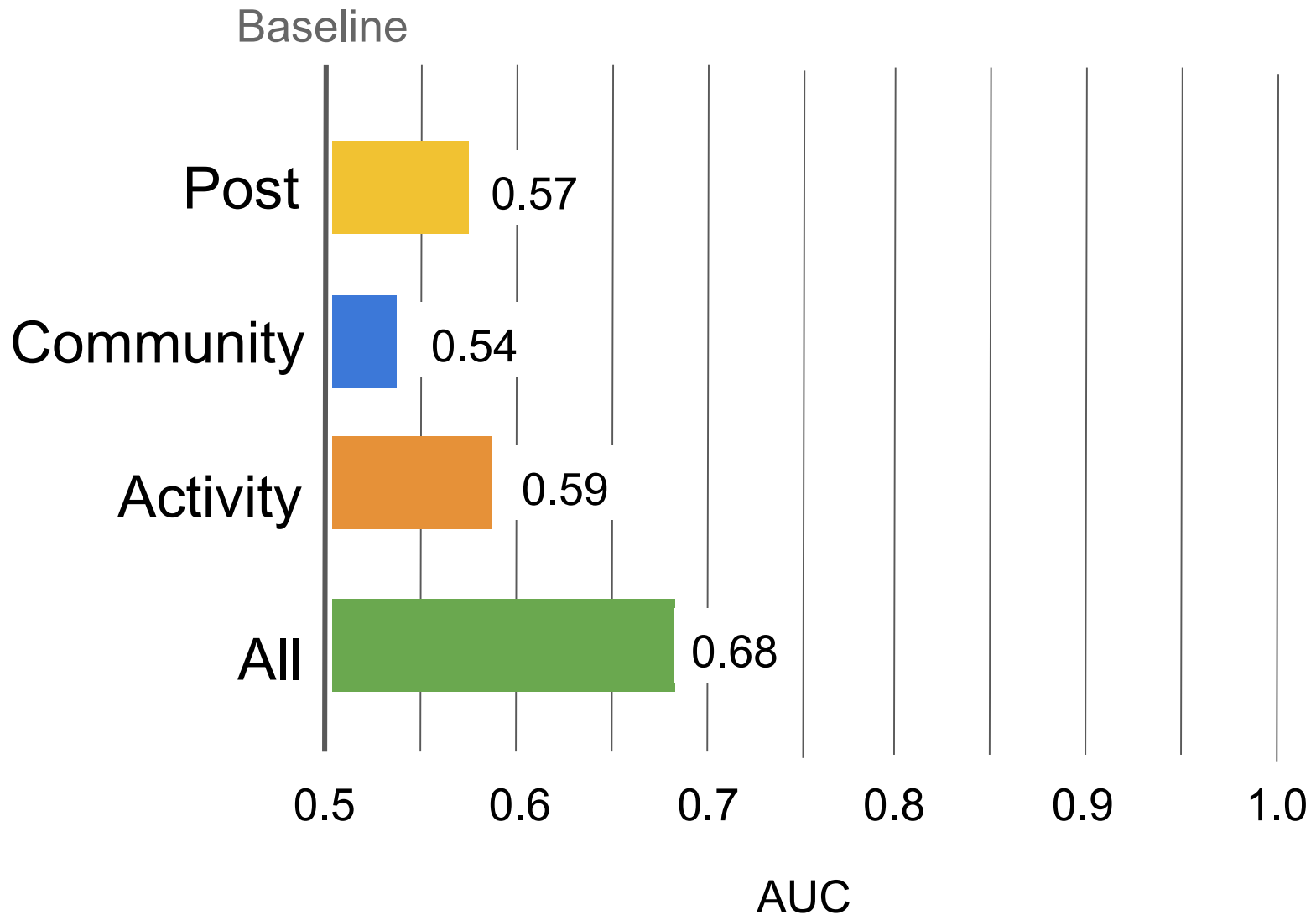
Features for detecting sockpuppets

- **Post:** count of words, characters; fraction of uppercases, punctuations, etc. LIWC features of text, readability, agreement
- **Community:** fraction of reported and deleted posts, upvotes, blocked?
- **Activity:** number of posts, time between posts, age of account

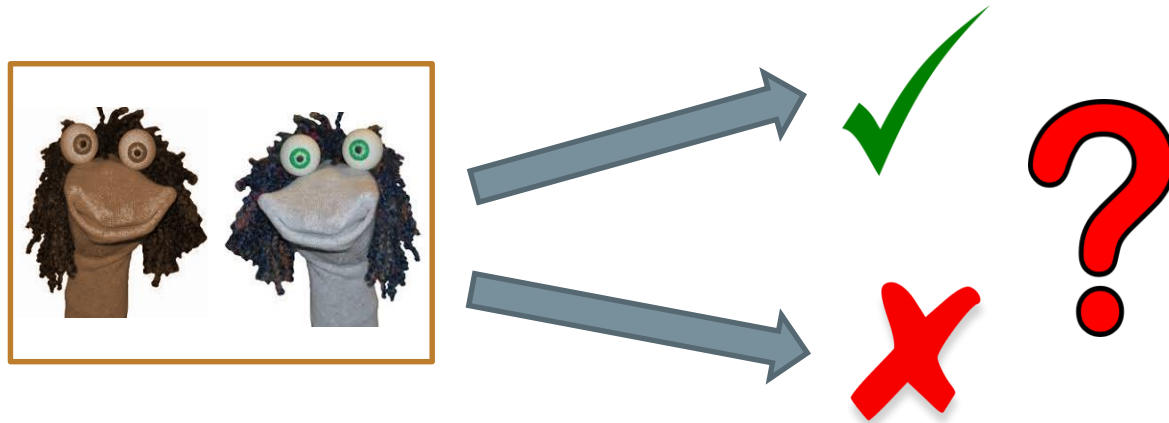
Is an account a sockpuppet?



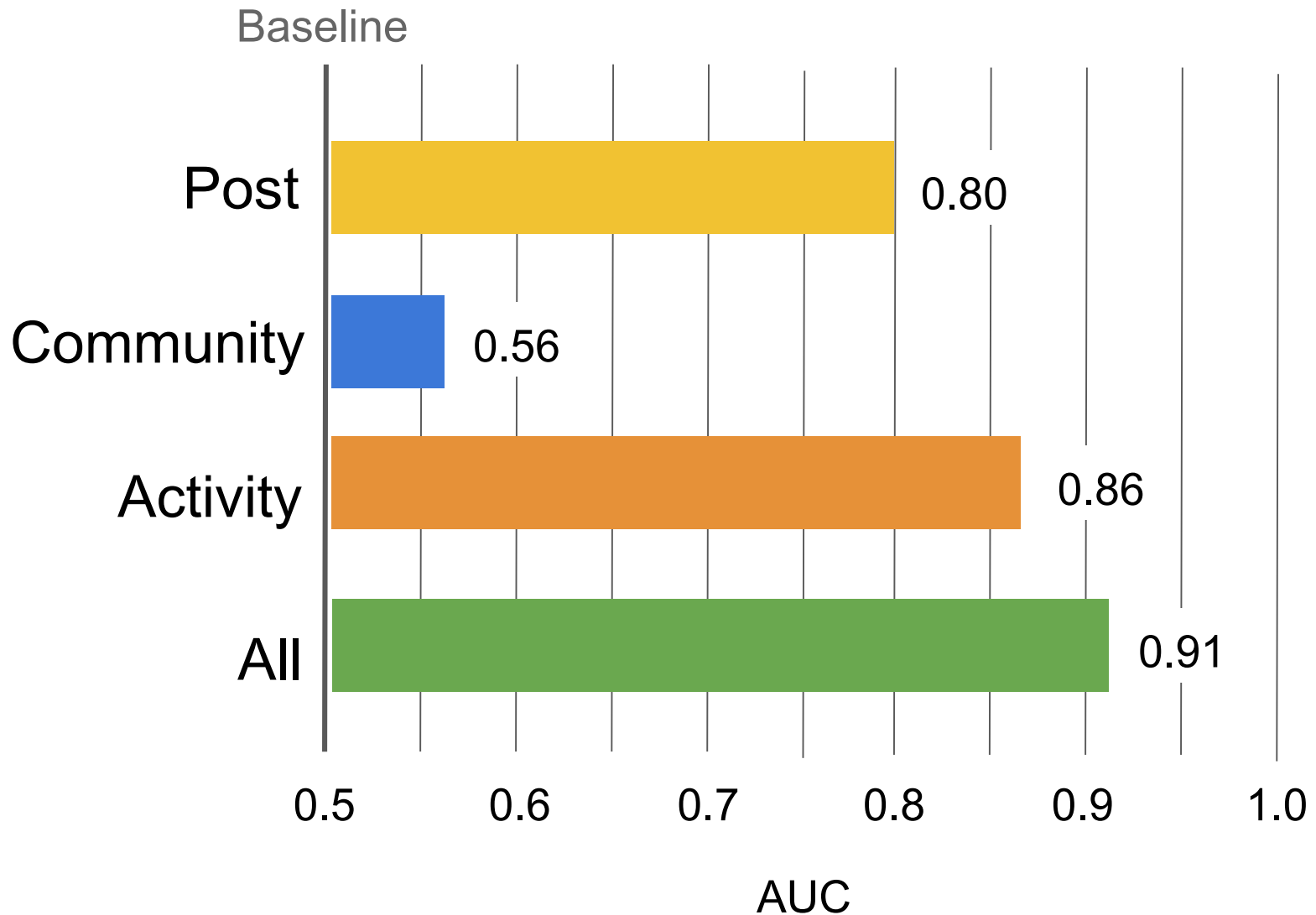
Online discussions



Do two accounts belong to the same person?



Online discussions



Summary: Sockpuppets

- **Sockpuppets:** Usage of multiple accounts, both for benign and malicious intent
- Sockpuppets write worse than non-sockpuppets
- Sockpuppet accounts help each other
- Sockpuppets can vary in how deceptive they are and how supportive they are
- Sockpuppets can be detected from what they post and how they post, but not efficiently from community feedback

References

Kumar, S., Cheng, J., Leskovec, J. & Subrahmanian, V.S. (2017). An Army of Me: Sockpuppets in Online Discussion Communities. Proceedings of the 26th International Conference Companion on World Wide Web. WWW 2017. <http://bit.ly/sockpaper>

Solorio, T., Hasan, R. & Mizan, M. (2013). A case study of sockpuppet detection in Wikipedia. LASM 2013.

Zheng, X., Lai, Y.M., Chow, K., CK, Lucas & Yiu, S.M. (2011). 2011 Seventh International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)

Bu, Z., Xia, Z. & Wang, J. (2013). A sock puppet detection algorithm on virtual spaces. Knowledge-Based Systems 2013

Liu, D., Wu, Q., Han, W. & Zhou, B. Sockpuppet gang detection on social media sites. Frontiers of Computer Science

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

References

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

A. Caspi and P. Gorsky. Online deception: Prevalence, motivation, and emotion. *CyberPsychology & Behavior*, 2006.

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.

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

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*. 2016.

End of Part 1