Inferred Profiles: Examining How People Understand and Control What Algorithms Infer about Them

Isaac Johnson
Northwestern University
Evanston, IL 60208, USA
isaacj@u.northwestern.edu

Brent Hecht
Northwestern University
Evanston, IL 60208, USA
bhecht@northwestern.edu

Abstract
In privacy research, an important and understudied audience of social media users are scientists and entities conducting social media surveillance. These groups often collect large amounts of public data to run through various algorithms to build *inferred profiles* for individuals – i.e. information about the person that they did not explicitly share. These inferred profiles are used for tasks ranging from conducting social science research to tracking protesters. Yet, there is little research on how people understand and can best control these inferred profiles given that there is often no direct benefit to the user of these inferences and sometimes quite negative consequences. In this position paper, we 1) motivate and outline important research questions regarding how people understand what can be algorithmically inferred about them, and, 2) discuss the design of tools to support more educated decisions about what people share online in the context of inference algorithms.

Author Keywords
Privacy; Algorithms; Public Data; Awareness.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous;

Introduction

With the rise of social media and greater sharing of personal content on the web, there has been much research on how people envision and manage privacy concerns online. For instance, researchers have studied what leads to individuals sharing information that they later regret [17], how well people estimate their online audience [3], and attempts to conform to norms as people navigate the complex ecosystem of social networks [19]. This research has largely defined a user's audience as the other people who are viewing her/his photos and posts. When research addresses how individuals think about other audiences when engaging with social networks – e.g. the social network companies themselves – it has largely focused on individuals' willingness to exchange information for some perceived direct benefit (e.g. privacy-personalization paradox [2]).

However, this past work largely ignores another growing audience of social media activity and content: entities engaged in social media surveillance and quantitative researchers whose goal is to leverage a user's public social media information to build an inferred profile about the user. For instance, the company GeoFeedia [13] tracked Black Lives Matter activists on sites like Facebook, Twitter, and Instagram and provided this data to law enforcement agencies. Similarly, social media users in China have to find creative ways to post human-interpretable messages that allow them to discuss sensitive subjects while not triggering government detection or censorship [8]. Inversely, during crises, people broadcasting information and calls for assistance often want to be detected and correctly categorized by algorithms [16]. The Obama administration gathered input about key policy concerns from Twitter [6] and other firms continue to analyze this public discourse (e.g. [18]). Within the literature, researchers have used public social media posts to build, refine, and apply algorithms that can effectively infer user characteristics such as gender [4], location [10], political affiliation [12], and interests [11].

Researchers and entities engaged in social media surveillance differ from the audiences typically considered in privacy work in that they are neither traditional social network contacts nor a company seeking to improve the user's experience. Instead, these audiences use algorithms that collect an individual's public data and draw inferences *for the researcher or entity's own use and benefit*.

We believe that research is needed to understand social media privacy in the context of inference algorithms. Based on prior work investigating folk theories of algorithms [5,15], it is unlikely that users are aware of the inferential power of these algorithms. A user who follows Black Lives Matter activists though does not post related content may not recognize that these network connections alone are often used in inferring political leanings [12]. Likewise, indicating interest in seemingly innocuous pages like "Thunderstorms" or "Hello Kitty" has been used for predicting intelligence and emotional stability, respectively [11]. An individual's inferred profile is often largely based on what is known about related individuals and, in turn, each person's actions affect what is inferred about others. This networked nature of algorithmic inference greatly complicates an individual's ability to understand the algorithmic implications of their actions.

In this position paper, we propose that this research should proceed in two threads:

- **Descriptive Research**: Study how people understand their "inferred profile" and how to best support individuals who wish to control what is and is not inferred.
- **Tool-building Research**: Build out the tools to give individuals access and a better understanding of what can be inferred about them. These tools should be action-oriented, helping users understand the implications of posting specific content or engaging in other activity on a social media site.

Descriptive Research

Much as Hamilton et al. [7] outlined key research questions around algorithmic interface awareness and understanding individuals' folk theories about these interfaces, we wish to propose several key research questions around algorithmic inference awareness:

- What folk theories do people have about the types of information algorithms can learn from them? How accurate is their understanding and how do they act on their theories?
- What types of information do people want algorithms to infer and not infer about them and in what contexts?
- How can people best support or protect themselves against these inferences?
- How can we provide individuals with actionable feedback on their social media presence so that they can effectively adjust if they so choose?

While research on the privacy-personalization paradox (e.g., [2]) has explored how people feel about personalization and the inferences that are made about them in that space, little work has explored how individuals feel about the inferences that are made by outside groups and that do not have direct benefits such personalized experiences. How individuals navigate privacy in this context will likely be variable and domain-specific (e.g., thinking about health data differently than political affiliations), so research will have to proceed carefully and be careful to not gloss over domain and individual differences.

Like many in the HCI research community, we have conducted research on the potential benefits of social-media-based inference. For instance, we have explored how to identify social media users local to an area so that researchers might more appropriately study phenomena such as societal happiness [9]. This work has also raised concerns about the negative uses of algorithms to track and make inferences about individuals through their online profiles. We believe that there is great opportunity to bring together those of us working in the inference algorithms community with those in the privacy community to guide research that bridges this space.

Tool-Building Research

Whereas access to privacy settings on social networks is largely improving, attempts to give an individual control over their "inferred profile" have been limited. For example, ProPublica recently explored what Facebook can infer about its users. The researchers provided a Chrome extension that directs users to a page provided by Facebook that lists some of the inferences that Facebook draws internally but noted

that Facebook exposes very little of the data that they actually have about individuals [1]. Kulshrestha et al. [12], work appearing in CSCW 2017, explore political inference and bias on Twitter and laudably co-released a platform that exposes what their algorithm would infer about a user's political leanings based on their Twitter network. Petkos et al. [14] describe a platform that they are building to provide feedback to a user about how much they are sharing and what could be inferred about them across several categories.

Social media users (and researchers) still lack the tools to broadly surface and interrogate these inferred profiles. Petkos et al. discuss surfacing the data behind any inferences that can be made, but it is also important to expose in advance how new posts or actions might impact the confidence and conclusion of these inferences because APIs allow surveillance entities to stream data in real-time.

Notably, we are aware that this is delicate work because providing inaccurate or underpowered algorithms could mislead individuals into believing that their profiles are inscrutable to algorithms. We are excited to contribute to the design of inference awareness and action tools, and we welcome feedback about the design of such a platform from others in the privacy community who have explored how to best support privacy awareness and action.

Conclusion

In this position paper, we advocated for the development of tools and research into how people understand and wish to control the inferred profiles built by researchers or entities conducting social media surveillance. We believe this to be an important and understudied aspect of privacy. As the HCI community has been involved in designing some of these inference algorithms, it is appropriate that we also seek to communicate the risks and benefits in an actionable and accessible manner. Given our background with inference algorithms, we consider ourselves well-situated to undertake some of this work. We are excited to gather input from others in the privacy community to better understand how to balance the importance of privacy in the face of increasingly powerful inference algorithms with the benefits of open communication.

References

1. Julie Angwin, Terry Parris Jr., and Surya Mattu. 2016. What Facebook Knows About You. *ProPublica*. Retrieved from https://www.propublica.org/article/breaking-the-black-box-what-facebook-knows-about-you
2. Naveen Farag Awad and Mayuram S. Krishnan. 2006. The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization. *MIS quarterly*: 13–28.
3. Michael S. Bernstein, Eytan Bakshy, Moira Burke, and Brian Karrer. 2013. Quantifying the invisible audience in social networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 21–30.
4. Xin Chen, Yu Wang, Eugene Agichtein, and Fusheng Wang. 2015. A Comparative Study of Demographic Attribute Inference in Twitter. In *ICWSM*.
5. Motahhare Eslami, Aimee Rickman, Kristen Vaccaro, Amirhossein Aleyasen, Andy Vuong, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig.

2015. "I always assumed that I wasn't really that close to [her]": Reasoning about Invisible Algorithms in News Feeds. In *CHI*, 153–162.

6. Klint Finley. 2013. Out in the Open: You Too Can Use Obama's Secret Social Media Weapon. *Wired*. Retrieved from https://www.wired.com/2013/10/thinkup/
7. Kevin Hamilton, Karrie Karahalios, Christian Sandvig, and Motahhare Eslami. 2014. A path to understanding the effects of algorithm awareness. In *CHI*, 631–642.
8. Chaya Hiruncharoenvate, Zhiyuan Lin, and Eric Gilbert. 2015. Algorithmically Bypassing Censorship on Sina Weibo with Nondeterministic Homophone Substitutions. In *Ninth International AAAI Conference on Web and Social Media*.
9. Isaac Johnson, Subhasree Sengupta, Johannes Schöning, and Brent Hecht. 2016. The Geography and Importance of Localness in Geotagged Social Media. *CHI*.
10. David Jurgens. 2013. That's What Friends Are For: Inferring Location in Online Social Media Platforms Based on Social Relationships. *ICWSM* 13: 273–282.
11. M. Kosinski, D. Stillwell, and T. Graepel. 2013. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences* 110, 15: 5802–5805.
12. Juhi Kulshrestha, Motahhare Eslami, Johnnatan Messias, Muhammad Bilal Zafar, Saptarshi Ghosh, IIEST Shibpur, India Krishna P. Gummadi, and Karrie Karahalios. 2017. Quantifying Search Bias: Investigating Sources of Bias for Political Searches in Social Media. *CSCW*.
13. Kalev Leetaru. 2016. Geofeedia Is Just The Tip Of The Iceberg: The Era Of Social Surveillence. *Forbes*. Retrieved from http://www.forbes.com/sites/kalevleetaru/2016/10/12/geofeedia-is-just-the-tip-of-the-iceberg-the-era-of-social-surveillence/
14. Georgios Petkos, Symeon Papadopoulos, and Yiannis Kompatsiaris. 2015. PScore: A Framework for Enhancing Privacy Awareness in Online Social Networks. 592–600.
15. Emilee Rader and Rebecca Gray. 2015. Understanding User Beliefs About Algorithmic Curation in the Facebook News Feed. In *CHI*, 173–182.
16. Kate Starbird and Jeannie Stamberger. 2010. Tweak the tweet: Leveraging microblogging proliferation with a prescriptive syntax to support citizen reporting. *ISCRAM*.
17. Yang Wang, Gregory Norcie, Saranga Komanduri, Alessandro Acquisti, Pedro Giovanni Leon, and Lorrie Faith Cranor. 2011. "I Regretted the Minute I Pressed Share": A Qualitative Study of Regrets on Facebook. In *SOUPS*, 10:1–10:16.
18. Heather Whaling. 2017. Why Women (and Men) Are Marching Today, According to Twitter Data. *Wired*. Retrieved from https://www.wired.com/2017/01/women-men-marching-today-according-twitter-data/
19. Xuan Zhao, Cliff Lampe, and Nicole B Ellison. 2016. The Social Media Ecology: User Perceptions, Strategies and Challenges. In *CHI*, 89–100.