"Hey Players, there is a problem...": On Attribute Inference Attacks against Videogamers

(Supplementary Document)

Linus Eisele, Giovanni Apruzzese
Liechtenstein Business School, University of Liechtenstein
{name.surname}@uni.li

Abstract—We focus on a subtle privacy issue that affects (potentially hundreds of) millions of videogamers: attribute inference attacks (AIA). Through AIA, evildoers can infer gamers' *private* attributes (e.g., age, gender, occupation) by leveraging in-game statistics that are *publicly available*. Despite some research efforts revealing the practicality of AIA in Dota2, such a threat has been mostly ignored by the overarching gaming community. This is a problem: AIA can only be mitigated through the cooperation of the entire gaming community—and this cooperation can only begin if all stakeholders *acknowledge* the threat of AIA.

We seek to promote such a positive change by raising the gaming ecosystem's awareness about AIA. First, we provide evidence that AIA have truly been overlooked in the gaming domain. Then, we scrutinize the gaming landscape, pinpointing (i) the games that are more prone to AIA, and (ii) the risk that some communities "inadvertedly" contribute to AIA. Finally, through an (ethical) user survey (n=516) resembling a fundamental data-collection step of AIA, we (iii) proactively assess the threat of AIA. We advocate gamers and developers to reflect upon our findings—which we disseminate in an educational campaign: the subtle threat of AIA cannot be countered solely by researchers.

> **Purpose:** This supplementary document extends our main paper by providing additional descriptions and analyses.

APPENDIX A
LITERATURE REVIEW

We describe the literature review in more detail (this Appendix is meant to extend §2 of the main paper).

A. IEEE CoG

We explain how we analysed the IEEE CoG papers, for which we adopt a systematic approach organized in **three phases**: paper collection, keyword inspection, and manual analysis.

- Firstly, we downloaded all the proceedings from IEEE Xplore, considering all editions of IEEE CoG, i.e., 2019. 2020, 2021, 2022, 2023. We consider *all* papers, including short papers—amounting to 682 papers.
- Secondly, we write a custom script which iterates over each paper, searching for terms "AIA", "attribute inference attack" and "privacy". We could not find any match for the first two, whereas for "privacy", we found 20 papers that had one match.
- Thirdly, we review all these 20 papers, checking *where* the term "privacy" occurs. For seven papers, it only occurs in references; for one paper in the appendix. For

six papers it is used only to describe the data collection procedure. For five papers, it is mentioned in a context that has nothing to do with the privacy of videogamers.

Explanation. Let us elaborate the last statement. In [1], it is used once in references, and once to refer to the privacy of citizens when considering using gamification approaches for driving. In [2], it used once in references, and once to state that "the browser's privacy settings may change our results". In [3] it is used once to refer to "privacy leaks in crowdsourcing" for blockchain applications. In [4] is is used three times to refer to companies which are "unwilling to release data due to trade secrets and privacy [...] when fraud may be contained in the data". In [5] it is used extensively, but because the paper analyses "how *bathrooms* in games convey the sense of 'privacy' expected by real-world ones." In summary, the only work which specifically used the term "privacy" in the context of "protecting videogamers' privacy" is [6], for which it occurs 4 times in the ethical section (and 3 times in the references). We visualize these results in Table I.

TABLE I: **Papers mentioning "privacy" in IEEE CoG –** We download all 682 papers which appeared in the IEEE CoG proceedings from 2019–2023, and search for occurrences (Occ) of the term "privacy". The "Location" column reports "where" in the each paper the term occurs: OOC=out of context; DC=data collection; Ref=References. Only [6] fits our vision.

Paper	Year	Occ	Location
Min et al. [3]	2019	1	OOC
Eidelberg et al. [7]	2019	1	DC
Parakh et al. [8]	2019	1	Ref
Santos et al. [9]	2019	1	DC
Christiansen et al. [10]	2019	1	Ref
Bjornsson et al. [11]	2021	1	Ref
Lima et al. [12]	2021	1	DC
Zeng et al. [13]	2022	1	Ref
Audibert et al. [14]	2023	1	DC
Di Paolo et al. [15]	2023	1	Ref
Gerster et al. [16]	2023	1	Appendix
Van et al. [2]	2019	2	Refs, OOC
McQuighan et al. [17]	2019	2	Refs
Melhart et al. [18]	2019	2	Refs
Massoud et al. [1]	2019	2	Refs, OOC
Tao et al. [19]	2020	2	DC
Tritscher et al. [4]	2021	3	OOC
Ferguson et al. [20]	2021	4	DC
Bowey et al. [6]	2021	7	Refs, Ethic
Antognoli et al. [5]	2021	33	OOC

B. Google Scholar

We follow a systematic approach for our literature review on Google Scholar. Recall that our goal is assessing the extent to which the theme of privacy is accounted for in papers that *(i)* deal with video games and *(ii)* discuss correlations/predictions/relationships between in-/off-game player data which *(iii)* could be maliciously exploited via publicly available tools (e.g., tracking websites). Let us explain how we do this.

- First, we frame the scope and boundaries, relevant keywords, and selection criteria for our search. Specific terms we looked for are combinations of: "(videogames, gaming, games, MMO, gamers, players) ∧ (attribute inference attacks, AIA, privacy, security, issues, violation, correlation, association, survey, profiling, prediction)", which we provided as input to Google Scholar.
- Then, we broadly analyze the papers returned by Google Scholar, starting from the output of the query. We filtered out results that were clearly outside our scope. Then we read at least the abstracts, introductions, and conclusions, and eliminate candidates that are not directly relevant to our goal. Whenever we found a paper we deemed appropriate, we use the snowballing method [21] and search for related works.
- Finally, we deeply review all relevant works we found. We specifically look for occurrences of the term "privacy" in the paper, scrutinizing the context in which it was used (if it even appeared).

To give some examples of some "excluded" papers, Martinovic et al. [22] mention the theme privacy, but such a privacy risk only existed for data collected by the developers of the game (e.g., keystrokes), which are not publicly available. Kroger et al. [23] do raise the attention on privacy issues pertaining to data-collection practices in games, but do not carry out any original user study through which novel correlations/associations are brought to light. Of the remaining papers (e.g., [24]), the term privacy hardly occurs, or the potential privacy risks related to the main contribution of the paper (i.e., the discovery of new correlations) is not emphasized. To note: this reviewing procedure is carried out by two authors who frequently interacted to discuss their findings and resolve discrepancies in their own judgment.

APPENDIX B
INVESTIGATION: COMMUNITIES

Let us provide more details on our community analysis. First, **we look for communities that have a large userbase** (i.e., > 1000 recent users). For instance, we exclude r/gaminghardware and /r/Gamingthoughts because they have barely 100 users. We set ourselves the target of finding 50 "large and active" communities, and whose userbase has interest in our considered 14 games; this requirement can be either explicit (e.g., name of the game in the title of the community, such as /r/wow/) or we derive with a qualitative analysis (e.g., for r/truegaming, there are many posts on our games). **We then inspect the guidelines**. Four cases can happen:

✓ posting a survey is generally allowed without approval (possibly under some conditions / properties that the survey must fulfill);

✗ posting a survey is explicitly not allowed; we will not proceed further.

? policies do not mention anything about surveys at all and give no clues as to whether they are (dis)allowed

☑ mod approval is required, or the policies make it not fully clear whether surveys can be posted (e.g., they may mention that certain types of content are not allowed)

As a general rule, if we had strong reason to believe that a survey was *not allowed*, we chose ☑ and waited for a response before deciding to post. In many cases, we sent follow-up emails to solicit a response. Then **if surveys are allowed, we will post it; if not, we will not post it;** otherwise we either:

- **contact the mod**. In which case, we will either receive a positive response (👍) and we will post the survey; or we will get a negative response (👎). If we do not receive a response after 14 days 🕐, we will post the survey if nothing was specified before (i.e., **?**), and we will not post the survey if mod approval was required (i.e., ☑). Or:
- if there is no explicit person to contact for such inquiries, we will post the survey right away.

After we post the survey, we wait. If it gets taken down (⚡) then it meant that surveys were not allowed (and we will delete any response we collected, if any). Otherwise, we will keep collecting responses. We show the complete categorization of the procedure above in Table II (which extends Table II of the main paper).

REFERENCES

[1] R. Massoud, F. Bellotti, R. Berta, A. De Gloria, and S. Poslad, "Eco-driving profiling and behavioral shifts using iot vehicular sensors combined with serious games," in *IEEE CoG*, 2019.

[2] W. van der Vegt, W. Westera, H. Kurvers, and E. Nyamsuren, "Portability of serious game software components," in *IEEE CoG*, 2019.

[3] T. Min, H. Wang, Y. Guo, and W. Cai, "Blockchain games: A survey," in *IEEE CoG*, 2019.

[4] J. Tritscher, A. Krause, D. Schlör, F. Gwinner, S. Von Mammen, and A. Hotho, "A financial game with opportunities for fraud," in *IEEE CoG*, 2021.

[5] D. Antognoli and J. Fisher, "The purposes and meanings of video game bathrooms," in *IEEE CoG*, 2021.

[6] J. T. Bowey, J. Frommel, B. Piller, and R. L. Mandryk, "Predicting beliefs from npc dialogues," in *IEEE CoG*, 2021.

[7] A. Eidelberg, C. Jacob, and J. Denzinger, "Using active probing by a game management ai to faster classify players," in *IEEE CoG*, 2019.

[8] A. Parakh, P. Chundi, and M. Subramaniam, "An approach towards designing problem networks in serious games," in *IEEE CoG*, 2019.

[9] C. P. Santos, N. C. M. F. van Gaans, V.-J. Khan, and P. Markopoulos, "Effects of advertisements and questionnaire interruptions on the player experience," in *IEEE CoG*, 2019.

[10] A. H. Christiansen, E. Gensby, and B. S. Weber, "Resolving simultaneity bias: using features to estimate causal effects in competitive games," in *IEEE CoG*, 2019.

[11] Y. Björnsson, S. Helgason, and A. Pálsson, "Searching for explainable solutions in sudoku," in *IEEE CoG*, 2021.

[12] L. Lima, P. Gouveia, P. Cardoso, and C. Pinto, "Never imagined i would work in the digital game industry," in *IEEE CoG*, 2021.

[13] F. Zeng, G. Xing, and G. Han, "Memory-augmented episodic value network," in *IEEE CoG*, 2022.

[14] J. Audibert, E. Gonzalez, R. Orlando, N. Wong, E. Agu, and M. Claypool, "Machine learning prediction of just dance exergame enjoyment from mobile sensor data," in *IEEE CoG*, 2023.

TABLE II: **Community Analysis –** We report the 50 communities we have analysed (having players of the 14 games boldfaced in Table I of the main paper). The first 23 communities are those that, after our analysis (see §IV-B of our main paper) can be classified as at risk of (inadvertently) contributing to AIA—due to either allowing surveys by default (red), or after messaging the mod (blue), or which did not remove our survey after we posted it (gray cells do not have specific mods to contact; yellow cells have mods, but did not respond to our request after 14 days). Communities with a white background (27) have not been considered since surveys were not accepted. ⚡ is when we posted the survey, but it was removed. 🕐 is when we did not receive any response after 14 days. ✓ is when surveys are explicitly allowed; ✗ when they are disallowed; 📝 is when mod approval is required, and ? when it is unclear. 👍 is when we received approval by the mods, and 👎 when we did not receive permission to post the survey.

Community (Large and Active)	Survey Allowed?	Admin Response?	Msgs Sent
truegaming	✓		
SampleSize	✓		
JoyFreak	✓		
Rainbow6	📝	👍	1
wow	📝	👍	1
leagueoflegends	📝	👍	6
Overwatch	📝	👍	4
VALORANT	📝	👍	3
youtubegaming	📝	👍	1
GameTheorists	?	🕐	1
videogames	?	🕐	1
consoles	?	🕐	1
AskGames	?	👍	1
MMORPG	?	👍	2
playstation	?	🕐	2
ubisoft	?	👍	1
Instant Gaming	?	👍	1
RocketLeague	?	👍	3
gamers	?	👍	1
PC Gamer	?	–	
COD Forums	?	–	
Valorant Forums	?	–	
GTA Forums	?	–	
xboxone	✗		
gamingsuggestions	✗		
gamingnews	✗		
esports	✗		
summonerschool	✗		
xboxseriesx	✗		
GamingLaptops	✗		
FortNiteBR	✗		
pcgaming	✗		
gaming	✗		
Games	✗		
Steam Community Hub	✗		
destiny2	📝	👎	5
gtaonline	📝	👎	4
CODWarzone	📝	👎	1
apexlegends	📝	👎	2
Steam	📝	👎	2
DestinyTheGame/	📝	👎	2
gamecollecting	📝	👎	1
Twitch	📝	👎	70
Dynamo Gaming	📝	🕐	1
GlobalOffensive	📝	🕐	4
privacy	📝	🕐	1
battlefield2042	📝	🕐	3
PUBATTLEGROUNDS	?	👎	1
Chillzone	?	🕐⚡	1
TheTechGame	?	⚡	

[15] D. DiPaola, J. Shen, R. Hu, S. Alghowinem, and C. Breazeal, "Dronescape: Designing an educational escape room for adult ai literacy," in *IEEE CoG*, 2023.

[16] R. Gerster, B. Swinkels, T. Streefkerk, T. Sjerps, K. van Melis, and R. B. Delft, "Little big data–a secondary school game raising awareness on data collection and analysis," in *IEEE CoG*, 2023.

[17] J. M. McQuighan, G. Bajwa, and J. M. Pittman, "B2ci 2019: The ieee brain to computer interface competition's gaming event," in *IEEE CoG*, 2019.

[18] D. Melhart, A. Azadvar, A. Canossa, A. Liapis, and G. N. Yannakakis, "Your gameplay says it all: Modelling motivation in tom clancy's the division," in *IEEE CoG*, 2019.

[19] J. Tao, Y. Xiong, S. Zhao, Y. Xu, J. Lin, R. Wu, and C. Fan, "Xai-driven explainable multi-view game cheating detection," in *IEEE CoG*, 2020.

[20] C. Ferguson, R. Lewis, C. Wilks, and R. Picard, "The guardians: Designing a game for long-term engagement with mental health therapy," in *IEEE CoG*, 2021.

[21] C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," in *EASE*, 2014.

[22] D. Martinovic, V. Ralevich, J. McDougall, and M. Perklin, ""you are what you play": Breaching privacy and identifying users in online gaming," in *IEEE PST*, 2014.

[23] J. L. Kröger, P. Raschke, J. Percy Campbell, and S. Ullrich, "Surveilling the gamers: Privacy impacts of the video game industry," *Entertainment Computing*, 2023.

[24] Z. Wang, A. Sapienza, A. Culotta, and E. Ferrara, "Personality and behavior in role-based online games," in *IEEE CoG*, 2019.