The 13th ACM Conference on Data and Application Security and Privacy

Attribute Inference Attacks in Online Multiplayer Video Games: A Case Study on Dota2

Pier Paolo Tricomi[1], Lisa Facciolo[1], Giovanni Apruzzese[2], Mauro Conti[1,3]

[1] University of Padova, Italy

[2] University of Liechtenstein, Liechtenstein

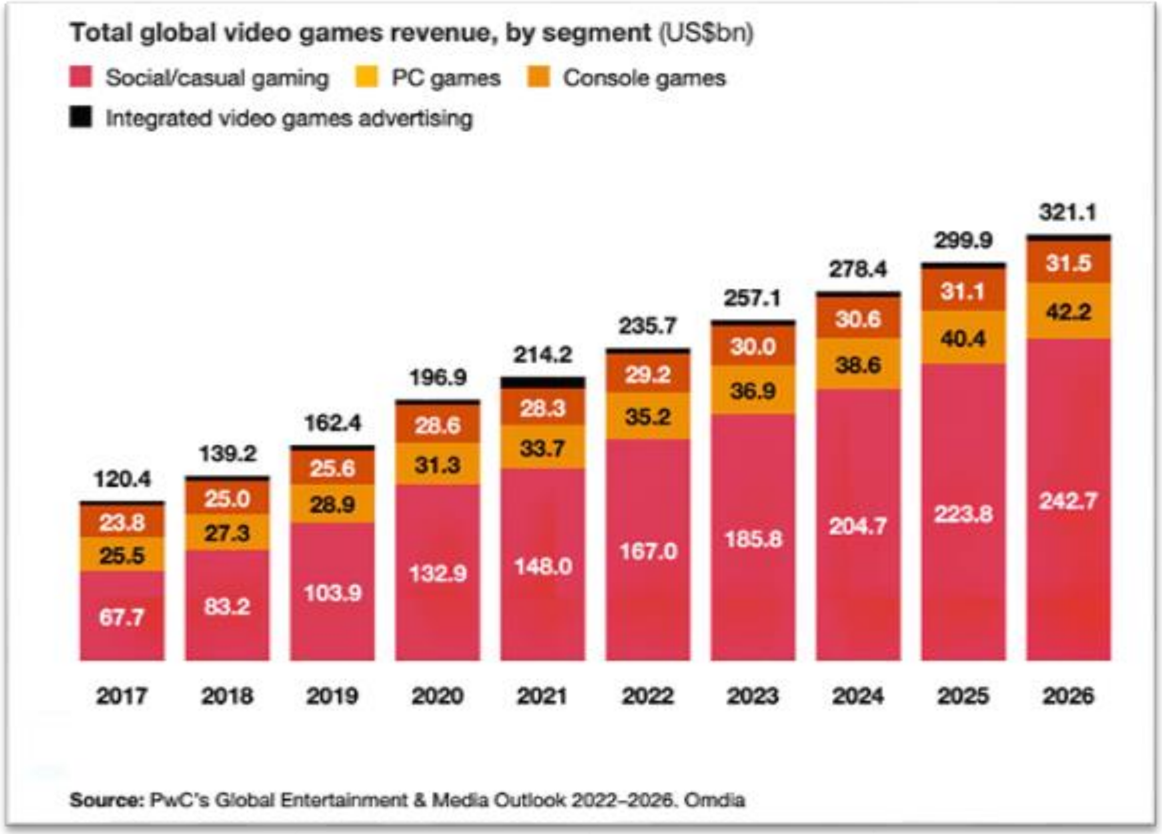
[3] Delft University of Technology, The Netherlands

April 24, 2023 - Charlotte, NC, United States

A bit of context...

Video Games (VG) are increasingly popular (3.2 Billion Gamers)

Some Competitive VG are denoted as "E-sports"



E-Sports

Some tournaments of such E-sports have very high prize-pools

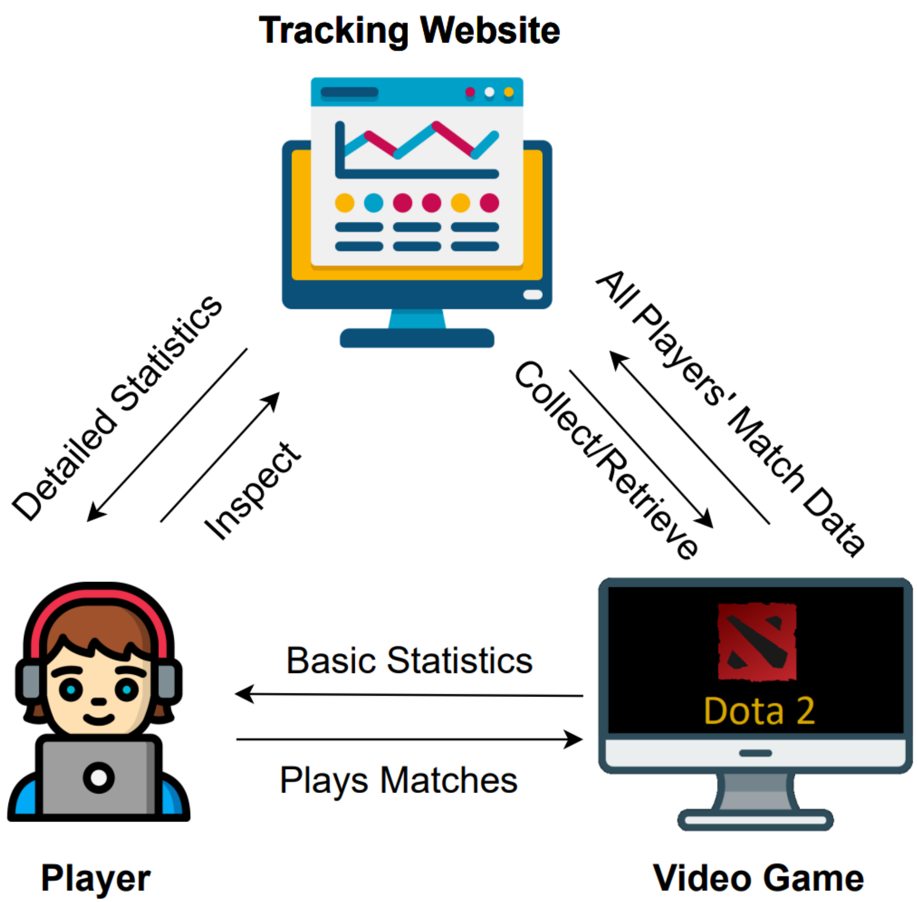
Such prizes attract a lot of players who "play-to-win" and want to get better...

Best way of improving at something? Learn from others or past mistakes!

... How?

Tracking Websites (TW)

Websites that track players' activities on Videogames, exposing statistics and learning resources



TW Examples – Player Overview

ALL OF THIS IS PUBLIC!

The STRATZ website is completely free to use

Dota 2 Matches Parsed

3·205·814·390

Player Profiles

82·712·882

STRATZ is a team of esports veterans who have come together to build the future of esports analytics. Starting with Valve Software's Dota 2, we store and parse data from every public match, and use it to create highly personalized, clear and concise interfaces for players to explore and learn from.

ALL OF THIS IS PUBLIC!

The STRATZ website is completely free to use

The player base wants TW data and statistics to be publicly available!

Reasons?

- Adapting to the trends(to win matches)
- Inspecting *other* players profiles to learn new strategies
- Gain visibility if they perform well (possibly hired by pro-teams)
- Participate/climb public ladders
- And many more!

STRATZ is a team of esports veterans who have come together to build the future of esports analytics. Starting with Valve Software's Dota 2, we store and parse data from every public match, and use it to create highly personalized, clear and concise interfaces for players to explore and learn from.

All such data is public, ok... so what?

Who cares if others know:

- How much I win...
- The hero I use the most...
- When I play my games...
- How fast I use my mouse and keyboard...
- If my playstyle is passive or aggressive...
- How I use the chat...

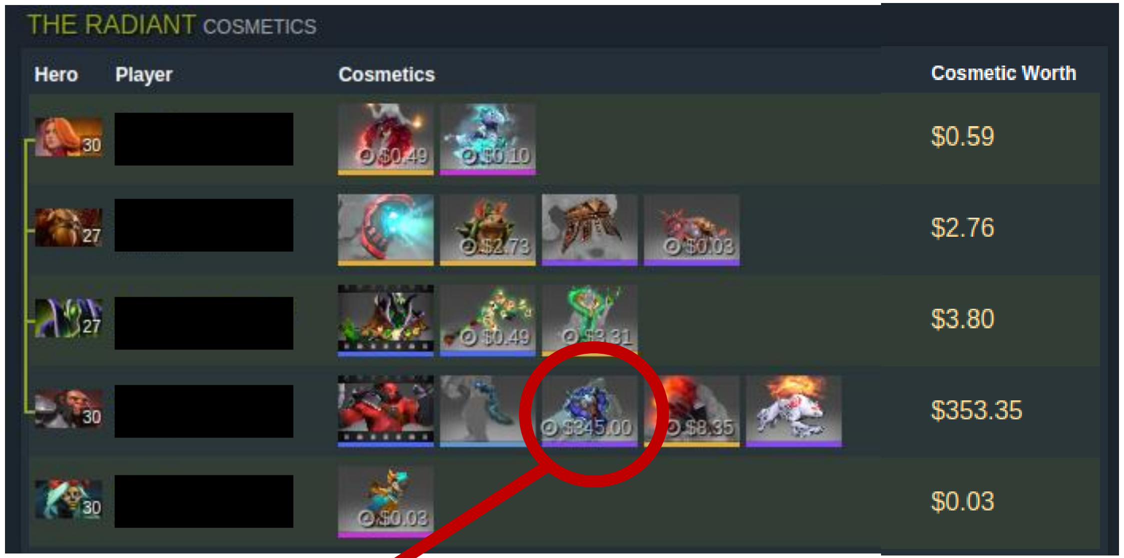
...Right?

TW Examples – Player Activity

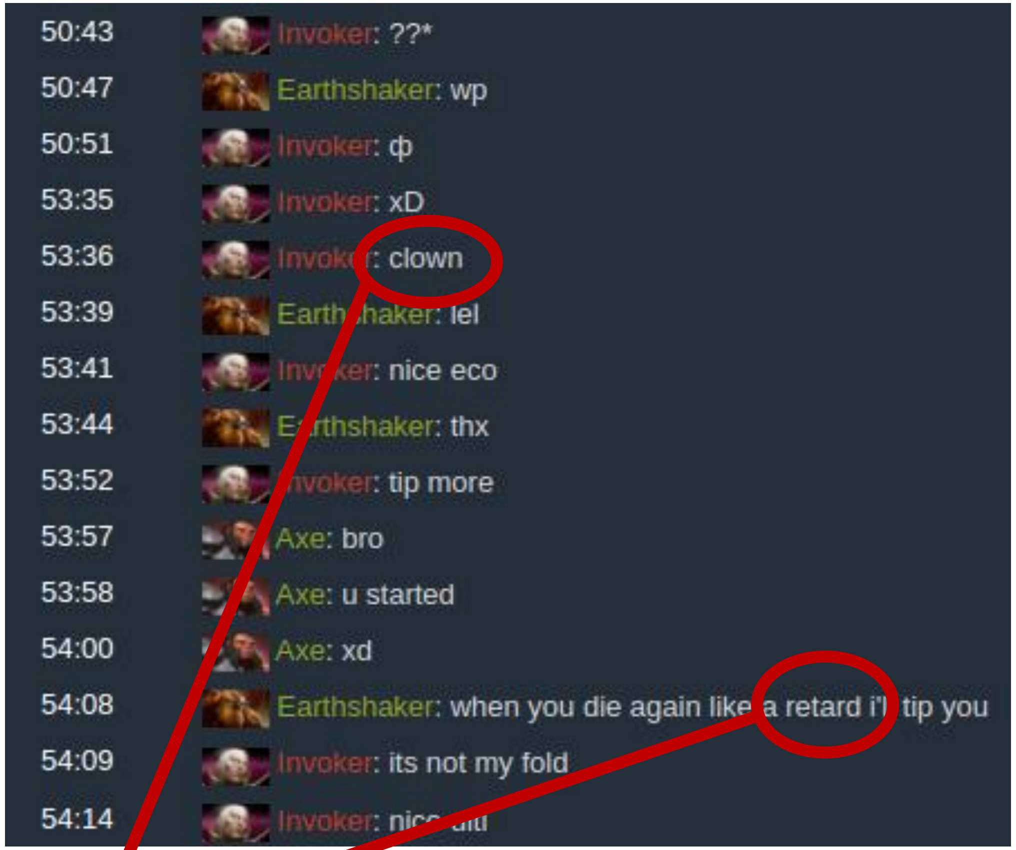
Holidays?

Evening
+
Weekend
=
Employee?

TW Examples - Cosmetics & Chat



345$ cosmetic item!!!
Rich Person?



Insults… High Neuroticism?

Players' In-game data availability exposes them to "Attribute Inference Attacks" (AIA)

($\neq$ Membership Inference Attack)

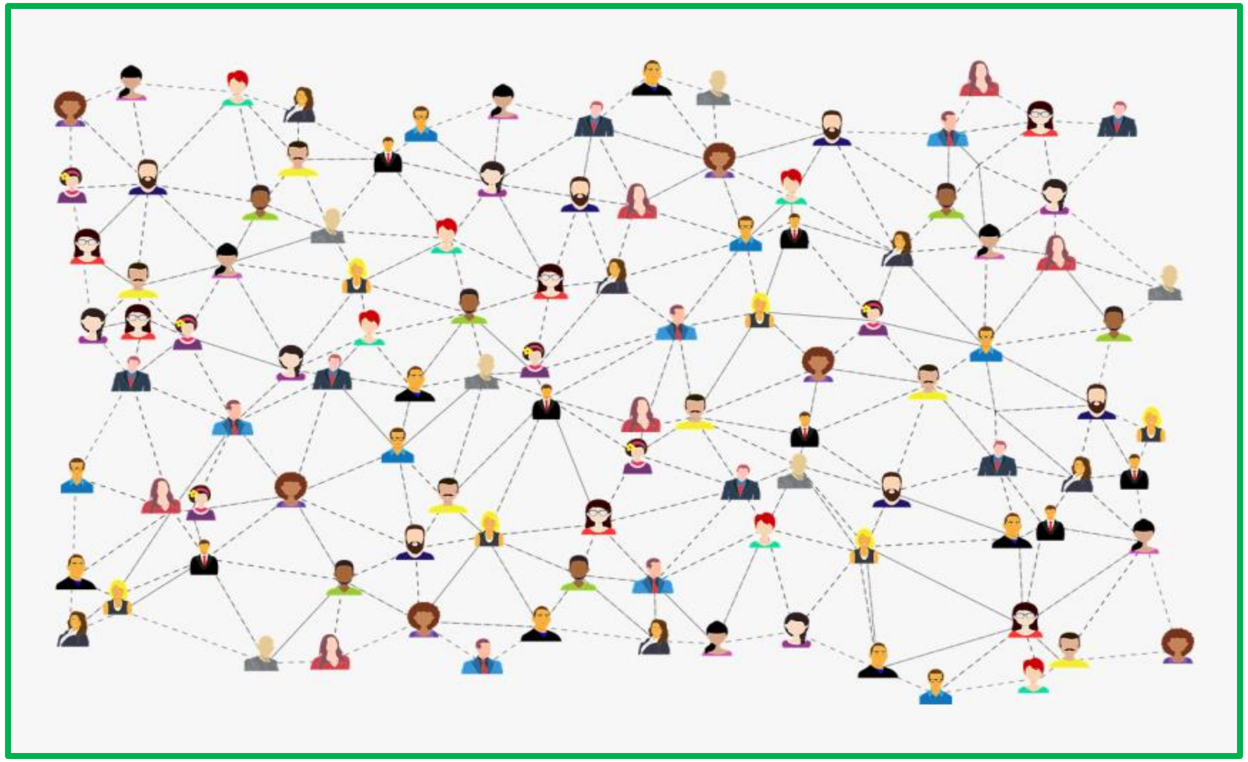
Goal: inferring private information on a given target by exploiting their publicly available data

Players' In-game data availability exposes them to "Attribute Inference Attacks" (AIA)

(≠ Membership Inference Attack)

Goal: inferring private information on a given target by exploiting their publicly available data

Public Data Available



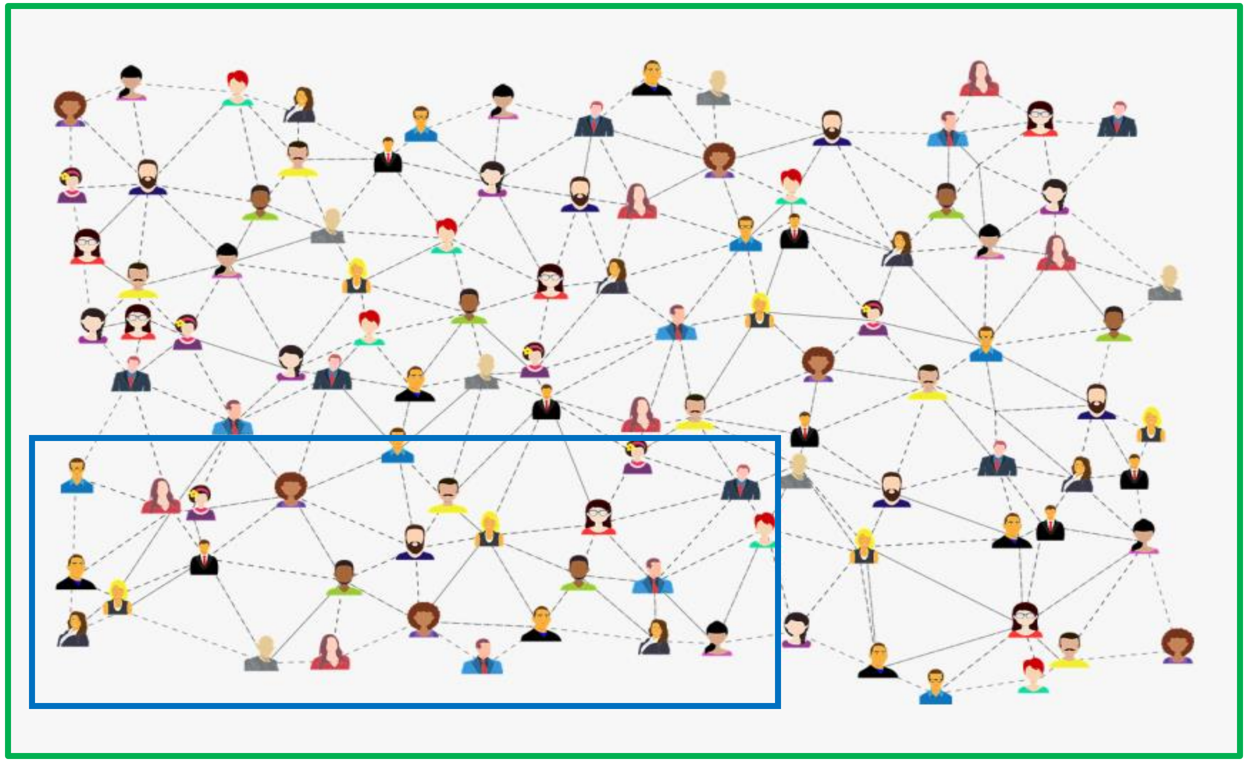
Players' In-game data availability exposes them to "Attribute Inference Attacks" (AIA)

(≠ Membership Inference Attack)

Goal: inferring private information on a given target by exploiting their publicly available data

Public Data Available

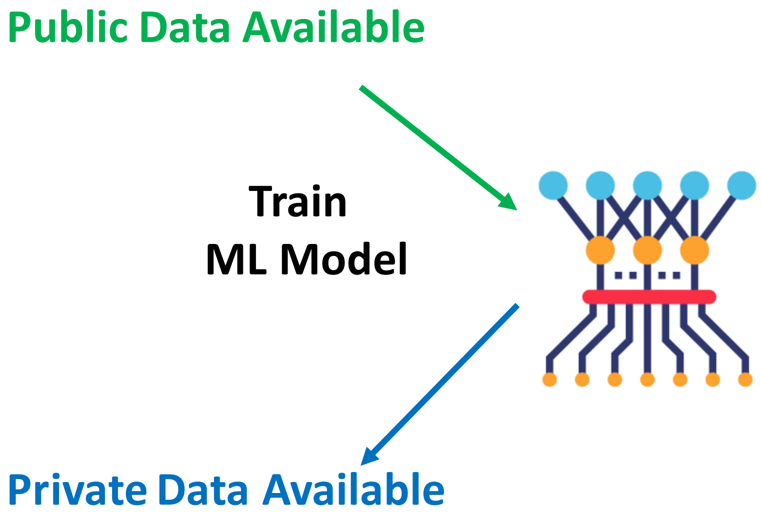
Private Data Available

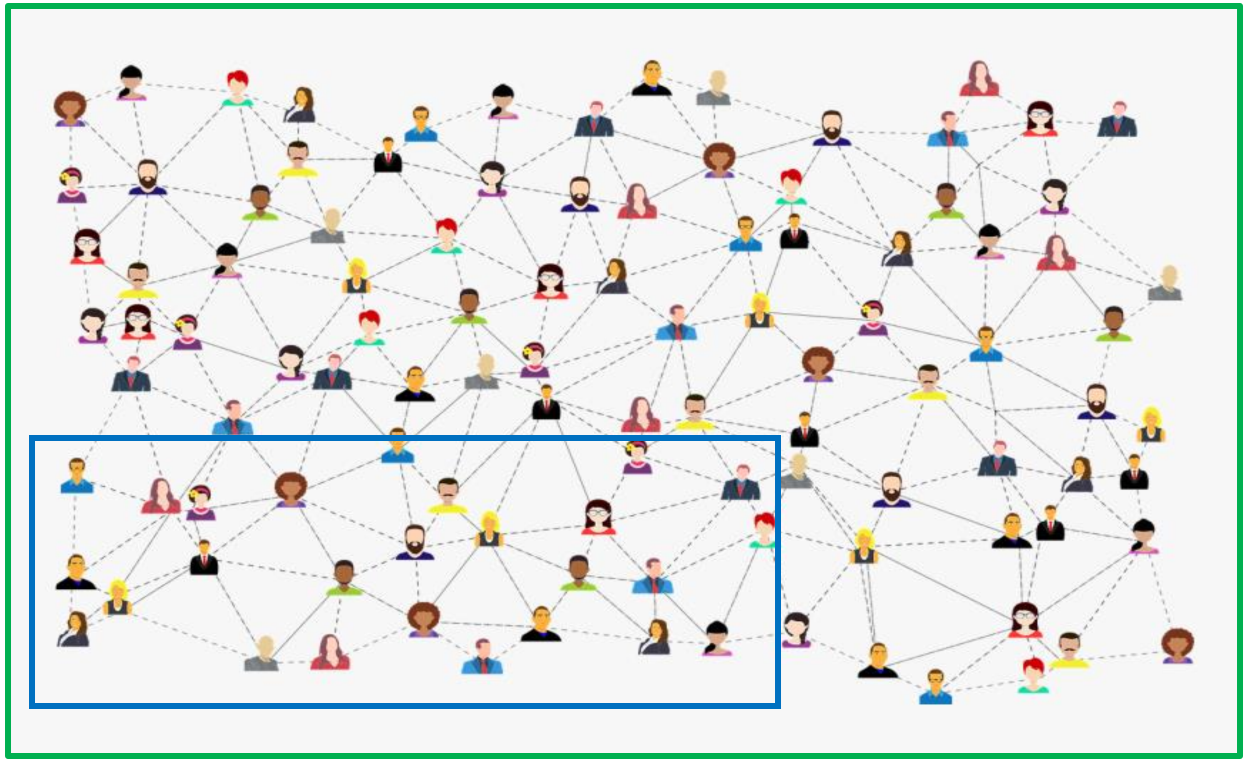


Players' In-game data availability exposes them to "Attribute Inference Attacks" (AIA)

(≠ Membership Inference Attack)

Goal: inferring private information on a given target by exploiting their publicly available data

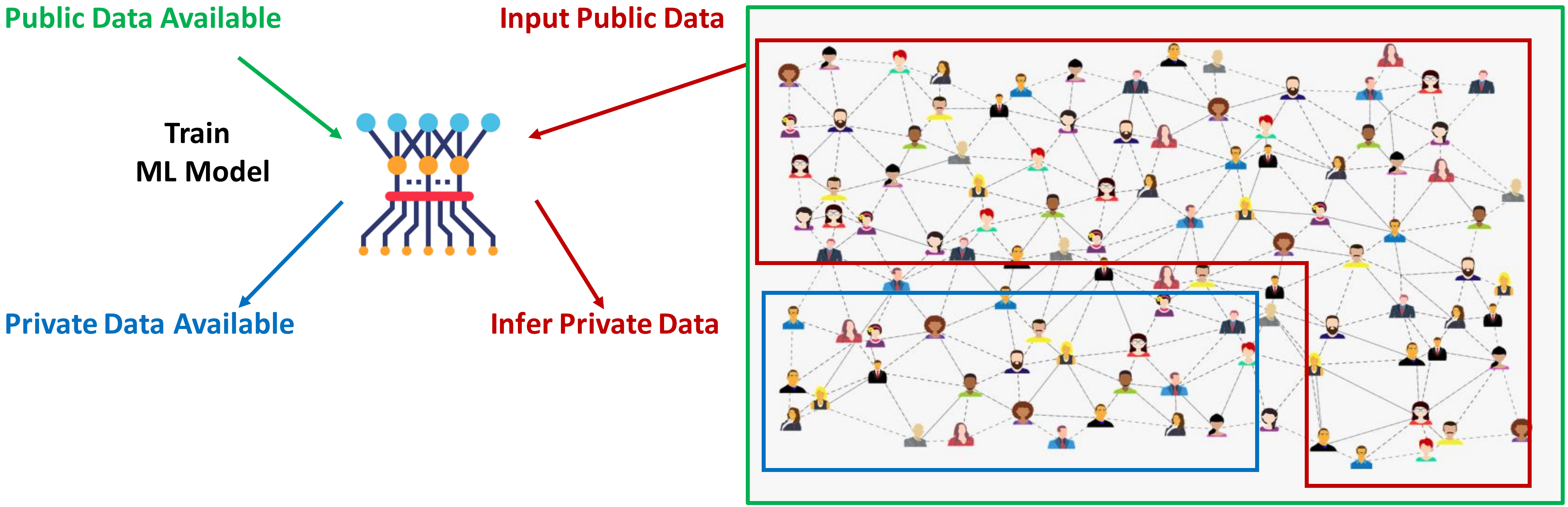




Players' In-game data availability exposes them to "Attribute Inference Attacks" (AIA)

(≠ Membership Inference Attack)

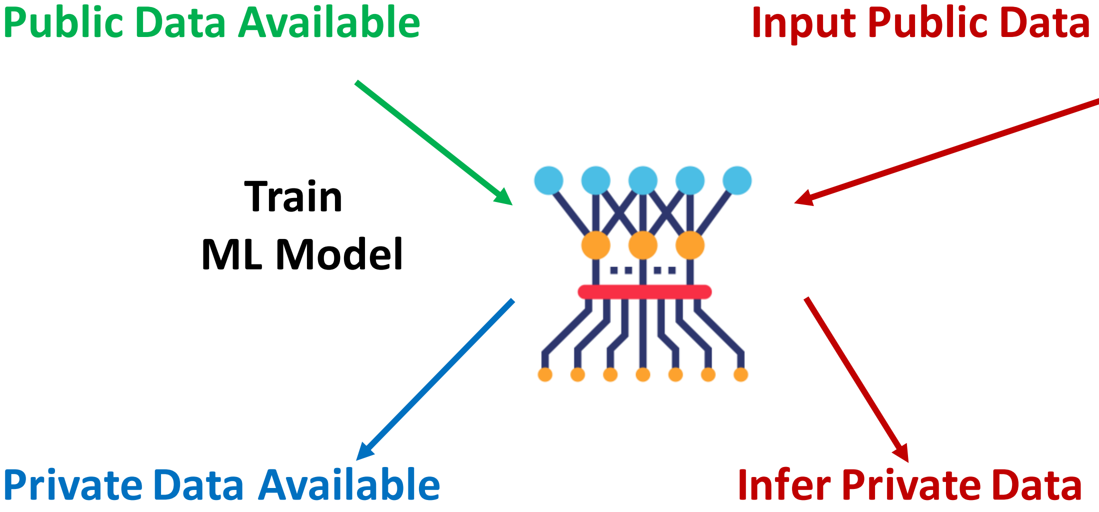
Goal: inferring private information on a given target by exploiting their publicly available data

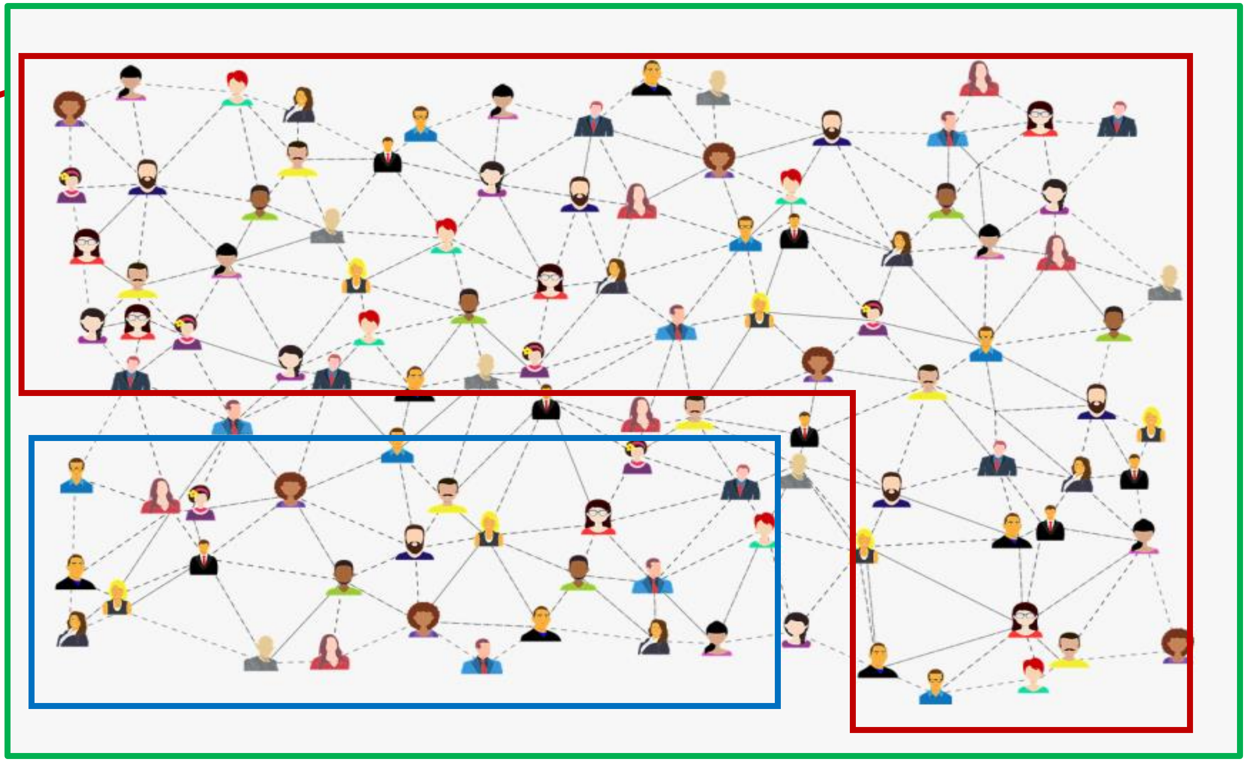


Players' In-game data availability exposes them to "Attribute Inference Attacks" (AIA)

(≠ Membership Inference Attack)

Goal: inferring private information on a given target by exploiting their publicly available data

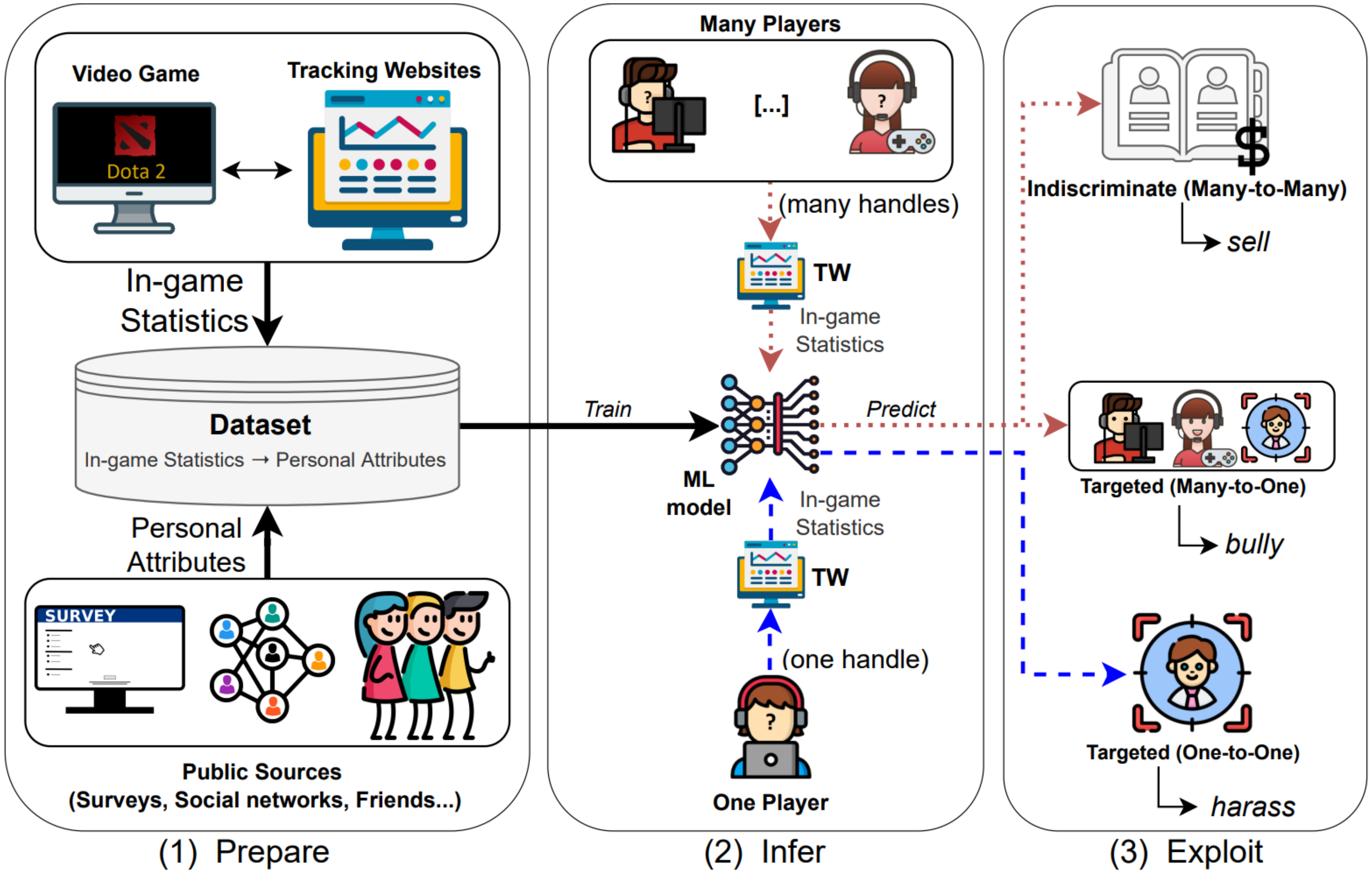




Examples (in the literature):

- Use Facebook likes to infer demographics
- Use video ratings to infer users' gender
- Use Tweets to infer users' locations

Our Proposed Threat Model



Our Assessment

- We proactively assess such a threat, because nobody ever did something similar in the E-sports ecosystem. We focus on Dota2

- We collected data of 484 Dota2 Players
  - Public in-game data from TW
  - Private data through an informed survey

- We found a correlation (!) between the players in-game statistics and their real life private info
  - Such a finding suggests that AIA can be successful!

- We (ethically) perform diverse AIA
  - Use 80% players to train ML models
  - Predict personal attributes of remaining 20% players

Data Collection

Informed survey: 625 answers from 62 different countries, **484 valid players**

In-game public data: OpenDota (free API), data from 26241 matches (one month)

Private Attributes (non-sensitive): Age, Gender, Occupation, Purchase Habits (Dota2 content), OCEAN personality

- Validated on previous survey of 29,351 Dota2 players

After preprocessing, we obtained 3 datasets:

- Player-based (P): one record per **Player's aggregated statistics** (484 samples x 187 features)
- Match-based (M): one record per **match, considering all matches** (26241 samples x 137 features)
- 'Distilled' Match-based ($\overline{M}$): same as M, but **at most 30 matches** per player (11117 samples x 160 features)
    - Reduce imbalance of games per person
    - Add features based on our domain knowledge (e.g., jargon, '?', audio messages)

Preliminary Results - Correlations

Table 8: Significant Correlations at different p-values in our three datasets. Each column reports a personal attribute in $\mathcal{A}$. Rows denote how many features in each dataset (either $\mathcal{M}$, $\overline{\mathcal{M}}$ or $\mathcal{P}$) achieve p below the target α (i.e., the correlations are statistically significant).

Dataset	*Metric*	α	gend.	age	occ.	purch.	extr.	agree.	consc.	neur.	open.
$\mathcal{M}$	Cram.	<0.01	17	17	15	18	13	18	17	16	13
	Cram.	0.05	18	19	15	18	14	19	18	19	14
	Cram.	0.1	18	19	17	19	15	19	19	19	16
	Spear.	0.01	–	88	–	51	44	52	22	70	36
	Spear.	0.05	–	95	–	65	57	59	35	85	50
	Spear.	0.1	–	99	–	73	62	67	43	87	59
$\overline{\mathcal{M}}$	Cram.	<0.01	16	12	12	11	15	10	10	14	8
	Cram.	0.05	18	17	18	15	17	11	14	15	11
	Cram.	0.1	18	17	18	15	18	14	15	20	13
	Spear.	0.01	–	95	–	43	53	38	25	60	27
	Spear.	0.05	–	104	–	63	65	54	40	82	47
	Spear.	0.1	–	108	–	69	73	64	53	90	58
$\mathcal{P}$	Cram.	<0.01	2	1	2	1	0	0	0	1	0
	Cram.	0.05	3	3	3	1	0	0	1	1	0
	Cram.	0.1	4	3	3	1	0	0	1	2	1
	Spear.	0.01	–	69	–	11	13	2	0	2	0
	Spear.	0.05	–	97	–	16	27	13	8	22	4
	Spear.	0.1	–	110	–	26	47	26	16	44	14

Examples

Age: Number of kills (-0.267)

Gender: Hero Gender (0.224)

Occupation: Paid Subscription (0.235)

Purchasing habits: Cosmetic prices (0.360)

Openness: Long-term strategy (0.103)

Conscientiousness: Fragile heroes (-0.105)

Extroversion: Chat usage (0.167)

Agreeableness: Win percentage (0.134)

Neuroticism: Denies (-0.125)

Simple AIA (Aggregated player data)

Idea: Use aggregated statistics of a player (readily available from TW) to infer private data
Method: Use the information contained in *P* to train ML models

Input: Aggregated data of many matches of player *x*

Output: Target Attribute of player *x*

Table 3: Impact of the *simple* AIA (based on $\mathcal{P}$) as measured by the F1-score. Rows report the attributes and columns our ML models (boldface denotes the best model for a given attribute).

	LR	*DT*	*RF*	*NN*	*Dummy*
gender	64.97±10.9	59.71±12.7	50.91±5.33	**67.24±13.4**	51.62±10.9
age	40.47±6.30	39.38±8.76	**44.08±6.17**	28.06±7.59	32.21±5.70
occup.	53.23±7.22	47.44±8.34	56.08±7.88	**59.89±7.15**	43.76±9.56
purch.	32.05±10.1	31.74±4.53	**34.40±8.20**	32.17±7.19	31.20±6.26
open.	28.94±5.94	**40.76±6.80**	32.6±7.77	30.89±7.60	29.59±2.04
consc.	26.52±5.65	33.87±8.78	**34.27±5.60**	23.83±8.18	33.23±8.94
extrav.	30.15±7.53	36.16±5.14	**36.49±5.56**	28.59±5.95	32.27±7.01
agreeab.	29.46±6.29	**34.11±8.58**	33.68±6.25	24.54±9.43	33.39±7.35
neurot.	32.38±6.56	**40.76±6.80**	32.6±7.74	31.6±8.30	30.07±4.46

One-Match AIA (ablation study)

Idea: A single match could be enough to infer players' private data (best/worst case scenario)
Method: Use the information contained in M or $\overline{\text{M}}$ to train ML models

Input: Single match of player *x*, using *naïve* or *expert* features

Output: Target Attribute of player *x*

Table 4: Impact of the *one-match* AIA (F1-score). Columns refer to the 'naive' attacker (using $\mathcal{M}$), 'expert' attacker (using $\overline{\mathcal{M}}$), and the Dummy (random guess). The expert attacker is always superior.

	Naive attacker (ablation study)	Expert attacker (domain knowledge)	Dummy (baseline)
gender	49.03±0.18	**58.47**±5.21	49.75±0.55
age	**43.72**±2.66	**56.82**±3.01	33.28±0.46
occup.	49.42±4.56	**68.42**±1.90	49.87±0.89
purch.	**35.61**±5.06	**49.71**±3.85	33.37±0.53
open.	32.26±3.75	**43.73**±2.96	33.48±0.41
consc.	29.49±3.63	**46.11**±3.20	32.88±0.62
extrav.	32.33±2.47	**46.82**±1.96	33.25±0.56
agreeab.	33.62±2.28	**45.36**±3.37	34.09±0.46
neurot.	27.39±4.78	**46.60**±2.72	33.65±0.58

Sophisticated AIA

Idea: Victim behavior is more likely to emerge if many of their matches are analyzed
Method: Average the predictions (post-processing) of One-Match AIA on many matches of the victim

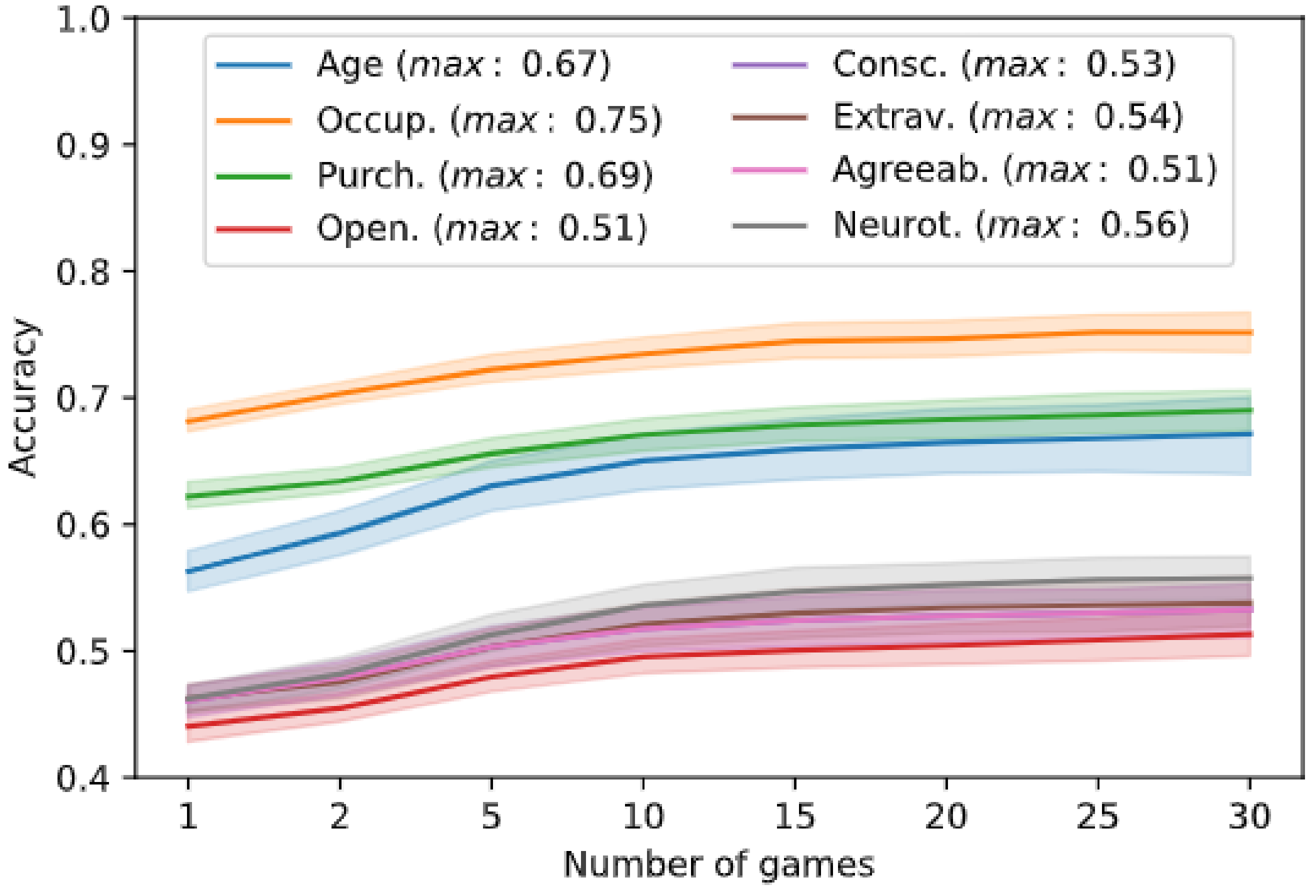
Input: Many matches of player *x*, using *expert* features

Output: Avg probability of Target Attribute for player *x*

Example: Is *x* male?

Match 1: probability = 0.1
Match 2: probability = 0.2
Match 3: probability = 0.8
Match 4: probability = 0.2

Avg: 0.325 -> x is female
(the error is filtered out)



Practical AIA (The True Threat)
Indiscriminate 'many-to-many' AIA

Idea: The attacker is satisfied with “not completely wrong” predictions

Method: Consider both first and second predictions as correct (via sophisticated AIA)

Table 6: Indiscriminate ‘many-to-many’ AIA (mid column). Compared to the baseline (cf. Fig. 5), the accuracy substantially increases.

	Sophisticated AIA (30 matches)	Indiscriminate AIA (30 matches)	Improvement
age	$67.15_{\pm 6.87}$	$\mathbf{89.15}_{\pm 4.66}$	+22.00%
purch.	$68.99_{\pm 3.81}$	$\mathbf{96.13}_{\pm 2.86}$	+27.14%
open.	$51.30_{\pm 3.87}$	$\mathbf{77.86}_{\pm 3.39}$	+26.56%
consc.	$53.24_{\pm 4.88}$	$\mathbf{80.19}_{\pm 4.12}$	+26.95%
extrav.	$53.78_{\pm 3.90}$	$\mathbf{81.51}_{\pm 4.40}$	+27.73%
agreeab.	$50.71_{\pm 4.65}$	$\mathbf{76.84}_{\pm 5.59}$	+26.13%
neurot.	$55.74_{\pm 3.88}$	$\mathbf{80.64}_{\pm 4.02}$	+24.90%

Practical AIA (The True Threat)
Targeted 'many-to-one' AIA

Idea: The attacker wants to be precise in finding a target, not in finding all of them

Method: Train and validate models to reach high precision (via sophisticated AIA)

$$\text{Precision} = \frac{TP}{TP + FP}$$

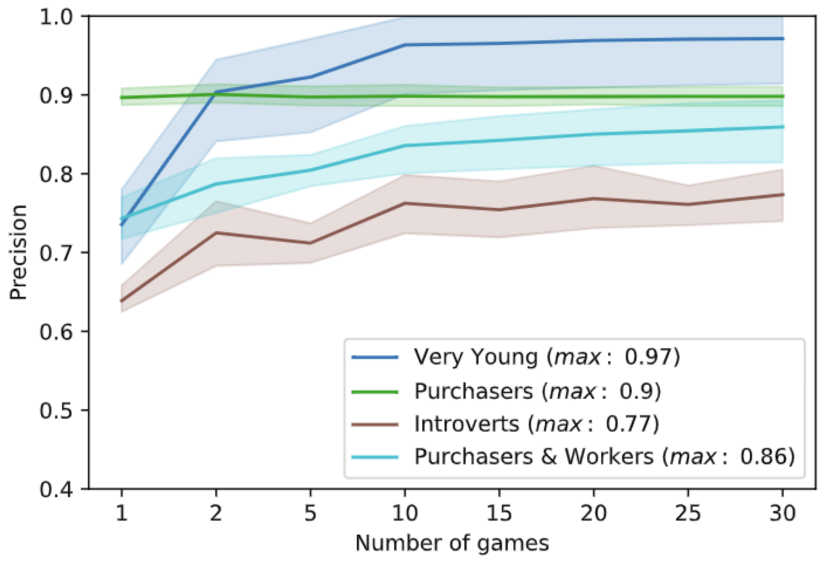


Fig. 6: Targeted 'many-to-one' AIA. We train our ML models by maximizing the *precision* on a single targeted class. Such AIA are very effective after analyzing 10 matches for each player in the test-set.

Countermeasures

- **Hard counters?** No!
  - The entire E-sport ecosystem would be disrupted
  - VG/TW Could remove the most correlated features… But relevant ones are so many, and others are likely to appear!

- **Compromises?** Yes!
  - The users should be informed that having their in-game statistics to be publicly accessible by TW exposes them to AIA
  - Access control rules
  - Turn TW into social networks
  - All of these require effort and collaboration between VG and TW (not easy!)

Extension to other E-Sports

- **What about other games?** Many E-sports share the same ecosystem with Dota2
  - AIA are theoretically possible also in other VG, but correlations need to be found first!
  - TW are not necessary, data come directly from VG!

Table 7: Overview of E-Sports VG. Numbers are taken from various sources [17, 20, 32, 52, 59].

	Release Year	*Genre*	*Monthly Players*	*Concurrent Players Avg*	*Playtime Avg*	*Age Range (PEGI rec.)*	*Tournament Revenue*	*Exemplary TW*	*Replay System*	*Max Players per Lobby*
League of Legends	2009	MOBA	127 M	700 K	832 H	11–50 (12+)	$93 M	lolprofile.net	Yes	10
CS:GO	2012	FPS	34 M	560 K	611H	13–40 (18+)	$134 M	csgostats.gg	Yes	18
Rocket League	2016	Sport	90 M	25 K	315 H	6–35 (3+)	$18 M	rltracker.pro	Yes	8
Fortnite	2017	Battle Royale	270 M	4 M	1800 H	6–54 (12+)	$121 M	fortnitetracker.com	Yes	100
PUBG	2018	Battle Royale	510 M	200 K	356 H	12–55 (16+)	$45 M	pubg.op.gg	Yes	100
Apex Legends	2019	Battle Royale	118 M	195 K	91 H	8–37 (16+)	$10 M	apex.tracker.gg	No	60
DOTA2	2013	MOBA	3.7 M	450 K	1700H	12–50 (12+)	$283 M	opendota.com	Yes	10

- **We sent an email to Valve** to inform them of such vulnerability.
  - We are unsure about whether they will take any action in the short-term

Takeaway

- Attribute Inference Attacks in video games are a concrete, feasible, and real threat!

- More research on video-games security is needed!

Pier Paolo Tricomi

Lisa Facciolo

Thank you!

Questions?

Giovanni Apruzzese

Mauro Conti

AIA Performances in Previous Work

Table 5: Results of prior work on AIA. Cells denote the value of a given 'Metric' for each of the attributes considered in our paper.

Prior Work	Metric	gend.	age	occup.	open.	consc.	extrav.	agreeab.	neurot.
Goelbeck [26]	MAE	–	–	–	0.09	0.10	0.14	0.11	0.13
Weinsberg [64]	AUC	0.84	–	–	–	–	–	–	–
Al [4]	Acc.	0.80	0.80	–	–	–	–	–	–
Chen [14]	AUC	0.82	0.61	–	–	–	–	–	–
Fang [23]	Acc.	0.80	0.73	0.25	–	–	–	–	–
Bunian [11]	Acc.	–	–	–	0.58	0.60	0.58	0.58	0.58
Yo [69]	Acc.	0.70	0.80	0.70	–	–	–	–	–
Mei [43]	MAE	–	0.09	–	–	–	–	–	–
Pijani [51]	F1	0.83	–	–	–	–	–	–	–
Zhang [71]	F1	0.74	0.38	0.13	–	–	–	–	–
Eidizadehakhcheloo [21]	AUC	0.95	0.98	–	–	–	–	–	–