# All Your Players Are Belong To Us – Reverse-Engineering Player Identities

Dennis Aumiller, Philip Hausner, Fabian Kaiser

DataFest Germany 2019

Heidelberg University
Computer Science

May 05, 2019

# Motivation

- Sensitive data in the dataset
- Anonymity and confidentiality of data should be a priority
  - → Let's see just how *anonymous* it is!
- The point is to show that it is entirely possible, not to inflict any harm

- Majority of data is GPS tracking info
  - → How can we exploit this?



Professional Sports Teams are Risking a Cybersecurity Own Goal

Figure: https://www.infosecurity-magazine.com/opinions/professional-sports-teams/



NOV 08 2018

**SECURITY**

# Sports Face Off Against Cybersecurity Threats

High-profile cyberattacks have put teams and other sports organizations on alert.

Figure: https://biztechmagazine.com/article/2018/11/sports-face-against-cybersecurity-threats

# Methodology

Idea: Utilize external player data to get back to the real identity

## Roster Matching

- Attendance of 12 players per event
  - Total squad size of 17 players

## Substitution Resolution

- Resolve conflicting assignments
- Determine substitutions from provided GPS data
  - Via movement variance

## External Data Sources

- Online match results
  - `https://world.rugby/`
  - `https://rwcsevens.com/`
  - ...
- Contains player line-ups and substitutions (with time)



Player 2



Player 3



Player 4

# Results

- Purely matching by the roster
  $\rightarrow$ 8 players' identities revealed!

- For the remaining players, all but two could be matched by substitution resolution
  $\rightarrow$ Easy resolution via VOD analysis

- Unsupervised recognition of substitutions possible (prototype precision: $> 90\%$)

**100% of players identified!**

# Questions

Thank you for your attention!