CROSS-DEVICE MEASUREMENT

PATCG Discussion Topic #58

Martin Thomson, 2022-06

SIMPLE GOAL

Count the number of times that an ad on site Y immediately preceded a purchase on site X in the same browser

ATTRIBUTION FUNCTION

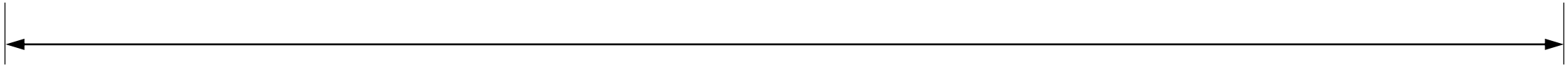
For each conversion on site X, find all ad events that:

- are from the same browser,
- precede the conversion, and
- meet some additional conditions

For those events, calculate a function over associated values

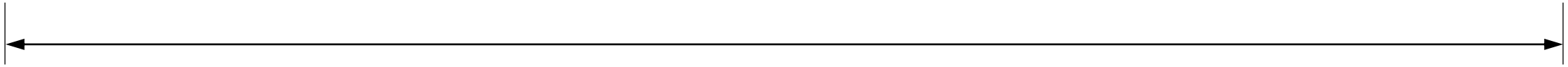
Aggregate the results of each calculation against the site in the ad event using a second function

SAME BROWSER, OR...



The same **person**?

PERSPECTIVES



Utility

Privacy

Competition

UTILITY ARGUMENT

Facebook data (2020) shows

- 7 day last-click attribution is ~30% cross-device
- 30 day multi-touch (view or click) is ~80% cross-device
- ¾ of advertisers observe most conversions across multiple devices

Google survey (2012) shows

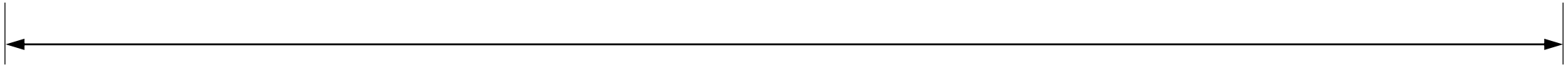
- 90% of people use multiple devices to accomplish tasks
- 67% continued online shopping on another device
- 51% used search when moving between devices for shopping

UTILITY, ML, EXTENT

Zimmeck et. al. (2017) found that ML predictions for interest in a topic improved with cross-device data:

- Gain attributed to use of more and more relevant features
- Estimated at least 67% (Desktop)/64% (Mobile) activity is tracked
- Estimated at least 20% of activity tracked cross-device (incl. Apps)

UTILITY

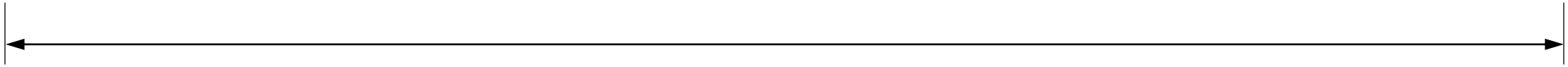


Measuring cross-device interactions improves performance

More events, more choice in attribution modelling

No real drawbacks

PRIVACY



More utility = worse privacy, right?

WHO CAN LINK DEVICES?

Browser/platform

- Same vendor – possible
- Different vendors – maybe

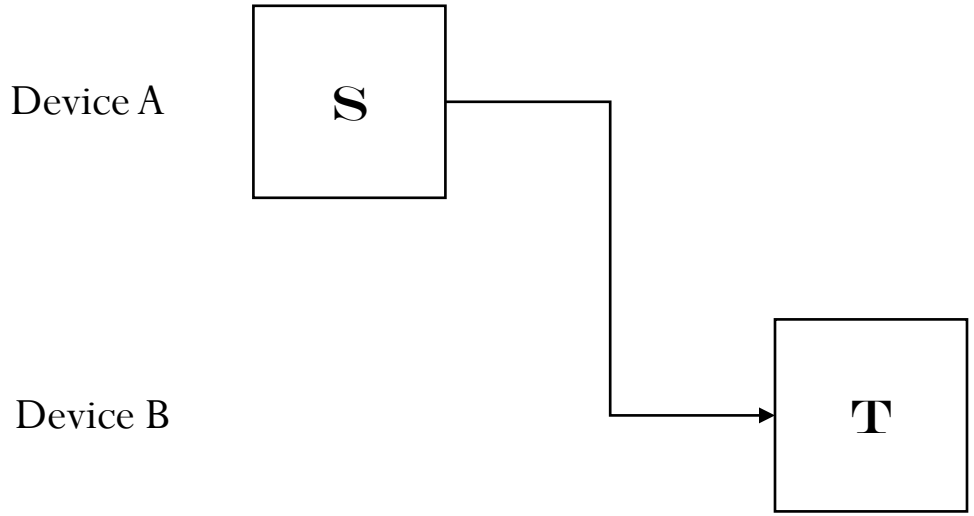
Applications with cross-device login

Applications with tracking heuristics (IP, …)

MORE UTILITY

Cross-device measurement makes more information available

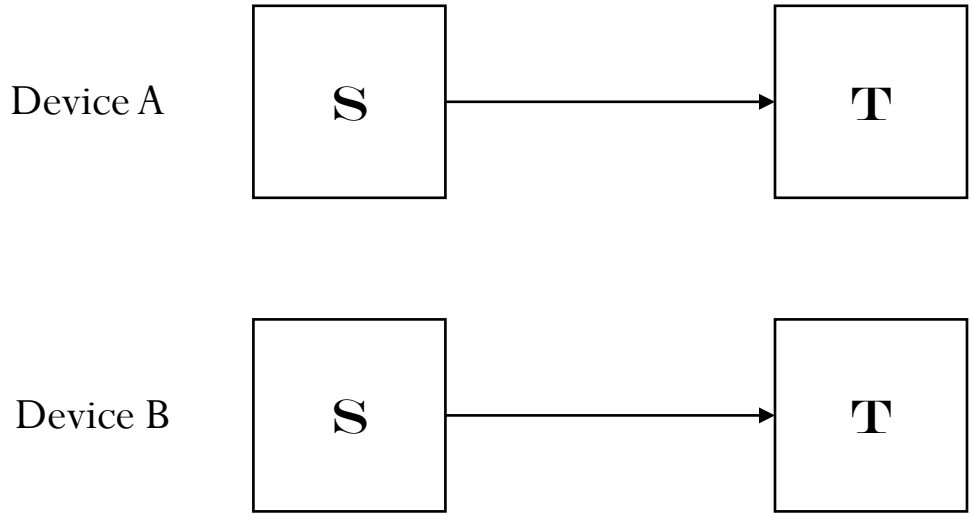
for linking events



MORE PRIVACY

Cross-device measurement makes more information available

for limiting disclosure



ABUSE

Abuse occurs when sites use measurement for tracking

- Finding worst-case characteristics describe abuse potential

If a site can link devices, but the measurement system cannot

- Site amplifies information gain by number of devices
- Measurement system remains ignorant of abuse

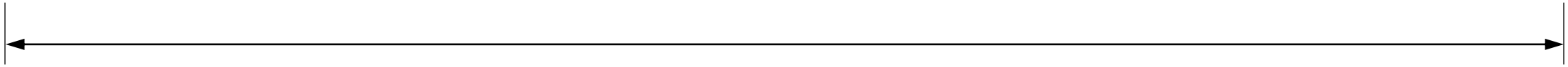
DEFENDING ABUSE

Differential privacy relies on having a bound on contributions

- Unlinked contribution is proportional to number of devices
- Without cross-device knowledge the bound is at best a guess
- Conservative guesses mean more noise ($\mathcal{O}(\sqrt{n})$)

Using device linkage enables tighter bounds

COMPETITION EFFECTS



Which sort of actor benefits most from different outcomes?

BASELINE

Lots of actors already in a position to link devices today

- Lots of deterministic linkage already (accounts, email, phone #)
- Heuristics are also in wide use for attribution (IP, time of day)

Coverage is likely not uniform

- Tracking countermeasures will render some ineffectual

LEAKING DEVICE REACH

The entity that provides device linkage reveals something

- Browser/platform or website

Abuse of the API might be used to get estimates of how many devices people have – as seen by the provider

CHANGE IS CHANGE

To the extent that cross-device measurement would entail change, that change has an effect

You get what you (can) measure

- Corollary: you can't get what you can't measure
- Example: Balance between search and display ads

DISCUSS