

ECE458/ECE750T27: Computer Security

3 Attack Cases – Cambridge Analytica

Dr. Kami Vaniea,
Electrical and Computer Engineering
kami.vaniea@uwaterloo.ca



UNIVERSITY OF
WATERLOO

FACULTY OF
ENGINEERING



First, the news...

- First 5 minutes we talk about something interesting and recent
- You will not be tested on the news part of lecture
- You may use news as an example on tests
- Why do this?
 1. Some students show up late for various good reasons
 2. Reward students who show up on time
 3. Important to see real world examples

3 Cases

- Verizon Supercookie
- Lenovo Superfish
- Cambridge Analytica

CAMBRIDGE ANALYTICA

Course concepts

- Access control
- Privacy
- Consent
- Law

Micro-targeting and political campaigns

- Cambridge Analytica enabled micro-targeting by political campaigns
- Micro-targeting:
 - Tell each voter what they most want to hear
 - Apps for door-to-door campaigners that tell them information about the home resident before they even knock
- Micro-targeting requires data about individuals to work

Obama 2012: The Most Micro-Targeted Campaign in History?

In his 2010 book “The Audacity to Win,” Obama 2008 campaign director David Plouffe explained that the goal of the campaign was not only to ensure high participation and [...]

by **MATTHEW C. NISBET**

May 2, 2012

IN HIS 2010 BOOK “The Audacity to Win,” Obama 2008 campaign director David Plouffe explained that the goal of the campaign was not only to ensure high participation and turn out from the Democratic base but also to expand the size of the electorate by mobilizing first-time voters among young adults and minorities. To reach these voters, the campaign focused on Internet and text message strategies, sending more than a 1 billion emails over the course of the campaign, and turning My.Barack.Obama.com into a powerful echo-chamber where all things campaign related were re-interpreted through the prism of a pro-Obama outlook.

In today’s Washington Post, Dan Eggen adds to a series of articles at the Post and NY Times detailing the Obama 2012 campaign’s even more advanced micro-targeting and Internet strategies. As Eggen writes, Obama’s dog Bo and pet lovers have been added to the mix of targeting appeals:

For the Obama campaign, pet lovers are just one niche among many, with specific appeals aimed at women, African Americans, students, military families and countless others. The result is a campaign that might be the most micro-targeted in history, attempting to use the power of the Web and social media to reach ever-thinner slices of the electorate.

Predicting protected information

- Researchers found out that:
 - “easily accessible digital records of behavior, Facebook Likes, can be used to automatically and accurately predict a range of highly sensitive personal attributes including: sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive substances, parental separation, age, and gender.”

Private traits and attributes are predictable from digital records of human behavior

Michal Kosinski , David Stillwell, and Thore Graepel [Authors Info & Affiliations](#)

Edited by Kenneth Wächter, University of California, Berkeley, CA, and approved February 12, 2013 (received for review October 29, 2012)

March 11, 2013 | 110 (15) 5802-5805 | <https://doi.org/10.1073/pnas.1218772110>

765,480 | 1,852



Abstract

We show that easily accessible digital records of behavior, Facebook Likes, can be used to automatically and accurately predict a range of highly sensitive personal attributes including: sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive substances, parental separation, age, and gender. The analysis presented is based on a dataset of over 58,000 volunteers who provided their Facebook Likes, detailed demographic profiles, and the results of several psychometric tests. The proposed model uses dimensionality reduction for preprocessing the Likes data, which are then entered into logistic/linear regression to predict individual psychodemographic profiles from Likes. The model correctly discriminates between homosexual and heterosexual men in 88% of cases, African Americans and Caucasian Americans in 95% of cases, and between Democrat and Republican in 85% of cases. For the personality trait “Openness,” prediction accuracy is close to the test-retest accuracy of a standard personality test. We give examples of associations between attributes and Likes and discuss implications for online personalization and privacy.

Idea behind Cambridge Analytica

- If data from a large number of people could be obtained:
 - Political views could be predicted with high accuracy
 - Advertisements could be micro-targeted
- Cambridge Analytica could then sell micro-targeted persuasion options to political campaigns worldwide
- Cambridge Analytica was used during:
 - 2016 US presidential election
 - UK Brexit campaign
 - Leave.EU: – ICO's investigation did not find that Cambridge Analytica used user data in the campaign
 - UKIP – unknown, but still being challenged in courts

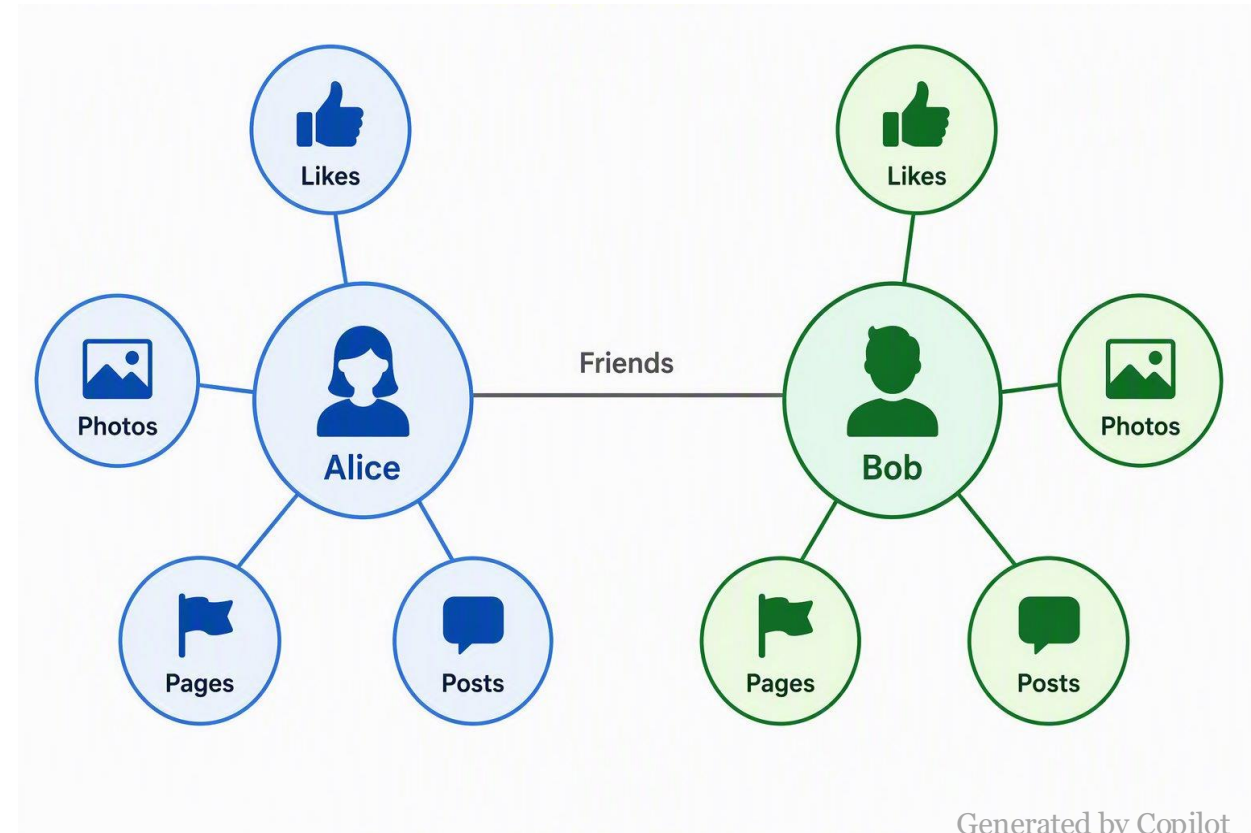
Facebook's Graph API v1.0

- Facebook wanted to enable social game play
- Games on FB were given access to a full list of the user's friends AND information on those friends
 - All friends, just not friends with the app
 - Info: likes, location, birthday
- Facebook policies stated that data was not to be downloaded and retained by add developers



Imagine a social network as a massive graph

- Users are nodes
- Friend relationships are edges
- Photos, likes, groups, pages, events, etc. are also nodes



Consent and queries

- User Alice installs an app FooBar
- Alice consents to FooBar accessing likes, photos, and friends
- FooBar is given an access token ALICE_TOKEN
- FooBar on behalf of Alice can access authorized content

Some graph queries:

`/me`

`/me/likes`

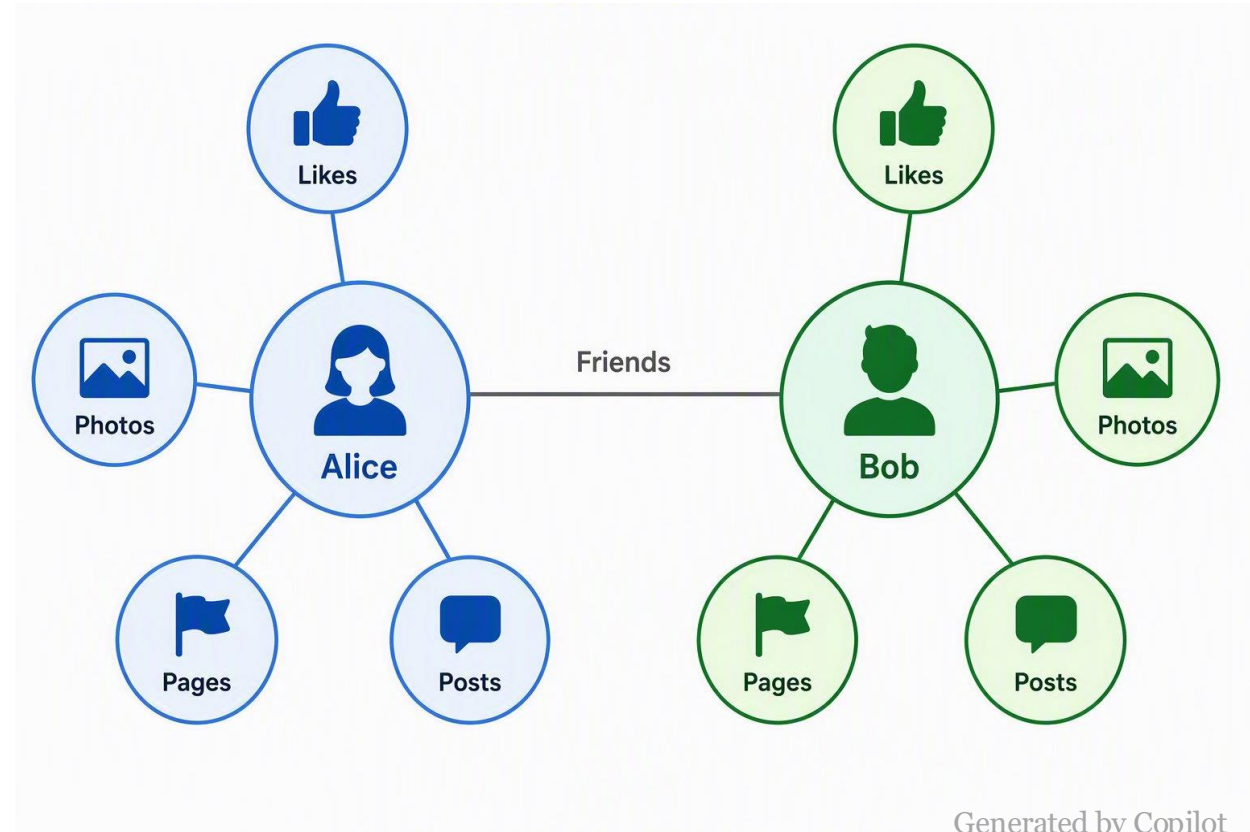
`/me/photos`

`/me/friends`

`https://graph.facebook.com/me/friends?access_token=ALICE_TOKEN`

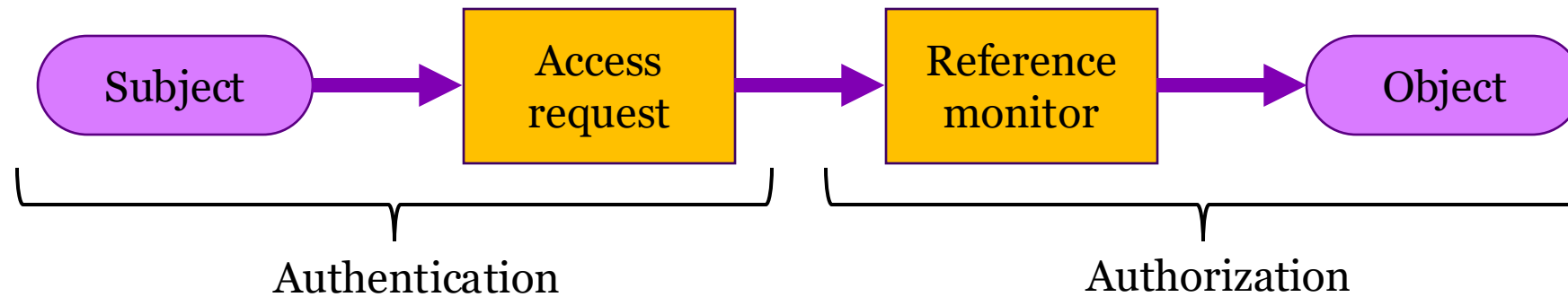
Think-pair-share: How should access by an app be managed?

- Discretionary Access Control (DAC)
 - The **owner** of a resource decides who may access that resource. Policy set on a case-by-case basis.
- Mandatory Access Control (MAC)
 - The decision of accessing resources is controlled **system-wide** by a uniform policy.
- Role Based Access Control (RBAC)
 - Similar to MAC, but users are assigned roles and permissions are granted to roles, not people.
- Capabilities
 - Single- or multi-use ticket for a <subject, object, access mode> triple



Access control

A guard controls whether a principal (the subject) is allowed access to a resource (the object).



Cambridge Analytica – the company

- ThisIsYourDigitalLife – Facebook app that conducted a personality test based on a range of questions
- App also collected lots of data about the user (posts, messages, email addresses, ...)
- And about friends (location, birth date, photographs, pages the friend liked)
- Had 320,000 direct users and up to 87 million “friends”
- Data was then used to provide modelled data about the app user and their friends
- Cambridge Analytica also paid a 3rd party survey company to get members of the public to install the app for <\$1

Impacting campaigns

- Data was used for micro-targeting of ads and messaging

In the run-up to the referendum vote in June 2016, AIQ ran 218 ads solely on behalf of Vote Leave and directed at email addresses on Facebook. In response to our information notice, Facebook stated that the

Consent: Did users consent?

- Users consented to installing the app and to the permissions
- Unfair and deceptive?

Joint investigation of Facebook, Inc. by the Privacy Commissioner of Canada and the Information and Privacy Commissioner for British Columbia

Report of findings

PIPEDA Findings #2019-002

April 25, 2019

Overview

In March 2018, in response to a complaint, the Office of the Privacy Commissioner of Canada (“OPC”) commenced an investigation into Facebook, Inc. (“Facebook”) relating to its compliance with the *Personal Information Protection and Electronic Documents Act* (“PIPEDA”) in the wake of revelations about Facebook’s disclosure of the personal information of certain of its users to a third-party application (the “TYDL App”)—information that was later used by third-parties for targeted political messaging. In April 2018, the OPC was joined by the Office of the Information and Privacy Commissioner for British Columbia (“OIPC BC”) and the investigation continued as a joint investigation. [1](#)

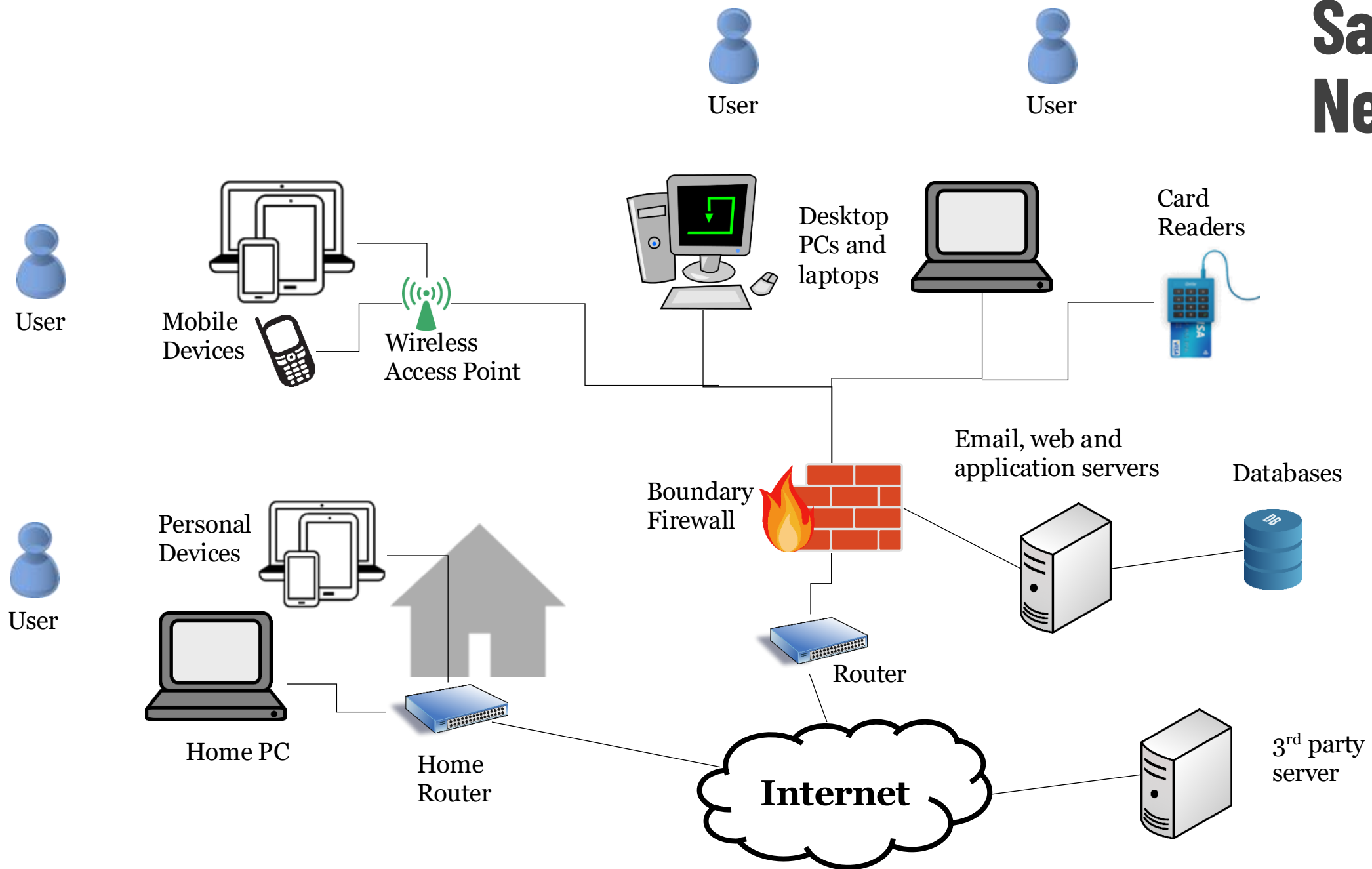
Our investigation focused on three general areas of concern under PIPEDA and the *Personal Information Protection Act (British Columbia)* (“PIPA”): (i) consent of users, both those who installed an app and their friends, whose information was disclosed by Facebook to apps, and in particular to the TYDL App; (ii) safeguards against unauthorized access, use and disclosure by apps; and (iii) accountability for the information under Facebook’s control.

To ensure a fair investigation of the facts, we have sought information and submissions from Facebook. We are disappointed that many of our questions have as yet gone unanswered or not answered to our satisfaction (i.e. they were incomplete, or otherwise deficient).

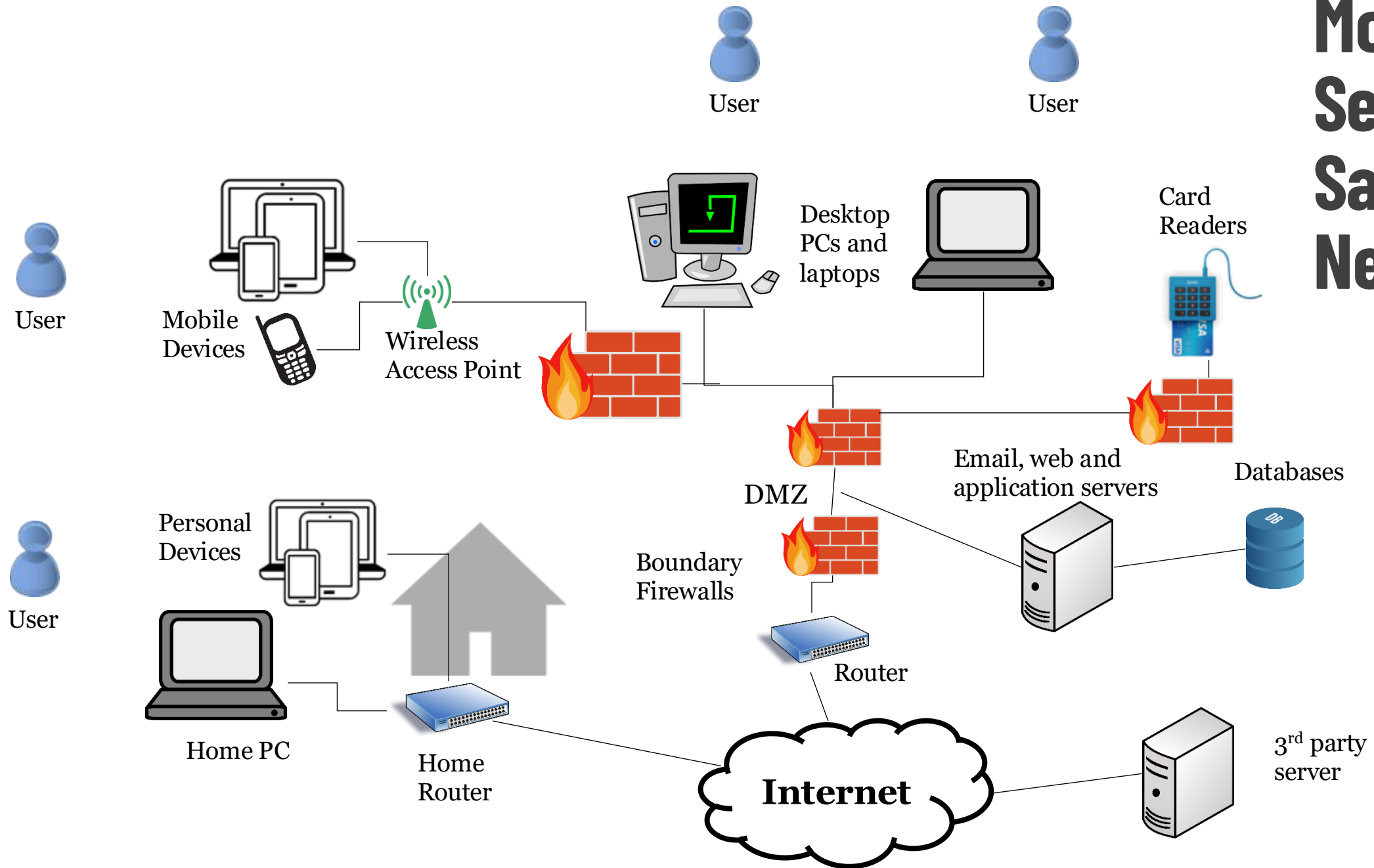
Based on the evidence gathered during this investigation, our findings can be summarized as follows:

EXTRA SLIDES FOR REFERENCE

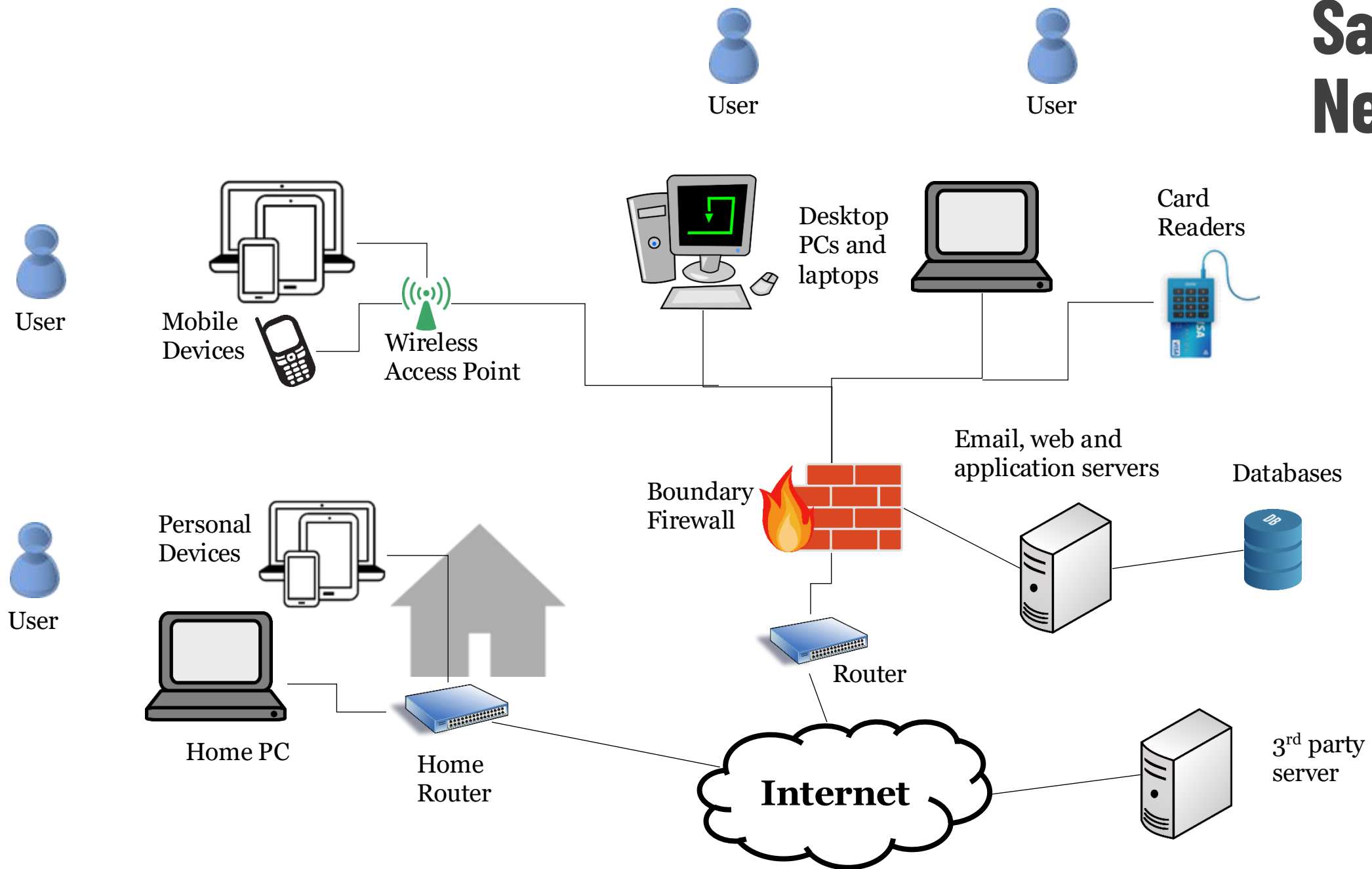
Sample Network



More Secure Sample Network

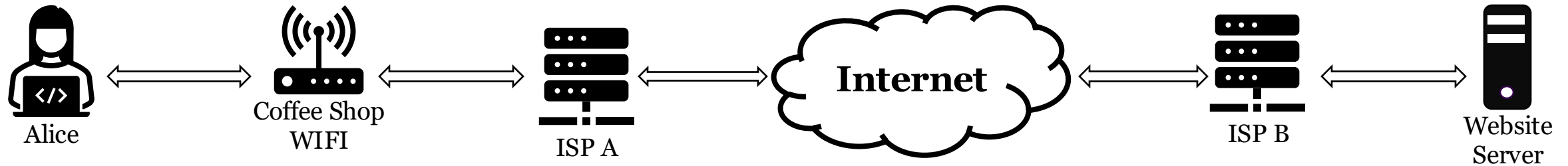


Sample Network



Sample connection: Alice loads a website

Alice visits: `http://example.com`



For each of the above icons, answer the following:

1. The name and/or IP address of the website Alice is visiting
2. The content of the webpage Alice is viewing
3. The IP address of Alice's computer (i.e. `ifconfig` or `ipconfig`)
4. The IP address of the Coffee Shop
5. Alice's Operating System