Phorm:
Function & Legal Failings

Dr Richard Clayton

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

fipr

LINX 61
19th May 2008

Overview

- What does it do ?
- How does it work ?
- Is it lawful ?

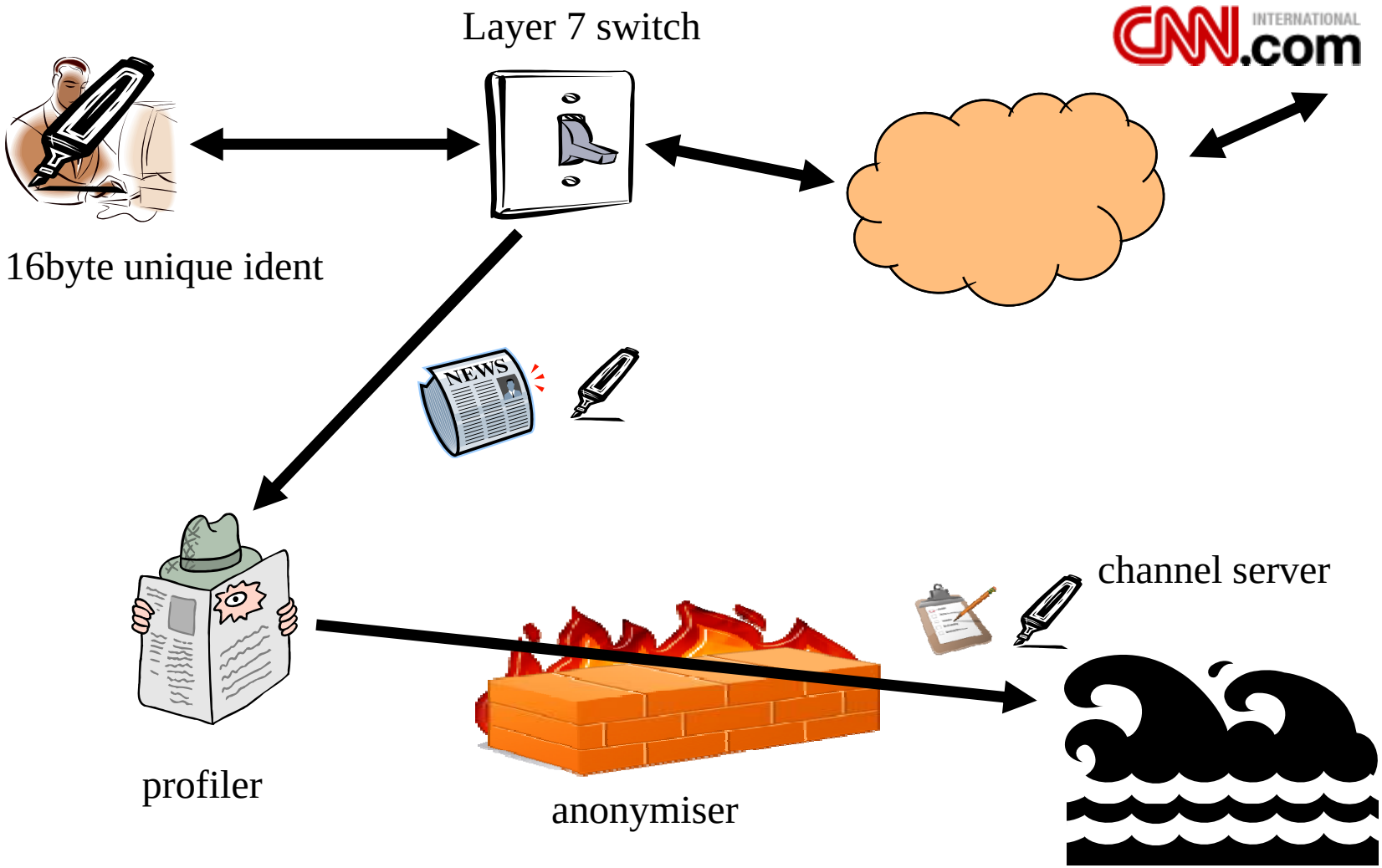
Behavioural advertising

- Advertising is big bu$ine$$!
- Google model is “put ads on relevant pages”
- Alternative approach is “show ads that are relevant to people who happen to visit”
  - DoubleClick tracks visits to participating sites by cookies (returned to DoubleClick)
  - Phorm inspects HTML on *(almost) all* visited pages to deduce nature of content, then serves relevant advert if you visit a participating site

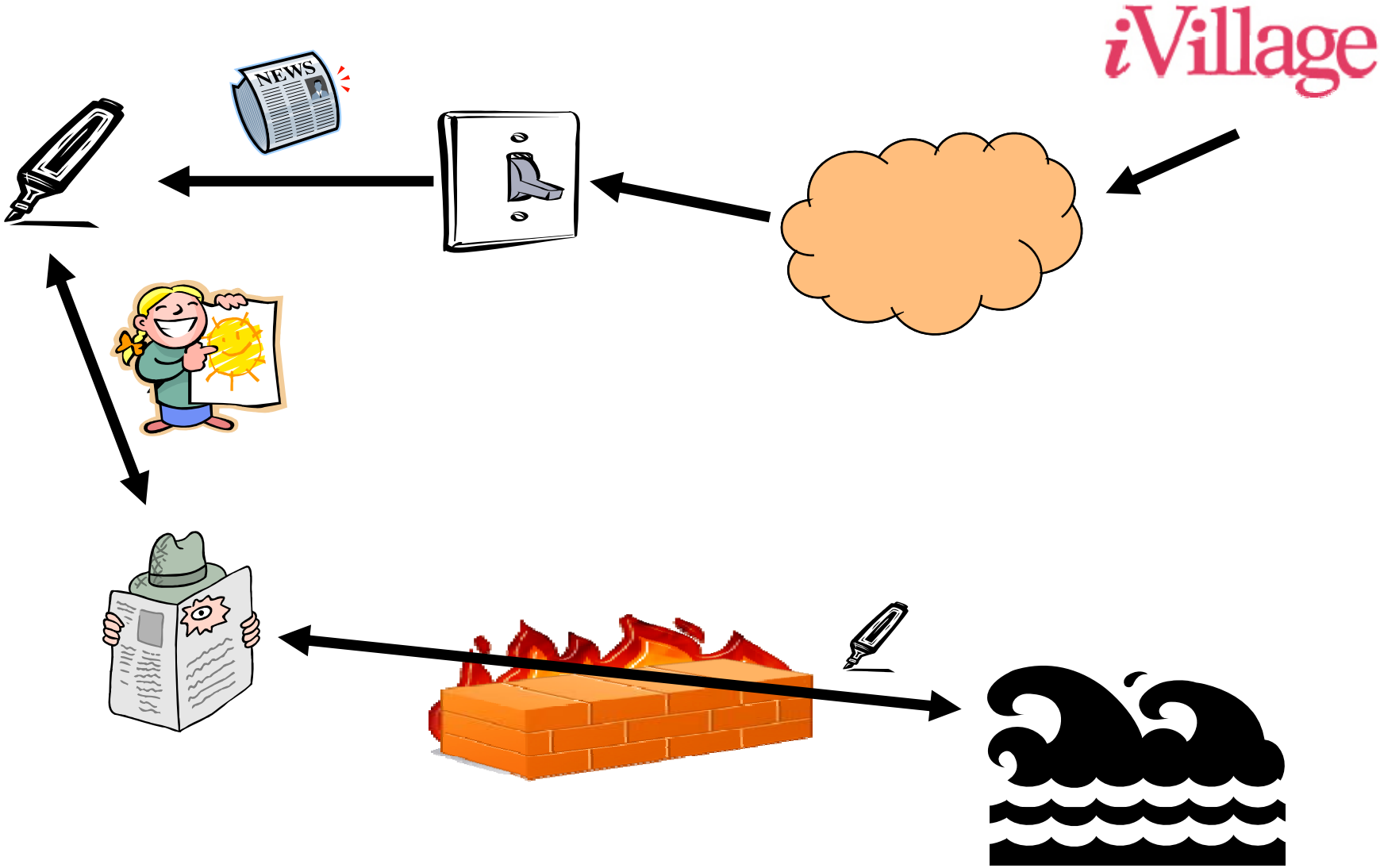
Advertisers

- Want to know what you do, not who you are
- Break people down into categories
  - ABC1, “empty nesters”, lots of fancy new names….
- So they can live with anonymity
- But will car adverts work on a book review site?
- Will adverts annoy?
  - what if you’ve already bought?
  - what if it’s supposed to be a surprise?

Phorm design #1



Phorm design #2



Distilled pages

attack
CleanFeed
format
inquiry
ISP
legal
packets
paper
PDF
system

content
document
event
partners
Phorm
PIA
privacy
School
system
Thinking

advertising
consumers
leading
OIX
online
Phorm
technology
Virgin
websites
Webwise

blog
comment
guardian
meeting
Offensive
Phorm
problem
reference
Unsuitable
written

This is what all those rubbish search engines used to do before Google came along!

Channel server

- Channel server also told of search terms
- Channel server also told of URLs
- Channel server only learns UID not IP address
- Channel server matches top 10 words etc against advert “channel”
- Only records the UID and time against channel
- Serves best (££) advert for matches like “has visited at least 3 travel sites in the past week”

Cookies (slide 1 of 2)

#1 **`GET cnn.com/index.htm`**

and a *fake cnn.com* responds:

`307 webwise.net/bind?cnn.com/index.html`

#2 **`GET webwise.net/bind?cnn.com/index.htm`**

system now allocates the user a 16byte UID

`307 webwise.net/bind-2?cnn.com/index.html`

accompanied by *webwise.net* cookie with UID

Cookies (slide 2 of 2)

#3 **`GET webwise.net/bind-2?cnn.com/index.html`**

so can check user is returning webwise cookies

`307 cnn.com/magic?cnn.com/index.html&UID`

#4 **`GET cnn.com/magic?cnn.com/index.html&UID`**

and a *fake cnn.com* machine responds

`307 cnn.com/index.html`

along with a ***cnn.com*** cookie containing UID

#5 **`GET cnn.com/index.html`**

Opting-out

- User can opt-out with a webwise.net cookie containing this request
- User can opt-out by refusing to return webwise.net cookies (or cnn.com cookies)
- User will have problems if they set webwise.net to resolve to 127.0.0.1
- ISPs looking at network level opt-outs
  - presumably RADIUS setting to select IP pool

Some safeguards etc.

- Only redirects recognised **`User-Agents`**
- Only considers **`text/html`** pages
- Doesn’t monitor HTTPS (duh!)
- Doesn’t monitor “simple auth”
- Doesn’t monitor “webmail sites”
- “Honours” **`robots.txt`**
- If nothing happening, switches itself off!

Legal failings &c

Has to be “opt in”

- May be processing “sensitive personal data” (religion, trade union, medical etc)
  `<h1>Union advice for vicars living with AIDS</h1>`
  - DPA requires an informed opt-in for this
- Information Commissioner says that Privacy and Electronic Communications Regulations requires an opt-in
- And besides, “Permission-based advertising” is the new black

Stability issues

- Access to new sites goes through four redirects
  - damages browser heuristics on wickedness
- webwise.net is now part of the CNI
  - breaks if webwise.net is “localhost” (127.0.0.1)
- Cookie based opt-out is not “fail safe”
  - standard advice on deleting cookies now wrong

Interception

- RIP 2000 requires permission from both ends of communication before disclosure
- s16 shows Phorm keywords do infringe
- Permission for data TO servers not given
- Permission for data FROM servers not given
- Permission from THIRD PARTIES not given
  - think “email” or “web forum”

Forging cookies

- Phorm impersonates domains to store cookies
- Clearly illegal under s1 of Fraud Act 2006
  - Phorm incites, the ISPs commit the offence
- ALSO leads to defamation claims
    - AND sometimes trademark infringement
  - if site says “we never use tracking cookies” and user inspects their file system then the user will conclude that the website owner is dishonest…

The public debate

Jan: HO says Phorm *may* be lawful under RIP

Feb: Phorm says its design is lawful

Mar: FIPR writes to regulators asking them to act

May: outlaw.com: “no harm caused to anyone”

Phorm now says:

> The article makes the important distinction between the letter and spirit of the law, as well as what often happens in practice. The author observes that Phorm is unlikely to break the spirit of the law and that many of the arguments against Phorm are based on technical breaches “and technical breaches are common”.

Cookie tossing

- Cookies removed by Phorm system
- But doesn’t prevent JavaScript (etc) access
  - hence websites can read Phorm identifiers
  - can sell linkage information (not in the EU!)
  - can fetch some Phorm adverts and then reverse engineer the user profile
- Can replace cookies with their own
  - wicked(?) website can opt you out (or in)

Privacy

- Privacy and Data Protection are not the same!
- Data Protection just mechanistic approach to controlling corporations with mainframes
  - and UK has minimal watered down variant
  - to a first approximation, anonymity fixes everything
- Privacy relates to controlled disclosure of information that matters TO YOU
  - your privacy is violated even if you are anonymous

Analogies (no cars!)

- Post Office opening all your letters, so you can get a better class of junk mail.
- TV company scanning everything on your bookshelf (and the magazines under the bed) so you can watch more interesting ads.
- Tesco collating what's on your shopping list so that you can be offered vegetarian menu when you go into McDonalds

My bottom line

- The proposition is a long way from proven
- It _has_ to be informed opt-in
- It does not improve stability
- As proposed, it is illegal to operate
- It will be defamatory
- It infringes your privacy
- Surveys suggest that customers will leave

?? So why on earth are the IP bothering ??

Phorm:
Function & Legal Failings

www.fipr.org

1998-2008

www.lightbluetouchpaper.org

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

fipr