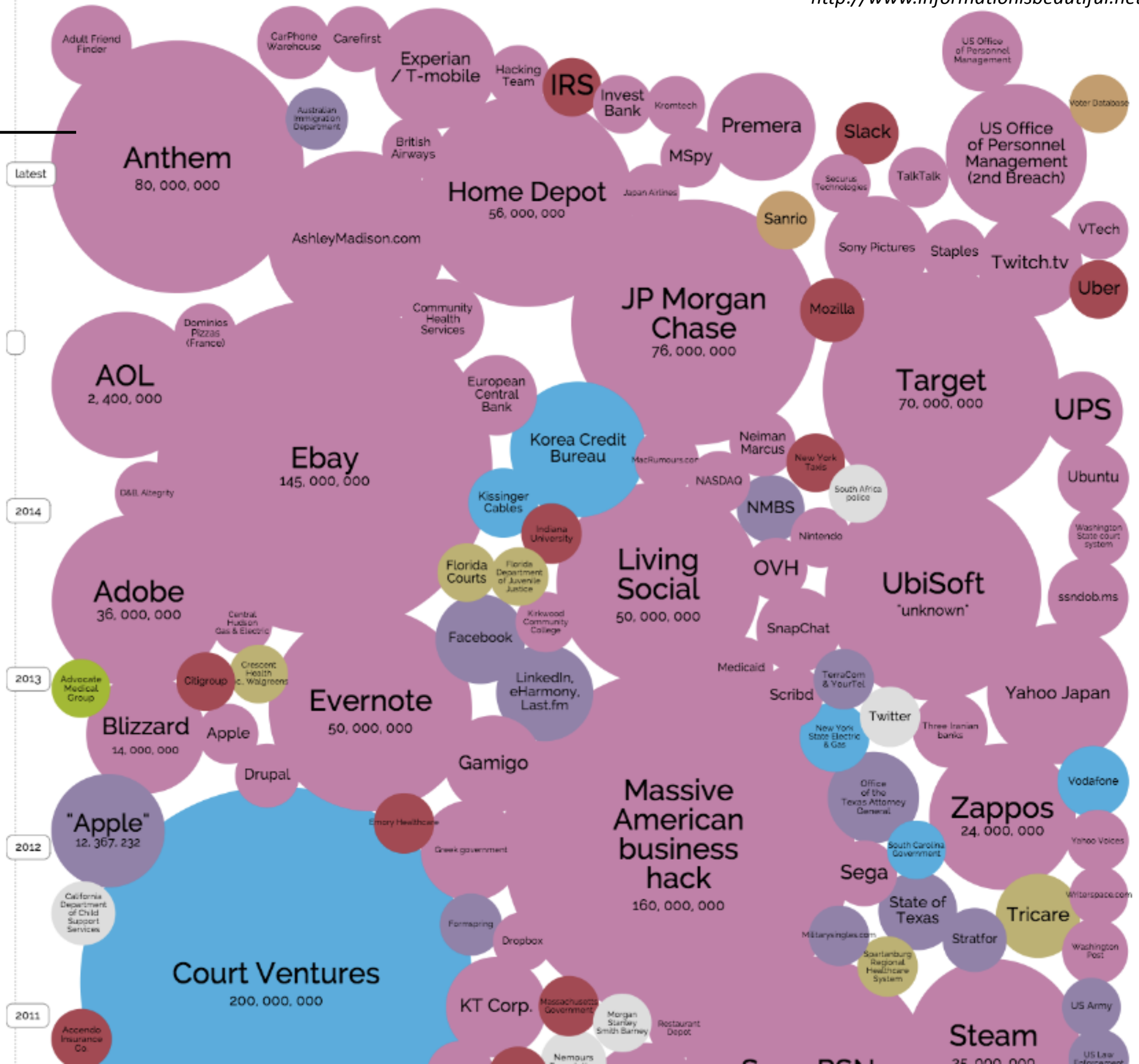


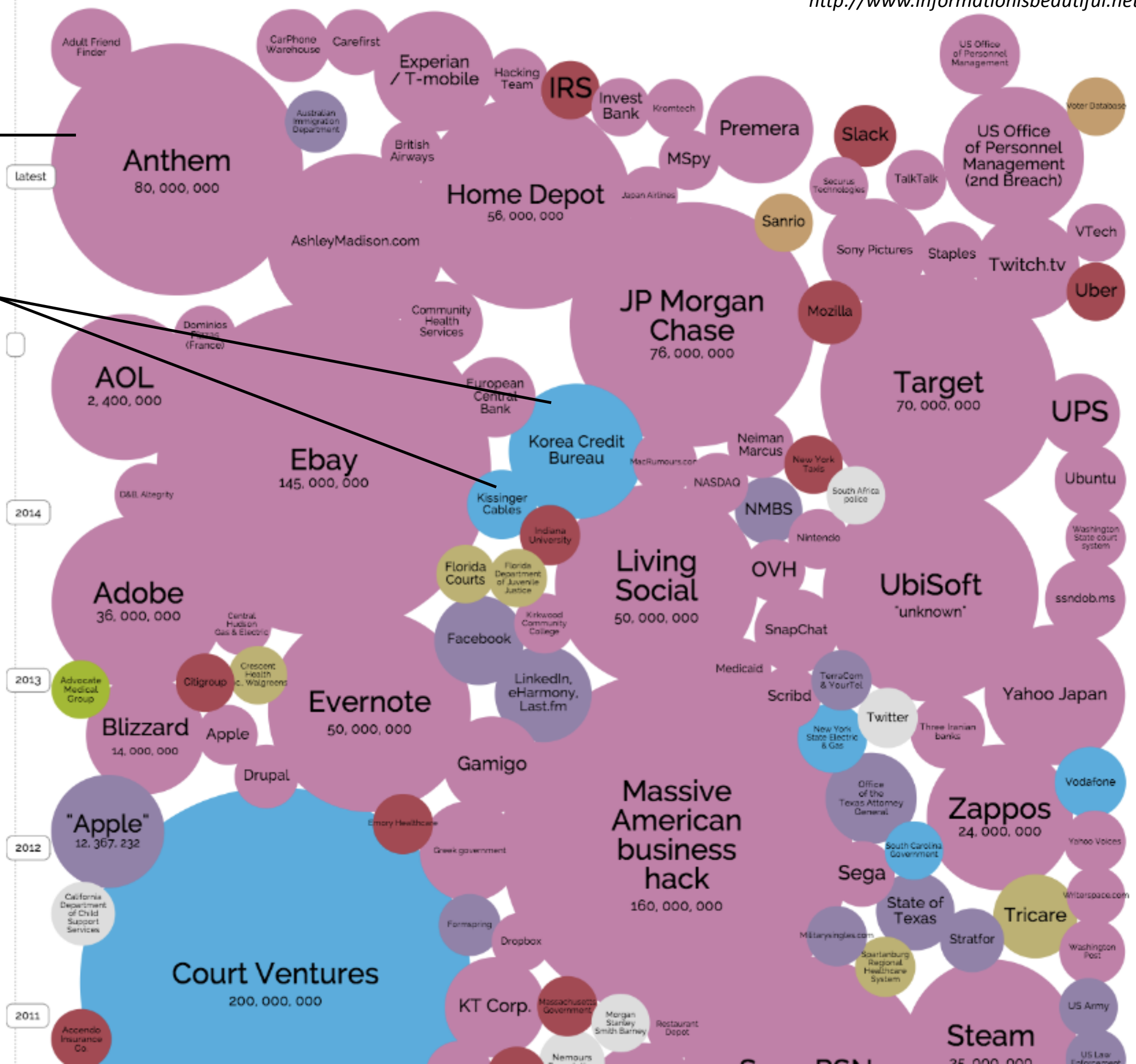
What Did We Learn From Recent Data Breaches?

Srdjan Čapkun

*Department of Computer Science
ETH Zurich*



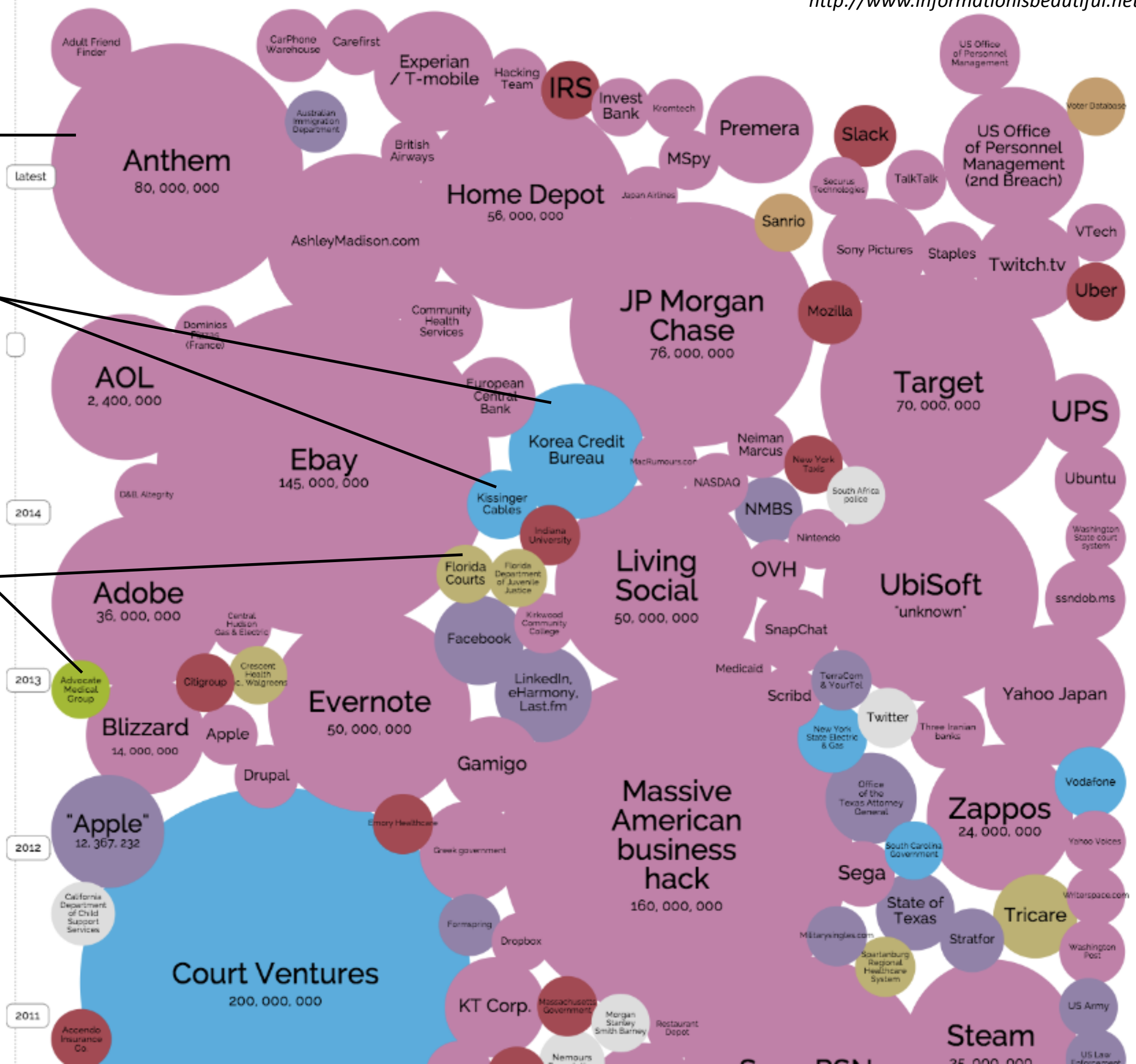
Inside job



Hacked

Inside job

Lost/stolen
media/pc

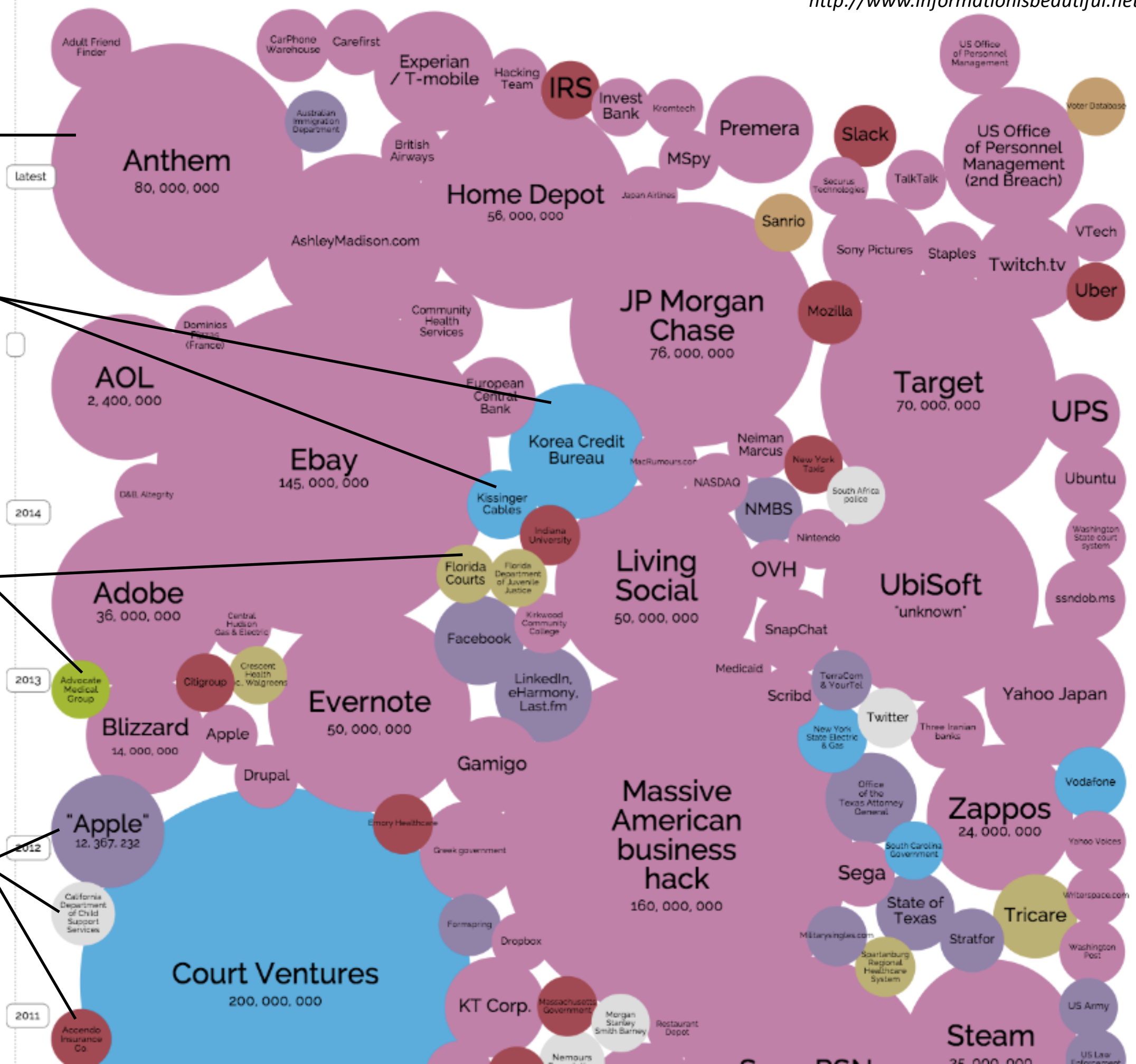


Hacked

Inside job

Lost/stolen
media/pc

- Accidentally published
- Configuration errors
- Poor security



VTech (2015)



- **SQL injection vulnerability** reveals 4.8 million customer email addresses, names and **weakly hashed passwords (MD5)**,
- **Leaked names and DOB of children, parents, photos, ...**
- VTech's account registration services also **did not use SSL/TLS**.

VTech (2015)



- **SQL injection vulnerability** reveals 4.8 million customer email addresses, names and **weakly hashed passwords (MD5)**,
- **Leaked names and DOB of children, parents, photos, ...**
- VTech's account registration services also **did not use SSL/TLS**.



Important Notice Learning Lodge Resumes Key Functions

The Learning Lodge® has been re-opened with the resumption of key functions. Customers of Learning Lodge connected products, with the exception of InnoTV™, InnoTab® MAX and some other products, are now able to securely register accounts for new products, manage their existing accounts, change their passwords and download content appropriate to their products.

Please click here for more information 

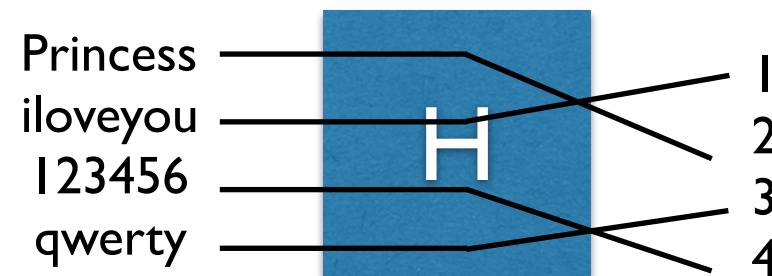
Ashley Madison (2015)

- **Undisclosed how the attack happened**
- 37 million customer records including **millions of account passwords made vulnerable by a bad hash implementation and weak passwords!**



Ashley Madison (2015)

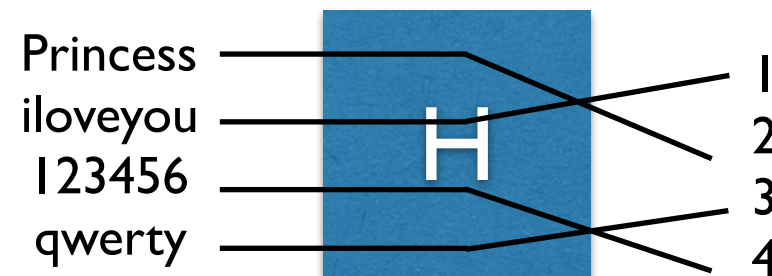
- **Undisclosed how the attack happened**
- 37 million customer records including **millions of account passwords made vulnerable by a bad hash implementation and weak passwords!**



Ashley Madison (2015)



- **Undisclosed how the attack happened**
- 37 million customer records including **millions of account passwords made vulnerable by a bad hash implementation and weak passwords!**



CynoSure analysis: 11 million hashes cracked

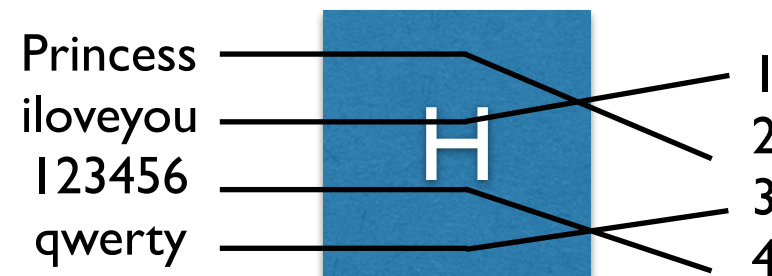
>630,000 people used usernames as passwords

MD5 hashed left around accidentally

<http://cynosureprime.blogspot.ch/2015/09/csp-our-take-on-cracked-am-passwords.html>

Ashley Madison (2015)

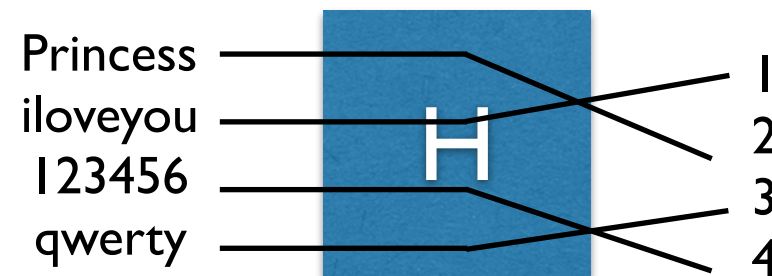
- **Undisclosed how the attack happened**
- 37 million customer records including **millions of account passwords made vulnerable by a bad hash implementation and weak passwords!**



Ashley Madison (2015)



- **Undisclosed how the attack happened**
- 37 million customer records including **millions of account passwords made vulnerable by a bad hash implementation and weak passwords!**



mypasswordispassword
superhardpassword
thebestpasswordever
thisisagoodpassword
likeimreallygoingtocheat
justcheckingitout
jusPryingthisout
goodguydoingthewrongthing
ishouldnotbedoingthis
ithinkilovemywife
thisiswrong
whaPhehellamidoing

US Office of Personnel Management (2015)



- Personnel records on 22 million current and former US federal employees (including some fingerprints)
- Attack: **Used contractor's stolen credentials to plant malware backdoor in the network.**
- Undetected for **343 days**
- Goal: intelligence, information gathering
- Discovery: Anomalous SSL traffic and a decryption tool found in the network, followed by a forensic investigation.

Why do we see so many breaches?

Why do we see so many breaches?

- Lot more data is being collected and it has value
Big Data = Big Value = Big Risk

Why do we see so many breaches?

- Lot more data is being collected and it has value
Big Data = Big Value = Big Risk

Why do we see so many breaches?

- Lot more data is being collected and it has value
Big Data = Big Value = Big Risk
- **Increased System Complexity and Connectivity**
 - Simple attacks **still work**.

Why do we see so many breaches?

- Lot more data is being collected and it has value
Big Data = Big Value = Big Risk
- **Increased System Complexity and Connectivity**
 - Simple attacks **still work**.

Why do we see so many breaches?

- Lot more data is being collected and it has value
Big Data = Big Value = Big Risk
- **Increased System Complexity and Connectivity**
 - Simple attacks **still work**.
- **Knowledge is *widely available***
 - Schools, Conferences, Tools widely available, and inexpensive
 - End-systems and protocols are reverse-engineered

Why do we see so many breaches?

- Lot more data is being collected and it has value
Big Data = Big Value = Big Risk
- **Increased System Complexity and Connectivity**
 - Simple attacks **still work**.
- **Knowledge is *widely available***
 - Schools, Conferences, Tools widely available, and inexpensive
 - End-systems and protocols are reverse-engineered

Why do we see so many breaches?

- Lot more data is being collected and it has value
Big Data = Big Value = Big Risk
- **Increased System Complexity and Connectivity**
 - Simple attacks **still work**.
- **Knowledge is *widely available***
 - Schools, Conferences, Tools widely available, and inexpensive
 - End-systems and protocols are reverse-engineered
- **Increasingly hostile environment**
 - Attacker is not only “on the channel”
 - Internal attackers (employees, subcontractors)
 - *“Post-Snowden attacker”:*
active subversion of systems, protocols and standards

Sensitive Data?

- Which data is sensitive?
- e.g., NSA is collecting metadata on phone calls ('just metadata')

Sensitive Data?

- Which data is sensitive?
- e.g., NSA is collecting metadata on phone calls ('just metadata')

Ex-NSA Director Admits 'We Kill People Based On Metadata' with Drone Strike

Wednesday, May 14, 2014 Wang Wei

 311  Like 1k  Share 1171  Tweet 230  Share 105  Share 2019



Sensitive Data?



ELECTRONIC FRONTIER FOUNDATION eff.org

30C3 – 30 December 2013

Why Metadata Matters

- They know you rang a phone sex service at 2:24 am and spoke for 18 minutes. But they don't know what you talked about.
- They know you called the suicide prevention hotline from the Golden Gate Bridge. But the topic of the call remains a secret.
- They know you spoke with an HIV testing service, then your doctor, then your health insurance company in the same hour. But they don't know what was discussed.

Let's assume we know which *information* is sensitive ...

- How do we share/release data sets WITHOUT leaking sensitive information (e.g., PII)?

Let's assume we know which *information* is sensitive ...

- How do we share/release data sets WITHOUT leaking sensitive information (e.g., PII)?
- e.g. AOL in 2006 released 'anonymized' search data of 650,000 users

Let's assume we know which *information* is sensitive ...

- How do we share/release data sets WITHOUT leaking sensitive information (e.g., PII)?
- e.g. AOL in 2006 released 'anonymized' search data of 650,000 users

A Face Is Exposed for AOL Searcher No. 4417749

By MICHAEL BARBARO and TOM ZELLER Jr.
Published: August 9, 2006

Buried in a list of 20 million Web search queries collected by AOL and recently released on the Internet is user No. 4417749. The number was assigned by the company to protect the searcher's anonymity, but it was not much of a shield.



No. 4417749 conducted hundreds of searches over a three-month period on topics ranging from “numb fingers” to “60 single men” to “dog that urinates on everything.”

And search by search, click by click, the identity of AOL

 SIGN IN TO E-MAIL THIS

 PRINT

 SINGLE PAGE

 REPRINTS

The New York Times

Let's assume we know which *information* is sensitive ...

- How do we share/release data sets WITHOUT leaking sensitive information (e.g., PII)?
- e.g. AOL in 2006 released 'anonymized' search data of 650,000 users

TechCrunch

AOL: "This was a screw up"

Posted Aug 7, 2006 by [Michael Arrington \(@arrington\)](#)

SHARES        

AOL is apologizing in the aftermath of [yesterday's story about their voluntary release of search data](#) on 650,000 users.



o. 4417749

AOL
e
's

SIGN IN TO E-MAIL THIS

PRINT

SINGLE PAGE

REPRINTS

eds of searches over a three-month period on topics ranging from "numb fingers" to "60 single men" to "dog that urinates on everything."

And search by search, click by click, the identity of AOL

The New York Times

Maybe We Should Delete Data?

Why would you want to delete data ?

- ***Once that data is not there, it can no longer be extracted***
- Legal/policy requirements (e.g., EU Data Protection Directive)

Maybe We Should Delete Data?

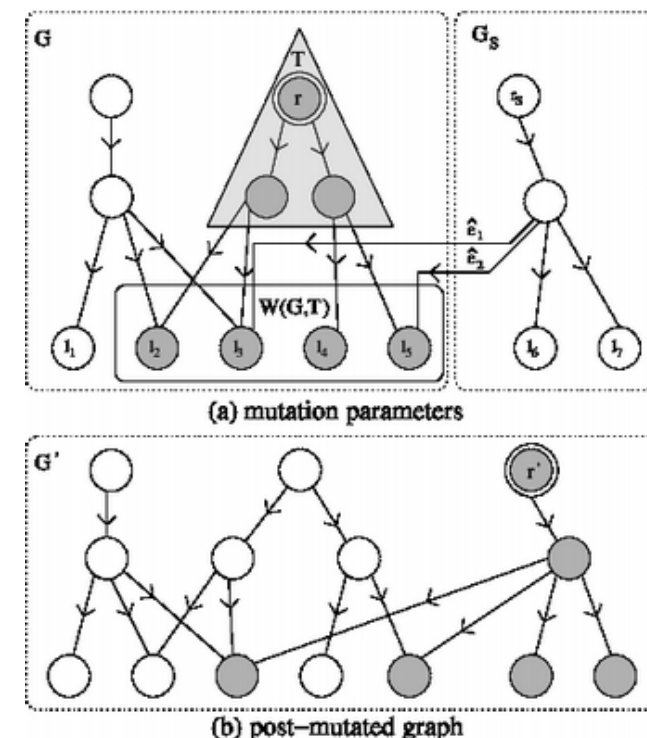
Why would you want to delete data ?

- *Once that data is not there, it can no longer be extracted*
- Legal/policy requirements (e.g., EU Data Protection Directive)
- *Our systems are not optimized to securely delete data (or metadata).*
- They are designed to be robust against failures and to re-purpose storage only when needed!

Maybe We Should Delete Data?

Why would you want to delete data ?

- ***Once that data is not there, it can no longer be extracted***
- Legal/policy requirements (e.g., EU Data Protection Directive)
- ***Our systems are not optimized to securely delete data (or metadata).***
- They are designed to be robust against failures and to re-purpose storage only when needed!
- A massive key management challenge.



Maybe We Should Delete Data?

Why would you want to delete data ?

- ***Once that data is not there, it can no longer be extracted***
- Legal/policy requirements (e.g., EU Data Protection Directive)
- ***Our systems are not optimized to securely delete data (or metadata).***
- They are designed to be robust against failures and to re-purpose storage only when needed!
- A massive key management challenge.

Belgian press reveals British hacking of Belgacom

Share 37 in Share 14 Tweet 33 Share 8

Published: 15/12/2014 - 07:08 | Updated: 15/12/2014 - 10:15 | 2

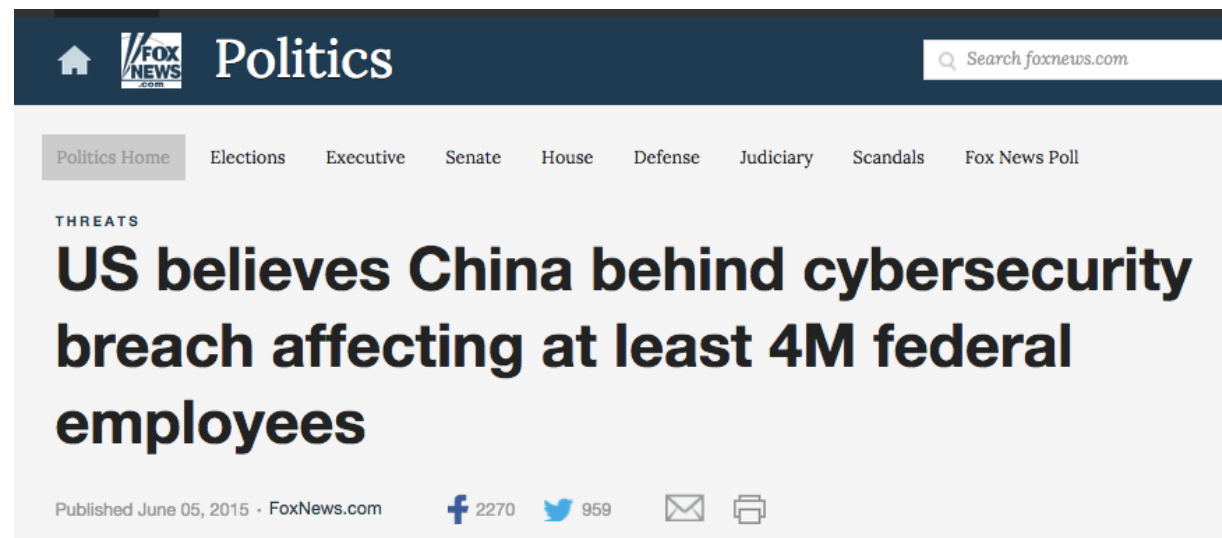


Maybe We Should Solve This Contractually?

What if the data leaks ... hard to prove who leaked it. *Attribution?*
Requires access to systems ... forensics ... crossing jurisdictions ...

Maybe We Should Solve This Contractually?

What if the data leaks ... hard to prove who leaked it. *Attribution?*
Requires access to systems ... forensics ... crossing jurisdictions ...



Maybe We Should Solve This Contractually?

What if the data leaks ... hard to prove who leaked it. *Attribution?*
Requires access to systems ... forensics ... crossing jurisdictions ...

News / Asia

Will the US-China Cybersecurity Pact Work?

 Print  Comment  Share:



President Barack Obama shakes hands with Chinese President Xi Jinping during their meeting in the Oval Office of the White House in Washington, Sept. 25, 2015.

Maybe We Should Solve This Contractually?

What if the data leaks ... hard to prove who leaked it. **Attribution?**
Requires access to systems ... forensics ... crossing jurisdictions ...

News / Asia

Will the US-China Cybersecurity Pact Work?

Print Comment Share:



President Barack Obama shakes hands with Chinese President Xi Jinping during their meeting in the Oval Office of the White House in Washington, Sept. 25, 2015.

Report: China Already Violated Cybersecurity Agreement with US

SHARE TWEET EMAIL



Obama and Chinese President Xi Jinping / AP

Tweet Google+ Reddit Pinterest

BY: Morgan Chalfant

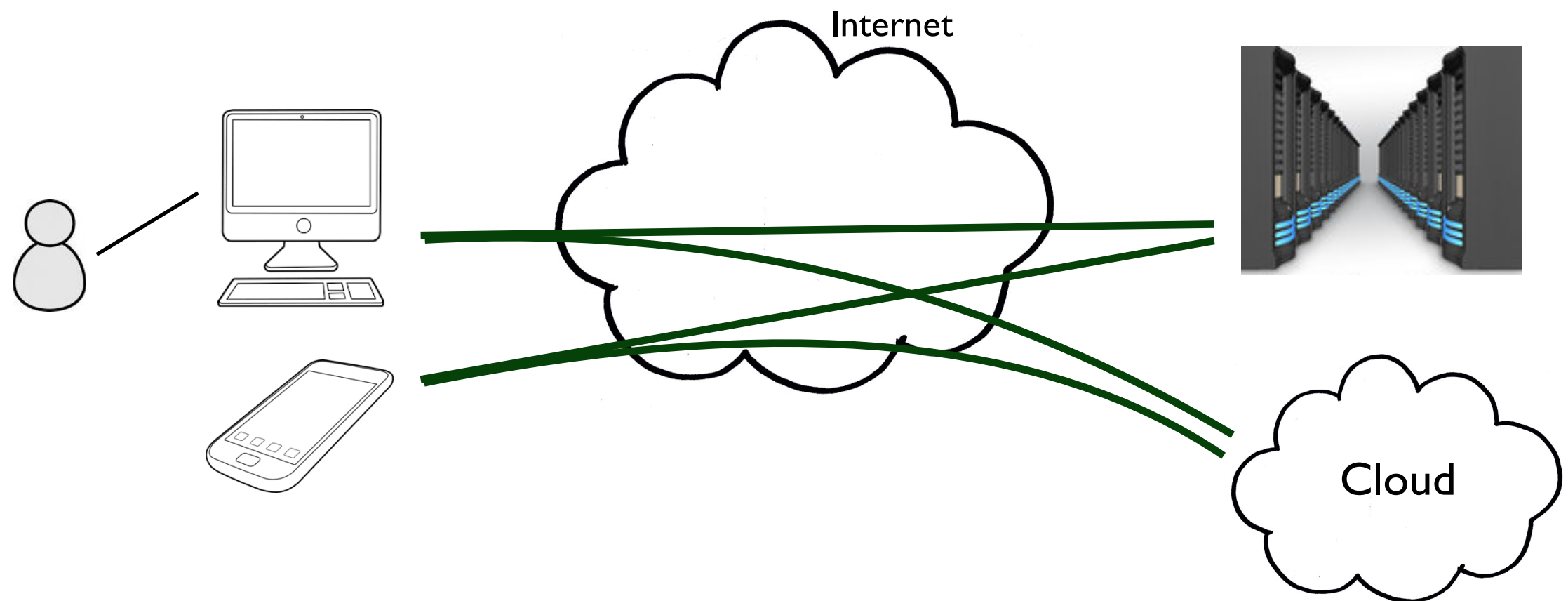
Follow @mchalfant16

October 19, 2015 11:33 am

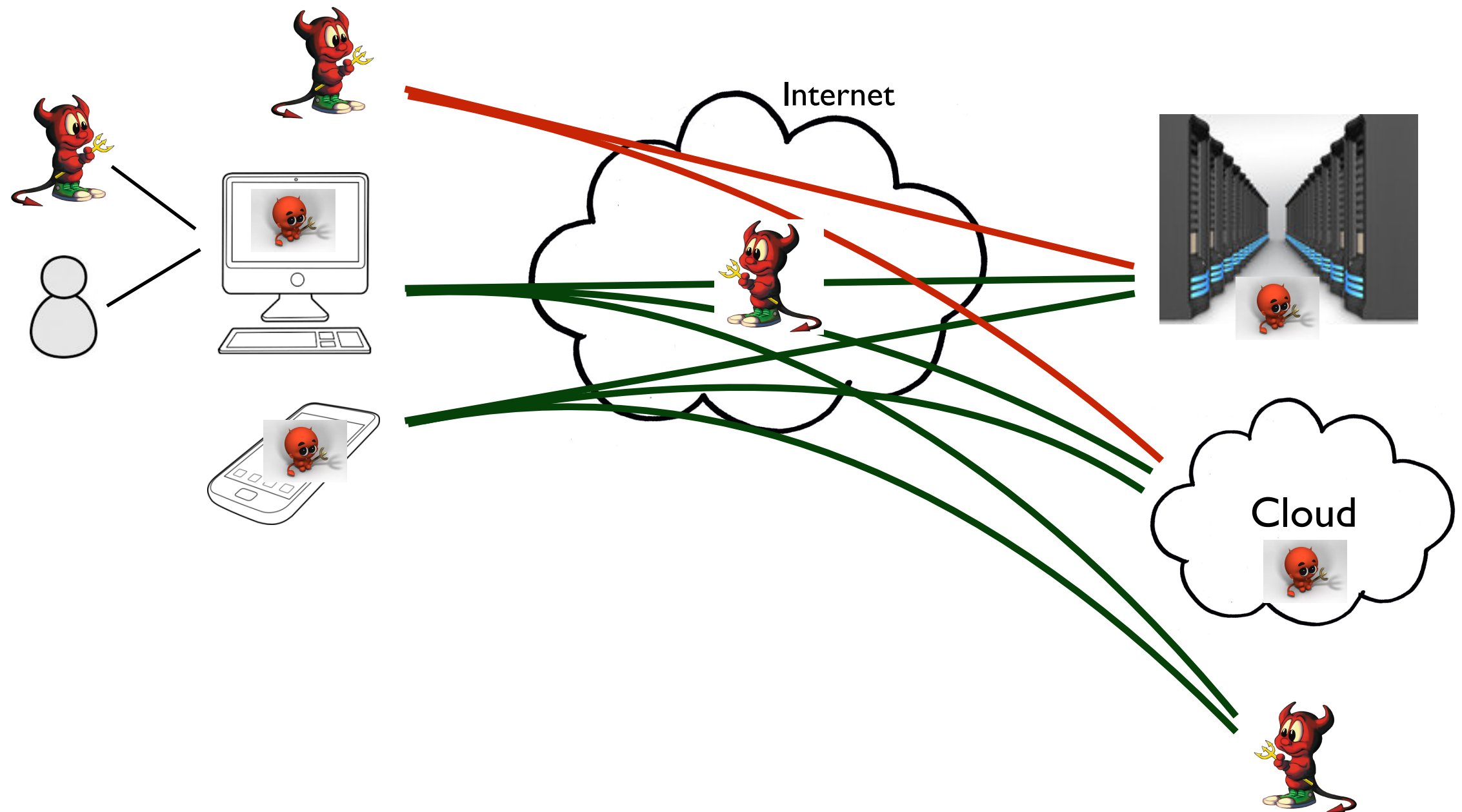
A U.S. cybersecurity firm that works with the government has evidence that Chinese government-linked hackers violated the cyber agreement reached between President Obama and Chinese President Xi Jinping less than a month ago.

The *Wall Street Journal* reported that CrowdStrike Inc. will announce Monday that some of its customers fell victim to unsuccessful cyberattacks that violated

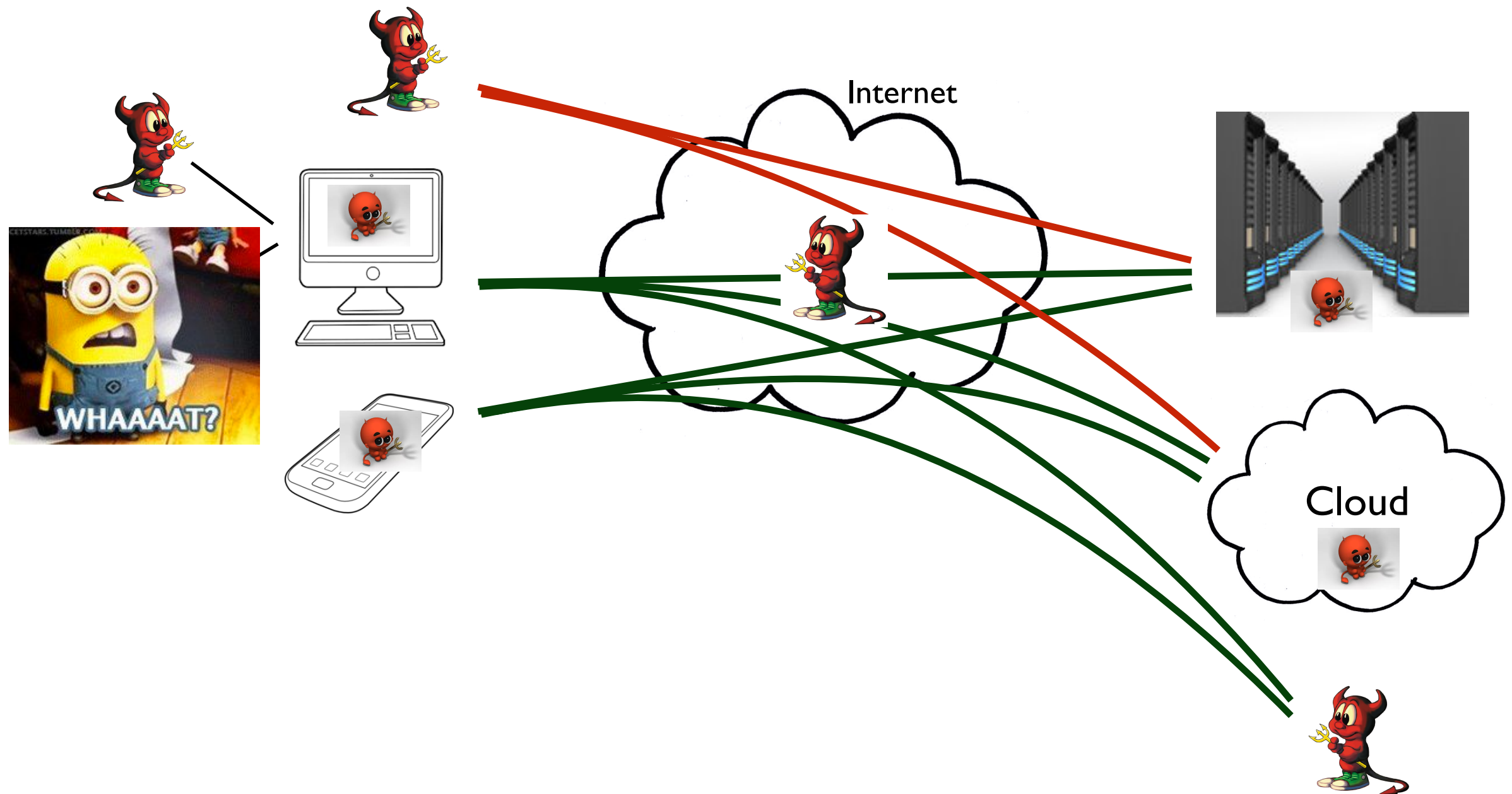
Increasingly Hostile Digital World



Increasingly Hostile Digital World



Increasingly Hostile Digital World



Problems on the Channel: *TLS ++*



Survey of the TLS vulnerabilities of the most popular websites

Attacks	Security			
	Insecure		Depends	Secure
Renegotiation attack	2.2% (−0.1%) support insecure renegotiation		0.8% (±0.0%) support both	94.1% (+0.3%) support secure renegotiation
RC4 attacks	0.4% (−0.1%) support only RC4 suites	12.3% (−1.4%) support RC4 suites used with modern browsers	40.6% (−1.8%) support some RC4 suites	47.1% (+3.1%) no support
BEAST attack (mitigated at client side with modern browsers)	N/A		89.2% (+1.2%) vulnerable if connect with old browser	N/A
CRIME attack	4.0% (−0.2%) vulnerable		N/A	N/A
Heartbleed	0.2% (±0.0%) vulnerable		N/A	N/A
ChangeCipherSpec injection attack	1.4% (−0.2%) vulnerable and exploitable		8.5% (−0.4%) vulnerable, not exploitable	88.9% (+0.4%) not vulnerable
POODLE attack against TLS (Original POODLE against SSL 3.0 is not included)	4.3% (−0.4%) vulnerable and exploitable		N/A	94.2% (+0.4%) not vulnerable
Protocol downgrade	32.6% (−1.0%) TLS_FALLBACK_SCSV not supported		N/A	54.4% (+1.3%) TLS_FALLBACK_SCSV supported

Problems on the Channel: *TLS ++*



Survey of the TLS vulnerabilities of the most popular websites

Attacks	Security			
	Insecure	Depends	Secure	Other
Renegotiation attack	2.2% (−0.1%) support insecure renegotiation	0.8% (±0.0%) support both	94.1% (+0.3%) support secure renegotiation	2.9% (−0.1%) no support
	0.4% (−0.1%)	12.3% (−1.4%)	40.6% (−1.8%)	47.1% (−2.1%)
Session 1: Transport Layer Security Transcript Collision Attacks: Breaking Authentication in TLS, IKE and SSH In response to high-profile attacks that exploit hash function collisions, software vendors have started to phase out the use of MD5 and SHA-1 in third-party digital signature applications such as X.509 certificates. However, weak hash constructions continue to be used in various cryptographic constructions within mainstream protocols such as TLS, IKE, and SSH, because practitioners argue that their use in these protocols relies only on second preimage resistance, and hence is unaffected by collisions. This paper systematically investigates and debunks this argument.				
Cha				N/A
attack	vulnerable and exploitable	vulnerable, not exploitable	not vulnerable	1.1% (+0.1%) unknown
POODLE attack against TLS (Original POODLE against SSL 3.0 is not included)	4.3% (−0.4%) vulnerable and exploitable	N/A	94.2% (+0.4%) not vulnerable	1.5% (±0.0%) unknown
Protocol downgrade	32.6% (−1.0%) TLS_FALLBACK_SCSV not supported	N/A	54.4% (+1.3%) TLS_FALLBACK_SCSV supported	13.0% (−0.3%) unknown

Problems in the End-Systems



The fact that something is small or embedded does not mean that it is not complex, does not process sensitive data and cannot be hacked!



Problems in the End-Systems



The fact that something is small or embedded does not mean that it is not complex, does not process sensitive data and cannot be hacked!



e.g. Target attack (70M customer names, credit/debit card numbers, expiration dates and CVVs)

Kaptoxa/BlackPOS malware runs on Windows POS terminals
(Home Depot Hit By Same Malware as Target)

Problems in the End-Systems

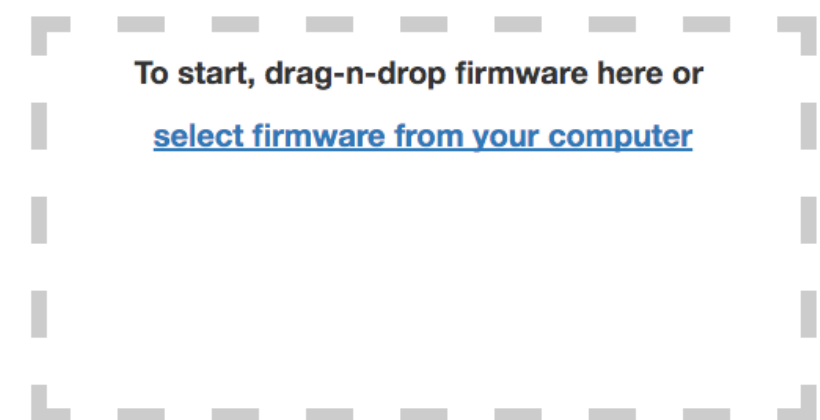
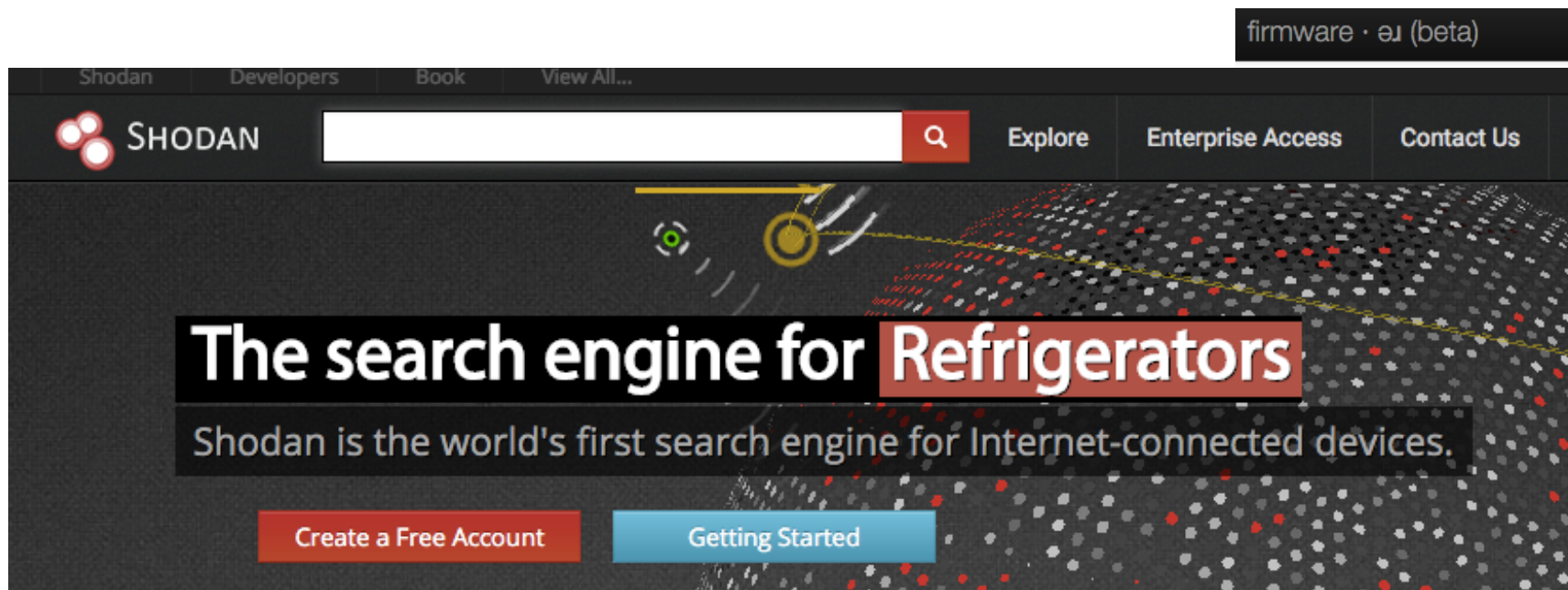


The fact that something is small or embedded does not mean that it is not complex, does not process sensitive data and cannot be hacked!



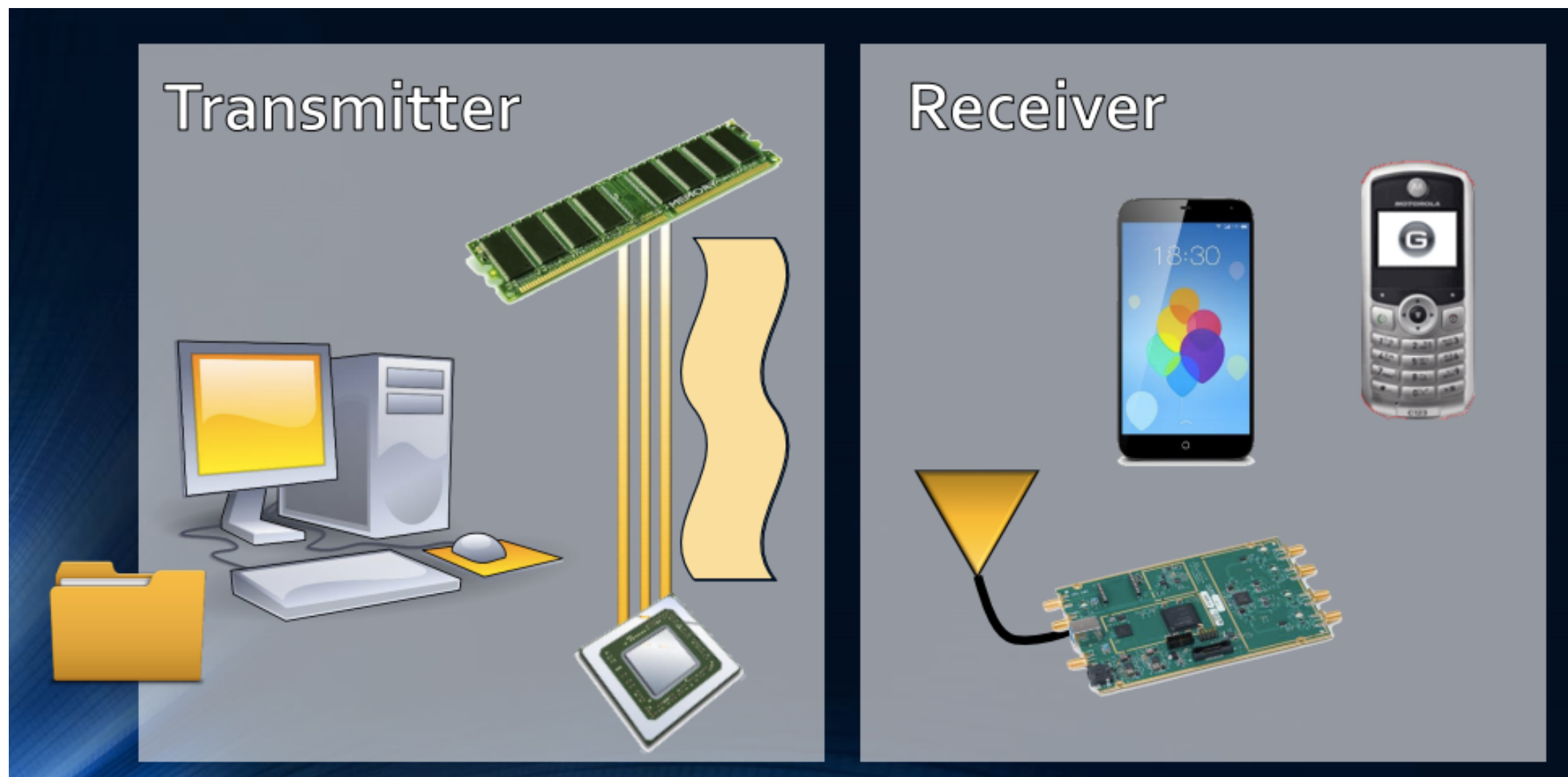
e.g. Target attack (70M customer names, credit/debit card numbers, expiration dates and CVVs)

Kaptoxa/BlackPOS malware runs on Windows POS terminals
(Home Depot Hit By Same Malware as Target)



Air-Gapped you say?

Data Exfiltration from Air-Gapped Computers over GSM Frequencies



Problems with Standards and Recommendations (and therefore Compliance ...)

Post-Snowden Cryptography ... (2013)

- Bullrun/Sigint: a program by the NSA to *"Insert vulnerabilities into commercial encryption systems"*, to *"influence policies, standards and specification for commercial public key technologies"* and to *"shape the worldwide commercial cryptography marketplace to make it more tractable to advanced cryptanalytic capabilities being developed by NSA/CSS"*. Example: Dual EC RNG
- Most companies and agencies follow the NIST recommendations

Problems with Standards and Recommendations (and therefore Compliance ...)

Post-Snowden Cryptography ... (2013)

- Bullrun/Sigint: a program by the NSA to *"Insert vulnerabilities into commercial encryption systems"*, to *"influence policies, standards and specification for commercial public key technologies"* and to *"shape the worldwide commercial cryptography marketplace to make it more tractable to advanced cryptanalytic capabilities being developed by NSA/CSS"*. Example: Dual EC RNG
- Most companies and agencies follow the NIST recommendations

	Undetectability	Lack of Conspiracy	Deniability	Ease of Use	Monitorability	Severity	Durability	Scale	Precision	Control
Lotus Notes	L	L	L	H	L	H	M	H	M	H
Dual EC	M	L	L	M	L	M	H	L	L	H
Debian PRNG	H	H	H	M	M	H	L	H	M	L
Heartbleed	H	H	H	M	M	H	M	H	M	L
Double Goto	H	H	H	M	H	H	L	M	M	L

Low L Medium M High H

Problems with Outsourcing of Data



Do we trust the cloud provider / jurisdiction?

Do we/they trust their employees?



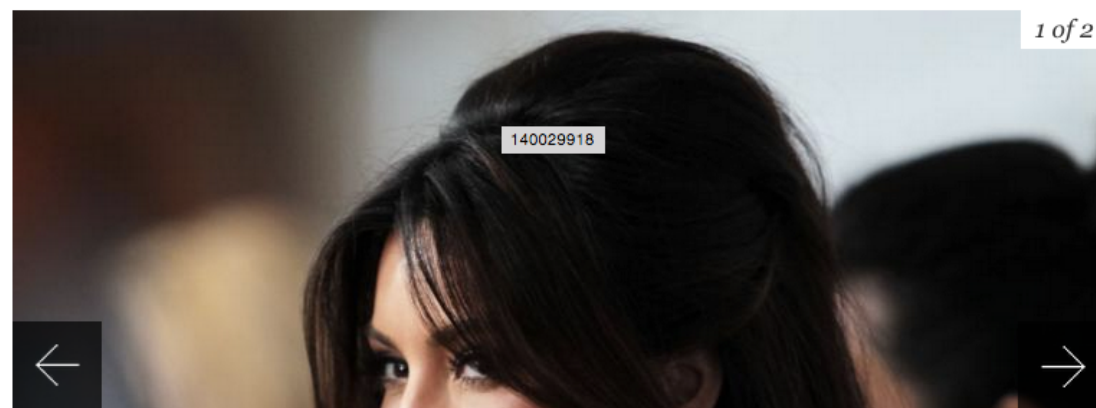
NSA's Grand Plan to Snowden-Proof Its Data Using the Cloud

Nextgov - 3 days ago

With NSA's widespread adoption of **cloud computing**, the spy agency may not have to.

iCloud Nude Leaks: 26 Celebrities Affected In The Nude Photo Scandal

By Amanda Remling [@AmandaTVScoop](#) on September 21 2014 1:19 PM EDT



Do we believe that they will secure our data / services better than we can?

We outsource more than data ...



- Virtually everything is ‘Virtualized’ and ‘Outsourced’
 - We store our data in the cloud (Dropbox, Google, ...)
 - We run servers in the cloud (systems, analytics) in the cloud
 - ***We run middleboxes (firewall, switches, VPN, ...) in the cloud***
 - ***We run our network services in the cloud***
 - ***We run our security services in the cloud (PKI, AAA, ...)***
 - ***We (will) run our (critical) infrastructure in clouds***

So how do we move forward and improve?

We should ...

We should ...

- Aim for security not compliance.
- Easy to be compliant, but not protect the system / data.

We should ...

- Aim for security not compliance.
- Easy to be compliant, but not protect the system / data.

We should ...

- Aim for security not compliance.
 - Easy to be compliant, but not protect the system / data.
- Be paranoid
 - Yes, this time “they” are really after you.

We should ...

- Aim for security not compliance.
 - Easy to be compliant, but not protect the system / data.
- Be paranoid
 - Yes, this time “they” are really after you.

We should ...

- Aim for security not compliance.
 - Easy to be compliant, but not protect the system / data.
- Be paranoid
 - Yes, this time “they” are really after you.
- Cover the basics
 - Most attacks still happen due to basic mistakes.
 - Very few truly sophisticated attacks (e.g., Stuxnet)

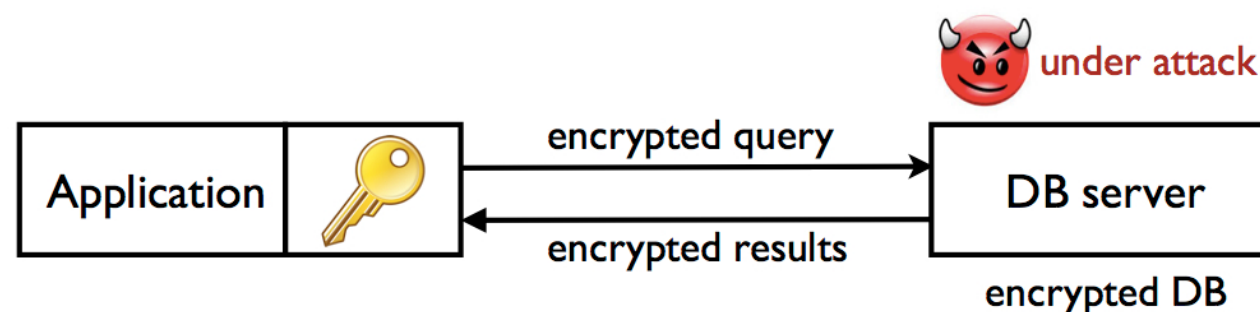
We should ...

- Aim for security not compliance.
 - Easy to be compliant, but not protect the system / data.
- Be paranoid
 - Yes, this time “they” are really after you.
- Cover the basics
 - Most attacks still happen due to basic mistakes.
 - Very few truly sophisticated attacks (e.g., Stuxnet)
- But then be creative
 - There are some cool security technologies out there - we should deploy them.
 - Diversify and enable defense in depth.

We Can Do Processing over Encrypted Data

Data **Confidentiality/Integrity**. Storing only encrypted data in the cloud? But how do we process / query encrypted data (efficiently)?

- e.g., CryptDB



```
mysql> select * from table_YYAXXURGXH;
```

database content for attacker with root access		
CKEWNKTJVoDET	WIWWHTQZBQo0PE	ISKNOMN0YSoAGG
12350108222818996780	14614281758989003640	;Li e0b(T 5
(0000701080[호 0f0000]0+0 0H"050\0000F0bv0T\$00-0400y0k000P02m000000"		00%60000'0<t4m0B0}010 0p0
		000c)0010c00000 0?mk00U0qi0\$00ne0C0000)0g"
15753844592636354766	16607196790268127689	R P-0i0c00u80]00}f Ea]
U]H=000 &00A0h080f0u000:0004000S0z00s00\0qZ006C0		y000U00b40e00j00c00)0V
13411675538537145840	12454649889312668381	000pK^v000000000I}0000
00R000"0`0J 00_000{tm00]009T0M		=f000g000I:0CR]000040og0+00g0+0z000M000{0(0000MKhiv000o0F000(0000C

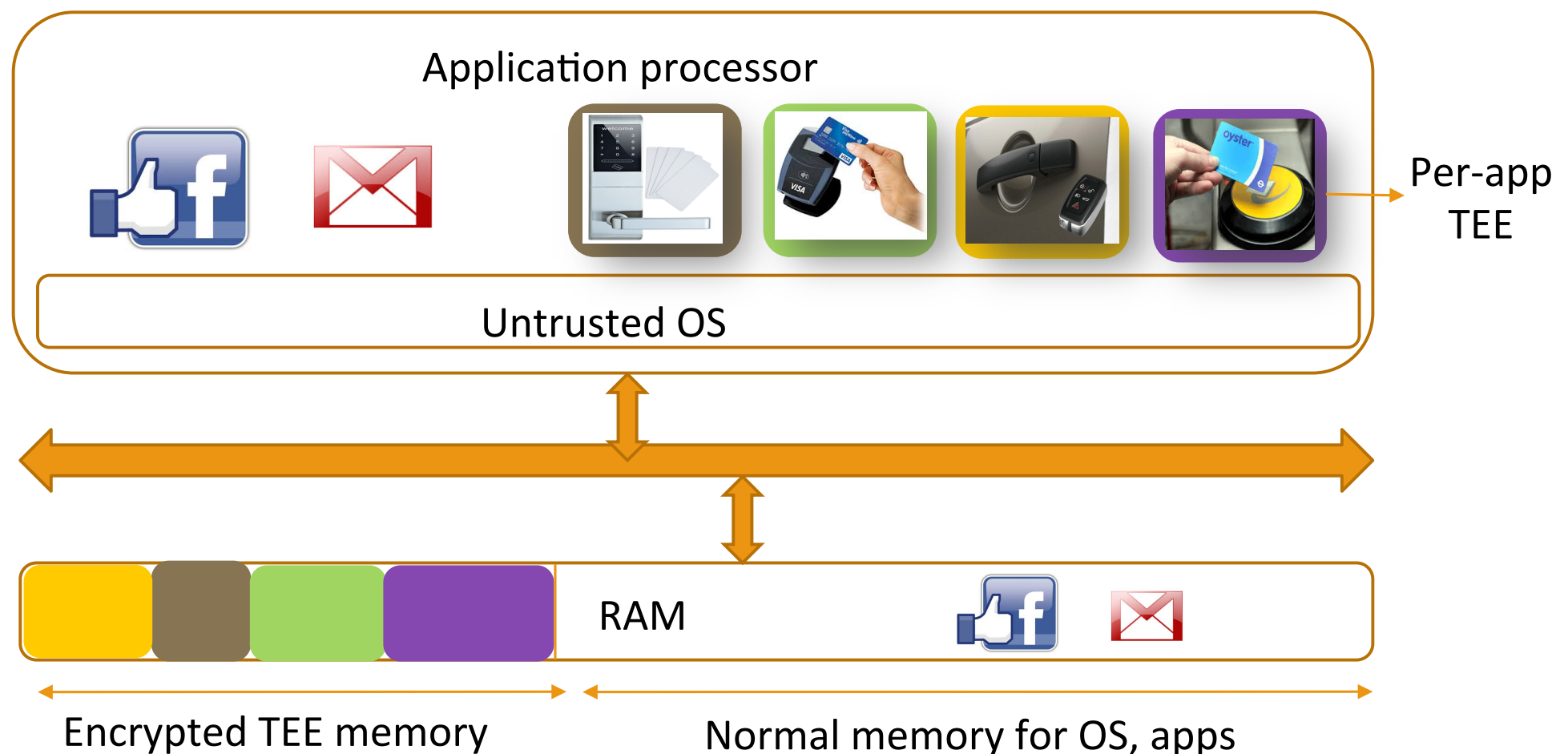
- Computing arbitrary functions over encrypted data
(*Fully homomorphic encryption*, C. Gentry, 2009)

We Can Isolate Sensitive Operations

It is not only about storage, what about running code?

(Launch/Run time Integrity) - Trusted Computing is back!

- e.g., Intel SGX promises isolated execution on an UNTRUSTED SYSTEM



And many many more ...

And many many more ...

- Privacy-preserving Machine Learning

And many many more ...

- Privacy-preserving Machine Learning
- Machine-Learning over Encrypted Data

And many many more ...

- Privacy-preserving Machine Learning
- Machine-Learning over Encrypted Data
- ...

And many many more ...

- Privacy-preserving Machine Learning
- Machine-Learning over Encrypted Data
- ...

And many many more ...

- Privacy-preserving Machine Learning
- Machine-Learning over Encrypted Data
- ...
- @ETH e.g.,
 - New Internet Architectures providing isolation and security. (<http://www.scion-architecture.net/>)
 - Usable 2nd factor authentication systems
<http://www.sound-proof.ch/>
 - Location-based Authentication and Access Control
www.securepositioning.com

Security@**ETH**zürich

Institute of Information Security (INFSEC), ETH Zurich

> 60 PhD students and senior researchers in security



www.informationsecurity.ethz.ch

www.zisc.ethz.ch

Basic Copyright Notice & Disclaimer

©2016 This presentation is copyright protected. All rights reserved. You may download or print out a hard copy for your private or internal use. You are not permitted to create any modifications or derivatives of this presentation without the prior written permission of the copyright owner.

This presentation is for information purposes only and contains non-binding indications. Any opinions or views expressed are of the author and do not necessarily represent those of Swiss Re. Swiss Re makes no warranties or representations as to the accuracy, comprehensiveness, timeliness or suitability of this presentation for a particular purpose. Anyone shall at its own risk interpret and employ this presentation without relying on it in isolation. In no event will Swiss Re be liable for any loss or damages of any kind, including any direct, indirect or consequential damages, arising out of or in connection with the use of this presentation.