De-anonymizing Users of French Political Forums

Dominique Bongard

0xcite LLC, Switzerland

dominique.bongard@0xcite.ch | @reversity | www.0xcite.ch

Agenda

- What is Gravatar and how does it work
- Privacy attacks on Gravatar
- Practical example of a French political forum

Part 1

What is Gravatar and how does it work

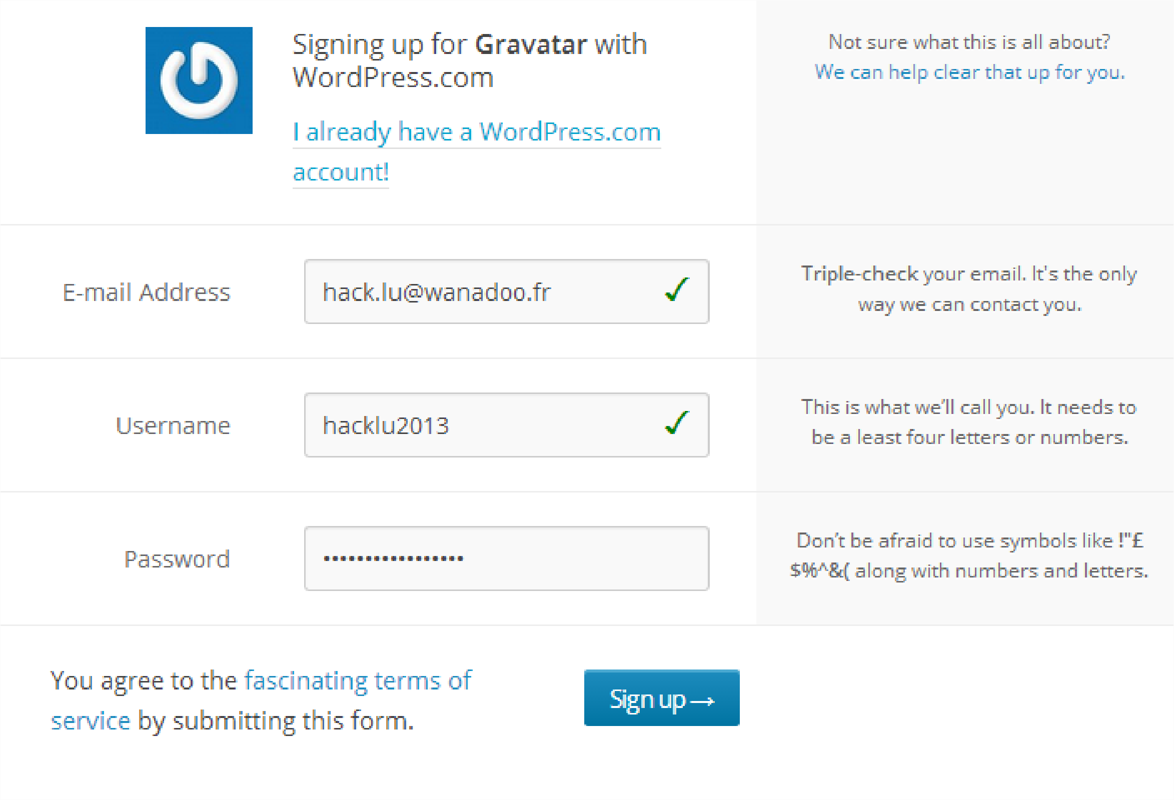
Globally Recognized Avatar

- Service which allows members of forums and blogs to automatically have the same profile picture on all participating sites
- Uses the MD5 hash of a person‘s email address as identifier
- Gravatar is owned by Automattic
- It is used by several major sites

Gravatar is used by several major services

WordPress
stackoverflow
github
SOCIAL CODING
hootsuite™
DISQUS
techdirt

How to create a Gravatar



Signing up for **Gravatar** with WordPress.com

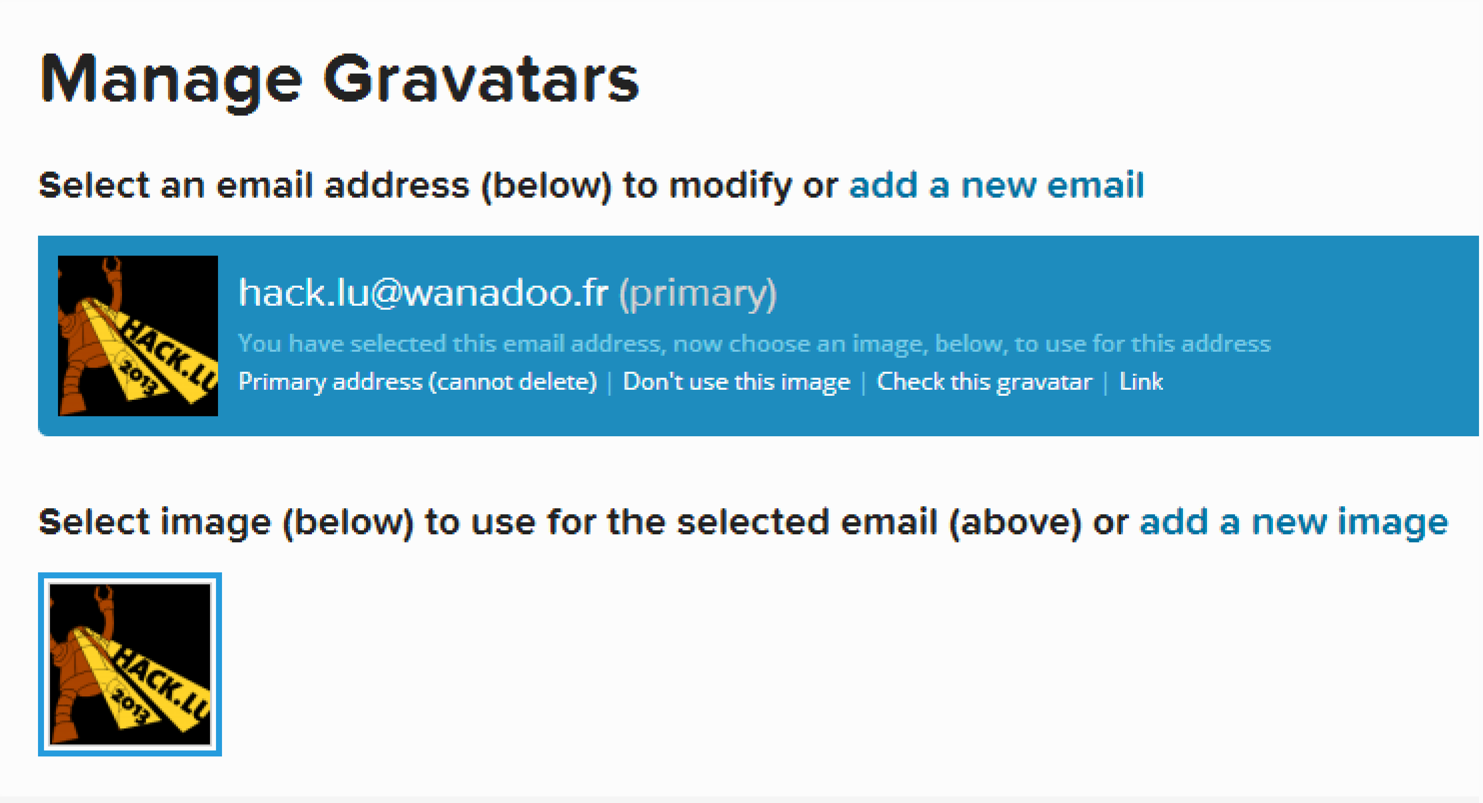
I already have a WordPress.com account!

Not sure what this is all about?
We can help clear that up for you.

E-mail Address	hack.lu@wanadoo.fr ✓	Triple-check your email. It's the only way we can contact you.
Username	hacklu2013 ✓	This is what we'll call you. It needs to be a least four letters or numbers.
Password	•••••••••••••••••	Don't be afraid to use symbols like !"£$%^&(along with numbers and letters.

You agree to the fascinating terms of service by submitting this form.

How to create a Gravatar



Manage Gravatars

Select an email address (below) to modify or add a new email

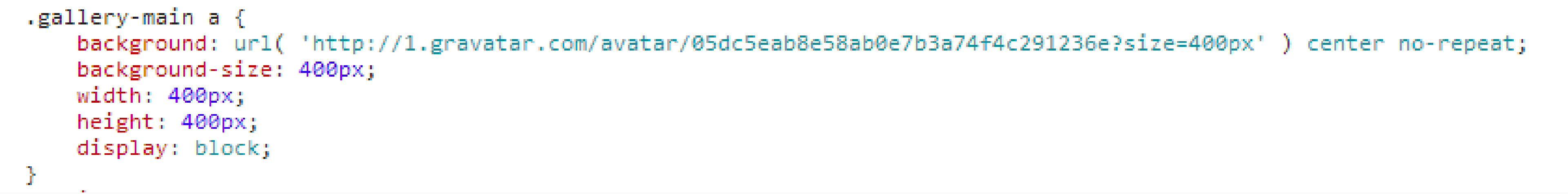
hack.lu@wanadoo.fr (primary)

You have selected this email address, now choose an image, below, to use for this address

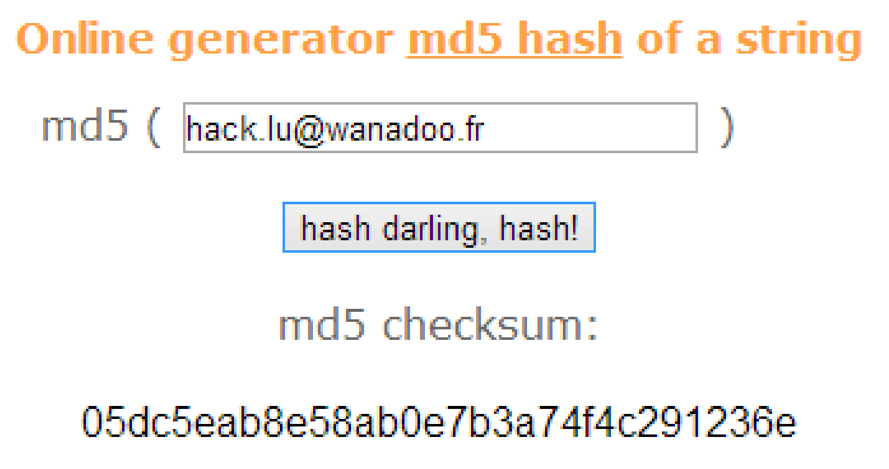
Primary address (cannot delete) | Don't use this image | Check this gravatar | Link

Select image (below) to use for the selected email (above) or add a new image

How Gravatars are displayed



```
.gallery-main a {
    background: url( 'http://1.gravatar.com/avatar/05dc5eab8e58ab0e7b3a74f4c291236e?size=400px' ) center no-repeat;
    background-size: 400px;
    width: 400px;
    height: 400px;
    display: block;
}
```



Online generator md5 hash of a string

md5 (hack.lu@wanadoo.fr)

md5 checksum:

05dc5eab8e58ab0e7b3a74f4c291236e

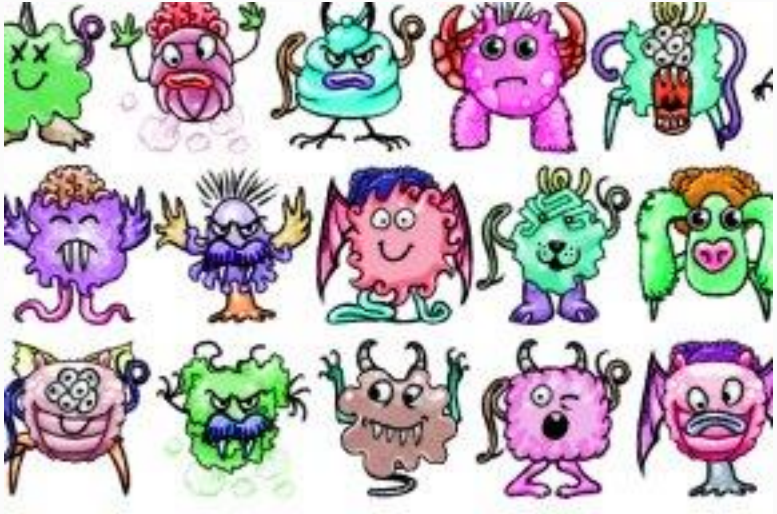
Gravatar Default Pictures

- Used for users who haven't registered an avatar
- Site administrators can also set a custom image
- **MD5 hashes are also displayed for users who didn't register with Gravatar!**

Gravatar Identicons

Gravatar Wavatars

Gravatar MonsterIDs

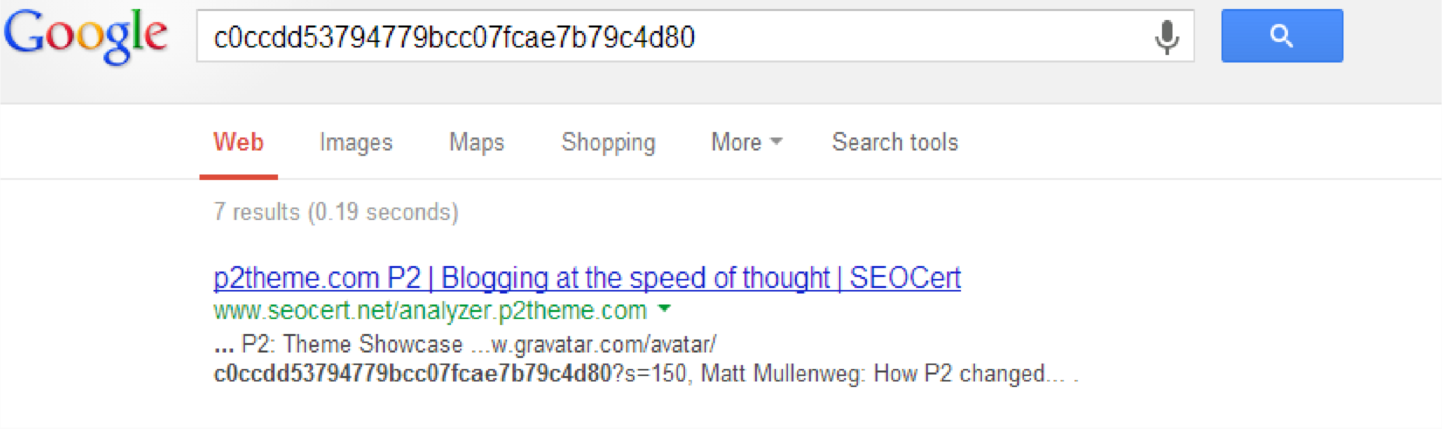


Part 2

Privacy attacks on Gravatar

Discover someone’s email address

matt.mullenweg@automattic.com	58f266c26cd28643c9f3ae42c858dfe5
mullenweg@automattic.com	9a68676b220b1357308951c3ce0b3911
matt@automattic.com	**c0ccdd53794779bcc07fcae7b79c4d80**



c0ccdd53794779bcc07fcae7b79c4d80

7 results (0.19 seconds)

p2theme.com P2 | Blogging at the speed of thought | SEOCert
www.seocert.net/analyzer.p2theme.com
... P2: Theme Showcase ...w.gravatar.com/avatar/
c0ccdd53794779bcc07fcae7b79c4d80?s=150, Matt Mullenweg: How P2 changed... .

Find the email address of commenters

- Use a password cracking software like Hashcat
- In 2008 Abell of developer.it recovered 10% of the email addresses of 80’000 stackoverflow.com users[1]
- Gravatar played it down with the following arguments
- An email address is not secret information
- The ressource tradeoff of cracking MD5 is not worth it for spammers
- There are easier ways to harvest email addresses

[1] http://www.developer.it/post/gravatars-why-publishing-your-email-s-hash-is-not-a-good-idea

Gravatar's defence (used to be in the FAQ)

> **"MD5 is plenty good for obfuscating the email address of users across the wire. if you're thinking of rainbow tables, those are all geared at passwords (which are generally shorter, and less globally different from one another) and not email addresses, furthermore they are geared at generating anything that matches the hash, NOT the original data being hashed. If you are thinking about being able to reproduce a collision, you still don't necessarily get the actual email address being hashed from the data generated to create the collision. In either case the work required to both construct and operate such a monstrocity would be prohibitively costly. If we left your password laying around in the open as a plain md5 hash someone might be able to find some data (not necessarily your password) which they could use to log in as you... Leaving your email address out as an md5 hash, however, is not going to cause a violent upsurge in the number of fake rolex watch emails that you get. Lets face it there are far more lucrative, easier, ways of getting email address. I hope this helps ease your mind."**

Rebuttal of Gravatar's stance

- Yes emails are longer than passwords but:
- A majority of addresses come from a few free providers
  - gmail.com, outlook.com, live.com, yahoo.com
- Email addresses are highly predictable
- First names, last names, dictionary words, hybrid
- GPU password crackers can try billions MD5 per second

Rebuttal of Gravatar's stance

- What if you are trying to stay anonymous ?

Part 3

Practical example of a French political forum

The current political context in France

French Government

- Presided by François Hollande since 2012
- Left wing social democrat (Parti Socialiste)
- Lowest satisfaction rate for a French President
- 23% in July 2013[1]

[1] http://www.rtl.fr/actualites/info/politique/article/la-cote-de-popularite-de-francois-hollande-au-plus-bas-7762951095

Controversial Reforms

- Same sex marriage
  - Huge protests (1 million persons in Paris according to the organizers[1])
- Surrogacy laws (GPA)
- Assisted Reproductive Technology laws (PMA)
- Voting rights for immigrants
- The recent «Leonarda» scandal

[1] http://www.lemonde.fr/societe/article/2013/05/26/des-manifestants-anti-mariage-homosexuel-interpelles-sur-les-champs-elysees_3417612_3224.html

Result

- Opponents vent anonymously on Internet forums
  - Especially far right wing forums

No constitutional right to free speech

- Forum posters can and do get sued for
- Racial hatred speech
- Inciting violent actions
- Libel
- You may lose your job for displaying you opinions
- It can also get you harrased or physically attacked

De-anonymisation of French political forums’ members

Several French political sites use Gravatar

- Members of such forums mostly use pseudonyms
- They have a high expectation of privacy
  - For the reasons detailed in the previous section
- Some savy posters register with disposable addresses
  - But ordinary people don‘t know how to create one and use their usual email address
  - Most forums require to verify the email address before posting

Risks caused by the recovery of addresses

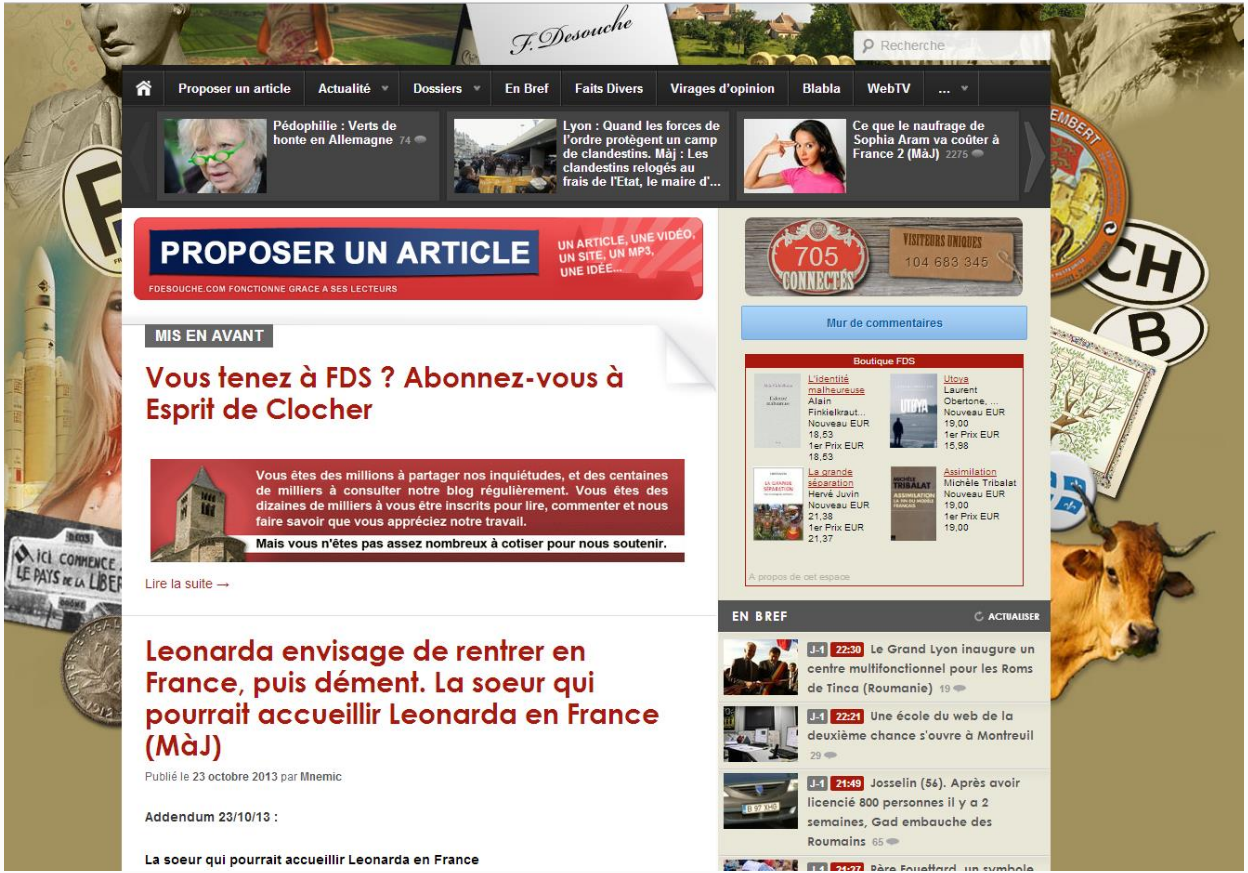
- The identity of many users can easily be discovered
  - Name, lastname in address
  - Same email address used to register an eBay profile
- The authorities can obtain the user's identity with a court order to the email provider
  - The email provider is often local (ISP) or with local activities (Google)
  - The site's administrator and the foreign web host would not have cooperated
- A political adversary can spearfish the users

Practical example of de-anonymisation

French political forums

- Until August, Fdesouche.com was using Gravatar
- On the far right wing of the spectrum while the current government is left wing
- Qualified as „very influent“ by „Arrêt sur images“ and
- Top1 political blog in France by „Le Figaro“
- The identity of its administator is suspected but not proven
- Lawsuits are regularly filed against the site because of comments posted by members

Fdesouche.com



Pédophilie : Verts de honte en Allemagne 74

Lyon : Quand les forces de l'ordre protègent un camp de clandestins. Màj : Les clandestins relogés au frais de l'Etat, le maire d'...

Ce que le naufrage de Sophia Aram va coûter à France 2 (MàJ) 2275

PROPOSER UN ARTICLE
UN ARTICLE, UNE VIDÉO, UN SITE, UN MP3, UNE IDÉE...
FDESOUCHE.COM FONCTIONNE GRACE A SES LECTEURS

705 CONNECTÉS

VISITEURS UNIQUES 104 683 345

Mur de commentaires

MIS EN AVANT

Vous tenez à FDS ? Abonnez-vous à Esprit de Clocher

Vous êtes des millions à partager nos inquiétudes, et des centaines de milliers à consulter notre blog régulièrement. Vous êtes des dizaines de milliers à vous être inscrits pour lire, commenter et nous faire savoir que vous appréciez notre travail.

Mais vous n'êtes pas assez nombreux à cotiser pour nous soutenir.

Lire la suite →

Boutique FDS

L'identité malheureuse
Alain Finkielkraut...
Nouveau EUR 18,53
1er Prix EUR 18,53

Utoya
Laurent Obertone, ...
Nouveau EUR 19,00
1er Prix EUR 15,98

La grande séparation
Hervé Juvin
Nouveau EUR 21,38
1er Prix EUR 21,37

Assimilation
Michèle Tribalat
Nouveau EUR 19,00
1er Prix EUR 19,00

A propos de cet espace

Leonarda envisage de rentrer en France, puis dément. La soeur qui pourrait accueillir Leonarda en France (MàJ)

Publié le 23 octobre 2013 par Mnemic

Addendum 23/10/13 :

La soeur qui pourrait accueillir Leonarda en France

EN BREF

ACTUALISER

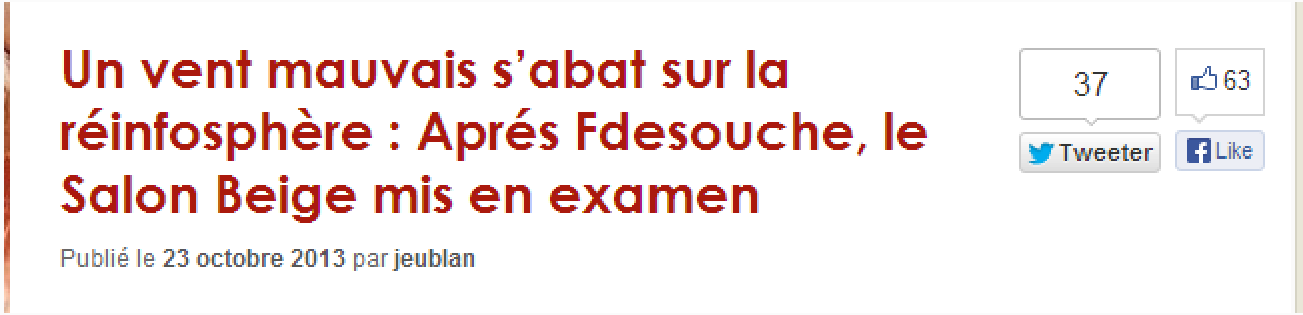
J-1 22:30 Le Grand Lyon inaugure un centre multifonctionnel pour les Roms de Tinca (Roumanie) 19

J-1 22:21 Une école du web de la deuxième chance s'ouvre à Montreuil 29

J-1 21:49 Josselin (56). Après avoir licencié 800 personnes il y a 2 semaines, Gad embauche des Roumains 65

J-1 21:27 Père Fouettard, un symbole

Fdesouche and Le Salon Beige prosecuted



Un vent mauvais s'abat sur la réinfosphère : Aprés Fdesouche, le Salon Beige mis en examen

Publié le **23 octobre 2013** par **jeublan**

Cracking of Gravatar hashes

- A custom crawler was written to acquire MD5 hashes
  - The site only kept 3 days of comments
  - Around 2400 different hashes were obtained
- A list of major email providers was established
- Cracking dictionaries were compiled
- Cracking rules were compiled
- A l33t cracking rig was built

Free webmail providers

gmail.com

aol.com

gmx.fr

hotmail.com

hotmail.fr

laposte.net

msn.com

live.fr

yahoo.fr

yahoo.com

ymail.com

outlook.com

bluewin.ch

voila.fr

ISP domains

aliceadsl.fr
club-internet.fr
libertysurf.fr
noos.fr
orange.fr
wanadoo.fr

cegetel.net
infonie.fr
neuf.fr
numericable.fr
sfr.fr

Disposable email addresses

get2mail.fr

yopmail.fr

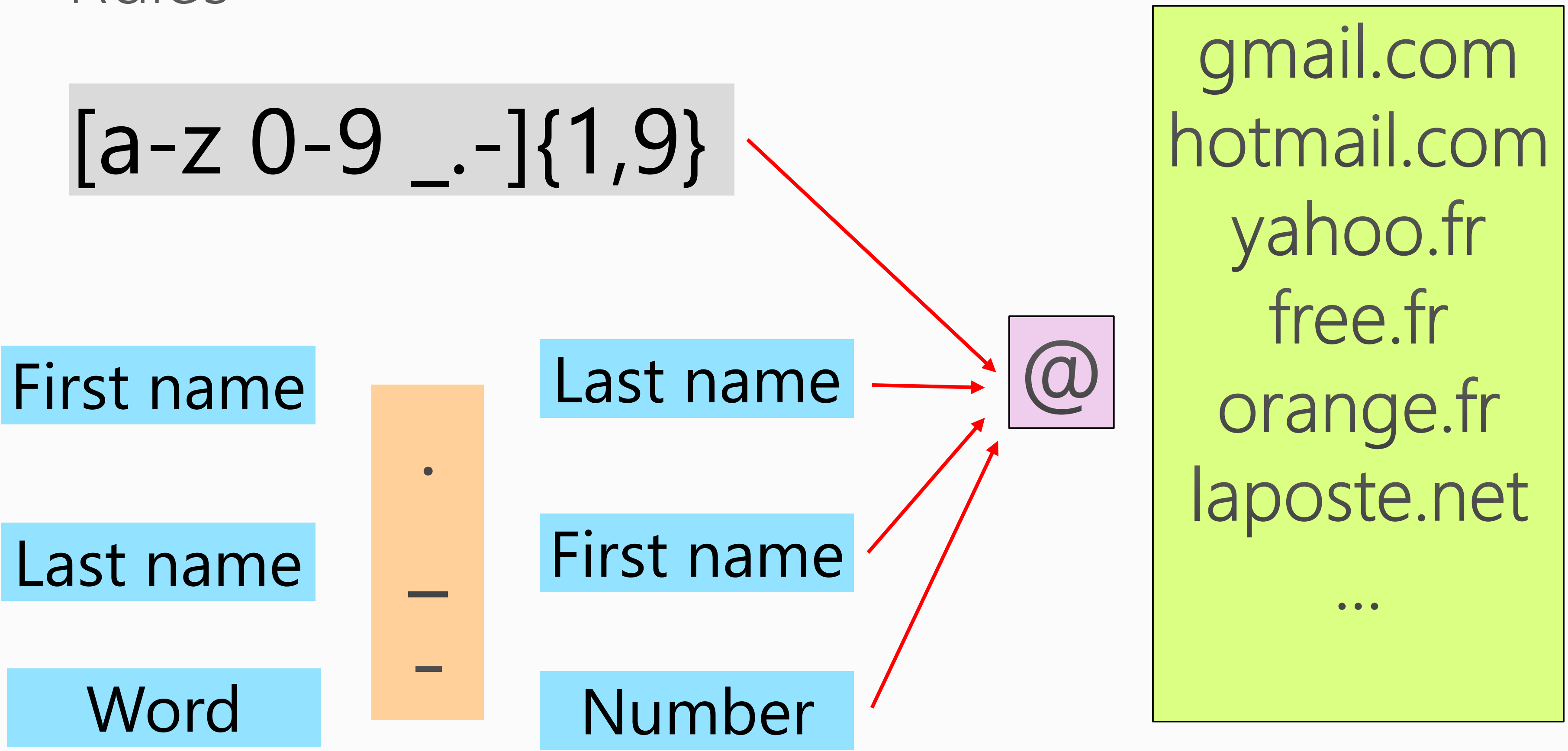
mailinator.com

yopmail.com

Dictionaries

- Fdesouche user nicknames
- French and English first names (Wikipedia, FB)
- French and English last names (Wikipedia, FB)
- French and English words
- French and English Wikipedia entries
- For sport team names, places…
- Numbers
- Current year and previous, usual birth years, ZIP codes

Rules



[a-z 0-9 _.-]{1,9}

First name

Last name

Word

.

_

-

Last name

First name

Number

@

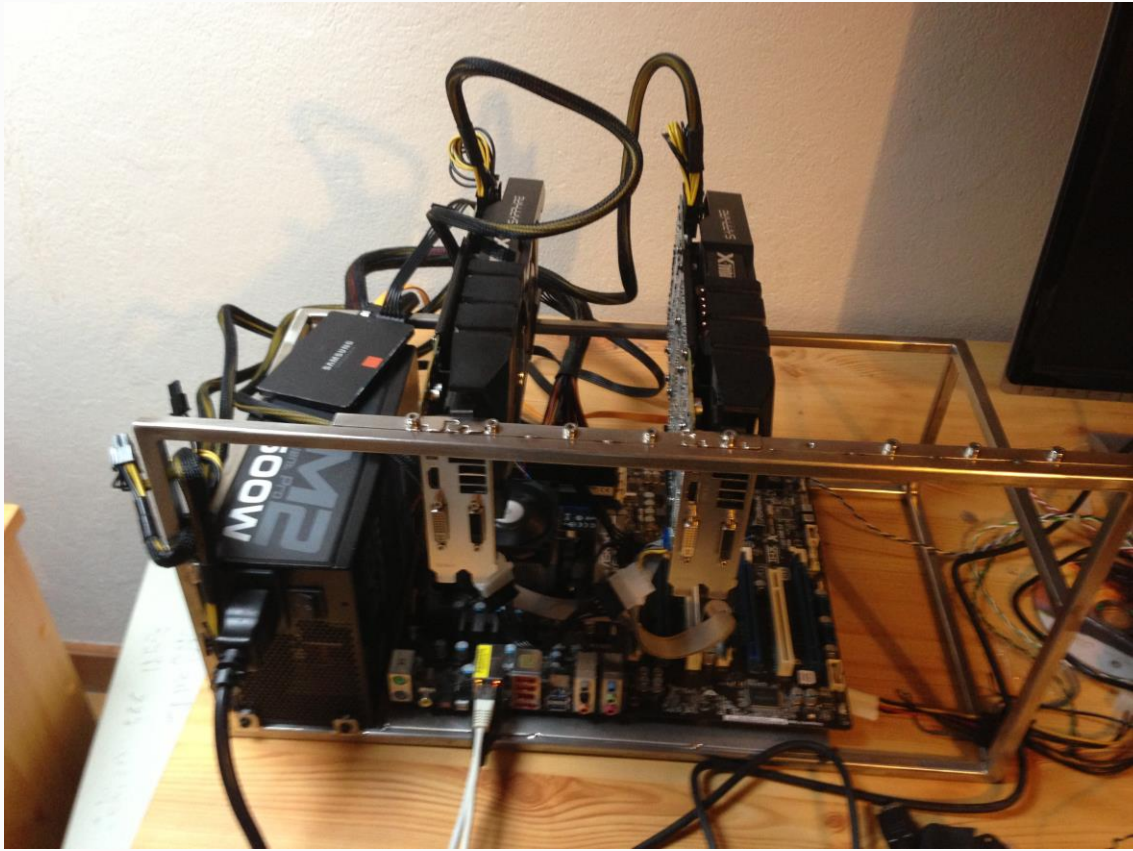
gmail.com
hotmail.com
yahoo.fr
free.fr
orange.fr
laposte.net
…

Cracking rig

- 4x AMD HD7970 GPU

Cracking rig

- 4x AMD HD7970 GPU
- Better pic but only 2 GPUs



Attack with oclHashcat 0.14

- **20%** of the email addresses recovered
- oclHashcat 0.14 was limited to 15 characters max
- «Good enough» for password cracking but not emails

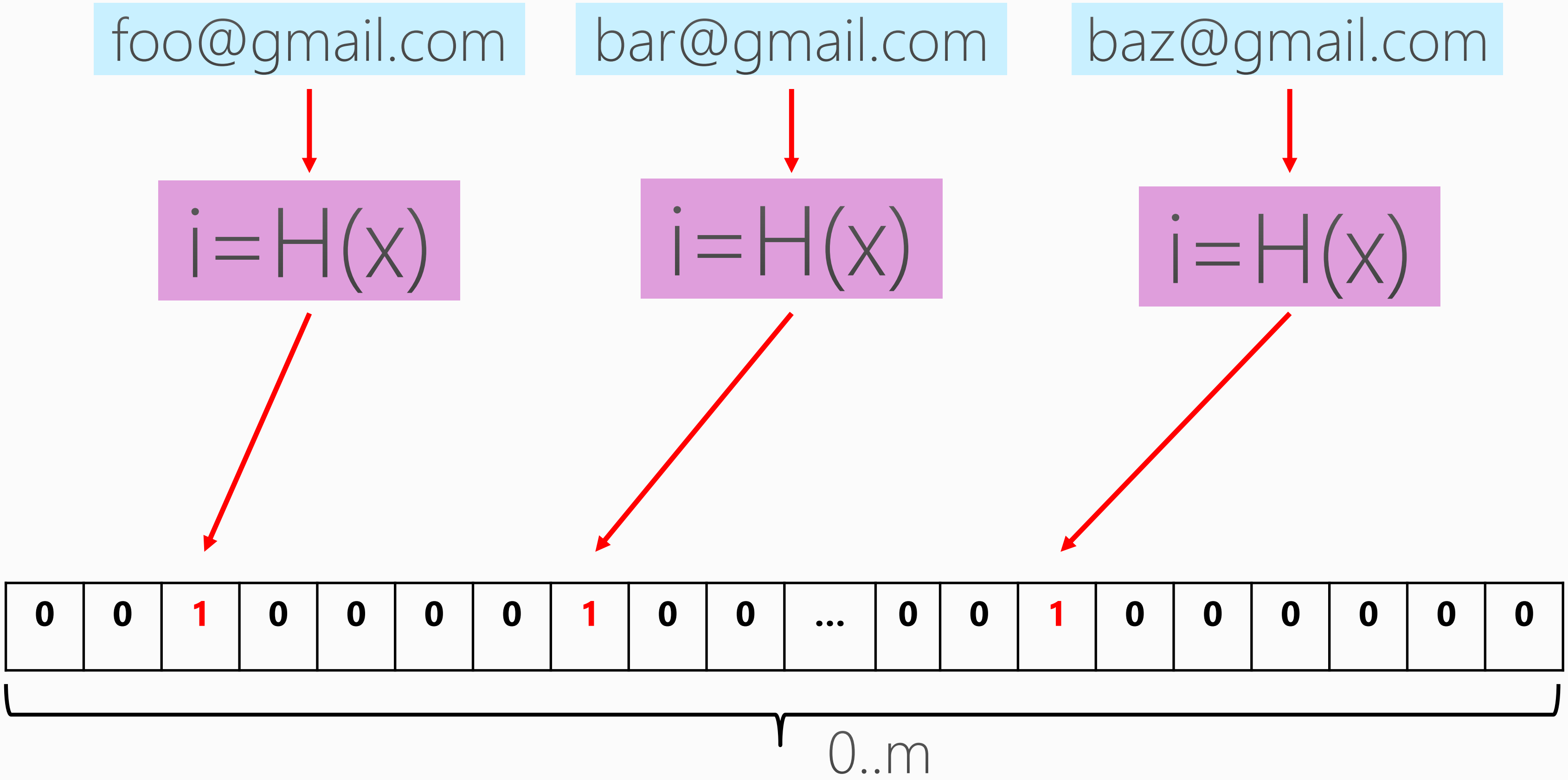
Attack with oclHashcat 0.15

- **45%** of the email addresses recovered
- oclHashcat 0.15 removes the 15 characters limitation
- Still cannot process all the custom email rules

Attack with custom cracking software

- Written in C++ with AMD's SDK
- Custom openCL kernel
- Candidate addresses are generated on GPU
- Hash pre-verification on GPU with Bloom filters
- Fast probabilistic set membership test
- Not especially optimized
- It's a very complex task

Bloom filter creation



foo@gmail.com

bar@gmail.com

baz@gmail.com

i=H(x)

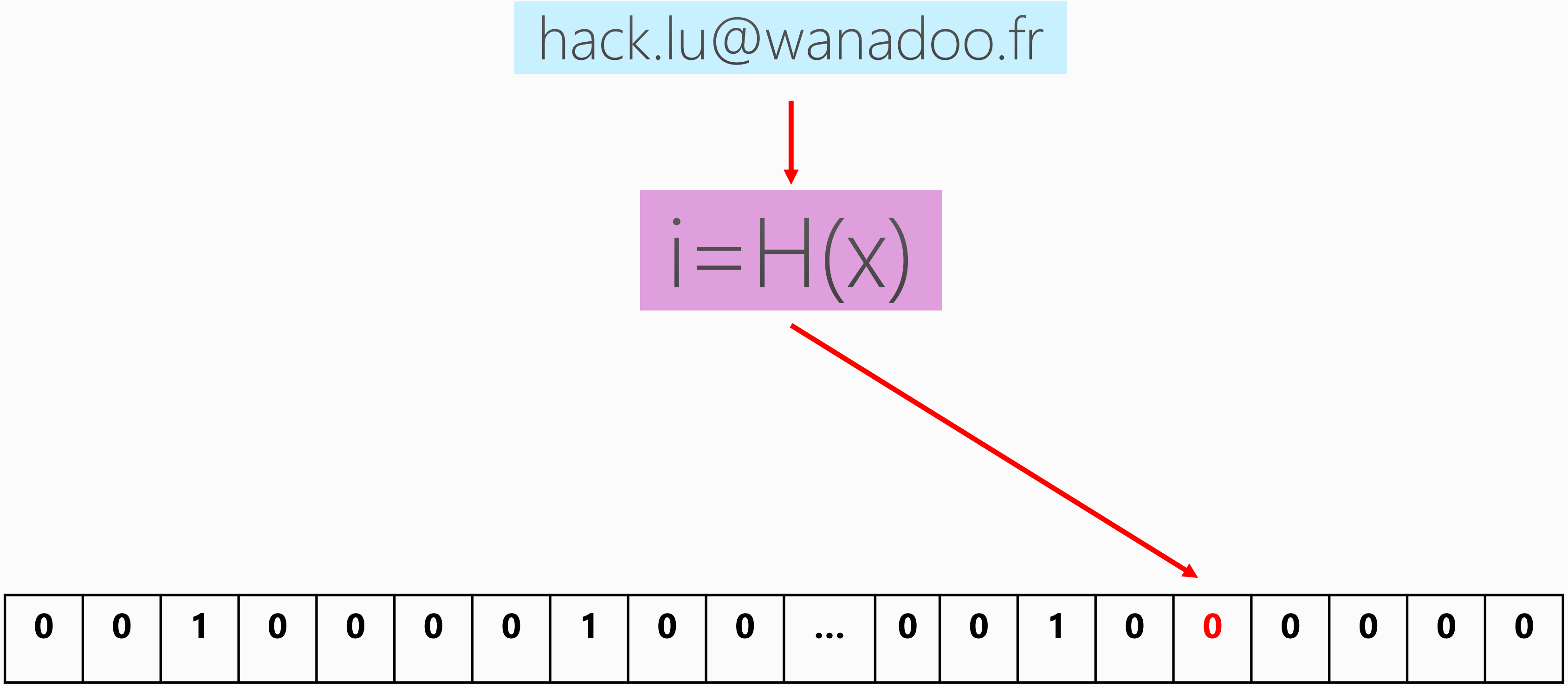
i=H(x)

i=H(x)

0	0	1	0	0	0	0	1	0	0	...	0	0	1	0	0	0	0	0	0

0..m

Bloom filter test



hack.lu@wanadoo.fr

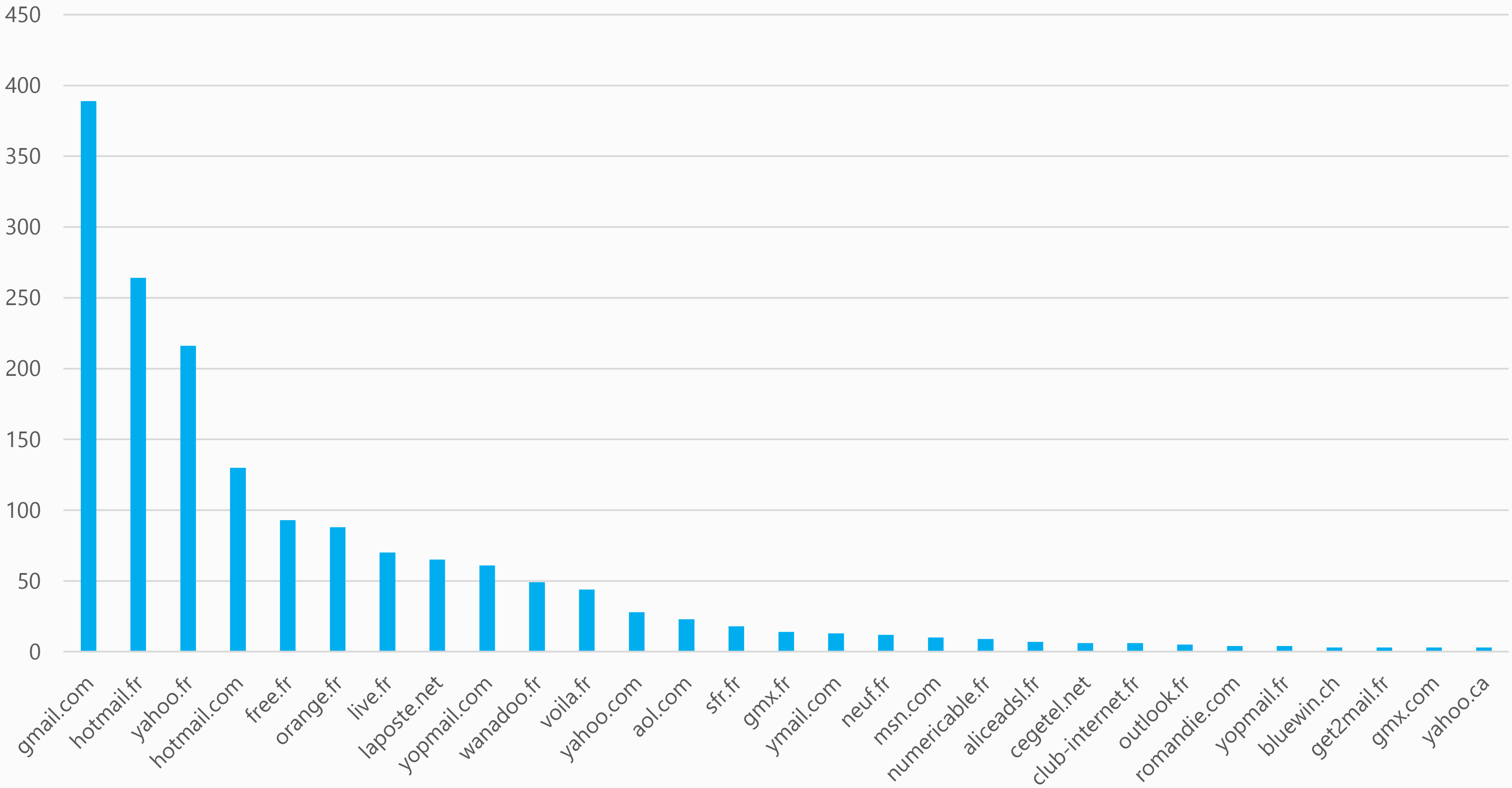
i=H(x)

0	0	1	0	0	0	0	1	0	0	…	0	0	1	0	0	0	0	0	0

Attack with custom cracking software

- **70%** of the email addresses recovered

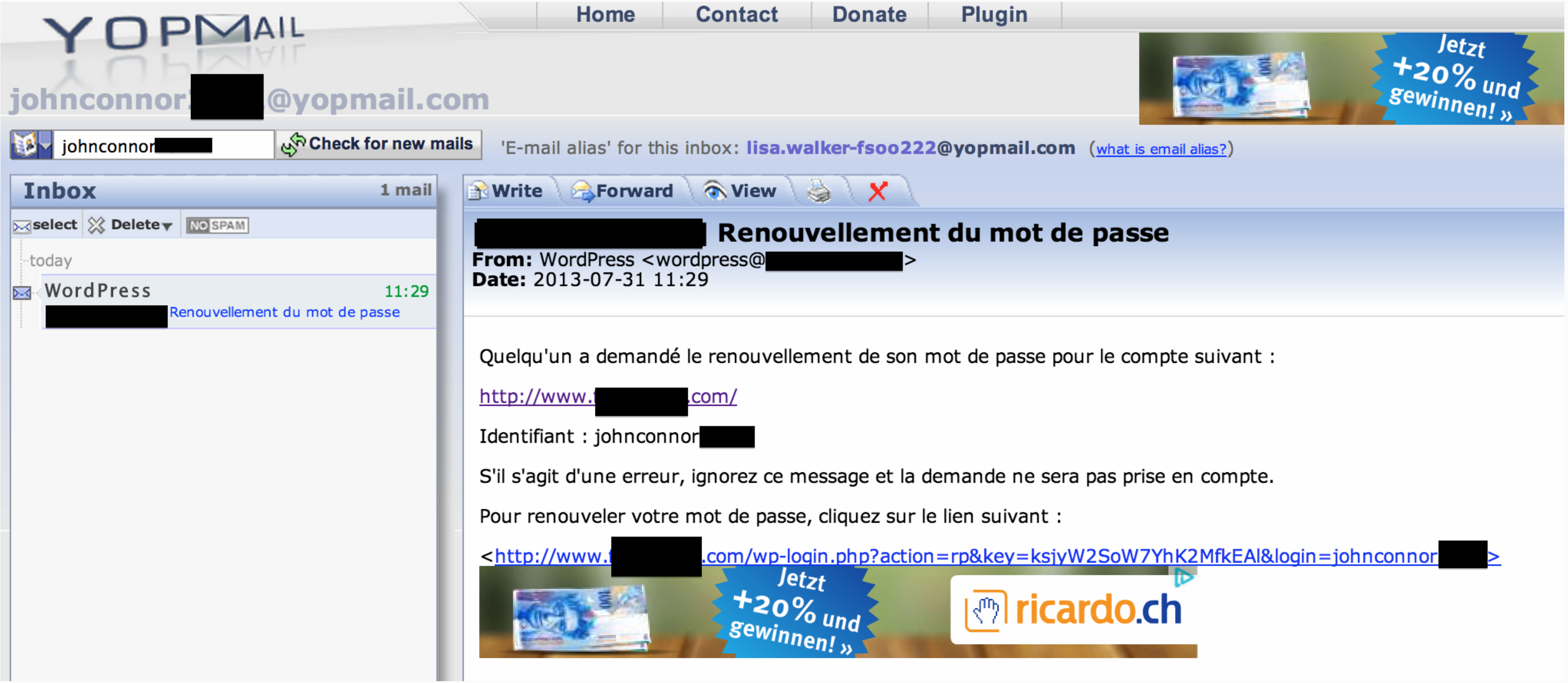
Distribution of domains



Some statistics on recovered addresses

Addresses containing probable real name:	**20%**
Addresses containing username:	10%
Addresses containing numbers:	30%
Addresses ending in numbers:	28%
Addresses containing punctuation:	26%
Addresses containing a dot:	20%
Addresses containing an underscore:	5%
Addresses containing a dash:	1.7%
Addresses with punctuation and numbers:	4%

New attack: hijacking accounts that use disposable emails



johnconnor[redacted]@yopmail.com

johnconnor[redacted]

'E-mail alias' for this inbox: **lisa.walker-fsoo222@yopmail.com** (what is email alias?)

WordPress 11:29
[redacted] Renouvellement du mot de passe

[redacted] Renouvellement du mot de passe
From: WordPress <wordpress@[redacted]>
Date: 2013-07-31 11:29

Quelqu'un a demandé le renouvellement de son mot de passe pour le compte suivant :

http://www.[redacted].com/

Identifiant : johnconnor[redacted]

S'il s'agit d'une erreur, ignorez ce message et la demande ne sera pas prise en compte.

Pour renouveler votre mot de passe, cliquez sur le lien suivant :

<http://www.[redacted].com/wp-login.php?action=rp&key=ksjyW2SoW7YhK2MfkEAl&login=johnconnor[redacted]>

Where are the missing 30% ?

- Anonymizers like «jetable.org»
- Which addresses are randomized and cannot be guessed
- Personal domains
- Less used email providers
- More complex email formats
- first.last<numbers>@domain.com

Acknowledgments

- I would like to thank the following people for their help
- Jens Steube (atom)
- Jean-Philippe Aumasson (veorq)
- Jeremi Gosney (epixoip)

Thanks for listening