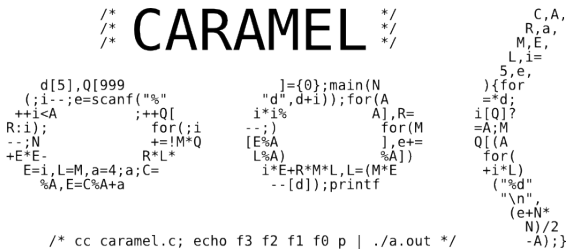
Crack me, I'm famous!

Cassage de phrases de passe célèbres

Hugo Labrande

SSTIC 2015, Rennes, 5 juin 2015



À propos

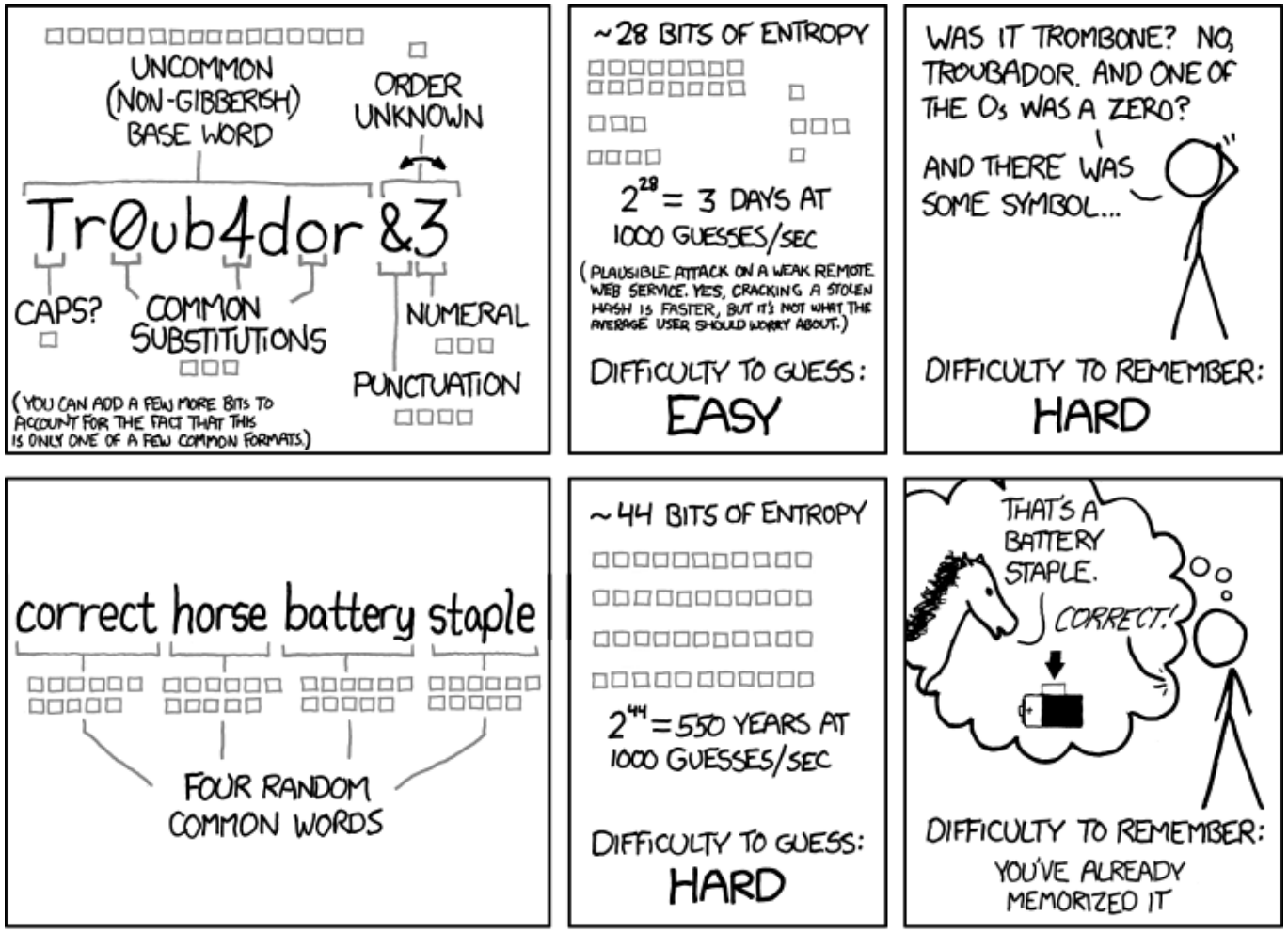
- Thésard en crypto (2ème année)
- Cotutelle entre la France et le Canada
- Crypto mathématique : théorie des nombres algorithmique + courbes elliptiques
  - Équipe CARAMEL (Nancy) : factorisation de RSA-768, cassage du log discret en petite caractéristique, Logjam, etc.
- Cette présentation (article court) : par curiosité sur mon temps libre
  - Quasi-n00b, jamais utilisé JtR ou sed ; ordi >6 ans
  - PoC : **un quasi-n00b peut casser 4 millions de phrases de passe célèbres**, 200,000 >16 chars
  - Dictionnaires compacts (65M candidats) et faciles à construire

1lomtjjzictcttsdkcs !

Mots de passe

- Mots de passe faibles : password, letmein, monkey, etc.
- Doivent être hachés + salés (sel unique) dans les BDD
  - Souvent pas salés ; ou ni l'un ni l'autre ! (RockYou, ViaPresse)
  - Failles de sécurité → BDD entières de hachés MD5 sur le Net
- Utiliser un password cracker (**John the Ripper**, ocl-Hashcat) :
  - Obtenir un dictionnaire : liste de tous les mots en anglais, liste de prénoms, etc.
  - Hacher chaque candidat ; dans la BDD ? Oui → craqué
  - Plus de candidats : **règles de modification** : « password » → « PASSWORD », « password0 », « pAsSwOrD », etc.
- Des millions d'essais par seconde (systèmes dédiés, GPU)

Phrases de passe (XKCD)



Phrases de passe

- Plus de possibilités et pas plus difficile à utiliser ?
  (controversé : plus de typos, toujours dur à mémoriser)
- Ou : phrase $\xrightarrow{\text{moyen mnémotechnique}}$ mot de passe
- Attaques connues :
  - Par combinaison : concaténer des mots du dictionnaire
  - Exploitation de motifs (« grammaire »), Tmesis
  - Chaînes de Markov (2005) : le 1er mot est « vélo », quelle proba que le 2ème soit « électrique » ou « moteur » ou « sapajou » ? $\rightarrow$ Générer des phrases probables

Dustin & Young (2010)

- Collecter autant de « vraies » phrases que possible
  - Wikipedia, IRC, projet Gutenberg (15,000 livres), Facebook, commentaires Youtube...
  - + les sous-phrases : « it was a dark and stormy night » → « it », « it was », « it was a »... « was », « was a », « was a dark »... → « darkandstormy », « stormynight », etc.
- 1,3 milliards de phrases, plusieurs TB de données
- Résultats :
  - « Des milliers » de mots de passe
  - (dicos pas publiés, quelles BDD utilisées ? → **peu d'info sur l'efficacité**)
  - crotalus atrox, eastofthesunwestofthemoon, Ph'nglui mglw'nafh Cthulhu R'lyeh wgah'nagl fhtagn !

Phrases célèbres

- Généralement découragé, parfois non : Le Monde, LifeHacker, Google « Ask Hamlet » (2bon2btitq)
- Kuo & al (2006) : « choisissez une phrase, prenez la première lettre de chaque mot »
  - Résultat : la phrase est célèbre pour $>$**50%** des users
- Dans Dustin & Young : grosse proportion de phrases connues
  - harderbetterfasterstronger, reesespeanutbuttercup, setthecontrolsfortheheartofthesun, turlututuchapeaupointu
  - Tout Gutenberg... pour ça ?
  - Bon, quand même : Ok thanks I get it now, ilovejustinbieber2305, ourweddinganniversary, Gates is an utter maggot.
- Faisons un dictionnaire de phrases connues

Merci Wikimedia Foundation

http://dumps.wikimedia.org

- Titres des articles de Wikipedia (13 millions)

Script

```
cat wikiarticles | tr "_" " " | sed "s/ (.*)$//;" >file1
cat file1 | sed "s/The //; s/La //; s/Le //;" >file2
sort -u file1 file2 >articles
```

- Citations sur Wikiquote (2 millions)
  - Astérisques, liens, noms de persos/auteur, etc.
  - Conventions moins homogènes que sur Wikipedia

Script

```
cat wikiquote | grep "^*[^*]" | sed "N; s/{{sic}}//g; P; D"
              | grep -v "{{" | etc.
```

Wikipedia et délimiteurs

Extraire les phrases célèbres d'un grand corpus ?

- Quotes are between quotes : si c'est entre « », gras ou italique, grandes chances que ça soit une phrase connue.
- $\rightarrow$ télécharger tout Wikipedia, garder ce qui apparaît entre ces délimiteurs
  - Script simple (remplacer ’”’’, ’’’, ’’ par " ; puis `grep '"'`)
  - Temps de traitement plutôt long (des dizaines de gigas !)
- 56 millions de candidats

Autres sources

On rajoute d'autres sources libres (GNU) :

- RapDict (wiki rap et culture hip-hop)
  - Beaucoup de spam... pas grave ! !
- FOLDOC (dictionnaire de vocabulaire d'informaticiens)

Ça marche sûrement aussi avec :

- Urban Dictionary, IMDB, Twitter, forums
- Mais récupérer tout ça $\simeq$ violation des ToS

Au final :

- 65 351 745 candidats
- Tout le monde peut en faire autant
  - Sources utilisées sont libres
  - Scripts faciles à écrire & rapides

Règles dans JtR

Dans John the Ripper :

- `/?p @?p >3 l $[2!37954860.?]`
- « Trois,quatre » → « troisquatre2 », « troisquatre ! », « troisquatre3 » etc.

Transformations pertinentes pour les phrases de passe :

- Enlever les espaces / les remplacer par des underscore
- Majuscules, minuscules, première lettre majuscule
- Supprimer la ponctuation, la ponctuation sauf ? !., les symboles
- Ajouter un chiffre/caractère à la fin

Toutes les combinaisons : 94 règles, que l'on rajoute aux règles **Wordlist** par défaut de JtR.

Les chiffres

- Temps de calcul : quelques semaines sur un vieux CPU (6 ans d'âge, 2GHz)
  - Faible puissance, pas de GPU : les pros mettront quelques heures
- Résultats :
  - 4 266 871 mots de passe de la base Korelogic
  - 1.2M basés sur des mnémoniques, mais seulement 300 >9 char
  - Plutôt orthogonal aux wordlist classiques (jusqu'à 40% d'originaux par rapport à GDictv2)

Contributions individuelles

- FOLDOC : 8b9e56c9b323f49036cf2d6219ccab83 : Perfect programmer syndrome
- RapDict : fcee6181358f14c75c97de615a7310f3 : Pudgee tha phat bastard
- Spam de RapDict : dd84c70ce416d07c2f4a58bffc17dd00 : penisenlargementthatworks
- Titres d'articles Wikipedia : f5b658c8f50e2218d025150e79a0578c : Lamborghini Murciélago LP 670-4 SuperVeloce
- Wikiquote : a2b8b746a12e54b7ffaeb956daf51f03 : When shall we three meet again ?
- Délimités dans Wikipedia : 1496a5a54e45a4e773a4aa3e4f57a5b3 : I could tell you but then I'd have to kill you

Vous les reconnaissez ?

fe7ab5639f5d0a1694d644fb488ab805 : mercurehotelkongresschemnitz
f7b2dc03c4c2d702fe0d371fa7f204b2 : aliciaenelpaisdelasmaravillas
0017d2da75ce52157ad1e11cbd9c4322 : teacher leave those kids alone
2a6df0eb9e9b481dbdc5229dbf8b044e : What the hell are you talking about ?

01c1bd4fe9e12be5e2821c5b1d89720a : Harry Potter and the Deathly Hallows Part 2 (43 char)
280f3868811d57d28c5cc312b6ff236f : Ghost in the Shell : S.A.C. Solid State Society (46 char)
cc76b63a63f70e63443229a67fdffe51 : superlativeextruderinterlocked-technologyexclusivenexus (54 char)
f145a74da3068a69897d93cadfc6a645 : judge a man by his questions rather than by his answers (55 char : record du monde ?)

Mnémoniques

Pour les mnémoniques, plus difficile :

- Créer un script qui prend la première lettre de chaque mot
  - En laissant la ponctuation, la casse et les chiffres intacts
- Utiliser « 94 règles+Wordlist »
- Beaucoup de « bruit » dans les résultats...
  - « A B C D E F », « C.R.E.A.M. », etc
  - Des fois, les initiales forment un mot (= bcp de résultats)
  - Des mots de passe bruteforçables (« Dab57 », « 7asdfg »)
  - $\rightarrow$ utiliser ce dico après les dicos standards
- ... mais aussi des résultats intéressants !
  - grep les plus longs, on retrouve à la main la phrase du dictionnaire qui correspond

Mnémoniques

- CSinmactaiat.
- mvpumelvqlveum
- tbtitbtwtbtewb
- wowowhiyhaynwa
- uuddlrlrababsS
- tamtihaehtadoiyp
- 1lomtjjzictcttsdkcs (19 char)

Mnémoniques

- CSinmactaiat. : « Computer science is no more about computers than astronomy is about telescopes. » - Dijikstra.
- mvpumelvqlveum : « Más vale perder un minuto en la vida que la vida en un minuto » - proverbe espagnol.
- tbtitbtwtbtewb : « The best there is, the best there was, the best there ever will be » - Bret "The Hitman" Hart.
- wowowhiyhaynwa : « Walk on, walk on, with hope in your heart And you'll never walk alone » - fans du Liverpool FC.
- uuddlrlrababsS : « Up Up Down Down Left Right Left Right A B A B select Start » - Konami Code
- tamtihaehtadoiyp : « There are more things in heaven and earth, Horatio, than are dreamt of in your philosophy. » - Hamlet.
- 1lomtjjzictcttsdkcs (19 char) : « Litwo ! Ojczyzno moja ! ty jestes jak zdrowie ; Ile cie trzeba cenic, ten tylko sie dowie, Kto cie stracil... » - ouverture de Pan Tadeusz, poème épique polonais.

Conclusion

Casser une phrase de passe célèbre c'est **super facile** grâce à Wikipédia

- Sources libres, scripts simples → dico de 65M+ de phrases célèbres
- Casse 4M+ mots de passe (des très longs) + des mnémoniques
- On peut raffiner
  - Plus de sources (non-libres)
  - Règles plus subtiles (fautes d'orthographe, l33t, autres moyens mnémotechniques...)

Merci de votre attention !

Questions ?

aiwsittdrcau
iejtmbdtiaeof
LDWMJOMTTKTP
fcpsitsgepgepgep
SINTIAUWHIAPLU
YOMISMAFUIMIRUTA
149162536496481100
ANOTHERONEBITESTHEDUST
IDONTWANTTHISANYMORE
WHATTHEFUCKISGOINGON
FOREIGNEXCHANGESTUDENT
NOWISTHEWINTEROFOURDISCONTENT
CHANGEYOURPASSWORD
IWANTTOGETMARRIED
athousandsplendidsuns
bilawalbhuttozardari
brplvidyalaya
lefruitvert
HOMOPHOBIAISGAY
hppavilionze4900
scrumpyjackcider
I'm so sorry.
alleaussermir
BackToFeature8
honraatuspadres
elcaballodesantiago
pikachuetpichu
SpottedMarshFrog9
Weltohnelicht

appellationdoriginecontrolee
assassinationofosamabinladen
amievergonnaseeyourfaceagain
CrockfordsClericalDirectory
Maxwell-boltzmann statistics
peanutbutterandjellysandwich
The other side of the tracks
call of duty modern warfare 3
Nobody expects the spanish inquisition !
les poules du couvent couvent
The Legend of Zelda : Hyrule Historia
Wikipedia, l'encyclopedie libre et gratuite
The quick brown fox jumps over the lazy eog
Pa's wijze lynx bezag vroom het fikse aquaduct
09 F9 11 02 9D 74 E3 5B D8 41 56 C5 63 56 88 C0
Cantami o diva del pelide Achille l'ira funesta
I could tell you but then I'd have to kill you
vocediperlablu7
sesameouvretoi6
DICTATORSHIPOFTHEPROLETARIAT
WINNER_WINNER_CHICKEN_DINNER!

youarelikelytobeeatenbyagrue
solongandthanksforallthefish
ithinkweshouldseeotherpeople
teacher leave those kids alone
All Your Base Are Belong To Us
Be sure to drink your ovaltine
freeyourmindandyourasswillfollow
I like big butts and I cannot lie
themagicwordsaresqueamishossifrage
The quick brown fox jumps over the lazy dog
putthatinyourpipeandsmokeit !
the cake is a lie
wmsteoopaaffwc
exercise, left as an
getcashrightnow
bestbadcreditpaydayloans
worldchampionshipwrestling
itsonlyrockandrollbutilikeit
breakonthroughtotheotherside
ithinkweshouldseeotherpeople
panneauxsolairesthermiques
saygoodnighttothebadguy
perfect programmer syndrome
realfriendshelpyoumovebodies
cashruleseverythingaroundme
letthelittlechildrencometome
Masque of the Red Death, The
WHAT THE—

Ordre des règles

Règles les plus utiles avec des passphrases :

- Supprimer espaces + minuscule (déjà dans JtR !)
- Supprimer espaces + ajouter chiffre (aussi bonne)
- Supprimer espaces et ponctuation + ajouter chiffre
- Supprimer espaces + ajouter chiffre + lowercase/uppercase
- Supprimer espaces

Apparemment personne n'utilise les underscores. Il faudrait une règle « majuscule sur le début de chaque mot », qui manque.

Autres bases de MD5

- Gamigo : alleaussermir // BackToFeature8 // honraatuspadres // elcaballodesantiago // pikachuetpichu // SpottedMarshFrog9 // Weltohnelicht // thazisogiamena2 // vocediperlablu7
- Aha : athousandsplendidsuns // bilawalbhuttozardari // brplvidyalaya // lefruitvert // HOMOPHOBIAISGAY // hppavilionze4900 // scrumpyjackcider // spinadipesce
- EHarmony : SINTIAUWHIAPLU // YOMISMAFUIMIRUTA ; 14916253649648110 0 // ANOTHERONEBITESTHEDUST // IDONTWANTTHISANYMORE // WHATTHEFUCKISGOINGON // FOREIGNEXCHANGESTUDENT // CHANGEYOURPASSWORD // NOWISTHEWINTEROFOURDISCONTENT // IWANTTOGETMARRIED

Autres bases de MD5 - mnémoniques

- whwhgtmgathn : « Whatever happens we have got // the Maxim gun and they have not » - *The Modern Traveller* (Hilaire Belloc) ; **Gamigo** database
- aiwsittdrcau : « All I wanna say is that they don't really care about us » – Michael Jackson ; **Aha** database
- wmkahemfh : « Wenn man keine Ahnung hat einfach mal Fresse halten » – Dieter Nuhr (« If you have no idea, just keep your face shut ») ; **Gamigo**
- LDWMJOMTTKTP : « Last dance with Mary Jane, one more time to kill the pain » – Tom Petty and the Heartbreakers ; **Eharmony**

All you need is love : EHarmony

ANOTHERONEBITESTHEDUST
CORAZONCONAGUJERITOS
IDONTWANTTHISANYMORE
LOVEISWHEREYOUFINDIT
HEALTHYRELATIONSHIPS
SNOWFALLINGONCEDARS
AVERYBEAUTIFULWOMAN
HEALTHYRELATIONSHIP
CONFEDERACYOFDUNCES
GODISGOODALLTHETIME
IMALOVERNOTAFIGHTER
LEAVEMETHEFUCKALONE
LONGDISTANCEVOYAGER
FORTRESSOFSOLITUDE
GIRLSHAPEDLOVEDRUG
ITMUSTHAVEBEENLOVE
SLEEPLESSINSEATTLE
IHATETHATYOULOVEME
THEBESTISYETTOCOME
SOMETHINGBEAUTIFUL
MYHEARTBELONGSTOME
IBELIEVEINYOUANDME
LETSSEEIFTHISWORKS

JESUSTAKETHEWHEEL
IWANTTOGETMARRIED
SINGLEANDLOVINGIT
LOVEANDFRIENDSHIP
THINGSWEDOFORLOVE
THEBESTFATHEREVER
MATCHMADEINHEAVEN
LOVESOMEONELIKEME
ARRANGEDMARRIAGE2
YOUFUCKINGASSHOLE
IWANTTHEFAIRYTALE
SOMEWHEREOUTTHERE
INMYDAUGHTERSEYES
HOPELESSROMANTIC!
SHAKESPEAREINLOVE
STARTINGOVERAGAIN
LOOKINGFORMRRIGHT
ICANFEELYOURANGER
PRINCESSBUTTERCUP
OILANDWATERDONTMIX
NOTHINGELSEMATTERS
LOVEISABATTLEFIELD

MARRIAGEMATERIAL
JESUSISMYHOMEBOY
WILLIFINDSOMEONE
HERESTOTHEFUTURE
THEGREATESTTHING
ALTOGETHERLOVELY
THEPROMISEOFLOVE
SECONDTIMEAROUND
THEPASTISTHEPAST
YOMISMAFUIMIRUTA
ILOVEJESUSCHRIST
LOVEYOURNEIGHBOR
LOVEISALLYOUNEED
TRULYMADLYDEEPLY
SINGLEANDLOOKING
LOOKINGFORTHEONE
BACHELOROFTHEYEAR
BEAUTYANDTHEBEAST
THECHILDRENSPLACE
AINTTALKINBOUTLOVE
JOSEPHINEBONAPARTE
AMATCHMADEINHEAVEN
GOVERNMENTOFGHANA