Looking at sentences as a password

Norbert Schmitz

December 3, 2012

Who am I

- M. Sc. IT-Security / Networks and Systems
- Masterthesis: "improved guessing on composite passwords"
- Supervisor: Markus Duermutt
- working as a in-house pentester in germany

Disclaimer

This presentation has nothing to do with my employer.
It is my personal opinion, based on my master thesis "Improved guessing of composite passwords"
During this work

- no computer and
- no personal data

were harmed in any way.

Passwords

We as researchers/pentester request from our users, that:

- the password is hard to guess
- therefore it shall be long (min. 10 or 12 signs)
- the password uses at least 3 out of four classes

which makes them hard to remember.

Possible solution

Some solutions our users are taking:

- write them down (bad)
- reuse passwords (even worser)
- choose easy to guess passwords (bad)
- use sentences/combined passwords (we will look at)

Password leaks

It all started with RockYou

- founded in 2005
- started to integrate games into facebook (2007)
- got hacked in December 2009
- 32.000.000+ passwords were leaked
- passwords were stored in clear text
- minimum size 5 chars
- asked to users to give the password for their social networks
- "We will not store them..."

Examples

Some Examples of the passwords

- 123456
- iloveyou
- ilovemanny
- bluedevils
- IHATEYOU1
- and many more

ups there are some sentences

Ways of cracking

- brute force
  - fast and early results
  - resonable for passwords up to 10 or 12 chars
- Rainbowtables
  - very good for single passwords
  - but only for up to around 8 characters
- dictionary attack
  - usually good and fast results
  - limited to the size of the dictionary + something
  - usually no sentences or combined passwords
  - usually some transformations are applied

Sentences

So what is a sentence?

- a combination of two or more words
- usually within a gramar context
- for this work I decided to ignore grammar

A new approach

is there a way to crack longer passwords

- which are a combination of words
- which are usually not in any dictionary
- which we have never seen before

A new approach

Requirements

- a set of passwords
- a dictionary

Process

- analyse each password with respect to the words in the dictionary
- learn how users are combining words based on their size
- improve the given dictionary based on what we learned

Requirements

the set of passwords

- RockYou

Dictionaries

- project gutenberg
- wikipedia
- ispell

Gutenberg and Wikipedia

- downloaded from the internet
- did some cleaning
- cutted the text into words
- counted them
- set a threshold to min. 200 occurences
- around 210.000 words

Ispell

- dictionary of a spell checker
- no cleaning was needed
- added a list of names
- around 140.000 words

Analyzer

- take a password
- isolate sequences of letters
- identify the words from the dictionary in this sequence by trying the bigger words first
- end up with something like this:
- iloveyou -¿ 1-4-3
- safe the result

What about 1337-speak

lets have a short look at 1337 speak

- found words like peps!!
- what does it mean?
  - peps!! - financial term with two !
  - pepsi! - a sode with one !
  - peps!i - two words
  - pepsii - three words or one word with the number ii

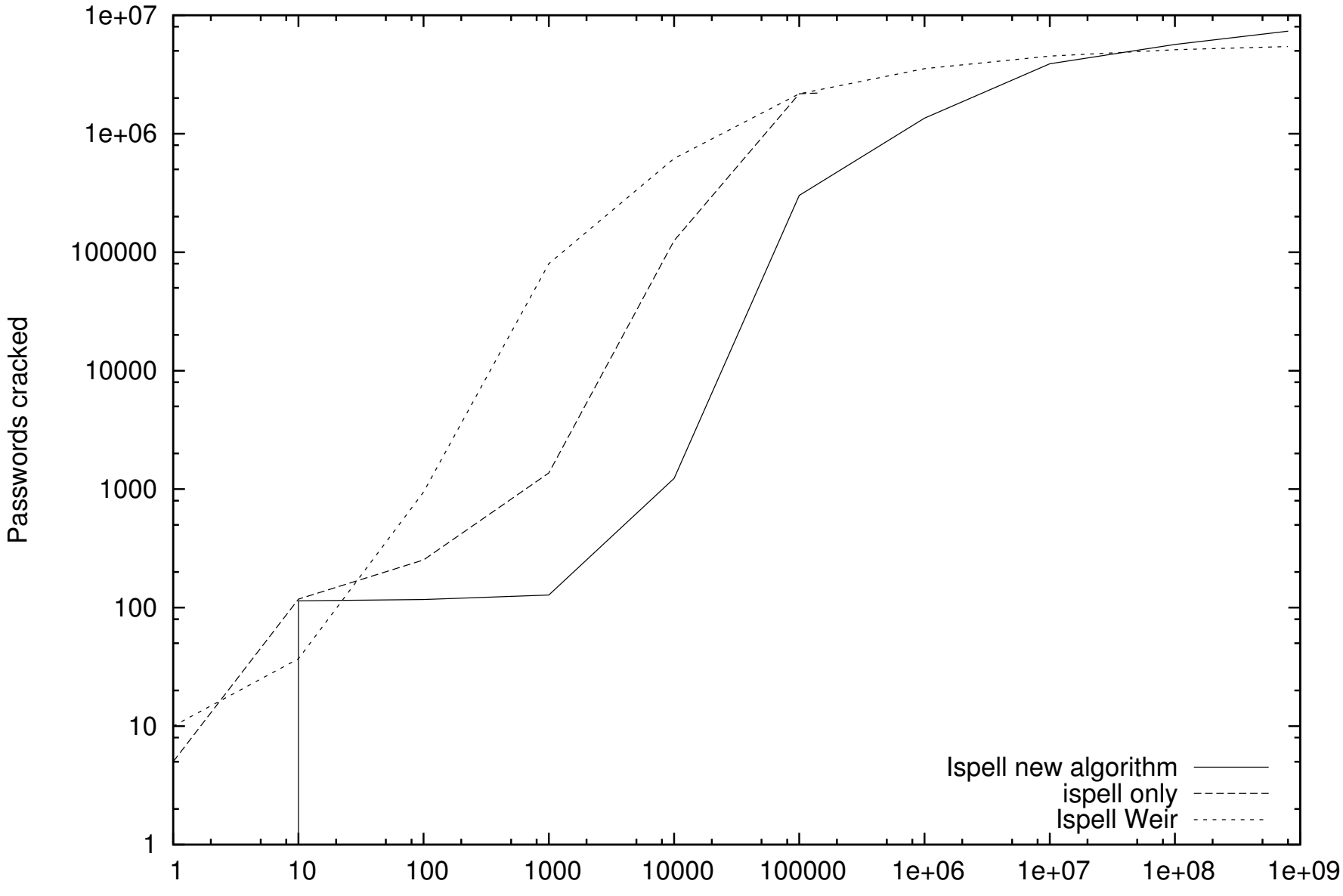
Create a new dictionary

- load the analysis
- calculate the possible size of the combinations
- set it into relation to the number of password we can crack with this
- sort it according to the relation and enhance the dictionary up to a certain given threshold

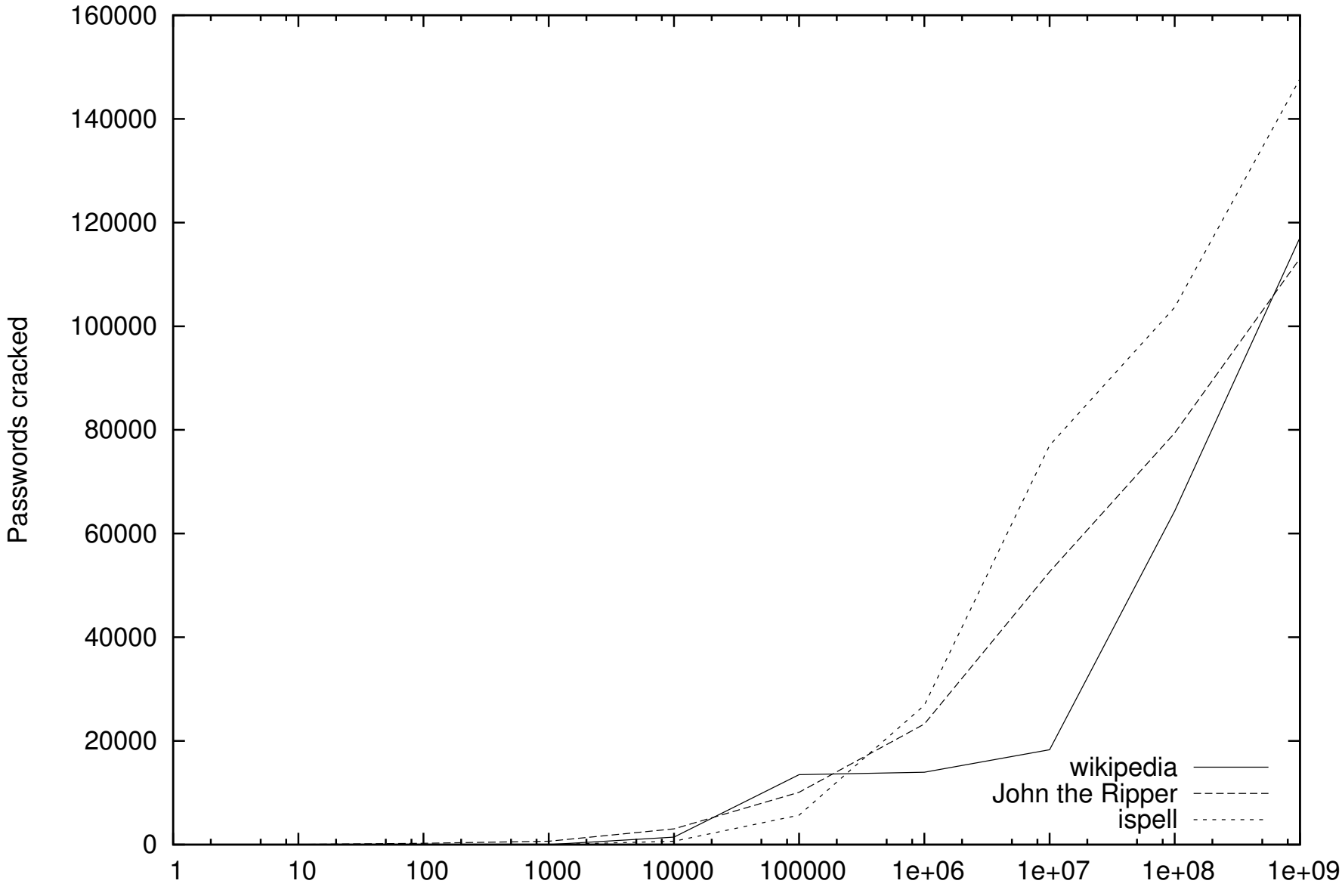
Cracking the leak

- used the work from Matt Weir to learn patterns for enhancing a dictionary
- applied these patterns to the words
- look the new words up in the leak

Results


Passwords cracked
1e+07
1e+06
100000
10000
1000
100
10
1
1
10
100
1000
10000
100000
1e+06
1e+07
1e+08
1e+09
Ispell new algorithm
ispell only
Ispell Weir

Results


Passwords cracked
160000
140000
120000
100000
80000
60000
40000
20000
0
1
10
100
1000
10000
100000
1e+06
1e+07
1e+08
1e+09
wikipedia
John the Ripper
ispell

146Mio passwords

- on twitter a torrent with 146 Mio password hashes was announced
- several researcher tried to break them
- I took one of the list of cracked password (122Mio paswords)
- run my modified dictionaries against it
- could crack additionally around 5 percent of the left over passwords

Conclusion

Some conclusions

- we have seen a new way to improve dictionaries
- the improved dictionaries resulted in 5 - 10 % more cracked passwords
- highly depending on the number of combined passwords

The source will be published in spring 2013 after some cleaning up

Time for

Questions?

pw@norbert-schmitz.de