Carnegie Mellon University

Effect of Grammar on Security of Long Passwords

Ashwini Rao*, Birendra Jha^, Gananand Kini
Carnegie Mellon University, ^Massachusetts Institute of Technology
*arao@cmu.edu

Motivation

- User-selected, text-based passwords are used widely
- Trade-off between **usability** and **security**
  password123 vs. *Bn$#76gHkl*
- Good compromise: longer + simpler composition requirements
  minimum 16 character password [1], passphrase [2]
- Users rely on memory aids to cope with increased length grammatical structures, postal address, URLs

Table 1: Examples of phrases in long password dataset

Category	Password Example	Phrase	Total
Simple	abiggerbetterpassword	a bigger better password	178
Substitution	thereisnomored0ts	there is no more dots	20
Extra Symbol	longestpasswordever8	longest password ever	70
Total out of 1434			268

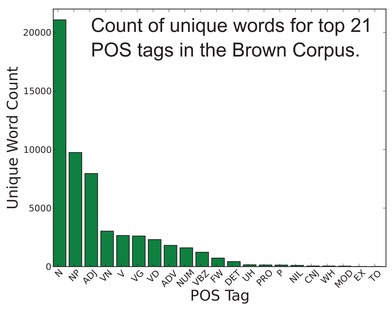
How does presence of structure affect security of passwords?

Are longer passwords always stronger?

Thr3r3 can only b3 #1! vs. *Superman is $uper str0ng!*

- Previous research focused on structural dependencies at the character level [1,3,4]
- We focus on **higher-order** structure - grammatical structure

Approach



Word counts are unevenly distributed among different POS tag types; important for password search space and guessing effort

Table 2: Brown Corpus statistics

Words	1161192
Unique Words	49815
Sentences	57340
Characters per Word	4.26
Words per Sentence	20.25
Unique Characters	58
Content Genres	15

- Parts-of-Speech (POS) tagging to capture structure within long passwords
  bigger better password ⟶ *Adjective Adjective Noun*
- Extract *probable sequence of POS tags* (*tag-rules*) from corpora, e.g. the Brown Corpus [6], long-password data set [1]
- Tag-rules used for
  estimating reduction in search space
  decrease in guessing effort
  building novel grammar-aware password cracker

Search Space

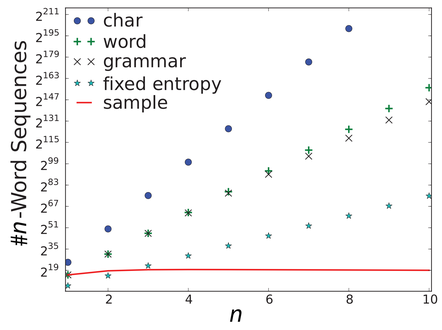


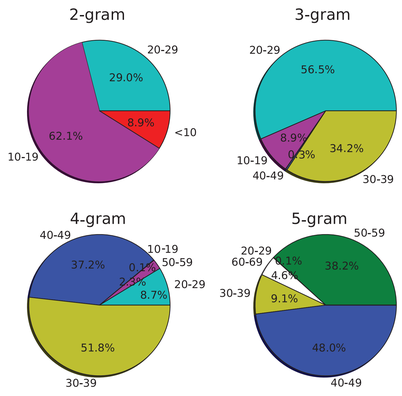
Table 3: Percent decrease in password search space

n-word	1	2	3	4	5
$\frac{grammar}{word}$ %	100	99.92	96.90	80.66	46.95
n-word	6	7	8	9	10
$\frac{grammar}{word}$ %	17.17	4.28	0.99	0.25	0.07

- Analytical model to estimate decrease in search space
- Numerical evaluation using the Brown Corpus

>50% decrease in search space for passwords of length 5 words due to the presence of structure

Guessing Effort



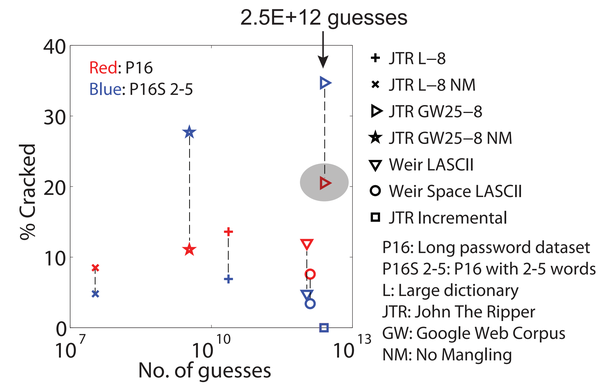
- Tag-rules divide password search space unevenly.
- Significant no. of tag-rules have very low strength
- Enforcing uniqueness [5] is not enough to ensure uniform distribution of password values when structures are present
- Optimization framework to estimate decrease in guessing effort due to presence of structures

$$\text{maximize } gain = \sum_{i=1}^{count(TS^{\text{grammar}})} r_i v_i$$

$$\text{subject to } \sum_{i=1}^{count(TS^{\text{grammar}})} r_i g_i \leq G$$

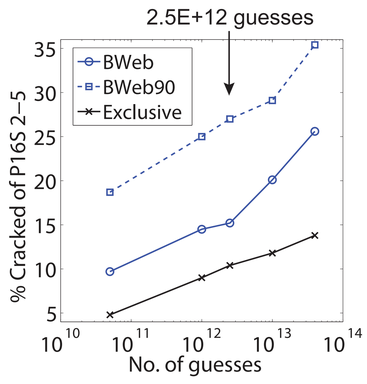
$$0 \leq r_i \leq 1$$

Limitations of current crackers



- Cannot generate longer passwords by automatically combining multiple words
- User has to manually add longer values to the dictionary
  implies better dictionary may improve cracking
- We used Google Web Corpus as dictionary
  corpus contains word sequences of length 1-5
  cracked **20.5% vs. 6%** (previous published result [7])

Grammar Aware Cracking



BWeb: Tag dictionaries created from words from Brown Corpus + small web corpus
BWeb90: BWeb with .9 cumulative probability cut-off for each tag dictionary
Exclusive: Passwords cracked only by grammar-aware cracker

- Emulates user behavior; automatically combines multiple words to create longer passwords
- Generates probable candidate guesses using tag-rules and individual dictionary of words for each POS tag
- Can target non-sensical phrases, different user groups

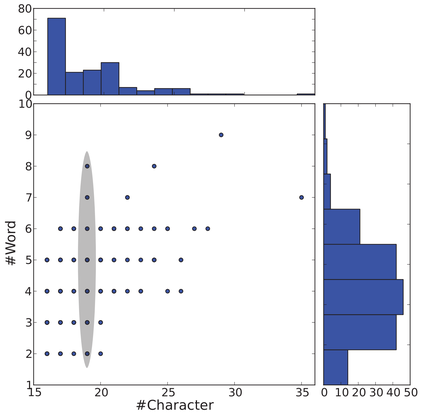
Cracked 10% passwords not cracked by other crackers
Outperforms Weir and JTR-Incremental
Consumes <10MB of storage compared >50GB for JTR-GW
Optimized dictionary can further improve performance

Policy Implications

Table 4: Comparison of passphrase strength

Passphrase	Tag-Rule	Guesses	Time
Th3r3 can only b3 #1!	EX MOD V DET PRO	1.3E12	22 min
Hammered asinine requirements.	VD ADJ N	12.6E12	3.5 h
Superman is $uper str0ng!	NP V ADJ ADV	12.3E15	142 d
My passw0rd is $uper str0ng!	PRO NP V ADJ ADV	1.7E18	56 yr

- Passphrase strength is not a direct function of number of characters or words (longer not necessarily stronger)
- Similar looking passphrases may differ in strength by orders of magnitude



- Users can adjust number of words and length of each word to meet password policy requirements
  implies different underlying structures and passwords of different strengths

Conclusions

- Long passwords are promising
- Need to understand effect of structure to achieve level of security and usability envisioned
- Long password policies and enforcement tools have to be structure cognizant
- More research is required

References

1. Kelley et al. Guess again (and again and again): Measuring password strength by simulating password-cracking algorithms. IEEE Symposium on Security and Privacy (2012)
2. Indiana University Passphrase Policy
3. Shay, et al. Encountering stronger password requirements: User attitudes and behaviors. Symposium on Usable Privacy and Security (2010)
4. Weir et al. Password Cracking Using Probabilistic Context-Free Grammars. IEEE Symposium on Security and Privacy (2009)
5. Schechter et al. Popularity is Everything: A new approach to protecting passwords from statistical-guessing attacks. USENIX Workshop on Hot Topics in Security (2010)
6. Kucera et al. Computational Analysis of Present-Day American English. Brown University Press (1967)
7. Brants et al. Web 1T 5-gram Version 1. Linguistic Data Consortium (2006)