UNIVERSITY OF AMSTERDAM

Generating probable password candidates for the offline assessment of Dutch domain password hashes

By Tom Broumels

Supervisor: P. Campers

Secura

Introduction

- Use and abuse of passwords common practice today[1]
- Passwords stored as hashes
- Cracking by trying: selecting likely password candidates
- NIST recommendation on using breach corpuses[2]

Secura → 0DF335A49BD7B40BE674EEE80A6FBADD

[1] Joseph Bonneau et al. “The quest to replace passwords: A framework for comparative evaluation of webauthentication schemes” , 2012

[2] Paul A Grassi et al. “NIST Special Publication 800-63b: Digital Identity Guidelines”, 2017

Assessing password hashes at Secura

- Password hash strength assessment part of security assessments, e.g., red teaming exercises
- Improved assessment can lead to shorter lead times or more complete results:
  - Finding more passwords in total
  - Finding more passwords first 30 minutes
- Frequently assessing hashes for Dutch clients

Ethical considerations

- Using breached passwords
  - Realistic assessments
  - Removal of e-mail addresses

- Validating research on hashes of active user accounts
  - Secured environment
  - Password hashes only

Research question

How do different password guessing algorithms compare in selecting probable password candidates for assessing offline Dutch domain password hashes?

Secura

Important related work

- No publications available on Dutch passwords
- **Human behaviour** related to password generation
- Password candidate generation:
  - Markov, **PCFG**, OMEN, PRINCE
  - **PassGAN**, NeuralNetwork
- Combining approaches: TarGuess

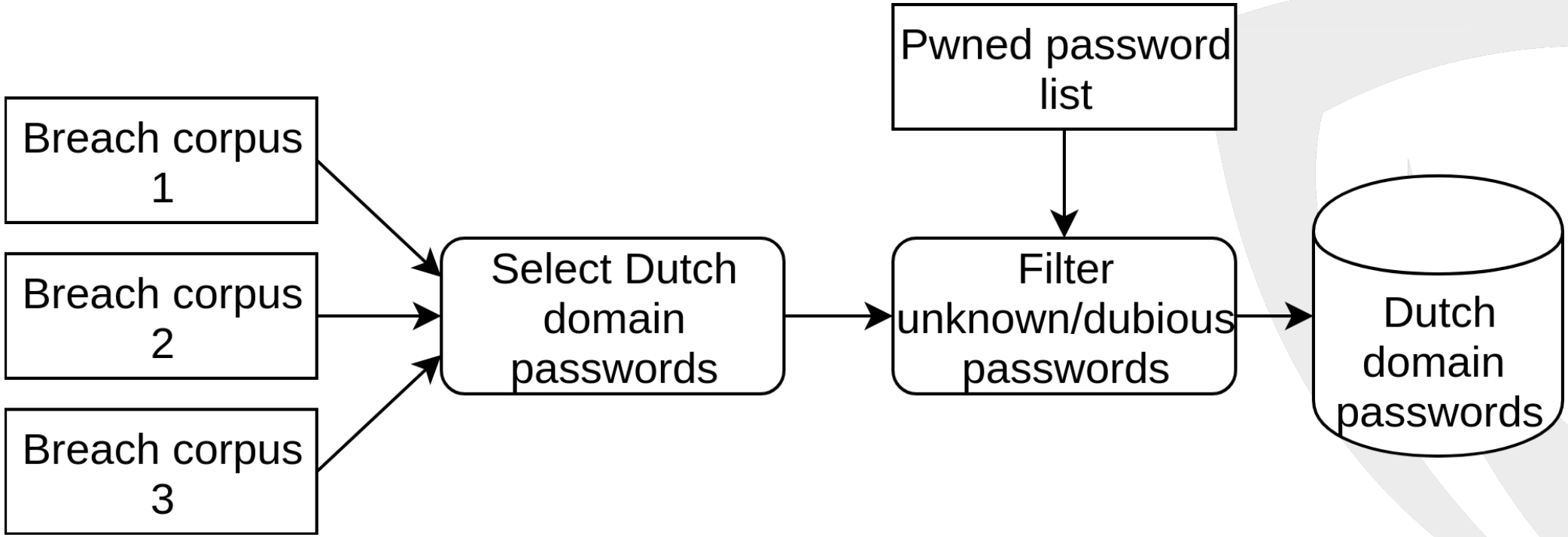
Research method

1. Dutch domain password selection
2. Selecting different password cracking approaches
3. Comparing approaches using experiments
4. Selecting a well performing approach

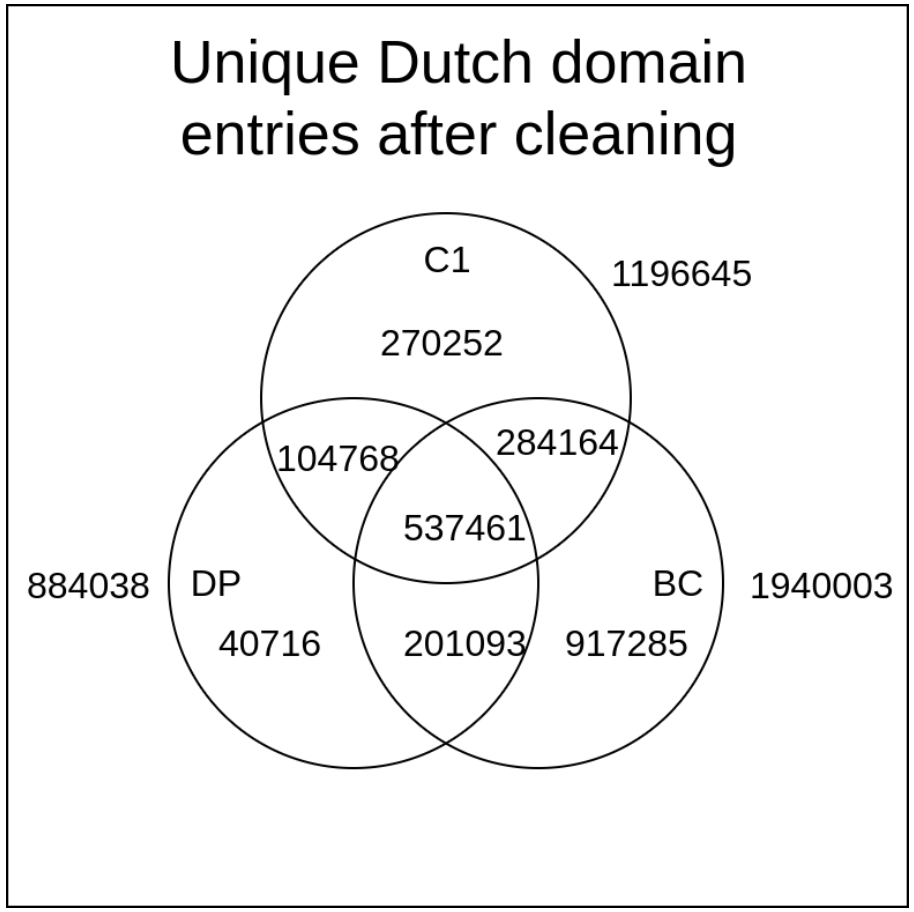
Research method

1. **Dutch domain password selection**
2. Selecting different password cracking approaches
3. Comparing approaches using experiments
4. Selecting a well performing approach

Dutch domain password selection



Dutch domain passwords



- Unique email/password entries: **3,424,464**
- Unique passwords: **2,355,739**
- 31,2% duplicates

Common Dutch domain passwords

Password	#	%
123456	8795	0.26
welkom	3950	0.12
SKIFFY	3708	0.11
welkom1	2547	0.07
123456789	2524	0.07
qwerty	2304	0.07
welkom01	2220	0.06
wachtwoord	2177	0.06
geheim	1792	0.05
amsterdam	1568	0.05

All passwords

Password	#	%
Welkom01	861	0.26
ka_dJKHJsy6	198	0.06
Welkom123	187	0.06
PPPrt30TA	152	0.05
Feyenoord1	139	0.04
P@ssw0rd	107	0.03
Amsterdam1	102	0.03
Hallo123	101	0.03
Wachtwoord1	94	0.03
Geheim01	76	0.02

8ULNS passwords

Research method

1. Dutch domain password selection
2. **Selecting different password cracking approaches**
3. Comparing approaches using experiments
4. Selecting a well performing approach

Approach 1 of 3: Human behaviour on password selection

- Alan S. Brown et al. (2004)
- Generating and remembering passwords
- Questionnaire 218 US students
- Common content of basewords (e.g., reference to self, relative, animal, personal interest, job)
- Common use of basewords (e.g., complete word used in 97% of the time)

Approach 2 of 3: Probabilistic Context-Free Grammars

- Matt Weir et al. (2009)
- Breaking up and recombining passwords
- Frequencies important

Welkom2020!

L_6 ⇒ welkom

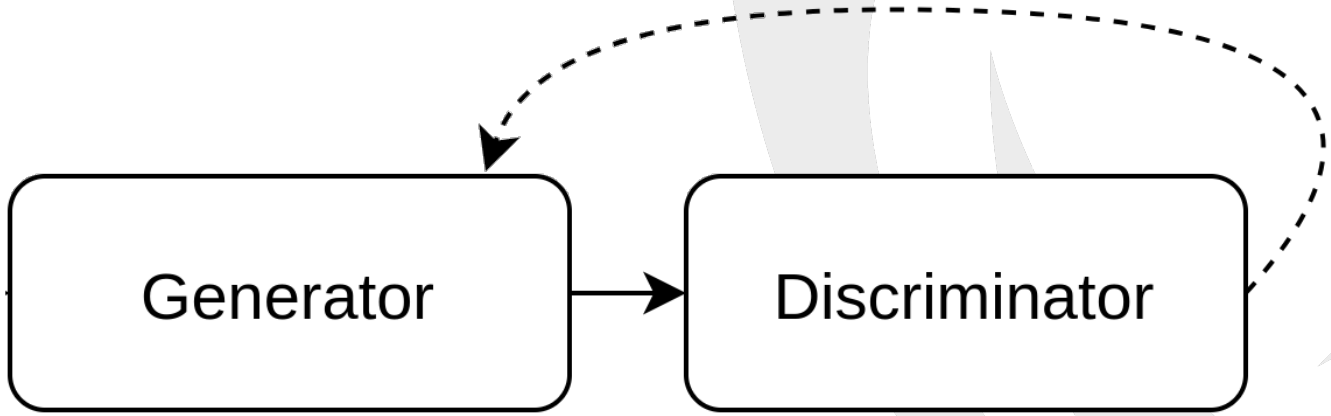
D_4 ⇒ 2020

S_1 ⇒ !

Rule ⇒ $L_6D_4S_1$

Approach 3 of 3: Generative Adversarial Network

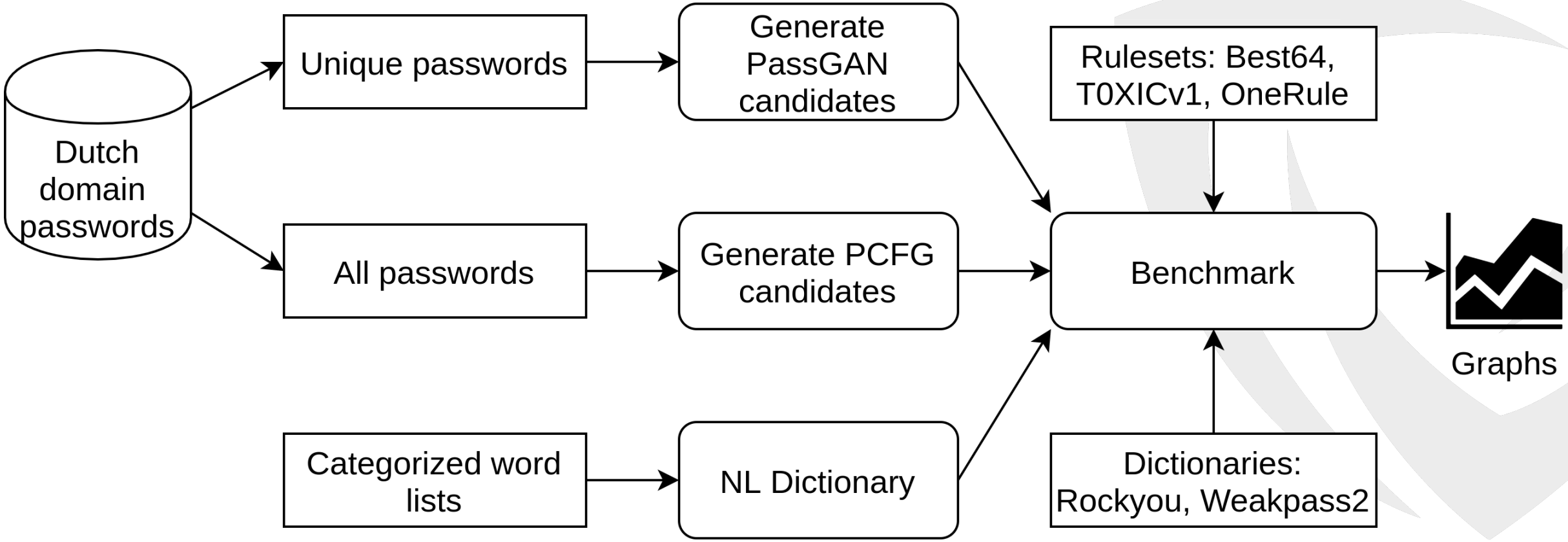
- Hitaj et al. (2019)
- Learning how to generate “passwords”.
- Machine learning based



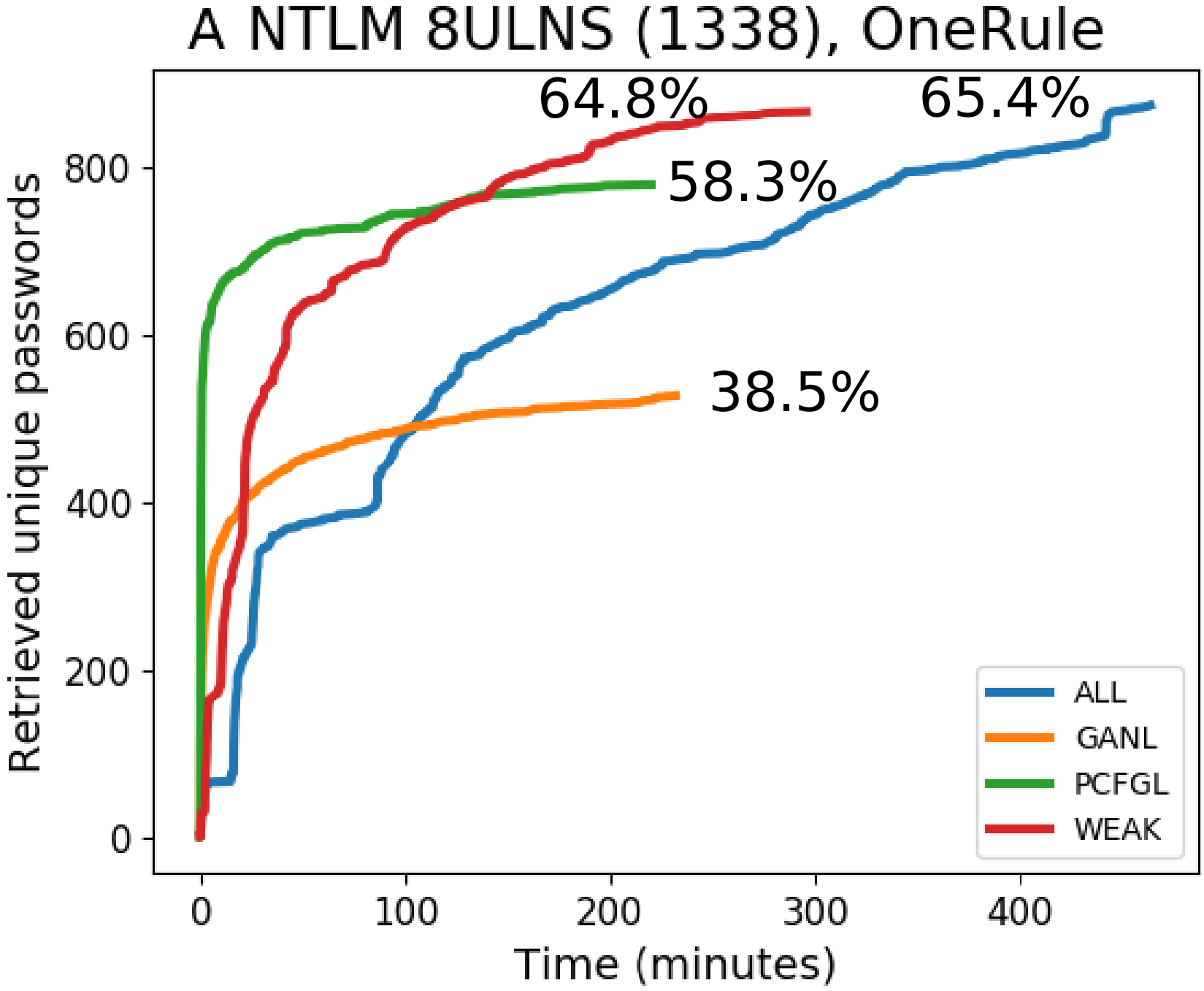
Research method

1. Dutch domain password selection
2. Selecting different password cracking approaches
3. **Comparing approaches using experiments**
4. Selecting a well performing approach

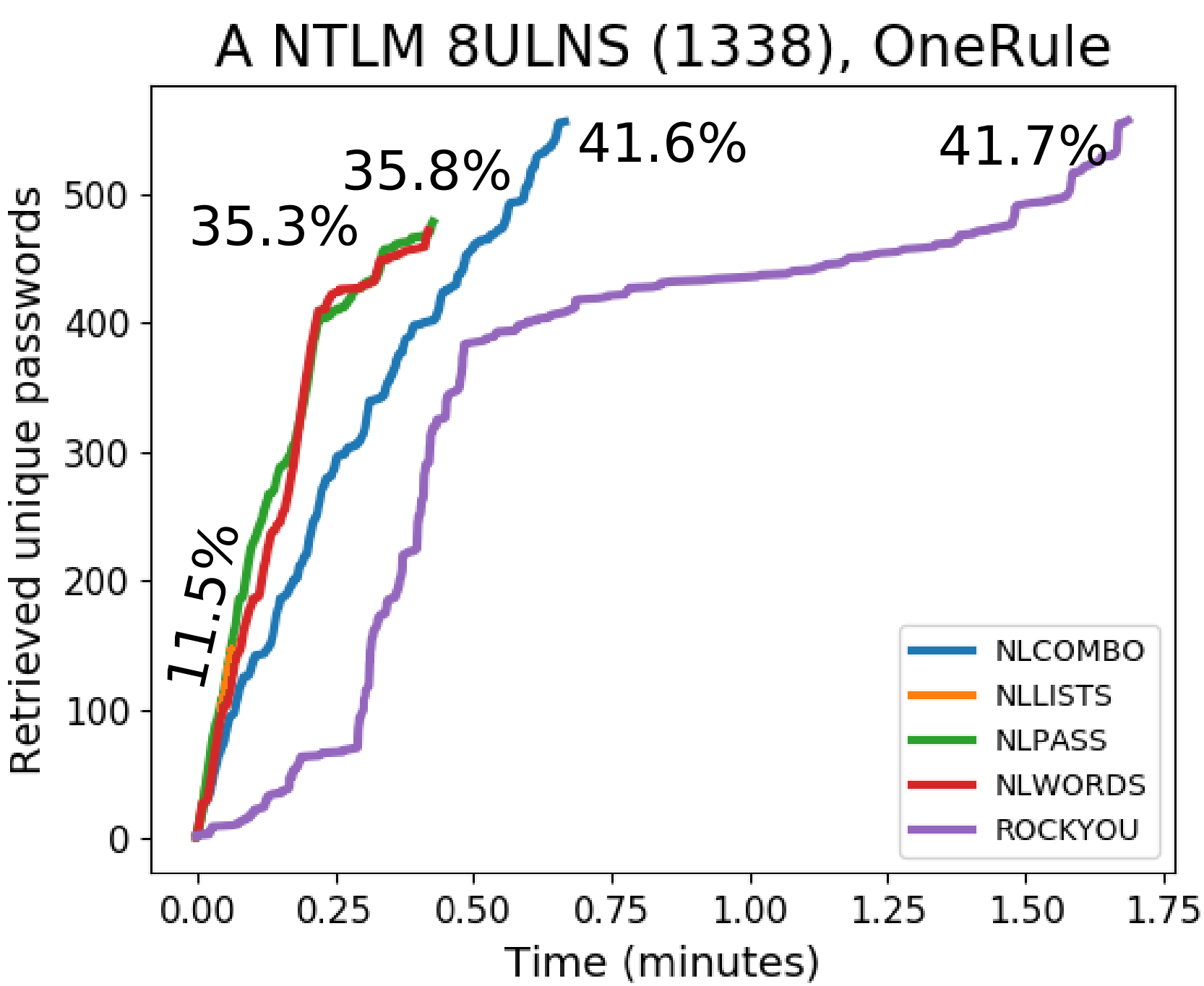
Comparing approaches in an experiment



Results


A NTLM 8ULNS (1338), OneRule
64.8%
65.4%
58.3%
38.5%
Retrieved unique passwords
0
200
400
600
800
0
100
200
300
400
Time (minutes)
ALL
GANL
PCFGL
WEAK

Results


A NTLM 8ULNS (1338), OneRule
Retrieved unique passwords
0
100
200
300
400
500
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
Time (minutes)
11.5%
35.3%
35.8%
41.6%
41.7%
NLCOMBO
NLLISTS
NLPASS
NLWORDS
ROCKYOU

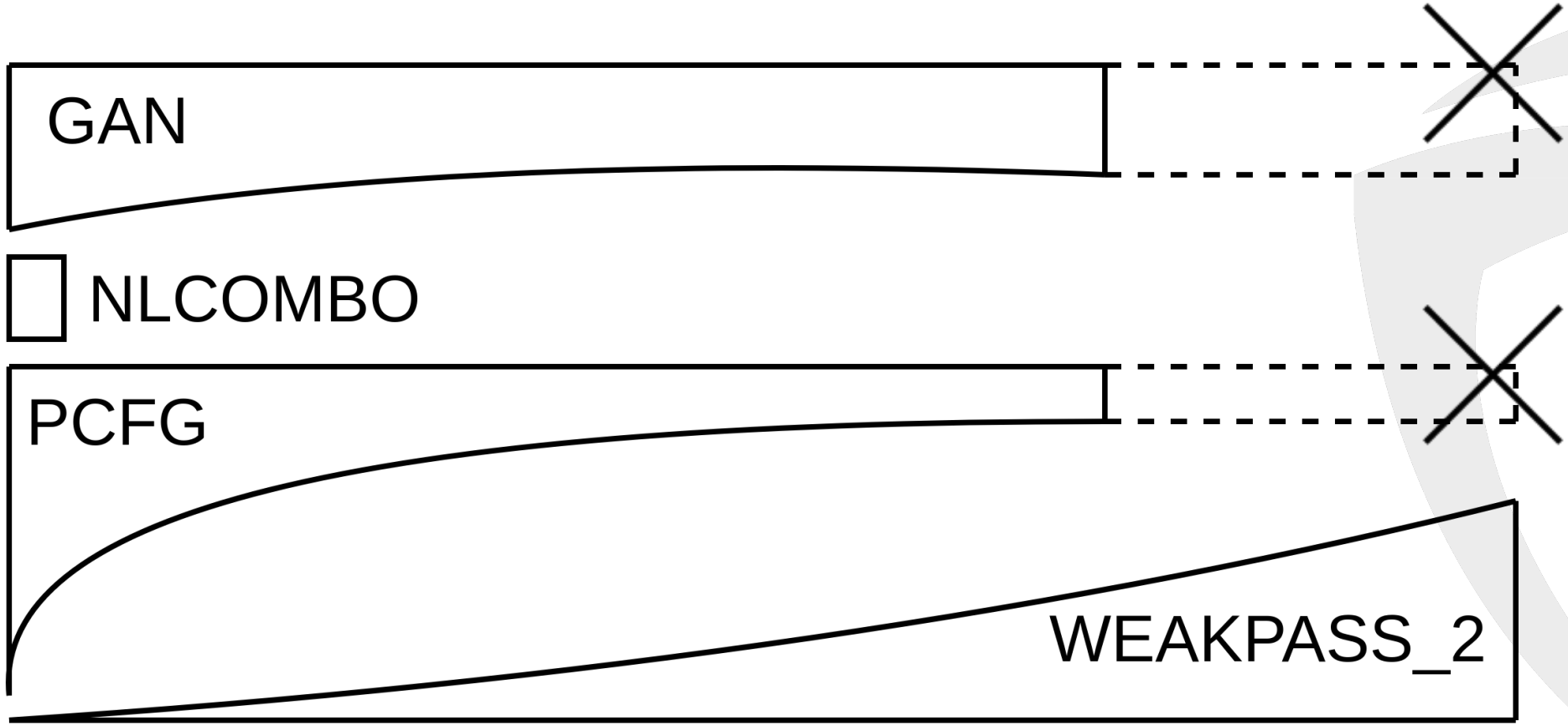
Research method

1. Dutch domain password selection
2. Selecting different password cracking approaches
3. Comparing approaches using experiments
4. **Selecting a well performing approach**

Combining approaches

Hashes	Hash type	Unique hashes	Ruleset	1	2	3	4	5
A	NTLM	1338	(none)	24.7%	30.7%	33.1%	33.6%	33.9%
A	NTLM	1338	Best64	37.0%	41.6%	43.0%	43.7%	44.2%
A	NTLM	1338	T0XICv1	58.4%	62.0%	63.5%	64.3%	64.8%
A	NTLM	1338	OneRule	65.4%	69.1%	70.0%	70.3%	70.6%

Combining approaches

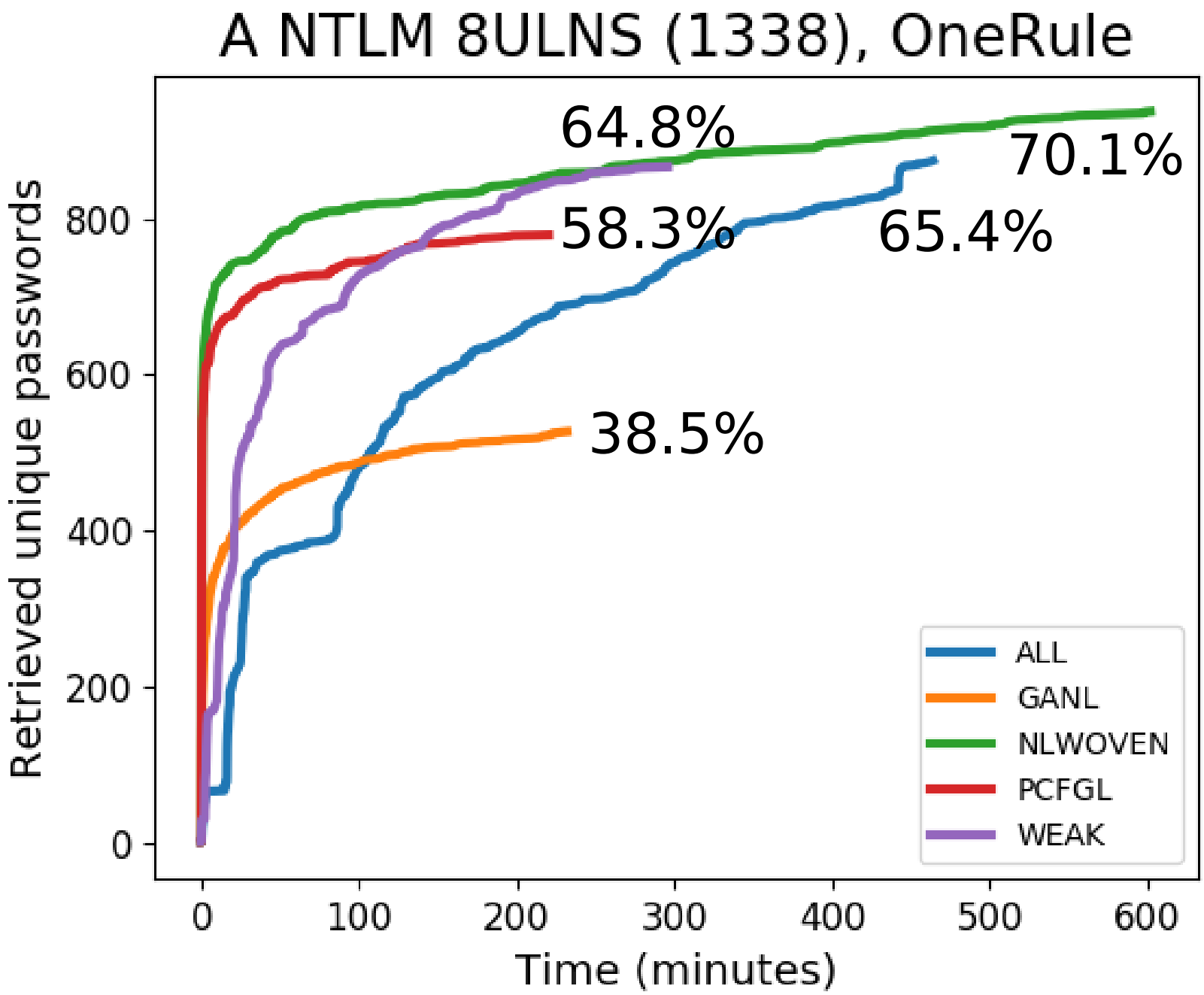


1. Removing duplicates (9.1%)
2. Merging dictionaries by “weaving”

Results

PCFGL	ALL	NLWOVEN
52.32%	25.71%	55.75%

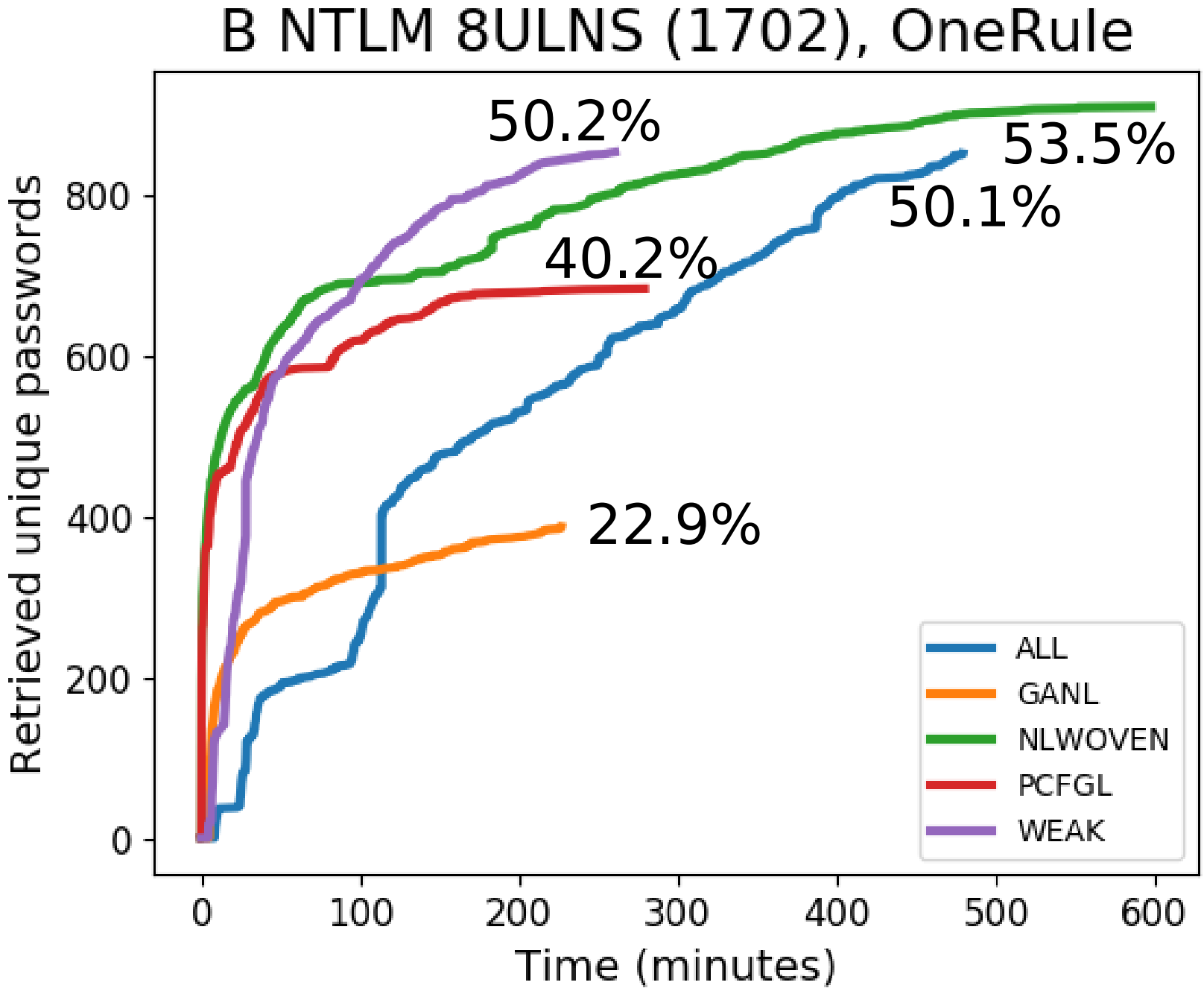
After 30 minutes



Results

PCFGL	ALL	NLWOVEN
30.79%	7.34%	32.9%

After 30 minutes



Research question

How do different password guessing algorithms compare in selecting probable password candidates for assessing offline Dutch domain password hashes?

Secura

Conclusions

- Single approaches perform well on one or two aspects:
  - total amount, amount 30 mins. or guesses per crack
- Combining approaches can:
  - Increase the total amount of passwords found:

    +7.2% and +6.7%
  - Increase the amount of passwords found within 30 mins: +117% and +348%
- Rulesets increase the amount of hashes found for all the selected approaches.

Discussion & Future work

- Dirty data in breach compilations
- Tests on two sets of hashes
- Consider adding organisation specific information
- Consider iterative cracking by using cracked passwords as input for further cracking

Acknowledgement

Secura

P. Campers

E. Slangen

R. Moonen

Scattered Secrets

J. van Beek

Secura

Questions

Conclusions:

- Single approaches perform well on one or two aspects
- Combining approaches can:
  - Increase the total amount of passwords found
  - Increase the amount of passwords found within 30 mins.
- Rulesets increase the amount of hashes found for all the selected approaches.

Secura

NIST Special Publication 800-63b: Digital Identity Guidelines

"When processing requests to establish and change memorized secrets, verifiers SHALL compare the prospective secrets against a list that contains values known to be commonly-used, expected, or compromised. For example, the list MAY include, but is not limited to:

- Passwords obtained from previous breach corpuses.
- Dictionary words.
- Repetitive or sequential characters (e.g. 'aaaaaa', '1234abcd').
- Context-specific words, such as the name of the service, the username, and derivatives thereof."

Benchmark results

Hashes	Ruleset	GANL	PCFGL	NLLISTS	NLWORDS	NLPASS	NLCOMBO	ROCKYOU	WEAK	ALL	NLWOVEN
A NTLM 1338	(none)	4.6%	24.4%	0.0%	0.2%	3.1%	3.1%	3.1%	22.2%	21.3%	33.2%
A NTLM 1338	Best64	10.8%	33.6%	1.6%	7.2%	9.9%	12.4%	12.0%	37.0%	36.6%	43.4%
A NTLM 1338	T0XICv1	26.6%	52.5%	5.4%	22.0%	25.4%	30.1%	32.1%	58.4%	58.4%	63.8%
A NTLM 1338	OneRule	39.5%	58.3%	11.5%	35.3%	35.8%	41.6%	41.7%	64.8%	65.4%	70.1%
B NTLM 1702	(none)	1.6%	7.9%	0.1%	0.1%	0.8%	0.8%	0.9%	9.0%	8.5%	13.1%
B NTLM 1702	Best64	3.8%	12.6%	0.4%	1.6%	2.4%	3.0%	4.0%	17.6%	17.3%	22.2%
B NTLM 1702	T0XICv1	13.6%	28.3%	1.4%	5.1%	7.6%	8.8%	12.3%	38.9%	38.7%	43.9%
B NTLM 1702	OneRule	22.9%	40.2%	4.0%	12.7%	12.8%	16.8%	19.9%	50.2%	50.1%	53.5%

Combining 2 approaches

	WEAK	ROCKYOU	NLCOMBO	GANL	PCFGL	ALL
WEAK	64.8%	64.8%	64.8%	66.9%	67.5%	65.8%
ROCKYOU	64.8%	41.7%	47.5%	52.5%	59.2%	65.4%
NLCOMBO	64.8%	47.5%	41.6%	53.1%	60.6%	65.4%
GANL	66.9%	52.5%	53.1%	39.5%	60.7%	67.3%
PCFGL	67.5%	59.2%	60.6%	60.7%	58.3%	67.9%
ALL	65.8%	65.4%	65.4%	67.3%	67.9%	65.4%

Common basewords used in Dutch domain passwords

Category	Matching passwords	Matching unique passwords	Elements in wordlist
First names	531337	267085	9348
Family names	203503	107539	9113
Pet names	159859	71978	646
Cities and townships	64118	30498	7120
Comic character names	43515	19593	774
Animals	32178	13445	4924
Payed soccer teams	28638	8130	310
...	...	...	...