Modeling Differences

Sudhir Aggarwal, Shiva Houshmand, Randy Flood
Florida State University
Department of Computer Science
E-Crime Investigative Technologies Lab
Tallahassee, Florida 32306

August 5-7, 2015

Modeling Differences: the problem

- **I know a user's password is alice123! and the user has changed this password. How do I make use of this information to crack the new password?**

- Try developing a conditional probability distribution. But, we do not have much data? And how does this help in defining a grammar?

- Try using Edit distance (Levenshtein distance) to find passwords close to the seed password. But how close is close?

- Try using transformational approach (s/1/2/, s/1/11/) where we use a set of regular expressions. Simple transformation seem ok but where do we draw the boundary?

Targeted Attack

- Integrating the information about the target into the probabilistic context-free grammar

- What if we are attacking a specific target that we have some information about

1. Only one old password of the user is accessible
2. Two or more successive passwords are available

Data Collection

- At least a pair of old and new passwords
- More data can help us in defining the transformations by understanding how people change their passwords.
- It could also allow us to use conditional probabilities.
- Used survey questionnaire to collect passwords

ECIT Survey

Create Account Login

Please read the terms of the Consent Form carefully.

By checking I Agree below and creating an accout, you confirm that you have read the above information, you have asked any questions you may have had and have received answers, and you consent to participate in the study.

Create an Account

Please create an account for use in this study. Use your FSU email address for your username. Assume you are creating an email account and you want your password to be strong enough. Try to create your password in a manner that you would normally do. You should take whatever steps you normally take to remember and protect your password. **DO NOT** provide passwords that you currently use for another service. All passwords will be saved and analyzed. DO NOT use this password elsewhere.

Email Address your@my.fsu.edu email

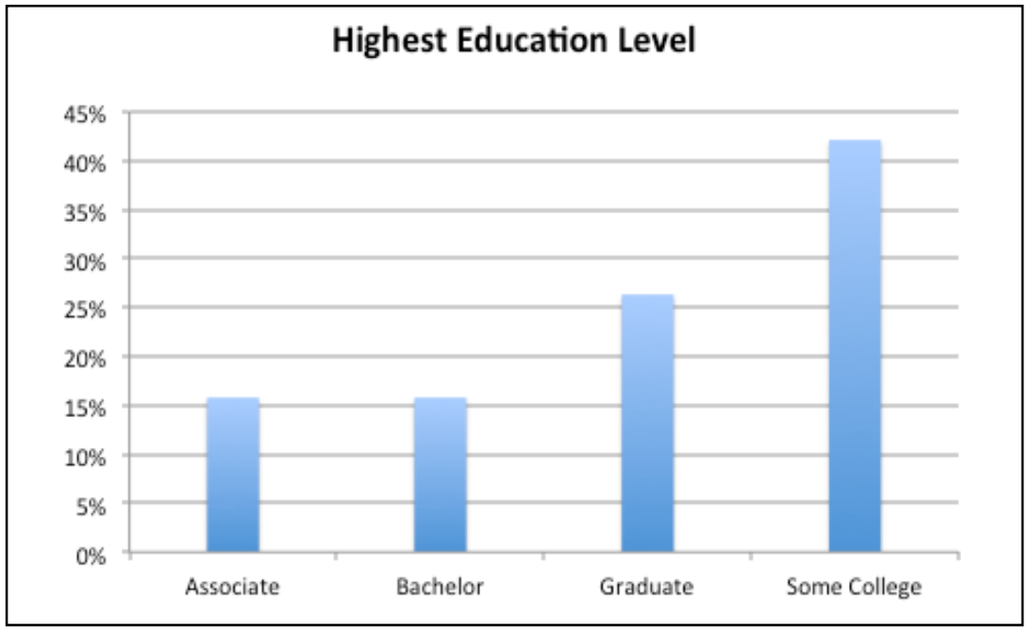
I agree to the terms and conditions stated above.

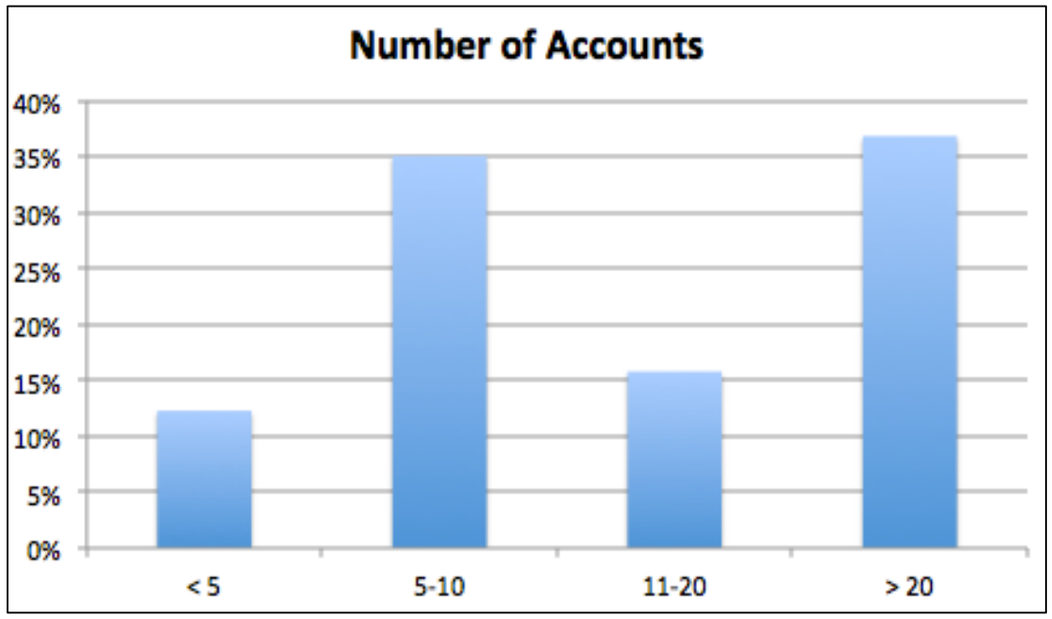
New Password Minimum of 8 characters

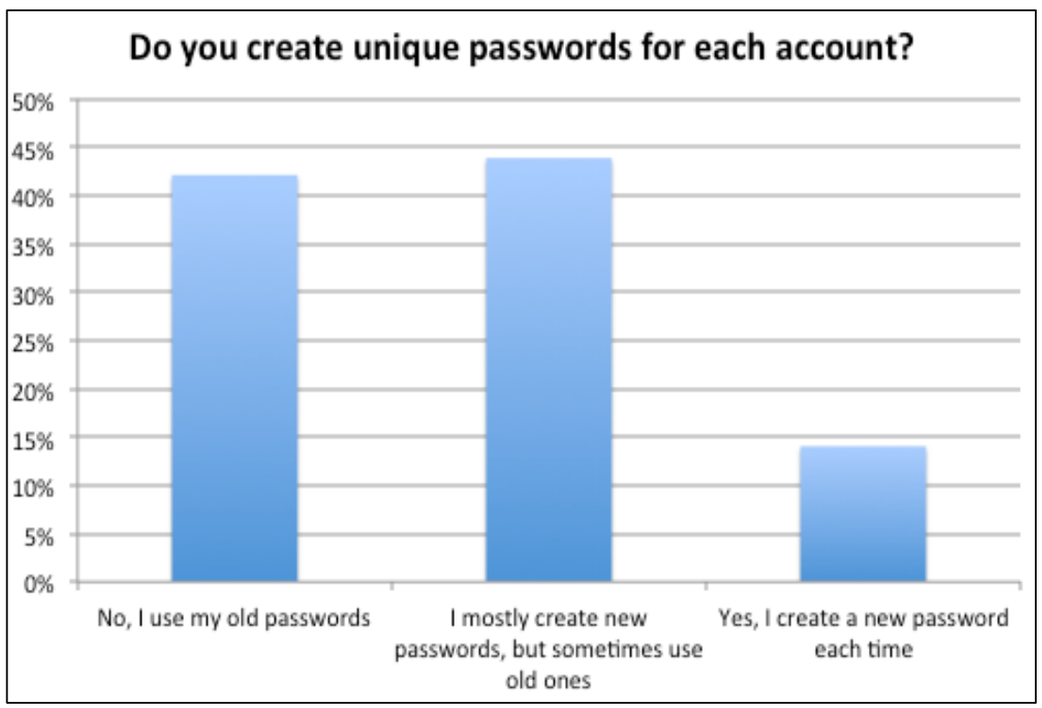
Confirm Confirm Password

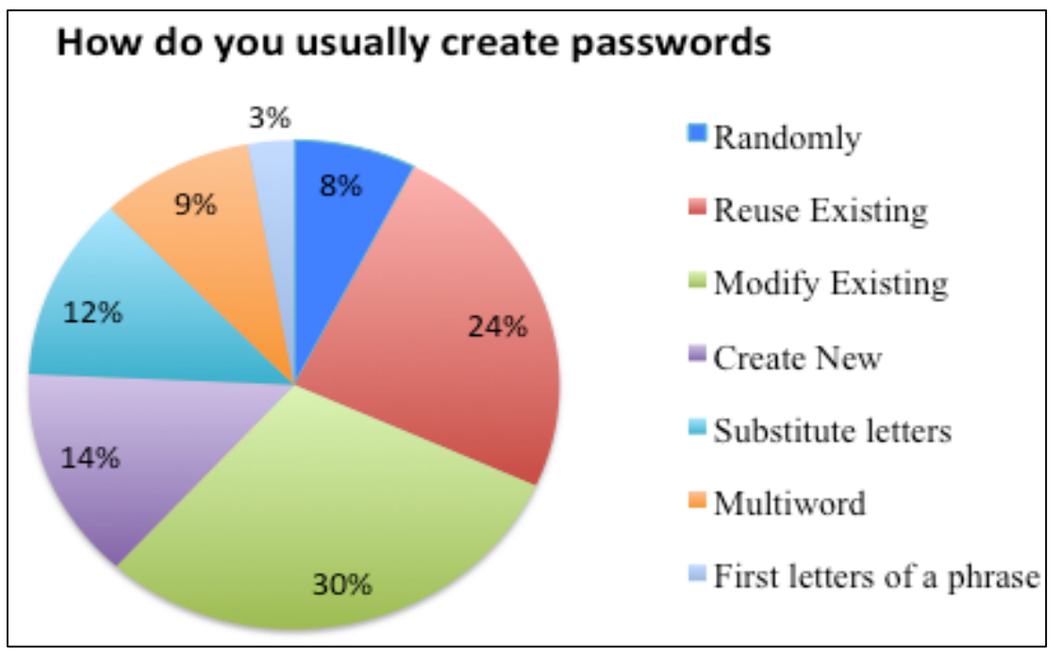
Create Account

Survey Result





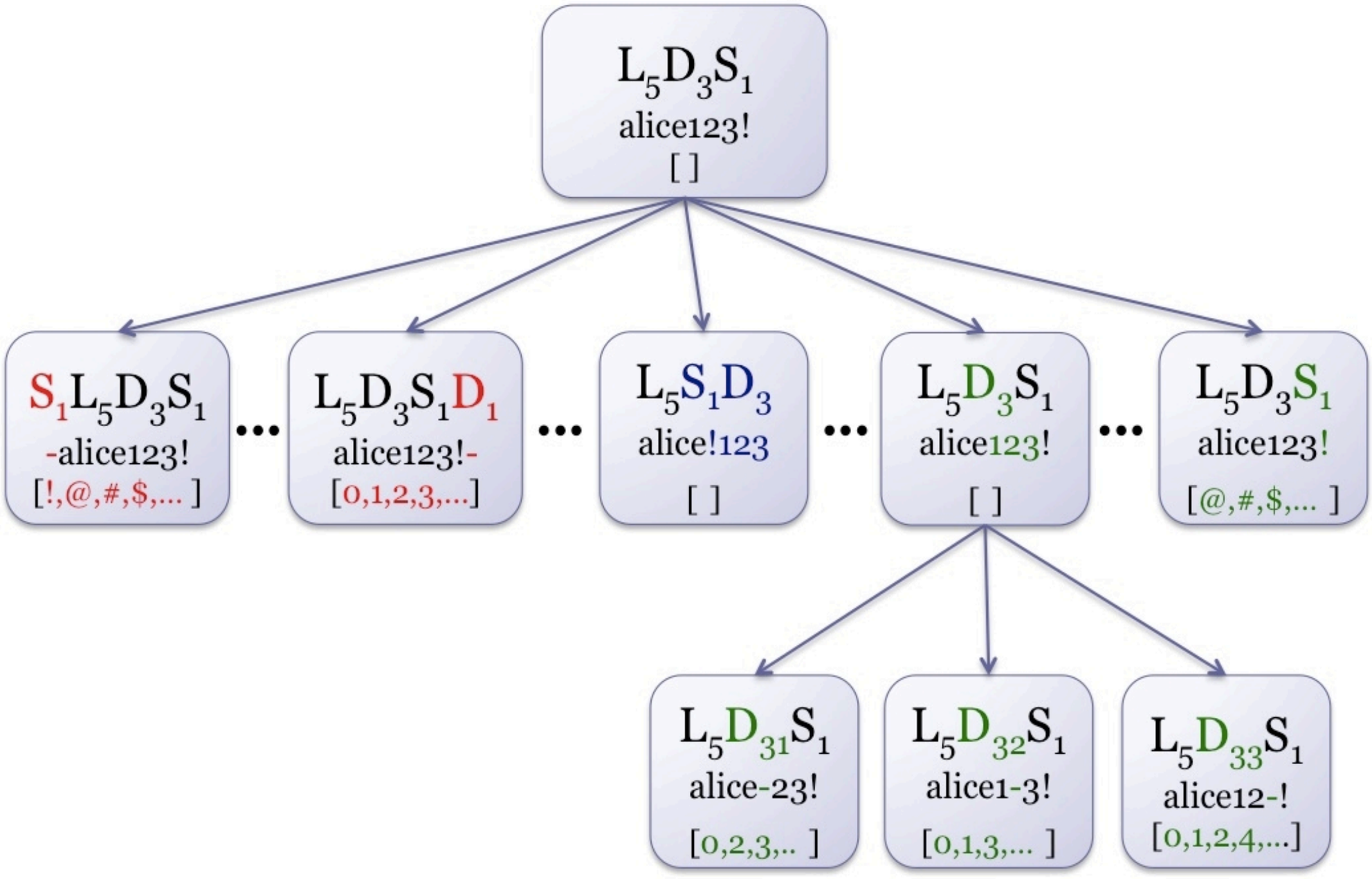




Distance Function

- ***Operations on the Base Structure:*** $L_5D_3S_1$
  - Insertion $L_5S_1D_3S_1$
    - There is no insertion of K_1 or X_1 or R_1
  - Deletion $L_5D_3\cancel{S_1}$
  - Transposition $D_3L_5S_1$

- ***Operations on the component*** D_3: 123
  - Insertion 1234
  - Deletion ~~1~~23
  - Substitution 129



Levenshtein Distance 1 Algorithm

Distance Function

- **Multiwords**
  - Insertion

X8: star,wars
star S_1 wars
star D_1 wars

  - Deletion of each word

thebigdog!
thedog!, bigdog!

  - Transposition

john,marry,bob
marry,john,bob
john,bob,marry
bob, marry, john

Generate Similar Password Guesses

- Generate all possible passwords within some edit distance from the old password.
- This process results in a grammar called Egrammar (Edit distance Grammar)

EGrammar for “alice123!”

Base structure	D_1	D_2	D_3		D_4		S_1		S_2			C_1	C_5
$L_5D_3S_1$	0	12	120	153	0123	1233	@	|	)!	!_	!|	L	LLLLL
L_5D_3	1	13	121	163	1123	1243	!	\	^!	!]	<!	U	LLLLU
$L_5S_1D_3$	2	23	122	173	2123	1253	?	.	!-	!=	@!		LLLUL
$S_1L_5D_3S_1$	3		123	183	3123	1263	/	_	!:	!^	!#		LLULL
$L_5D_3S_2$	4		124	193	4123	1273	}	#	=!	!!	!"		LULLL
$L_5D_4S_1$	5		125	023	5123	1283	:	$	!{	!(	!,		ULLLL
$L_5S_1D_3S_1$	6		126	223	6123	1293	+	]	_!	:!	#!		
$L_5D_3S_1L_1$	7		127	323	7123	1230	{	~	.!	!$	!%		
$L_5D_3L_1S_1$	8		128	423	8123	1231	*	>	{!	[!	!/		
$L_5D_3S_1D_1$	9		129	523	9123	1232	<	,	!'	!}	!)		
L_5S_1			103	623	1023	1234	(	=	'!	![	$!		
$D_3L_5S_1$			113	723	1223	1235	%	^	(!	!+	`!		
$D_1L_5D_3S_1$			133	823	1323	1236	"	'	!`	+!	!.		
$L_5D_2S_1$			143	923	1423	1237	)	;	?!	!~	&!		
					1523	1238	`	[	%!	!<	|		
					1623	1239	-	&	*!	~!	!\		
					1723				!?	!*	]!		
					1823				!;	-!	}!		
					1923				!&	\!	/!		
					1203				,!	"!	;!		
					1213				!>	!@	>!		

Determining Password Changes

- If we have two previous passwords of the user, we can determine changes made between the passwords
- Predict the new password based on previous changes

Distance Matrix: “alice123!$” “12alice$!”

		a	**l**	**i**	**c**	**e**	**1**	**2**	**3**	**!**	**$**
	0	1	2	3	4	5	6	7	8	9	10
1	1	1	2	3	4	5	5	6	7	8	9
2	2	2	2	3	4	5	6	5	6	7	8
a	3	2	3	3	4	5	6	6	6	7	8
l	4	3	2	3	4	5	6	7	7	7	8
i	5	4	3	2	3	4	5	6	7	8	8
c	6	5	4	3	2	3	4	5	6	7	8
e	7	6	5	4	3	2	3	4	5	6	7
$	8	7	6	5	4	3	3	4	5	6	6
!	9	8	7	6	5	4	4	4	5	5	6

Calculating Damerau-Levenshtein Edit Distance

“alice123!$” “12alice$!”

		a	**l**	**i**	**c**	**e**	**1**	**2**	**3**	**!**	**$**
	0	1	2	3	4	5	6	7	8	9	10
1	1	1	2	3	4	5	5	6	7	8	9
2	2	2	2	3	4	5	6	5	6	7	8
a	3	2	3	3	4	5	6	6	6	7	8
l	4	3	2	3	4	5	6	7	7	7	8
i	5	4	3	2	3	4	5	6	7	8	8
c	6	5	4	3	2	3	4	5	6	7	8
e	7	6	5	4	3	2	3	4	5	6	7
$	8	7	6	5	4	3	3	4	5	6	6
!	9	8	7	6	5	4	4	4	5	5	6

Hierarchical algorithm
Simple base structures

alice123!$ → LDS
12alice$! → DLS

		L	D	S
	0	1	2	3
D	1	1	1	2
L	2	1	1	2
S	3	2	2	1

t

n

Level 1 edit distance = 1

Transpose back: 123alice!$, 12alice$!

		1	2	3	a	l	i	c	e	!	$
	0	1	2	3	4	5	6	7	8	9	10
1	1	0	1	2	3	4	5	6	7	8	9
2	2	1	0	1	2	3	4	5	6	7	8
a	3	2	1	1	1	2	3	4	5	6	7
l	4	3	2	2	2	1	2	3	4	5	6
i	5	4	3	3	3	2	1	2	3	4	5
c	6	5	4	4	4	3	2	1	2	3	4
e	7	6	5	5	5	4	3	2	1	2	3
$	8	7	6	6	6	5	4	3	2	2	2
!	9	8	7	7	7	6	5	4	3	2	2

n n d n n n n n t

Level 2 edit distance = 2

Determining Password Changes

- Total edit distance = level1 + level2
  = 1 + 2 = 3

alice123!$ → 12alice$!

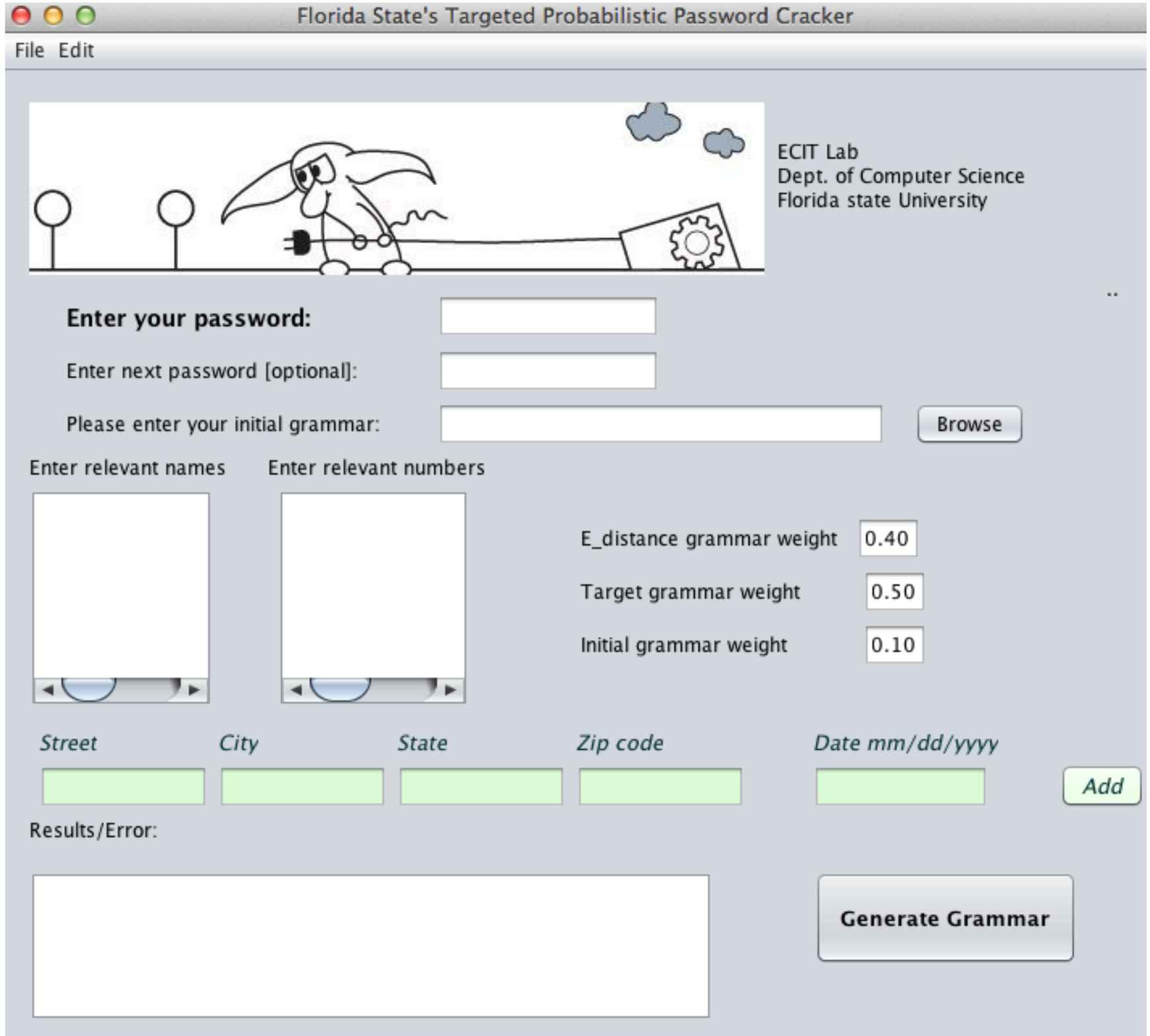
1. Transposition D and L components
   123alice!$
2. Remove 3 from D_3
   12alice!$
3. Transpose inside S_2
   12alice$!

Creating Targeted Grammar

- Get the information from Levenshtein edit distance
  - Increment / decrement a number by 1:
    - Example: old passwords: (alice125, alice126), we add [127,L_5D_3] to TGrammar

- Insertion of the same digit:
  - Example: old passwords: (alice, alice5), we add [55, L_5D_2] to TGrammar.

- Capitalization of alpha strings


Florida State's Targeted Probabilistic Password Cracker
File Edit
ECIT Lab
Dept. of Computer Science
Florida state University
Enter your password:
Enter next password [optional]:
Please enter your initial grammar:
Browse
Enter relevant names
Enter relevant numbers
E_distance grammar weight
0.40
Target grammar weight
0.50
Initial grammar weight
0.10
Street
City
State
Zip code
Date mm/dd/yyyy
Add
Results/Error:
Generate Grammar

How should I generate guesses?

- Use the edit 1 grammar. But I want to generate other guesses also. After all, the user might not have made small changes and might even have chosen a totally different password!
- This led us to the idea of merging probabilistic context free grammars. We can actually combine two different grammars and by extension any number of grammars!

The *Merge* of two grammars

- Let G_1 and G_2 be two probabilistic context-free grammars based on our structures of base structures and component structures. We construct a new grammar G_3 that we define as the *merge* of G_1 and G_2 and we represent it as:

$$G_3 = \alpha\, G_1 + (1- \alpha)\, G \qquad \text{where } 0 \leq \alpha \leq 1$$

- Consider a grammar rule R in G_1 or G_2. Let the probability of R in G_1 be r_1 and the probability of R in G_2 be r_2. (Note that if R is not in a grammar its probability is viewed as 0.) Then the probability r_3 of R in G_3 is:

$$r_3 = \alpha\, r_1 + (1 - \alpha)\, r_2$$

$L_5D_3S_1$	0.25
$L_5S_1D_3$	0.25
$L_5D_4S_1$	0.25
$L_5D_3S_2$	0.25
123	0.25
124	0.25
125	0.25
133	0.25
12	0.5
13	0.5
1234	0.5
1235	0.5
!	0.2
@	0.2
#	0.2
$	0.2
%	0.2
!!	0.33
!#	0.33
!@	0.33

Edit 1 Grammar

$W_1 = 0.8$

+

$L_4D_2S_1$	0.5
$L_3D_3S_2$	0.3
$L_5D_3S_1$	0.07
$L_6D_4S_2$	0.05
$L_8D_2S_1$	0.05
$L_5D_3S_2$	0.03
999	0.6
111	0.3
123	0.1
88	0.5
11	0.5
5656	0.5
1234	0.3
0909	0.2
!	0.4
)	0.3
?	0.2
%	0.1
!!	0.3
##	0.3
$#	0.2
!#	0.2

Initial Grammar

$W_2 = 0.2$

=

$L_5D_3S_1$	0.214
$L_5D_3S_2$	0.206
$L_5D_4S_1$	0.2
$L_5S_1D_3$	0.2
$L_4D_2S_1$	0.1
$L_3D_3S_2$	0.06
$L_6D_4S_2$	0.01
$L_8D_2S_1$	0.01
123	0.22
124	0.2
125	0.2
133	0.2
999	0.12
111	0.06
12	0.4
13	0.4
88	0.1
11	0.1
1234	0.46
1235	0.4
5656	0.1
0909	0.04
!	0.24
%	0.18
#	0.16
$	0.16
@	0.16
)	0.06
?	0.04
!!	0.324
!#	0.304
!@	0.264
##	0.06
$#	0.04

Input password:
pluto1995

pluto1995	pluto1915	1995Pluto
Pluto1995	pluto5995	1995plutO
plutO1995	pluto1395	1995plUto
plUto1995	pluto1195	1995pluTo
pluTo1995	pluto1095	1995pLuto
pLuto1995	pluto4995	2pluto1995
1995pluto	pluto1795	3pluto1995
1995	pluto9995	4pluto1995
pluto1985	pluto8995	7pluto1995
pluto1990	pluto0995	0pluto1995
pluto1975	pluto2995	5pluto1995
pluto1991	pluto1695	8pluto1995
pluto1994	pluto1595	9pluto1995
pluto1993	pluto1495	6pluto1995
pluto1992	pluto1895	pluto1234
pluto1999	pluto7995	1q2w3e4r
pluto1996	pluto3995	pluto!1995
pluto1998	pluto6995	123456
pluto1965	pluto1995e	pluto1995!
pluto1997	pluto1995r	!pluto1995
pluto1955	pluto1995s	pluto@1995
pluto1945	qwerty	pluto1995@
pluto1935	pluto1995E	@pluto1995
pluto1925	pluto1995R	pluto2008
pluto1295	pluto1995S	pluto2009
pluto1905	1pluto1995	pluto_1995

Testing Result

Old password	New Password	#of Guesses Targeted Attack	# of Guesses Regular Attack	Grammar
tharaborithor	thorborithara	--	--	--
Simba144!	@Simba2523	734,505,973	--	MGrammar
$unGl@$$220	$unGl@$$110	4,070	--	MGrammar
research!	Research!	554	5,059,949,503	EGrammar
starWars@123	star#Ecit@123	2,227,558	--	EGrammar
thebigblackdogjumps	blackdogmoretime	--	--	--
Ahk@1453	Ahk#1453	12,026	--	EGrammar
qpalzm73	qpalzm73*	1,810	--	EGrammar
pluto1995	boonepluto	--	--	--
caramba10	caramba12	14	11,424,542	MGrammar
Elvis1993!	Professional1993!2	--	--	--
Pepper88	peppergator88	128,197,109	2,563,504,751	Mgrammar

Testing Result

Old password	New Password	#of Guesses Targeted Attack	# of Guesses Regular Attack	Grammar
ganxiedajiA1!!	1ganxiedajiA	7,794	--	Mgrammar
88dolphins!	55dolphins!	38,503	--	MGrammar
kannj2013!	kannj2013	97	--	EGrammar
!FSU$qr335	!FSU$qr335mcddt	--	--	--
vballgrl77	schatzimae	--	--	--
nickc1007	corkn1007	--	--	--
sunflower12	sunflower13	202	119,336,969	EGrammar
meg51899	Meg51899*	5,381	--	EGrammar
Research1	research11	206	23,728,452	EGrammar
Gleek1993	Gleek1985	9,661	1,994,709,669	MGrammar
Oaklea0441	Oaklea0112	91,014	--	MGrammar

Old password1	Old password2	New password	Number of Guesses made to crack	Merged Or Edit distance grammar
russell	-	RUSSELL	1	Edit distance
russell1	-	russell	1	Edit distance
abc2009	-	pm2009	4,334,388	Merged
maverick	-	maverick7	118	Edit distance
dreamhope	-	hopehope	-	Merged
hopeful	-	hopeful1	14	Edit distance
starwars	-	starwars1	17	Edit distance
sweetie	-	sweetie1	20	Edit distance
krishna	-	krishnap	-	Merged
hope77	-	hope22	2,111	Merged
bland0608	-	plat0608	136,066,042	Merged
milena	-	Milena	4	Edit distance
milena	-	milene	-	Edit distance
bluemoon1	bluemoon2	bluemoon3	1	Edit distance
moonlight	-	redmoonlight	-	Merged
1writer	-	writer	1	Edit distance
1blackcat	-	blackcat	1	Edit distance
starwars	starwars5	starwars55	1	Edit distance
sweety	-	SWEETY	308	Merged
groove5721	-	Katie5721	-	Merged
171995	-	may171995	47,881,797	Merged
skymoon7	-	moon7sky	-	Merged
chomsky$po	-	po$chomsky	-	Merged
gamegreen	-	greendoc	-	Merged
d30023286	-	30023286	1	Edit distance
081983lori	-	081983	1	Edit distance
243currier	-	24378443	-	Merged
19632439	-	19632007	-	Merged
blackhawk	-	black7out	-	Merged

Thanks!
Questions/Comments?