

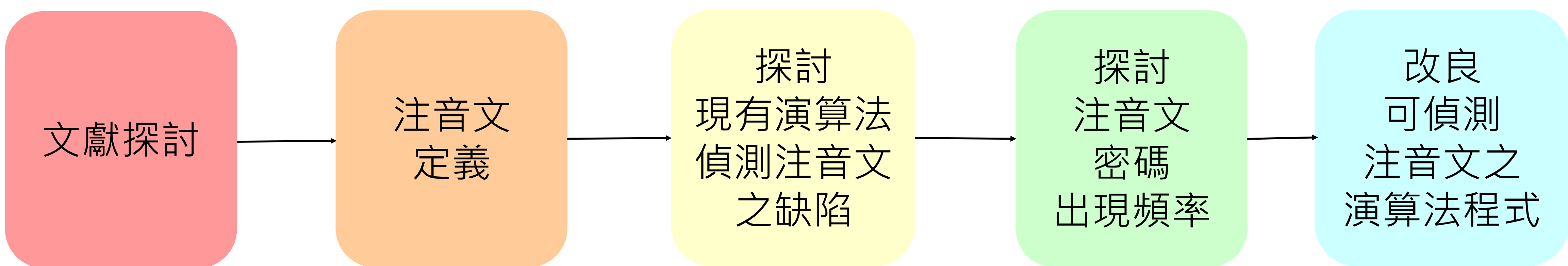
# 研究動機與目的

之前在社群網站曾看到一篇工程師貼出的文章，詢問為何“ji32k7au4a83”常被當作密碼使用？一看電腦鍵盤才知道，這組密碼是用注音文打出來的，而對應的中文是「我的密碼」。這讓我們對注音文密碼產生了興趣，注音文打出來乍看像亂碼，有一定的規律存在。但注音文密碼在各個網站建置新帳號密碼時，並沒有被列入密碼強度偵測判斷的程式中。因此，我們希望可以寫出一個程式來偵測密碼是否為注音文並輸出強度，藉此提醒使用者盡量不要使用注音文密碼當作帳號的密碼，協助使用者設定更為安全的密碼。

## 研究設備及器材

1. 筆記型電腦  
CPU : IntelCoreTMi5-8250U 1.80GHz  
RAM : 8.00GB  
作業系統 : Windows10 家用版 x64
  2. C++  
程式語言，使用版本為C++14
  3. Ubuntu  
Linux系統，使用版本為19.04

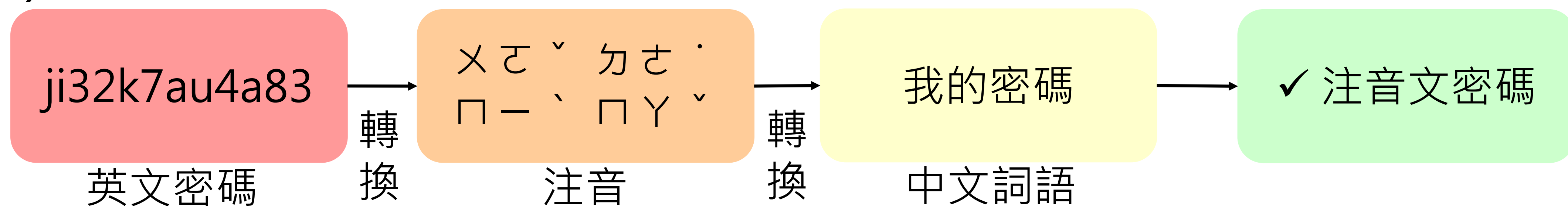
# 研究架構



## 研究過程及方法

## 1.名詞解釋&定義

(1) 注音文密碼 → 兩個注音單元連續



## (2) 注音單元

為可以轉換為中文單字的注音符號，例如：臺灣中的「臺」，其注音為ㄊㄞˊ，而ㄊ+ㄞ+ˊ即為合法的注音單元。

### (3) 密碼複雜度

一個密碼對抗猜測或暴力破解的有效程度，密碼複雜度越高，表示密碼愈難破解。

## 2.各個網站密碼準則

## (1)臺灣十大網站

台灣十大常用網站	jj32k7au4a83	ji32k7 au4a83	i32k7au4a83j	是否偵測注音文
1 google.com	強			無法偵測
2 youtube.com	與 (1) 共用密碼系統			與 (1) 共用密碼系統
3 pixnet.net	僅顯示密碼過短			無法偵測
4 ettoday.net	不會顯示強度，但會在變更前發送手機驗證碼			不會顯示
5 yahoo.com	不會顯示強度，會在變更密碼前要求再次登入			不會顯示
6 ltn.com.tw	無會員制度			無會員制度
7 google.com.tw	強			無法偵測
8 setn.com	無會員制度			無會員制度
9 facebook.com	中	強	強	無法偵測
10 momoshop.com.tw	不會顯示			不會顯示

## (2)世界十大網站

世界十大常用網站		ji32k7au4a83	ji32k7 au4a83	i32k7au4a83j	是否偵測注音文
1	google.com	強			無法偵測
2	youtube.com	與 (1) 共用密碼系統			與 (1) 共用密碼系統
3	baidu.com	僅會顯示密碼過短			無法偵測
4	tmall.com	不會顯示			不會顯示
5	qq.com	不會顯示			不會顯示
6	sohu.com	不會顯示，但會在修改密碼前要求手機驗證			不會顯示
7	facebook.com	中	強	強	無法偵測
8	taobao.com	與 (4) 共用密碼系統			與 (4) 共用密碼系統
9	wikipedia.org	僅會輸出密碼太短			無法偵測
10	vahoo.com	不會顯示強度，會在變更密碼前要求再次登入			不會顯示

### 3.密碼選用

### (1) 注音文密碼

原密碼	轉為注音	中文	預期種類	調整後	外洩次數
vm6jo3xu4	ㄊㄩˋ ㄇㄨˋ ㄐㄧㄣˋ ㄨˋ ㄎㄞˋ	徐偉立	名字	Vm6jo3xu4	6
ej03xu3admin	ㄍㄨˋ ㄇㄨˋ ㄐㄧㄣˋ ㄨˋ admin	管理admin	中文+英文	Ej03xu3admin	4
1u3ru41p3	ㄅㄨˋ ㄨˋ ㄩˋ ㄞˋ ㄅㄨˋ ㄕㄨˋ	筆記本	日常用字	1U3ru41p3	34
ji394su3	ㄐㄧˋ ㄙㄨˋ ㄕㄨˋ ㄙㄨˋ	我愛你	日常用字	Ji394su3	23,918
ji3cp3qj94	ㄐㄧˋ ㄙㄨˋ ㄕㄨˋ ㄕㄨˋ ㄕㄨˋ ㄕㄨˋ	我很帥	日常用字	Ji3cp3qj94	439

修改後的注音文密碼為將密碼的第一個英文字母改為大寫。修改前後的注音文密碼會一同與強密碼比較。

## (2) 強密碼

利用Google新建帳號設定密碼時，會出現的建議密碼，選十組進行比較。

aipdxZVU3qyu5WE	xnZtfvwUmws6Q8U	6dFni9dgerF8biZ	Lv96yJHxcn9h9Tw	8ZaQHxHUdb3Khe8
2YQAEksXriJU9ge	d3ZCNAGYiik3HvQ	qYUaewD85QCJRLE	JCLcTqdQKzGK9AG	3LwRjAqjb6jGWF2



# 研究結果及討論

## 1. 注音文密碼特性

我們偵測了Github下載的國外外洩密碼庫，發現有25筆注音文密碼。  
此25組注音文密碼特性：

分類	數量	舉例1	舉例2	舉例3
常用詞語	12	1u4xm35l4qu04 ( 畢旅照片 )	cj6xu6 ( 狐狸 )	au4a83 ( 密碼 )
名字	7	tp6qo4hp6 ( 陳佩岑 )	ck6rup xj/6 ( 何金龍 )	vm6u/4vu ( 徐映晰 )
自戀語詞	3	ji3cl3gj94 ( 我好帥 )	cl3gj94 ( 好帥 )	ji3cp3gj94 ( 我很帥 )
有關愛情	3	94cp3xo4 ( 愛很累 )	ji394su3 ( 我愛你 )	yjo494qo4fu6 ( 最愛佩琪 )

## 2. 偵測注音文密碼強度的演算法與程式

以下為原選用之10組強密碼與10組注音文密碼之強度檢測比較：

項目	強密碼	注音文密碼 ( 原10組 )
分數 ( 滿分4分 )	4分	2.1分
猜測數 ( $\log_{10}$ )	24.0651	9.1623
密碼複雜度	79.9421	30.6424

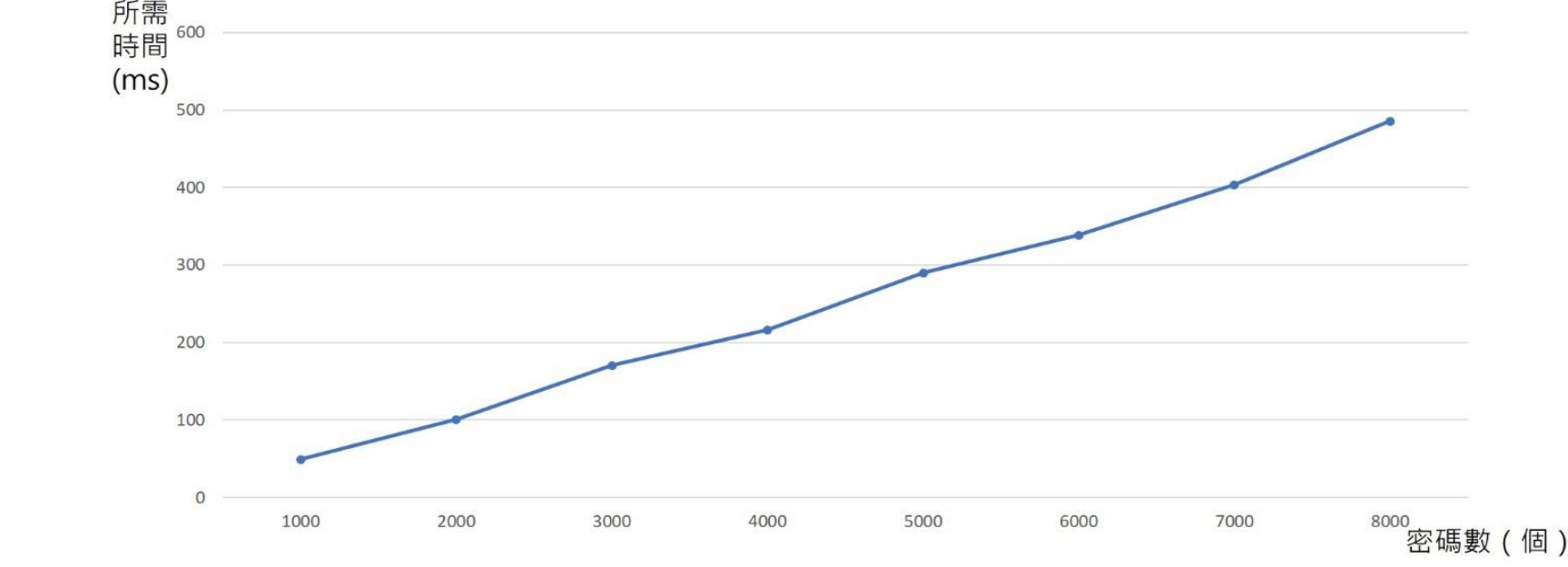
為更了解本演算法對判斷注音文密碼強度的效用，我們對國外密碼庫新偵測到的25組注音文密碼進行檢測，以下為檢測結果：

項目	注音文密碼 ( 新25組 )
分數 ( 滿分4分 )	1.4分
猜測數 ( $\log_{10}$ )	8.551
密碼複雜度	28.339

我們可以從兩表得知，本研究所研發的演算法除了可以將注音文密碼的強度降低為中至中弱的強度，同時不會影響強密碼，是一個十分有效的演算法。

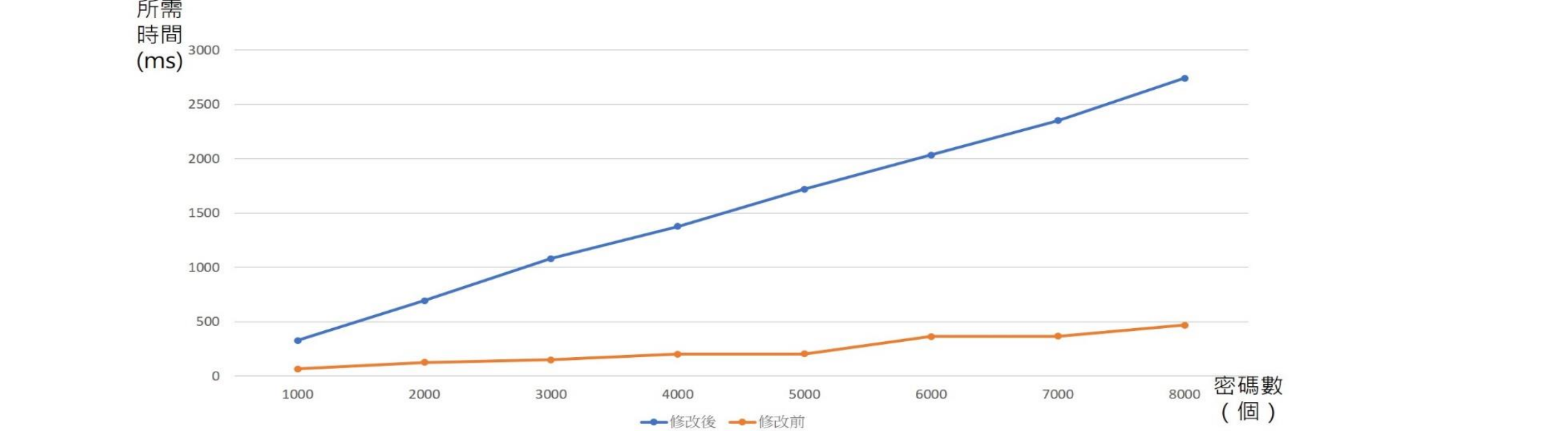
## 3. 效能分析

### (1) 注音文偵測程式



當密碼個數為n，注音文偵測分析程式效率為O(n)。

### (2) 考量注音文密碼強度程式



當密碼個數為n，偵測注音文密碼強度程式效率為O(n)。

### (3) 注音文密碼強度程式單筆偵測速度

平均計算出一個注音文密碼強度所花時間為0.68毫秒。由此可知，本演算法不僅有效，速度也十分快速，是一效能十分高的演算法。

# 結論

- 本研究研發之注音文密碼偵測演算法可將注音文密碼所顯示的強度降低，並加上警語，且效率佳，但強密碼的顯示強度依然不變，以協助使用者設定更為安全的密碼。
- 從國外外洩密碼庫偵測使用者常用來設定注音文密碼規則如下：  
(1) 直接將注音打出來 (2) 將注音文與英文混雜。  
注音文密碼的取材來源為：  
(1) 常用詞語 ( 通常是名詞 ) (2) 自己的名字 (3) 自戀語詞 (4) 有關愛情。
- 此模式及演算法可套用至各種拼音文字，可進一步協助使用者設定更安全的密碼。

# 參考資料

- Ding Wang and Ping Wang, *Peking University* ; Debiao He, *Wuhan University* ; Yuan Tian, *University of Virginia* .2019 “Birthday, Name and Bifacial-security: Understanding Passwords of Chinese Web Users” In Proceedings of the 28<sup>th</sup> UNENIX Security Symposium, pages 1537-1554
- Xavier de Carné de Carnavalet and Mohammad Mannan, Concordia Institute for Information Systems Engineering .2014 “From Very Weak to Very Strong : Analyzing Password-Strength Meters”
- Daniel Lowe Wheeler, Dropbox Inc. 2016“zxcvbn: Low-Budget Password Strength Estimation” In Proceedings of the 25<sup>th</sup> UNENIX Security Symposium, pages 157-173