

# 資工系實務專題研究計畫成果報告

## Meaningful Chinese CAPTCHA 中文人機辨識碼

專題編號：098CSIE-S004

執行期限：97 年 1 學期至 98 年 1 學期

指導教授：楊士萱 教授

專題計劃參與人員：方偉德、楊羚

### 一、中文摘要

CAPTCHA「Completely Automated Public Turing test to tell Computers and Humans Apart」人機辨識碼。現在網路系統發達，在留言、註冊帳號、下載檔案等功能中，加入這 CAPTCHA 的機制，以防止被透過程式，造成大量的 spam 留言，註冊、下載；而為了防止這些困擾，我們製作了人機辨識碼，以防止機器人惡意的灌水和註冊。

### 二、緣由與目的

目前市面上的人機辨識碼，都是採用數字與英文字的結合，或是搭配背後的雜點做防護，這類型的人機辨識碼已經十分普及，而且也早已被破解，所以只能用不同的雜訊組合作搭配，或是用照片或是數學問題來製作辨識系統，但往往人類要進入此網站之前，都需要花很大的力氣，才能證明自己是人，而不是機器人。

目的：我們製作的人機辨識碼，使用的是**有意義的中文字詞**，加上背後的扭曲與雜訊做防護，當人類在進入系統時，可以**根據人類學習過有意義的詞來判斷中文組合字的意思，因此驗證碼的雜訊系數可提更高，以防止機器人不斷辨識、猜測而破解。**

### 三、專題目標

專題目標希望在網路上應用，Server 端有利用到資料庫時，或是網站需要基本防護時，可利用驗證碼驗證。

#### ●防護的方式：

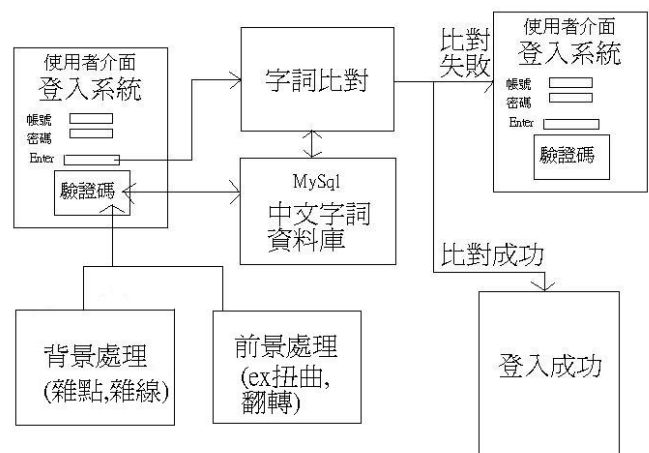
1. 希望利用雜點、斜線、扭曲、大小、字體和傾斜的方式讓文字辨識軟體無法辨識。

目前資料庫有百餘字詞(可再增加／刪減)，讓機器人無法隨機猜測字詞(中文字詞有萬種搭配)。

### 四、開發工具

1. AppServ
2. MySql
3. Dreamweaver
4. Flash
5. IE

### 五、系統架構



Chinese CAPTCHA 系統架構圖

說明：

以使用者登入介面為範例，登入時須輸入帳號、密碼及驗證碼。驗證碼的图片由背景雜訊處理加上從資料庫隨機抽出字詞做合併。使用者輸入完驗證碼後，會傳到背後做字詞比對，輸入正確會到登入成功頁面，輸入失敗會要求使用者重新輸入。

### 六、使用技術

#### (一)使用者介面

使用者介面包括登入系統及註冊系統，兩部分皆有驗證碼之驗證，在client端呼叫captcha5.php產生驗證碼的图片。註冊系統中設有防呆處理，當會員輸入資料格式有誤時，會顯示警告ex. 生日格式須依照民國/年/月輸入。

#### (二)背景處理

1. 雜點：在底圖上隨機位置產生 10000 個黑色像素點
2. 雜線：底圖上隨機位置產生 60 度到-120 度之黑色直線
3. 弧線：底圖上隨機位置產生 10 條大小，方向不等的弧線

### (三)前景處理

1. 字體變化：隨機產生下列四種不同字體。
2. 字體大小：隨機產生 23~30 之間大小的字體。
3. 字體傾斜：隨機產生-60 度~60 度之間角度的傾斜字體。
4. 扭曲：利用 `imagesetpixel()` 的函式，透過我們撰寫的旋波數學公式( $\cos$  和  $\sin$  的波形)來作旋波的扭曲變形。
5. 位置變化：由初始位置  $(x, y) = (\text{rand}(0, 50), 0)$  開始， $y$  軸隨機產生座標，而  $x$  軸則每字間隔 55 之距離產生座標，為了避免字詞疊在一起，導致使用者難以辨識。

(四)中文字詞資料庫：目前有百餘筆字詞。

### (五)字詞比對：

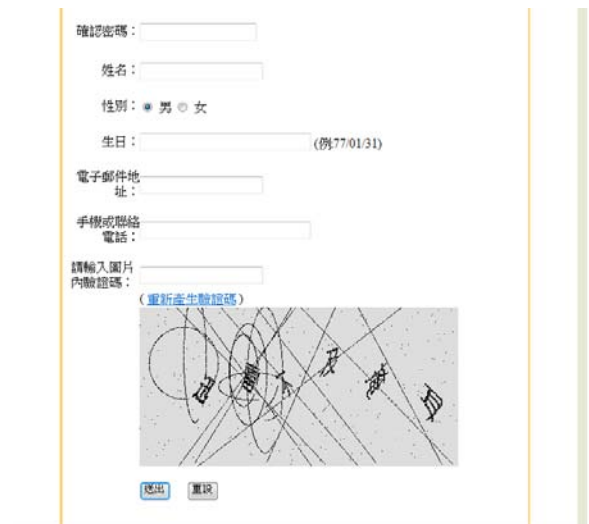
使用 SESSION 儲存資料庫提取的字詞，並把使用者輸入驗證碼的內容，存入 POST 傳回最比對，比對正確便登入（註冊）成功，比對不相符則會擋住使用者的頁面，並不會保存註冊之資料，要求使用者在輸入一次。

## 七、使用者主要介面

主要驗證碼應用介面：



login.php 的登入畫面



register.php 的註冊畫面

## 八、完成後之測試成果

雖然把扭曲系數調很高，但一般人皆可根據前後字意猜測答案，另外我們也會有”重新整理驗證圖片”之功能，因為使用者並不是百分之百都能看得懂驗證碼的內容，但又避免機器人因按”重新整理驗證圖片”之功能，而記起資料庫之內容，我們讓這重新整理的功能設定為 5 次，當重新整理超過 5 次，便會鎖住資料庫的內容，在圖片上顯示”error”，我們在機器人與人類判斷中達平衡，不但可防止破解，也讓使用者能順利驗證人類的身分。

## 九、參考文獻

- 【1】作者：廖奕齊, 碩士論文：A CAPTCHA Mechanism By Exchanging Image Block
- 【2】PHP5 網管實驗室 [www.php5.idv.tw](http://www.php5.idv.tw)
- 【3】PWNTCHA [caca.zoy.org/wiki/PWNTcha](http://caca.zoy.org/wiki/PWNTcha)
- 【4】My SQL 5 徹底研究 Michael Kofler
- 【5】挑戰 PHP/MYSQL 文淵閣工作室編著

