第四章　實驗步驟

本研究的實驗步驟共有分五大項，分別是建立新聞資料、將檔案加密、加入驗證碼、解密前的檔案驗證、將檔案解密其功能說明如下：

一、建立新聞資料

加密新聞資料之前我們必須取得預先處理好的新聞資料，新聞資料表的欄位組成如圖 4-1 其欄位說明如圖 4-2：

Paper

PaperID	PaperTitle	PaperClass	PaperAuthor	PaperText	PaperTime	PaperFile	PaperEncrypt
41	翁奇楠命案兇嫌	社會	東森新聞記者謝		/13 下午 03:50:39	NOHITATUSBKW	
42	怒線民拿錢不辦	社會	中廣新聞網		/13 下午 03:51:52	QLCORXVBOXLI	
43	駕船繞世界，加	社會	中廣新聞網		/13 下午 03:52:21	KHBGYCKJMEM	
44	吳敦義：扁當總	社會	張德厚		/13 下午 03:53:33	QOEXQNKDUL	
45	黑道倫理 獄中過	社會	陳凱勛、林欣儀		/13 下午 03:53:52	PVAFCDIDANQO	
46	伏襲翁奇楠 廖嫌	社會	陳凱勛、林欣儀		/13 下午 03:55:02	ESXNCTKMMFI	
47	睡醒後就失憶 她	科技	羅彥傑		/13 下午 03:59:42	PEXCLHW	
48	挑戰公審會 邀社	科技	中央社記者溫貴		/13 下午 04:00:19	HRGCBIUHGMGJ	
49	俄模搭飛機隆乳	科技	中央社記者溫貴		/13 下午 04:02:15	XIOCQKYDXQJEI	
50	南非邪教要信徒	體育	中廣新聞網		/13 下午 04:02:57	OXNKVVRSZIM	
51	國際特赦籲吉爾	體育	黃啓霖		/13 下午 04:03:38	RELOUOLNGPMF	

圖 4-1　新聞資料表

新聞編號	標題	類別	作者	內文	建立時間	索引	加密與否

圖 4-2　新聞資料欄位

新聞編號：為系統自動編號。新聞標題：每一筆新聞的標題儲存於此。新聞類別：從程式做類別管理或是從資料庫直接建立資料，如圖 4-3。新聞作者：該新聞資料的原始作者。新聞內文：新聞內文是藉由後面的新聞索引讀入額外儲存的檔案內容，文章預設是儲存成.txt 格式如圖 4-4。

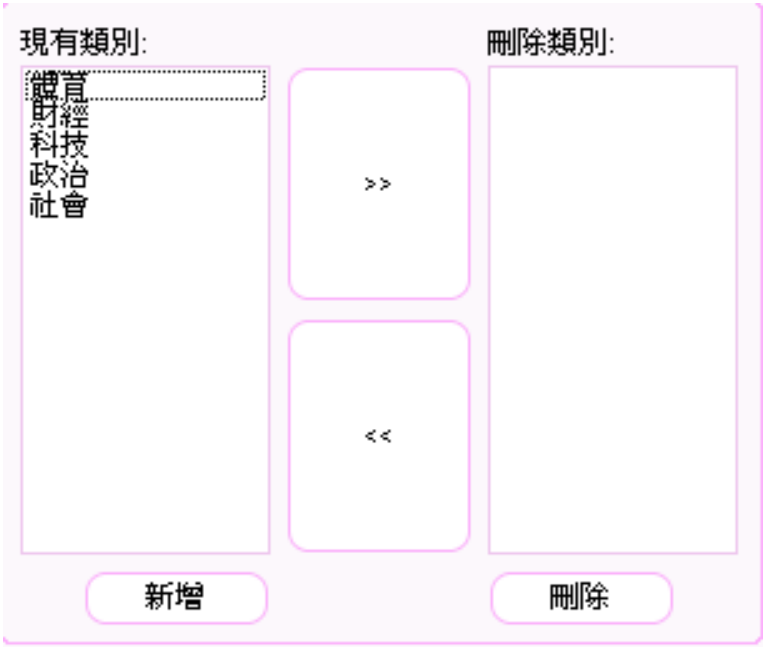

圖 4-3　類別管理

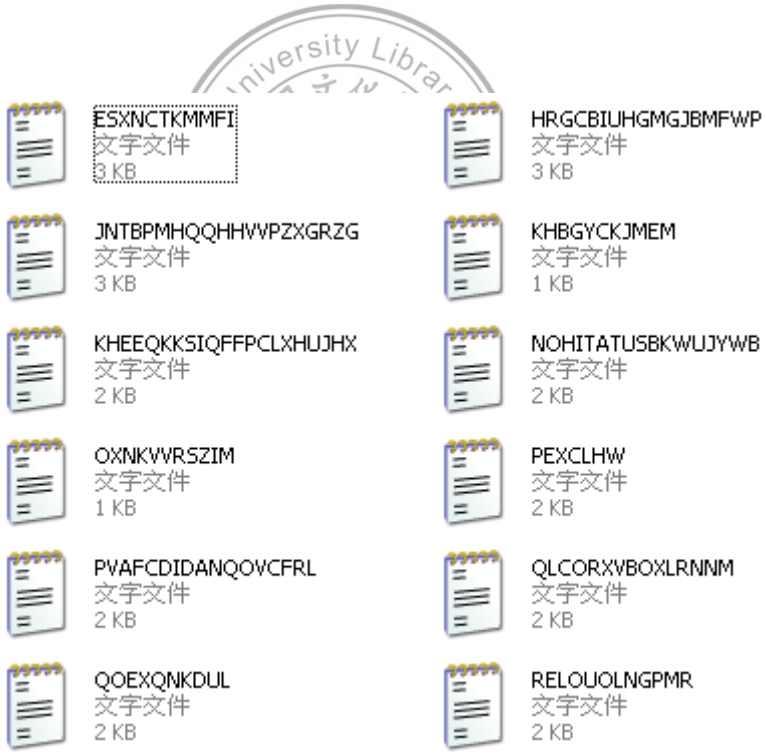

圖 4-4　新聞資料儲存格式

新聞建立時間：每一筆新聞資料建立的時間。新聞索引：如同圖 4-3 所示，每一筆新聞資料在儲存的時候會由系統給予

索引碼，每篇文章的索引碼為檔案名稱，檔案名稱為 10~20 個英文字，藉由亂數名稱來初步防範非法入侵者的搜索，新聞索引也是解密時用來計算檔案位置的依據。加密與否：加密過的文章在資料庫會顯示已加密。

二、將檔案加密

檔案的加密分為以下十個步驟，分別是獲取新聞資料、設定加密資料表、分割欄位、旋轉、左移、位置交換、輸出方向改變、建立驗證碼表、創建加密後的檔案、插入左移表、驗證碼表、加密資料表，詳細說明如下：

(一)獲取新聞資料

使用預先分類處理好的新聞檔案或是新增一筆資料如圖 4-5。

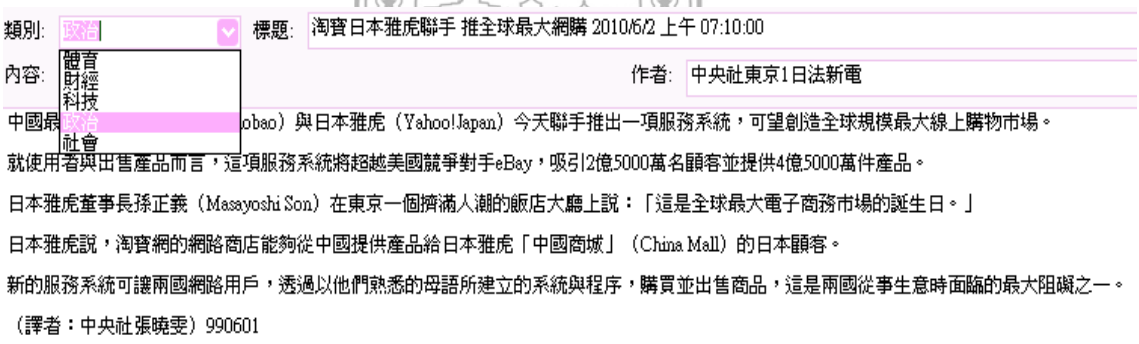

圖 4-5　新增新聞資料

預設類別為體育新聞、財經新聞、科技新聞、政治新聞、社會新聞。

(二)獲取新聞資料

選取新聞資料，並且設定加密資料表如圖 4-6。

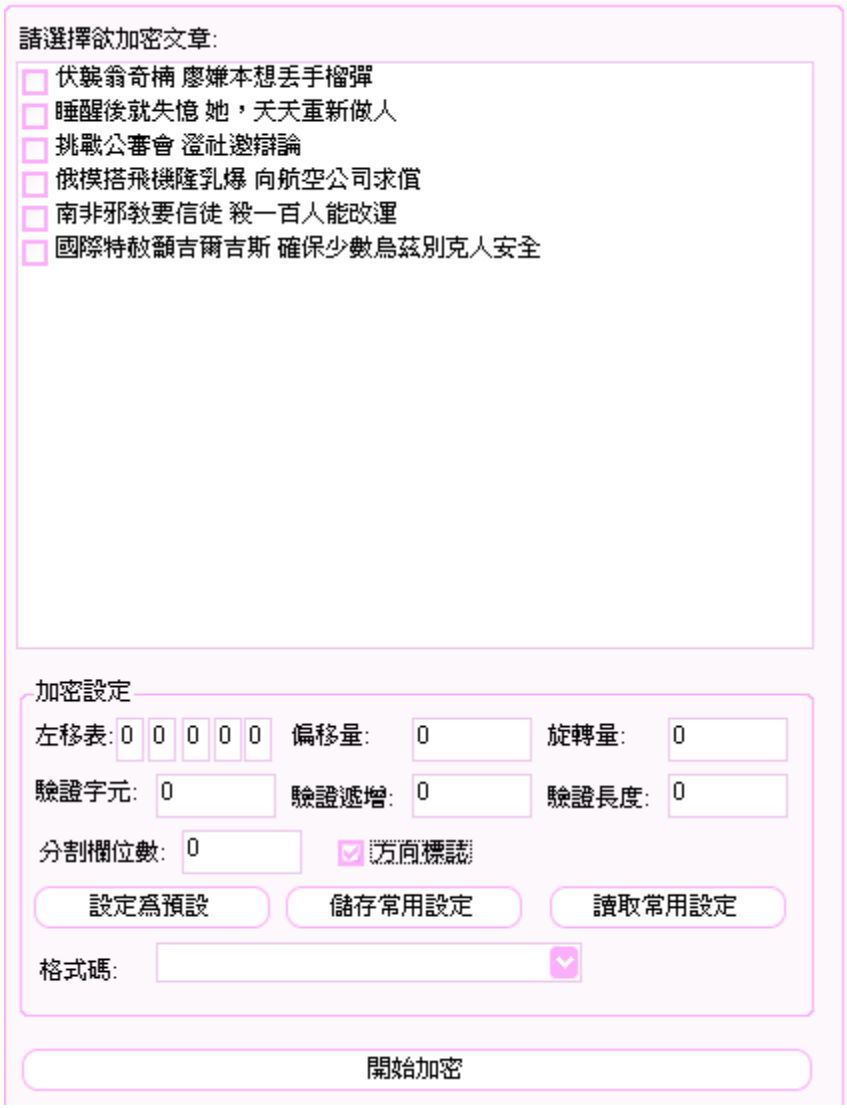

圖 4-6　設定加密資料表頁

之後設定好各欄位並選擇格式碼以及方向標誌(預設方向標誌是向右輸出)如圖 4-7。

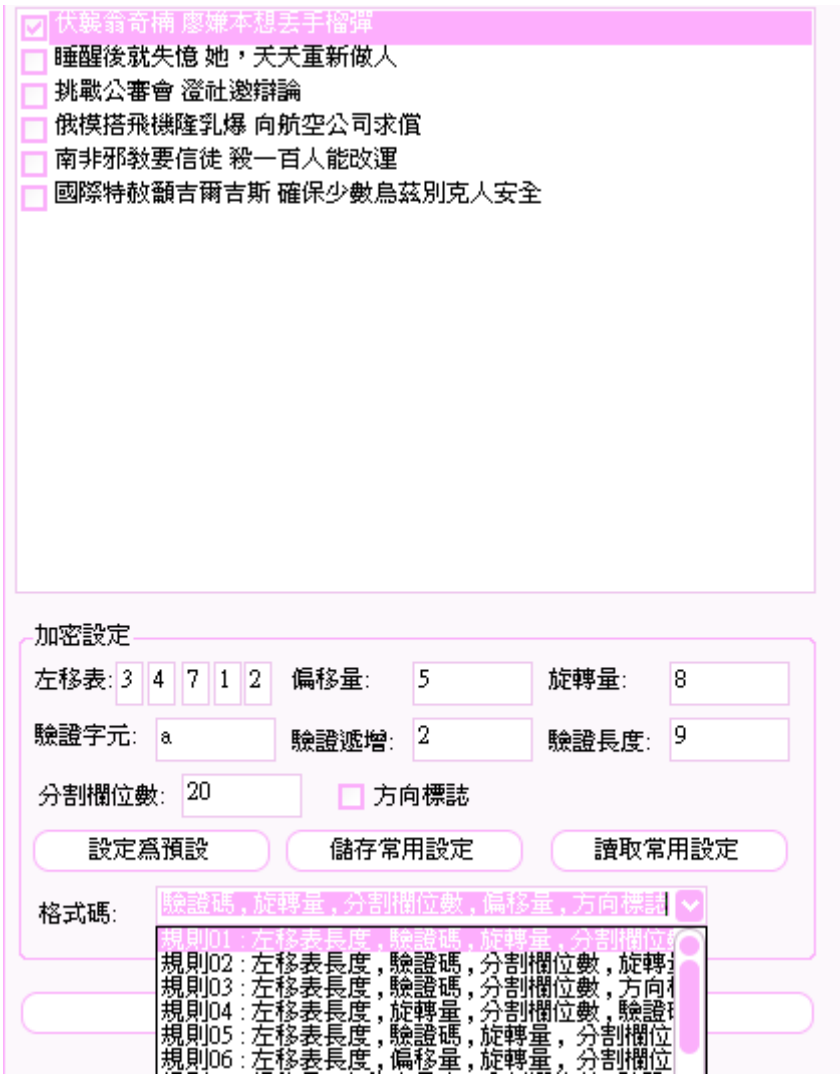

圖 4-7　設定加密資料表

依照格式碼的不同，分別將加密表單鍵值以不同順序儲存，如圖 4-7，格式碼是左移表長度、旋轉量、偏移量、驗證碼、分割欄位數、方向標誌去排列組合，共有 6 階乘種變化。

(三)分割欄位

我們在加密時是先將文章讀入一個陣列以藉此固定位置，確定文章的開頭與內文的位置，之後將其按照加密資料表的分割欄位數來分割區塊，最後區塊不足時加入多

餘符號，產生多餘符號表。然後開始加密，如圖 4-8。

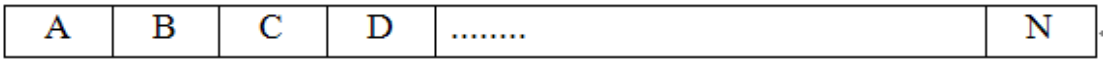

圖 4-8　分割後的明文

(四)旋轉

取得加密資料表中旋轉量(RB)，將多餘符號表中區塊從起始的區塊重複向左或向右旋轉。如圖 4-9，從區塊 A 開始向左向右旋轉。

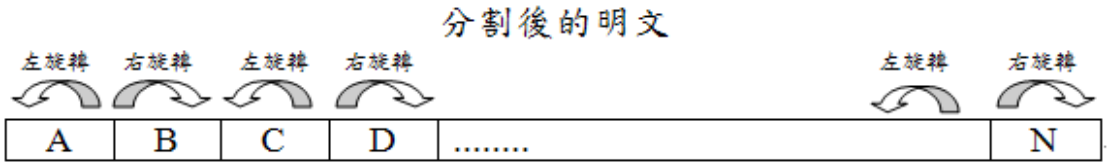

圖 4-9　檔案旋轉

區塊 A 向左旋轉 RB 位元組，區塊 B 向右旋轉 RB 位元組，依序從檔案開頭到結尾，取得旋轉後符號表。

(五)左移

取得加密資料表中左移表長度及左移表的值，根據設定的值，把每個位元左移 n 個位元。如圖 4-10。

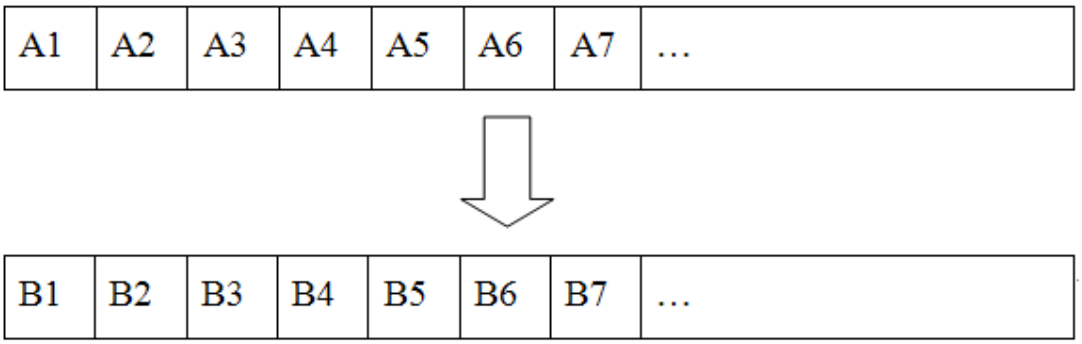

圖 4-10　左移

左移表為 7346123，B1 為 A1 左移 7 個位元；B2 為 A2 左移 3 個位元；B3 為 A3 左移 4 個位元；B4 為 A4 左移 6 個位元；B5 為 A5 左移 1 個位元；B6 為 A6 左移 2 位元；B7 為 A7 左移 3 個位元。重複直到旋轉後符號表結束為止。得到左移後符號表。

(六)位置交換

從加密資料表獲得偏移量。建立位移後符號表。將左移後符號表，順序每隔偏移量取出，加到位移後符號表後面，一直到左移後符號表結束。把偏移量減少 1，重複做上一個步驟。直到偏移量等於 0，得到位移後符號表。

以偏移量 4 為例，如圖 4-11。

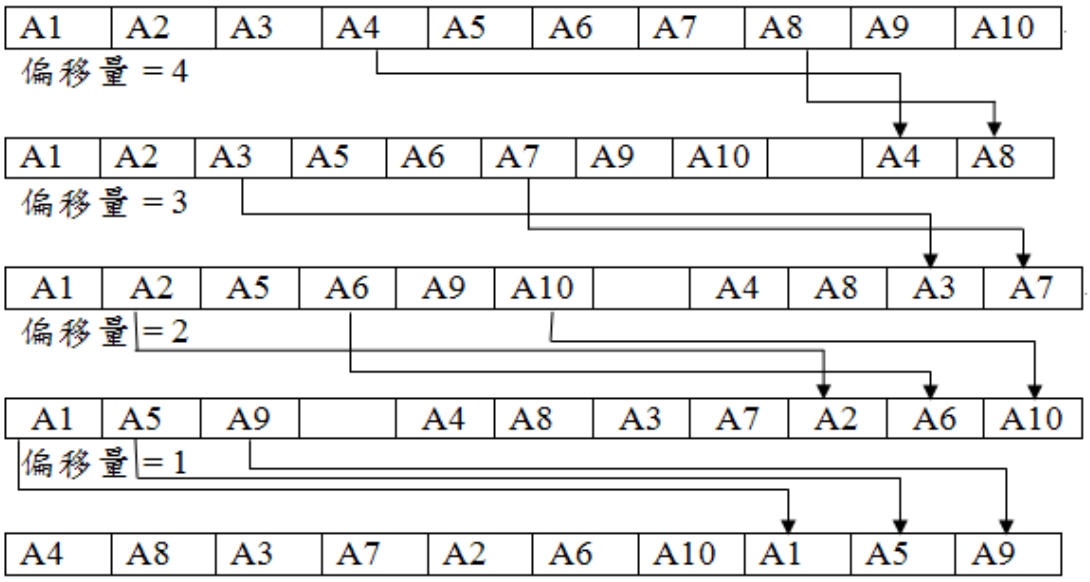

圖 4-11　位置交換過程

(七)輸出方向改變

加密時選擇輸出方向標誌。如果方向標誌已經設定了，就把位移後符號表反轉，如圖 4-12，反轉完成後獲得反轉後的符號表。

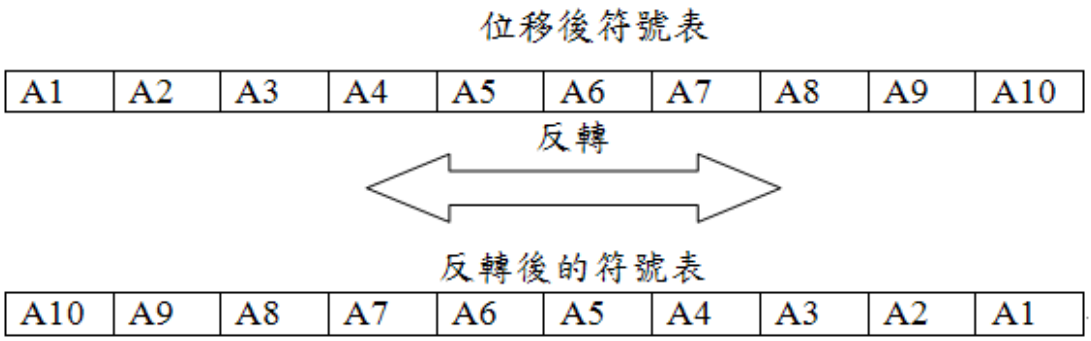

圖 4-12　輸出方向改變過程

(八)建立驗證碼表

驗證碼的儲存格式如圖 4-13。

編碼	偏離量	增加的值	編碼數
A	3	2	4

圖 4-13　驗證碼儲存格式

假設密文內容 ： How are you？

把密文分割成之後存進符號表如圖 4-14。

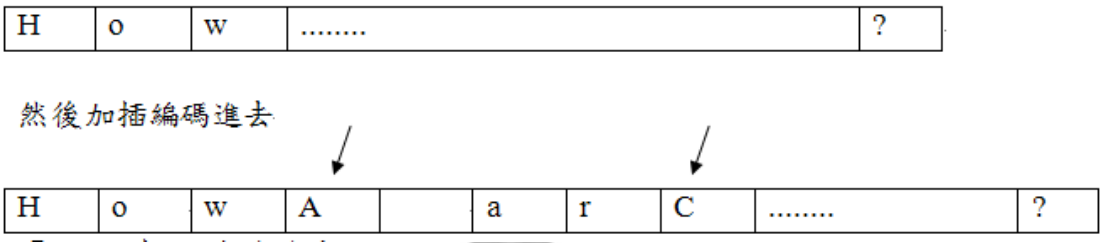

圖 4-14　插入驗證碼過程

編碼 a 偏移量 3 增加的值 2 代表在從開頭算第三個字元插入編碼，之後每 3 個字元插入第 a+2(*n)個編碼，邊碼數為總共插入的編碼總數。

(九)創建加密後的檔案

經過分割欄位、旋轉、左移、位置交換、輸出方向改變等六個步驟之後，我們輸出加密過的檔案。

(十)插入左移表、驗證碼表、加密資料表

透過關鍵代碼和檔案長度，我們計算新聞資料位置 1 並且插入左移表到轉後的符號表。

我們計算新聞資料位置 2，並且加入驗證碼表到轉後的符號表。

我們計算新聞資料位置 3，並且把加密資料表加入到轉後的符號表，就創建了加密新聞資料，如圖 4-15。

新密文格式

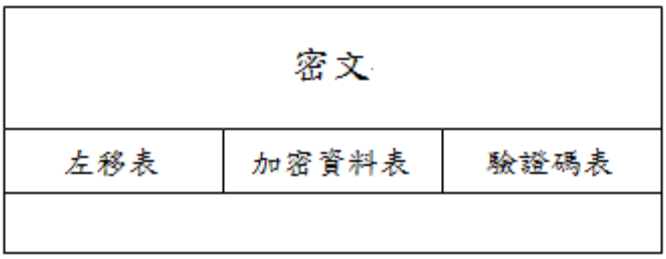

圖 4-15　新密文格式

圖 4-16 是加密完成的範例，加密完成後的密文是由一連串無法直接解譯的亂碼所組成的，如圖 4-17。

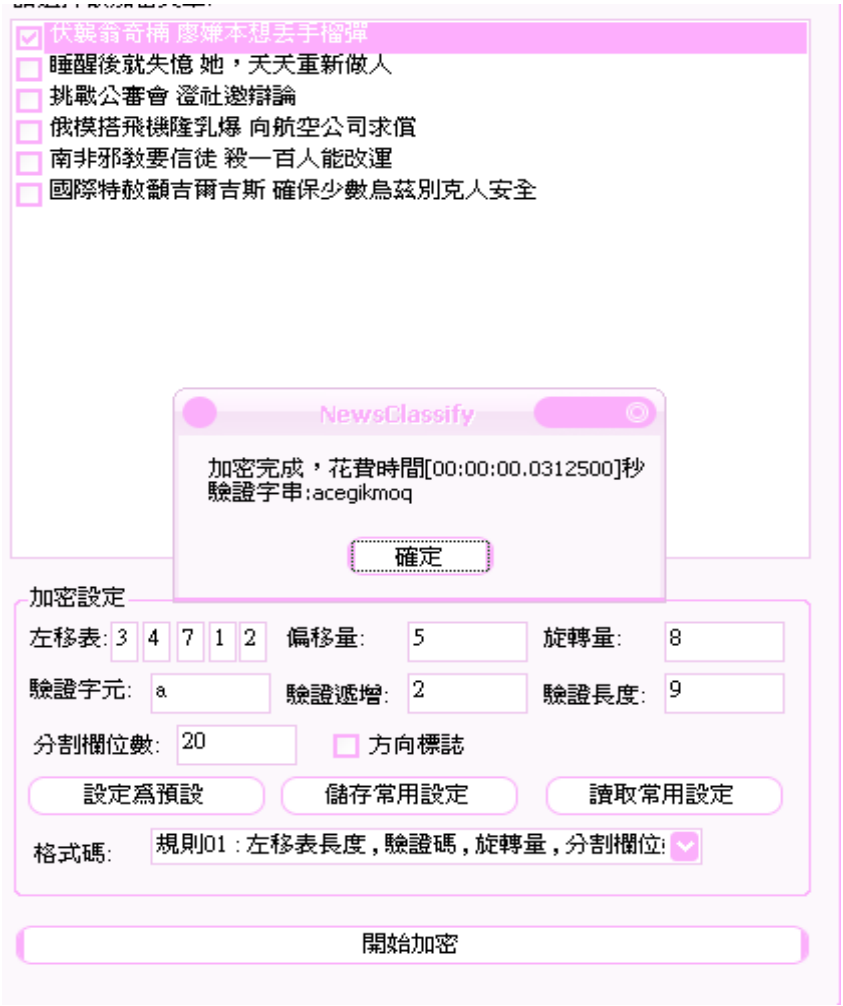

圖 4-16　加密完成

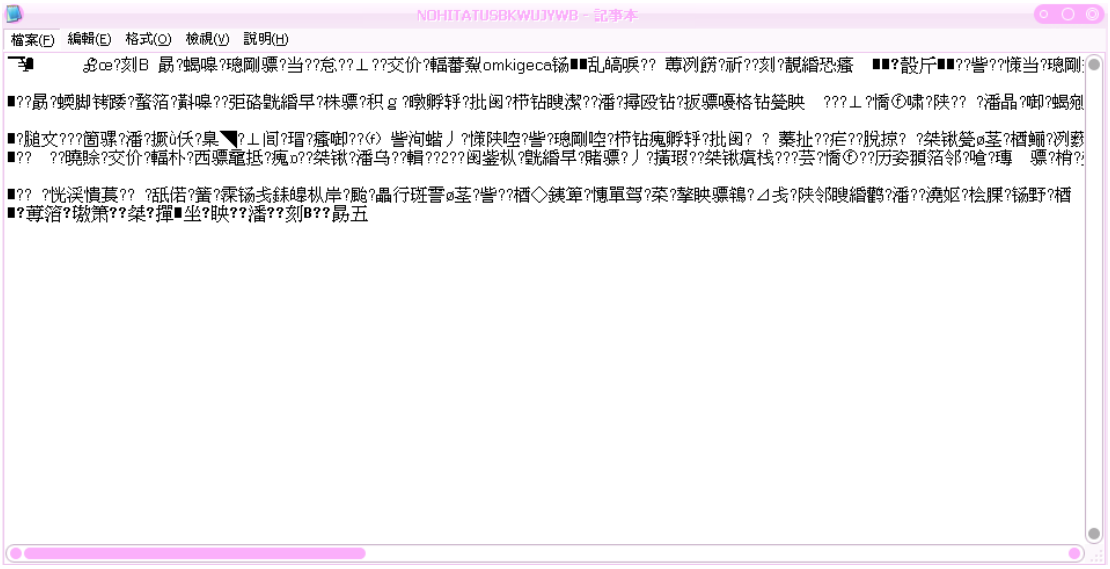

圖 4-17　加密過的密文

三、檔案驗證

在解密之前，我們要將檔案進行驗證，才能知道他是不是被竄改，驗證的步驟由獲得加密後的檔案、從密文中計算出檔案位置、將檔案解密、資料驗證所組成，詳細說明如下：

(一)獲得加密新聞資料

首先獲得加密新聞資料如圖 4-18。

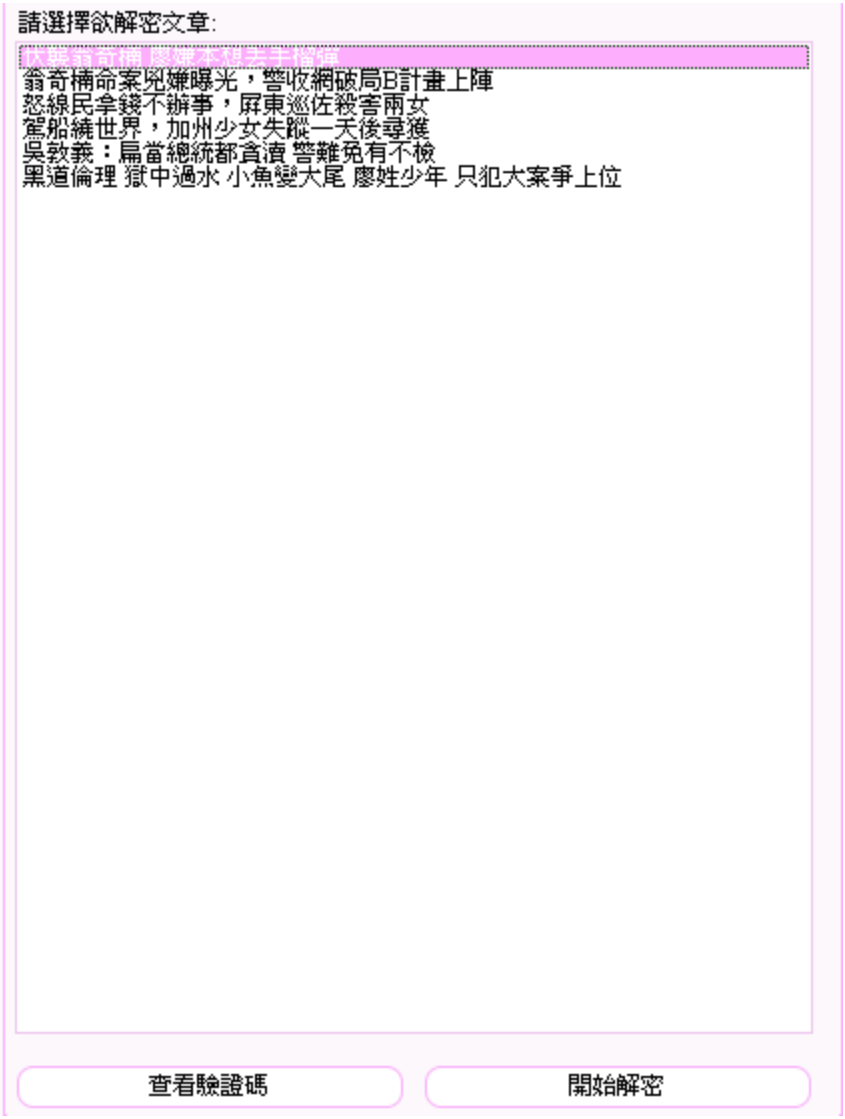

圖 4-18　獲得加密的檔案

(二)從新密文中計算出檔案位置，驗證步驟如下：

1. 獲得加密資料表。

   (1)獲得加密新聞資料。

   (2)從新聞資訊資料庫獲得檔案長度和關鍵代碼。

   (3)從檔案長度、關鍵碼來計算新聞資料位置 3。

   (4)獲得加密資料表。

2. 獲得加密資料表。

   (1)從加密資料表獲得驗證碼如圖 4-19，建立驗證碼表。

如圖 4-20。

編碼	偏移量	增加的值	編碼數
a	3	2	9

圖 4-19　驗證碼

a	c	e	g	i	k	m	o	q

圖 4-20　驗證碼表

(2)從檔案長度和驗證碼，我們計算出新聞資料位置 2。

(3)從新聞資料位置 2 和驗證碼，我們得到檢驗碼表。

(4)如果檢驗碼表和存放驗證碼表相同，儲存的資料就是原始沒有被更動過的資料。

如範例所示可以得出第一個編碼為 a，藉由增加的值可以得知編碼規則為 a+2(n)。

則提取出來的編碼為 a、c、e、g、i、k、m、o、q。將提取出的檢驗碼表和圖 4-20 的驗證碼表比對。

驗證會有四種狀況，分別是驗證成功、驗證碼長度錯誤、驗證碼不符合產生規則、無法讀取驗證碼。

規則上相符合就是資料沒有遭到竄改，如圖 4-21，4-22 為驗證碼長度錯誤，4-23 為驗證碼不符合產生規則，4-24 為無法讀取驗證碼。

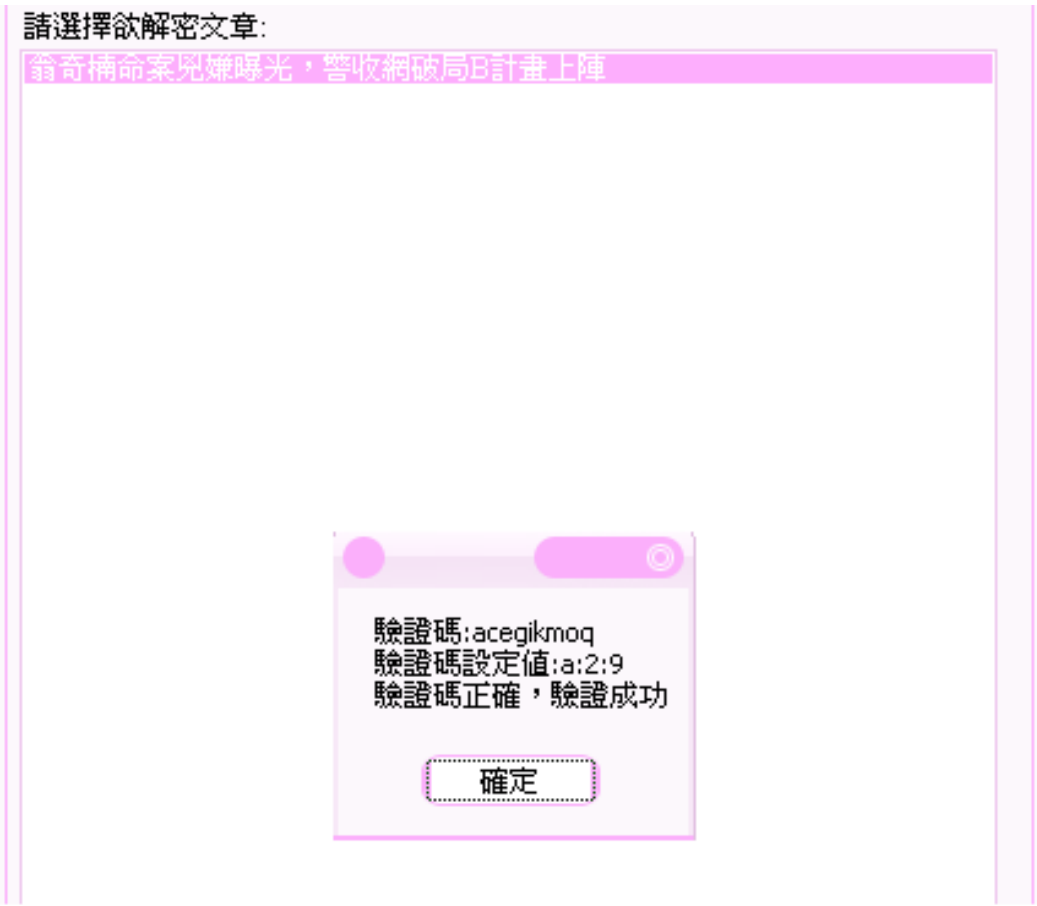

圖 4-21　驗證碼正確，驗證成功

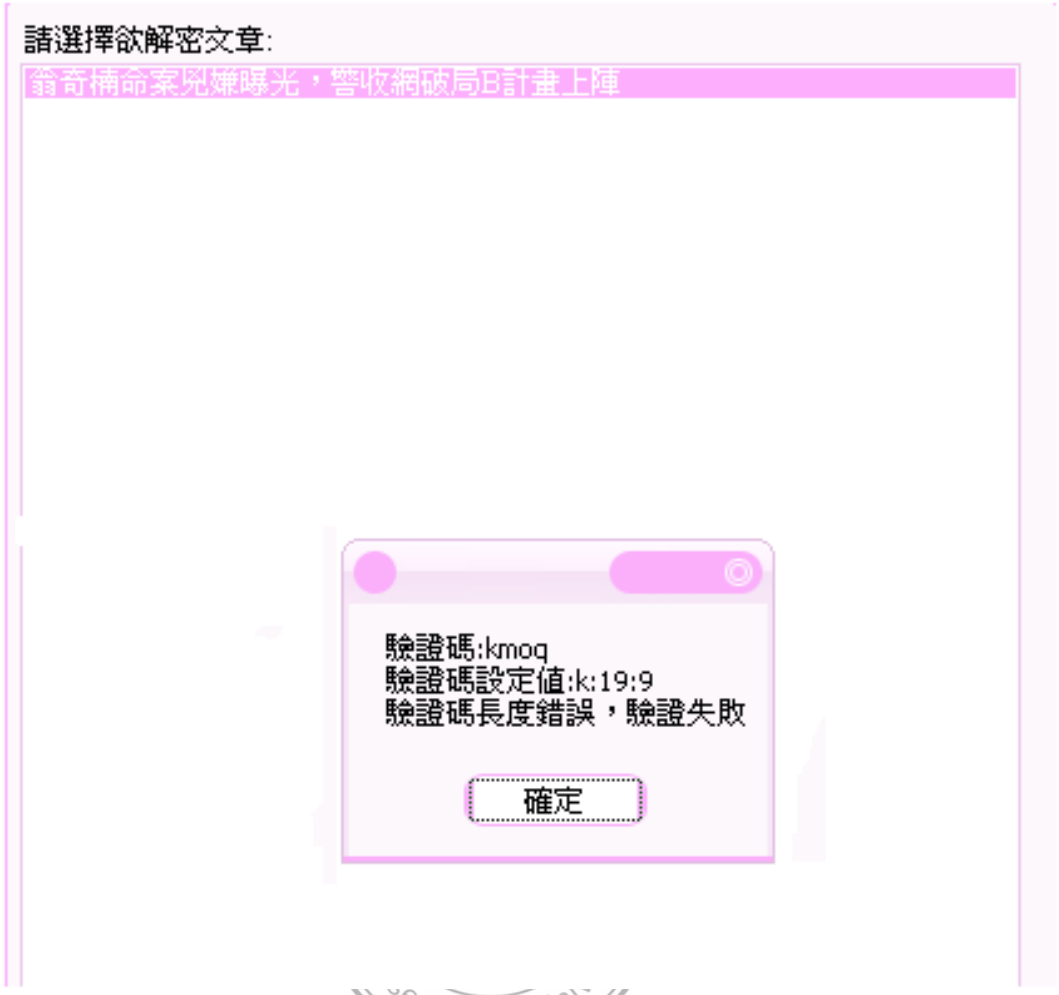

圖 4-22　驗證碼長度錯誤，驗證失敗

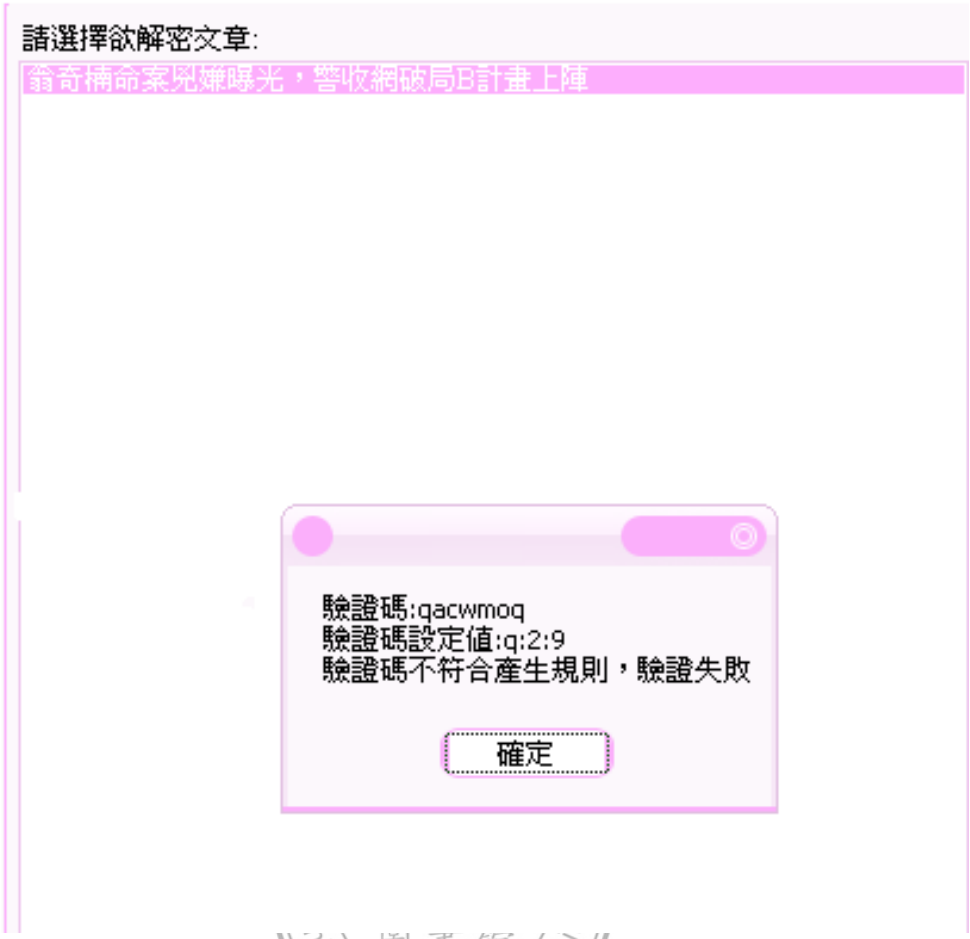

圖 4-23　驗證碼不符合產生規則，驗證失敗

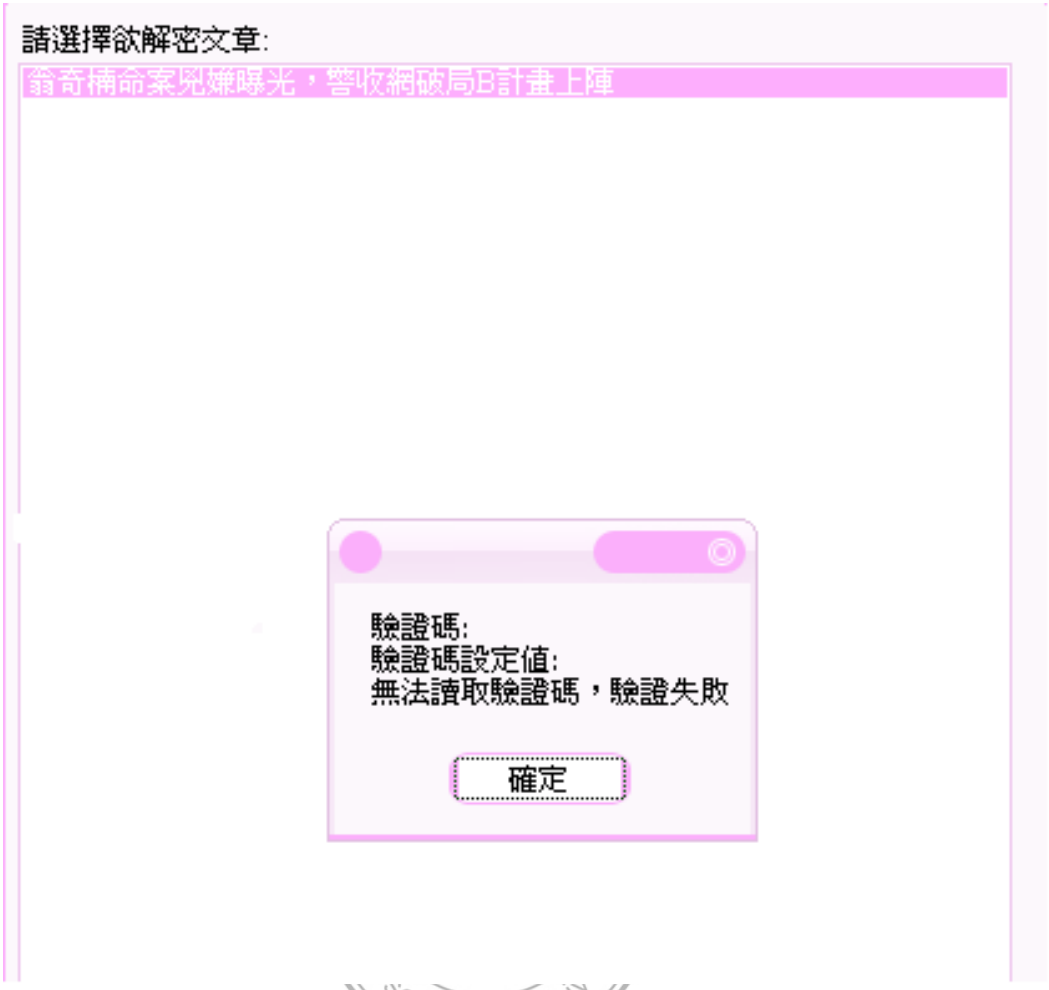

圖 4-24　無法讀取驗證碼，驗證失敗

三、將檔案解密

解密就是把加密步驟顛倒過來，成功解密的話會如圖 4-25。

圖 4-25　解密成功

如果資料遭到竄改如圖 4-26。

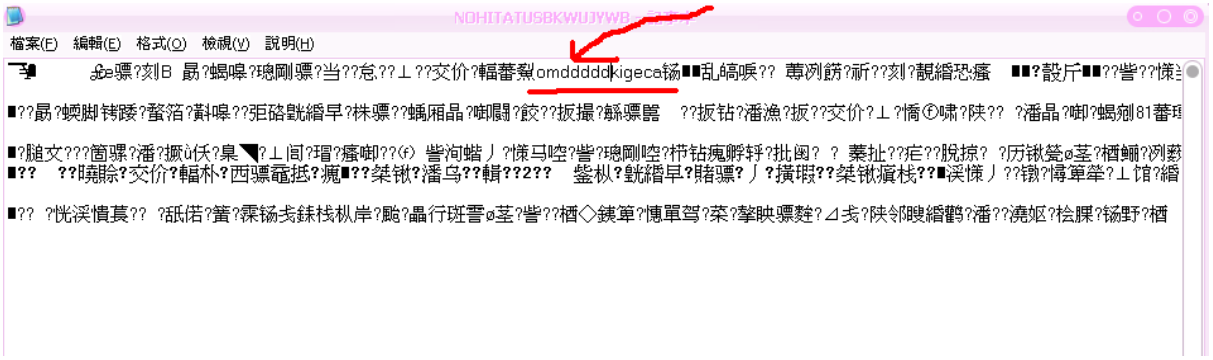

圖 4-26　資料遭到竄改

驗證失敗，解密就不會成功，也就無法發佈，如圖 4-27。

圖 4-27　解密失敗

這時候就需要動用到備份系統，將受損的新聞資料回復，被竄改的資料則遭到廢棄(因為他是亂碼，沒有任何實用價值)。