

智慧型 WAP 網站自動化管理系統之研究

黃鈺珊、曾秋蓉

中華大學 資訊工程研究所
新竹市東香里東香 30 號

黃國禎

國立暨南國際大學 資訊管理系
南投縣 545 埔里鎮大學路 1 號
E-mail: gjhwang@ncnu.edu.tw

摘要

無線通訊技術與行動電話功能的不斷創新與改進，已使得行動電話上網成為未來的趨勢。目前絕大多數的網站均以 HTML 為主要語言，而在提供行動電話上網時則需要使用 WML 語言，同時要提供兩種不同語言的網站對於一般公司或單位將造成兩倍的人力負擔。因此，若能將 HTML 資訊快速地轉換至 WML 的領域中，不但能刺激無線通訊的發展，也能為業者帶來無限的商機。但是由於 HTML 網頁內容過於複雜，若直接轉換成 WML 網頁，過多的資訊將不適合在行動電話上瀏覽。在本研究中，我們提出一套智慧型 WAP 網站自動化管理系統，可自動在系統管理者維護 HTML 網站的同時，自動維護相對應的 WML 網站，因此，可降低人力的負擔。同時，該系統不但可精確的將 HTML 網頁轉換成適當的 WML 網頁，更可依系統管理者的要求，適當的精簡網頁內容而不影響其原意。本研究的成果，不僅可以精簡維護 HTML 及 WML 網站所需的人力，更可提供真正適用於 WAP 行動電話個性的資訊內容，對未來行動電話上網的推廣將有很大的幫助。

關鍵詞：WAP，WML，網站管理，智慧型系統

一、簡介

為了實現"網路無線"的架構，易利信

(Ericsson)、摩托羅拉(Motorola)、諾基亞(Nokia)及Phone.com(Unwired Planet)等通信大廠於1997年12月成立無線應用協定論壇(WAP Forum)，提出了WAP(Wireless Application Protocol) [3]無線應用軟體協定，該應用協定主要是為數位行動電話(GSM)與其他無線終端機裝置(PDA、IA等)，提供無線通訊與資訊服務，並執行Internet網路存取服務的開放標準。WAP所使用的協定類似HTTP的Internet協定，但主要針對無線通訊設備所開發的，因為無線通訊設備頻寬有限、螢幕也較小，因此需要制定專門的協定來支援。只要行動電話能支援WAP，就可以經由無線通訊網路存取網際網路資源，更可以透過行動電話在戶外進行各項交易。WAP所使用的語言是WML(Wireless Markup Language)標記式語言，其地位相當於個人電腦上面所使用的HTML程式語言。

WAP對行動電話或其他無線裝置而言，就像是個人電腦上的IE或Netscape瀏覽器。行動電話的使用者可透過支援WAP協定的行動電話，在Internet上存取所需的資訊。網頁資訊提供者只需了解簡單的WML語法，即可為WAP行動電話用戶提供加值服務，增加產業的商機。

至2000年底為止，台灣擁有行動電話用戶已超過1000萬人，遠超過以個人電腦上網人數，預估在公元2003年使用WAP行動電話上網的人數將超過個人電腦。對一般使用者而言，可以利用行動電話等無線設備獲得即時性、機動性的生活資訊，例如收發E-mail、即

時股價、行動銀行、新聞快報、星座資訊、天氣預報及訂購電影票等服務；對企業而言，由於 WAP 可提供安全、個人化和互動式的通訊媒介，即使遊走在各個地點與各項會議之間，仍可隨時與公司內部的網路連線，掌握瞬息萬變的資訊，因此可作為與顧客和員工聯繫的最佳工具。

WAP 主要是建構在現有的網路架構上，將網際網路上的資訊經由 WAP Gateway 傳送到無線網路伺服器，以提供行動電話用戶連線使用。由於行動電話的螢幕小、不易表達複雜的表格及色彩炫麗的介面，再加上受到記憶體及計算效能的限制，每個 WML 文件只允許 1.4Kbytes 的大小，因此無法處理過大的文件；再者，對於無線上網的消費者來說，太長的文字敘述並不適於透過小螢幕長時間瀏覽，因此若只是將 HTML 語法轉成 WML 語法，未來透過行動電話瀏覽資訊將有很大的困難。

在本研究中，我們提出一套智慧型 WAP 網站自動化管理系統，可自動在系統管理者維護 HTML 網站的同時，自動維護相對應的 WML 網站，因此，可降低人力的負擔。同時，該系統不但可精確的將 HTML 網頁轉換成適當的 WML 網頁，更可依系統管理者的要求，適當的精簡網頁內容而不影響其原意。本研究的成果，不僅可以精簡維護 HTML 及 WML 網站所需的人力，更可提供真正適用於 WAP 行動電話個性的資訊內容，對未來行動電話上網的推廣將有很大的幫助。

二、 相關研究及技術

在 HTML 轉換成 WML 的過程無法用一對一的語法直接替代，因為 HTML 網頁標籤極為複雜，經常不具有對稱性，因此分析上有所困難；而 WML 的語法對標籤的對稱性要求

非常嚴謹，稍有錯誤，瀏覽器就無法顯示網頁內容；因此，在轉換時必須針對 HTML 與 WML 的元件進行分析其相容性，找出相對應的轉換原則。

Spyglass 出了一套自動轉換系統[4]，將每一個 HTML 元件定義好一種轉換模式，例如遇到表格時就直接將標籤轉換成 WML 的表格，因此可以在很短的時間內完成轉換。但其轉換後將會失去 HTML 文件本身的階層結構性，也因為沒有對文件做切割，可能造成轉換後的文件太大，使得 WML 的網頁無法瀏覽。

Oracle 的 Portal-to-go 是屬於一套靜態轉換系統[5]，將文件內容區分為一個個的元件，讓使用者選擇需要轉換的元件，並可加入轉換規則來幫助系統作更佳的轉換。此方式的缺點在於轉換過程中需要人力介入。

國內由政治大學資訊科學系，所開發的 HTML 與 WML 轉換與整合開發系統[1]，即針對 WAP 行動電話的限制及 WML 的能力，設計了一套合適的轉換原則。該系統提供了一個 HTML2WML 轉換器，可利用一般的瀏覽器如 IE、Netscape 來檢視整群 WML 網頁，但目前對 Colspan 和 Rowspan 屬性的表格並未處理，同時在資訊減量方面只刪除了註解及 WML 所不支援的標籤，並沒有精簡文章的內容。

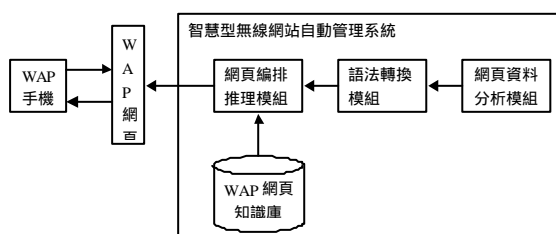
交通大學資訊工程研究所提出的 HTML 文件至 WML 文件之自動轉換系統[2]，運用了 Compiler 技術，先將 HTML 文件分析為一樹狀架構，並根據不同的元件找出最適宜的轉換方式，以保留原來 HTML 的結構階層性，並經過一個轉碼器轉換中文的編碼。但其轉換中並沒有針對 Form 做轉換，使得網頁缺乏與使用者的互動，且並無法精簡資訊。

雖然目前所提出的 HTML 轉換 WML 文件的系統已具備基本的功能，但是受限於行動電話的螢幕及容量大小以及通訊頻寬的問

題，在使用行動電話瀏覽網頁的過程將造成時間及金錢的負擔。因此，在下一節中，我們將介紹一套智慧型 WAP 網站自動化管理系統，該系統可以自動在系統管理者維護 HTML 網站的同時，自動維護相對應的 WML 網站，因此，可降低人力的負擔。同時，該系統不但可精確的將 HTML 網頁轉換成適當的 WML 網頁，更可依系統管理者的要求，適當的精簡網頁內容而不影響其原意。本研究的成果，不僅可以精簡維護 HTML 及 WML 網站所需的人力，更可提供真正適用於 WAP 行動電話個性的資訊內容，對未來行動電話上網的推廣將有很大的幫助。

三、系統架構

行動電話上網主要是透過無線通訊連上 WAP Gateway，再透過有線電纜連上 Internet 來存取網路，其中 WAP Gateway 可看成一個轉繼站，負責接收行動電話使用 WAP 通訊協定傳遞過來的訊息，並透過 HTTP 通訊協定來上網，取得網頁訊息，經由『智慧型 WAP 網站自動化管理系統』(架構如圖一所示)的網頁資料分析、語法轉換以及網頁編排推理模組，來產生內容精確的 WAP 網頁。我們使用 WAP 網頁知識庫來為網頁內容進行排版。最後再將這些網頁訊息回傳給行動電話。以下我們將對圖一的每個元件作詳細的介紹：



圖一：智慧型 WAP 網站自動化管理系統架構

● 網頁資料分析模組

分析網頁內容對建構 WAP 網站是一個很

重要的工作。適當的分析並精簡網頁內容，不僅可以提昇瀏覽效率，更可以增加使用者上網的意願。由於行動電話並不適合瀏覽篇幅過大的文章，因此，我們依據關鍵字及數字對文章內容的影響程度，來判斷每個句子是否刪除或保留，如此一來，精簡後的文章不但資訊量大減，且內容與原本文章的主題尚有一定的關聯性。

● 語法轉換模組

本模組將先行去除 WML 無法支援的標籤，例如含有無法解析的電子檔的 hypelink、Script 語言、IMG 圖檔、字型、文件中的註解等標籤，並將標籤統一先轉成小寫，將 WML 所不能接受的標籤互相交叉或不對稱重新修正，重整文件的架構，並且將 HTML 的標籤轉換成相對應的 WML 標籤。

● 網頁排版推理模組

本模組將網頁內容分析及精簡之結果，利用 WAP 網頁知識庫中關於行動電話及排版的知識，例如螢幕解析度、選項數目、每行字數等，為網頁進行最佳化的內容編排工作，其中 WAP 網頁知識庫是以推理法則 (rules) 來表達排版知識，此方式的最大優點是修改及擴充容易，對於變化很大的無線通訊技術將有良好的適應能力。

四、演算法

本節中，我們將詳細介紹到如何將現有的 HTML 網頁轉換成行動電話可瀏覽的 WML 網頁的方法；以及如何在失去文件的本意下刪簡文件的內容。在轉換的過程中，不但不會影響原本 HTML 網頁的資訊，也不會破壞網頁上的鏈結，因此網站的架構將被完整的保留。

(一)、網頁內容精簡化演算法

在精簡文件內容的過程中，我們將依據以下三個原則來決定文章精簡的程度：

(1). 數字的敏感度 $N_s(i)$

在文件當中，數字通常具有重要的意義，因此當句子中出現數字時，將必須考慮到數字對文件的重要程度有多少，以免刪除重要的資訊，我們將之稱為數字的敏感度。對於數字敏感度我們將之區分為五個等級，詳見下表 4.1。

表 4.1 數字敏感度 $N_s(i)$

Input N_{sf}	$N_s(i)$	說明
絕對敏感	1.0	該文件中的數字具絕對的重要性，不可刪除
非常敏感	0.8	該文件中的數字具一定程度的重要性，盡量避免刪除
敏感	0.6	該文件中的數字具重要性，應避免刪除
些微敏感	0.4	該文件中的數字具些微重要性，可列入刪除與否之參考
不敏感	0.2	該文件中的數字不具重要性，可不列入刪除與否之參考
* 無數字	0	無數字

(2). 關鍵字的重要性 $K_w(i)$

對文件內容而言，有些詞句代表著文件內容的要義，因此當句子中出現這些關鍵字時，必定代表著其對文件內容具有相當程度的重要性。對於關鍵字的重要性我們將之區分為五個等級，詳見表 4.2。

表 4.2 關鍵字重要性 $K_w(i)$

Input K_{wf}	$K_w(i)$	說明
絕對重要	1.0	該文件中的關鍵字具絕對的重要性，不可刪除
非常重要	0.8	該文件中的關鍵字具一定程度的重要性，盡量避免刪除
重要	0.6	該文件中的關鍵字具重要性，應避免刪除

些微重要	0.4	該文件中的關鍵字具些微重要性，可列入刪除與否之參考
不重要	0.2	該文件中的關鍵字不具重要性，可不列入刪除與否之參考
* 無關鍵字	0	無關鍵字

(3). 精簡的重視度 C_p

使用者希望文件在不失去本意之下，能將文章精簡到何種程度，我們將之稱為精簡的重視度。對於精簡的重視度我們將之區分為五個等級，詳見下表 4.3。

表 4.3 精簡重視度 C_p

Input C_{pf}	C_p	說明
絕對重視	1.0	對文件的精簡的要求絕對的重視
非常重視	0.8	對文件的精簡程度非常的重視
重視	0.6	重視文件的精簡程度
些微重視	0.4	些微重視文件的精簡程度
不重視	0.2	不重視文件的精簡程度
* 完全忽略	0	不精簡文件

依據使用者對文件的了解，讓使用者來決定文件中的數字敏感度、關鍵字對文件的重要性及對文件精簡重是的程度，以作為計算出文件中句子是否應該保留的依據。其演算法如下：

$Index[0] = 0$; $Index[i]$ 代表第 i 句的無關指標，亦即與上一個含關鍵字的句子相關的句數

$= (1 - C_p) * B_{cp}$; 代表無關句的最大容忍程度，若某句的無關指標超過 $Index[i]$ ，代表該句已至可刪除的程度；精簡重視指

標 Bcp 代表精簡程度在句子中的影響範圍 (句數)

FOR i=1 TO N DO ; N 表示句子的總數

BEGIN

Index[i] = Index[i-1] - Ns(i) * $a_{Ns} - Kw$
 (i) * a_{Kw+1}

; 考慮數字敏感度 Ns 及 a_{Ns} , 及關鍵字重要性 Kw 及 a_{Kw} , 每往下移動一句, 基本的距離加 1

; 數字影響指標 a_{Ns} 代表數字在句子中的影響範圍 (句數)

; 關鍵字影響指標 a_{Kw} 代表關鍵字在句子中的影響範圍 (句數)

IF Index[i] > THEN; 若 Index[i] 大於 (無關句的容忍句數)

BEGIN

Flag (i) = 1 ; Flag (i) 代表第 i 句是否刪除的指標, 當 Flag (i) 設為 1, 代表第 i 句可刪除

Index[i] = Index[i] - 1 ; 假設該句已刪除, 則下一句與相關句的距離減 1

END

ELSE Flag (i) = 0 ; Flag (i) 設為 0 , 代表保留該句

END

FOR i=1 TO N DO

IF Flag (i) = 1 THEN ; 如果指標為 1 則表示是刪除第 I 句

Remove Sentence (i)

以表 4.4 中的句子為例, 假設目前處理第 1 句, 數字敏感度為敏感 (Ns=0.2)、關鍵字重要性為重要 (Kw=0.8)、精簡重視度為普通 (Cp=0.8) 且 Index[0]=Q $a_{Ns} = 2$ $a_{Kw} = 2$ Bcp=2、關鍵字為 『 In-Law』 或 『 in-law』, 假設目前處理第 1 句, 因為第 1 句中有數字及沒有關鍵字, 因此 Ns(1)=0.2、 Kw(1)=0 , Index[0]=0 , 則 Index[1]=Index[0]-Ns(1)* $a_{Ns} - Kw(1)$ * $a_{Kw+1} = 0.6$, = (1-Cp) *Bcp=0.4, 由於 Index[1] > , 所以第 1 句刪除, 因為第 1 句被刪除, 所以 Index[1] 被修正為 Index[1]-1; 接著處理第 2 句, 因為第 2 句中無數字但有關鍵字, 因此 Ns(2)=0、 Kw(2)=0.8 , 則 Index[2]=Index[1]-Ns(2)* $a_{Ns} - Kw(2)$ * $a_{Kw+1} = -1$, = (1-Cp) *Bcp=0.4, 由於 Index[2] < , 所以保留第二句。結果如表 4.5。

表 4.4 精簡範例

<p>Monday August 13 5:35 PM ET</p> <p>In-Law Conflicts May Have Long-Term Effects</p> <p>By Alan Mozes</p> <p>NEW YORK (Reuters Health) - Apparently, the work of a meddling mother-in-law or father-in-law is never done: Conflicts with in-laws can diminish the perception of marital bliss even 20 years after a couple takes their wedding vows, according to new study findings.</p> <p>“What was the most surprising was that in-laws still had an influence on the couples even though they had been married for such a long</p>
--

period of time," said study lead author Dr. Chalandra M. Bryant. "I thought that was remarkable."

Bryant and her colleagues at Iowa State University in Ames questioned more than 450 rural families over a 4-year period to assess how successful they perceived their marriages to be and what kind of relationship they maintained with their in-laws.

The couples all had adolescent children and had been married an average of almost two decades.

Marital success was defined as a combination of relationship satisfaction, stability and commitment, while the quality of in-law relations was defined as the degree of ongoing conflict, tension and disagreement.

Reporting in the August issue of the Journal of Marriage and Family, Bryant's team notes that for many couples, discord with any of the in-laws is associated with a diminished view of the success of their marriage.

The researchers note that over time, in-laws wield a significant influence over their children's marriage, enduring well beyond the vulnerable early years of marriage.

The investigators also found that wives appeared to have a better relationship with their in-laws if they perceived their marriage to be strong. They pointed out, however, that husbands who thought their marriage was secure did not have relatively improved contact with their spouse's parents.

The study authors suggest that parents might develop different types of gender expectations with regards to marriage--placing more pressure on the husband to maintain a healthy marriage.

Bryant also pointed out that the rural couples in the study might have had considerably closer contact with their in-laws than urban couples would, perhaps affecting the type of relationship that developed.

SOURCE: Journal of Marriage and Family 2001; 63: 614-626.

In-Law Conflicts May Have Long-Term Effects ,
By Alan Mozes , the work of a meddling mother-in-law or father-in-law is never done: Conflicts with in-laws can diminish the perception of marital bliss even 20 years after a couple takes their wedding vows , according to new study findings , "What was the most surprising was that in-laws still had an influence on the couples even though they had been married for such a long period of time , " said study lead author Dr , Bryant and her colleagues at Iowa State University in Ames questioned more than 450 rural families over a 4-year period to assess how successful they perceived their marriages to be and what kind of relationship they maintained with their in-laws , The couples all had adolescent children and had been married an average of almost two decades , discord with any of the in-laws is associated with a diminished view of the success of their marriage , in-laws wield a significant influence over their children's marriage , enduring well beyond the vulnerable early years of marriage , The investigators also found that wives appeared to have a better relationship with their in-laws if they perceived their marriage to be strong , Bryant also pointed out that the rural couples in the study might have had considerably closer contact with their in-laws than urban couples would , perhaps affecting the type of relationship that developed ,

(二)、網頁語法分析與轉換

在網頁精簡化之後，我們將對精簡後的網頁進行語法的轉換，可分為兩大部分：

- **整理 HTML 文件**

整理 HTML 文件包括兩個重點：去除 WML 無法處理的敘述及重整文件架構使標籤對稱。由於 HTML 文件包含文字及 HTML 標籤、各式的 Script 語言以及各種電子檔案(圖形檔、執行檔、文件檔...等)，而對於行動通訊設備而言，因受限於計算能力及記憶體大小，無法解析除了 wml 文件、wml script 及 WBMP Type0 的圖形之外的檔案。所以在轉換前必須先將含有無法解析的電子檔的 hypelink、Script 語言、IMG 圖檔、字型、文件中的註解等標籤先行去除，並將標籤統一轉

表 4.5 精簡結果

成小寫，以便在比對標籤時能正確且快速找到相對應的標籤。另外目前各種 Browser 對 HTML 的容錯性很高，就算是不嚴謹的語法，照樣能解析整個 HTML 文件；而 WML 的格式較嚴謹，不能接受標籤互相交叉或不對稱的情形，所以必須重整文件架構，使其標籤能符合 WML 的格式。在表 4.6 中，我們提出一套 HTML 與 WML 排版的對應關係。

表 4.6 HTML 與 WML 之排版標籤對應表

HTML	WML
P	P
BR	BR
B	B
I	I
U	U
Strong	B
Small	Small
Big	Big
H1~6	Strong , B
Em , Dfn , Cite , Code , Kbd , Blink , Var , Q , Address , blockquote e	em

● 分析 "TABLE"、"FRAME" 與 "FORM"

由於行動電話螢幕太小，在處理 HTML 中的 "TABLE"、"FRAME" 與 "FORM" 指令時，當行數較多時，使用者必須一個字一個字的向下瀏覽，這樣的方式並不方便，而句子

的意義將會變的較難理解。因此，以下我們將提出解決的方法。

在表格的處理方面，我們不採取直接轉換的方式，而是將依行動電話的特性及 WML 所支援的語法以下列四種方式處理：

- (1) 當表格的 column 數小於 2 的時候，不會有瀏覽頁面的問題產生，故直接將 HTML 的表格直接轉換成 WML 表格標籤。
- (2) 因為行動電話螢幕的大小問題，所以在行動電話上的瀏覽方式多以點選為主，所以我們也將 HTML 中的表格以 "select" 標籤來表示。
- (3) 當上述的方法無法解決表格轉換的問題時，則直接將表格中的文字轉為一般條列文字內容。
- (4) 由於 WML 的 TABLE 標籤並沒有支援 colspan 和 rowspan 的屬性；因此，我們將依 rowspan 或 colspan 的 value 和所在位置，加入 rowspan 或 colspan 欄位中的文字，將 TABLE 改成是沒有 rowspan 或 colspan 的 TABLE，再去判斷此 TABLE 適用上述三種方法中的哪一種，如此將不會失去 TABLE 中欄位與欄位間原本的關聯性。

一般 HTML 的網頁，若是含有框架的架構，通常是由兩個以上的網頁文件建構而成。在使用者以 WAP 手機要求網頁的時候，我們會預先分析此文件是否包含框架，若含有框架，則先分析此文件並找出哪一個框架所包含的選項會帶出另外的框架內容，如此便可得知哪個文件是首頁，再依據使用者所選擇的鏈結，下載所要瀏覽的網頁，並轉換成 WML 文件。

通常 HTML 僅有單向的資訊傳遞，缺少使用者之間的互動，為了強化 HTML 與使用者的互動，通常網頁設計者以 FORM 指令輔以 CGI、APPLET、PHP、ASP 等程式語言或網路資料庫語言來達成。由於 WAP 也可以透過 CGI、PHP 等程式語言來與使用者溝通，故

我們提供了相關的轉換。HTML和 WML 都是用相同的觀念來傳送表單的變數，只是 WML 的 INPUT 標籤並沒有像 HTML 有那麼多的型態，所以我們依 WML 所支援的標籤來轉換 HTML。

五、系統實作與測試

我們使用具有跨平台特性的 JAVA 語言來作為開發程式的工具，並以 Nokia WAP Toolkit 模擬器來解讀 WML 文件與顯示 WML 網頁。以下各子節中將介紹一些應用實例。

(一)、具有表格的文件

關於表格部分，我們曾說明依行動電話的特性，及 WML 所支援的語法，將 HTML 中的 table 分成四種方式來表現在行動電話上，分別為：直接轉換、用 WML 中的 select 的方式來表現、將 table 中每一格文字依序輸出、colspan 和 rowspan 等四種情形。以下我們將分別以這四種情形來顯示出轉換的結果：

● WML 表格

因為表格的 column 數小於 2，所以直接將 HTML 的表格轉換成 WML 表格標籤，結果如圖 5.1。

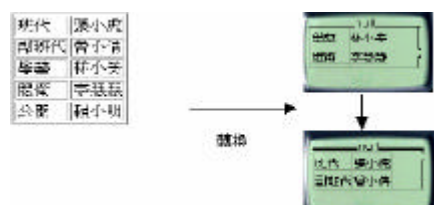


圖 5.1 表格直接轉換結果

● 選項以 "select" 標籤呈現

圖 5.2 的表格中，因為沒有 colspan 和 rowspan 的問題、且其 column 數大於 2，我

們判斷出其第一列中具有標題，因此我們以 Select 的方式來呈現這個表格，其轉換結果為圖 5.3。

師資介紹			
姓名	職稱	最高學歷	研究領域
蔡英仁	教授	加拿大新佈倫斯威克大學	天然物全合成 催化不對稱反應
		化學博士	
黃文樞	教授	美國奧勒岡州立大學	無機電化學、有機金屬合成
		化學博士	
夏友平	教授	美國伊利諾理工學院	環境化學、理論計算
		化學博士	
戴建夫	教授	美國匹茲堡大學	生物有機化學、生物轉換
		化學博士	

圖 5.2 具有標題的 HTML 表格

我們將表格第一行的資料顯示成一個個的鏈結，如圖 5.3 中的圖 (A)，以取得該列的資料。以蔡英仁為例：當我們選擇蔡英仁時，將會鏈結到圖 (B) 顯示出職稱、最高學歷、研究領域，以供選取想了解的資料，如圖 (C) (D) 所示。

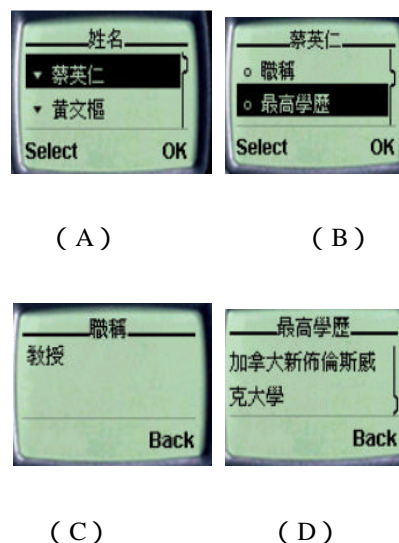


圖 5.3 轉換後的結果

● 條列文字

圖 5.4 的表格中，因為沒有 colspan 和 rowspan 的問題、且其 column 數大於 2，我們發現其表格中的欄位都是鏈結，因此我們將表格中的每一格文字依序輸出，其轉換結果為圖 5.5。

中文歌曲 英文歌曲 日文歌曲 動畫&電玩				
傷心	不老的傳說	檸檬樹	姐妹	對面的女孩看過來
擁擠	分享	起飛到淚水	高遠真	我和你來吹風
海馬龍	倩女幽魂	健康歌	李香蘭	一千個傷心的理由
渴望	狂野之城	紅豆	龍小兒	

圖 5.4 沒有結構的 HTML 表格



圖 5.5 轉換後的結果

● Colspan 和 Rowspan 的問題

rowspan 或 colspan 的 value 代表其所橫跨的欄位數，假設 value=2，我們就將其切割為兩個欄位，並填入 rowspan 或 colspan 欄位中的文字，將 table 修正為沒有 rowspan 或 colspan 的表格，再去判斷此表格適合用何種轉換的方法來呈現。

● Colspan

如圖 5.6，Colspan=2 因此我們將此欄位分割為兩格，並填入【常用研究相關網站】，修正後的表格就不再有 Colspan 的屬性存在。結果如圖 5.7 所示。

常用研究相關網站	
<ul style="list-style-type: none"> ● 臺大圖書館首頁 ● 台大館藏Tulpe ● EBSCO資料庫 ● 全國科技網路 	<ul style="list-style-type: none"> ● FirstSearch ● COMEXDEX工程資料庫 ● 全國圖書資訊網 ● 有限元素法

圖 5.6 具有 Colspan 的 HTML 表格

常用研究相關網站	常用研究相關網站
<ul style="list-style-type: none"> ● 臺大圖書館首頁 ● 台大館藏Tulpe ● EBSCO資料庫 ● 全國科技網路 	<ul style="list-style-type: none"> ● FirstSearch ● COMEXDEX工程資料庫 ● 全國圖書資訊網 ● 有限元素法

圖 5.7 修正後的 HTML 表格

● Rowspan

如圖 5.8，第一個 Rowspan=3，因此我們將此欄位分割為三格，並填入【環工】；第二個 Rowspan=2，我們將此欄位分割為兩格並填入【運工】，使修正後的表格就不具有 Rowspan 的屬性存在。結果如圖 5.9 所示。

組別	姓名	職稱	分數	學歷	電子郵件信箱
環工	歐陽鴻聲	教授	4654	國家工學博士	cfouyang@ncem.ncu.edu.tw
	曾迪華	教授	4655	美國普渡大學博士	dhseag@ncem.ncu.edu.tw
	廖遠良	教授	4662	美國密蘇里羅拉校區博士	slhaw@ncem.ncu.edu.tw
運工	李俊福	教授	4658	美國密西根州立大學博士	fllee@ncem.ncu.edu.tw
	張木彬	教授	4663	美國伊利諾大學香檳校區博士	mbchang@ncem.ncu.edu.tw

圖 5.8 具有 Rowspan 的 HTML 表格

組別	姓名	職稱	分數	學歷	電子郵件信箱
環工	歐陽鴻聲	教授	4654	國家工學博士	cfouyang@ncem.ncu.edu.tw
環工	曾迪華	教授	4655	美國普渡大學博士	dhseag@ncem.ncu.edu.tw
環工	廖遠良	教授	4662	美國密蘇里羅拉校區博士	slhaw@ncem.ncu.edu.tw
運工	李俊福	教授	4658	美國密西根州立大學博士	fllee@ncem.ncu.edu.tw
運工	張木彬	教授	4663	美國伊利諾大學香檳校區博士	mbchang@ncem.ncu.edu.tw

圖 5.9 修正後的 HTML 表格

(二)、具有 Frame 的文件

以圖 5.10 為例，我們分析此文件發現兩個框架彼此間的關係為：左邊框架選項會帶出右邊框架的內容，因此首頁應該為左邊框架內的網頁。



圖 5.10 具有 Frame 的 HTML 網頁圖與轉換後的結果

例如選擇【金庸小說】，則會秀出其所對應的網頁內容，結果如圖 5.1 所示。

(三)、 Form 的轉換

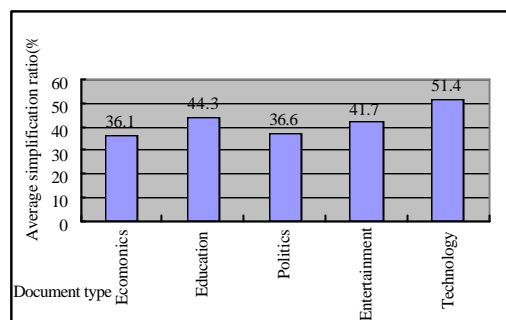
我們依據表單中的輸入型態轉換成 WML 所支援的輸入型態。如圖 5.12，姓名欄的輸入型態為 TEXT，則相對應 WML 的輸入型態為 INPUT，其轉換後的結果如圖 5.13 所示。



圖 5.12 具有 Form 的 HTML 文件

六、 系統的整體效率

為了去評估整個智慧型 WAP 網站自動化管理系統的效率，我們針對政治、教育、經濟、娛樂以及科技等五種不同型態的文件作了一個精簡化的實驗。每種型態的文件各取 20 篇，其字數介於 600 至 5000 左右。在精簡化的過程中，我們對每一篇文章作重複的精簡，依照 Kw 的值從 1.0 至 0 遞減，直到文件失去原本的意義，並紀錄文章精簡過程中保持原意義的最佳精簡比例。在處理完所有的文件後，獲得文件的平均最佳精簡化比例如圖二所示。整體而言，文件平均精簡的比例都超過 35%；這樣的數據可約略看出智慧型 WAP 網站自動化管理系統將可減少超過 35% 的文件內容，並且不影響文件的原始要義以及文件的可讀性。



圖二. 五種型態的文件之平均精簡比例

七、 結論

在本論文中，我們提出一套智慧型 WAP 網站自動化管理系統，這套系統具備了以下的優點：可自動在系統管理者維護 HTML 網站的同時，自動維護相對應的 WML 網站，因此，可降低人力的負擔。同時，該系統不但可精確的將 HTML 網頁轉換成適當的 WML 網頁，更可依系統管理者的要求，適當的精簡網頁內容而不影響其原意。本研究的成果，不僅可以精簡維護 HTML 及 WML 網站所需的人力，更可提供真正適用於 WAP 行動電話個性的資訊內容，對未來行動電話上網的推廣將有很大的幫助。

文字內容抽象化及網頁個人化將是我們未來研究的目標。文字內容抽象化就是將文章內容以精要的一句話或幾行文字表示，讓使用者了解文章大致的內容為何；在網頁個人化方面，我們希望系統能自動的依據使用者的瀏覽習慣，分析出其瀏覽模式，接著系統會自動為使用者設計出屬於自己的個人化首頁，將使用者不常使用的鏈結隱藏，以減少使用者搜尋的時間。

參考文獻

- [1] 連耀南,陳宗儀,何彥德,吳曉峰,蔡子傑,張

宏慶.HTML 與 WML 之間的轉換與整合
開發系統,中華民國九十年第七屆行動計
算研討會論文集 , D67-D71 頁

[2] 徐元瑛曾建超,HTML 文件至 WML 文件
之自動轉換系統 , 中華民國九十年第七
屆行動計算研討會論文,D59-D66 頁

[3] WAP Forum ; <http://www.wapforum.org/>

[4] Spyglass Inside Mobile Data – White
Paper ; <http://www.spyglass.com/>

[5] Oracle Portal-to-Go , Oracle Business
White Paper ; <http://www.oracle.com>