健康資訊網頁之可參考度適性化評估

Adaptive Referencability Estimation for Healthcare Web Pages

李芊葳
Chian-Wei Li
慈濟大學醫學資訊學系（所）
95325111@stmail.tcu.edu.tw

劉瑞瓏
Rey-Long Liu
慈濟大學醫學資訊學系(所）
rlliutcu@mail.tcu.edu.tw

摘要

網際網路已成為民眾獲取健康資訊的重要來源之一。然而在醫療領域中，資訊之可參考度特別重要，包括資訊之完整性、可靠度、合適之深淺度、語言、及領域等。據此，本文提出一個資訊技術 MyRR，自動依據使用者指定之高參考度網頁，將搜尋引擎回傳結果進行可參考度之重新排序，以方便使用者閱讀符合自己需求且具高參考價值之資訊。實驗證實，MyRR 可大幅提升醫學資訊網頁之可參考度辨識效能。此外，MyRR 之計算有別以往，不需蒐集整個全球資訊網和費時之迭代運算，在日新月異之潮流中，更能即時因應使用者之需求。

關鍵字：健康資訊、資訊可參考度、適性化可參考度評估

Abstract

The Internet has been an important healthcare information source for general users and healthcare professionals. Information referencability is particularly essential in healthcare. It considers integrity, reliability, skill levels, languages, and domains of healthcare information. Therefore, we propose a technology MyRR, which intelligently re-ranks healthcare web pages based on their referencability, which is adaptively estimated based on a set of web pages that deserves reference by individual users. Users may thus read good information to fit their needs. Experimental results show that MyRR can significantly improve the performance of referencability estimation. Moreover, MyRR achieves the estimation without conducting a time-consuming crawl on the entire WWW.

Keywords: Healthcare information, Information referencability, Adaptive referencability estimation.

1、前言

在現今以病人為主之醫療照護思維中，病患及民眾已漸主動積極地閱讀及研究切身相關之保健疾病資訊，而網際網路成為重要之資訊來源[3]。然而，在網路資訊充斥及品質良莠不齊的情形下，資訊之可參考度相當重要。

參酌 Risk、Dzenowagis[1]所回顧之網際網路健康資訊品質研究，以及行政院衛生署優良健康網站之評鑑指標，網頁之可參考度實為網頁品質之重要維度。一般而言，可參考度之主要考量維度包括（1）網頁需符合使用者所關注之特定領域，（2）網頁內容需載明文章著作出處或撰文資訊，且（3）具完整豐富及可讀之網頁內容。低參考度之健康資訊，除了很難為使用者所用，甚至可能造成不良誤導。

1.1、問題定義

本研究開發一個資訊技術，讓每一位使用者自行指定自己信賴且具參考價值之網頁集，系統自動為使用者計算此網頁集鏈結之資訊空間中各網頁之可參考度。當使用者於一般檢索系統輸入查詢語句時，此預先計算好之可參考度即可用來將回傳之網頁進行重新排序，讓使用者更可輕易地發現符合自己需求且具較高參考度之醫療健康資訊。

1.2、研究動機

網際網路醫療健康資訊的便利取得，直接影響使用者對醫療的認知與生命的安全[3][5]，故衍生出眾多資訊品質評估之方法[8]。然而，目前一般搜尋引擎回傳的結果沒有針對健康領域資訊之可參考度作衡量，造成一些不甚值得參考之資訊夾雜其中，除了埋沒了真正

有用的資訊，更有可能誤導使用者。在相關研究方面，多僅以網頁間互相鏈結之情形為依據，其中較常見之作法為 PageRank[6]及 HITS[4]。

然而，PageRank 依網頁間鏈結情形來計算各網頁被使用者閱讀之機率。一個網頁之 PageRank 分數越高即代表其被使用者閱讀之機率較高，故亦隱喻其越值得閱讀。HITS 則同時計算網頁之指示度（Hub）及權威度（Authority），一個網頁之指示度值高表示網頁向外鏈結至許多權威高之網頁，權威度值高表示網頁被許多網頁推薦瀏覽。PageRank 及 HITS 兩者均僅以一般網頁間之鏈結為考量，網頁重要性乃依全球資訊網中鏈結情形來決定，故屬一般性層面之網頁評估，無法針對特定領域將網頁可參考程度做較細部之分析，亦無法為每個使用者量身訂製其偏好之網頁可參考度。

以 Google[1]為例，於 2006 年 9 月 6 日鍵入「發燒」為醫療健康查詢語句時，前 10 筆回傳結果中有 2 筆即非屬醫療健康領域[2]，其中亦有網頁語言不適合國人閱讀者，但其 PageRank 值卻相當高。

由PageRank鏈結分析衍生出之網頁個人化排序研究，如Haveliwala[7]根據使用者於搜尋引擎下達查詢語句之主題排序文章、和Jeh、Widom[2]由給定之單一偏好網站進行WWW之個人化評估，皆是以提升檢索結果準確度，從個人化角度為出發點。但需爬取整個全球資訊網後，進行迭代運算，待數值趨於平穩後，才能得知網頁之重要性分數。對於偏好日新月異的使用者而言，整個歷程費時且效率差，較無法因應使用者之需求。故我們需要更多之研究以快速及適性化地為每位使用者提供具參考價值之網頁資訊。

1.3、研究方法與貢獻

本研究提出個人化可參考度評估技術 MyRR（My Referencability Ranking），依據個別使用者加入之主觀認定具高參考度之網頁集來進行鏈結分析。經網頁之可參考度評估後，即可於線上為使用者調整搜尋引擎回傳結果，將可參考度較高之網頁排在前面，以利使用者輕易掌握有意義之資訊。

本研究並實際驗證 MyRR 之成效，利用 MyRR 計算之網頁可參考度來重新排序 Google 網頁回傳之結果。研究結果發現，此重新排序結果比 Google 更接近專家評估之可參考度順序，故 MyRR 可讓使用者輕易掌握高參考度之資訊。此貢獻讓高品質健康資訊更能貼近一般使用者，進而促進健康。

2、網頁可參考度之適性化評估

MyRR 包含離線作業與線上作業，主要流程如 Figure 1 所示。離線作業之目的是事先進行網頁可參考度適性化評估，並儲存於資料庫中；而線上作業之目的是當使用者至搜尋引擎輸入欲查詢之語句後，MyRR 即可將回傳頁面之網頁依可參考度來重新排序。以下將詳細說明 MyRR 之操作流程。

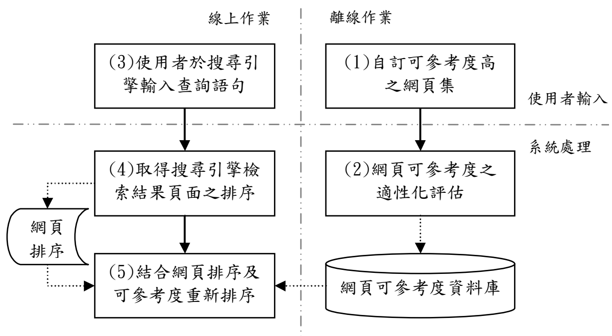

Figure 1 MyRR 之操作流程圖

2.1、離線作業

在離線作業時，MyRR 接受使用者自訂可參考度高之網頁集（Anchor）作為爬取網頁之起始點。爬取後之龐大網頁集經過鏈結分析可得每個網頁之可參考度並存入網頁可參考度資料庫中。

網頁可參考度之鏈結分析流程如 Table 1 所示。MyRR 是以深度優先（Depth first）的方式，自給定的 Anchor（Step 1）開始抓取網頁（Step 2.1）。針對每一個網頁，MyRR 記錄三項訊息：（1）該網頁之網址（p），（2）該網頁可傳遞予其所鏈結之網頁之可參考度（r），及（3）該網頁是否正被展開取其向外鏈結中（e，True 代表正被展開，False 代表非正被展開）。當一個網頁

1 Google 網址 http://www.google.com.tw/。
2 排序第八名為聯合新聞網之一日「發燒」新聞、國內要聞專欄（PageRank＝4）；排序第九名為非常發燒網屬簡體字網站，以販售各式音樂器材等、並設有「發燒」論壇（PageRank＝5）；（PageRank 滿分為 10）

Table 1 MyRR algorithm

```
Algorithm MyRR (anchor, R)
Input: (1) anchor is the URL of a starting web page trusted and designated by the
       user, and (2) R is a set {(q, rank_q)}, where q is a web page, rank_q is the
       AdaptiveRank value of q;
Effect: Starting from anchor, (1) add (n, rank_n) into R for each new web page n
       found, and (2) update (q, rank_q) for each existing web page q;
Begin
(1) PushStack(S, (anchor, 1.0, False));
(2) While S is not empty do
   (2.1) (p, r, e) ← PopTop(S);
   (2.2) If (e is False) then      // not expanded ➔ Try to expand it
      (2.2.1) PushStack(S, (p, r, True));    // cached for avoid cyclic expansion
      (2.2.2) RankInc = β×r;
      (2.2.3) If (RankInc ≥ δ) then
         (2.2.3.1) For each hyperlink q in p, do
         (2.2.3.1.1) If (q, r, True) is not in S (∀ r), then    // avoid cyclic expansion
            (2.2.3.1.1.1) If (q, rank_q) ∈ R (for some rank_q),
               (2.2.3.1.1.1.1) rank_q ← rank_q + RankInc;
            (2.2.3.1.1.2) Else R ← R ∪ {(q, RankInc)};
            (2.2.3.1.1.3) PushStack(S, (q, RankInc, False));
End.
```

為非正被展開中且其可參考度經傳遞折扣（β，Step 2.2.2）大於等於一門檻值（Step 2.2.3）時，爬取該網頁之向外鏈結。如果此向外鏈結之網頁非正處理中之網頁始能列入考慮（Step 2.2.3.1.1）。此作法是為避免網頁間互相迴圈式鏈結，而影響可參考度之正確評估。最後，更新並在資料庫中記錄每個網頁之可參考度（Step 2.2.3.1.1.1 及 2.2.3.1.1.2）。

網頁可參考度是以傳遞的方式來衡量：當一個網頁 p 內有鏈結至網頁 q 時，若 p 之可參考度高，則 q 可獲得較高之參考度加分，因其隱含 p 推薦 q 之意，由此觀念進行網頁之可參考度傳遞，越直接由 p 鏈結之網頁其參考度越受影響；反之，越是間接由 p 鏈結之網頁，其可參考度越不受 p 影響，故以β作為傳遞網頁參考度之折扣（Step 2.2.2），以δ為傳遞網頁參考度之門檻（Step 2.2.3），衡量網頁之參考度。

以 Figure 2 為例，p 為起始網頁(anchor page)，向外鏈結至網頁 q，網頁 g 為 q 之向外鏈結網頁。假設網頁 p 獲得可參考度加分（r）為 1.0，則經折扣 0.5（β）

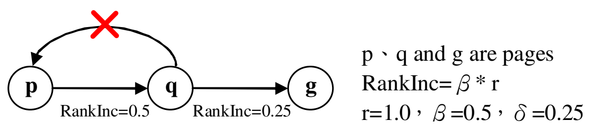

Figure 2 網頁傳導可參考度之示意圖

後，傳遞 0.5（=1.0*0.5）給 q（RankInc），以增加其可參考度值。q 再傳遞 0.25（=0.5*0.5）給 g，但不傳遞給 p，因 p 屬正展開中之網頁故不列入考慮。最後，因 g 之傳遞值 0.125（=0.25*0.5）小於傳遞網頁參考度之門檻（δ）0.25，便不再向外傳遞。故 p、q、g 經傳遞後，分別獲得可參考度值加分為 1.0、0.5、0.25。

2.2、線上作業

在線上作業時，使用者可至搜尋引擎（如 Google）輸入查詢語句，MyRR 針對回傳之頁面，將各網頁依可參考度由大到小重新排序。一個網頁 p 可參考度之計算方式如下：

（1）若 p 已存在網頁可參考度資料庫中，則直接取用可參考度。

（2）若 p 尚未儲存於資料庫，但資料庫中存在與 p 相同網域之網頁群，則加總平均該網頁群之可參考度作為 p 之可參考度。

（3）若 p 和其同網域之網頁群皆不存在資料庫時，網頁之可參考度為 0。

（4）當多個網頁之參考度相等時，排序時以 Google 之排序順位為依據。

由於網頁爬取之範圍未完全涵蓋全球資訊網，基於同一網站內之網頁水準相似的概念，採取模糊比對網址（上述方式 2)讓相同網域下之網頁 p 繼承網域之平均可參考度，降低網頁爬取不完全之影響。當網頁 p 與同網域網頁之可參考度皆未納入資料庫時（方式 3），表示 p 可能較不具參考性，故給予 0 分。最後，同參考度得分之網頁則以原 Google 之排序順位來排列（方式 4）。

3、系統實作與驗證

為實際驗證 MyRR 之成效，我們實作 MyRR 並進行一個實務環境下之實驗。經實驗證實 MyRR 不僅較 Google 更貼近真實之網頁可參考度，亦可適性化地調整網頁之重要性。整體評估流程如 Figure 3 所示。本節詳細介紹實驗之過程並討論其結果。

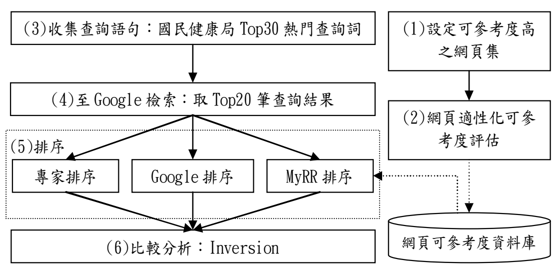

Figure 3 評估流程圖

3.1、實驗環境

高參考度網頁集（Anchor）之設定方法，是將民國九十一年至九十五年行政院衛生署優良健康網站[3]中六大類型之獲獎網站共 131 個網站，濾除重覆網站 39 個後，取 92 個作為實驗之高可參考度網頁集（Figure 3 之 Step 1）。自此網頁集開始以深度優先方式爬取並進行網頁可參考度評估（Step 2），參數設定為β=0.5、δ=0.1。MyRR 歷經約 40 小時之離線作業總得 63107 個網頁，其參考度分數存入網頁可參考度資料庫，作為線上檢索時重新排序搜尋引擎回傳網頁之依據。

在線上查詢語句方面，我們根據國民健康局網站之熱門查詢關鍵字排行榜[4]，取前 30 名熱門查詢詞[5]作為醫療健康領域之查詢語句（Step 3）。至搜尋引擎 Google 檢索「所有網頁」[6]（Step 4），並記錄每個查詢語句之前 20 筆搜尋結果，共計 600（=30*20）個網頁作為實驗資料。

3.2、驗證方式

每個查詢語句之回傳結果為一組驗證資料，每組驗證資料均有三種排序方法（Step 5）：（1）依據專家給定之參考度分數進行排序（是為實驗之專家排序）；（2)Google 原先回傳頁面之排序(是為 Google 排序)；（3）依 MyRR 之可參考度評估來排序（是為 MyRR 排序）。

我們以專家排序為標準來衡量 Google 排序及 MyRR 排序之優劣。故為公正獲得專家排序，驗證準則參考行政院衛生署優良健康資訊網站之評鑑指標[7]，針對網頁可參考度之考量維度為主軸設計問卷，問卷內容共計 10 小題，每題答為是得一分，答為否得零分，滿分十分，問卷內容如 Figure 4 所示。

此問卷即為網頁可參考度之主要評估維度，由 3 名醫學專業人士針對實驗資料集之網頁進行可參考度評估。每個網頁之得分加總平均後，即為該網頁之可參考度程度，分數越高表示資訊較有參考價值，越值得閱讀。

1	是否與查詢語句相關之醫療資訊
（一）醫療健康網頁相關資訊之可參考度	
2	文章是否由醫療專業人士所撰寫
3	文章是否註明作者姓名、服務單位
4	內容是否註明聯繫方式 (如電話或電子郵件)
5	文章是否註明發表日期 (或最新修改日期)
6	是否註明網站維護者資料 (如管理者姓名、電話、E-mail)
（二）網頁內容整體呈現之完整性及豐富度	
7	資訊是否完整
8	是否有依專業程度分別陳述
9	資訊之使用語言是否適合本國人閱讀
10	有無引述明確之參考文獻 (包含作者、發表時間、文獻名稱)

Figure 4 網頁可參考度之問卷內容

故以專家排序為標準，我們依據 Google 排序及 MyRR 排序之錯置情形（Inversion）來衡量優劣（Figure 3 之 Step 6）。Inversion 之定義如下：

$$I_R=\sum_{r=1}^{n} I_r \;;\quad I_r=\sum_{k=1}^{r-1} i_{k,r}$$

其中 $i_{k,r}\begin{cases}1, \text{if expert ranks the } k^{th} \text{ site after the } i^{th} \text{ site}\\ 0, \text{otherwise}\end{cases}$

R：某一查詢語句之回傳結果排名集；

I_R：排名集 R 相對於專家排序之 Iinversion 值；

n：回傳結果之筆數；

I_r：計算排序第 r 位所產生之 Inversion 值；

$i_{k,r}$：當第 k 位之排序值大於第 r 位時，$i_{k,r}$ 為 1，否則為 0。

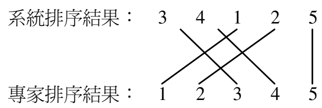

Figure 5 Inversion 之計算範例

以 Figure 5 為例，假設由左至右共 5 筆網址（n=5），專家排序網址之正確順序為 12345，而系統將其排為 34125（R），此排序所產生之 Inversion 分別為 0、0、2、2、及 0，故 Inversion 之總和為 4(IR=0+0+2+2+0)。Inversion 之值越大表示與正確排序順位差距越大，較不能提供合宜之資訊排名；反之，Inversion 之值越小則越接近正確排序之順位，更能將適當之資訊優先列予使用者，有助獲得具參考價值之資訊。故實驗結果研究以 Inversion 為主要之驗證準則。分別比較 MyRR

3 行政院衛生署優良健康網站 http://awards.doh.gov.tw/。
4 行政院衛生署國民健康局之熱門查詢詞 http://www.bhp.doh.gov.tw/BHP/index.jsp
5 前 30 名辭彙：癌症、罕見疾病、發燒、牙齒、保健、糖尿病、十大死因、高血壓肥胖、氣喘、新生兒篩檢、中老年保健、乳癌、BMI、癌症地圖、徵才、戒菸、檳榔、視力保健、大腸癌、子宮頸癌、社區健康營造、憂鬱症、兒童健康手冊、健康、視力、健康檢查、癌症登記、肺結核、戒煙。（95/09/06）
6 檢索日期（95/09/06）
7 行政院衛生署優良健康網站之評鑑指標 http://awards.doh.gov.tw/award_2.asp（95/8/22）

排序、Google 排序何者最接近專家之排序、所提供之網頁排序最能反映真實之資訊可參考度。

3.3、結果分析與討論

MyRR 和 Google 在 30 組查詢語句資料集中之表現如 Figure 6 所示。橫軸為各組查詢語句[8]，縱軸表在同一查詢語句下，MyRR 和專家排序間之 Inversion 及 Google 和專家排序間之 Inversion。

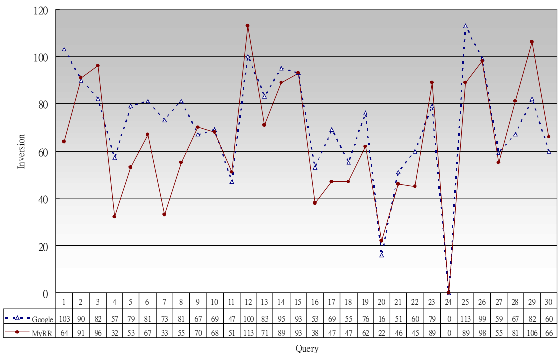

Figure 6 總體資料集 MyRR 與 Google 之 Inversion

整體來說，30 組查詢語句中，MyRR 有 18 組表現優於 Google，2 組持平，10 組略差於 Google。

此外，由於對使用者而言，檢索系統回傳之結果前幾筆越能提供合適訊息，故我們進一步觀察專家評判前 5 名之 176 筆資料群[9]（Figure 7），MyRR 在 30 組查詢語句中有 13 組表現優於 Google，8 組持平，9 組不如於 Google。

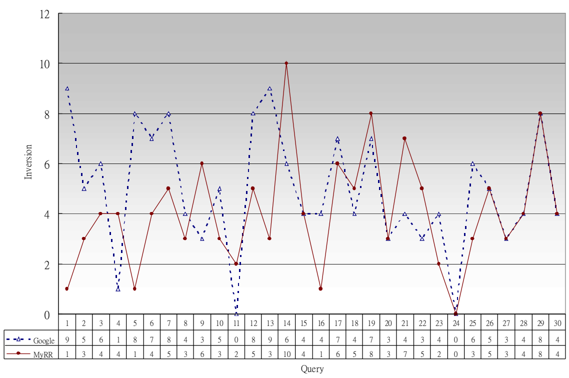

Figure 7 Top5 資料集 MyRR 與 Google 之 Inversion

8 對應編號 1 至 30 之查詢語句為：BMI、十大死因、大腸癌、子宮頸癌、中老年保健、牙齒、戒菸、戒煙、罕見疾病、乳癌、兒童健康手冊、社區健康營造、肺結核、肥胖、保健、氣喘、高血壓、健康、健康檢查、發燒、視力、視力保健、新生兒篩檢、微才、憂鬱症、糖尿病、癌症、癌症地圖、癌症登記、檳榔。

9專家評為同名者亦納入考量，如查詢語句 21「視力」中有 2 筆皆為第五名，則「視力」組之將有 6 筆資料。

如 Table 2 所示，總體平均而言，平均之 Inversion：

Table 2 MyRR 與 Google 之 Inversion 結果比較

	總體資料集 (600 筆)	Top5 資料集 (176 筆)
	平均數 Inversion	平均數 Inversion
MyRR 排序	64.57	4.07
Google 排序	71.30	4.97
MyRR 之改進幅度	9%	18%

MyRR 為 64.57(1937/30)，Google 為 71.30(2139/30)，MyRR 改進 Google 約 9%之效能。前 5 名資料之平均 Inversion，MyRR 為 4.07（122/30），Google 為 4.97（149/30），MyRR 改進 Google 之幅度達 18%。由於被專家評判為前 5 名之網頁通常是較明確具高參考度之網站，故此前 5 名資料群之大幅改進，意味著使用者更不需將 20 筆資料完全瀏覽，即可於前 5 筆中發現參考價值高且合其偏好之資訊，縮短尋找可用資訊之時間，此貢獻深具實務應用之意義。

我們進一步以查詢語句 BMI（Body Mass Index，身體質量指數）為例，比較 Google、MyRR、專家（Experts）之排序情形（Figure 8）。橫軸為 20 筆回傳網頁，縱軸為各排序法之順位。此例中，MyRR 與專家之 Inversion 為 64，Google 與專家之 Inversion 為 103。

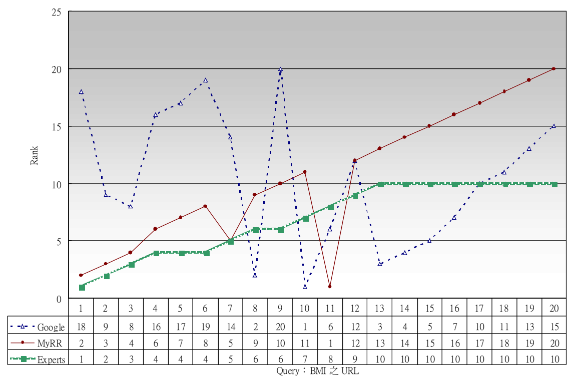

Figure 8 查詢語句：BMI 之排序

如 Figure 8 所示，MyRR 之排序曲線較 Google 貼近專家排序之曲線。值得注意的是，專家評定之前五筆具參考價值之健康資訊網頁，Google 卻將之排序於第 8 名之後，甚至超過 15 名，此況除了很難替使用者呈現適切之網頁排序，造成閱讀上之負擔外，更可能提供不恰當之網頁予使用者參考。反觀 MyRR，藉由可參考度高網頁集之設定，經網頁之鏈結分析適性地給予網頁可參考度，將參考價值高之網頁優先列於前半段，對健康資訊網頁呈現的排序確具實質之貢獻，大幅改善 Google 之排序，有助使用者快速獲得參考價值

高且符合使用者閱讀特性之健康資訊。

3.4、未來展望

目前網頁之爬取數量未完全涵蓋整個全球資訊網，僅以給定之高參考度網頁集、參數設定等，作為網頁進行爬取與可參考度評估之依據。由於網頁集和參數之設定會影響網頁爬取之數量，故部份網頁之可參考度尚無值，視為 0，而造成 MyRR 於部份查詢語句表現欠佳。此外，因僅透過網頁鏈結分析網頁之可參考度，難免會爬取到不該納入考量之網頁（如廣告），導致某些網頁之可參考度高於部分真正較具參考價值之網頁，產生非預期之效益，干擾可參考度之排序。未來將針對以上兩大問題做進一步之改良。

較可行之改良方法包括：（1）增加具可參考度網頁集之網頁數量，或調整參數之設定，以計算更深且廣之網頁可參考度。（2）改良網址之模糊比對，以便至資料庫取得相近網頁之可參考度，減少網頁無可參考度之機率。（3）杜絕廣告鏈結之干擾方面，可從網頁黑名單之設置、網頁 Html 標籤（tag）之識別、網頁內容關鍵詞之比對等著手。

4、結論

網際網路健康資訊直接影響使用者對於醫療的認知及決策，其可讀性、品質、及完整性均相當重要。為避免不甚可參考之資訊夾雜其中誤導使用者，本文提出一個資訊技術 MyRR，透過健康資訊網頁間之交互鏈結推薦關係，衡量網頁之可參考度。使用者可自行指定符合自身需求之高參考度網頁集（合適之語言、領域、品質完整、難易等），MyRR 即可以此為依據，評估網頁之可參考度。

實驗證實，在給定之歷年得獎之優良健康網站作為高參考度網頁集下，MyRR 可成功地進行健康資訊領域可參考度之評估，並重新排序檢索系統之回傳結果。MyRR 平均可改善 Google 9%~18%之效益，將可高參考度之網頁優先排序於前，不僅讓使用者減少瀏覽過濾所需資訊之時間，優先提供可參考度高之網頁，亦可減少使用者瀏覽到不適切資訊造成誤導之機會。MyRR 系統上線提供服務後，使用者即可輕易獲得參考價值高且符合其閱讀習慣之網頁資訊，真正達到資訊因人制宜之效益。

5、誌謝

本研究承國科會研究計劃補助（計劃編號：NSC 94-2218-E-320-006），謹此誌謝。

參考文獻

[1] A. Risk and J. Dzenowagis, “Review of internet health information quality initiatives”, *J Med Internet Res,* October 2001.

[2] G. Jeh and J. Widom, “Scaling personalized web search”, *In Proceedings of the 12th international World Wide Web conference,* 271 – 279, 2003.

[3] H. A. Liszka, T. E. Steyer, W. J. Hueston, “Virtual Medical Care: How Are Our Patients Using Online Health Information?”, *Journal of Community Health*, Vol. 31, No. 5, 368-378, 2006

[4] J. Kleinberg, “Authoritative sources in a hyperlinked environment”, *Journal of the ACM,* Vol. 46, Issue 5, 604-632, September 1999.

[5] K. S. Shuyler and K. M. Knight, “What Are Patients Seeking When They Turn to the Internet? Qualitative Content Analysis of Questions Asked by Visitors to an Orthopaedics Web Site”, *Journal of Medical Internet Research,* Vol. 5, No. 4, 2003

[6] L. Page, S. Brin, R. Motwani, and T. Winograd, “The PageRank citation ranking: bringing order to the Web”, *Technical report, Stanford University Database Group,* 1998.

[7] T. H. Haveliwala, “Topic-sensitive PageRank: a context-sensitive ranking algorithm for Web search”, *IEEE Transactions on Knowledge and Data Engineering,* Vol. 15, No. 4, 784-796, 2003.

[8] P. Wilson, “How to find the good and avoid the bad or ugly: a short guide to tools for rating quality of health information on the internet”, *British Medical Journal,* Vol. 324, No.7337, 598-600, 2002