

نمذجة وتحليل وصلات روابط افتراضية بين مواقع المراكز التابعة لجامعة الموصل

طلال فاضل حسين

باسل يونس الخياط

كلية علوم الحاسوب والرياضيات

جامعة الموصل

تاريخ قبول البحث: 2013/01/30

تاريخ استلام البحث: 2012/10/07

ABSTRACT

This paper deals with PageRanks which are used in the analysis of web links. It displays an introduction for the concept PageRanks and also some basic theoretical aspects related to it. It exposures three common methods for the purpose of calculating PageRanks: the direct method, the Markovian method and the power method. In order of fortification, these methods are applied on Internet sites being considered for the centers of the University of Mosul. Because the lack of realistic data, we dealt with hypothetical data for modeling and analysis of the links between sites centers affiliated to the University of Mosul. It is clear from the practical application of the three methods that they give consistent results.

الملخص

يتناول هذا البحث رتب بيج التي تستخدم في مجال تحليل وصلات الويب. ويستهل بعرض لمفهوم رتب بيج وكذلك بعض الجوانب النظرية الأساسية المتعلقة به. وتعرض ثلاث طرائق شائعة لغرض حساب رتب بيج: هي الطريقة المباشرة وطريقة الأتمودج الماركوفي وطريقة القوة. ومن أجل الإغناء التطبيقي تتم دراسة مواقع الانترنت للمراكز التابعة لجامعة الموصل. وبالنظر لعدم توفر بيانات واقعية فيتم التعامل مع بيانات افتراضية لنمذجة وتحليل وصلات الروابط بين مواقع المراكز التابعة لجامعة الموصل. ويتضح من التطبيق العملي إن الطرائق الثلاث تعطي نتائج منسجمة.

1- مقدمة:

رتبة بيج PageRank هي خوارزمية تحليل وصلة Link Analysis Algorithm، ولدت في جامعة ستانفورد Stanford University وسميت على اسم مبتكرها Larry Page، ثم طورت من قبل Sergey Brin بوصفها جزء من مشروع بحث لنوع جديد من محركات البحث، وقد نشرت أول الأوراق البحثية عنها في العام 1998، علما بان Larry Page و Sergey Brin كانا المؤسسين لشركة Google [2]. تستخدم هذه الخوارزمية من قبل ماكنة بحث الإنترنت لشركة Google والتي تُخصّص قيمة عددية (وزناً عددياً) إلى كلّ عنصر من مجموعة وصلة تشعبية Hyperlinked من الوثائق Documents، مثل الشبكة العالمية World Wide Web، لغرض "قياس" الأهمية النسبية ضمن المجموعة. هذه الخوارزمية يمكن أن تُطبق على أيّ مجموعة من الكيانات بالتقديرات والإشارات المتبادلة. إن الوزن العددي الذي يُخصّص إلى أيّ عنصر (صفحة) مُعطى، وليكن E، يُدعى أيضاً رتبة بيج العنصر E ويرمز له PR(E).

تُحسب رتبة بيج عن طريق حل معادلة من 500 مليون متغير وأكثر من ملياري عبارة. وتعتمد رتبة بيج على البنية الترابطية التي يتميز بها الإنترنت ويستخدمها أداة لتنظيم عمله، بمعنى أن المحرك Google يفسر ارتباطاً من الصفحة أ إلى الصفحة ب على أنه "تصويت" قامت به الصفحة أ لمصلحة الصفحة ب. ويقوم Google بأهمية الصفحة بعدد التصويتات التي يتلقاها. كما أنه يحل أيضاً الصفحة التي تجري التصويت. فإذا كانت الصفحات التي تصوّت "مهمة"، يعطيها تصويتها ذلك وزناً أكبر، ويجعل الصفحات الأخرى التي تصوت لها "مهمة" أيضاً. إن رتبة بيج يُمكن أن تُحسب لحشود من الوثائق ومن أيّ حجم كانت. وهي تقتصر بان التوزيع يُقسّم بانتظام بين كلّ الوثائق في الحشد في بداية العملية الحسابية.

تصف شركة Google رتبة بيج بأنها تعكس رتبة بيج وجهة نظرها عن أهمية صفحات الويب باعتبار أكثر من 500 مليون متغير و 2 بليون علاقة متبادلة. إن الصفحات التي يعتقد بأنها صفحات مهمة تستلم رتبة بيج أعلى ومرجحة للظهور في أعلى نتائج البحث، كذلك فإن رتبة بيج تعطي الأهمية لكلّ صفحة تجمع صوتاً، كأصوات من بعض الصفحات التي تعدّ نوات قيمة أعظم. لذا تُعطى الصفحة الموصولة (المرتبطة) Linked Page قيمة أعظم. وعادة تستخدم مناهج واقعية للمساعدة على تحسين نوعية التقصي وخلق منتجات مفيدة، وتُستعمل تقنيات خاصة بالويب لتقرير أهمية كل صفحة، أما الأعمال المنجزة في جامعة الموصل ذات العلاقة بالإنترنت فلم نجد سوى تقرير دبلوم في قسم علوم الحاسوب قدمه الباحث فتحي (2011)، إذ قام بتقديم نموذج يعتني بأبرز معايير تقييم المواقع الالكترونية الشائعة والمتفق عليها من قبل جمهور الباحثين، ليسترشد بها في تقييم المواقع الالكترونية ويساعد في رفع مستويات الجودة، ويسهم في جعل إعداد المواقع وإدارتها عملاً علمياً وتعليمياً.

2- نظرية رتب بيج PageRank Theory:

في هذا المبحث نقدم خوارزمية مهمة جداً تستخدم من قبل جوجل في ترتيب صفحات الويب في الإنترنت في تصفح الإنترنت، عادة يستخدم المتصفحون محركات البحث للعثور على صفحات الويب ذات الصلة لتلبية احتياجاتهم. وفي كثير من الأحيان يمكن أن يكون هناك آلاف من صفحات الويب التي لها صلة باحتياجاتهم. لذا تبرز الحاجة لوضع قائمة للصفحات الويب وفق أهميتها وينبغي أيضاً أن يتم تحديث هذه القائمة بانتظام وبشكل متكرر وبالتالي فإنه من المهم أن تسعى لخوارزمية سريعة لحوسبة الصفحات حسب أهميتها وكذلك إقلال زمن التحديث. وتكمن صعوبة هذه المشكلة بسبب الحجم الكبير لصفحات الويب في شبكة الإنترنت فضلاً عن النمو السريع في هذا الحجم [11].

لقد ابتكر الباحثون [8] رتب بيج بوصفها القيم العددية التي تفيد في تحديد أهمية صفحات الويب وذلك بالاستفادة من فكرة استخدمها الباحث [6] في مجال تحديد أهمية المجالات في المكتبات يسمى عامل التأثير impact factor، إذ أن عامل التأثير يمثل معدل عدد الاقتباسات للبحوث المنشورة مؤخراً في المجلة. تعرّف رتب بيج على النحو الآتي: إذا كان N يمثل عدد صفحات الويب الكلية، فإن مصفوفة الارتباط التشعبي تعرّف كما يأتي:

$$I_{ij} = \begin{cases} 1/k & \text{إذا كانت الصفحة } i \text{ صادرة عن الصفحة } j \\ 0 & \text{في خلاف ذلك} \end{cases}$$

حيث k إن يمثل عدد جميع الصفحات الصادرة عن الصفحة j .

إن قيمة رتبة بيج تعكس مقدار فرصة (حظ) الخادم (server) العشوائي الذي سيحط على الصفحة من خلال النقر على وصلة. إن المصفوفة L التي تضم العناصر L_{ij} يمكن عددها مصفوفة انتقالية، وتسمى مصفوفة الوصلة التشعبية Hyperlink Matrix.

إن الفكرة الأولية لرتب بيج يمكن تمثيلها بوصفها التوزيع الاحتمالي probability distribution لمستخدم انترنت/انترنت يقوم بالنقر على وصلات عشوائيا لتمثيل أرجحية الوصول إلى وصلة معينة، لذا فإن قيم رتب بيج تكون محصورة بين 0 و 1. إن رتبة بيج يُمكن أن تُحسب لحشود من الوثائق ومن أي حجم كانت. وهي تفترض أن التوزيع يُقسّم بانتظام بين كل الوثائق في الحشد في بداية العملية الحسابية. ونقدم في أدناه ثلاث طرائق شائعة لحساب رتبة بيج.

2-1 الطريقة المباشرة

لو كانت A_1, \dots, A_N تمثل مواقع مربوطة في شبكة انترنت/انترنت وبفرض أن $L(A_1), \dots, L(A_N)$ تمثل الروابط الصادرة من A_1, \dots, A_N ، على التوالي، ولو كانت B هي أية صفحة في تلك الشبكة وكانت S_B تمثل المجموعة جميع الروابط التي ترتبط بالصفحة B ، فإن رتبة بيج للصفحة B في تلك الشبكة، يرمز لها $PR(B)$ ، وتعرّف على النحو الآتي [4]:

$$PR(B) = \sum_{B \in S_B} \frac{PR(B)}{L(B)}, \quad \dots(1)$$

ولو كان d يمثل عامل التثبيت (التخميد) damping factor (أي احتمالية استمرار المستخدم في خطوة معينة) فيمكن إعادة كتابة المعادلة السابقة على النحو الآتي [4]:

$$PR(B) = \frac{1-d}{N} + d \sum_{B \in S_B} \frac{PR(B)}{L(B)}, \quad \dots(2)$$

إذا أن $0 < d < 1$. ونود الإشارة إلى أن القيمة المستخدمة عادة من عامل التثبيت هي حوالي $d=0.85$. لو فرضنا الآن أن المتجه العمودي ذو البعد $N \times 1$ هو كما يأتي:

$$R = \begin{bmatrix} PR(A_1) \\ \vdots \\ PR(A_N) \end{bmatrix} \quad \dots(3)$$

يمثل متجه رتب بيج المطلوب إيجادها، إذ أن $\sum_{i=1}^N PR(A_i) = 1$. فيمكن تحويل المعادلة السابقة إلى معادلة

مصفوفات وكما يأتي:

$$R = d \begin{bmatrix} l(A_1, A_1) & l(A_1, A_2) & \dots & l(A_1, A_N) \\ l(A_2, A_1) & l(A_2, A_2) & \dots & l(A_2, A_N) \\ \vdots & \vdots & \ddots & \vdots \\ l(A_N, A_1) & l(A_N, A_2) & \dots & l(A_N, A_N) \end{bmatrix} R + \frac{(1-d)}{N} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \dots(4)$$

حيث أن $l(A_i, A_j)$ هي دالة تجاور adjacency function تمثل نسبة الانتقالات من A_i إلى A_j وان $l(A_i, A_j) = 0$ عند عدم وجود اتصال بين A_i و A_j . وهذه الدالة تكون معيارية مجموع كل صف فيها يساوي 1، أي تحقق القيد الآتي:

$$\sum_{i=1}^N l(A_i, A_j) = 1. \quad \dots(5)$$

كما يمكن كتابة المعادلة (4) على النحو الآتي:

$$R = dLR + \frac{(1-d)}{N} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \dots(6)$$

إذ أن

$$L = \begin{bmatrix} l(A_1, A_1) & l(A_1, A_2) & \dots & l(A_1, A_N) \\ l(A_2, A_1) & l(A_2, A_2) & \dots & l(A_2, A_N) \\ \vdots & \vdots & \ddots & \vdots \\ l(A_N, A_1) & l(A_N, A_2) & \dots & l(A_N, A_N) \end{bmatrix} \quad \dots(7)$$

وتسمى المصفوفة L بمصفوفة الارتباط التبعي hyperlink matrix. ولحل المعادلة رقم (6) يمكن تحول هذه المسألة إلى مسألة متجه ذاتي eigenvector لغرض تسهيل المعالجة الرياضية وتقليل أخطاء التنوير، [9].

2-2 الطريقة غير المباشرة

تستند هذه الطريقة على النمذجة الماركوفية لمحركات البحث Search Engines Markovian Modeling of. لو فرضنا أن $L_{ii} > 0$ لجميع قيم i . فهذا يعني انه لكل صفحة ويب هناك وصلة تُشير إليها. لذا فإن L يُمكن عدها مصفوفة انتقالية لسلسلة ماركوفية ووفق أنموذج المشي العشوائي Random Walk. وهذا يعني أن المتصفح قد عدّ ماشياً عشوائياً في حين أن صفحات الويب Webpage قد عدت حالات States لسلسلة ماركوفية. وبذلك يمكن نمذجة رتب بيج وفق الأنموذج الماركوفي، حيث إن الحالات states تمثل الصفحات pages، والانتقالات transitions تمثل الروابط (الوصلات) بين الصفحات، [2].

يقال للسلسلة الماركوفية بأنها غير قابلة للاختزال Irreducible إذا كان من الممكن الوصول إلى أي حالة فيها من أي حالة أخرى، [1]. إن الشرط الأساسي الواجب توفره للتأكد من أن التوزيع المتزن موجود هو أن تكون السلسلة الماركوفية غير قابلة للاختزال Irreducible وغير دورية. وفي تطبيقات محركات البحث فربما تكون السلسلة الماركوفية قابلة للاختزال Reducible، كما أن حجوم المصفوفات الانتقالية تكون عادة ضخمة جداً. من ناحية أخرى فإن الوقت المطلوب لحساب رتبة بيج من قبل Google لا يتجاوز بضعة أيام. لذا فإن الطريقة السابقة (الطريقة المباشرة) في حساب رتب بيج (الاحتمالات المتزنة) ليست مرغوب فيها لكونها لا تلبى احتياجات هذه المسألة. لذا اهتم الباحثون في إيجاد بدائل، وهناك عدة طرائق متوافرة حالياً لمعالجة هذه المسألة.

وبفرض أن السلسلة الماركوفية غير قابلة للاختزال Irreducible وغير دورية، فإن التوزيع المتزن يعرف رياضياً على النحو الآتي، [1]

$$\underline{\pi} = (\pi_0, \pi_1, \dots, \pi_m) = \lim_{n \rightarrow \infty} P^n \quad \dots(8)$$

حيث إن P تمثل مصفوفة الاحتمالات الانتقالية في الأنموذج الماركوفي. وبخصوص صفحات الويب فإن التوزيع المتزن يكون موجوداً وأن π_i تمثل نسبة الوقت الذي يقوم به المشي العشوائي Random Walker (المتصفح Surfer) بزيارة المحرك (صفحة الويب Webpage) i. إن القيمة الأكبر من π_i تشير إلى أن صفحة الويب الأكثر أهمية هي الصفحة i. وعلى هذا الأساس فإن رتبة بيج لصفحة الويب i سوف تعرف لكي تكون π_i ، وسنطلق على هذه الطريقة في إيجاد رتبة بيج بالطريقة غير المباشرة.

2-3 طريقة القوة

تستخدم حالياً طريقة القوة power method على نطاق واسع لمعالجة مسألة رتبة بيج. إن طريقة القوة هي طريقة تكرارية iterative method لإيجاد أكبر قيمة ذاتية largest eigenvalue بالقيمة المطلقة وكذلك إيجاد المتجه الذاتي eigenvector المقابل، [7]. إن هذه الطريقة تقدم حلاً عملياً لمسألة رتبة بيج، إذ عندما يكون عدد صفحات الويب أكثر من 80 مليون صفحة وعند المعلمة $d = 0.85$ فيمكنها إنجاز الحل خلال 50 تكرار فقط. للمعالجات الرياضية وتفصيل طريقة القوة [5]. ونشير إلى أن طريقة القوة هي التي تعتمدها Google بهذا المجال. والخوارزمية الآتية تلخص طريقة القوة لإيجاد رتب بيج.

الخوارزمية (1): طريقة القوة لإيجاد رتب بيج:

الخطوة (1): إدخال المصفوفة الانتقالية P.

الخطوة (2): حساب المصفوفة Q من العلاقة الآتية:

$$Q = d \begin{bmatrix} l(A_1, A_1) & l(A_1, A_2) & \dots & l(A_1, A_N) \\ l(A_2, A_1) & l(A_2, A_2) & \dots & l(A_2, A_N) \\ \dots & \dots & \dots & \dots \\ l(A_N, A_1) & l(A_N, A_2) & \dots & l(A_N, A_N) \end{bmatrix} P + \frac{(1-d)}{N} \begin{bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{bmatrix} \quad \dots(9)$$

الخطوة (3): إيجاد مقلوب transpose المصفوفة Q، أي Q'.

الخطوة (4): إيجاد القيم الذاتية والمتجهات الذاتية للمصفوفة Q'، [10].

الخطوة (5): اخذ المتجه الذاتي الأول (الذي يقابل القيمة الذاتية الأولى والتي تساوي دائماً 1). وبقسمة كل عنصر من عناصر هذا المتجه على مجموع العناصر (أي تطبيع Normalization) نحصل على رتب بيج.

3- الروابط الافتراضية في جامعة الموصل وتحليل وصلاتها:

يوجد في جامعة الموصل 7 مراكز تابعة لرئاسة الجامعة وهي:

1- مركز الحاسوب والانترنت وسنرمز لهذا المركز $q_1 = \text{Webpage 1}$.

2- مركز التحسس النائي وسنرمز لهذا المركز $q_2 = \text{Webpage 2}$.

- 3- مركز طرائق التدريس وسنرمز لهذا المركز $Webpage 3 = q_3$.
- 4- مركز دراسات الموصل وسنرمز لهذا المركز $Webpage 4 = q_4$.
- 5- مركز الدراسات الإقليمية وسنرمز لهذا المركز $Webpage 5 = q_5$.
- 6- مركز بحوث السودان والموارد المائية وسنرمز لهذا المركز $Webpage 6 = q_6$.
- 7- مركز البيئة والسيطرة على التلوث وسنرمز لهذا المركز $Webpage 7 = q_7$.

بالنظر لعدم توفر بيانات واقعية حول الروابط بين المراكز التابعة لجامعة الموصل فقد قام الباحث بالاتصال ببعض الأساتذة ذوي الخبرة في جامعة الموصل حول رؤيتهم للروابط المفروضة بين هذه المراكز، وفي ضوء ذلك افترض أن الانتقالات هي كما مبينة في أدناه:

$Webpage 1 \rightarrow 1, 1 \rightarrow 2, 1 \rightarrow 3, 1 \rightarrow 4, 1 \rightarrow 5, 1 \rightarrow 6, 1 \rightarrow 7,$

$Webpage 2 \rightarrow 1, 2 \rightarrow 2, 2 \rightarrow 3, 2 \rightarrow 6, 2 \rightarrow 7$

$Webpage 3 \rightarrow 1, 3 \rightarrow 3.$

$Webpage 4 \rightarrow 1, 4 \rightarrow 4, 4 \rightarrow 5$

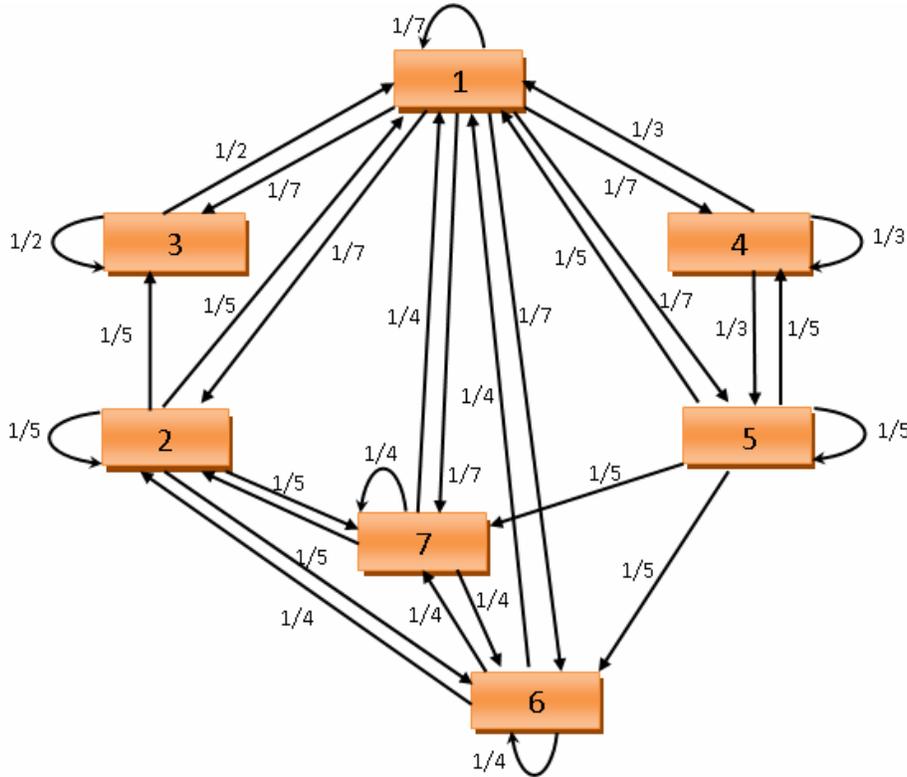
$Webpage 5 \rightarrow 1, 5 \rightarrow 4, 5 \rightarrow 5, 5 \rightarrow 6, 5 \rightarrow 7$

$Webpage 6 \rightarrow 1, 6 \rightarrow 2, 6 \rightarrow 6, 6 \rightarrow 7,$

$Webpage 7 \rightarrow 1, 7 \rightarrow 2, 7 \rightarrow 6, 7 \rightarrow 7.$

والشكل التالي يوضح المخطط الانتقالي بين صفحات الويب الكلية المفروضة الخاصة بالمراكز التابعة

لجامعة الموصل.



الشكل (1). المخطط الانتقالي بين صفحات الويب الكلية الخاصة بالمراكز التابعة لجامعة الموصل.

والهدف من هذا البحث هو نمذجة وتحليل الوصلات بين صفحات الانترنت لهذه المراكز. ونقوم الآن بإيجاد رتب
بيج باستخدام الطرائق الثلاث المذكورة سابقا:

أولا: الطريقة المباشرة:

بالتعويض في المعادلة (4)، نجد أن:

$$\begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \\ q_7 \end{pmatrix} = d \begin{pmatrix} 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \\ 1/5 & 1/5 & 1/5 & 0 & 0 & 1/5 & 1/5 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 & 0 & 0 \\ 1/5 & 0 & 0 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/4 & 1/4 & 0 & 0 & 0 & 1/4 & 1/4 \\ 1/4 & 1/4 & 0 & 0 & 0 & 1/4 & 1/4 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \\ q_7 \end{pmatrix} + \frac{(1-d)}{7} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \dots(10)$$

أي أن:

$$\begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \\ q_7 \end{pmatrix} = \begin{pmatrix} \frac{d}{7}(q_1 + q_2 + q_3 + q_4 + q_5 + q_6 + q_7) \\ \frac{0.85}{5}(q_1 + q_2 + q_3 + q_6 + q_7) \\ \frac{0.85}{2}(q_1 + q_3) \\ \frac{0.85}{3}(q_1 + q_4 + q_5) \\ \frac{0.85}{5}(q_1 + q_4 + q_5 + q_6 + q_7) \\ \frac{0.85}{4}(q_1 + q_2 + q_6 + q_7) \\ \frac{0.85}{4}(q_1 + q_2 + q_6 + q_7) \end{pmatrix} + \begin{pmatrix} \frac{1-d}{7} \\ \frac{1-d}{7} \\ \frac{1-d}{7} \\ \frac{1-d}{7} \\ \frac{1-d}{7} \\ \frac{1-d}{7} \\ \frac{0.15}{7} \end{pmatrix} \quad \dots(11)$$

وبعد تبسيط المعادلة (11)، نحصل على المعادلات الآتية:

$$q_1 = \frac{d}{7}(q_1 + q_2 + q_3 + q_4 + q_5 + q_6 + q_7) + \frac{1-d}{7}$$

$$q_2 = \frac{d}{5}(q_1 + q_2 + q_3 + q_6 + q_7) + \frac{1-d}{7}$$

$$q_3 = \frac{d}{2}(q_1 + q_3) + \frac{1-d}{7}$$

$$q_4 = \frac{d}{3}(q_1 + q_4 + q_5) + \frac{1-d}{7}$$

$$q_5 = d(q_1 + q_4 + q_5 + q_6 + q_7) + \frac{1-d}{7}$$

$$q_6 = \frac{d}{4}(q_1 + q_2 + q_6 + q_7) + \frac{1-d}{7}$$

$$q_7 = \frac{d}{4}(q_1 + q_2 + q_6 + q_7) + \frac{1-d}{7}$$

وبتبسيط المعادلات الأخيرة نجد أن:

$$\frac{7-d}{7}q_1 - \frac{d}{7}(q_2 + q_3 + q_4 + q_5 + q_6 + q_7) = \frac{1-d}{7}$$

$$\frac{5-d}{5}q_2 - \frac{d}{5}(q_1 + q_3 + q_6 + q_7) = \frac{1-d}{7}$$

$$\frac{2-d}{2}q_3 - \frac{d}{2}q_1 = \frac{1-d}{7}$$

$$\frac{3-d}{3}q_4 - \frac{d}{3}(q_1 + q_5) = \frac{1-d}{7}$$

$$(1-d)q_5 - d(q_1 + q_4 + q_6 + q_7) = \frac{1-d}{7}$$

$$\frac{4-d}{4}q_6 - \frac{d}{4}(q_1 + q_2 + q_7) = \frac{1-d}{7}$$

$$\frac{4-d}{4}q_7 - \frac{d}{4}(q_1 + q_2 + q_6) = \frac{1-d}{7}$$

ولحل المعادلات السبع الأخيرة، وبعد الاستفاد من القيد $\sum_{i=1}^3 q_i = 1$ ، نحول المعادلات الخطية السبع إلى

مصفوفات بالصيغة العامة المعروفة والتي هي بالشكل العام $QX = B$ ، فيكون الحل العددي هو بالشكل $X = Q^{-1}B$ ، إذ أن $Q = (q_1 \ q_2 \ q_3 \ q_4 \ q_5 \ q_6 \ q_7)$. ولما كانت Q هي مصفوفة غير مربعة، لذا يمكن إيجاد المعكوس المزيّف Pseudo Inverse، ويمكن الحصول عليه من الدالة الحاسوبية pinv. والجدول رقم (1) يلخص النتائج التي تم الحصول عليها عند $d=0.85$:

الجدول (1). رتب بيچ للروابط الافتراضية في جامعة الموصل وفق الطريقة المباشرة.

Webpage	1	2	3	4	5	6	7
رتبة بيچ	-0.0327	0.0378	0.0131	-0.1415	-0.4009	0.0392	0.0392

ومن الجدول رقم (1) نستنتج أن:

Webpage6 = Webpage 7 > Webpage 2 > Webpage3 > Webpage1 > Webpage4 > Webpage5

إذ أن القيم الأكبر تشير إلى صفحات الويب الأكثر أهمية.

ثانياً: الطريقة غير المباشرة (استخدام الأتمودج الماركوفي):

كما هو معروف فإن العناصر الأساسية للأتمودج الماركوفي هي فضاء الحالة والمصفوفة الانتقالية،

[1]. إن فضاء الحالة، والذي يمثل المجموعة التي تضم صفحات الويب الكلية الخاصة بالمراكز التابعة لجامعة

الموصل، سيكون $S = \{1,2,3,4,5,6,7\}$. أما المصفوفة الانتقالية فتستكون على النحو الآتي:

$$P = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{matrix} & \begin{pmatrix} 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \\ 1/5 & 1/5 & 1/5 & 0 & 0 & 1/5 & 1/5 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 & 0 & 0 \\ 1/5 & 0 & 0 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/4 & 1/4 & 0 & 0 & 0 & 1/4 & 1/4 \\ 1/4 & 1/4 & 0 & 0 & 0 & 1/4 & 1/4 \end{pmatrix} \end{matrix}$$

وكما هو واضح فان مجموع كل صف في المصفوفة السابقة يساوي 1.

إن رتبة بيج تمثل التوزيع المتزن للأنموذج الماركوفي والذي يمكن الحصول عليه من حل المعادلة الآتية، [1].
 $\pi(I - P) = 0$... (12)

حيث أن:

$$I = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

وان $\pi = (\pi_1 \pi_2 \pi_3 \pi_4 \pi_5 \pi_6 \pi_7)$ يمثل متجه التوزيع المتزن وفي الوقت نفسه يمثل رتب بيج لصفحات الويب الكلية الخاصة بالمراكز التابعة لجامعة الموصل. وبحل المعادلة الأخيرة حصلنا على الحل الآتي:

$$\pi = (0.2508 \ 0.1451 \ 0.1297 \ 0.0768 \ 0.0768 \ 0.1604 \ 0.1604)$$

وباعتبار أن الاحتمالات المتزنة تمثل رتب صفحات الويب السبع، فان الجدول رقم (2) يلخص النتائج التي تم الحصول عليها وفق هذه الطريقة:

الجدول (2). رتب بيج للروابط الافتراضية في جامعة الموصل وفق الطريقة غير المباشرة.

Webpage	1	2	3	4	5	6	7
رتبة بيج	0.2508	0.1451	0.1297	0.0768	0.0768	0.1604	0.1604

ومن الجدول رقم (2) نستنتج أن رتب صفحات الويب ستكون (حسب أهميتها من حيث الاتصالات) على النحو الآتي:

$$\text{Webpage 1} > \text{Webpage 6} = \text{Webpage 7} > \text{Webpage 2} > \text{Webpage 3} > \text{Webpage 4} = \text{Webpage 5}$$

ثالثاً: باستخدام طريقة القوة:

بفرض أن $d=0.85$ وباستخدام طريقة القوة تم إيجاد القيم الذاتية للمصفوفة الانتقالية P وكانت القيمة

الذاتية الأولى 1، وكان المتجه الذاتي المقابل لهذه القيمة الذاتية:

$$[0.5958 \ 0.3533 \ 0.3242 \ 0.2312 \ 0.2312 \ 0.3926 \ 0.3926]'$$

وبتحويل هذا المتجه الذاتي إلى متجه معياري وذلك بقسمته على مجموع عناصره (2.5207) تم الحصول على المتجه الآتي:

$$[0.2364 \ 0.1401 \ 0.1286 \ 0.0917 \ 0.0917 \ 0.1557 \ 0.1557]'$$

وهذا المتجه يمثل رتب بيج لصفحات الويب الكلية الخاصة بالمراكز التابعة لجامعة الموصل وهو موضح في الجدول رقم (3):

الجدول (3). رتب بيج للروابط الافتراضية في جامعة الموصل وفق طريقة القوة.

Webpage	1	2	3	4	5	6	7
رتبة بيج	0.2364	0.1401	0.1286	0.0917	0.0917	0.1557	0.1557

ومن الجدول رقم (3) نستنتج أن:

$$\text{Webpage1} > \text{Webpage 6} = \text{Webpage 7} > \text{Webpage2} > \text{Webpage3} > \text{Webpage4} = \text{Webpage5}$$

4- الاستنتاجات والمناقشة:

الجدول رقم (4) يوضح نتائج رتب بيج للروابط الافتراضية في جامعة الموصل والتي تم الحصول عليها من الطرائق الثلاثة.

الجدول (4). رتب بيج للروابط الافتراضية والتي تم الحصول عليها من الطرائق الثلاثة.

7	6	5	4	3	2	1	صفحة الويب حسب الأهمية
5	4	1	3	2	7	6	الطريقة المباشرة
5	4	3	2	7	6	1	الطريقة غير المباشرة
5	4	3	2	7	6	1	طريقة القوة

وكما هو واضح فان نتائج طريقة القوة مطابقة تماما لنتائج الأنموذج الماركوفي. أما نتائج الطريقة المباشرة فتختلف قليلا عن الطريقتين السابقتين. ونشير إلى أن نتائج الطريقة المباشرة تعتمد على القيمة المختارة من المعلمة d فضلا عن أنها تتطلب حسابات كثيرة، لذا لا يوصى باستعمالها وبخاصة عندما يكون عدد الصفحات كبير .

مما تقدم نجد أن رتب بيج لصفحات الويب الكلية الخاصة بالمراكز التابعة لجامعة الموصل وفق

الروابط المفروضة بين هذه المراكز ستكون على النحو الآتي:

1- مركز الحاسوب والانترنت رتبته الأولى.

2- مركز بحوث السدود والموارد المائية رتبته الثانية.

3- مركز البيئة والسيطرة على التلوث رتبته الثالثة.

4- مركز التحسس النائي رتبته الرابعة.

5- مركز طرائق التدريس رتبته الخامسة.

6- مركز دراسات الموصل رتبته السادسة.

7- مركز الدراسات الإقليمية رتبته السابعة.

المصادر

- [1] الخياط، باسل يونس ذنون (2011a) "النمذجة الماركوفية مع تطبيقات عملية- الجزء الأول"، دار ابن الأثير للطباعة والنشر، جامعة الموصل.
- [2] الخياط، باسل يونس ذنون (2011b) "النمذجة الماركوفية مع تطبيقات عملية- الجزء الثاني"، دار ابن الأثير للطباعة والنشر، جامعة الموصل.
- [3] حسو، مهى عبد الرحمن و فتحي، فتحي غازي (2011) "دراسة معايير تقييم المواقع الالكترونية"، رسالة دبلوم عالي غير منشورة، كلية علوم الحاسوب والرياضيات، جامعة الموصل.
- [4] Brin, S.; Page, L. (1998) "**The anatomy of a large-scale hypertextual Web search engine**". Computer Networks and ISDN Systems 30.
- [5] Ching, Wai-Ki and Ng, Michael K. (2006) "**Markov Chains: Models, Algorithms and Applications**", Springer Science and Business Media, Inc.
- [6] Garfield E (1972) "**Citation Analysis as a Tool in Journal Evaluation**", Science, 178.
- [7] Golub G and van Loan C (1989) "**Matrix Computations**", The John Hopkins University Press.
- [8] Page L, Brin S, Motwani R and Winograd T (1998), "**The PageRank Citation Ranking: Bring Order to the Web**", Technical Report, Stanford University.
- [9] Ricardo, Henry (2010) "**A Modern Introduction To Linear Algebra**" (1st ed.), CRC Press.
- [10] Ronald W. Shonkwiler (2009) "**An Introduction to Parallel and Vector Scientific Computation**", Georgia Institute of Technology pp. 206-230.
- [11] Wenpu Xing, Ali Ghorbani (2004) "**Weighted PageRank Algorithm**", Proceedings of the Second Annual Conference on Communication Networks and Services Research (CNSR'04) IEEE computer society.