المجلة العراقية للعلوم الإحصائية (24) 2013
ص ص[1-16]

طريقة مقترحة لقياس نشاط مواقع الإنترنت مع التطبيق على موقع جامعة الموصل

أ.د.باسل يونس ذنون* سارة غانم محمود**

الملخص

يتضمن هذا البحث اقتراح طريقة جديدة لقياس نشاط مواقع الإنترنت المرتبطة بشبكة معينة. ويتم تطبيق هذه الطريقة على مواقع الإنترنت المرتبطة بجامعة الموصل. تعتمد الطريقة المقترحة أساسا على تقنية تحليل المركبات الرئيسة التي من خلالها تستخلص العوامل الأكثر تأثيراً في نشاط مواقع الإنترنت. وتتلخص الطريقة المقترحة بتحويل المركبات الرئيسة إلى معادلة تصنيفية تأخذ بالحسبان أهمية كل مركبة رئيسة مؤثرة. وقد اُستخدمت الطريقة المقترحة لتصنيف مواقع الكليات المرتبطة بموقع جامعة الموصل وأعطت نتائج مشجعة.

A Suggested Method for Measuring the Performance of Web Sites with Application to the Web Site of Mosul University

Abstract

This paper deals with a new proposed method to measure the activity of internet sites linked to a certain network. This method is then applied on the Web sites associated with the University of Mosul. The proposed method depends mainly on the principal components technique and from which is extracted the most effect factors on the activity of Web sites. It transforms the principal components to a classification equation taking into account the importance of each influential principal component. It has been used to classify websites of colleges associated with the University of Mosul, according to the proposed method, and it gives encouraging results.

* أستاذ / قسم الرياضيات / كلية علوم الحاسوب والرياضيات / جامعة الموصل

**طالبة ماجستير/ قسم الرياضيات / كلية علوم الحاسوب والرياضيات / جامعة الموصل

تاريخ استلام البحث 2013/3/5 ـــــــــــــــــــــــــــ **تأريخ القبول** 2013/6/9

1- مقدمة:

لقد أولى الإنسان منذ القدم اهتماما بالغا بالتصنيف وكانت بداية اهتمامه بتصنيف الكائنات إلى حية وغير وحية، وتطور هذا الاهتمام مابرح إلى إن تحول الى علم أولى الفلاسفة والمفكرون مجهودات قيمة لإرساء أُسس علم التصنيف ويعرف التصنيف على أنه تمثيل البيانات على شكل مجموعات مقسمة كل مجموعة تشترك بصفات وميزات مشتركة، تختلف طرق التصنيف بحسب الهدف المراد منها وحسب طبيعة البيانات التي يتم العمل عليها وذلك من اجل ترتيب الأشياء إلى أقسام معينة أو أصناف لسهولة تجميعها وفهمها. وقد يتم التصنيف عن طريق الإنسان بواسطة الإدراك والحس البشري أو أن يتم عن طريق الحاسوب عبر أنظمة متخصصة، إن ميادين التصنيف واسعة وله تطبيقات في شتى مجالات الحياة فله تطبيقات طبية وهندسية والكترونية وحتى مجالات إنسانية واقتصادية، [Bishop, 2006] و [Theodoridis, 2008] و[Theodoridis, 2010].

2- الدراسات سابقة :

لقد زاد استخدام الإنترنت بشكل كبير إلى أن وصل في عامي 1996 و 1997 إلى مستويات عالية جداً. ويرجع الفضل بهذا النمو إلى عدم وجود الإدارة المركزية وكذلك فإن البروتوكولات الخاصة بالإنترنت تشجع الأشخاص والشركات على تطوير الأنظمة وبيعها وكذلك فهي تمنع أية شركة من الغلو في السيطرة على الشبكة، بدأت شركات الاتصالات الخدمة على خدمة الجوال ISP بواسطة شبكة الهاتف عام 1995،(الخياط b 2012 ، Wikipedia)، وقد دفع هذا كله المفكرين والعلماء إلى التفكير بتصنيف محتويات صفحات الويب التي أصبح من البديهي القيام بذلك، وقد تم ذلك بالفعل عن طريق بحوث علمية عديدة. اذ قدم الباحثان Susan Dumais و Hao Chen في بحثهما المنشور والموسوم "Hierarchical Classification of Web Content" طريقة لتصنيف كم هائل من المعلومات على مواقع بحث عملاقة لتصنيف محتوياتها، وقد استخدمت طريقة التعليم الموجه آلة الدعم الموجه Support vector machine اختصاراً (SVM).

وفي عام 2007 قام الباحثان Xiaoguang Qi و Brian D.Davison بتصنيف محتوى صفحات الويب لمساعدة المستخدم على إيجاد المعلومات التي يحتاجها بسرعة ودقة جيدة، وقدم استخدمت طريقة الجار القريب KNN وذلك في ورقتهما البحثية الموسومة "Web Page Classification: Features and Algorithms"، حيث تم استخلاص ما يتم إدخاله من كلمات من قبل المستخدم. وفي العام 2006 قدم الباحثون Dou Shen وآخرون ورقة بحثية بعنوان " Building Bridges for Web Query Classification"، حيث قاموا بتصنيف استفسارات المستخدمين وتحديد رتبة الصفحة Page Rank للنتائج التي ستظهر.

أما بخصوص تصنيف الجامعات والمعاهد فهناك معايير عالمية للقيام بالتصنيف، فإن من أهم التصنايف العالمية الويب ماتركس (Web metrics Ranking of World Universities)، بدأ هذا التصنيف سنة 2004بتصنيف 16000جامعة ، يهدف هذا التصنيف بالدرجة الأولى إلى حث الجهات الأكاديمية في العالم لتقديم ما لديها من أنشطة علمية تعكس مستواها العلمي المتميز على الإنترنت وليس ترتيباً أو تصنيفاً للجامعات، بل ترتيباً لموقع الجامعة .(Ranking Web)ويتم عمل هذا التصنيف في شهر يناير ويوليو من كل سنة ، ويعتمد على قياس أداء الجامعات من خلال مواقعها الإلكترونية ضمن المعايير التالية (الحجم - الإشارة إلى الأبحاث - الأثر العام).

إن التصنيف الدولي للمواقع الإلكترونية للجامعات والكليات على الشبكة العالمية 4ICUs هو تصنيف استرالي عالمي، يشبه تصنيف الويب ماتركس الأسباني ولكن يهتم بقياس مدى شهرة المواقع الإلكترونية للجامعات التي نالت الاعتراف أو الاعتماد الأكاديمي من منظمات أو هيئات دولية.و عادة يتم إعلان تصنيف ICUs4 كل ستة أشهر، ويطلب من كل الكليات والجامعات المشاركة في هذا التصنيف إضافة وتحديثها بياناتها شهرياً. علما أن هذا التصنيف يحتوي على 9000 كلية وجامعة يتم تصنيفها وفقا لشهرة موقعها الإلكتروني على شبكة الإنترنت لدى 200 دولة، ويهدف هذا التصنيف إلى ترتيب الكليات والجامعات العالمية وفق شهرة وجماهيرية الموقع الإلكتروني للجامعات بشكل تقريبي، وهو لا يصنف مؤسسات التعليم العالي بناء على جودة التعليم أو مستوى الخدمات المقدمة ، ويعتمد التصنيف على ثلاثة مقاييس موضوعية ومستقلة على شبكة الإنترنت مستخلصة من ثلاثة محركات بحث وهي (تصنيف صفحة الجوجل - الروابط الداخلية بالياهو -تصنيف مرور اليكسا).

إما بخصوص جامعة الموصل فإن مركز الحاسوب والإنترنت يقوم بجهود طيبة بخصوص تصنيف ملحقات المواقع ونشاطاتها من خلال إصدار تقييمات دورية لتقييم أداء نشاطات مواقع الكليات المرتبطة بالموقع الرسمي للجامعة. وكما جاء في كتاب مركز الحاسوب والإنترنت المرقم 1707/49/9 فإن التقييم الصادر عنه يهتم بقياس مقدار "تفاعل الكليات والمراكز والمديريات التابعة لجامعة الموصل بموقعها الألكتروني ضمن نطاق الموقع الرسمي لجامعة الموصل". وعلى الرغم من ان هذا التقييم يمثل جانبا مهما بوصفة المعطيات لهذه المنظومة الحيوية، فأننا سنحاول في هذا البحث القيام بتطوير هذا القياس وذلك باضافة عامل الاستجابة لتلك المعطيات من قبل الزوار والمستفيدين. يهدف هذا البحث إلى اقتراح طريقة لتصنيف مواقع الكليات على أساس النشاط الذي يقوم به العاملون فضلا عن تفاعل المستخدمين مع موقع كل كلية.

3- تحليل المركبات الرئيسة Principle Component Analysis:

يعد تحليل المركبات الرئيسة من التقنيات التي تُستخدم في مجال التعرف على الأنماط ويهدف إلى تقليل عدد المتغيرات المتعلقة بالمسألة من خلال توليد متغيرات جديدة، تسمى المركبات الرئيسة، وبحيث تكون هذه المتغيرات الجديدة مسؤولة عن نسبة كبيرة من الاختلاف الموجود في المتغيرات. إن المركبات الرئيسة هي عبارة عن تركيبات خطية Linear Combinations من المتغيرات الأصلية تتميز بكونها غير مترابطة فيما بينها ومرتَبة حسب الأهمية من حيث نسبة السيطرة على الاختلاف الموجود في المتغيرات. إن احد أهم الأسباب في استخدام تحليل المركبات الرئيسة العثور على اصغر مجموعة من هذه المركبات والتي يمكنها أن تسيطر على نسبة كبيرة من الاختلاف الموجود فيما بين البيانات، [Richard, 1988].

يعود اكتشاف تحليل المركبات إلى الإحصائي كارل بيرسون في العام 1901، ويهدف إلى تحويل المتغيرات المعتمدة إلى عدد اقل من المركبات غير المعتمدة على بعضها البعض، بمعنى آخر الحصول على متغيرات جديدة مُرتَّبة حسب الأهمية وعلى شكل تركيبات خطية للمتغيرات الأصلية وغير مترابطة هندسيا،[Richard, 1988] [Theodoridis, 2010].

إن المركبات الرئيسية عبارة عن تركيبات خطية للمتغيرات العشوائية $X_1, X_2, \ldots, X_p$ غير مرتبطة مع بعضها البعض وتبايناتها اكبر ما يمكن ويرمز لها PC. فلو فرضنا أن المتجه $X' = [X_1, X_2, \ldots, X_p]$ له مصفوفة التغاير Covariance Matrix :

$$\sum = \begin{bmatrix} V(X_1) & \cdots & C(X_1, X_p) \\ \vdots & \ddots & \vdots \\ C(X_p, X_1) & \cdots & V(X_p) \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{bmatrix} \qquad (1)$$

وبفرض أن مصفوفة التغاير لها القيم الذاتية $\lambda_1, \lambda_2, \ldots, \lambda_p$، حيث أن $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$. نفرض أيضا أن $\underline{L}_i' = (l_{1i}, l_{2i}, \ldots, l_{pi})$ هو تركيب خطي للمتجه $X' = [X_1, X_2, \ldots, X_p]$، فإن المركبات الرئيسة تُعرَّف على النحو الآتي،[Richard, 1988] [Theodoridis, 2010]:

$$Y_1 = l_{11}X_1 + l_{21}X_2 + \cdots + l_{p1}X_p$$

$$Y_2 = l_2' \underline{X} = l_{12}X_1 + l_{22}X_2 + \cdots + l_{p2}X_p$$

⋮

$$Y_p = \underline{l}_p^{'} \underline{X} = l_{1p}X_1 + l_{2p}X_2 + \cdots + l_{pp}X_p \qquad (2)$$

أي إن

$$\therefore Y_i = \sum_{j=1}^{p} l_{ij}X_j \qquad ,i = 1,2,\ldots,p \qquad (3)$$

إذ إن تباينات هذه المركبات الرئيسة تحسب كما يأتي،[Theodoridis,2010]، :[Richard, 1988]

$$V(y_i) = V(\sum_{j=1}^{p} l_{ij}X_j) \qquad (4)$$

$$= \sum_{j=1}^{p} L_{ji}^2 V(X_j) + \sum_{k\neq j}\sum L_{ji}L_{ki}C(X_j, X_i)$$

ويمكن كتابة العلاقة الأخيرة بصيغة مصفوفات وكما يأتي، [Richard, 1988]:

$$\left.\begin{array}{r} V(y_i) = \underline{L}_i^{'}\sum \underline{L}_i, i = 1,2,\ldots,p \\ C(y_i, y_k) = \underline{L}_i^{'}\sum \underline{L}_k, i,k = 1,2,\ldots,p; \end{array}\right\} \qquad (5)$$

إن التركيب خطي لـ $\underline{L}_i^{'}\underline{X}$, $i = 1,2,\ldots,p$ ، يحقق القيد الآتي [Keinosuke, 1990]:

maximizes $V(\underline{L}_i^{'}\underline{X})$

subject to $\underline{L}_i^{'}\underline{L}_i = 1$

and $C(\underline{L}_i^{'}\underline{X}, \underline{L}_k^{'}\underline{X}) = 0$, for all k<i

لنفرض أن القيم الذاتية والمتجهات الذاتية لمصفوفة التغاير $\sum$ ممثلة على النحو الآتي: $(\lambda_1, e_1),(\lambda_2, e_2),\ldots,(\lambda_p, e_p)$ ، إذ إن $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$.

فإن المركبة الرئيسة $i = 1,2,\ldots,p$ ، يمكن حسابها بدلالة القيم الذاتية وعلى النحو الآتي، [Mueller, 2002]:

$$Y_i = \underline{e}_i^{'}\underline{X} = e_{1i}X_1 + e_{2i}X_2 + \cdots + e_{pi}X_p \quad ; \quad i = 1,2,\ldots,p$$

إذ إن:

$$V(Y_i) = \underline{e}_i' \sum \underline{e}_i = \lambda_i \quad ; \quad i = 1,2,\ldots, p \qquad (6)$$

$$C(Y_i, Y_k) = \underline{e}_i' \sum \underline{e}_i = 0 \ ; \ \text{for all } i \neq k. \qquad (7)$$

إن استخدام المركبات الرئيسة يتيح لنا تحقيق خاصيتين على درجة كبيرة من الأهمية من الناحية العملية، وهما [النشواتي، 2010]:

1- إن التباين الكلي في المركبات الرئيسة يساوي التباين الكلي في المتغيرات الأصلية، ويساوي أيضاً مجموع القيم الذاتية لمصفوفة التغاير، أي إن:

$$\sum_{i=1}^{p} V(X_i) = \sum_{i=1}^{p} \lambda_i = \sum_{i=1}^{p} V(Y_i)$$

2- إن معامل الارتباط بين أي زوج من المركبات الرئيسة يساوي صفرا. أي إن معامل الارتباط بين Y_i and Y_k لجميع قيم $i \neq k.$. أي إن:

$$\rho_{Y_{k,Y_i}} = 0, \forall i \neq k.$$

3- معامل الارتباط بين Y_i and Y_k:

$$\rho_{Y_{k,Y_i}} = \frac{e_{ki}\sqrt{\lambda_i}}{\sigma_{kk}} \quad ; \quad i,k = 1,2,\ldots, p.$$

إذ إن $\sigma_{kk} = V(X_k)$.

4-مواقع الإنترنت وتحديد أهميتها:

يُعدُّ تصنيف مواقع الإنترنت من المسائل التي تلاقي اهتمامات واسعة من قبل العديد من الجهات ذات العلاقة وذلك للنمو الهائل بشبكات الإنترنت وكثرة المواقع الالكترونية بسبب الطلب الكبير عليها من قبل شتى الجهات الرسمية والشعبية. إن تصنيف مواقع الإنترنت يساعد الجهات الرسمية، الحكومية والأهلية،على تحسين مستوى الإدارة لمواقع الويب الخاصة بهم بزيادة جودة المعلومات أو الأخبار المنشورة ودقتها والانفراد بمواضيع خاصة بهم دون باقي مواقع الويب. وبخصوص الجامعات والكليات فإن تصنيف مواقع الإنترنت يعدّ من الأمور ذات الأهمية لعلاقته الوثيقة بالتواصل المعرفي. كما يعدّ تصنيف الجامعات مهما عالميا لمعرفة ماهية عمل الجامعات وآلية التطور بها، لذا يتم تصنيفها سنويا ومن خلال هذا التصنيف يتم تقييمها علميا. مثال على ذلك التقييم التابع لهيئة البحوث العليا في أسبانيا

ويب ماتركس Webometrics حيث يتم تصنيف الجامعات حسب معايير معينة تؤخذ من مواقع الجامعات المصنَّفة، ومن خلال هذا التصنيف العالمي ترتب الجامعات "حسب الأهمية"، [Mueller, 2003و[Wikipedia] . وبخصوص جامعة الموصل فإنها تقوم بإصدار تسلسلات دورية لتقييم نشاطات مواقع الإنترنت المرتبطة بها، حيث يتم تقييم مواقع الكليات عن طريق مركز الحاسوب والإنترنت في جامعة الموصل لبيان فعاليات المواقع، ويتم ذلك التقييم عبر أخذ عدد الأخبار المنشورة على الموقع وعدد القوائم المضافة لكل موقع خلال فترة معينة يحددها مسؤولو المركز. ويُصدر مركز الحاسوب والإنترنت في جامعة الموصل دوريا تقييما للمواقع، وقد صدر المركز إلى حد الآن تقييمين ، انظر الجدولين 1 و 2.

الجدول(1): التقييم الأول لمواقع الإنترنت المرتبطة بجامعة الموصل لغاية 2012/9/4.

ت	اسم الكلية	عدد القوائم المضافة	عدد الإخبار المنشورة	مجموع النقاط	التقييم
1	تربية بنات	3	43	46	جيد جدا
2	الزراعة والغابات	9	25	34	جيد
3	طب الموصل	5	27	32	جيد
4	الحقوق	6	23	29	متوسط
5	هندسة الالكترونيات	1	21	22	متوسط
6	التربية الرياضية	1	18	19	متوسط
6(مكرر)	طب الأسنان	1	18	19	متوسط
7	التربية الأساسية	لا يوجد	18	18	متوسط
7(مكرر)	العلوم السياسية	لا يوجد	18	18	متوسط
8	العلوم	8	9	17	متوسط
9	الإدارة واقتصاد	لا يوجد	15	15	مقبول
10	الصيدلة	4	10	14	مقبول
11	طب نينوى	2	9	11	مقبول
12	الآثار	لا يوجد	10	10	ضعيف
13	الآداب	1	4	5	ضعيف
14	الطب البيطري	لا يوجد	4	4	ضعيف
14(مكرر)	علوم الحاسوب والرياضيات	لا يوجد	4	4	ضعيف
14(مكرر)	الفنون جميلة	لا يوجد	4	4	ضعيف
14(مكرر)	الهندسة	1	3	4	ضعيف
14(مكرر)	علوم البيئة وتقاناتها	1	3	4	ضعيف
14(مكرر)	العلوم الإسلامية	1	3	4	ضعيف
15	التمريض	لا يوجد	3	3	ضعيف
16	التربية	لا يوجد	2	2	ضعيف

الجدول(2): التقييم الثاني لمواقع الإنترنت المرتبطة بجامعة الموصل 2012/9/17-2012/11/14.

ت	اسم الكلية	عدد القوائم المضافة	عدد الإخبار المنشورة	مجموع النقاط	التقييم
1	الزراعة والغابات	17	95	112	امتياز
2	الآداب	1	86	87	امتياز
3	علوم الحاسوب والرياضيات	23	55	78	جيد جدا
4	الحقوق	9	56	65	جيد
4(مكرر)	التربية	14	51	65	جيد
5	الإدارة واقتصاد	10	51	61	جيد
6	تربية بنات	2	58	60	جيد
7	التربية الأساسية	لا يوجد	58	58	متوسط
8	الطب البيطري	4	48	52	متوسط
9	طب الأسنان	4	38	42	مقبول
10	العلوم السياسية	لا يوجد	40	40	مقبول
11	التمريض	16	17	33	مقبول
12	العلوم	6	24	30	مقبول
13	الآثار	3	24	27	ضعيف
14	طب الموصل	2	24	26	ضعيف
15	طب نينوى	2	23	25	ضعيف
16	هندسة الالكترونيات	1	18	19	ضعيف
17	الصيدلة	1	15	16	ضعيف
18	الفنون الجميلة	لا يوجد	11	11	ضعيف
18(مكرر)	التربية الرياضية	لا يوجد	11	11	ضعيف
19	علوم البيئة وتقاناتها	1	3	4	ضعيف
20	العلوم الإسلامية	لا يوجد	3	3	ضعيف
21	الهندسة	لا يوجد	لا يوجد	لا يوجد	ضعيف

ومن خلال مقارنة التقييمين السابقين يتضح ما يأتي:

1- هناك تباين واضح بين التقييمين على الرغم من تقارب المدة الزمنية بينهما لمعظم المواقع والكليات التابعة لجامعة الموصل، ماعدا كليات الزراعة والغابات والآثار وعلوم البيئة وتقاناتها والعلوم الإسلامية.

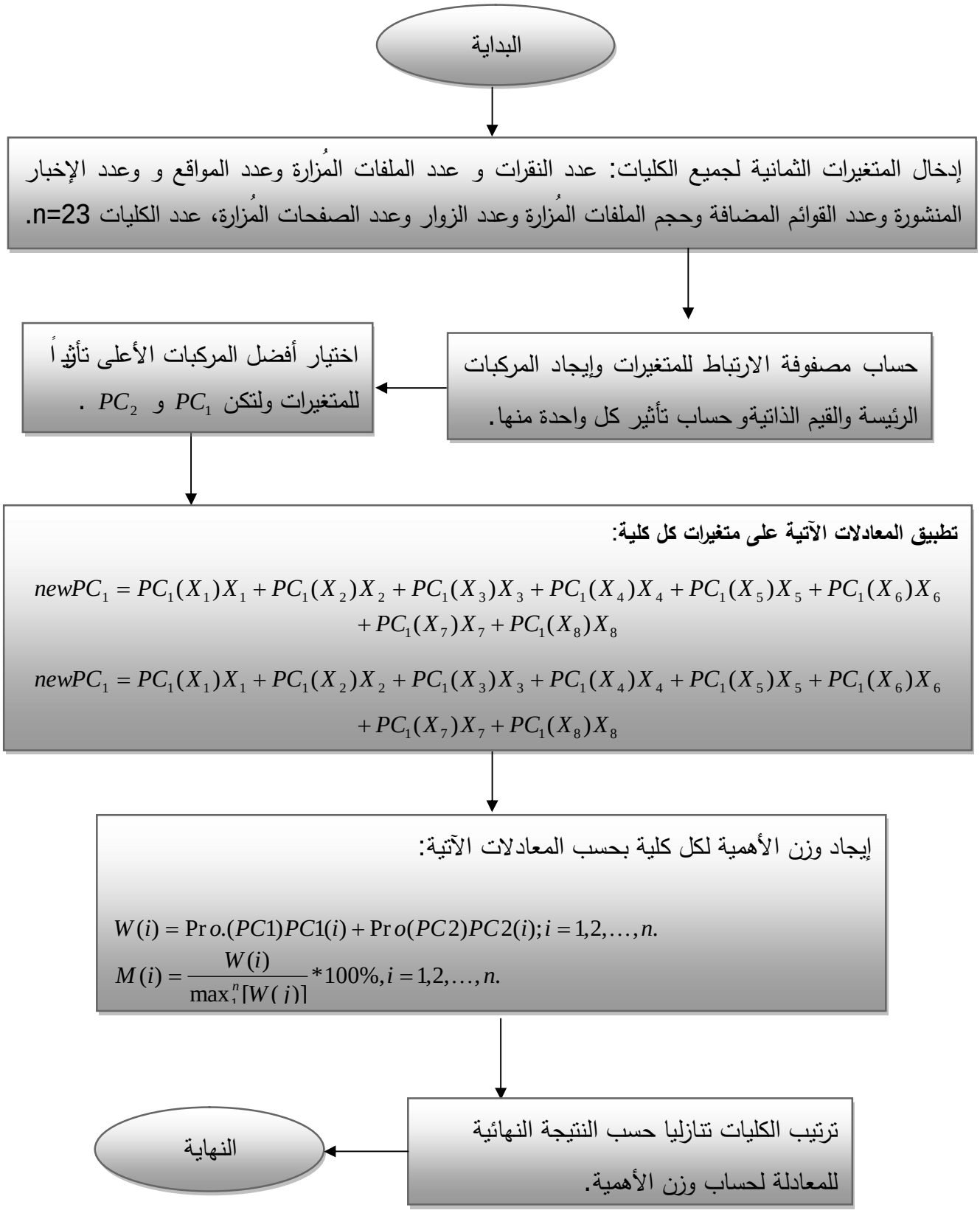
2- يلاحظ أن كلية الزراعة والغابات حظيت بتقييم متقدم في كلا التقييمين في حين أن كلية العلوم الإسلامية حظيت بتقييم متدن في كليهما.

من خلال تفحصنا لعناصر التقييمين يتضح لنا أنهما يرتكزان أساسا على متغيرين هما عدد القوائم المضافة وعدد الأخبار المنشورة في الموقع أي على فعالية أو نشاط القائمين على الموقع ونقترح إضافة نشاط الموقع من قبل المستخدمين مثل عدد المواقع (sites) التي يتم الدخول إليها من مواقع كل كلية وحجم الملفات المُزارة خلال الفترة المعينة مقاس بـ Kbytes وعدد الزوار لكل موقع من مواقع الكلية visitors وعدد الصفحات المُزارة من قبل المستخدمين وعدد الملفات التي يتم الدخول إليها من خلال المواقع الأخرى.

يتركز هدفنا في هذا البحث على قياس التطور الزمني الملاحظ في المواقع والكليات التابعة لجامعة الموصل من اجل تقييم نشاطاتها الالكترونية وذلك من خلال تفاعل المستخدمين مع هذه المواقع، ونعتقد بأن لهذا التقييم أهمية كبيرة للتواصل المعرفي والوعي الالكتروني.

4- الخوارزمية المقترحة:

المخطط الآتي يوضح الخوارزمية المقترحة لغرض قياس نشاط مواقع الإنترنت.



6- الطريقة المقترحة لتحديد أهمية مواقع الإنترنت لمواقع في جامعة الموصل:

تُقترح في هذا البحث طريقة جديدة لتحديد نشاط مواقع كليات جامعة الموصل فيما يخص نشاط المستخدمين وذلك بالاعتماد على معادلة رياضية مُشتقة من أول مركبتين رئيستين. ومن خلال هذه المعادلة يتم تصنيف المواقع في جامعة الموصل وفق البيانات المبينة في أدناه خلال الفترة من 2012/9/17 إلى غاية 2012/11/14، وهي الفترة نفسها التي اختيارها مركز الحاسوب والإنترنت في التقييم الثاني لتقييم أعمال المواقع الكترونيا، أما المتغيرات التي ستستخدم لغرض التقييم فنقترح أن تكون كما مبين في أدناه، إذ إن هذه المتغيرات المقترحة تعطي مقاييس واقعية عن فعالية المواقع، أي أن يكون هناك تصنيف إضافي إلى تصنيف مركز الحاسوب والانترنت ليشمل فعالية ونشاط المواقع من قبل القائمين على الموقع ومستخدميه:

1- عدد المواقع (sites) التي تم الدخول إليها إلى مواقع كل كلية، ويرمز له X1.
2- حجم الملفات المُزارة خلال الفترة المعينة مقاس بـ Kbytes، ويرمز له X2.
3- عدد الزوار لكل موقع من مواقع الكلية visitors خلال الفترة المعينة، ويرمز له X3.
4- عدد الصفحات المُزارة من قبل المستخدمين خلال الفترة المعينة، ويرمز له X4.
5- عدد الملفات التي تم الدخول إليها من خلال مواقع الكليات خلال الفترة المعينة، ويرمز له X5.
6- عدد الضربات (click) لكل موقع خلال الفترة المعينة، ويرمز له X6.
7- عدد القوائم المضافة لكل موقع خلال الفترة المعينة، ويرمز له X7.
8- عدد الأخبار المنشورة خلال الفترة المعينة، ويرمز له X8، [9].

إن التقييم المقترح يعتمد على بيانات موثُقة من مركز الحاسوب والإنترنت و التي سحبت من موقع "cpanel" وهو اختصاراً لـ control panel ويعني لوحة تحكم رسومية لمواقع الويب، وتعمل هذه اللوحة من على الويب ومصممة لتسهيل إدارة المواقع. إن لوحة cpanel تتحكم بجميع جوانب الموقع من خلال واجهته، وهي عبارة عن برنامج احتكاري تو عه الشركة المالكة حصراً، وقد صممته خدمات استضافة الو يب التجارية. لذا فإن الشركة لا تقدم أسعا ا مخفضة للاستعمال الشخصي، وعلى الرغم من هذا فإن أصحاب المنظمات غير الربحية، مثل المنشآت التعليمية والخيرية، بإمكانهم طلب رخصة استعمال مجانا أو بأسعار مخفضة، [Wikipedia].

لقد قام مركز الحاسوب والإنترنت مشكورا بالمساعدة بالحصول على البيانات اللازمة من لوحة التحكم cpanal استجابة لكتاب تسهيل المهمة الصادر عن كلية علوم الحاسوب والرياضيات ذي الرقم 5153 والمؤرخ في 2012/11/6.

وبالعودة إلى المتغيرات المقترحة لقياس نشاط مواقع الإنترنت، وبما أن هذه المتغيرات متفاوتة بالمديات، ومن اجل تعييرها Standardization، قمنا بتقسيم كل متغير على اكبر قيمة فيه. كما تم إجراء المعالجة الحاسوبية بواسطة كل من MINTAB 13 و MATLAB R2012a .

أنجز الجزء الأول من التحليل حسب المعادلة (2) وبواسطة MINTAB13 وحصلنا على النتائج الآتية:

Principal Component Analysis: sites; Kbytes; visits; pages; files; Hits; list; news

Eigen analysis of the Correlation Matrix

الجدول (3): القيم الذاتية والتأثير لكل مركبة والتراكمات.

Eigen value	4.2092	2.0965	0.2363	0.4330	0.2165	0.1472	0.0460	0.0154
Proportion	0.526	0.262	0.105	0.054	0.027	0.018	0.006	0.002
Cumulative	0.526	0.788	0.893	0.947	0.974	0.992	0.998	1.000

الجدول (4): المركبات الثمانية للمتغيرات المدخلة.

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
X1	0.479	0.001-	0.003	0.030	0.327	0.030	0.042	0.812
X2	0.468	0.025-	0.031	0.062-	0.454	0.328	0.468	0.493-
X3	0.445	0.134-	0.101	0.042	0.595-	0.482-	0.430	0.030-
X4	0.458	0.071-	0.007	0.076-	0.457-	0.558	0.503-	0.078-
X5	0.191	0.542-	0.343-	0.480	0.242	0.309-	0.359-	0.197-
X6	0.280	0.449	0.487	0.068-	0.226	0.431-	0.442-	0.216-
X7	0.034	0.583	0.166-	0.763	0.121-	0.150	0.111	0.010-
X8	0.166	0.376	0.779-	0.414-	0.001-	0.209-	0.072-	0.067-

يتضح من النتائج الأخيرة أن المركبة الأولى مسؤولة عن 52.6% من التباين الكلي، في حين إن المركبة الثانية مسؤولة عن 26.2% من التباين الكلي، وبذلك تكون هاتان المركبتان مسؤوليتين عن 78.8% من التباين الكلي. إن المعادلة الرياضية لكل من هاتين المركبتين هي كما مبين في أدناه:

$$PC1 = 0.479X1 + 0.468X2 + 0.445X3 + 0.458X4 + 0.191X5 + 0.280X6 + 0.034X7 + 0.166X8,$$

$$Var(PC1) = 4.2093.$$

$$PC2 = -0.001X1 - 0.025X2 - 0.134X3 - 0.071X4 - 0.542X5 + 0.449X6 + 0.583X7 + 0.376X8,$$

$$Var(PC2) = 2.0965.$$

لو عددنا الآن نسبة مسؤولية كل مركبة من التباين بوصفها وزن لأهميتها، فإننا نقترح أن يكون وزن أهمية كل موقع إنترنت يمثل المعدل الموزون لهاتين المركبتين بأوزان تعادل نسبة مسؤوليتها من التباين الكلي. أي إن المعادلة المقترحة تكون على النحو الأتي:

$$W(i) = 0.526PC1(i) + 0.262PC2(i); i = 1,2,\ldots,n. \quad (3)$$

ولغرض جعل أوزان الكليات من 100%، نقترح حساب درجة التقييم النسبية لكل موقع من المعادلة الآتية:

$$M(i) = \frac{W(i)}{\max_1^n [W(j)]} * 100\%, i = 1,2,\ldots,n. \quad (4)$$

وبذلك تكون درجات التقييم النسبية للمواقع تنازلية: $M(1) = 100 \geq M(2) \geq M(3) \geq \cdots \geq M(n)$.

7- الاستنتاجات والنتائج ومناقشتها:

لقد تم حساب أوزان مواقع الإنترنت المرتبطة بمركز الحاسوب والإنترنت بجامعة الموصل وفق المعادلة (3)، كما تم حساب درجة التقييم النسبية لكل موقع من المعادلة (4). والجدول الآتي يبين خلاصة درجات التقييم التي تم الحصول عليها لكل موقع.

الجدول(5): تقييم مواقع كليات جامعة الموصل وفق الطريقة المقترحة.

ت	الكليات	المركبة الأولى	المركبة الثانية	وزن الأهمية	درجة التقييم النسبية
1	كلية علوم الحاسوب والرياضيات	1.857	1.053	1.253	100
2	كلية طب الموصل	1.951	0.238	1.088	87
3	كلية الزراعة والغابات	1.518	0.792	1.006	80
4	كلية التربية الأساسية	2.127	-0.48	0.994	79
5	كلية الحقوق	1.594	0.329	0.925	74
6	كلية الصيدلة	1.773	-0.09	0.908	73
7	كلية العلوم	1.589	-0.19	0.786	63
8	كلية التربية	1.168	0.435	0.728	58
9	كلية الإدارة واقتصاد	1.146	0.409	0.71	57
10	كلية الآداب	1.284	0.063	0.692	55
11	كلية التربية للبنات	1.124	0.328	0.677	54
12	كلية الطب البيطري	1.232	0.077	0.668	53
13	كلية العلوم السياسية	1.255	-0.07	0.642	51
14	كلية طب الأسنان	1.06	0.234	0.619	49
15	كلية طب نينوى	1.164	-0	0.612	49
16	كلية الآثار	1.225	-0.27	0.574	46
17	كلية العلوم الإسلامية	1.2	-0.34	0.542	43
18	كلية التربية الرياضية	1.11	-0.26	0.515	41
19	كلية هندسة الالكترونيات	0.946	0.051	0.511	41
20	كلية علوم البيئة وتقاناتها	0.998	-0.21	0.469	37
21	كلية الهندسة	0.977	-0.29	0.437	35
22	كلية التمريض	0.536	0.415	0.391	31
23	كلية الفنون الجميلة	0.744	-0.15	0.352	28

ومن الجدول السابق نستنتج ما يأتي:

1- يلاحظ أن هناك تباينا بين تقييم مركز الحاسوب والإنترنت الذي يعتمد على فعالية القائمين على الموقع والتقييم المقترح للكليات والذي يعتمد على نشاط القائمين على الموقع ومستخدميه ماعدا بعض الكليات التي حافظت على مواقعها المتقدمة في التقييمين وهي: كلية علوم الحاسوب والرياضيات وكلية الزراعة والغابات وكلية الحقوق.

2- هنالك كليات كانت تقييماتها متدنية في كلا التقييمين وهي: كلية الهندسة وكلية الفنون الجميلة وكلية علوم البيئة وتقاناتها وكلية التربية الرياضية.

3- هنالك كليات كانت وفق تقييم مركز الحاسوب والإنترنت قد حصلت على مراكز متدنية في حين أنها وفق التصنيف الجديد حصلت على مراتب متقدمة وهي: كلية طب الموصل وكلية الصيدلة وكلية العلوم.

8- التوصيات:

نوصي باستخدام الطريقة المقترحة لقياس نشاط مواقع كليات جامعة الموصل فضلا عن التقييم المعتمد من قبل مركز الحاسوب والانترنت وكذلك باقي المراكز التابعة له. كما نوصي باعتماد هذه الطريقة لقياس النشاطات الالكترونية لباقي الجامعات العراقية.

9-شكر وتقدير:

يتقدم الباحثان بالشكر والثناء إلى مركز الحاسوب والإنترنت بجامعة الموصل للمعلومات القيمة التي جهزونا بها عن المواقع وكذلك البيانات.

المصادر

1-الخياط، باسل يونس ذنون (2011b) "النمذجة الماركوفية مع تطبيقات عملية- الجزء الثاني "، دار ابن الأثير للطباعة والنشر، جامعة الموصل.

2-النشواتي هشام مصطفى ،"التحليل الإحصائي متعدد المتغيرات باستخدام حزم ساس (SAS)"، جامعة الملك سعود، 2010.

3-Bishop, Christopher; "Pattern Recognition and Machine Learning", Berlin: Springer, 2006.

4-Dou Shen, Jian-Tao Sun, Qiang Yang and Zheng Chen.(2006) ''Building Bridges for Web Query Classification''.

5-Theodoridis, Sergios, Koutroumbas, Konstantinos, "Pattern Recognition", (4th ed.). Boston: Academic Press,2008.

6-Theodoridis Scrgios, Konstantions Koutroumbas , "An Introduction to Pattern Recognition AMATLAB Approach", Elsevier Inc.,2010.

7-Keinosuke F'ukunaga; " Introduction to Statistical Pattern Recognition", Second Edition, Academic Press all rights reserved,1990.

8-Mueller Jahn Paul, "Web Matrix Developer's Guide", John paul muekker, 2003.

9-Richard A. Johnson, Dean W. Wichern; "Applied Multivariate Statistical Analysis", Prentice-Hall, Inc UK,1988.

10-Sussan Dumais and Hao Chen. '' Hierarchical Classification of Web Contant''.

11-Webb R. Andrew,; "Statistical Pattern Recognition", Second Edition, QinetiQ Ltd., Malvern UK, 2002.

12-http://ar.wikipedia.org/wiki/CPanel.

13-http://ar.wikipedia.org/wiki/تاريخ_الشبكة_العنكبوتية_العالمية.

14-http://cpanel.net/.

15-Xiaoguang Qi and Brian D. Davison.(2007) '' Web Page Classification: Features and Algorithms'', Department of Computer Science & Engineering Lehigh Unversity.