Use of Common-Word Order Syntactic Similarity Metric for Evaluating Syllabus Coverage of a Question Paper

Dimple V. Paul[1], Jyoti D. Pawar[2]
[1]Department of Computer Science,
D. M's College of Science, Commerce and Arts
Mapusa, Goa, India
[2]Department of Computer Science and Technology
Goa University
Taleigao, Goa, India
dimplevp@rediffmail.com, jyotidpawar@gmail.com

ABSTRACT: *Syllabuses are used to ensure consistency between educational institutions. A modularized syllabus contains weightages assigned to different units of a subject. Different criteria like Bloom's taxonomy, learning outcomes etc., have been used for evaluating the syllabus coverage of a question paper. But we have not come across any work that focuses on syntactic text similarity evaluation of unit contents with the question contents in order to estimate the syllabus coverage of a question paper. Hence in this paper we address the problem of measuring the syllabus coverage of an examination question paper by using the order based word-to-word syntactic similarity metric. Text preprocessing techniques are used to extract multiple words and its associated locations from textual contents in the question paper and also in the respective syllabus file. Comparison of word order vectors of units with word order vectors of questions results in generation of the corresponding common word pair question vector and common word pair syllabus vector. The common word pair vectors assist in computing the similarity measure between question vector and unit vector, representing the similarity measures in a question-to-unit similarity matrix and selecting the maximal similarity measure among the set of computed common word pair vectors. The maximal similarity measures are used as a guideline in grouping the unit-wise questions, matching its weightage against Syllabus File and evaluating the syllabus coverage of the question paper. The result of syllabus coverage evaluation can be used as a guideline by the subject expert or question paper setter or question paper moderator to revise the questions of examination question paper accordingly.*

Keywords: Word Order Similarity, Syllabus Coverage, Unit Weightage, Syntactical Similarity Measure, Educational Taxonomy

Received: 2 March 2014, Revised 1 April 2014, Accepted 7 March 2014

1. Introduction

The effectiveness of an educational system basically depends upon the usefulness and success of its examination system. Examination is an organized system of evaluating and estimating the academic abilities of students and it provides feedback to the instructors to improve upon their teaching methodology [1-2]. It can also be considered as a means of measuring knowledge, skills, feelings, intelligence, or aptitude of a student or a group. Written examination is a conventional yet a universal tool to evaluate the student's performance in the educational area. The efficiency of written examination in assessing student's ability is very much dependent on the questions included in the examination question paper [3]. The set of questions for an examination question paper are generally selected by question paper setter or question paper moderator who need not necessarily be the instructor who teaches the same subject. Paper setters are not assured to have the knowledge about the importance of unit weight ages which act as a base in allocating the percentage of marks under each unit and thereby providing restriction to the selection of limited number of questions satisfying the unit weightage under each unit. Accordingly, the paper setter selected question set does not guarantee to provide proper syllabus coverage. In order to overcome this limitation, we propose an order based word-to-word syntactic similarity metric. The metric acts as a guideline in computing the similarity measure of common word pair vectors of question vector and unit vector, support in representing the similarity measures in a question-to-unit similarity matrix, assist in identifying the maximal similarity measures among the computed question-to-unit match and finally apply it for evaluating the syllabus coverage of the question paper. This paper is organized as follows. Section 2 discusses the literature review. Methodology adopted is explained in section 3. The problem statement and experimental results are given in section 4 and 5 respectively. Section6 discuss the performance evaluation and finally section 7 concludes the paper.

2. Literature Review

A syllabus file provides assistance to the instructors in knowing what must be taught and what is not required. It highlight the information about how to plan for teaching a subject, how to evaluate and monitor students' performance in the subject, and how to allocate time and resources to areas in which more learning is required. The particular structure that a syllabus includes varies greatly with the type of course that it details. A syllabus can serve students as a model of professional thinking and writing. Each unit in the syllabus file is given a weightage that correspond to the number of lecture hours to be used by the instructor to teach that unit. The weightage also indicates the importance assigned to that unit which is used by the instructor to decide on the depth to which the topics in that unit should be covered, considered by the paper setter to decide on the allocation of marks under each unit and used by the students to allocate time-schedule for each unit while preparing for an examination. Syllabus file is always acting as a benchmark while selecting the questions for the generation of examination question papers. Question papers can include either open-ended or closed-ended questions matching to the syllabus file. The easiest type of questions is closed questions or multiple-choice questions. However, multiple-choice questions cannot determine the skills of students in writing and expressing. At present, educators prefer to have essay questions to grade more realistically the students' skills. Open questions are considered to be the most appropriate, because they are the most natural and they produce a better degree of thought. They help to evaluate the understanding of ideas, the students' ability to organize material and develop reasoning, and to evaluate the originality of proper thoughts. Use of open-ended question evaluation tools is good for understanding the different cognitive skills. Several methodologies have been proposed to solve the problems in automatic evaluation of open-ended and closed-ended questions. Some of them are summarized as follows-

In [4] Chang et al. made a comparative study between the different scoring methods. They also studied the different types of exams and their effect on reducing the possibility of guessing in multiple choice questions. In multiple choice question type, the evaluation by using the set of correct answers is the traditional method. But this method does not respect the order of answers [5]. There is also the evaluation by using vector concept. It is more complicated but does not respect the order of the answer so that the solution need not have to be exactly similar to the template of the model answer [6]. Reference [7] proposed a fuzzy cognitive map to determine the concept dependencies. It applies the network graphic representation. Fuzzy concepts are used to represent domain concepts' knowledge dependencies and adaptive learning system knowledge representation. It also represents the concept's impact strength over the other related concepts.

Existing methods for computing text similarity have focused mainly on either large documents or individual words [8]. In [3] the fairness of a question paper is evaluated by measuring the relevance of its questions with the individual verbs in different cognitive levels of Bloom's Taxonomy. Bloom's taxonomy is a classification system of educational objectives based on the level of student understanding necessary for achievement or mastery. Educational researcher Benjamin Bloom and colleagues have suggested six different cognitive stages with its corresponding verbs in order to represent the intellectual activity at different stages of learning such as Knowledge, Comprehension, Application, Analysis, Synthesis and Evaluation [9]. Reference [10] implemented the prototype of a knowledge-based system that assists instructors in preparing a course syllabus which not only contains all vital components of a typical syllabus but also is prepared in accordance with the pedagogical principles especially Bloom's taxonomy in a consistent format prescribed by the institution. The relevant information made available through this system allows instructors to phrase each question for assessing associated learning outcome. The proposed system was found to be useful for instructors in developing and generating Bloom's taxonomy compliant formative and summative assessment instruments.

The Vector Space Model (VSM) is a popular information retrieval system implementation which facilitates the representation of a set of sentences as vectors in the term space [11]. Text pre-processing techniques are used for extracting the words and their associated locations from contents of question paper and syllabus file. A domain based WordNet is proposed to get incorporated in this work to handle the synonyms of the question content as well as the syllabus content. The taxonomy verbs occurring in the syllabus content as well as in the question content are identified by matching them against the Bloom's taxonomy verbs and are eliminated at the pre-processing stage. The extracted words are used for the generation of question vectors and syllabus vectors. Similarity matrix consisting of a set of question vectors and a set of syllabus vectors is a two dimensional matrix representing the pair-wise similarity of question vector with unit vector. Pair-wise similarity computation can be performed on different similarity measures. Cosine Similarity Coefficient is a common measure [12] to assign a similarity score to each pair of compared question vectors and syllabus vectors. But the major limitation of cosine which represent the terms as bags of words is that the underlying sequential information provided by the ordering of the words is typically getting lost. In [13] we performed the syllabus fairness evaluation using cosine similarity metric. The results obtained were not promising as *tf-idf* calculation could not assign higher weights to the terms that frequently occur within the syllabus content. The inclusion of *idf* in the term-weight calculation made considerable decrease in the weights of many of the frequent terms even though they were relevant. Hence we proposed to focus on computing the similarity between question content and syllabus content on the basis of word-to-word syntactic similarity metric. Sentences containing the same words but in different orders may result in very different meaning. It is easy to manually process such word order information. However the incorporation of word order into computational methods for understanding natural language is a challenging task. Word order similarity can reflect the order of a pair of the same words in two sentences. It indicates the word order regardless of the location of the word pair in an individual sentence. Word order similarity measure between two sentence vectors is calculated as a normalized difference of word order. The measure is sensitive to the distance between two words of the word pair. If the distance increases, the measure decreases. The word order similarity measure is generally 1.0 for identically ordered word vectors and is in the range of 0.0 to 1.0 for non-ordered or partially ordered vectors. Word order similarity measure remains as one of the best measure to find the similarity between sequential words in sentences and is increasing its popularity due to its simple interpretation and easy computation.

3. Methodology

The procedure for finding similarity between question content and syllabus content using word order similarity measure follows the steps as below-

- Pre-processing of Question Content and Syllabus Content
- Computing Question-Vs-Syllabus Similarity Matrix
- Displaying Syllabus Coverage of Question Paper

A brief description of the approaches used for performing each of the above steps is given below-

3.1 Pre-processing of Question Content and Syllabus Content

The five sub-steps involved in pre-processing the question content and respective syllabus content is as follows:

3.1.1 Tokenization

The set of question contents of a question paper as well as the set of unit-wise contents of syllabus are treated as sequence of words (or ordered words), which are then partitioned into a list of words with their associated locations.

3.1.2 Filtering Stop Words

Stop words are frequently occurring, insignificant words within the question content and also in the syllabus content and are eliminated.

3.1.3 Filtering Taxonomy Verbs

The taxonomy verbs within the question content as well as the syllabus contents are identified and eliminated. Details of verb and question examples that represent intellectual activity at each level of Bloom's taxonomy can be found in [3].

3.1.4 Stemming Words

Stemming is a heuristic process of cutting off the ends of words of question content as well as syllabus content for getting the correct root form of the word. There are various word stemmers available for English text and the most commonly used Porter stemmer is considered.

3.1.5 Normalization

The idea behind normalization is to convert all words which mean the same, but written in different forms (e.g. CPU and C.P.U) into the same form. We are using the following techniques for performing word normalization-

- Lowercase the words
- Remove special characters

3.2 Computing Question-Vs-Syllabus Similarity Matrix

The similarity matrix computation is carried out by considering the matrix representation of vectors which is a natural extension of the existing VSM. Matrix representation considers the questions of a question paper as the row headers and units of the syllabus file as the column headers of the matrix. Each question is represented as a vector of question-words and each unit of the syllabus file is considered as a vector of syllabus-words. In the multidimensional matrix of N questions and M units say $N \times M$ matrix, each pair of question-word vector and syllabus-word vector gets compared to determine how identical they are by using word order similarity measure. The word order similarity between question vector and syllabus vector say $QV1$ and $SV1$ is computed in a two step process. In the first step, the common words in $QV1$ and $SV1$ are identified and are inserted in the same order in which they appear into two other common word vectors say $QCWV1$ and $SCWV1$. In the second step, a unique index number that represent the order in which the words appear in $QCWV1$, is identified for each word in $QCWV1$ and is inserted to the corresponding index vector say $QINV1$. Based on these unique index numbers in $QCWV1$, we also assign respective unique index numbers to words in $SWV1$ and insert it to the corresponding index vector say $SINV1$. The word order similarity of question-vector and syllabus-vector is calculated by using the following formula

$$\text{Similarity}\,(QV1, SV1) = (1 - (norm\,(QV1, SV1)) \tag{1}$$

where, norm ($QV1$, $SV1$) which normalizes the difference of common word order pairs in $QV1$ and $SV1$ can be expanded in the following manner.

$$\text{norm}\,(QV1, SV1) = \frac{\sum_{i=1}^{n} |\,(QINV1\,[i] - SINV1[i]\,)\,|}{\sum_{i=1}^{n} |\,(QINV1\,[i] + SINV1[i]\,)\,|} \tag{2}$$

Sample of a similarity matrix with computed pair-wise similarity say smx, y for n questions and m units is represented in Table 1 below. The computation of similarity of n questions with m units is carried out by calculating the similarity of $n \times m$ pairs of question vectors and syllabus vectors.

	***SV*1**	***SV*2**	***SV*3**	***SV*4**	***SV*5**	**…**	***SV*m**
***QV*1**	sm_{11}	sm_{12}	sm_{13}	sm_{14}	sm_{15}	…	sm_{1m}
***QV*2**	sm_{21}	sm_{22}	sm_{23}	sm_{24}	sm_{25}	…	sm_{2m}
***QV*3**	sm_{31}	sm_{32}	sm_{33}	sm_{34}	sm_{35}	…	sm_{3m}
***QV*4**	sm_{41}	sm_{42}	sm_{43}	sm_{44}	sm_{45}		sm_{4m}
***QV*5**	sm_{51}	sm_{52}	sm_{53}	sm_{54}	sm_{55}	…	sm_{5m}
…	…	…	…	…		…	…
$\boldsymbol{QV_n}$							sm_{mm}

Table 1. Similarity Matrix Representation

3.3 Displaying Syllabus Coverage of Question Paper

The computed similarity matrix is used to generate unit-wise question groups, calculate unit-wise question groups' weightage and compare the calculated weightage against the actual unit-weightage in the syllabus file. The syllabus coverage evaluation carried out using the similarity matrix is considered as a good measure by the question paper setter or question paper moderator to revise the questions of a question paper. Choosing suitable threshold value for similarity computation is a difficult task and it is problem dependent. We have considered 0.50 as the threshold value for better recall of questions while generating question groups that satisfy the paper setter specified requirements. Precision, Recall and F-measure are commonly used as the metrics to evaluate the accuracy of predictions and the coverage of accurate pairs of comparisons in the question to syllabus matching system. They are computed as

$$\text{Precision } (P) = \frac{\text{Number of Relevant } QV1 \text{ to } SV1 \text{ Matches Retrieved by Tool}}{\text{Total Number of } QV1 \text{ to } SV1 \text{ Matches Retrieved by Tool}} \quad (3)$$

$$\text{Recall } (R) = \frac{\text{Number of Relevant } QV1 \text{ to } SV1 \text{ Matches Retrieved by Tool}}{\text{Total number of } QV1 \text{ to } SV1 \text{ Matches Given by Paper Setter}} \quad (4)$$

$$\text{F-measure} = \frac{(2 \times P \times R)}{P + R} \quad (5)$$

4. Problem Formulation

4.1 Problem Statement

Given a question paper of subject S consisting of N questions represented as $QP\ (S) = \{qst_1, qst_2, \ldots, qst_N\}$ and a syllabus file of S consisting of M units represented as $SF\ (S) = \{unt_1, unt_2, \ldots, unt_M\}$, the problem is to find unit-wise similar question groups UQG_1, UQG_1 ,.....,UQG_k. A question qst_i can be said to belong to unt_j if *similarity* $(qst_i, unt_j) >= \partial$ "*where*" ∂ is the user input threshold value to find the similarity.

The similarity (qst_i , unt_j) function could use any of the similarity measures available. We have used word order similarity to perform the experimental study. The main modules of this algorithm are shown in Figure1.

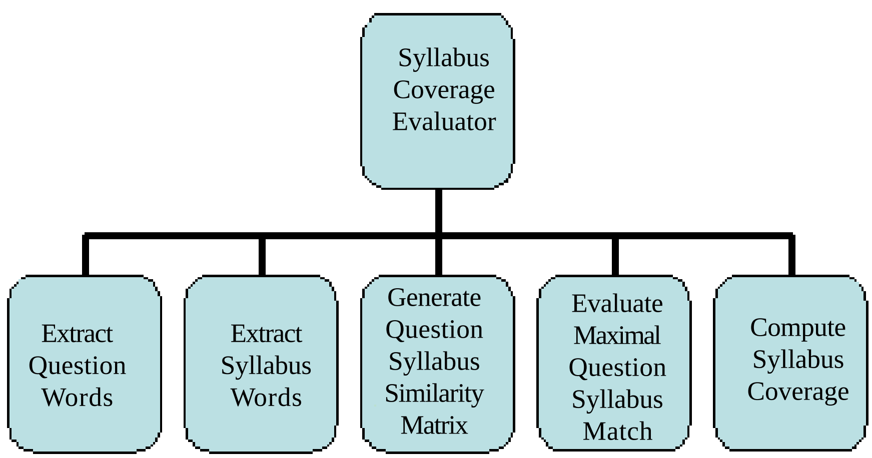


Figure 1. Main modules of Syllabus-Fairness Evaluator

The brief details of modules are presented below–

4.1.1 Extract-Question-Words

Input qst_i (i = 1 to N) and for each qst_i it extracts words with their associated locations, $qt_{ij}[j]$ (j =1 to N_i)

4.1.2 Extract-Syllabus-Terms

Input unt_j (j =1 to M) and for each unt_i in the syllabus file, it extracts words with their associated locations, $st_{jk}[k]$ (k =1 to M_j).

4.1.3 Generate-Question-Syllabus-Similarity-Matrix

Input question-words qt_{ij} (j =1 to N_i) for all qst_i (i =1 to N) and also syllabus-words st_{jk} (k =1 to M_j) for all unt_j (j =1 to M).

subject_quest_cd	questionpaper_cd	subject_quest_descr
qp7quest1	qp7	Explain fake websites.how do we recognize them
qp7quest10	qp7	browser security
qp7quest11	qp7	regular update virus definition
qp7quest12	qp7	meaning of file share and piracy issues
qp7quest13	qp7	multimedia elements and any four with example
qp7quest14	qp7	security issues in mcommerce
qp7quest15	qp7	internet live stream and bitrate in multimedia
qp7quest16	qp7	credit card payment transactions with its steps
qp7quest17	qp7	meaning of ecommerce technology and a note on two of them
qp7quest18	qp7	multimedia appilcations
qp7quest19	qp7	shopping cart and its advanatge in etransactions
qp7quest2	qp7	use of codec in multimedia .
qp7quest20	qp7	lossy and lossless compression in multimedia
qp7quest21	qp7	website.explain static and dynamic website
qp7quest22	qp7	different email protocol used for sending and receiving
qp7quest23	qp7	url. the relation of url and domain names
qp7quest24	qp7	usage of web server and web client.
qp7quest25	qp7	pki cryptography in web security
qp7quest26	qp7	multiuser chat and social network
qp7quest27	qp7	three virus types
qp7quest28	qp7	the role of user generated content known as moodle and blogs
qp7quest3	qp7	ebanking and phishing
qp7quest4	qp7	http and https protocols used in internet
qp7quest5	qp7	information distribution using pdf
qp7quest6	qp7	multicast in multimedia
qp7quest7	qp7	virus attack prevention methods

Record: 119 of 174 No Filter Search

Figure 2. Sample dataset of IT Question paper with 28 questions

For each pair of question-words qst_i and syllabus-words unt_j compute similarity (qst_i unt_j) for i =1 to N and j =1 to M using any standard similarity measuring scheme. Represent the result as a Question-Syllabus-Similarity-Matrix.

4.1.4 Evaluate-Best-Question-Syllabus-Match

For each question in the Question-Syllabus-Similarity-Matrix, it finds the highest value of similarity among the set of computed similarity (qst_i , unt_j) >= ∂, for i = to N and j =1 to M. If the highest value of similarity does not get identified for a question, then the question is considered to be indirectly associated with the syllabus and is represented as indirect question else the question is considered to be directly associated with the syllabus and is represented as direct question.

4.1.5 Compute-Syllabus-Coverage

Under each unit, it performs the summation of the marks of direct questions and represents the result of summation as unit-direct-question-mark. Also it identifies whether the unit-direct-question-mark of each unit *satisfies-the-unit-weightage* of the syllabus file. The term *satisfies-the-unit-weightage* means that the unit-direct-question-mark is less-than-or-equal-to (<=) the unit-weightage. If the unit-direct-question-mark is greater-than-or-equal-to (>=) the unit-weightage, then the value of unit-direct-question-mark gets replaced with the value of unit-weightage. This replacement process is carried out to limit the value of unit-direct-question-mark to the extent to which it matches with the unit-weightage. At the next stage, it adds up the unit-direct-question-mark of all the units and represents the result of addition as direct-question-weightage. Using the direct-question-weightage, it computes the Syllabus Coverage of the question paper. The computed coverage is represented as Poor or Average or Good or Excellent depending upon whether the percentage of direct-question-weightage falls in the range of 0–40 or 41–60 or 61–80 or 81–100 respectively.

4.2 Algorithm for Syllabus-Coverage-Evaluator

4.2.1 Algorithm for Question-Word-Extraction

subject_unit	subject_cd	unit_content
	sub5	authoring tool,definition,internet live stream,multimedia elements,multimedia applications and products,image types,audio types,video types, an tool,creation and conversion,codecs,compression,transmission,bitrates,multicast,hi definition
sub5unit2	sub5	ecommerce technology,introduction of ecommerce,ecommerce today,ecommerce enablers,epayment systems,ecash, credit card payment, debit ca gateway, example, fake websites,digital certification spoof,eretail,ecarts,shopping cart,eshopping examples,ebay,amazon,music istore, ebanking,online banks,phishing, security issues in mcommerce, definition of mcommerce,mcommerce usage
sub5unit3	sub5	static website,dynamic website,web server, web security,web client,url,domain name,domain name definition,domain name,domain name format protocol,email,definition,email usage, email protocol,email clients,security,spam, phishing,information distribution,pdf,news feed,news readers
sub5unit4	sub5	virus types,define virus,malware,malware, symptoms of attack,virus attack,virus attack prevention,antivirus and spyware detectors,define antivirus spyware detection software, antivirus software installation, regular update virus definition,browser security,virus attachments,web security,securit identification,hacking types,hacking, effects, examples,pki cryptography, pki definition, digital signature, digital certificates
sub5unit5	sub5	,piracy issues,user generated content,blogs,wikis,twitter,youtube, flickr, moodle,collaboration,social networks, multi user chat, application exampl share, p2p, torrents, protocols, examples

Figure 3. Sample dataset of IT Syllabus File

// Extract Question-Words and its associated locations by stop-word removal, taxonomy verb removal and stemming
$QT = \{ \}$
For $i = 1$ to N
Extract words from qst_i and store it in array qt_i[ ]
Remove the stop-words from qt_i[ ]
Remove taxonomy verbs from qt_i[ ]
Extract the stem of each word in qt_i[ ]
$QT = QT \cup qt_i$ [ ]
End For
Output of Question-Word-Extraction, $QT = \{qt_1[\,], qt_2[\,], qt_3[\,], \ldots, qt_N[\,]\}$

4.2.2 Algorithm for Syllabus-Term-Extraction
// Extract Syllabus-Words and its associated locations by stop-word removal, taxonomy verb removal and stemming
$ST = \{ \}$
For $j = 1$ to M
Extract words from unt_j and store it in array st_j[ ]
Remove the stop-words from st_j [ ]
Extract the stem of each term in st_j [ ]
$ST = ST \cup st_j$[ ]
End For
Output of Syllabus-Word-Extraction, $ST = \{st_1[\,], st_2[\,], st_3[\,], \ldots, st_M[\,]\}$

4.2.3 Algorithm for Syllabus-Coverage-Evaluation
//Evaluate-Syllabus-Coverage
Input: $QT = \{qt_1[\,], qt_2[\,], qt_3[\,], \ldots, qt_N[\,]\}$ // set of questions in the question paper

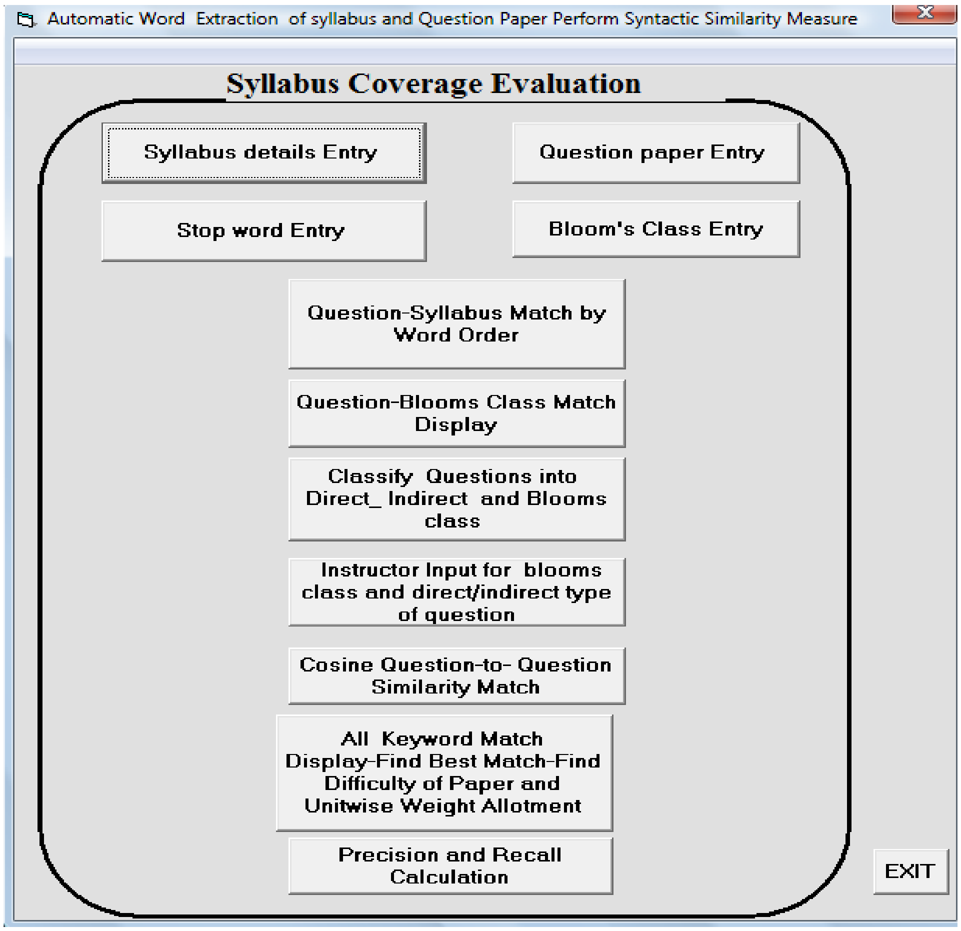


Figure 4. Screenshot with Paper Setter Input for Syllabus Coverage Evaluation

$ST = \{st_1[\ ], st_2[\ ], st_3[\ ],\ldots, st_M[\ ]\}$ // set of units in the syllabus file

where

$qt_i = \{qt_{i1}, qt_{i2}, qt_{i3},\ldots, qt_{ip}\}$ for $p = 1$ *to count* (qst_i *terms*) //set of terms in question i

$st_j = \{st_{j1}, st_{j2}, st_{j3},\ldots, st_{jq}\}$ for $q = 1$ to count (unt_j terms) //set of terms in syllabus j

$N = \{qst_1, qst_2,\ldots, qst_N\}$ // number of selected questions

$M = \{unt_1, unt_2,\ldots, unt_M\}$ // number of selected units

Threshold = ∂ // threshold value for similarity computation

Output: *k Unit_Question_Group* $UQG_1, UQG_2,\ldots, UQG_k$ where UQG_k consist of a set of qst_i' questions of *QP* (*S*)
// Form unit-wise question groups and verify its syllabus coverage

Begin
// Initialization
cnt = 0 //counter for number of question-groups
direct_question_percentage = 0 //count percentage of direct questions
unit_question_set = [ ]
// Unitwise-Question-Group-Formulation
//Compare the unit-wise marks of questions of QP(S) with the corresponding unit-wise weightages in the syllabus
For *i* =1 to *M*
    cnt = *cnt* + 1
// Formulate unit-wise new question groups

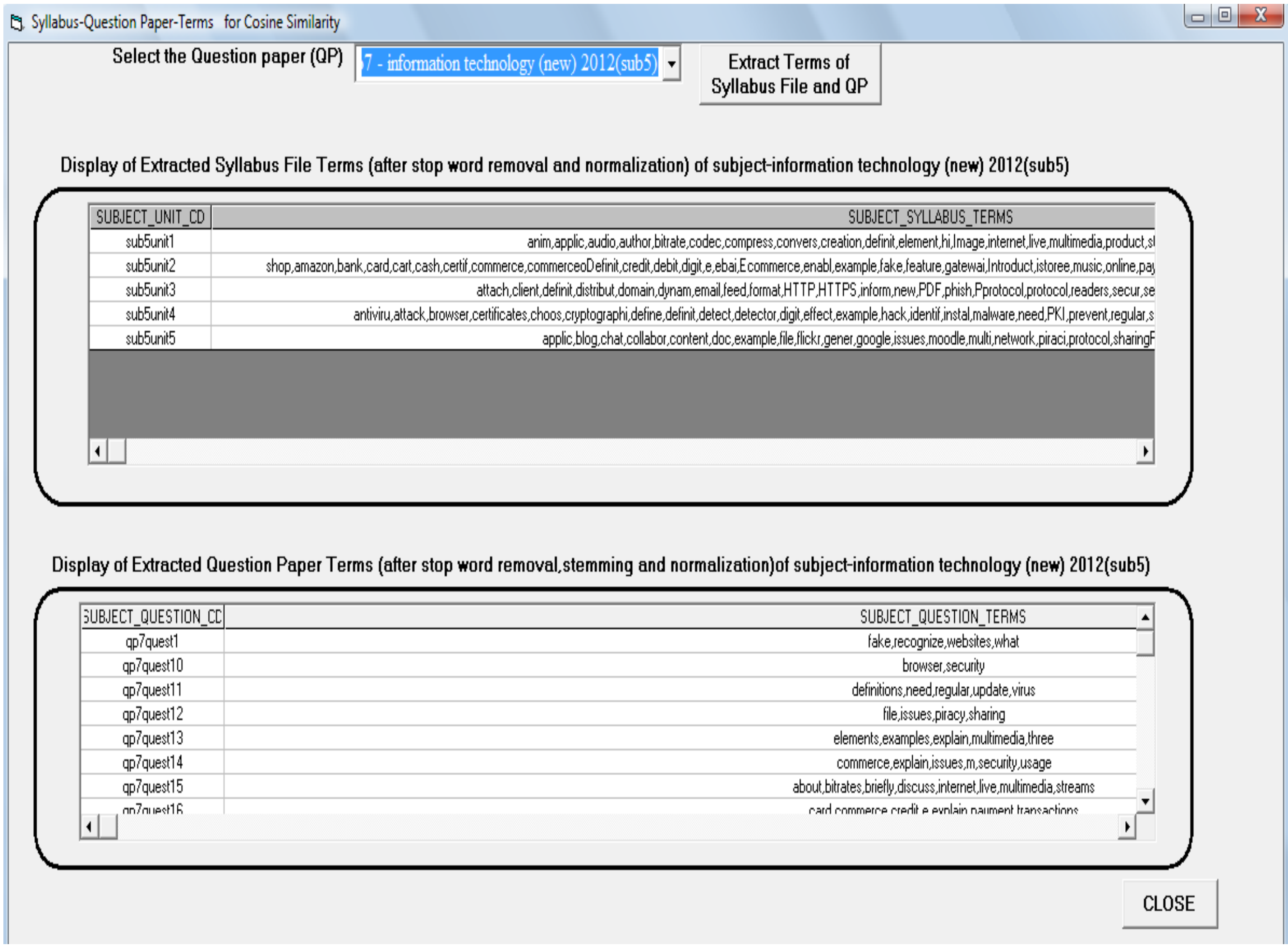


Figure 5. Extracted list of Terms of Syllabus File and Question Paper

UQG_i = New_Unit_Question_Group (unt_i , cnt)
For j =1 to N
If qst_j not in unit_question_set then

// Maximal-Question-Syllabus-Match using Similarity-Matrix
If similarity (qst_j , unt_i) >= ∂ *then*
temp = similarity (qst_j , unt_i)
For k =1 to M
If similarity (qst_j ,unt_k) > temp then
Exit for
End if

// Appending questions to each question-groupsiteratively
Add qst_j to New_Unit_Question_Group p
unit_question_set = unit_question_set + qst_j
End For
End If
End If
End For

End For

// Evaluate Syllabus-Coverage using marks of unit-wise questions in question-groups
For i = 1 to *cnt*
*A*ccept Unit_Question_Group (unt_i , i)
$marks\_unt_i$ = *sum* (*marks of all questions of Unit_Question_Group* (unt_i ,i))
If $marks\_unt_i$ <= *syllabus-weight*(unt_i) *then*
direct_question_perc = *direct_question_perc* + $marks\_unt_i$
Else
direct_question_perc = *direct_question_perc* + *syllabus-weight* (unt_i)
End If
End For
End

5. Implementation Details

5.1 Hardware and Software Platform Used

Implementation is done using Microsoft Visual Basic .NET as Front End Tool and SQL Server as Back End Tool on a 2GHz processor with 1GB RAM.

5.1.1 Datasets used

The question paper of the first year of three year bachelor's degree course of computer science (B.Sc Computer Science) for Information technology subject examination at Goa University contains 28 questions. The syllabus file for this subject includes 5 units. Details of experimental data used for similarity computation is as follows-

a) S = sub5 = Information technology (IT)

b) QP (S) = qp7 = {$quest$1, $quest$2,…, $quest$10,.., $quest$28}

c) SF (S) = {$unit$1, $unit$2,.., $unit$9}

d) ∂ = 0.75

e) Sample Dataset of IT Question Paper and IT syllabus File is displayed in Figure 2 and Figure 3 respectively.

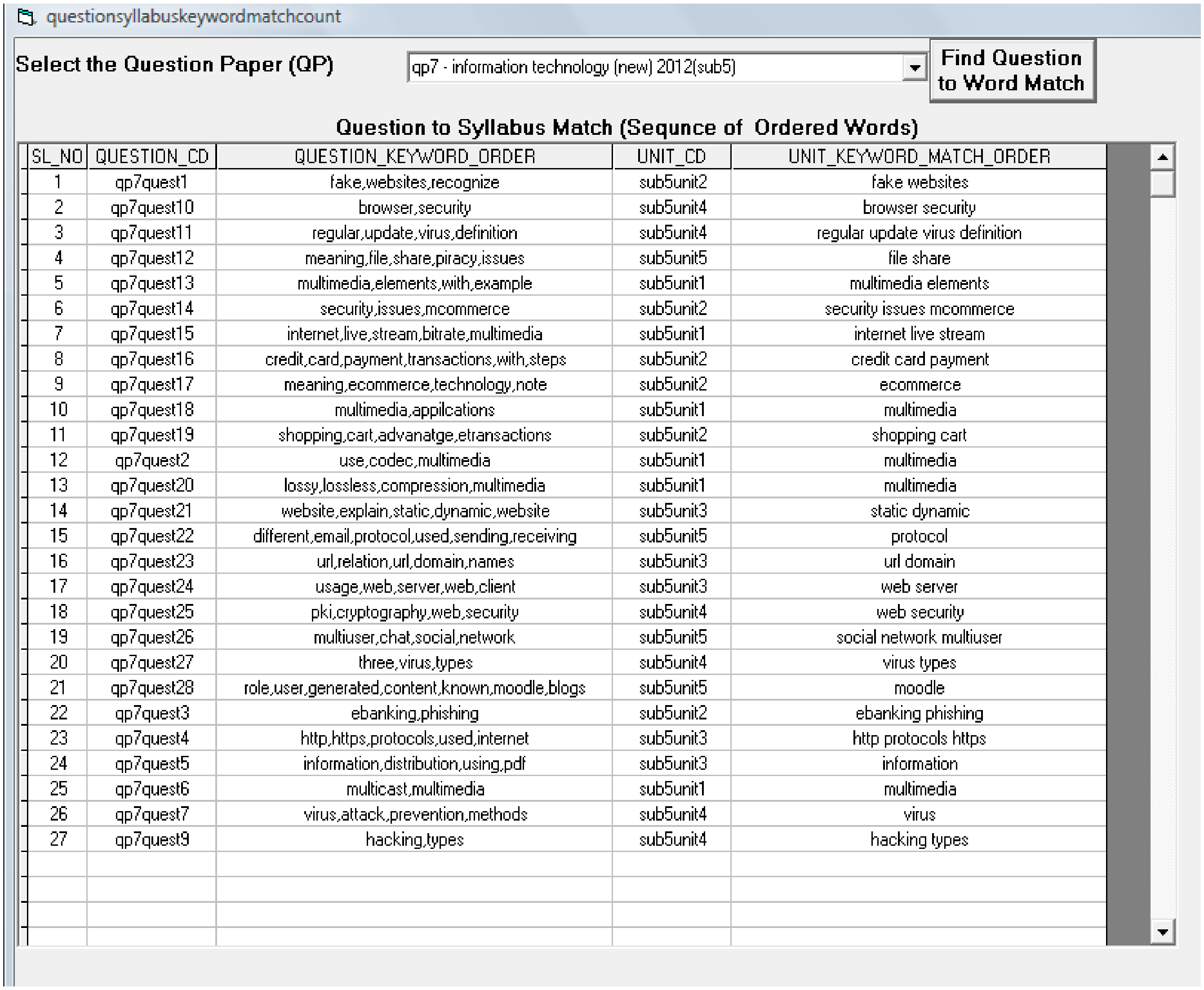
f) Figure 4 presents the screenshot displaying paper setter input for syllabus Coverage Evaluation. It provides the facility for entering the general details such as details of Syllabus File, Question Paper details, Stop Words and Bloom's Class details.

g) A snapshot of the set of units {sub5unit1, sub5unit2} with its extracted list of terms anim, applic, audio,…},{shop, amazon, bank, card,…},….,{attach, client, definit, distribut,…} etc., and set of questions {qp7quest1, qp7quest10, qp7quest11…, qp7quest28} with its extracted list of terms {fake, website, what}, {browser, security}, {definition, need, regular, update, virus}, {file, issues, piracy, sharing} etc., for qp7 is displayed in Figure 5. Extraction of terms from IT subject's syllabus file and IT subject's question paper were carried out by performing four different pre-processing stages such as Tokenization, Stop Word Removal, Filtering Taxonomy Verbs and Normalization of Words.

5.2 Results Obtained

Figure 6 below shows a sample screen shot of the process of generation of the common word pair vectors by matching each question vector with different unit vectors. Each of the questions among the 28 questions in the QP of IT was compared against 5 units in the IT syllabus file. IT-Word-Order-Question-Syllabus-Match generated 28 × 5 combination of values, and among them only the common word pair vectors were identified. The maximal similarity measures among the common word pair vectors were computed for each question in such a way that the question was finally matched to the corresponding unit to which it has the maximal similarity measure.

Figure 7 below shows a sample screen shot of the similarity matrix computation using word order syntactic similarity measure. Each of the questions among the 28 questions in the QP of IT was compared against the 5 units in the IT syllabus file. IT-Word-Order-Similarity-Matrix generated 28 × 5 combination of values using word order similarity. If no similarity exists between a pair of IT question-terms and IT unit-terms, word order similarity returned a value of zero and in every other case; it returned a value in the range of 0.0-1.0.



SL_NO	QUESTION_CD	QUESTION_KEYWORD_ORDER	UNIT_CD	UNIT_KEYWORD_MATCH_ORDER
1	qp7quest1	fake,websites,recognize	sub5unit2	fake websites
2	qp7quest10	browser,security	sub5unit4	browser security
3	qp7quest11	regular,update,virus,definition	sub5unit4	regular update virus definition
4	qp7quest12	meaning,file,share,piracy,issues	sub5unit5	file share
5	qp7quest13	multimedia,elements,with,example	sub5unit1	multimedia elements
6	qp7quest14	security,issues,mcommerce	sub5unit2	security issues mcommerce
7	qp7quest15	internet,live,stream,bitrate,multimedia	sub5unit1	internet live stream
8	qp7quest16	credit,card,payment,transactions,with,steps	sub5unit2	credit card payment
9	qp7quest17	meaning,ecommerce,technology,note	sub5unit2	ecommerce
10	qp7quest18	multimedia,appilcations	sub5unit1	multimedia
11	qp7quest19	shopping,cart,advanatge,etransactions	sub5unit2	shopping cart
12	qp7quest2	use,codec,multimedia	sub5unit1	multimedia
13	qp7quest20	lossy,lossless,compression,multimedia	sub5unit1	multimedia
14	qp7quest21	website,explain,static,dynamic,website	sub5unit3	static dynamic
15	qp7quest22	different,email,protocol,used,sending,receiving	sub5unit5	protocol
16	qp7quest23	url,relation,url,domain,names	sub5unit3	url domain
17	qp7quest24	usage,web,server,web,client	sub5unit3	web server
18	qp7quest25	pki,cryptography,web,security	sub5unit4	web security
19	qp7quest26	multiuser,chat,social,network	sub5unit5	social network multiuser
20	qp7quest27	three,virus,types	sub5unit4	virus types
21	qp7quest28	role,user,generated,content,known,moodle,blogs	sub5unit5	moodle
22	qp7quest3	ebanking,phishing	sub5unit2	ebanking phishing
23	qp7quest4	http,https,protocols,used,internet	sub5unit3	http protocols https
24	qp7quest5	information,distribution,using,pdf	sub5unit3	information
25	qp7quest6	multicast,multimedia	sub5unit1	multimedia
26	qp7quest7	virus,attack,prevention,methods	sub5unit4	virus
27	qp7quest9	hacking,types	sub5unit4	hacking types

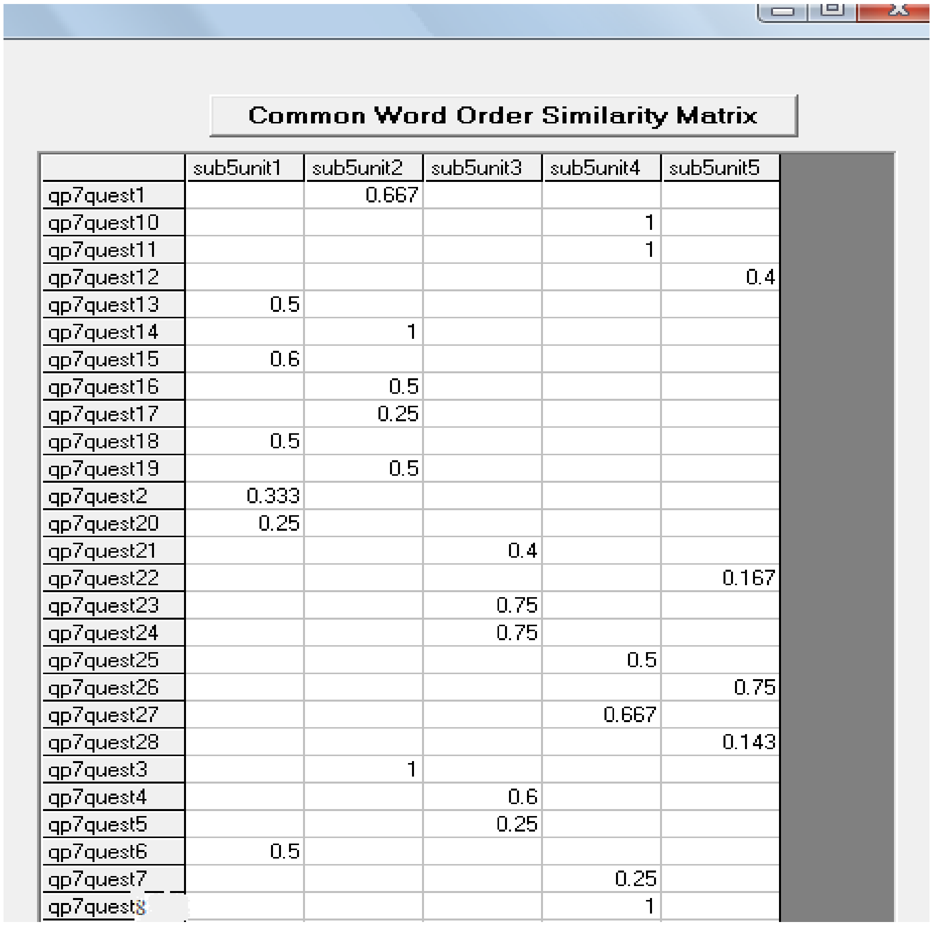
Figure 6. Computed Maximal Similarity Measure for Question to Syllabus Match

Figure 8 below represents the process of generation of IT-unit-question-groups. For each question in the IT-Common-Word-Order-Similarity-Matrix, the highest value of similarity among a set of pair-wise *similarity* $(IT - quest_i, IT - unit_j) >= 75$, for i = 1 to 28 and j =1 to 5 were found. When the highest value of similarity could get computed for an *IT-quest*, the question was termed as direct; else the question was represented as indirect.

Figure 8 below shows the evaluated measure of syllabus coverage. Under each *IT-unit*, marks of direct questions were added up and were named as IT-unit-direct-question-mark. Whether or not the IT-unit-direct-question-mark of each unit could satisfy the IT-unit-weightage of IT-syllabus file was identified and the IT-unit-direct-question-mark is updated accordingly. Summation of

Question Paper Code	Precision	Recall	F-measure
IT 2012	0.80	0.83	0.81
IT 2010	0.78	0.81	0.79
IT 2009	0.77	0.82	0.79
IT 2008	0.75	0.81	0.77

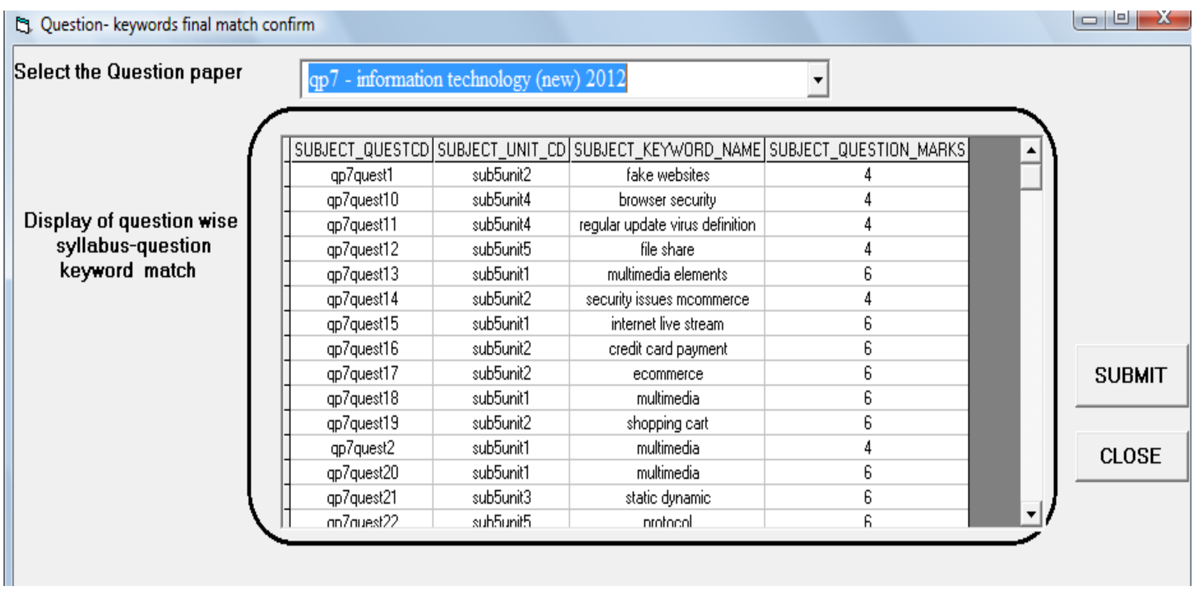
Table 3. Performance Evaluation of It_Syllabus_Coverage



Common Word Order Similarity Matrix

	sub5unit1	sub5unit2	sub5unit3	sub5unit4	sub5unit5
qp7quest1		0.667			
qp7quest10				1	
qp7quest11				1	
qp7quest12					0.4
qp7quest13	0.5				
qp7quest14		1			
qp7quest15	0.6				
qp7quest16		0.5			
qp7quest17		0.25			
qp7quest18	0.5				
qp7quest19		0.5			
qp7quest2	0.333				
qp7quest20	0.25				
qp7quest21			0.4		
qp7quest22					0.167
qp7quest23			0.75		
qp7quest24			0.75		
qp7quest25				0.5	
qp7quest26					0.75
qp7quest27				0.667	
qp7quest28					0.143
qp7quest3		1			
qp7quest4			0.6		
qp7quest5			0.25		
qp7quest6	0.5				
qp7quest7				0.25	
qp7quest8				1	

Figure 7. Common Word Order Similarity Matrix (Question to Syllabus Match)



SUBJECT_QUESTCD	SUBJECT_UNIT_CD	SUBJECT_KEYWORD_NAME	SUBJECT_QUESTION_MARKS
qp7quest1	sub5unit2	fake websites	4
qp7quest10	sub5unit4	browser security	4
qp7quest11	sub5unit4	regular update virus definition	4
qp7quest12	sub5unit5	file share	4
qp7quest13	sub5unit1	multimedia elements	6
qp7quest14	sub5unit2	security issues mcommerce	4
qp7quest15	sub5unit1	internet live stream	6
qp7quest16	sub5unit2	credit card payment	6
qp7quest17	sub5unit2	ecommerce	6
qp7quest18	sub5unit1	multimedia	6
qp7quest19	sub5unit2	shopping cart	6
qp7quest2	sub5unit1	multimedia	4
qp7quest20	sub5unit1	multimedia	6
qp7quest21	sub5unit3	static dynamic	6
qp7quest22	sub5unit5	protocol	6

Figure 8. Iterative stages of unit-question-groups formation

the IT-unit-direct-question-mark of all the units was carried out to generate SE-direct-question-weightage. Using IT-direct-question-weightage, IT-Syllabus-Coverage was computed and was represented as "*Good*", as the percentage of SE-direct-question-weightage was in the range of 61-80. Performance analysis of the results indicates that word-order similarity is a good

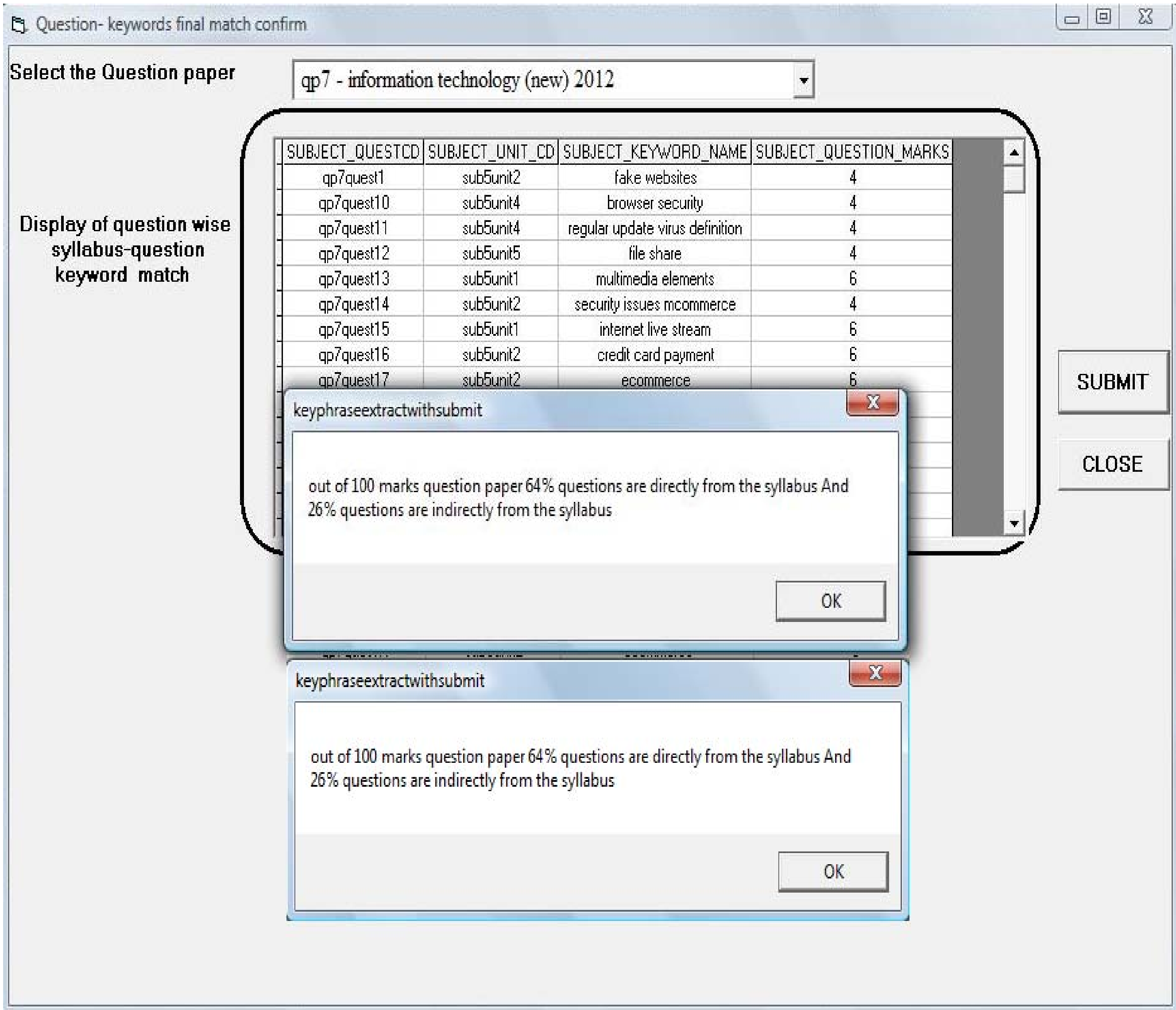


Figure 9. Computed syllabus-Coverage Measure using unit-question-groups

measure in grouping similar questions.

6. Performance Evaluation

In order to carry out the performance evaluation of the word-order- similarity based IT-syllabus-Coverage measure; we computed Precision, Recall and F-measure values for the IT question-syllabus match. We have taken into consideration the IT question papers of four different years. Each question paper consist roughly about 30 questions which were getting matched with the corresponding IT syllabus content. The result of computation is shown in Table 3 below.

7. Conclusion and Future Work

This paper focused on a new approach for syllabus coverage evaluation of a question paper by performing ($n \times m$) pair-wise question-vector and syllabus-vector comparisons. Similarity matrix computation was carried out using common word order similarity which is a commonly used similarity measure for short documents. Even though word order similarity has a limitation of identifying the word locations, this is not a major concern in our work as we deal with short text documents. Results obtained indicate that word order similarity is a good measure in formulating unit-wise question groups. The formulated question groups has been found successful in identifying the syllabus coverage of a question paper by comparing the generated unit-wise question groups' weightages against the actual unit-weightage specified in the syllabus file. The generated question groups are useful in situations where novice instructors or the question paper setter or question paper moderator needs to evaluate the syllabus coverage of a question paper and revise the questions of examination question paper accordingly. The primary objective of this study was to identify the effectiveness of statistical measures in formulating similar question groups and evaluating the coverage of a question paper. Our future work will focus on replacing the statistical approaches of similarity matrix generation by latent semantic approaches.

References

[1] Chang, S-H., Lin, P-C., Lin, Z-C. (2007). Measures of partial knowledge and unexpected responses in multiple-choice tests, International Forum of Educational Technology & Society (IFETS), p. 95–109.

[2] Jinping Liu., Dongxu Wu. (2009). The Use of Data Mining in Contents of the Examination and the Relevance of Teaching Research, First International Workshop on Education Technology and Computer Science, ETCS 2009 Wuhan, Hubei, p. 745 – 748.

[3] Karl O. Jones., Janice Harland. (2009). Relationship between Examination Questions and Bloom's Taxonomy, 39th ASEE/IEEE Frontiers in Education Conference, San Antonio, October 18- 21.

[4] Chang, S-H., Lin, P-C., Lin, Z-C. (2007). Measures of Partial Knowledge and Unexpected Responses in Multiple-Choice Tests, International Forum of Educational Technology & Society (IFETS), p. 95–109.

[5] Fasuga. R. (2010). Automatic Evaluation of Created Tasks in e-learning Environment, 10th International Conference on Intelligent Systems Design and Applications (ISDA).

[6] Wong, K. M., Raghavan, V. V. (1984). Vector Space Model of Information Retrieval: A Reevaluation, *In*: Proceedings of the 7th annual international ACM SIGIR conference on Research and development in information retrieval, SIGIR, p. 167-185, Swinton, UK, British Computer Society.

[7] Konstantina Chrysafiadi., Maria Virvou. (2013). A Knowledge Representation Approach using Fuzzy Cognitive Maps for Better Navigation Support in an Adaptive Learning System, SpringerPlus, *A Springer pen Journal, Springer Plus*, 2, 81.

[8] Yutaka Matsuo., Mitsuru Ishizuka. (2004). Keyword Extraction from a Single Document using Word Co-occurrence Statistical Information, *Journal on Artificial Intelligence Tools*, 13 (1) 157-169 © World Scientific Publishing Company.

[9] Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview, *Theory into Practice*, 41 (4) 212-219.

[10] Paice Chris, D. (1994). An Evaluation Method for Stemming Algorithms, *In*: Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, p. 42-50.

[11] Naseer Ahmed, Anwar M. A., Abdurahem Mohammed Al Ameen (2012).Course Syllabus and Assessment Instrument Development System: An Integrated Approach, *International Journal for Cross-Disciplinary Subjects in Education* (IJCDSE), 3 (3).

[12] Aizawa, A. (2000).The feature Quantity: An Information-theoretic Perspective of tfidf-like Measures, *In*: Proceedings of the 23rd ACM SIGIR Conference on Research and Development in Information Retrieval, p. 104–111.

[13] Dimple V. Paul, Jyoti D. Pawar. (2014). A Syllabus-Fairness Measure for Evaluating Open-Ended Questions, Advances in Signal Processing and Intelligent Recognition Systems, Springer International Publishing Switzerland, 264, p. 57-71.