A Personalised Group-Based Recommendation Approach for Web Search in E-Learning

Mohammad Mustaneer Rahman, and Nor Aniza Abdullah, Member, IEEE
Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur 50603 Malaysia.

Corresponding author: Nor Aniza Abdullah (e-mail: noraniza@um.edu.my)

This work is fully supported by the University of Malaya Research Grant (RP032D-16SBS) and the Postgraduate Research Grant (PG276-2015B).

ABSTRACT The unprecedented growth of the Internet, its pervasive accessibility and ease of use has increased students’ dependencies on the Web for quick search and retrieval of learning resources. However, current search engines tend to rely on the correct keywords. This excludes other characteristics such as the individual’s learning capability and readiness for specific learning materials. As a result, the same set of search-keywords deliver the same search results. This situation hinders the optimisation of the Web search engines in supporting the heterogeneity of its users in their learning endeavours. This paper aims to address the issue. It attempts to augment Web search engines with personalised recommendations of search results which match students’ learning competencies and behaviours. The results drawn from our experiments suggest that our novel approach can provide a notable improvement in terms of performance and satisfaction for the students.

INDEX TERMS E-learning, group-based recommendation, personalised web search, recommender system, students profiling.

I. INTRODUCTION
The progression of multimedia and computer technologies has inspired both the industry experts/professionals and researchers into developing a more fascinating environment for learning which can help to improve students’ learning apprehension [1], [2]. E-learning, which refers to learning through the use of electronic tools usually on the Internet such as web search engines, in current situations, enables learners of all ages, competencies and preferences, to look for information or knowledge “anywhere and anytime” [3], [4]. It also allows learners instant access to particular information [5]. Nevertheless, behaviours, outlooks, learning styles, and potentials may vary from one student to another. These differences can eventually make an impact on their learning processes. Viewed from that perspective, it can be seen that giving learners the same learning materials may not create similar learning experiences. Instead, it may reduce their learning performances [5].

Keywords are required in making a web search and it appears that learners or students are given the same keywords when making their search through any major commercial search engines such as Google, Yahoo or Bing). Without doubt, this process is likely to return the same set of search results for all the students in the group regardless of their learning proficiencies [5]–[8]. Moreover, it was identified that when a dominant search engine was used to fetch educational materials, four results were found profoundly educative among the topmost 50 returned documents [9]. However, the relevancy of the results mainly relies on the right combination of keywords used in the query. Many students, especially novice learners often struggle with finding the right keywords particularly when they are new to the learning topics [6], [10], [11]. In addition, students may also have difficulties in selecting the most relevant links from the huge number of link results returned by the search engines [5], [6], [12]. Moreover, commercial search engines often place sponsored advertisements in the form of CPM (Cost per thousand viewers), CTR (Click-through rates), CPC (Cost per click), or CPA (Cost per action) over relevant items which also distract students from choosing the right sources of content from the returned search results [6], [13], [14]. In fact, previous investigations have found that users only consider the first 10 to 20 links of the results returned by search engines [15]. Inadvertently, this lands some learners with difficulties in searching for contents that match their requirements. Related to this is that search engines are often assessed with respect to the relevancy of web pages to
particular inquiries [16] whereas in reality, the search request may barely represent a portion of the user’s overall information requirement. This is because the same query may require different sets of information, depending on users’ needs. In this context, user’s profile information should be considered together with their queries by Web links. This is particularly important in the educational environment where students have different educational backgrounds and learning behaviours which indirectly, influence their learning progress and acceptability. As an example, a tutorial video prepared for an advanced level of students is not going to be suitable for a novice learner due to dissimilarity in level of understanding among them. It appears that the relevancy of the search results do not always convey users’ satisfaction [17]. An ideal search result could be one that is based on a personal evaluation of the user’s daily information needs derived from the returned links or contents of the Web search engine [17]. Similarly, in a collaborative learning environment, an assessment taken from a group of similar learner profiles can reveal more significant input for the learners rather than just the relevancy of the webpages [18], [19].

Also, worth noting is how personalised service on the Web can increase sales and this is one aspect of technology that needs to be emphasised on Web-based learning. Usually, when rendering personalised services, majority of the personalised systems ponder on user interests, user preferences and user’s browsing behaviours [20]. Literature [20] has also pointed out that the best performance comes from personalisation algorithms where a big chunk of data about an individual is ready to be used [21]. Using data that are gathered from various individuals is one way to augment the individual’s profile [19], [22]–[24]. In addition, the behavioural information of a group of people can also be acquired to be used to refine the personalisation outcome. This is especially true for certain groups and inquiries related to certain groups [19], [25]–[29]. By learning how other people undertook the same task previously, these individual users can also improve their own searching experiences through Web search engines [12]. Therefore, synthesizing one individual user’s data with the data of other members in a similar group’s profile can amplify their personalised search and searching experience.

There are many search applications which support individual users with their shopping, travelling, entertainment or personal requirements. One example is Agoda.com which recommends personalised tourist places and hotels for users. Another is Amazon.com that recommends products to its consumers and YouTube.com that makes personalised video recommendation [30], [31]. In comparison to the learning or teaching domain, it appears that there is hardly any application that is available in supporting e-learners to find learning materials that match their learning competencies when using Web search engines. Researchers [12] have shown that users’ profound feeling towards the need for search guidance is higher when they are looking for learning materials within unfamiliar topics or complex to understand the materials. Therefore, it is essential to augment the Web search engines with the ability to provide personalised recommendations of search results in the e-learning domain.

In this paper, we present a personalised group-based recommendation approach for Web search in e-learning. The primary motivation of this study is to present an adaptive e-learning method for students of different learning capabilities when using the popular search engines. To achieve this, a Web search recommender system was developed as a gateway between the Google search engine and the institutional e-learning portal so as to enable the search engine to deliver personalised search results as recommendations for students based on their individual needs. The rest of this paper is organised as follows. An overview encompassing other related works on finding learning materials from the Web using search engines as stated in the literature is provided. This is followed by an explanation of the framework and methodology employed to develop the proposed system. A section illustrating the experimental setup follows before the findings are discussed and concluded.

II. LITERATURE REVIEW
Most educational programmes and online systems offer similar learning tools to all students [32]. In order to accommodate customised learning materials for the individual students, upon investigating their profiles, researchers have strived to come up with numerous recommendation systems or adaptive learning mechanisms [1], [32]–[39]. Studies [40] show that about 80% of students use and prefer search engines as educational tools. Nonetheless, very few studies have been conducted to investigate how the popular search engines can be utilised to leverage students by providing them with personalised e-learning materials.

LOBSTER is a specialised search assistant tool based on the Google search engine. It was developed to help teachers find learning objects (LO) more easily [6]. The number of instances where teachers were able to obtain the required learning objects fairly quickly through LOBSTER has increased significantly compared with just using Google search engine independently. LOBSTER contains a few special characteristics such as term suggestions and a bilingual search which improve users search results. The major augmentation this study was to present a set of assistive elements that operate simultaneously with the Google search engine in assisting teachers to find LOs throughout the whole search process. Teachers receive assistance from the time they start drafting their inquiry by using the text boxes such as bilingual topics related term suggestion, clustering of search results according to
required language and LO types, advanced searches and suggestions for appropriate query terms. Although researchers have shown that the LOBSTER search assistant may help teachers to find learning materials more efficiently than using Google, the main characteristic missing from the system is the personalisation of learning materials based on individual users’ profiles. As a result, LOBSTER also returns same results regardless of the users’ learning competencies.

Many popular Web search engines such as Yahoo, Google, and Bing provide term suggestions for queries. When a search has been performed, users will obtain a list of returned search results and suggestions of relevant queries with some sponsored advertisements. However, it does not display additional information with regards to the personalised recommendations [6], [41].

In 2013, Microsoft started a beta version of its “Bing For Schools” programme which promises to provide the K-12 students with access to an ad-free search activity that comes with added privacy protections, content filters and a couple of functional and specific learning attributes [42]. The primary motivation of the programme was to provide a separate search engine for educational purposes and it is opened to all K-12 schools in the United States of America (U.S.). Microsoft revealed that the programme had previously attracted schools in the five largest districts of the U.S. and is currently used by about 4.5 million students. It was also asserted that the programme in total has served around 35 million ad-free queries. Given that the student population of this programme is 4.5 million, this means that there is just under eight queries per student, which is not exactly a very high number since most people do more than eight Google searches per day. Microsoft, however, reported that the programme is currently growing by more than a million queries per day [43].

Alongside the most used search engines (such as Google, Yahoo, Bing) are a few more search engines which are quite beneficial in assisting students to figure out the Web contents. Table 1 is an overview and comparison of the search engines which were specifically designed for educational purposes. The findings noted from literature revealed that:

1) Web search engines provide limited capabilities in personalising the search results according to the student’s profile even though the Web search engines were the most popular tool used by students in searching for educational materials.

2) Majority of research looking at adaptive learning approaches had concentrated mainly on the personalisation of learning materials in various e-learning systems but not on the Web search engines.

3) Several studies have examined the use of the Web search engines as e-learning tools but thus far, to the best of our knowledge no study was performed in providing personalised recommendations to students using Web search engines.

4) Groupization algorithm has not yet been implemented in personalising the delivery of learning materials through Web search engines.

### Table 1

<table>
<thead>
<tr>
<th>Name</th>
<th>Overview</th>
<th>Dynamic profiling</th>
<th>Personalized recommendation</th>
<th>Search the entire web/resources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Google Scholar</strong></td>
<td>Powered by Google, this search engine makes search simple to attain abstracts, citations, papers, and different scholarly literature.</td>
<td>No</td>
<td>No</td>
<td>Limited</td>
</tr>
<tr>
<td><strong>iSEEK</strong></td>
<td>Created particularly for educators and students, this search engine is non-commercial and presents results that are editor-reviewed from government sites, universities, and other non-commercial providers [44].</td>
<td>No</td>
<td>No</td>
<td>Limited</td>
</tr>
<tr>
<td><strong>DMOZ</strong></td>
<td>Technically it is not a search engine since it does not index the whole DMOZ, Web, or the Open Directory Project, it allows the searchable gateway to millions of links which are collected by volunteers who are specialists in their areas. [45]</td>
<td>No</td>
<td>No</td>
<td>Limited</td>
</tr>
<tr>
<td><strong>LOBSTER</strong></td>
<td>A specialized search assistant tool based on the Google search engine, comprising some advanced features to help teachers finding learning objects more easily in contrast with just directly using Google search engine [6].</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Search Guide (SG)</strong></td>
<td>A tool used as a Search Guide exhibits the trails of searches (pages bookmarked, results clicked, queries issued) from three past users who performed the task [12].</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Proposed system</strong></td>
<td>A Personalized Group-based Recommendation system to assist students to find more relevant personalized content using Google search engine.</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
III. DEVELOPMENT OF A PERSONALISED GROUP-BASED WEB SEARCH RECOMMENDATION SYSTEM FOR E-LEARNING

This study proposes a personalised group-based recommendation system for Web search. It comprises a Web interface that acts as a gateway between the Google search engine and the institution’s Web portal. Fig. 1 demonstrates the flow diagram of the online process. The Web portal provides a secure login and authentication in the proposed system. When a user logs into the system, the proposed system will automatically identify the user through the institution’s Student Management Information System (Student MIS) and via a local server record. When a student submits a query to the system, it goes to both the Google search engine and the system’s local server. As usual, Google will return a list of links that satisfies the keywords used in the query. If every student uses the same query, the returned links will be the same for all of them. However, through the proposed system, the local server will also return a list of recommended links that are personalised based on the student’s profile.

In the profiling process, students are categorised into three groups: Beginner, Intermediate, and Master. This is derived from their academic records and learning behaviours. The students’ profiles are automatically updated based on their learning behaviours when searching for e-learning materials using the Google search engine through the proposed system. The behavioural data were obtained from the students’ Web browsing history and session logs when using the system. Moreover, when the system detects any change in the academic records, it also fetches the updated information and it dynamically updates the user’s profile. For each type of student group, the students’ implicit feedbacks such as link selections, search activities and log histories are analysed and then stored in the local server. By consistently comparing this information with other students who carry similar profiles, a dynamic link recommendation outcome can be achieved.

A Web crawler is used to dynamically update the students’ profiles and content ranking lists. This is based on the changes noted in their academic records, learning behaviours and other related contextual records. The crawling is a cyclical process. The crawler visits both the Student MIS and the local Web server in every 72 hours’ time span. The goal is to fetch any changes appeared in the students’ profiles and ranking lists. Any change exhibited will be sent to a profile analyser to be processed so that the students’ profiles can be dynamically updated. The analyser also identifies changes that occur in the students’ log histories and session data before sending the information to a ranking analyser to be processed for re-ranking. The process aims to present the right set of recommendation links to the appropriate group of students.

**TABLE II CATEGORIZATION OF LEARNER PROFILE DATA**

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic</td>
<td>Previous and present academic records.</td>
</tr>
<tr>
<td>Information</td>
<td></td>
</tr>
<tr>
<td>Behaviour</td>
<td>No of login, search per login, page size</td>
</tr>
<tr>
<td>Information</td>
<td></td>
</tr>
<tr>
<td>Contextual</td>
<td>Changes in student academic performance and</td>
</tr>
<tr>
<td>Information</td>
<td>behavioural activities context</td>
</tr>
</tbody>
</table>

A. STUDENT PROFILING MODULE

This module aims to construct and maintain the profile of each individual student. It allows the system to understand the different learning needs and capabilities of each individual student. It then uses this information to improve the relevancy of the returned Web search results by selecting the most relevant personalised links. This is achieved by prioritising the links according to each profile. To accomplish this, the module has two functional components: Academic Record Analyser and Behavioural Activity Analyser.

Learner profile is the key component of any e-learning system [46]. Therefore, this study uses the most well-known user profiling model, the IEEE PAPI standard [47] to model the student’s profile. This model is further expanded by adding a learner profile design model [46]. It considers the following elements: Academic, Behavioural and Contextual record as shown in Table 2.

The academic record is used to measure the individual student’s past and present academic performance. This information is derived from the Student MIS. The system also observes the student’s learning behavioural activities through his/her browsing histories and session data. The goal is to assess the interest level and the activeness of the
student towards learning while using the Web search engine. Any change noted in the student’s academic and behavioural record will result in his/her profile update. The system also identifies other coherent contexts [48] by tracking any change exhibited in his/her academic and behavioural records for a time span and not just a few sessions or over a few days. The time span may range crossing months of data as contextual information begins to correspond to the dynamic profile needs. Figure 2 demonstrates the processes involved in the student’s profiling module.

1) ACADEMIC RECORD ANALYZER

The Academic Record Analyzer module is responsible for identifying the signed-in students by retrieving their profiles from the Student MIS. It mainly retrieves students’ previous and current academic information and then stores the information in the local server for further processing. Subsequently, the module calculates the standard T-Scores extracted from the raw scores of the academic records for each student using the grading policy used by the University of Texas at Austin [49]. The T-Score is one form of standardised test statistics that transform individual raw scores into standardised forms of scores for ease of comparison. It provides a constant of the mean and standard deviation on any set of data [50], [51]. Furthermore, the T-Score also reduces natural variations that occur within grade points thereby rendering a way to ascertain whether the scores are high or low. By averaging the calculated standard T-Scores, an average score is achieved for each student. This helps to classify the students’ profiles. In this study, we define it as Knowledge Point (KP).

2) BEHAVIOURAL ACTIVITY ANALYSER

The Behavioural Activity Analyser module is responsible for continuously monitoring and capturing the students’ learning behaviours. This is accomplished through their Web search activities while they use the proposed system. Each student’s learning activities are recorded and stored individually as extracted from their browsing histories and session logs. All information is stored inside the local server. The data captured comprises the number of times they login, the number of search per login, the issued queries, the selected documents, page names, page sizes, average scrolls, links clicked, and time spent. Subsequent to that, the stored activities are analysed so as to classify the students’ level of interest towards learning while searching for the relevant learning materials from the Web search engine.

Due to the dynamism of students’ learning behaviours through Web search, their behaviours and interests can only be interpreted within a fuzzy setting. Fuzzy rules are generated by applying the decision trees proposed by Sendhilkumar [52]. The fuzzy rules permit the students’ learning behaviours to be classified into four levels: Low, Low Medium, Medium and High. The classification also regards the pattern of the students’ behaviours as representing the students’ level of interest towards learning.

3) STUDENT PROFILE CLASSIFICATION

From the earlier sections, the student’s knowledge point (KP) and level of interest are obtained. These outputs require further processing so as to enable every student’s profiles to be classified. The students’ profiles are classified according to their academic performance (knowledge point) and learning behaviours (level of interest towards learning). In order to achieve this, the extended classification rule is applied [5], [10], [52]. A decision tree model is created from the data set by using the C4.5 algorithm [53]. This enables the students to be further classified into three groups (Beginner, Intermediate and Master) as shown in Figure 3. The decision trees can be utilised for inducing classification models usually alluded to as a statistical classifier. The C4.5 algorithm is also a popular algorithm that has appeared in the Top 10 Algorithms for Data Mining [54]. Studies [55] have confirmed that the C4.5 algorithm is the most preferred and most powerful method in use. In this study, the decision tree (Figure 3) is translated into the following rules by using the student’s Knowledge Point (KP) and interest level:

1) Students with KP higher than 80 will be assigned a Master Class.
2) Students with KP less than 63 will be assigned a Beginner Class.
3) Students with KP<=80, >=63 are classified based on the following:
   - If his/her interest level is Low, then he/she will be assigned a Beginner Class.
   - If his/her interest level is Low Medium, then he/she will be assigned an Intermediate Class.
   - If his/her interest level is Medium, then he/she will be assigned an Intermediate class.
   - If his/her interest level is High, then he/she will be assigned a Master Class.

**FIGURE 3. Decision tree for student profile classification.**

**B. CONTENT RE-RANKING MODULE**

For a given query, a Web search engine having personalised features can return different search results for dissimilar individuals. Alternatively, it can match the search outcomes in a different way according to the user’s intent [56] based on the user’s information needs and preferences [57]. However, it is very challenging to collect user’s data which should be rich enough that can understand the user’s precise needs and preferences. One way to address this challenging task is by combining related data gathered from other individuals with similar profiles when building personalisation features. In the ‘Groupization’ technique, personalisation is established by assigning higher weights to pages that are apposite with the predominance of the members of the group. This is achieved by harmonising with every group member’s document term frequencies and Web histories [23]–[26], [29]. The requirement for group recommendations has emerged in many situations such as suggesting a destination of travel for a family to spend a holiday break, a movie for a group of friends to watch together, a good restaurant for co-workers to have a weekday lunch or even a set of learning materials to assist students of similar profiles in understanding a topic. Intuitively, items that are interesting to a member of a group can also be interesting to other members of the same group. Furthermore, students learn better in well-structured cooperative group environment rather than in the traditionally structured classroom environment [58].

Groupization is one of the methods that can improve the value of collaborative search tools. This can be performed by using personalisation with shared interests [19], [29]. In the collaborative filtering technique, the recommender system suggests items depending on how identical users desire the item. It endeavours to uncover others with similar or related concerns/interests [18], [59]. However, the matches are constructed based solely on queries [18] or on suggestions or community connectedness not rankings [59]. Groupization takes into account the shared interests of the group members so as to generate better personalisation results (document ranking level) [29].

The shared information is usually insignificant or unavailable to search engines which encompasses information such as type and length of members’ relationship among the search engine users [29]. Hence, the available search engines are unable to identify information regarding the relationships among the group members due to unavailability of users’ public data and privacy concern. On the other hand, our proposed personalised group-based recommendation approach is able to use the advantage of the Groupization algorithm to discover different users’ profiles within the homogeneous group of users (students) in an educational institution environment. This process helps to strengthen the approach in delivering more accurate personalised recommendations within the e-learning domain.

This study employs the Groupization technique in the content analyser module (Figure 4) for the purpose of re-ranking the returned search results into a custom sequence and prioritize them in a way that is more pertinent to the group members based on member similarity and level of preference. The motivation behind using the technique is to improve the ranking of the Web search results thereby making it more relevant according to student’s profile. To achieve this, the Groupization algorithm performs group assessments on the Web contents. Specific groups are given priorities to certain contents based on what is useful for group members of similar profiles. To execute Groupization on a set of search results returned from the Google search engine for a search query, few steps are involved. First, a personalisation score is calculated for every single returned search result for every member of a group. Next, the Groupization score is calculated as the grand total of the personalisation scores of every member of the group. Subsequently, a weighted combination of the
Groupization score and the search result’s original rank is considered. This helps to conserve valuable information that is utilised by the Web search engine i.e. the result’s “authoritativeness” [23], [60].

In this study, let p1, p2, p3...pi represent the set of pages returned by a search query from a search engine (where i=1,2,3…n). Let N be the total number of students in a sample, ni is the total number of students in the sample who are classified as Beginner, Intermediate or Master. In this context, R is the total number of students who visited pi, and ri is the total number of students who visited page pi and who are Beginners, Intermediate or Master. The weight of the page pi is,

$$W_{pi} = \log \frac{(R_i + 0.5)(N - n_i + 0.5)}{(n_i + 0.5)(R - r_i + 0.5)}$$

A higher page weight indicates higher ranking. The top five most relevant recommendation results are presented since students spend more time on documents and learn better when the size of the search engine result pages decreases [61]. Different types of recommendation links will be presented to different groups of students as is exemplified by:

- **Individual.** This group comprises the learning behaviour of the current student. Within this group, related search tasks, queries and link selections are taken entirely from the prevailing user’s long-term history.
- **Group (Global).** Within this group, similar tasks, queries, and link selections are taken from each person’s Web search history, in this case, all students involved in the experiment.
- **Group (Class).** Within this group, similar tasks, queries, and link selections are gathered from each person’s search history from a particular class such as Beginner, Intermediate or Master.

**FIGURE 4. Content re-ranking module.**

**IV. EVALUATION METHOD**

To appraise the ease of use, usefulness, and effectiveness of the proposed personalised Web search recommendation system, an experimentation was designed and conducted on the undergraduate students of our university. A prototype was developed using PHP and MySQL and then hosted in the said university’s local server. Google API was used to integrate the Google search engine with the prototype. For the purpose of the experiments, the prototype was made available for the participants to access the system by creating individual accounts.

**A. PARTICIPANTS**

In total, 70 students were approached to participate in the experiment. The second-year students from the Faculty of Computer Science and Information Technology were chosen as participants. The intention of working with Computer Science students was derived from the assumption that they were more technically sound than others in using a search engine. The experiment aims to measure how easy these students can search for their learning materials from the Google search engine by using our proposed system. Among the 70 invited students, 60 students participated in the experiment. Four groups were made by randomly selecting 15 students regardless their academic records from the 60 participants to participate in four different sessions.

**B. MATERIAL**

In the experiment, the participants were requested to use the Google search engine through our developed prototype to solve their given problems. The problem comprises 25 multipurpose questions which deal with the advanced level of JAVA programming. The decision to ask the students to solve the JAVA advanced level programming problems is because the users’ interactions with the search engine is greater when it involves more complex tasks [12]. For further intricate tasks, participants expect assistance from the Web search engine to find the information that can help them to solve the given problems. It is expected that participants would profit from the more complicated tasks. Using the advanced level of difficulty will also bring more willingness among the participants to use a Web search engine to find the desired materials to solve their given problems, which is the key factor in this experiment. Once the participants have completed all the 25 questions within a given time frame, they were asked to complete a questionnaire form. This helps to determine their level of acceptance towards the proposed system. This questionnaire consists of 12 questions as shown in Table 3.

The questionnaire was adapted from the Technology Acceptance Model (TAM) proposed by Davis [62]. A TAM questionnaire model is used to explain how eagerly a user comes across the cold face of technology before accepting and using it [63], which has been used in several studies.
This instrument measures the user’s perception on the usefulness and the ease of use of the system. Ease of use is determined as the degree to which a person perceives the specific system would be effortless to use. Usefulness is determined by the extent to which a person anticipates that adopting a specific system would improve the performance of his/her task. In this study, questions Q1-Q6 measure the participants’ perception regarding the ease of use of the proposed system while searching for learning materials. Similarly, questions Q7-Q12 measure the participants’ perception about the usefulness of the proposed system. Each of the participants filled a Google survey form of all the statements using the 5-point Likert scale with values varying from (1) completely agree to (5) completely disagree.

<table>
<thead>
<tr>
<th>Participants’ Perceptions of Ease of Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q-01</td>
</tr>
<tr>
<td>Q-02</td>
</tr>
<tr>
<td>Q-03</td>
</tr>
<tr>
<td>Q-04</td>
</tr>
<tr>
<td>Q-05</td>
</tr>
<tr>
<td>Q-06</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participants’ Perceptions of Usefulness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q-07</td>
</tr>
<tr>
<td>Q-08</td>
</tr>
<tr>
<td>Q-09</td>
</tr>
<tr>
<td>Q-10</td>
</tr>
<tr>
<td>Q-11</td>
</tr>
<tr>
<td>Q-12</td>
</tr>
</tbody>
</table>

**C. PROCEDURE**

The experiments were conducted in four different sessions. Each session belongs to each succeeding group. Each session begins by asking the participants to solve all the 25 problems in 40 minutes. The single restraint placed on the participants throughout the session was to search for the solutions by employing the Google search engine that is accessible from the proposed system. The browsing history, session logs and time to complete the test were recorded. After solving the given problems, participants were requested to respond to a set of questionnaires on Google survey form. The same sets of problems and questionnaires were given to all the subsequent groups. Once the first group finishes their session, the other three sessions were conducted with the remaining groups successively.

There are two phases of data collections. First is the students’ personal and academic information which were automatically retrieved from the Students MIS and stored in a local server. Second is the students’ learning activities which were obtained from their browsing histories and session data were collected while they were using the proposed system.

**V. RESULTS AND DISCUSSION**

**A. ANALYSIS OF DEVELOPMENT OUTCOME**

To recall, the vastness and diversity of the contents in the Web has spurred the need to revisit the concept of one search results to fit all users’ proposition of the current Web search engine. Even though there are indeed noteworthy personalised search applications to support users with their shopping, travelling, entertainment or personal requirements, we believe that one community that also urges urgent consideration is the students where recommendations for suitable e-learning materials from the Web need to be addressed. To match this gap, we designed and developed a novel personalised Group-based recommendation system for the Web search engine users. It aims to deliver a more personalised search result recommendation that matches individual students’ profiles. To measure the reliability and acceptability of the proposed system, we conducted experiments with a group of second year undergraduate students from the Faculty of Computer Science and Information Technology. In this study, the users’ level of satisfaction in terms of the ease of use and usefulness of the system were measured. We also evaluated the searching performance and time taken by the users of four groups while using the proposed system. The search results’ accuracy such as the precision and recall evaluation were, however, excluded from this study.

During the experiments, students were given a set of problems. They were given access to the developed prototype system to search for resources that can help them to address those problems. The proposed system permitted the students to get personalised recommended links of the returned search results that matched their queries (such as ‘TreeSet and SortedSet’) and their learning profiles (such as Beginner, Intermediate) as shown in Figure 5. Note that the left side of Figure 5 displays the typical search results returned by Google search engine which are the same for all students. On the other hand, the recommended links returned by the system (as shown on the right side), appears to be dissimilar for different types of students’ profiles. Each profile belongs to one of the following three groups: Beginner, Intermediate, and Master.

The proposed system presents the top five personalised recommended links based on the group the students belong to. For instance, the same queries submitted by students from Intermediate and Master group will respectively receive different results as shown in Figure 6. It is worthy to note that the order of the recommended links is sorted based on the most relevant weighted page. For instance, ‘Java Collection Tutorial - Java Sorted Set’ link appears in
2nd order in the recommended list for the Intermediate profile while it is unavailable from the Master profile. This is because ‘Groupization’ algorithm assigns higher weight to it as the page is preferred by many members of the Intermediate group. On the other hand, the page is not preferred by most members of the Master group hence lower weight is assigned to it, so much lower than the weights of the top five links in the recommended list for the Master profile. The weight is determined by the total number of hits (descending order) made by students within the same group. The group is determined by the students’

![Search Interface and recommendations list display.](image)

**FIGURE 5.** Search Interface and recommendations list display.

**FIGURE 6.** Display of personalised recommendation list for (a) Intermediate profile, and (b) Master profile.
academic record assessments and their learning behaviours detected while using the system. Therefore, students of different groups will receive different sets of recommended links based on the different arrangement of priorities.

B. ANALYSIS OF STUDENTS’ ACCEPTANCE

In order to determine students’ acceptance of the proposed system, a survey was conducted using the TAM questionnaire. In this study, to measure the scale of reliability of the questionnaire set, we performed Cronbach’s Alpha test. In statistics, Cronbach’s Alpha is widely used to estimate of the reliability of an instrument. The test shows that our questionnaire has high reliability score which is around 95%. To recall, the focus of the survey was to gauge the level of ease of use and the usefulness of the proposed system. Table 4 demonstrates the results. The outcome of the one-way ANOVA test and the post-hoc test of the four groups of participants are presented. A one-way ANOVA is preferred more than the T-Test because this experiment compares averages of more than two groups. The Tukey HSD post-hoc test indicates which groups were significantly different from the others. For instance, ‘a<b’ and ‘b<c’ indicate that mean score of group one is significantly lower than group two and group two’s mean score is significantly lower than group 3, and so forth. The results obtained for both perceived ease of use (F = 6.11, p < .05) and perceived usefulness (F = 4.19, p < .05) denote that there were significant differences between the means of all the four groups. In both dimensions in terms of perceived ease of use and perceived usefulness,

TABLE IV
ANOVA RESULTS OF THE FOUR GROUPS (TAM QUESTIONNAIRE)

<table>
<thead>
<tr>
<th>Scale</th>
<th>Groups</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>F</th>
<th>Post hoc tests</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of use</td>
<td>Group 1 (a)</td>
<td>15</td>
<td>3.57</td>
<td>0.95</td>
<td>6.11</td>
<td>a &lt; b</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Group 2 (b)</td>
<td>15</td>
<td>3.91</td>
<td>0.78</td>
<td></td>
<td>b &lt; c</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 3 (c)</td>
<td>15</td>
<td>4.26</td>
<td>0.63</td>
<td></td>
<td>c &lt; d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 4 (d)</td>
<td>15</td>
<td>4.66</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usefulness</td>
<td>Group 1 (a)</td>
<td>15</td>
<td>3.72</td>
<td>0.86</td>
<td>4.19</td>
<td>a &lt; b</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Group 2 (b)</td>
<td>15</td>
<td>3.93</td>
<td>0.76</td>
<td></td>
<td>b &lt; c</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 3 (c)</td>
<td>15</td>
<td>4.34</td>
<td>0.63</td>
<td></td>
<td>c &lt; d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 4 (d)</td>
<td>15</td>
<td>4.53</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p<0.05.

highest degree of satisfaction in terms of perceived ease of use and perceived usefulness in comparison with the rest of the groups. When group four was asked to find solutions to the given questions, the personalised group-based recommendation technique enabled them to find solutions more easily and faster than the rest of the groups. This is because the system has gathered sufficient information from the completion of the previous sessions of the three groups thereby enabling the system to produce better recommendations of links for their solutions. On the other hand, when group one participated in the first session, they did not get any recommendation at all since the system had not accumulated enough information to produce recommendations. The initial group had to search for answers based on the default results returned by the Google search engine. Therefore, their level of satisfaction in respect of perceived usefulness and perceived ease of use was modest. Similarly, group three disseminated better level of satisfaction as compared to group two. Figure 7 shows that most of the students identified the proposed recommendation system as very easy to use and useful. The proposed system had helped them to find the relevant learning materials hence they were able to solve the given problems successfully and effortlessly. They were able to solve the given problems successfully and effortlessly.

FIGURE 7. The result of overall students’ perception towards the proposed system.

TABLE V
ANOVA RESULTS OF THE FOUR GROUPS (SEARCHING TIME)

<table>
<thead>
<tr>
<th>Scale</th>
<th>Groups</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>F</th>
<th>Post hoc tests</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search time</td>
<td>Group 1 (a)</td>
<td>15</td>
<td>34.60</td>
<td>3.97</td>
<td>6.11</td>
<td>a &gt; b</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Group 2 (b)</td>
<td>15</td>
<td>30.73</td>
<td>5.53</td>
<td></td>
<td>b &gt; c</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 3 (c)</td>
<td>15</td>
<td>24.40</td>
<td>4.50</td>
<td></td>
<td>c &gt; d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 4 (d)</td>
<td>15</td>
<td>21.26</td>
<td>4.21</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p<0.05.
C. ANALYSIS OF SEARCHING TIME

Table 5 shows the result obtained from the ANOVA test of the searching time for the four groups. We investigated the period of time the students of each group took to search for answer materials so as to complete their given tasks. The ANOVA result (searching time) shows that there was significant time difference among the four groups ($p<0.05$). For instance, the mean searching time for group one is 3.87 minutes longer than group two, and the mean searching time for group three is 6.33 minutes shorter than group two, and so forth. Group four found their desired learning materials within the shortest time. Each of the groups, starting from group one to group four, had shown a decrease in searching time as each group received better recommendations over time, for finding worthy materials to solve the given problems. Group one did not benefit from the personalised recommendation due to insufficient information gathered for the recommendation. In this respect, they were dependent on the default links returned by Google search. Only a few participants from the group were facilitated through run-time searching history that were created by other participants from the same group in the same session. Group four utilised the shortest time to solve the problems as they were able to receive quality recommendations from the proposed system. This implies that the personalised group-based recommendation approach had a significant impact in reducing students’ search time for relevant e-learning materials using the Web search engine.

D. ANALYSIS OF SEARCHING PREFERANCE

In order to determine the students’ preference towards the links displayed in each search results, we monitored their links selections. Figure 8 shows the total number of clicks on the Google returned links and the recommended links of each experimental group. When comparing group one to group four, it was observed that the number of clicks identified on the recommendation links increases while the number of clicks on Google returned links decreases. This exhibits an inverse relationship between recommendation links and the Google returned links when compared chronologically among the groups. One possible reason for the high number of clicks on the links displayed by Google is due to the students’ familiarity with Google as compared to the recommended links section. The exciting finding obtained from this experiment shows that there is an incremental interest to select the recommended links. As mentioned in the earlier part of this paper, as the participations of each group increases, their browsing histories also become richer hence making it possible for the proposed system to recommend links that were personalised to their profiles.

VI. DISCUSSION AND CONCLUSION

Although search engines have been extensively used by the students for educational purposes, they deliver similar contents regardless of students’ profiles [66]. This is not beneficial to the students because the same contents may not meet the requirements of every student. Web search results that are personalised to a student’s learning profile is therefore, necessary. In this study, a personalised group-based recommendation approach for Web search in e-learning was proposed. We designed an adaptive system that comprises dynamic profiling and content re-ranking mechanisms that will cater to students in finding Web-based learning materials based on their academic records and learning behaviours. The proposed system augments the Google search engine with the ability to recommend and prioritise the top five most suitable links to students depending on their personal profiles.

The results from the experimental study revealed that students from group four had benefitted the most from the personalised recommendation links provided by the proposed system. They gained a significantly better learning comprehension as compared to the earlier three groups through greater Web search experiences. Students from the preceding three groups had achieved better performance as compared to their immediately preceding groups (for instance, group three achieved better performance than group two, and so on). This finding implies that the personalised group-based recommendation system is helpful in improving the effectiveness of the students’ Web search and learning performance.

The survey results also indicate that the highest percentage of students who perceived the proposed system to be useful were those from group four with 52% strongly agree and 48% agree. This is followed by group three, group two and group one (see Fig. 7) respectively. Further to this, 98% of the students from group four found the system easy to use (66% strongly agree and 32% agree) when searching for relevant learning materials (see Fig. 7). The students’ perception about the ease of use of the system was noted as 54% from group one, 78% from group 2 and 88.5% from group three (see Fig. 7). This finding suggests that the personalised group-based recommendation approach was able to make the process of searching for learning materials using the Web search engine more...
effective thereby less exhaustive. In other words, it offers better support and promotes successful collaborative Web based learning attitude.

To provide a personalised Web search, the recommender system needs to access users' personal profiles, description of user interest and user behaviour [67], [68]. However, it is challenging for any recommender system to collect users’ data [68], thus more innovative approaches need to be considered. In this study, the proposed system employed in the educational institutional environment where student data can be obtained from the Students’ Management Information System. This therefore enriches the data of the students to construct rich user profile. By utilising a rich user profile and by identifying students with similar learning potential and attitude, the search engine can provide a more personalised recommendation of the search results. However, the current setup of the proposed system has a limitation because the custom Google Search API limits to only 100 free search queries per day. Nonetheless, the most challenging task is to encourage students to use Web search engine through an institutional e-learning portal. Therefore, future work should consider the use of add-ons into the respective browsers for delivering a seamless personalised recommended links to students. Additionally, the possibility of incorporating students’ social identities from their social networks for richer students’ profiles and more desirable recommendation outcomes would also be an appealing topic to explore.

ACKNOWLEDGMENT
This study is supported by the University of Malaya Research Grant (RP032D-16SBS) and the Postgraduate Research Grant (PG276-2015B).

REFERENCES


M.M. Rahman and N.A. Abdullah: Personalised Group-Based Recommendation Approach

MOHAMMAD MUSTANEER RAHMAN is currently working as a graduate research assistant with the Faculty of Computer Science and Information Technology, University Malaya, Kuala Lumpur, Malaysia. He received his B.S degree on computer science and engineering from Khulna University, Khulna, Bangladesh. He is currently pursuing the master’s degree in computer science at University Malaya. He has also six years of industry experience in software development. His research interests include E-Learning, Recommender Systems, Interactive Learning, Technology Enhanced Learning, Spatial Crowdsourcing, Decision Support System, Machine Learning and Big data.

NOR ANIZA ABDULLAH is an Associate Professor at the Faculty of Computer Science and Information Technology, University of Malaya, Malaysia. She graduated with Bachelor degree (hons) in Computer Science from the University of Malaya. She has a Master degree in Interactive Multimedia from the Westminster University, London. She obtained her PhD degree in Computer Science from the Southampton University, United Kingdom. She has authored or co-authored over 50 refereed publications in international journals, book chapters, and conferences. She supervised several Ph.D. and Master’s students in University of Malaya, Malaysia. She also co-supervised several Master by Research students in the Moratuwa University of Sri Lanka. She serves as reviewer for several ISI-indexed journals. Her research interest is in Personalized and Adaptive Learning, Recommender System, Decision Support System, Big Data Analytics, and Content-based Image/Video Retrieval.