INTERNATIONAL JOURNAL OF COMMUNICATION SYSTEMS
Int. J. Commun. Syst. 2013; **26**:732–746
Published online 3 April 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/dac.2331

Intelligent route generation: discovery and search of correlation between shared resources

Neil Y. Yen[1,*,†], Runhe Huang[2], Jianhua Ma[2], Qun Jin[3] and Timothy K. Shih[4]

[1]*School of Computer Science and Engineering University of Aizu Fukusima, Japan*
[2]*Faculty of Computer & Information Sciences Hosei University Tokyo, Japan*
[3]*Deparment of Human Informatics and Cognitive Sciences Saitama, Japan*
[4]*Department of Computer Science & Information Engineering National Central University Taoyuan, Taiwan*

SUMMARY

Sharing information and resources on the Internet has become an important activity for education. The use of ubiquitous devices makes it possible for learning participants to be engaged in an open and connected social environment, and also allows the learning activities to be performed at any time and any place. In this study, the discovery of correlation among shared resources is concentrated. A hypothetical scenario is considered that the information, such as photos and thoughts, is applicable to be shared with implicit context (i.e., geographical information) by learners through a practical implementation, PadSCORM, on a mobile device. Two major contributions are achieved. First, the correlations among resources are determined through usage experiences mining and geographical information adjustment. It then assists learners in filtering out redundant information by highlighting the significance of resources. Second, an intelligent searching algorithm is proposed to visualize adaptive routes to facilitate search process and to enrich the learning activity. The empirical experiments revealing the feasibility and performance (e.g., accuracy and effectiveness) are conducted in the universities in North Taiwan. Copyright © 2012 John Wiley & Sons, Ltd.

Received 27 September 2011; Revised 26 January 2012; Accepted 27 January 2012

KEY WORDS: pervasive computing; social network analysis; information filtering; result enrichment; ubiquitous learning

1. INTRODUCTION

As a promising approach, social network analysis has been recognized as an efficient way to clarify the complex interactions among individuals and/or objects. It enables implicit, or indirect, relations to be found through the mining of common intersections [1]. The concept can also be applied to every domain in which the interactions may occur by considering customized factors and metrics [2]. Hence, from a perspective of knowledge sharing, it raises the emerging issues about patterning collective intelligence [3].

Social learning, especially in forms of sharing and discussing, is becoming an emerging paradigm presently. Social learning in a ubiquitous scenario can make learning activities available without any limitations from the environment. Participants can share their contexts (e.g., thought, location, emotion, etc.) simply through a portable device. For instance, learners may strengthen their domain abilities by finding out some relevant information, or may collect resources for different learning purposes. It is almost the same as the instructors in creating the learning materials. In this situation, a

*Correspondence to: Neil Yen, School of Computer Science and Engineering , University of Aizu Fukusima, Japan.
†E-mail: neil219@gmail.com

well-designed resource discovery mechanism that meets such requirement may be needed to achieve high efficiency and flexibility.

In earlier works [4], a series of integrated services have been developed to assist users in performing corresponding learning activities. An outdoor adventure game was implemented to support learning procedures based on the use of geographical information and radio frequency identification technologies. To increase the efficiency of learning resource discovery, an intelligent search guidance algorithm [5] was developed. In this study, based on a systematic reexamination, we concentrate on the issues of correlations among shared resources. The resources are considered as any digital file, such as text, photo, and/or combinations, which are generated simply through portable devices (iPad in this study). The methods based on social network analysis are applied to make explicit representation of the resources shared by learners in a ubiquitous environment. The weighted graph is utilized to illustrate the possible correlations, which are quantified for further usage. Then the user's context information (i.e., geographical location) is utilized to generate the adaptive route in accordance with the selected resource on the map. The route, identifying the directional information, represents only the first level connection to avoid the noise [6] while displaying the information on a device. Learners are expected to build knowledge corresponding to specific topics and area information while performing the learning activities.

1.1. Motivation and contributions

Considering the massive information sharing channels, a systematic method that achieves efficient management and reuse is required. Although methods to social network analysis have provided preliminary solutions to implicit relations discovery, a few issues on interactions, especially among objects, are still left unsolved while being applied for further usage. The resources do not exist independently because they are the major media that prompt the interactions. And thus, interactions might be led in accordance with different types of correlation among resources. In addition, the search service makes use of customized metrics, such as ranking or recommendation rules, to assist users in obtaining desired resources. However, the obtained resources are isolated, and that is, one query is made for one best fit resource (and other independent resources with lower priorities). If users would like to search for other related resources, few queries are required. This may discourage continuity while forming specific knowledge.

To cope with the mentioned situation, two major contributions are achieved in this study. First, the vague interactions among resources are clarified and utilized to construct an object network where the main participants are shared resources. A few metrics based on social network analysis and time-series data mining are proposed to quantify the correlation for further reuse. Second, as a practical contribution, an automated search mechanism is implemented to generate possible routes related to selected resource. The route is generated based on not only existing correlations but the relative physical distance between user and resources. The achievements can be regarded as a preliminary solution to facilitate information discovery and summarize the collective intelligence, and being the supplementary materials for specific knowledge formation.

The rest of this paper is organized as follows. Related issues and works are discussed in Section 2. The proposed methods and corresponding metrics and automated algorithms are introduced in Section 3. Section 4 represents the implemented results. The experiment results and performance are illustrated and discussed in Section 5. We conclude this work and address future work in Section 6.

2. RELATED WORK

Three major issues correlated to the study are walked through in this section. First the technologies and scenarios of ubiquitous learning are identified. The issues to information filtering are then discussed with the focus on learning content delivery. The use of social network analysis to achieve implicit correlation mining is addressed in the last.

DOI: 10.1002/dac

2.1. Applied technologies in U-learning environments

Wireless technologies, such as radio frequency identification and GPS, have been applied for facilitating instructional processes by creating a highly interactive environment [7, 8]. A system was developed to support the foreign language learning. The task-based strategy was applied to increase the learning motivation in an outdoor learning environment [9]. As an extension, Chen *et al.* [10] constructed a mobile-learning system for outdoor learning activity support, where the learning process of each learner can be observed utilizing the mobile devices. Learners will be divided into several groups and the instructor will ask learners to discover different kinds of birds and find out relevant information. Hsu *et al.* [11] also utilized the same strategy to construct an asynchronous self-regulated learning environment. Learners can get involved in the learning scenario created by instructors through this strategy. Moreover, instructors can also realize the learning status or learning history of each learner, and, if necessary, can give feedback to learners instantly based on individual needs.

Summarizing the above issues, the benefits of using portable devices with ubiquitous technologies are addressed, and represent the importance of learning content (such as the delivered tasks and the given instructions) delivery. However, the issue on how the information is retrieved by learners themselves has not yet been addressed.

2.2. Information retrieval and filtering

As stated, the retrieval issue on appropriate information has become an open issue in performing ubiquitous learning activity. To provide relevant/adaptive information, the filtering methodologies are often adopted to solve these issues and, in general, can be discussed in three categories: (1) user-based filtering — the method classifies users based on the similarity factors including users' interests or hobbies, to generate the user model [12]. Also, the system can provide resources adaptive to specific user groups; (2) item-based filtering — with the increasing number of users, the user-based filtering method exposes the problem of excessive computation time. According to Ref. [13], an item similar to another item in which a user is interested may attract this user's attention. It can provide more stable recommendations based on the similarity between items; (3) model-based filtering — this type of recommendation concentrates on providing solutions to the limitations (e.g., information scarcity and scalability) remaining from the above two methods, which may cause the provision of inaccurate results. In brief, this method adopts the advantage from user profile, the same as the user models created by user-based filtering, and trains the following usage experiences to make prediction [14, 15].

2.3. Web search

The filtering methods are always implemented in the search service. A well-structured search service should include three major parts including a friendly user interface that allows users to give customized requests, the process logic where the core algorithms exist, and the rich data sources that can be processed. An interactive interface [16] that achieves multimodal concentrations, such as basic textual input, tag and hyperlink navigation, and the visual processing, was implemented in a WordNet-based ontology repository. As a technical extension, the query expansion [17] mechanism was proposed to facilitate the search experience and avoid ambiguous queries. Users, especially those novice to specific domains, would be given guidance, such as potential keywords, based on their initial query. To make an efficient expansion of the query, two major issues are concerned. First, the representation of user's query is required. A semantic approach [18] was proposed to identify the input and extract essential text combinations. With the representative text set from query, the in-text tagging algorithm [19] for obtaining representative keyword sets were also widely discussed to achieve better matching results.

2.4. Methods based on social network analysis

Social network analysis concentrates on the discovery of implicit correlation via the explicit exchange of resources among participants. The participant, including the human user and/or object,

DOI: 10.1002/dac

ranges from individual to group [20]. In general, the structure of the social network can be divided into two types, homogenous and heterogeneous, in accordance with their internal attributes. To obtain expected information from such networks, it is necessary to take some factors defined in a graph into consideration by using the method of data mining for implicit information and/or pattern discovery [21]. The representative methods for determining the significance of nodes and links include PageRank [22], Hyperlink-Induced Topic Search [23], and EigenRumor [24], whose return values can be regarded as weights of measuring the degree of centrality. The semantic methods [25, 26] are also widely used for describing interdependency within the social structure. The semantics of a node is modeled using its surrounding labeled network structure, which is represented by the sequences of labels (i.e., paths) together with some statistical dependency measures associated with them. A scalable mining method [27] is also applied to assist researchers in discovering and generating the possible connections.

Technology has been widely applied to the learning environment. It provides comprehensive support (especially seamless [28]) in accordance with the characteristic of the learning paradigm. Social media is an outstanding instance that offers a well-interactive channel for learners to share and obtain knowledge. In this study, we employ the concept of social network analysis and web search to facilitate the implicit correlation discovery among shared resources and to generate the adaptive routes while performing ubiquitous learning activity.

3. MODELING SHARED RESOURCES AND INTELLIGENT SEARCH ROUTE GENERATION

The general model that is utilized to identify correlations among shared resources is addressed in this section. It begins with the basic definitions, the formation of graph model, and the metrics that is applied for graph quantification. The automated algorithm that generates the adaptive routes is then discussed.

3.1. Definition of correlation types

Interactions, especially discussions in the form of post-and-reply, in social networks lead to collective knowledge. The process of discussion identifies the resource, any digital files (e.g., text, photo, or combination), raised by information sharing via mobile devices or any web services [29, 30] in the ubiquitous environment. The resource is applicable for specific reuses, which depends on the possible correlations summarized in four categories including:

- **Independency**
  Independency reveals the isolated existence of resource. The resource performs a specific activity or achieve a specific target without external support (e.g., prerequisite in learning) from others.
- **Reference**
  Reference addresses the concept of equivalence. The resource represents similar aims while being labeled as reference between connected resources.
- **Sequence**
  Sequence identifies the concept of coexistence. It reveals that the resource, often the last element in a sequence, is triggered if and only if the previous element is adopted.
- **Equivalency**
  Equivalency identifies the concept of peer existence. The resource labeled with this correlation can be utilized as an alternation of others that have similar attributes.

The correlations are then utilized as basic parameters to generate the connections (or edges in a network) among resources.

DOI: 10.1002/dac

3.2. A Graph for resource representation

With the definition, the graph model [31] is then developed as a weighted digraph expressed by

$$G = (V, E, W) \tag{1}$$

The notation V indicates the vertice set, $V = \{v_1, v_2, \ldots, v_n\}$, and is expressed by

$$V = \left\{x_i|_{x \in V, i \in N}\right\}$$

The notation E represents the edge, appending with specific correlation, which connects the vertices in V so that we have $E : V \times V$ with the general expression

$$E = \left\{e_{ij} = \left(x_i, x_j\right)|_{\forall x \in V, i \neq j, j \in N}\right\}$$

The weight W associated with the edge is developed to identify the strength of specific relation. It is represented as $W = \{w_1, w_2, \ldots, w_n\}$ and refers to a real number, $W : E \to R$, denoted by

$$W = \left\{w_{ij} = \left|e_{ij}\right||_{w_{ij} \in [-1.0, 1.0]}\right\}$$

The value of the connection weight decides the direction between vertices and equals to the absolute value of e_{ij} represented as $\left|e_{ij}\right|$. If w_{ij} is positive, the direction is from v_i to v_j and, on the contrary, is from v_j to v_i. For the case of 0, it represents no direct correlation between resources; that is, other resources are required to make the connection.

3.3. Highlighting the importance of resources

To quantify the weight of connection, two factors, frequency of use and its corresponding time information, are considered. The equation, $W(E)$, that implements the function, $W : E \to R$, can be expressed as follows:

$$W(E) = \sum_{e \in w^+} f(e)_s \cdot t(e)_s + \sum_{e \in w^-} f(e)_s \cdot t(e)_s \tag{2}$$

$$w^+(v) = \{e \in E | w(e) = (v, k), k \in V\}, \forall v \in V$$

$$w^-(v) = \{e \in E | w(e) = (k, v), k \in V\}, \forall v \in V$$

where $f(e)_s$ represents the frequency of use within the period s and $t(e)_s$ is the corresponding weighted coefficient. The two operators, $w^+(v)$ and $w^-(v)$, are utilized to sum the weight of opposite direction of the relation.

Considering the time information, two models, tilted time window model [32] and time fading model [33], are employed with the focus on integration of timeframe. The notation t is utilized to represent the timestamp of the event that the relations take place. We then have the basic definition of the time information as $t : t \in \mathbb{T}$ where $\mathbb{T}$ represents the length of the time, initializing with the system service, and $t = \{t_0, t_1, t_2, \ldots, t_n\}$ indicates the specific timestamp revealing the beginning and end of the resources that have been shared.

We then divide the length of selected timeframe into smaller units to determine the exact weighted coefficient of each period within the timeframe. For example, the unit 'hour' can be used if the selected length is a day while 'day' is also acceptable if the original length is set to 'month' and/or 'year'. The weighted coefficient, appended on the timeframe, can be obtained dynamically in accordance with the period a connection exists and the occurrence frequency a connection has. The weighted coefficient is utilized to highlight the significant information during the specific period within the timeframe.

DOI: 10.1002/dac

With the definition, the weighted coefficient of each period within the given timeframe can be obtained. The following equation is applied to implement the function $t\,(e)$:

$$t\,(e)_s = \frac{D_{q-s+1}}{\sum\limits_{s=1}^{q} D_s} \tag{3}$$

where D represents the unit, q is the sum of separate units and s is the indexer.

After the time information, we go further to discuss the quantification of the frequency of use among the shared resources. Here, the frequency of use of the resources concentrates on the co-occurrences, with other resources, the resource contains while performing a specific learning activity. The Hebbian algorithm is employed to quantify the frequency of use and highlight the significant connections. The general expression of the proposed method is implemented as shown in Equation (4).

$$f\,(e)_s = \frac{\sum \left(C\,\left(e_{ij}\right) \cdot (1 + H_{\text{coe}})\right)}{m} \tag{4}$$

where $C\,\left(e_{ij}\right)$ represents the involved correlation, m identifies the number of the homogenous correlation the selected connection e_{ij} has.

The notation $f(e)$ is interpreted as the pattern, from v_i to v_j, applied in specific learning activity. The pattern is characterized as the visit and its corresponding succession situation. The mechanism records the visit history while a recommended path is adopted by user. Note that the sequence of the pattern concerns (e.g., $v_i \to v_j \neq v_j \to v_i$). The succession situation then refers to the times that the pattern is adopted without violation. To give normalization to the mentioned factors (visit and succession situation), the learning algorithm revised based on Hebbian rule [5] is proposed. The Hebbian rule states that the strength (or weight in this work) of connection shall be raised if correlated nodes are triggered in accordance with existing patterns. On the contrary, it shall be reduced if the pattern is violated (i.e., only one of them is triggered). This approach is applied dynamically to adjust the weight of link and to highlight the patterns often used. Thus, in Equation (4), the additional weight, considering frequency of usage, is added to $f(e)$ by H_{coe}, which can be expressed as

$$H_{\text{coe}} = \phi \sum \overrightarrow{e_{ij}}\,\overrightarrow{e_{ij}}^{T} \tag{5}$$

The connection from v_i to v_j is converted to the vector, $\overrightarrow{e_{ij}}$, with the length representing the frequency of use. Note that the weight does not cause the difference on the length. To avoid the extreme value, the vector $\overrightarrow{e_{ij}}$ is then projected onto a virtual coordinate space to obtain additional increment or reduction to w_{ij} under normalization. That is, the obtained value will be relative to every other resource connected to the datum resource. An equilibrium coefficient ϕ is set in accordance with the directional relation (or usage direction) between resources. Its value will be within –1 and 1, but will not be equal to 0 because the existing pattern identifies that there is at least one existing connection between the designed resources on the graph. With the normalization, the value of H_{coe} is also regarded as the credibility of specific patterns (e.g., the significant connection).

3.4. Intelligent route generation

After the construction of graph model for shared resources, we then go further to generate the route based on specific resource (or one selected by the user). The overall scenario is illustrated in Figure 1. The shared resources (circles with black or red color) are considered as separate dots on the map. The route is generated while one of displayed resources (red circle for instance) is selected. In this study, we concentrate on the generation of first-level connection (e.g., the next connected resource(s)). That is, only the resources that have direct correlations are connected. To achieve the goal, two major factors are concerned. The first one is the existing connections (or correlations) among resources recorded in accordance with the concrete usage. The second one is the distance between connected resources (i.e., selected resource and resource 1, 2, and 3 in Figure 1) and the geographical information of the user. The distance is concerned because the scenario is performed in

DOI: 10.1002/dac

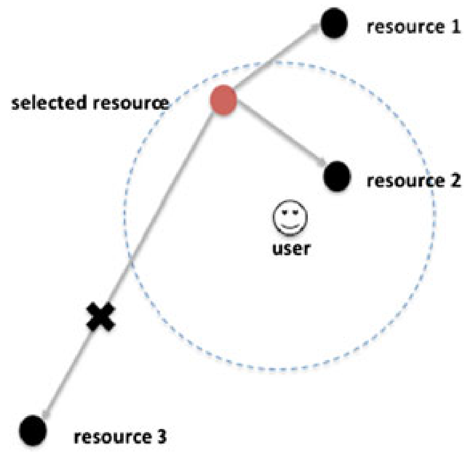

Figure 1. Illustration of the distance between resources.

Input: A selected resource v_i
Output: One or more connected resources C_v

Step 1: Find resource $v_j \in V$ connected directly to v_i.
Step 2: Find entry resource Set $G_i \subset G \bullet \forall e_{ij}, j \in \mathbb{N}$. Let E_i denotes the *Edge Bag* and make $E_i \leftarrow E_i \cup e_{ij}$.
Step 3: For $\forall v_j \in G_i$, find connected resources, $v_k \in V$. Let $G_i \leftarrow G_i \cup v_k$, and $E_i \leftarrow E_i \cup e_{jk}$.
Step 4: Use edges and its containing correlations in E_i to generate the continuous route S_v starting from v_i.
Step 5: Check user's geographical information, $l_{geo} = (l_x, l_y)$, compute the distance, d, between user and v_i.
Step 6: Take d as the radium to generate a tolerant range χ and filter out those resources v_j connected with v_i that are not within the range.
Step 7: Return C_v according to the given resource v_i, and display the first-level connection

Figure 2. Algorithm for connected resources retrieval.

a ubiquitous environment. In addition, it is necessary to reduce noises (i.e., unusual and/or redundant connections).

In Figure 1, three resources (resource 1, resource 2, and resource 3) and corresponding connections are obtained according to existing correlations. Once the geographical information of a user is determined, the connections would be modified to satisfy the tolerable moving distance. And thus, resource 3 is removed.

The algorithm (see Figure 2) to achieve automated route generation and optimization is illustrated as follows.

The algorithm represents the procedure of route generation from the existing shared resources starting from a selected resource v_i to a set of connected resources (C_v). Seven steps are involved. In Step 1, the resource v_j in V, which has direct connection with v_i, is found based on existing correlations. In the beginning, four categories of correlations are utilized. The correlation(s) between each connected resource v_j is then determined in Step 2. The *Bag* is adopted instead of the *Set* because it is possible for connected resources to contain more than one correlation. After that, the correlations in E_i are sorted by the priority of correlation (Sequence > Reference > Independency).

DOI: 10.1002/dac

The priority is determined in accordance with the number of resources the correlation could lead to. For example, the Sequence belongs to mandatory correlation, and that is, two resources, at least, are obtained once this correlation exist. For the rest, they are optional and are applicable to be obtained separately. In Steps 3 and 4, the subgraph G_i is generated to record the candidate resources until the similarity between default attributes of compared resources reach large disparity. The algorithm that computes the similarity coefficient between resources is shown in Figure 3.

The cosine similarity function is employed as the basic metric. The elements defined in IEEE learning object metadata are adopted as attributes, which were automatically assigned [4] in accordance with its content while being shared. The obtained coefficient is situated between 0 and 1, and the tolerable threshold for similarity matching is set to 0.80 by default.

In Step 5, the location information l_{geo} is recorded in the form of a two-dimensional coordinate system. With the information, the distance d between selected resource and user is then determined. We then compute the appropriate distance d to be utilized to obtain the route from v_i. The method to compute the distance is applied by Equation (6)

$$d = 2r \cdot \sin^{-1} \Delta, \tag{6}$$

where

$$\Delta = \sin^2 \theta + \cos l_x^A \cos l_x^B \sin^2 \theta'$$

and

$$\theta = \frac{l_x^B - l_x^A}{2}, \theta' = \frac{l_y^B - l_y^A}{2}$$

The variable r is the fixed length and known as the radium of the earth. The obtained distance d is then considered as a new radium to draw a great circle with the resource v_i. The area of the drawn circle is considered as the tolerant range for the user to reach. After that, the resources in G_i are compared with v_i by Equation (7) to filter out those resources connected but not within the range. The rest resources will be included in the set C_v to be returned.

$$d_{\mathrm{direct}} = \frac{\overrightarrow{v_i} \cdot \overrightarrow{v_j}}{\left|\overrightarrow{v_i}\right| \left|\overrightarrow{v_j}\right|}. \tag{7}$$

4. IMPLEMENTATION

The implementation phase is addressed in this section, which contains two parts. In the beginning, the concept with regard to the interface design is discussed. An example showing the concrete scenario of the implemented algorithm is then given.

Input: resources v_i and v_j
Output: numeric similarity coefficient κ

Step 1: Initialize default attributes for comparison. Set $E_{v_i}[r_w][s_w]$, $E_{v_j}[r_z][s_z]$ where $r, s \in \{IEEE\ LOM\}$
Step 2: Check the attributes that v_i and v_j contain.
For each element in E_{v_i} and E_{v_j}
Check r_w and $r_z \in$ same category
$\kappa += \sum \frac{\overrightarrow{s_w s_z}}{|s_w||s_z|}$
continue and check the rest categories
End For
Step 3: Return κ where $0 \leq \kappa \leq 1$

Figure 3. Algorithm for similarity calculation.

DOI: 10.1002/dac

The principle [34] is followed to design the user interface, which is shown in Figure 4.

The interface is composed by two portions: the control panel, including the search tool and annotation tool (A in the figure) and the display area (B in the figure). The control panel is responsible for the actions operated by users. An annotation tool deals with the input from the user. It allows the user to leave information (or resources) simply through typing in the text area (A-2 in the figure) or/with a link to external resource (A-3 in the figure). A search tool is developed to assist discovery of annotations. The display area is then responsible for the representation. Note that the annotations correlated to the selected zone on the map are presented in this study. The mentioned zone is determined by the user in accordance with individual preference.

The proposed algorithm to generate adaptive route is implemented on portable devices (i.e., Apple iPad and iPhone in this study) with the preinstalled application named PadSCORM. The mentioned device is taken as the client to assist resource sharing while performing specific activities.

A concrete example is utilized to represent the implementation result. The scenario is assumed to be that users are asked to familiarize with the environment through the collection of resources shared by previous users and to share their thought and/or information for further reuse.

The interface of PadSCORM, which is based on the simulator, is shown in Figure 5. Two portions corresponding to the design are revealed. The left-hand side (A in Figure 4) is the control panel. It provides three main functions including resource searching, sharing (through annotation on the map), and exchanging (in the form of reply), to the users while the right-hand side (B in Figure 4) illustrates the display area.

In Figure 5, six resources, created by previous users, in correspondence to the given map size are obtained through the default search query 'history.' We assume that the resource no. 1 is selected as an entry to generate the route(s), which identifies the connection between resources shown on the map, and that is, the possible sequence is revealed. Note that the routes will generate adaptively based on the selection. And thus, in this example, the routes exist as no.1–no.2, no.1–no.3 and no.1–no.6 in accordance with user selection (i.e., no. 1) and constraints mentioned in the previous section. The split lines are utilized to visualize the resources that have appropriate information with the selected resource. Note that the directional relation is also concerned. For instance, there exists a sequence from resource no. 1 to resource no. 2 in Figure 5.

Figure 6 then represents another instance of the selection of resource no. 2. In this case, two connected resources are shown. As mentioned in previous sections, we concentrate on the generation of first-level connection. Although the complete graph is generated, only the selected resource is utilized as a new query to search the graph. In this situation, the follow-up information can be shown to learners in an efficient way.

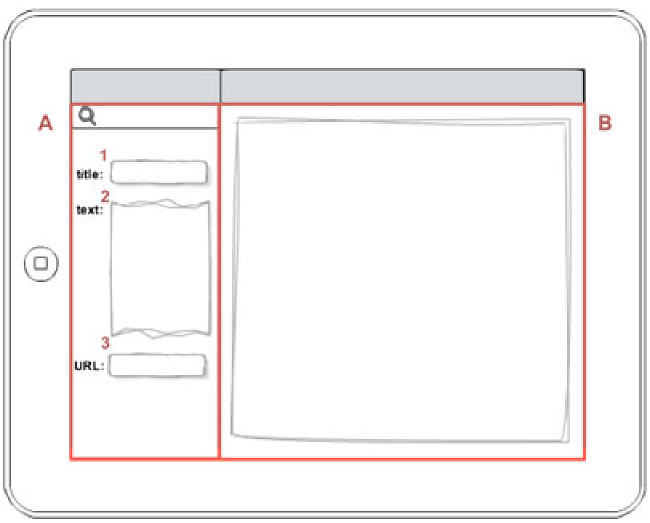

Figure 4. User interface design of PadSCORM.

DOI: 10.1002/dac

Figure 5. Illustration to the user interface design.

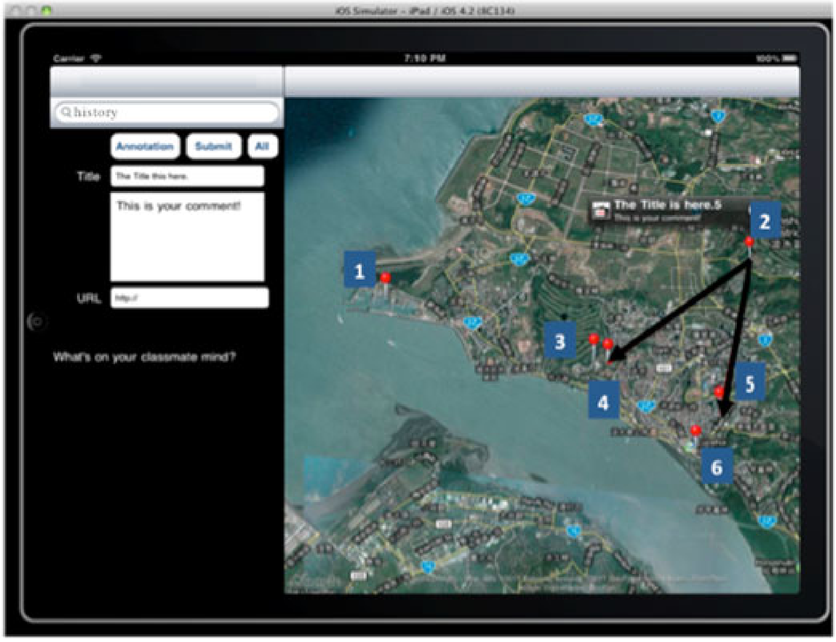

Figure 6. Instance of resource selection and generated route.

5. EVALUATION AND EXPERIMENT

An experiment is conducted to evaluate the performance of retrieved connections based on Precision–Recall [35]. The selected resource is considered as a query to obtain relevant resources.

The following metrics are applied:

$$\text{Precision}=\frac{\text{Number of resources retrieved}}{\text{Number of retrieved resources}}$$

$$\text{Recall}=\frac{\text{Number of relevant resources retrieved}}{\text{Number of relevant resources}}$$

In this experiment, not all of the relevant resources are taken into consideration. That is, the resources within a specific geographical range (i.e., Tamshui, North Taiwan as an instance in the study) are concentrated. The experiment results, the P–R curves generated by three sets of raw data, reveal that the performance of connected resources retrieval (also known as the generated split line) reach the baseline. The baseline (i.e., dashed line in Figure 7) is based on the average of all the relevant resources in the database.

The empirical study was conducted to evaluate the usability of the proposed automated route generation algorithm. In this experiment, we pay emphasis on the feedback from users. The usage data were collected from 77 college students. Three major factors were concerned while accessing the implemented system. The first is the 'number of query given', which identifies the expected queries from end users. The second is 'number of resources obtained', which identifies the possible results that can be obtained per query. And the third is 'number of routes generated', which represents the richness of supplementary information. The results evaluated based on raw data and clustered data are shown in Table I.

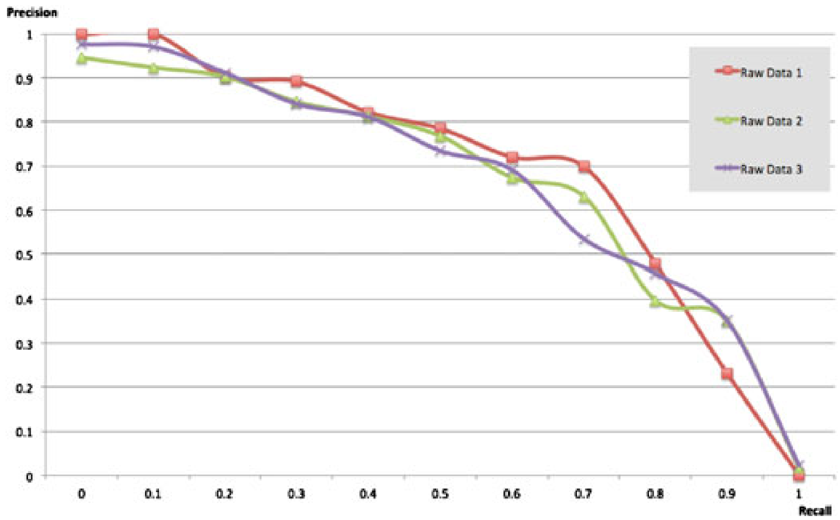

Figure 7. P–R evaluation of retrieved resources.

Table I. Usability evaluation results.

						(Unit: Item)
	Raw data			Clustered data		
Test NO.	Query	Resources	Route	Query	Resources	Route
1	219	701	306	219	7.60	3.06
2	346	6.32	3.67	346	7.13	1.54
3	451	765	4.39	271	641	2.27
4	4.04	7.10	470	4.04	6.79	319
5	315	6.77	2.31	315	703	414
Avg	3.47	6.97	3.63	3.11	6.99	2.84

DOI: 10.1002/dac

According to the experiment results, we can observe the following: (1) each query can obtain around seven resources; (2) each resource can connect to three other related resources; and (3) users can obtain expected resources and/or domain knowledge in less than four separate queries.

Following the previous experiment, we go further to examine the effectiveness of proposed automated algorithm. Two substantial assumptions are given:

- Assumption 1: The generated route may lead to confusing information display. That is, the query times in the posttest are more than one in the pretest.
- Assumption 2: The generated route may lead to rich information display. That is, the query times in the posttest are less than one in the pretest.

The one-tailed test is applied to verify the assumptions where α at 0.05 and DOF at 76. According to the t-distribution table, which assesses whether the means of two groups statistically differ from each other, DOF $= 76$ maps to a t value of 1.668. According to the formula, we obtained the value $t = 5.018$, which is greater than 1.665. Thus, Assumption 2 is acceptable and reveals that the generated routes can not only enrich the original resource but reduce the queries, around 1.2 per time in this study, in searching for supplementary resources.

	Pretest	Posttest
Mean	3.4	2.2
Observations	77	77
Pearson correlation coefficient	0.361	—
Degree of freedom	76	—
t Statistic	5.018	—
t Critical one-tail	1.668	—

6. CONCLUSION AND FUTURE WORK

The shared resources have become popular and, in fact, plenteous in a ubiquitous environment. For the educational perspective, the shared resources sometimes are considered as a kind of learning material while performing specific activities. To achieve efficient usage of the shared resources, an automated algorithm has been proposed to represent the correlations among them, and to search the implicit route in accordance with the correlation and user's geographical context. The algorithm generates the adaptive route based on the selected resource on the map, and connects the resources by using the split line representing the directional information. Learners are applicable to follow the route to build the knowledge regarding specific topics addressed in learning activities.

The major contributions of the study are summarized as follows. First, the method based on social network analysis is applied to represent the shared resources in a ubiquitous learning environment. Second, the automated algorithm is utilized to identify the useful and reachable resources relative to learners' context, and to generate the corresponding route(s).

The implemented application (i.e., PadSCORM) is considered as a preliminary step to discover the correlations among user-generated content. Although educational theories have shown to us how sequence of resources are designed to fit instructional purposes, a quantitative approach to measure effectiveness in the generation of learning activities is still difficult. We will try to give explicit description to the shared resources to achieve specific instructional purposes, and continuously provide solutions to the learning support.

ACKNOWLEDGEMENT

The work was jointly supported by the Institute for Information Industry, Taiwan, the Research Center for Science & Technology for Learning of the University System of Taiwan, 2010–2012 Waseda University Advanced Research Center for Human Sciences Project 'Distributed Collaborative Knowledge Information Sharing Systems for Growing Campus,' and the 2010-2011 RONPAKU Program of JSPS (Japan Society for the Promotion of Science).

REFERENCES

1. Palla G, Barabasi AL, Vicsek T. Quantifying social group evolution. *Nature* 2007; **446**:664–667.
2. Kumar R, Ragbavan P, Rajagopalan S, Tomkins A. The web and social networks. *IEEE Computer* 2002; **35**:32–36.
3. O'Reilly T. What Is Web 2.0: Design patterns and business models for the next generation of software. *International Journal of Digital Economics* 2007; **65**:17–37.
4. Shih TK, Wang TH, Chang CY, Kao TC, Hamilton D. Ubiquitous e-learning with multimodal multimedia devices. *IEEE Transactions on Multimedia* 2007; **9**:487–499.
5. Yen NY, Shih TK, Chao LR, Jin Q. Ranking metrics and search guidance for learning object repository. *IEEE Transactions on Learning Technologies* 2010; **3**:250–264.
6. Landay JA, Myers BA. Sketching interfaces: Toward more human interface design. *IEEE Computer* 2001; **34**:56–64.
7. Xiong NX, Cao M, Vasilakos AV, Yang LT, Yang F. An energy-efficient scheme in next-generation sensor networks. *International Journal of Communication Systems* 2010; **23**:1189–1200.
8. Liu K, Lee VCS. Performance analysis of data scheduling algorithms for multi-item requests in multi-channel broadcast environments. *International Journal of Communication Systems* 2010; **23**:529–542.
9. Ogata H, Yin C, Paredes RG, Oishi Y, Ueda T. Supporting mobile language learning outside classrooms construction. *International Conference on Advanced Leaning Technologies*, Kerkrade, The Netherlands, 2006; 928–32.
10. Chen YS, Kao TC, Sheu JP. A mobile learning system for scaffolding bird watching learning. *Journal of Computer Assisted Learning* 2003; **19**:347–359.
11. Hsu JM, Lai YS, Yu PT. Using the RFIDs to construct the ubiquitous self-learning environment for understanding the plants in the schoolyard. *IEEE International Conference on Wireless, Mobile, and Ubiquitous Technology in Education*, 2008; 210–212.
12. Herlocker J, Konstan J, Riedl J. An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms. *Information Retrieval* 2002; **5**:287–310.
13. Sarwar B, Karypis G, Konstan J, Riedl J. Item-based collaborative filtering recommendation algorithms. *World Wide Web Conference*, 2001; 285–295.
14. Suryavansh B, Shiri N, Mudur S. Improving the effectiveness of model based recommender systems for highly sparse and noisy web usage data. *IEEE/WIC/ACM International Conference on Web Intelligence*, 2005; 618–621.
15. Umyarov A, Tuzhilin A. Improving collaborative filtering recommendations using external data. *IEEE International Conference on Data Mining*, 2008; 618–627.
16. Pyshkin E, Kuznetsov A. Approaches for web search user interfaces: How to improve the search quality for various types of information. *Journal of Convergence* 2010; **1**:1–8.
17. Klyuev V, Yokoyama A. Web query expansion: A strategy utilising japanese WordNet. *Journal of Convergence* 2010; **1**:23–28.
18. Klyuev V, Oleshchuk V. Semantic Retrieval: An approach to representing, searching and summarising text documents. *International Journal of Information Technology, Communications and Convergence* 2011; **1**:221–234.
19. Kim J, Jin DS, Kim KY, Choe HS. Automatic in-text keyword tagging based on information retrieval. *Journal of Information Processing Systems* 2009; **5**:159–166.
20. Haythornthwalte C. Social network analysis: An approach and technique for the study of information exchange. *Library & Information Science Research* 1996; **18**:323–342.
21. Jensen D, Neville J. Data mining in social networks. In *National Academy of Sciences Symposium on Dynamic Social Network Analysis*, 2002a.
22. Langville AN, Meyer CD. *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton Univ. Press: New Jersey, USA, 2006.
23. Kleinberg J. Authoritative sources in a hyperlinked environment. *Journal of the ACM* 1999; **46**:604–632.
24. Fujimura K, Inoue T, Sugisaki M. The EigenRumor algorithm for ranking blogs. *World Wide Web Conference*, Chiba, Japan, 2005; 59–74.
25. Weng CY, Chu WT, Wu JL. RoleNet: Movie analysis from the perspective of social netowrks. *IEEE Transactions on Multimedia* 2009; **11**:256–271.
26. Wolf T, Schroter A, Damian D, Panjer LD, Nguyen THD. Mining task-based social networks to explore collaboration in software teams. *IEEE Software* 2009; **26**:58–66.
27. Albert R, Barabashi AL. Emergence of scaling in random networks. *Sciences* 1999; **286**:509–512.
28. Lai YX, Lai CF, Hu CC, Chao HC, Huang YM. A personalized mobile IPTV system with seamless video reconstruction algorithm in cloud networks. *International Journal of Communication Systems* 2011; **24**:1375–1387.
29. Kryvinska N, Thanh DV, Strauss C. Integrated management platform for seamless services provisioning in converged network. *International Journal of Information Technology, Communications and Convergence* 2010; **1**:77–91.
30. Oh JM, Moon N. A cultural dimensions model based on smart phone applications. *Journal of Information Processing Systems* 2011; **7**:209–220.
31. Tolsdorf CC. Social networks, support, and coping: An explorato-ry study. *Family Process* 1976; **15**:407–417.
32. Chen Y, Dong G, Han J, Wah BH, Wang J. Multidimensional regression analysis of time-series data streams. *International Conference on Very Large Databases*, 2002; 323–324.
33. Chang JH, Lee WS. Finding recent frequent itemsets adaptively over online data streams. *ACM International Conference on Knowledge Discovery and Data Mining*, 2003; 487–492.

DOI: 10.1002/dac

34. iOS Humane Interface Guidelines. http://developer.apple.com/library/ios/#documentation/UserExperience/Conceptual/MobileHIG/Introduction/Introduction.html, retrieved on 14 Dec, 2011.
35. Järvelin K, Kekäläinen J. IR evaluation methods for retrieving highly relevant documents. *ACM International Conference on Information Retrieval*, Athens, Greece, July 24-28, 2000; 41–48.

AUTHORS' BIOGRAPHIES

Neil Y. Yen received his doctorates in Human Sciences at Waseda University, Japan, and in Engineering at Tamkang University, Taiwan, in 2012. He is an assistant professor in the School of Computer Science and Engineering at the University of Aizu, Japan. His research interests are primarily in the scope of Web information retrieval, social computing. and knowledge engineering. Contact him at ny-yen@u-aizu.ac.jp.

Runhe Huang is a professor in the Faculty of Computer and Information Sciences at Hosei University, Japan. Her research fields include multiagent systems, computational intelligence, and ubiquitous intelligence computing. Huang has a PhD in Computer Science and Mathematics from the University of the West of England. Contact her at rhuang@hosei.ac.jp.

Jianhua Ma is a professor and head of the Department of Digital Media in the Faculty of Computer and Information Sciences at Hosei University, Japan. His research interests include multimedia ubiquitous computing, social computing, and cyber intelligence. Ma has a PhD in Information and Electronics Engineering from Xidian University. Contact him at jianhua@hosei.ac.jp.

Qun Jin is a tenured full professor in the Networked Information Systems Laboratory, Department of Human Informatics and Cognitive Sciences, Faculty of Human Sciences, Waseda University, Japan. He was engaged extensively in research work in computer science, information systems, and social and human informatics. He seeks to exploit the rich interdependence between theory and practice in his work with interdisciplinary and integrated approaches. His recent research interests include user-centric ubiquitous computing sustainable and secure information environments, user modeling, behavior informatics, information search and recommendation, human-computer interaction, e-learning support, and computing for well-being. He is a senior member of the IEEE. More information about him can be found on http://www.f.waseda.jp/jin/.

DOI: 10.1002/dac

Dr. Timothy K. Shih is a Professor at the National Central University, Taiwan. He was the Dean of College of Computer Science, Asia University, Taiwan, and the Department Chair of the CSIE Department at Tamkang University, Taiwan. Dr. Shih is a Fellow of the Institution of Engineering and Technology (IET). In addition, he is a senior member of ACM and a senior member of IEEE. Dr. Shih also joined the Educational Activities Board of the Computer Society. His current research interests include multimedia computing and distance learning. Dr. Shih has edited many books and published over 470 papers and book chapters, as well as participated in many international academic activities, including the organization of more than 60 international conferences. He was the founder and co-editor-in-chief of the International Journal of Distance Education Technologies, published by the Idea Group Publishing, USA. Dr. Shih is an associate editor of the IEEE Transactions on Learning Technologies. He was an associate editor of the ACM Transactions on Internet Technology and an associate editor of the IEEE Transactions on Multimedia. Dr. Shih has received many research awards, including research awards from National Science Council of Taiwan, IIAS research award from Germany, HSSS award from Greece, Brandon Hall award from USA, and several best paper awards from international conferences. Dr. Shih has been invited to give more than 35 keynote speeches and plenary talks in international conferences, as well as tutorials in IEEE ICME 2001 and 2006 and ACM Multimedia 2002 and 2007.