60% accurate. We can summarize that the number of leaves for TreeQ does not make any impact on the accuracy of the system while the LVQ with a small number of NOC shows a lower performance when tested against a big corpus.

Making a corpus, we have to consider and clearly train the corpus with the right genre. Based on our experience, we usually have a problem with the accuracy of the TreeQ and LVQ which show a bad performance since we trained the corpus with the songs by artist selection. We observed that one artist may produce a lot of styles of music even if he/she is classified as a country artist. Choosing the right song for training the corpus is very important for system performance.

For the system itself, we can say that it can perform very accurately if given clearly trained songs. We found some limitations of using this system such as the input format of the song (wav format) - we have to make sure that the sampling rate is the same for all training songs (44.1 kHz). We also tried to build a huge corpus by training almost 250 songs for the corpus, which meant we would have a very big training corpus since the HCOPY module can only work with audio (.wav) format. When creating a big corpus, we had to use a powerful computer since building a bigger corpus required a lot of time to generate the feature vector for the corpus.

5. DISCUSSIONS AND OUTLINE OF A RELEVANCE FEEDBACK -BASED MUSIC RETRIEVAL SYSTEM

In principle, classification and retrieval are inherently similar tasks. Throughout our experiments with TreeQ and LVQ based classification, we discovered that TreeQ based experiments yield faster response times than those of LVQ; and therefore, TreeQ–based systems maybe appropriate for online (real-time) music retrieval task. On the other hand, LVQ-based experiments consistently yield better accuracy than those using TreeQ; and therefore, LVQ-based systems may be more appropriate in music classification tasks, since music classification can be generally performed off-line.

Interactive retrieval such as web-based *Query by Humming* systems requires a faster response. Due to the inherent complexity in multimedia information processing, we propose a Relevance Feedback -based interactive system as outlined in Fig. 4. We envisage that with current state-of-the-art technology in processing music data, such a relevance feedback – based system may lead us to a usable interactive retrieval system. The proposed RF-based system works in 2 stages: Firstly, the user's initial query is matched against the entire music database to find the best matches (i.e., to maximize recall). Secondly, based on the user feedback (i.e., user selection), the scope of the query is restricted into a smaller domain (i.e., the specific genre) to improve precision.

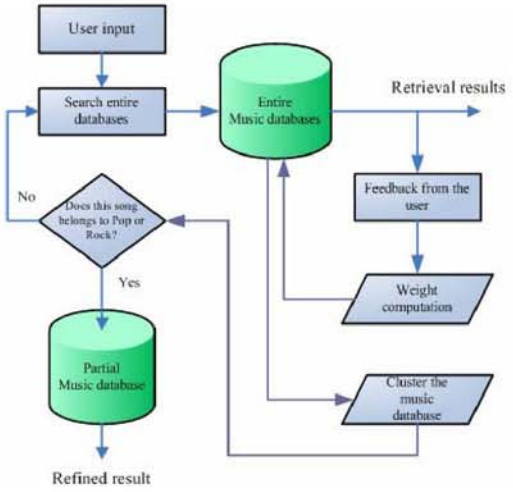

Fig. 4 Flow chart of interactive music retrieval using relevance feedback from user

6. CONCLUSION

We have studied the automatic classification of music signals according to their genres. An important part of the work was to build a highly accurate music genre classification system. This was a time-consuming task but worth the effort, since the amount of training data is an important factor when evaluating classification systems. The classification accuracy was found to depend on how clearly we trained the system to recognize the music genre.

We can conclude that any inaccuracies may involve several issues. First, modern Thai music is very difficult to classify as Pop, Rock or Country because the styles of playing are very similar to each other. We have to train the system with the right songs. Second, the variation of NOC of LVQ can be a factor that decreases the systems performance, since we observed that LVQ with NOC 200, which was tested in phase III against 50 songs*5 genre, showed a lower performance (60%) with these parameters.

As future work, we are planning to put together a TreeQ-based classification engine with the proposed relevance feedback loop (c.f. Fig. 4) for interactive retrieval of Thai music using a *Query by Humming* approach. We are confident that a system with this technique can reach the best usability in terms of precision and recall as justified earlier.

REFERENCES

[1] J. Foote, "A similarity measure for automatic audio classification," in Proc. AAAI 1997 Spring Symposium on Intelligent Integration and Use of Text, Image, Video, and Audio Corpora, March 1997.

[2] H. Soltau, T. Schultz, M. Westphal and A. Waibel "Recognition of music types," in Proc. International Conference on Acoustic, Speech, and Signal Processing, Lansdowne, Virginia, USA, Feb. 1998.

[3] B. Logan, *Mel-Frequency-Cepstral-Coefficients-for-music-modeling*," in Proc. International Symposium on Music Information Retrieval (ISMIR), Plymouth, MA, USA, Oct. 2000.

[4] G. Tzanetakis, G. Essl and P. Cook, "Automatic musical genre classification of audio signals," in Proc. International Symposium on Music Information Retrieval (ISMIR), Bloomington, IN, USA, pp. 205-210, Oct. 2001.

[5] The TreeQ package, http://treeq.sourceforge.net

[6] T. Kohonen, "The Learning Vector Quantization Program Package," Technical Report A30, Helsinki University of Technology, Laboratory of Computer and Information Science, FIN-02150 Espoo, Finland, 1996.

[7] The HTK Toolkit, http://htk.eng.cam.ac.uk/

[8] D. Reynolds, R. Rose, *Robust-Text-Independent-Speaker-Identification-Using-Gaussian-Mixture-Speaker-Models*," IEEE Transactions on Speech and Audio Processing, vol. 3, no. 1, pp. 72-83, 1995

[9] W. Chai and B. Vercoe: "Folk Music Classification Using Hidden Markov Models," in Proc. International Conference on Artificial Intelligence, June 2001.

[10] C. McKay and I. Fujinaga, "Automatic genre classification using large high-level musical feature sets," in The 5th International Conference on Music Information Retrieval (ISMIR), Barcelona, Spain, October 2004.

[11] T. Kohonen, "Self-Organizing Maps (SOM)," volume 30 of Springer Series in Information Sciences. Springer, Berlin, 3rd edition, 2001.

[12] L. Lu and H. J. Zhang, "Content analysis for audio classification and segmentation," IEEE Transactions on Speech and Audio Processing, vol. 10, no. 7, 2002.

International Conference on Information and Communication Technology
ICICT 2007, 7-9 March 2007, Dhaka, Bangladesh

USING PUBLICATIONS AND DOMAIN KNOWLEDGE TO BUILD RESEARCH PROFILES: AN APPLICATION IN AUTOMATIC REVIEWER ASSIGNMENT

Humayun Kabir Biswas, Md Maruf Hasan

School of Technology, Shinawatra University
1010 Viphavadi Rangsit Road, Bangkok 10900, Thailand
E-mail: mhkbiswas@yahoo.com; maruf@shinawatra.ac.th

ABSTRACT

Peer-review has been a common practice for quality control in scholarly publications for decades. The ubiquity of the Internet and, subsequently, the availability of easy-to-use Web-based systems (both free and commercial) has made the peer-review process fast, cost-effective and convenient. In a typical scenario, authors upload papers online and manually assign topic-areas; reviewers also sign up by letting the system know about their area of expertise. A rudimentary Paper-Reviewer matching is usually performed by the system and validated by the Program-Chair (for conferences) or by the Editor-In-Chief (for journals). As argued in relevant literature, the peer-review process suffers from several flaws including author's or reviewer's bias in choosing topic-areas and expertise, as well as inter-reviewer agreement, etc. In this research, we explore automatic reviewer assignment for papers by solely considering the content of the papers and the true profile of the reviewers.

In this research, we experimented with *three* approaches to calculate paper-reviewer relevance using the Vector Space Model. We used a set of 10 papers, 30 reviewers and the real paper-reviewer assignment information from a real-conference; and justifed the result of automatic paper-reviewer assignment based on the above three approaches. We noticed that the overlap between real-assignment and automatic-assignment is poor (with only 55-66% of the reviewers being in common). Such a result was not surprising to us, since we are aware that reviewers often express their frustrations claiming that some papers assigned them are not in line with their preferences and expertise. The data-set we used was rather small and suffered from data-sparseness problem and therefore we tried to analyze the automatic-assignment rationales through unbiased human judgment to identify the effect of the above-mentioned approaches in automatic reviewer assignment. We concluded that combining domain-knowledge with automatically extracted keywords (i.e., ontology-driven topic inference using automatically-extracted keywords) could potentially identify the most relevant candidate-reviewers for a paper.

1. INTRODUCTION

With the advent of the Internet and the WWW, the peer-review process has been highly optimized, and we no longer need to submit paper via snail-mail and opt for the same throughout the entire process of decision making, thus shortening the overall turnover time. In scholarly publishing process, when an author submits his or her scholarly work to a journal or a conference for publication, the Chief-Editor or the Program-Chair distributes the paper to potential reviewers for comments and rating. Based on the review comments a neutral decision is made about whether to or not to include the paper in the journal or in the conference. It is frustrating to see that peer-review has been the standard procedure for scholarly publication, but only 8% of the scholars believed that the peer-reviewing process is an efficient practice [1]. In this paper, we will focus on **(a)** building researcher's profile using the researcher's publications and domain ontology, **(b)** topic extraction using free text of a paper and ontology-driven inference, and finally, **(c)** automatically assigning reviewers to papers by measuring topic similarity.

In this research, we used the powerful vector space modelling (VSM) technique and machine learning tools for extracting features such as keyphrases, and Semantic Web technology such as ontology-driven topic inference to facilitate efficient Paper-Reviewer assignment. Both Reviewer's Profiles and Papers are identified using automatic methods and matching is also done automatically to avoid human-bias and to find the best possible Paper-Reviewer matches. Traditionally, authors usually assign keywords from a pre-defined topic-set, and the reviewers also do the same to specify his or her expertise [2]. A rudimentary matching is then carried out between a paper and a group of reviewers based on rudimentary

matching and human-judgment. This practice has its own inherent flaws. By automatically extracting relevant features from a paper and the reviewer's profile (expertise) from his or her publications, we could locate the right reviewers for a particular efficiently. This addresses the quality issues and eventually enhances the peer-review process. To a greater extent, our approach is capable of overcoming unexpected review outcomes (such as less-credible or biased review comments) due to a reviewer's inability to assess an assigned paper in a particular domain he or she is less familiar with. We attempt to justify our approach of paper-reviewer assignment using real paper-reviewer assignment datasets taken from a real conference and through human judgment.

2. RELATED WORK

The Vector Space Model (VSM) is a powerful tool in representing documents, and is widely used in Information Retrieval [3,4,5]. In VSM, any document can be represented as a vector in an n-dimensional term space. Each element of the vector represents a distinct term (or term-based feature) occurred in the document. Such a vector is usually called term vector, which is a bag-of-words representation of a document in the term-hyperspace. The same idea can be extended to represent the profile of a reviewer as long as we can crawl and capture a set of publications written by a person/reviewer from common sources such as, WWW, Google Scholar, ACM Digital Library, Citeseer and other repositories. We adopted the vector space representation to represent a paper, as well as to represent reviewers (using a set of documents written by the reviewer and gathered by crawling the Web or digital archives). Several relevance ranking algorithms are proposed in IR literature. For simplicity, we have chosen the MySQL Relevance Ranking algorithm as explained below.

2.1 MySQL Relevance Ranking for VSM Model

MySQL uses a local and a global term weight and normalizes the weight as explained below [4]. Term weights can be calculated using a function such as, Equation 2.1:

$$W_{i,j} = F(L_{i,j}, G_i, N_j) \tag{2.1}$$

which can also be represented as Equation 2.2:

$$W_{i,j} = L_{i,j} * G_i * N_j \tag{2.2}$$

Equation *2.1* and *2.2* account for local ($L_{i,j}$), global (G_i) and normalization (N_j) information. Local weights are functions of how many times each term occurs in a document, global weights are functions of how many documents in the collection contains the term appear, and the normalization factor corrects the discrepancies in the lengths of the documents. In the *classic "TF.IDF"* representation, *L, G* and *N* are defined as follows [5].

$$L_{i,j} = tf_{i,j}$$

$$G_i = \log\left(\frac{D}{d_i}\right)$$

$$N_j = 1$$

Therefore, the well-known *TF.IDF* weighting scheme is written as:

$$W_{i,j} = tf_{i,j} * \log\left(\frac{D}{d_i}\right).$$

In the MySQL implementation, however, normalization is performed heuristically. A typical application uses database tables consisting of *N* rows where each row corresponds to a document. *Local* and *Global* weights as well as *Normalization* are defined as follows:

1. $L_{ij} = (\log(dtf) + 1)/sumdtf$; i.e., local information based on logarithmic term counts
2. $G_i = \log((N - nf)/nf)$; i.e., global information based on probabilistic IDF
3. $N_j = U/(1 + 0.0115 * U)$; i.e., normalization based on a *pivot* value of 0.0115

In the above equations, *dtf* = number of times a term, *i* appeared in a document *j*; *sumdtf* = sum of (log (*dtf*) +1)'s for all terms in the same document; *U* = number of unique terms in the document collection; *N* = total number of documents in the collection; and *nf* = number of documents (rows) containing the term. The 0.0115 quantity is a *pivot value* and is determined heuristically.

Therefore, Equation 2.2 can also be rewritten as follows (Equation 2.3):

$$W_{ij} = (\log(dtf) + 1)/sumdtf * \log((N - nf)/nf) * U/(1 + 0.0115 * U) \tag{2.3}$$

We used Equation 2.3 to compute the term weights for words. In the Information Retrieval context, once the term weights are determined we need a ranking

function to measure the similarity between the query and documents vector [5]. The rank or relevance (R) is computed as the *dot product* of the document (term vector) and the query frequency of the words in the query (query vector, *qf*). The rank R is computed as follows:

$$R = W_{ij} \bullet qf \ .$$

In a paper-reviewer assignment scenario, the relevance between a paper and the reviewers can easily be modeled in line with the relevance between a query and a document since reviewers are authors themselves, and we can collect and merge their writings (set of publications) into a document.

We envisaged that full-text representation of a paper or a reviewer in a bag-of-word modelling tool (e.g., VSM) is highly rudimentary. Therefore, we also considered using machine learning -based Keyword extraction and represent both papers and reviewers in terms of *keyword-vectors* (based on the automatically extracted keyphrases). Furthermore, we also used domain-ontology based topic-inference to represent papers and reviewers more precisely over a vector space comprising of *topic-vectors*.

2.2 Automatic Keyphrase Extraction

The Keyphrase Extraction Tool, KEA [6] automatically extracts a number of keyphrases from full text using machine learning algorithm. The set of all candidate phrases in a document are identified using elementary lexical processing, features are computed for each candidate, and machine learning is used to generate a classifier or model that determines which candidates should be considered as keyphrases. The following figure (Fig.1) represents the two major steps used in KEA for automatically extracting keyphrases. In the first step, training is conducted using a set of training documents to build a classification model. In the second step, we use the trained model to extract potential keyphrases from document. As explained in KEA literatures, this tool is capable of extracting keyphrases in a domain independent fashion.

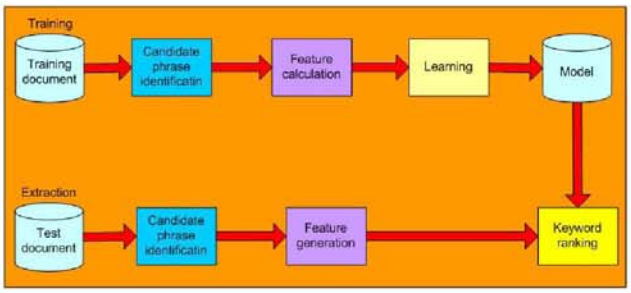

Fig.1: The Training and Extraction Process of KEA Algorithm

The training technique of the KEA algorithm uses a set of training documents for which the author's keyphrases are known. For each training document, candidate phrases are identified and their feature values are calculated using the TF.IDF and other features, such as the first occurrence position of that phrase [6]. Once the training is completed, the KEA tool can extract keyphrases from any new document based on the trained model.

Our second set of experimental results is based on a VSM representation of Keyphrases extracted from a particular paper, and form reviewers' publications (a set of documents). We computed the d similarity between the paper vectors and the reviewers' profile vector.

In the third leg of our experiment, we tried to use ontology to infer topic by means of these extracted keyphrases as explained further in the following section.

2.3 ACM Classification Hierarchy and Ontology

We developed a domain-ontology based on the ACM Computing Classification System (CCS) to make ontology-driven inferences and to perform mapping between a set of keyphrases and a set of ACM Classes (specific topics). Such specific topic-areas represent a paper or a person (reviewer's expertise) more specifically when compared to its raw term-based or Keyphrase-based representations. The ontology-schema is as shown in Figure 2. The topic inference is made through `hasKeyword` attribute and class relationship in ACM CCS.

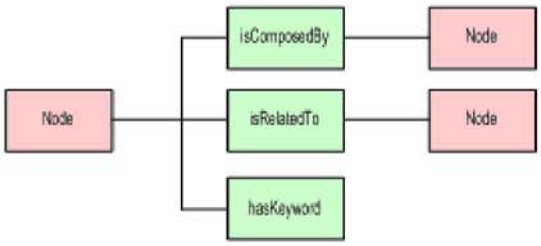

Fig. 2: Schema used for ACM CCS ontology

As an illustrative example, when we query the domain ontology using a single keyword, *concurrency*, we infer three potential topics: Operating System, Database and Foundation of Programming. However, when we query the ontology using a set of keyphrases such as *concurrency* and *transaction*, we managed to uniquely identify some specific topics, such as "Database". Such ontology-driven inference and the VSM representation helps us to match papers with their potential reviewers in a more precise manner on the topic space [7, 8].

2.4 Reviewer's Profile

For a reviewer, we use Web Crawler and Search Engines to collect a sizeable amount of text authored or co-authored by that person such as his or her publications from public repositories (Google Scholar, ACM Digital Library, Citeseer, etc.). Using the powerful VSM framework, we then conveniently build a reviewer's profile using *three* different ways: in terms of *free-text*, in terms of *extracted keywords* from the free-text, and finally, in terms of *inferred topics* using the domain ontology.

3. EXPERIMENTS AND EVALUATION

As mentioned earlier, in this research, we experimented with three approaches to calculate paper-reviewer relevance using VSM. First, we use free-text in representing papers (particular paper) and reviewer profiles (using a set of documents) into weighted vectors, and compute similarity among those vectors to identify the most relevant reviewers for a particular paper to ensure that a paper is assigned to the best possible reviewers for the best possible feedback. Next, we use automatic key-phrase extraction tools (such as KEA) to automatically extract keywords from the text corpora to calculate relevance between a paper and potential reviewers. Finally, we use domain-ontology and make use of ontology-driven inference to map automatically extracted keywords into topics and re-calculate paper-reviewer relevance on the topic-space.

We used a set of 10 papers, 30 reviewers and real paper-reviewer assignment data from a real conference to justify the outcome of automatic paper-reviewer assignment based on the above three approaches. The paper-reviewer assignment results are explained in the following subsections.

3.1 Experimental Results Based on Free-text

Table 1 represents the reviewer-assignment results based on free-text analysis for all 10 papers and 30 reviewers. Only the top-5 relevant reviewers for each paper are listed in the table along with their relevance values. The relevance value represents the similarity between a paper and a reviewer.

Table 1: Automatic Reviewer Assignment Based on Free-text

	Automatic Reviewers Assignment Based on Free Text with Relevancy Score														
Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy
12	18	41.14186	42	33	132.7411	30	13	100.247	73	45	42.15672	126	63	11.57207
	67	37.43856		56	58.27338		69	51.52956		57	41.9151		23	10.97387
	58	34.97052		22	55.60698		63	50.63344		63	40.06683		69	9.40585
	48	34.85569		70	41.40937		67	43.77578		56	34.37975		57	7.87127
	56	34.71021		18	28.19793		56	40.70989		22	25.6588		24	6.60767
27	63	17.54986	45	45	11.54079	62	33	68.27545	80	45	22.28067	140	67	152.8208
	23	14.699		13	11.50635		69	65.50259		57	22.05835		13	83.4708
	69	10.71853		57	11.05328		22	55.19514		22	19.97994		2	79.83838
	4	9.70572		23	10.28196		42	53.29727		63	14.56347		23	71.96418
	24	9.53715		63	9.9756		57	50.14878		68	10.91462		69	70.70271

3.2 Experimental Results Based on Keywords

Tables 2, 3, and 4 (respectively) represent the reviewer-assignment results where relevance values are calculated by means of 5, 10 and 50 extracted keywords. The empty cells represent instances where no suitable reviewers are found.

Table 2: Automatic Reviewer Assignment Based on 5 Keywords

	Automatic Reviewers Assignment Based on Five Keywords with Relevancy Score														
Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy
12	56	0.9463	42			30	57	4.49512	73	57	9.41226	126	4	7.07592
	58	0.9463					69	1.9683		2	3.0111		23	7.00772
	45	0.8483					22	1.94933		48	3.0111		63	6.24712
	67	0.8483					56	1.75098		58	3.0111		24	6.04033
	70	0.84012					48	1.75098		45	2.69927		69	2.32503
27	63	4.86.32	45	56	2.45186	62	42	6.3684	80	22	5.59304	140	48	1.75098
	4	3.53796		69	2.19794					57	3.1842		56	1.75098
	23	3.50386		59	1.50555								28	1.55452
	69	3.13082		33	1.47791								13	1.38858
	24	3.02016		64	1.34963									

Table 3: Automatic Reviewer Assignment Based on 10 Keywords

	Automatic Reviewers Assignment Based on Ten Keywords with Relevancy Score														
Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy
12	33	1.84725	42	24	2.47778	51	57	4.81481	73	57	10.93395	126	23	6.63706
	68	1.82973		22	1.68931		67	2.83135		48	2.86064		24	6.55048
	45	1.58517		69	1.68197		2	2.65461		58	1.796		4	6.49211
	58	1.46876		47	1.67117		56	2.63283		2	1.59499		63	5.87524
	56	1.37452		63	1.61412		24	1.75769		68	1.59499		69	5.58221
27	23	8.63253	45	36	3.81269	62	42	6.57848	80	22	8.4727	140	56	6.28352
	69	8.36732		2	3.47928		68	5.30922		57	3.98362		2	5.30922
	63	8.34469		70	2.21344		70	4.42688					33	3.35987
	47	6.91282		69	1.92285		69	2.81197					47	3.34234
	24	6.55048		56	1.91286								48	1.56311

Table 4: Automatic Reviewer Assignment Based on 50 keywords

	Automatic Reviewers Assignment Based on Fifty Keywords with Relevancy Score														
Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy	Paper Id	Rev. Id	Relevancy
12	57	6.05446	42	18	11.15598	30	13	12.98474	73	57	17.68021	126	69	18.78354
	59	4.6759		33	10.86462		57	10.10047		56	8.99499		63	16.75486
	2	4.07912		56	9.28441		67	7.41376		45	7.68605		23	15.50736
	13	4.04392		59	7.02499		48	6.44526		65	7.31847		24	12.74805
	18	4.00312		63	6.07723		63	6.29839		63	6.47883		4	11.20283
27	63	58.50779	45	57	10.18364	62	42	13.77097	80	22	21.88629	140	67	11.09377
	23	31.74881		45	7.00824		69	9.20802		57	10.99353		2	10.41137
	69	30.13826		23	6.96825		57	8.2346		68	8.22123		23	8.02754
	57	20.73111		70	6.89256		68	8.01763		37	6.17665		56	7.79112
	65	17.07643		2	5.32812		63	7.37857		36	5.48863		57	7.18632

3.3 Experimental Results Based on Ontology-Driven Topics Inference

In this section, we explain the experimental results of automatic reviewer assignment based on the topics inferred through the ACM CCS Ontology (Fig. 2). For instance, for Paper ID, 12, the inferred ACM CCS topics are *Software, Data, Information System, Computer Applications and Computing Milieux.* The Top-5 relevant reviewers suggested by the system are reviewers with IDs: 58, 13, 18, 67 and 2 (in order of relevance). Our manual verification of results (using human judgment) confirms that these Reviewer's profiles includes one or more of the topics of the Paper ID, 12. The inconsistency (non-overlap in the paper-reviewer assignment results) is probably due to the fact that we used the paper-reviewer assignment dataset from a very *specific* conference (an International Conference on Digital Library) which includes

specific topics (such as Social Issues in Digital Libraries, Legal Issues in DL and so on). At the same time, the ACM CCS we used cover broad Computer Science domain. Our future agenda is, therefore, to make a major revision of the domain ontology, which portrays a detailed coverage of the domain in question.

3.4 Comparison among Free-text, Extracted Keywords, and Ontology-driven Topic Inference

We have compared the results obtained from the three experiments described above (based on 10 papers and 30 reviewers' data sets) in detail. When we considered the free-text based approach for finding potential reviewers, the overlap of real reviewer assignment and automatic assignment is (66.6%) – a little better than that of the keyword-based (55%) and ontology-driven topic-based approach (55%). Such discrepancies (low overlap rate) can be further justified by the fact that our ontology is more generic (computer science) than the conference domain (digital libraries). Moreover, for some reviewers we failed to capture a sizeable amount of text reflecting his or her true profile. Nevertheless, through informal interviews with the reviewers we were also able to ascertain that most of the reviewers felt that they had been assigned one or more papers which did not conform to their preferences.

The following Table 5 represents the *true* subset of reviewer-assignment made at the conference. For the subset we used in this experiment, a set of 10 papers were assigned to a total of 29 reviewers as shown in the table. The Program-Chair rudimentarily assigned each paper to be reviewed by three reviewers, except for Paper ID 42 using available facilities in a conference management system.

Table 5: Reviewer Assignment in ADL Conference

Reviewer Assignment in Real ICADL 2005 Conference									
Paper Id	Rev. Id	Paper Id	Rev. Id	Paper Id	Rev. Id	Paper Id	Rev. Id	Paper Id	Rev. Id
12	51	42	13	50	3	73	42	126	23
	56		22		23		45		36
	58				56		57		47
27	4	45	38	62	28	80	33	140	18
	54		47		42		57		48
	59		59		57		61		61

In our experiments, we retained Top-5 relevant reviewers for each paper. The overlapping is calculated based on the criteria that at least 1 of the 3 manually-assigned reviewers in the real conference is also judged relevant by the automatic assignment methods (i.e, 5 reviewers).

4. DISCUSSIONS AND FUTURE WORK

Although the experimental results are yet to be comprehensive, when we validate reviewer assignment through unbiased human judgments, we noticed that in most cases automatic paper-reviewer assignment with the help of free-text but augmented with keyword extraction and ontology-driven topic mapping provides more relevant paper-reviewer matches. Our approach of paper-reviewer assignment is automatically done and therefore do not suffer from human bias and therefore, may improve the quality of the peer-review. We have plans to refine our ontology as well as crawling and matching approach to find a viable solution to the paper-reviewer assignment problem. We are currently working with a bigger dataset to further validate our approach explained in this paper. Furthermore, we also have plans to integrate our research results in automatic peer-review assignment with open-source Conference and Journal Management systems such as *OpenConf* or *Open Journal System* [2].

REFERENCES

[1] David Kaplan, "How to Fix Peer Review," Reprinted with kind permission of The Scientist. Volume 19, Issue 1; Page 10. January 6, 2005.

[2] OpenConf: The OpenConf Conference Management System: Available at: http://www.zakongroup.com/technology/openconf.shtml

[3] Hendratmo Widyantoro, "Dynamic Modeling and Learning User Profile in Personalized News Agent," available at: http://www.lib.umi.com

[4] Garcia, E., "Mysql implementation," Available at: http://www.miislita.com/term-vector/term-vector-5-mysql.html

[5] Dik L. Lee, Huei Chuang, Kent Seamons, "Document Ranking and the Vector-Space Model," Available at: http://www.cs.ust.hk/faculty/dlee/Papers/ir/ieee-sw-rank.pdf

[6] Ian. H.Witten, Paynter G.W., Frank E., Gutwin C. and Nevill-Manning, "KEA: Practical Autometic Keyphrase Extraction," Available at: http://www.nzdl.org/Kea/Nevill-et-al-1999-DL99-poster.pdf

[7] Sabrina Tiun, Rosni Abdullah, Tang Enya Kong, "Automatic Topic Identification Using Ontology Hierarchy," Available at: http://www.springerlink.com/content/tu4881ql8ehp0qq2/

[8] Miguel Ferreira, Ana Alice Baptista, "The use of Taxonomies as a way to achieve Interoperability and improve Resource Discovery in DSpace-based Repositories," Available at: https://repositorium.sdum.uminho.pt/bitstream/1822/873/1/paper-25.pdf

Intelligent P2P VoIP through Extension of Existing Protocols

Suchaya Sanghan and Md Maruf Hasan
School of Technology, Shinawatra University, Thailand
suchaya@scpgroups.com, maruf@shinawatra.ac.th

***Abstract* — The main propose of the *Intelligent P2P VoIP Architecture* through extension of existing protocols is to enable a VoIP user tracking and calling other users anywhere on the Net using just their e-mail IDs. We propose a pure P2P VoIP architecture which decentralizes the mechanism of call establishment and termination as well as user tracking without going through a centralized VoIP proxy or redirection server. We extend and adapted asynchronous multimedia communication protocols (namely, SMTP) and VoIP related protocols (namely, SIP) to facilitate synchronous voice communication over the Internet. Callee's location is probed intelligently using his or her SMTP server log.**

Present VoIP systems are generally based on a hybrid architecture where a SIP proxy or redirection server controls every voice call. Since all voice calls are mediated by a central server, such a system suffers from scalability and a single point of failure problems. Moreover, in such a system users are tied with a particular VoIP service provider. Decentralization of the existing hybrid architecture into the proposed architecture, explained in this report, frees the user from having to subscribe to a VoIP provider.

In this research, we focus on using open source software components to setup an enterprise VoIP system to demonstrate the proposed architecture. Nevertheless, the approach can be extended easily beyond the enterprise intranet.

***Keywords* — Internet Telephony; Voice over IP; Internet Protocols**

1. Introduction

VoIP (Voice over IP) is a rapidly expanding technology. More and more VoIP components are being developed, while existing VoIP technologies are being deployed at a rapid and still increasing pace. This growth is fueled by two goals: decreasing costs, and increasing revenues and efficiency.

Network and service providers see VoIP technology as a means of reducing their cost of offering existing voice-based services along with new value-added multimedia services. Service providers also view VoIP infrastructure as an economical base on which to build new revenue-generating services. As deployment of VoIP technology becomes widespread and part of a shared competitive landscape, this second goal will become more important, with service providers working to increase their market shares. However, this trend violates the philosophy of freedom behind the Internet. Internet users should be able to use VoIP as they have been using E-mail over the decades without having to pay to or tied with a particular VoIP provider.

1.1. Motivation

Although VoIP system becomes widespread and many VoIP providers offering services in the market, almost all existing VoIP architectures still have a common revenue generating strategy using a central Call Processing Server.

Current VoIP architecture requires user's communications to be mediated by a central VoIP server. Among others, the single server poses the threat of single-point of failure and privacy and freedom similar to those we often experience with Instant Messaging.

This VoIP architecture advertently includes a middle-man mechanism (the central server) where users are tied with the policy of the provider – payment to the provider is made directly or indirectly. Most common forms of indirect payment are advertising, and compromise of personal information, etc.

The presence of a *central server* in the VoIP architecture is probably due to capturing CDR (Call Detail Record) information for billing purposes. Technically speaking, VoIP calls can be made without having to go through such a central server. We attempt to eliminate such central server in the proposed VoIP system in this paper[1].

1.2. Our Approach

The oldest Internet application, E-mail is based on SMTP protocol. The SMTP protocol only supports text based E-mail. For asynchronous multimedia communication (e.g., multimedia E-mail), the Internet Engineering Task Force (IETF) proposed MIME extensions rather than modifying the original SMTP protocol. The IETF subsequently proposed Session Initiation Protocol (SIP; and RTP, RTCP etc.) to

[1] Deployment of IPv6 may offer more freedom for Internet-based communication.

facilitate asynchronous multimedia communication (such as, VoIP and Video Conferencing). The IETF is well-known for its philosophy of freedom and adherence of non-proprietary but innovative protocols and standards. Unfortunately, due to the commercially motivated implementation of present-day VoIP services, we failed to enjoy synchronous multimedia communication (e.g., VoIP) as freely as we have been enjoying the asynchronous multimedia (e.g., E-mail).

In our work, we tried to analyze relevant Internet protocols and implemented a VoIP system where we can completely bypass the revenue-generating legacy of the current VoIP frameworks where users need to go through a VoIP provider to make Internet calls. We extended existing protocols and implemented light-weight SIP client and server modules to implement a pure P2P VoIP system as explained in the rest of the paper.

To build the system that decentralizes the existing VoIP architectures into a pure peer-to-peer VoIP architecture, we need to eliminate the Central VoIP Call Processing Server (Figure 1) by having both the SIP client and server modules (a modified version) on end-user's computers. Intelligent use of existing SMTP server (Figure 2) is used to infer Callee's IP address heuristically from SMTP server-logs. We adopted open-source VoIP suites [1][2] for our purpose and developed required components on our own [3] to implement and test our idea.

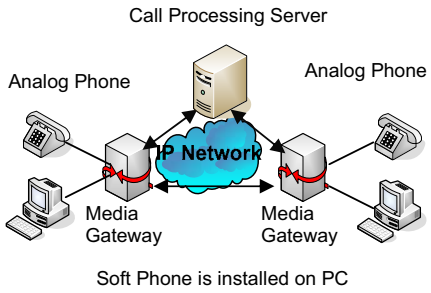

Figure 1. Existing VoIP System - Hybrid Architecture

Figure 1 shows the existing VoIP architecture; whereas, Figure 2 explains the proposed architecture which decentralizes the Call Processing Server by having a modified SIP client and server installed on each user's computer and by making intelligent use of Callee's SMTP server-log to obtain location information to initiate a VoIP call. Caller consults the Callee's SMTP server to obtain plausible IP addresses of the Callee. Callee's location is probed heuristically using IP addresses and timestamps obtained from SMTP server logs on the fly as explained in Figure 2. Call parameters are then extracted from the SIP server running on the Callee's computer.

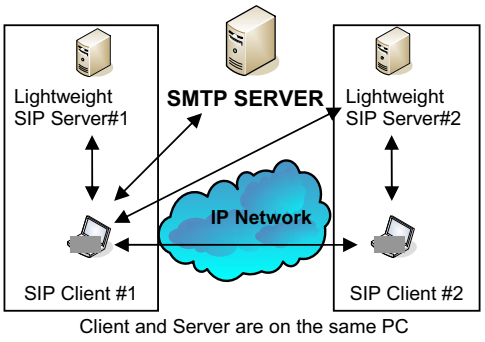

Figure 2. Proposed VoIP System – Pure P2P Architecture

2. Overview of the Proposed VoIP System

In this section we explain the proposed VoIP system with specific scenarios.

2.1. System Architecture

Our proposed VoIP system architecture decentralizes the Call Processing Server by implementing a light-weight SIP Client and Server module on each end. On each user's terminal, we have both SIP client (Softphone) and a modified SIP server (Call Processing Server) installed which is quite similar to SIP Redirect mechanism [7]. The modified SIP client on the Caller's side contacts Callee's SMTP server and subsequently contacts the Callee's SIP and RTP to locate connection parameters, and finally initiate a call. This architecture essentially eliminates the need of a VoIP provider. Our approach works with an assumption that users use their computers to check e-mail periodically; and we can heuristically gather information about the user's location by intelligently traversing the SMTP server-log. Please note here that the called party is uniquely identified by his or her E-mail address. SMTP logs give location information (IP, Time stamps, etc.); and finally, a call is initiated directly between the peers using SIP handshaking and RTP/RTCP based communication. Figure 2 is a simplified diagram of the proposed VoIP architecture. The proposed architecture is further illustrated below using 3 specific scenarios.

2.2. Three scenarios of the proposed system

(1) Standby

A user needs to run our program to sign into his or her terminal. The VoIP system (Call Processing Server) remains standby and listens to incoming call requests (Figure 3). This scenario is very similar to Online Chat but a user is logging into the server running on his or her own computer (not a remote commercial server).

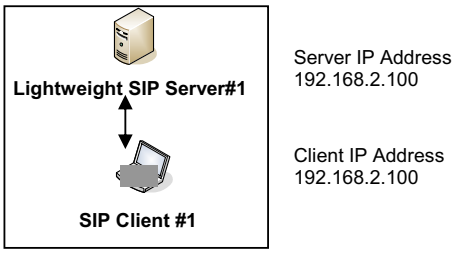

Figure 3. Standby Scenario

(2) Callee available at the last-known (most recent) IP address

From the last scenario, when user signs up and in standby mode, user will get his own IP address in that location. When user changes the location from Faculty#1 to Faculty#2, the user's IP address will be changed on both SIP Client and Light Weight SIP Server. Different locations, different IP addresses. SMTP Server will know the user's IP address when user logs into the SMTP server (Email Checking). User's information such as IP address, Timestamp, Name, and etc. are kept in SMTP logs of the SMTP Server. In a normal situation, most recent IP address from SMTP logs is always the last location where the user is most likely to be present. Caller can get the most recent IP address of Callee from the SMTP Server by retrieving selected information from the SMTP logs. With the help of timestamp, we identify the most recent IP that the user used. When Caller knows the Callee's IP address from SMTP logs, Caller probes the Callee's Call Processing Server (running on the destination's PC) using the IP address that Caller determined. A SIP session establishment with the Callee may subsequently be successful. We explain this scenario with a sequence diagram as follows:

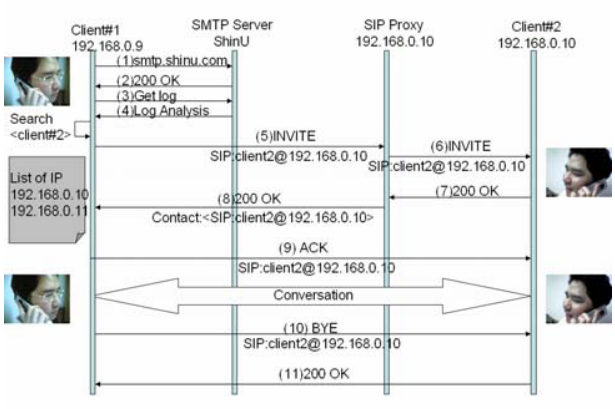

Figure 4. Sequence diagram for "Callee available at the last-known IP address"

From Figure 4, User#1 wants to call User#2. User#1 consults directly to the User#2's SMTP Server (smtp.shinawatra.ac.th)[2] to get the User#2 information from SMTP logs about IP addresses and Timestamp of User#2. Those information can identify the last location and most recent IP address that User#2 logged in - this is the most likely location where the Callee might be available. After User#1 gets the most recent IP address from SMTP logs, User#1 uses the most recent IP address (192.168.0.10) that User#2 used to login into the SMTP Server to establish the connection between both of them. User#1 sends the **INVITE** message to the SIP Server of User#2 and then SIP Server of User#2 forwards the **INVITE** message to the User#2's SIP phone. When User#2 picks up the phone, acknowledgement from User#2 will automatically be acknowledged back to User#2's Server and back to User#1. User#1 sends acknowledgement directly to User#2 to confirm. Both of them establish a voice connection (RTP session). After finishing a call, User#1 sends **BYE** message back directly to User#2 to terminate the connection. User#2 acknowledges back directly by **OK** message to User#1 to confirm the connection termination.

(3) Callee not available at the most-recent IP or Callee logged in from multiple locations

In the last scenario (2), the Callee was available at the last-known location. However, it may not be case all the time. There are also possibilities that a user is concurrently checking mail from multiple locations. In such a case, we use a *Trial and Error* –based polling. The most recent IP address in the SMTP logs will be chosen first when the Caller wants to call as it should be the normal situation.

Callee may not reply when the Caller calls to the most recent IP address. Our System will iteratively probe the next possible location (the next IP address based on timestamp) where the Callee may be available. The process will continue until a response is received or for all known IPs. That is,

when the Callee responses, a SIP Session is initiated. Otherwise, the Caller will keep trying to probe the next location iteratively. This scenario is explained below with a sequence diagram with 2 locations for the Calee. .

[2] A script is written for the SMTP server which periodically dump {user, IP, timestamp} info for our use.

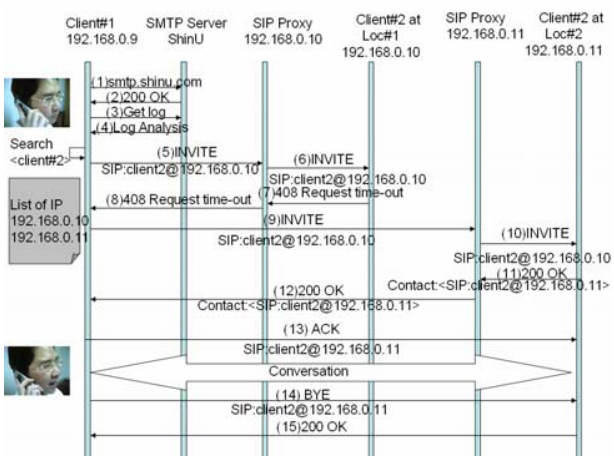

Figure 5. Sequence diagram for "Callee logged in at 2 distinct locations"

From Figure 5, User#1 wants to call User#2. User#1 consults directly with the User#2's SMTP Server (smtp.shinawatra.ac.th) to get the User#2's information from SMTP logs about IP addresses and Timestamp of User#2. Those information can indicate the most recent location and most recent IP address that User#2 logged in - this is the most likely location where is the Callee might be available. After User#1 gets the most recent IP address from SMTP logs, User#1 uses the most recent IP address (192.168.0.10) that User#2 used to login into the SMTP Server to establish the connection between both of them. User#1 sends the **INVITE** message to the SIP Server of User#2 and then SIP Server of User#2 forwards the **INVITE** message to the User#2's SIP phone. User#2 does not pick up the phone in the pre-defined time, User#2 will automatically send **Request Timeout** message back to User#2's Server and back to User#1. After User#1 knows that User#2 is not available at the first location, User#1 tries the next IP address that User#2 used to login. User#1 sends the **INVITE** message to the SIP Server of User#2 at another location (192.168.0.11) and then SIP Server of User#2 forward the **INVITE** message to the User#2's SIP phone. When User#2 picks up the phone, acknowledgement from User#2 will automatically be acknowledged back to User#2's Server and back to User#1. User#1 sends acknowledgement directly to User#2 to confirm. Both of them establish a voice connection (RTP session) assuming that the Callee is available at Location#2. After finishing a call, User#1 sends **BYE** message back directly to User#2 to terminate the connection. User#2 acknowledges back directly by **OK** message to User#1 to confirm the connection termination.

Nevertheless, since the Caller knows all the Callee's IP addresses from the Callee's SMTP logs, Caller can easily send probes to *all* Callee's locations by sending the **INVITE** message to all locations *in parallel*. The Cellee's location that Caller gets the response will be the location where Callee is. Caller will also get the **Request Timeout** message from other locations. Caller does not need to poll each location one by one sequentially, but can broadcast the request to all locations at one time.

2.3. Implemented System at Work

In our experiment, we used open source software Asterisk and X-Lite [1][2]. We modified the open source modules as necessary and developed necessary modules to make them work together under the proposed architecture.

The entire source code is available with detail installation guideline from our web site:

http://webmail.padawan.shinawatra.ac.th/~it601/ohm.

Followings are some screen-shots and analysis on how the proposed P2P VoIP system works.

(1) Initialization and Signing-in

Start up the program first and then click "Open Phone button" to turn on the SIP Softphone and SIP Server (Call Processing Server). The SIP phone will automatically login into its server (SIP Server on the same PC) and wait for any incoming call. See from Figure 6.

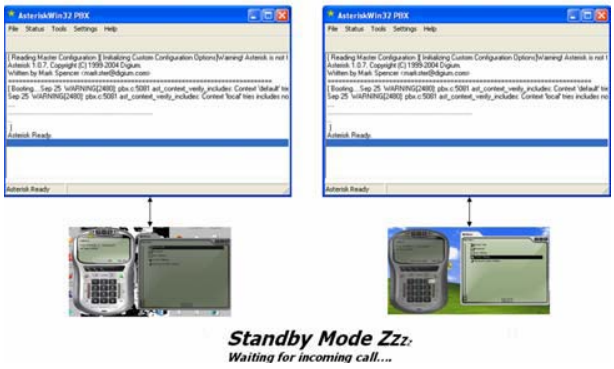

Figure 6. Standby mode of proposed system

(2) Calling and Termination

When a user wants to place a call to others, he or she specify the Callee's E-mail ID and clicks "load button" to load the log file from SMTP Server and find information about the Callee's location as shown in Figure 7.

Figure 7. Search for the IP addresses of destination

User click "Start IP Phone button" to call to the destination by using the first IP address on the row that is the most recent login IP address of the receiver.

This scenario is shown in Figure 8. *Client#1* is the one who wants to call *Client#2*. So, he logins to *Client#2*'s Server (Suchaya) by using address 61.90.12.151 and makes a call to *Client#2*.

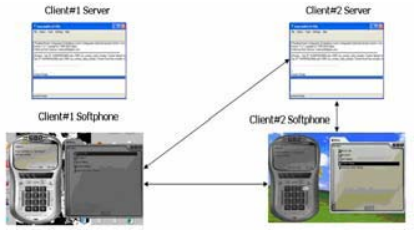

Figure 8. Client#1 calls Client#2

If the destination is not available at the most recent IP address or login on many places, *Client#1* will try to connect to the destination by using the IP addresses from the list, one by one, until he gets the acknowledge back from the destination like the Figure 9.

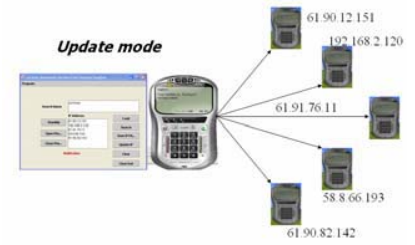

Figure 9. Update mode to search the location that is available

When the user finishes a call, they just click at the "Standby button" to login into their own server and wait for incoming call like Figure 6. When they want to turn off the system, just one click at the "Close phone button", SIP Softphone and SIP Server will be automatically shut down.

3. Conclusions and Future Work

We made use open source components to set up an enterprise VoIP system and made necessary modifications to existing protocols to demonstrate a middle-man free pure P2P VoIP framework.

Unlike asynchronous multimedia communication (such as, multimedia E-mail), due to huge commercial interests synchronous multimedia communications (such as VoIP) have been trapped into a provider-dominated service. In this paper, we explained how to avoid such middle-man scenario in VoIP to achieve a pure P2P VoIP framework. Our experiments did not address some crucial issues such as NAT and Firewall etc. We are also aware that people may be reluctant to allow access to SMTP logs due to privacy concerns. We only tried to demonstrate that we can use modifications and extensions of existing protocols to make VoIP a middle-man free and ubiquitous service for all Internet users.

The proposed approach has significant opportunity for revenue generation through hardware-based implementation such as manufacturing IP phones using open source OS kernel, VoIP components along with the approach explained in this paper.

ACKNOWLEDGMENTS

This work is supported by Thailand Research Fund (TRF Grant: MRG4880112) awarded to Dr. Md Maruf Hasan. He wishes to thank his mentor, Dr. Ekawit Nantajeewarawat, and his colleague, Dr. Chutiporn Anutariya for their valuable comments.

REFERENCES

[1] Mark Spencer (2005). The Open Source PBX. Digium, Inc. http://www.Asterisk.org (accessed November 11, 2005).

[2] Xten Networks, Inc. (2004). Download X-Lite Softphone. Xten Networks, Inc. http://www.xten.com/index.php?menu=download (accessed October 29, 2005).

[3] Suchaya Sanghan (2005). Intelligent P2P VoIP through Extension of Existing Protocols. Shinawatra University. http://webmail.padawan.shinawatra.ac.th/~it601/ohm.

[4] Digium, Inc. (2005). Tutorials. Digium, Inc. http://www.asteriskguru.com/ (accessed November 12, 2005).

[5] Sun Microsystems, Inc. (1994-2006). Downloads. Sun Microsystems, Inc. http://developers.sun.com/resources/downloads.html (accessed December 10, 2005).

[6] Foundry Networks, Inc. (2005). Foundry Voice Over IP. Foundry Networks, Inc. http://www.foundrynet.com/solutions/voip/index.html (accessed October 28, 2005).

[7] Gonzalo Camarillo, SIP DEMYSTIFIED, McGraw-Hill Publishing, 2001.

[8] M. Handley, H. Schulzrinne, E. Schooler, J. Rosenberg (1999). RFC 2543 Session Initiation Protocol. SunSITE.dk. http://rfc.sunsite.dk/rfc/rfc2543.html (accessed October 29, 2005)

[9] U. Varshney, A. Snow, M. McGivern, and C. Howard, "Voice over IP", Communication of ACM 45 (1), January 2002, pp. 89-96.

[10] S.R. Ahuja and JR. Ensor, "VoIP: What is it Good for?", ACM Queue vol. 2, no. 6, September 2004, pp. 48-55.

[11] A. Singh, P. Mahadevan, and A. Acharya, "Using session initiation protocol to build context-aware VoIP support for multiplayer networked games.", ACM Press New York, August 2004, pp. 98-105.

[12] Y. Amir, C. Danilov, S. Goose, D. Hedqvist, and A. Terzis, "1-800-OVERLAYS: using overlay networks to improve VoIP quality.", NOSSDAV '05. ACM Press, June 2005, pp. 51-56.

From Collaborative Video Library to Annotated Learning Object Repository: Using Annotated Video Library in Personalized E-Learning

Md Maruf Hasan[1], Nophadol Jekjantuk[1], Yenruedee Chanwirawong[1], Ekawit Nantajeewarawat[2]
[1]Shinawatra University, Thailand
[2]Sirindhorn International Institute of Technology, Thammasat University, Thailand
maruf@shinawatra.ac.th; nophadol_bkk@hotmail.com; yenruedeec@isd.th.ibm.com; ekawit@siit.tu.ac.th

Abstract – **With the proliferation of digital and video cameras, personal collection of multimedia materials such as amateur video-clips are abundant now-a-days. Most of these multimedia materials may be useful to others if they are shared and can be located easily. Semantic Web technologies hold promise to organize and re-use such non-textual information effectively. However, annotation of multimedia contents is a tedious task. Notwithstanding, as we observe growing number of community collaborations on the present Web, such content annotation can be done through online collaboration. In this research, we investigate the development of a collaborative video annotation system using open technologies and tools where people can upload, annotate and share their personal multimedia collections efficiently. We also examine how the contents acquired and annotated with this collaborative system can be transformed into (1) Reusable E-Learning Contents for personalized learning; as well as, (2) Dynamic Digital Library with exploratory search and retrieval facilities using state-of-the-art Semantic Web technologies.**

I. INTRODUCTION

Lecture materials such as those available under the MIT's Open Course Ware (OCW) initiative inherently consists of multimedia contents including video lectures and other non-textual objects [1]. Due to increased use of WWW in teaching and learning, e-learning contents have been increasingly available on the Internet. In this paper, we will use foreign language pedagogy and other intuitive (but simplified) examples to demonstrate the effectiveness of using collaborative annotation and Semantic Web technologies for sharing and re-using E-learning contents for personalized learning.

Videos play an important role in developing foreign language skills and apprehending foreign culture. Video clips for situational dialogue such as, shopping, dining, and opening a bank account etc. are highly reusable. For instance, a video-clip depicting the scenario of opening a bank account in English may easily be reused for learning French, and vice versa. Agencies involved in promotion of cultural exchange and tourism usually produce vast amount of multimedia materials. However, most of these materials are often available with a few main-stream languages. Therefore, people without any knowledge in such languages fail to make better use of such materials. By adding metadata, transcription and subtitle to video contents, we can facilitate broader use of multimedia materials. For instance, a documentary video about Thai Folk Dance is commonly available in Thai, English or some other main-stream languages. However, with metadata annotations, transcription and subtitle in other languages; such materials can be equally useful for people who are not proficient in main-stream languages.

Over the last decades, with the ubiquity of the Internet and the growing popularity of the Web, we have witnessed growing number of successful small community collaborations on the WWW [2]. The success of such collaborations largely depends on the development of a sophisticated system (a collaboration platform) with easy-to-use interface and tool-support. In this paper, we investigate the development of a *Collaborative Video Annotation System* which can facilitate collaborative learning of foreign language and promotion of cultural exchange. We argue that with the increasing availability of Internet and bandwidth, digital divide is not always due to technological barrier but often due to our inability to integrate human sharing spirit and goodwill along with the technological advances.

II. SCENARIO OF FOREIGN LANGUAGE LEARNING

Main-stream foreign language learning materials have a huge market and therefore, publishers worldwide have been actively publishing study materials for the target market. Main-stream foreign language such as English has plenty of learning materials available in many native languages. However, it is truly unfortunate (and undesirable) that a Vietnamese often needs to learn Thai using learning materials written in English, and vice versa. The market-driven phenomenon left some languages far behind from the others. Our language and culture may or may not have a feasible market but they are our unique assets, and we ought to find ways to promote them using enabling technologies. By successfully utilizing the potential of the Internet and human spirit of collaboration, it is possible to promote our language and culture in an online collaborative fashion. The development of our video annotation system is inspired by

M. Iskander (ed.), Innovations in E-learning, Instruction Technology, Assessment, and Engineering Education, 255–260.

language education and cultural exchange by making use of the potentials of ubiquitous Web and online virtual collaboration.

Unlike some technological disciplines, foreign language teaching and learning follows a typical pattern and rather static. Off-the-shelf foreign language learning kits generally consist of lessons organized into different proficiency *levels.* For instance, the *Beginner's Level* consists of lessons with basic alphabet-set, phonetics and writing system along with greetings, basic vocabulary, simple grammars and conversations. The *Intermediate Level* materials go beyond the basics and include further lessons and practices with the help of situational dialogues and further syntactic, semantic and pragmatic notes and drills. *Advanced Level* learning materials are less restrictive and often go beyond typical textbook lessons – it is not uncommon to use real newspaper articles, movies, TV news or documentaries to teach different linguistics and cultural facets to the advanced learners. In many cases, we noticed that a vast majority of these off-the-shelf software products are mere digitization of their predecessors - the video-based learning kits. However, there are some exceptions. For instance, ELLIS [3] – an English learning software tries to integrate ESL pedagogical research with computer and networking technologies. Such materials are often expensive and beyond the purchasing power of individuals.

In terms of contents, we noticed many similarities across foreign language learning materials. Unless commercially motivated, contents such as alphabet chart, pronunciation audio files, grammar flashcards, situational dialogue videos could have been reused effectively for teaching that foreign language to any native language community through some customization and annotation.

Screenplay [4], a Japanese publisher uses foreign films to help Japanese learners improve their foreign language ability. Using their Web site and multimedia products, a user can also conduct associative search to see a word in the real context. This is not only effective as endorsed by scholarly research findings [5], but also a fun activity which promotes the understanding of culture, etiquette and social practice beyond mere language learning through memorization of grammar rules, substitution drills, and the like.

III. Scenario of Cultural Exchange

Governments and agencies across the world do invest huge amount of money and efforts in producing documentaries and other multimedia materials for the promotion of culture and tourism. However, their focus remains limited to some targeted audiences (and therefore, limited to a few main-stream languages). For instance, most documentaries and videos developed in Japan are often available in Japanese, English, French, and Chinese, etc. Unless properly annotated with metadata, transcripts, subtitle or soundtrack, it is unlikely to find those materials useful for an audience who do not understand any of these languages.

With our decades of experience with other online collaborative projects, such as Project Gutenberg [6] and Aozora Bunko [7] - where copyright-free books are digitized and even translated by volunteers worldwide, we expect that such promotional video materials can equally be annotated collaboratively and free-of-charge as long as an easy-to-use collaborative system is available and online volunteers are supported with easy-to-use interface and tools.

IV. Copyright and Other Issues

It is a fact that commercially developed materials may not be available free, but it is inevitable that some promotional materials from *not-for-profit* agencies and governments may gradually become available publicly. We also rely on amateur personal videos taken by digital and video cameras as a good starting point. The Open Video Project [8] is a general digital library of publicly available videos from unrestricted domains – ranging from classroom lectures to public service or documentary videos – may also be imported and annotated in our system. However, at the moment, we target to gather and annotate video and multimedia materials in a *restricted domain* through small-community collaboration. Our primary goal is to promote cultural exchange and facilitate foreign language education while making full use of the ubiquitous connectivity of the Internet and the growing human spirit of virtual collaboration on the WWW. In doing so, we focus on developing an easy-to-use collaborative video annotation system which we have explained in this paper in detail. We also refer to our other publications on creation of sharable and reusable E-Learning materials using a component-oriented approach [9] – where we demonstrated that such annotated contents with proper metadata, and relevant structural, semantic and pedagogical information, it is possible to generate customized and personalized course materials for *any* discipline (not restricted to foreign language learning).

In any voluntary online collaboration, *motivation* plays a crucial role. The success of our initiative therefore, remains on the enthusiasm of volunteers who take pride and care about their own language and culture. The evidence of the steady growth of collaborative projects on today's Internet and WWW (e.g., the Open Video Project as mentioned earlier) is encouraging. Moreover, a collaborative project such as this one requires a *critical-mass* of volunteers and contents to become successful. It is obvious that unless we have a handful number of volunteers or a sizeable amount of contents, our initiative may not receive sufficient attention quickly.

With an initial support from *Asia-Pacific Telecommunity* (APT) and *Thailand Research Fund* (TRF), we have initiated this project and developed a fully working prototype which is currently operational at http://apt.shinawatra.ac.th/video/. Some of our international collaborators are consistently supporting us by developing or accumulating useful contents for promotion of foreign language and culture on the cyberspace. We envisage that we will continuously receive supports from governments and agencies to develop this project further.

In the rest of the paper, we will explain the implementation details and major features of this video annotation system followed by their effective application in E-Learning.

V. The Collaborative Video Annotation System

The video annotation system we developed is fully Web-based. Users interact with the system using their Java-enabled Web-browsers without having to resort to any video plug-in. We choose *MediaFrame* [10] – an open-source software for streaming video manipulation and playback. We use MediaFrame's JavaScript API, PHP scripts and MySQL database to develop our video annotation system. The technical complexities of the system remain hidden to the users since users interact with the system using an easy-to-use Web interface.

A. MediaFrame: Streaming Media Player

MediaFrame is an open-source streaming media platform in Java™ which provides a fast, easy to implement, and extremely small applet that enables web users to view audio/video contents without having to rely on external player or plug-ins. MediaFrame does not require special servers, software or programming knowledge. Integration of MediaFrame with other applications becomes easy due to its well-defined JavaScript API and its support to Mpeg-1 and Mpeg-4 video formats.

MediaFrame can deliver audio and video contents over the Internet in either streaming or progressive download mode. It is also capable of stretching media by up to 60% of its original size without a significant loss of quality. This enables us to simulate higher bit rates without the associated costs and to operate effectively in both broadband and narrowband environments. We want our system to be reachable by people in the remote area where the infrastructure is not as good, or may have to resort to satellite links for a connection. MediaFrame is capable of detecting user's connection speed in real-time and delivers media accordingly.

A full-featured *Digital Rights Management* (DRM) system is also integrated with the core of MediaFrame allowing media files to be locked to a specific domain and expired over time. This feature is crucial in disseminating copyrighted/restricted contents.

Moreover, MediaFrame is fully JavaScript controllable; making it simple to build our own control set and to integrate it with our video annotation system seamlessly. Using the JavaScript API, we could also conveniently add transcripts and subtitles to video materials. The multiple-playback-state feature also allows us to tightly integrate MediaFrame for effective video annotation and playback.

B. Metadata Annotation

For non-textual materials, such as videos, descriptive metadata plays an important role in exploratory search and browsing [11]. MPEG-7, LOM and SCROM [12] are sophisticated metadata sets for multimedia content annotation. However, we adopt a simpler subset of descriptive metadata for the system in order to avoid fears in the ordinary users. At the moment, the system allows users to annotate a video-clip with *Title, Genre, Original-Language, Keywords, Descriptions* and some other essential metadata. Whenever applicable, structural metadata such as *Video Format, Duration, Color,* etc. are extracted automatically from the video file.

C. Supplementary Annotations

The video annotation system allows a user to step through a video-clip and to add transcripts along with its original soundtracks and transcripts. User can also select transcript in an available language and translate it into a target language. During playback, users can choose from available transcription-languages to be displayed as subtitle with the video (cf. Fig 1-4).

It should be noted that for effective learning of foreign language using video, transcriptions need to be further annotated with extra information such as, pronunciation and grammatical annotations or relevant cultural notes. Such features are currently being implemented.

D. Some Representative Video-clips and Their Annotations

In this subsection, we explain a series of video-clips from our system, which summarize the effectiveness of our system in the context of language learning and cultural exchange.

Fig. 1 shows a situational dialog video-clip from ELLIS played with English (Fig. 1a) and Thai subtitle (Fig. 1b). Fig. 2 shows amateur video-clips of Kashima Jingu Shrine in Japan (Fig. 2a), and about Japanese Paper Art (Origami, Fig. 2b) – although the original soundtrack is in Japanese, English subtitles are made available through online annotation, and displayed accordingly. Fig. 3 shows a video tutorial from JEITA about *Natural Language Processing* (NLP) technology (originally with Japanese soundtrack) but is being played with English subtitles since such information has already become available through online annotation. In doing so, we made it possible for non-Japanese audience to benefit from this tutorial

For all these video-clips, descriptive metadata are also available in the system (annotated collaboratively and indexed automatically) to facilitate exploratory search and navigation [11]. Users can use keyword search (analytical strategy) or browse (partition strategy) the video collection effectively with the help of relevant metadata.

Fig. 1a. Situational dialogue in a bank (played with English subtitle).

Fig. 1b. Situational dialogue in a bank (played with annotated subtitle in Thai).

Fig. 2a. Amateur video-clip: Kashima Jingu Shrine in Japan (original soundtrack in Japanese but played with annotated English subtitle).

Fig. 2b. Amateur video-clip: Japanese Origami or Paper Art (original soundtrack in Japanese but played with annotated English subtitle).

Fig. 3. A video tutorial developed by JEITA on *Natural Language Processing* – originally in Japanese, but we added English subtitles and descriptive metadata online.

Finally, Fig. 4 shows an example video annotation interface. Video is displayed with playback control, time-stamp and other information. In the process of filling in the *Time* and *Subtitle* boxes below, annotating user can choose reference subtitle from available source languages, and control the video playback with control buttons as necessary. All other annotation interfaces are similar browser-based, and therefore, easy-to use.

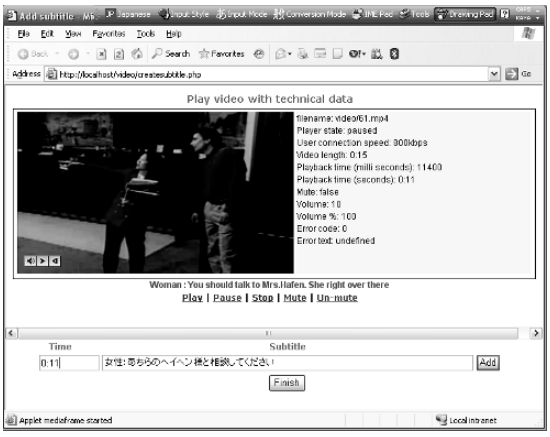

Fig. 4. Subtitle annotation interface: User can choose a source language and add subtitle in a target language. Timestamp and Playback Control are displayed as video-clip is being played.

VI. Applications of Annotated Video Library

The video annotation system serves as a *platform* for *small community collaboration* online as well as an *input channel* of contents. Multimedia contents gathered, annotated and manipulated in this way should be used (disseminated) in useful applications. The following two subsections explain

two unique applications in personalized E-Learning and dynamic Digital Library.

A. Intelligent Dissemination of E-Learning Contents

The success of gathering and disseminating sharable, reusable and customizable e-learning contents depends on developing an easy-to-use collaborative system – a system that not only hides the complexities behind a simple form-like web interface (for submission, annotation, retrieval, and the like), *but also* organizes and manipulates contents in a structured and efficient manner so that intelligent inferences can also be made. For example, content-level dependencies should be propagated and preserved at lesson or course level and so on. Semantic Web technologies including metadata and domain-ontology are used in the background so that both humans and software agents can equally effectively access and manipulate the annotated contents from the video annotation server.

We proposed a *3-Tier Architecture* for component-oriented E-Learning Content Management [9] as shown in Fig. 5. We use ontology to organize contents and ontology-based reasoning to make inferences about contents such as content dependency or other pedagogical attributes. Given that we have gathered and annotated contents with their pedagogical attributes, our current E-Learning prototype can respond to user's content needs intelligently. For instance, finding contents for *"10-hour* lessons for *introductory* Japanese focusing on *greetings and everyday conversation",* or a "*42 hours* of course materials for a *System-oriented* Database course*"* - can be created on-the-fly as long as annotated contents are available.

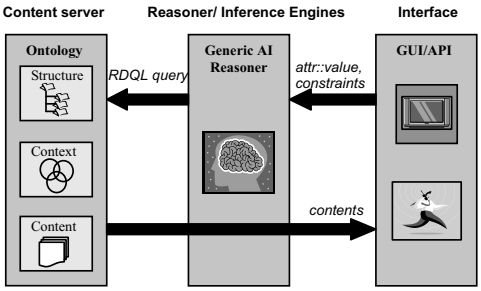

Fig. 5. *3*-Tier E-Learning Content Management Architecture.

A currently working prototype is available at http://apt.shinawatra.ac.th/ecms/. Internally, we organize the contents in an OWL-ontology based on the annotated metadata. Our current prototype uses *a Semantic Web Reasoner* (OWL-based Pellet, [13]) to identify appropriate contents from the ontology that satisfy user's criteria. Users interact with the system using Web-based interface. Fig. 6a and Fig. 6b explain how user may specify different constraints and find appropriate contents that satisfy their criteria and content dependency from a *Database* related course contents.

For example, a teacher or a learner trying to retrieve a *20-hour Foundation Course in Database* only needs to specify the criteria through the Web-based interface (Fig. 6a). User criteria and content dependencies are then verified using a G*eneric AI Reasoner* to locate the appropriate contents. The user is then presented with the appropriate contents in an organized fashion (Fig. 6b).

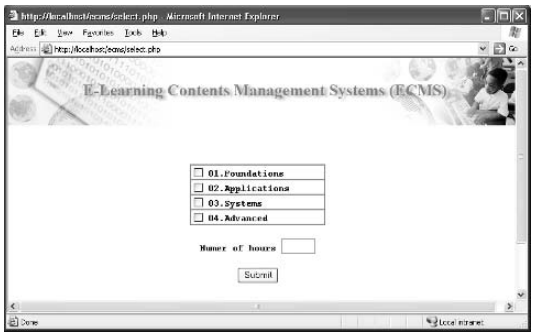

Fig. 6a. Search interface of E-Learning contents (example shows contents search for a *Database* course contents).

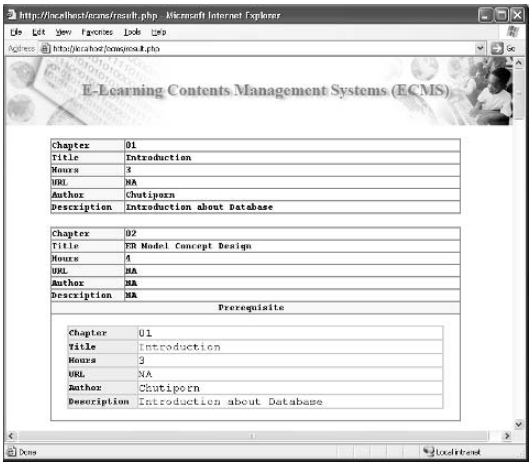

Fig. 6b. E-Learning contents retrieved and presented in an organized fashion Output reflects the criteria/constrains specified by user as well as content-level dependencies.

B. Dynamic Digital Library

Using Semantic Web technologies and exploratory search and browsing interfaces as explained in [11] and tested on Open-Video collection [8], we plan to disseminate our multimedia-contents arriving from the video annotation system (i.e., live *input channel*) in the form of *Dynamic Digital Library*. For example, a user interested in Temples and Pagodas in a particular region will define a *Library Template* by specifying the nature of contents (such as, video-clips of Temples and Pagodas in Thailand can be easily specified since our contents are annotated with descriptive metadata and organized in an ontology). Contents and their organizations in such dynamic libraries are automatically updated as new

materials are deposited or contents are updated, annotated or reorganized on the video annotation server.

VII. CONCLUSIONS

In this paper, we outlined an operational prototype of a collaborative video annotation system developed for foreign language education and cultural exchange in mind. We often ignore the fact that majority of Internet users are not proficient in foreign language and therefore, materials available on the WWW in a foreign language has little or no use to those people. Multimedia materials, such as videos are relatively easy to understand or appreciate without *complete* translation (since contextual, visual and other cues are present). Nevertheless, it is extremely difficult to locate non-textual materials on the Web using today's keyword-dependent *hit-or-miss* search engines which heavily rely on textual indexing. Metadata annotation is therefore essential and can be done effectively in small-group collaboration. We do admit that we are yet to conduct any experimental evaluation of the proposed system to justify the effectiveness of this approach. However, our future work will eventually attempt such evaluations and vigorous live trials.

We have also outlined an ontology-based *E-Learning Content Management* prototype for *personalized* E-Learning. We have plans to organize our video collections using Semantic Web technologies to support personalized foreign language education. We explained how our annotated video library with its multimedia contents can serve as (1) an *E-Learning Content Server* - storing both contents and structures (contents with annotated pedagogical attributes in an ontology) to facilitate personalized E-learning; and, (2) a *Digital Library Content Server* - storing amateur videos as well as professional multimedia (contents with annotated language and cultural attributes in an ontology) to support cultural exchange and foreign language learning tasks.

ACKNOWLEDGMENT

This research is supported by *Thailand Research Fund* (grant No. MRG4880112) and by *Asia-Pacific Telecommunity* (APT-HRD grant 2005).

REFERENCES

[1] MIT Open Course Ware Project: http://ocw.mit.edu/

[2] Johnson, C.H., "A survey of current research on online communities of practice," Internet and Higher Education Vol.4, pp.45-60.

[3] ELLIS Inc.: http://www.ellis.com/

[4] Screenplay: http://www.screenplay.co.jp/

[5] Cordilo, D.S. (1997). Using a Foreign Film to Improve Second Language Proficiency: Video vs. Interactive Multimedia, Journal of Educational Technology Systems, Vol.25 no.2 pp.169-77.

[6] Project Gutenberg: http://promo.net/pg/

[7] Aozora Bunko Project: http://www.aozora.gr.jp/

[8] Open Video Project: http://www.open-video.org/

[9] Hasan, M.M., S. Yamamoto, Y. Fujino and W. Chujo, "Towards Sharable and Reusable Contents using Semantic Web Technologies for Personalized Foreign Language Learning," IEICE General Conference 2006, Tokyo, Japan; D8: Artificial Intelligence and Knowledge Processing Track - Paper No.: D-8-1; pp.84, 2006

[10] MediaFrame: http://Mediaframe.org/

[11] Marchionini, G., "Exploratory Search: From Finding to Understanding," Communications of ACM, Vol.49, No.4, pp.41-46.

[12] Godwin-Jones R., "EMERGING TECHNOLOGIES - Learning Objects: Scorn or SCORM?" Language Learning & Technology, Vol.8, No.2, pp.7-12

[13] Pellet Reasoner: http://pellet.owldl.com/

Building a Collaborative Annotated Video Library for Language Education and Cultural Exchange

Md Maruf Hasan[1], Ekawit Nantajeewarawat[2], Thanh Tung Nguyen[3], Wataru Chujo[4]
[1]Shinawatra University, Thailand
[2]Sirindhorn International Institute of Technology, Thammasat University, Thailand
[3]Radio-Electronic Association of Vietnam
[4]National Institute of Information and Communications Technology, Japan
maruf@shinawatra.ac.th; ekawit@siit.tu.ac.th; cst.rev@fpt.vn; chujo@nict.go.jp

Abstract:

With the proliferation of digital and video cameras, personal collection of multimedia materials such as amateur video-clips are abundant now-a-days. Most of these multimedia materials may be useful to others if they are shared and can be located easily. Semantic Web technologies hold promise to organize and re-use such non-textual information in a structured fashion. However, annotation of such multimedia contents is a tidy task. Notwithstanding, on the present day world wide web, small community collaborations are on the rise. In this paper, we investigate the development of a collaborative video annotation system using open technologies and tools where people can upload, annotate and share their personal multimedia collections with ease. Support for *exploratory search* is also available for users to locate desired contents. We also explain how such an annotated video library would promote language education and cultural exchange.

1.0 Introduction

Multimedia contents such as videos play an important role in developing foreign language skills and apprehending foreign culture. Video clips for situational dialogue such as, shopping, dining, and opening a bank account are highly reusable. For instance, a video-clip depicting the scenario of opening a bank account in English may easily be reused for learning French, and vice versa. Agencies involved in promotion of cultural exchange and tourism usually produce vast amount of multimedia materials. However, most of these materials are often limited to a few main-stream languages. Therefore, people without any knowledge in such language fail to make better use of such materials. By adding metadata, transcription and subtitle to video contents, we can facilitate broader use of multimedia materials. For instance, a documentary video about Thai Folk Dance is commonly available in Thai, English or other main-stream languages. However, with metadata annotations, transcription and subtitle; such a material can be equally useful for people who are not proficient in those languages.

Over the last decade, with the ubiquity of the Internet and the growing popularity of the Web, we have witnessed growing number of successful small community collaborations on the WWW (Johnson, 2003). The success of such collaborations largely depends on the development of a sophisticated system (a collaboration platform) with easy-to-use interface and tool-support. In this paper, we investigate the development of a *Collaborative Video Annotation System* which can facilitate collaborative learning of foreign language and

promotion of cultural exchange. We argue that with the increasing availability of Internet and bandwidth, digital divide is not always due to technological barrier but often due to our inability to integrate human sharing spirit and goodwill along with the technological advances.

2.0 Foreign Language Learning Scenario

Main-stream foreign language learning materials have a huge market and therefore, publishers worldwide have been actively publishing study materials for the target market. Main-stream foreign language such as English has plenty of materials available in many native languages. It is truly unfortunate and undesirable that a Vietnamese learner often needs to learn Thai using materials written in English, and vice versa. The market-driven phenomenon left some languages far behind from the others. Our language and culture may or may not have a feasible market but they are our unique assets, and we ought to find ways to promote them using enabling technologies. By successfully utilizing the potential of the Internet and human spirit of collaboration, it is possible to promote our language and culture in an online collaborative fashion. The development of our video annotation system is inspired by language education and cultural exchange by making use of the potentials of ubiquitous Web and online virtual collaboration.

Unlike technological disciplines, foreign language teaching and learning follows typical pattern and rather static. Off-the-shelf foreign language learning kits generally consist of lessons organized into different proficiency *levels.* For instance, the *Beginner's Level* consists of lessons with basic alphabet-set, phonetics and writing system along with greetings, basic vocabulary, simple grammars and conversations. The *Intermediate Level* materials go beyond the basics and include further lessons and practices with the help of situational dialogues and further syntactic, semantic and pragmatic notes. *Advanced Level* learning materials are less restrictive and could go beyond typical textbook lessons – it is not uncommon to use real newspaper articles, movies or TV news or documentaries to teach different linguistics and cultural facets to the advanced learners. In many cases, we noticed that a vast majority of these off-the-shelf software products are mere digitization of it predecessor, video-based learning kits. However, there are exceptions. For instance, ELLIS (Ellis, 2006) – an English learning software tries to integrate ESL pedagogical research with computer and networking technologies. Such materials are often expensive and beyond the purchasing power of individuals.

In terms of contents, we noticed many similarities across foreign language learning materials. Unless commercially motivated, contents such as alphabet chart, pronunciation audio files, grammar flashcards, situational dialogue videos could have been reused effectively for teaching that foreign language to any native language community through customization and annotation.

Screenplay (Screenplay, 2006), a Japanese publisher uses foreign films to help Japanese learners improve their foreign language ability. Using their Web site and multimedia products, a user can also conduct associative search to see a word in real context. This is not only effective as evidenced by scholarly research (Cardilo, 1997), but also a fun activity which promotes the understanding of culture, etiquette and practices beyond mere language learning through grammar rules, substitution exercises and the like.

3.0 Cultural Exchange Scenario

Governments and agencies across the world do invest plenty of money and efforts in producing documentaries and other multimedia materials for the promotion of culture and tourism. However, their focus remains mainly to be the targeted audiences (and, unfortunately, limited to a few main-stream languages). For instance, most documentaries and videos developed in Japan are often available in Japanese, English, French, and Chinese, etc. Unless properly annotated with metadata, transcripts, subtitle or soundtrack, it is unlikely to find those materials useful for an audience who do not understand any of these languages.

With our decades of experience with other online collaborative projects, such as Project Gutenberg (Gutenberg, 2006) and Aozora Bunko (Aozora, 2006) - where copyright-free books are digitized and even translated by volunteers worldwide, we can expect that such promotional video materials can equally be annotated collaboratively and free-of-charge as long as an easy-to-use collaborative system is available and online volunteers are supported with easy-to-use interface and tools.

4.0 Copyright and Other Issues

It is a fact that commercially developed materials may not be available free, but it is inevitable that some promotional materials from *not-for-profit* agencies and governments may gradually become available publicly. We also rely on amateur personal videos taken by digital and video cameras as a good starting point. The Open Video Project (Open-Video, 2006) is a general digital library of publicly available videos from unrestricted domains - ranging from classroom lectures to public service or documentary videos). We focused on gathering and annotating video and multimedia materials in a *restricted domain* through small-community collaboration. Our target is to promote cultural exchange and facilitate foreign language education while making full use of the ubiquitous connectivity of the Internet and the growing human spirit of collaboration on the WWW.

In doing so, we focused on developing an easy-to-use collaborative video annotation system which we have explained in this paper in detail. We also refer to our other publication on creation of sharable and reusable E-Learning materials using a component-oriented approach (Hasan et al., 2006). We argue that using such multimedia contents with proper metadata, and relevant structural, semantic and pedagogical information, it is possible to generate customized and personalized course materials for foreign language learning community. The same approach is also applicable to other E-Learning scenarios.

We are aware that for any volunteer-based collaboration, motivation plays a crucial role. The success of our initiative therefore, remains on the enthusiasm of volunteers who take pride and care about their own language and culture. The evidence of the steady growth of collaborative projects on today's Internet and WWW (e.g., the Open Video Project) is encouraging. Moreover, a collaborative project such as this one requires a critical-mass of volunteers and contents to become successful. It is obvious that unless we have a handful number of volunteers or a sizeable amount of contents, our initiative may not receive sufficient attention quickly. With an initial support from Asia Pacific Telecommunity (APT), we have initiated this project and successfully developed a fully working prototype currently available at http://apt.shinawatra.ac.th/video/. Some of our international collaborators are consistently supporting us by developing or accumulating useful contents for promotion of foreign language and culture on the cyberspace. We envisage that we will continuously receive supports from governments and agencies to develop this project further.

In the rest of the paper, we will explain the implementation and feature of this video annotation system.

5.0 Outline and Features of the Video Annotation System

The video annotation system is fully Web-based. Users interact with the system using their Java-enabled Web-browsers without having to resort to any video plug-in. We choose *MediaFrame* (MediaFrame, 2006) – an open-source software for streaming video manipulation and playback. We use MediaFrame's Javascript API, PHP scripts and MySQL database to develop the Web-based video annotation system. The technical complexities of the system remain transparent to the users.

5.1 MediaFrame Streaming Media Player

MediaFrame is an open-source streaming media platform in Java™ which provides a fast, easy to implement, and extremely small applet that enables web users to view audio/video contents

without having to rely on external player or plug-ins. MediaFrame does not require special servers, software or programming knowledge. Integration of MediaFrame with other applications becomes easy due to its well-defined JavaScript API and its supports to Mpeg-1 and Mpeg-4 video formats.

MediaFrame can deliver audio and video content over the Internet in either streaming or progressive download mode. It is also capable of stretching media by up to 60% of its original size without a significant loss of quality. This enables us to simulate higher bit rates without the associated costs and to operate effectively in both broadband and narrowband environments. We want our system to be reachable by people in the remote area where the infrastructure is not as good or may have to resort to satellite links for a connection. MediaFrame is capable of detecting user's connection speed in real-time and delivers media accordingly.

A full-featured Digital Rights Management (DRM) system is also integrated with the core of MediaFrame allowing media files to be locked to a specific domain and expired over time. Moreover, MediaFrame is fully JavaScript controllable; making it simple to build our own control set and to integrate it with our video annotation system seamlessly. Using the JavaScript API, we could also conveniently add transcripts and subtitles to video materials and the multiple-playback-states allow us to tightly integrate MediaFrame for effective video annotation and playback.

5.2 Descriptive Metadata Annotation

For non-textual materials, such as videos, descriptive metadata plays an important role in exploratory search and browsing (Marchionini, 2006). MPEG-7, LOM and SCROM (Godwin-Jones, 2004) are sophisticated metadata for multimedia materials. However, we adopt a simpler subset of descriptive metadata for the system in order to avoid fears in the ordinary users. At the moment, the system allows users to annotate a video-clip with Title, Genre, Original-Language, Keywords, Descriptions and other metadata. Whenever applicable, structural metadata such as Video Format, Duration, Color, etc. are extracted automatically from the video file.

5.3 Supplementary Annotation

Our system allows a user to step through a video-clip and to add transcripts along with its original soundtracks and transcripts. User can also select transcript in an available language and translate it into a target language. During playback, users can choose available transcripts to be displayed as subtitle with the video.

It should be noted that for effective learning of foreign language using video, transcripts need to be further annotated with extra information such as, grammatical annotation or relevant cultural notes. Such features are currently being implemented.

5.4 Some Representative Video-clips and Annotations

In this section, we explain a series of video-clips from our system which explains the potentials of our system for language learning and cultural exchange.

Figure 1 shows a video clip from ELLIS played with English (a) and Thai subtitle (b). Figure 2 shows amateur video-clips of (a) Kashima Jingu Shrine in Japan and (b) about Japanese Paper Art (Origami) – although the original soundtrack is in Japanese, English subtitles are made available through online collaboration, and displayed accordingly. Figure 3 shows a video tutorial from JEITA about Natural Language Processing technology (originally with Japanese soundtrack) but is being played with English subtitles since such information has already become available.

It should be noted that for all these video-clips, descriptive metadata is also available in the system – annotated collaboratively and indexed automatically to facilitate exploratory search and retrieval. Users can use keyword search (analytical strategy) or browse (partition-strategy) the video collection effectively with the help of relevant metadata.

Figure 1a. Situational dialogue in a Bank: with English subtitle

Figure 1b. Situational dialogue in a Bank: with Thai subtitle

Figure 2a. Amateur Video Clips: Kashima Jingu Shrine in Japan

Figure 2b. Amateur Video Clips: Japanese Origami or Paper Art

Figure 3. A video tutorial developed by JEITA on *Natural Language Processing* – originally in Japanese, but we added English subtitles and descriptive metadata online.

Finally, Figure 4 shows an example annotation interface. Video is displayed with playback control, time-stamp and other information. In the process of filling in the *Time* and *Subtitle* boxes below, annotating user can choose reference subtitle from available source languages, and control the video playback with control buttons as necessary. Other annotation interfaces are similar browser-based easy-to use ones.

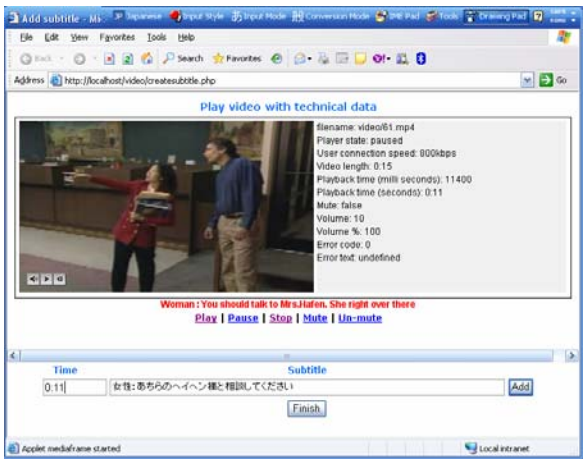

Figure 4. Subtitle annotation interface: Time-stamp and Playback Control are displayed as video-clip is being played.

6.0 Conclusion and Future Work

In this paper, we outlined an operational prototype of a collaborative video annotation system developed for foreign language education and cultural exchange in mind. We often ignore the fact that majority of Internet users are not proficient in foreign language and therefore, materials available on the WWW in a foreign language has little or no use to those people. Multimedia materials, such as videos are relatively easy to understand or appreciate without *complete* translation since contextual, visual and other cues are present there. However, it is extremely difficult to locate non-textual materials on the Web using today's *hit-or-miss* search engines which heavily rely on textual indexing. Metadata annotation is therefore essential and can be done effectively in small-group collaboration. We do admit that we are yet to conduct any experimental evaluation of the proposed system to justify the effectiveness of this approach. However, our future work will eventually attempt such evaluations and vigorous live trials.

We also plan to organize our video collection using Semantic Web technologies (such as with the help of an ontology) so that the video collection can be further manipulated efficiently and inferences can be made. For instance, a foreign language learner or a teacher accessing the video collection could be able to locate and organize relevant video materials on-the fly. Organization and distribution of the annotated multimedia materials using simple digital library interface are also planned.

We envisage that the annotated video library with its multimedia contents will serve as (1) an E-Learning Content Server - storing both contents and structures (contents with pedagogical attributes and ontology) to facilitate personalized E-learning; and, (2) a Video Library for Cultural Exchange – storing amateur videos and multimedia (contents with cultural attributes and ontology) to facilitate digital library services. Some of our collaborators are also keen to develop Mobile APIs for such E-Learning and Digital Library services to facilitate access to these multimedia contents using satellite links and handheld devices.

Acknowledgment

We like to thank **Asia Pacific Telecommunity** (APT) for its generous support through the ICT-HRD program. We also thank **Thai Research Fund** (TRF) for its TRF grant (No. MRG4880112). Mr. Nophadol Jekjantuk - a Shinawatara University student helped us developing and testing the video annotation system. We also acknowledge valuable comments from Dr. Suravuth Pratishthananda at Shinawatra University; Mr. Teerayuth Boonchote and his colleagues at Shin Sattelite Plc; Mr. Shin-ichi Yamamoto and Dr. Yoshiyuki Fujino at NICT-Japan towards adapting the system on satellite links and portable devices.

References

Aozora (2006). Aozora Bunko Web site; http://www.aozora.gr.jp/; In Japanese; Last accessed May 2006.

Cordilo, D.S. (1997). Using a Foreign Film to Improve Second Language Proficiency: Video vs. Interactive Multimedia, Journal of Educational Technology Systems, Vol.25 no.2 pp.169-77. http://www.albany.edu/ltl/martin/first.html

Ellis (2006) ELLIS Inc. Web Site, http://www.ellis.com/; Last accessed May 2006.

Godwin-Jones, Robert (2004). EMERGING TECHNOLOGIES - Learning Objects: Scorn or SCORM? Language Learning & Technology, Vol.8, No.2, pp.7-12

Gutenberg (2006). Project Gutenberg Official Site http://promo.net/pg/; Last accessed May 2006.

Hasan, M.M., S. Yamamoto, Y. Fujino and W. Chujo (2006). Towards Sharable and Reusable Contents using Semantic Web Technologies for Personalized Foreign Language Learning. IEICE General Conference - 2006, Tokyo, Japan; D8: Artificial Intelligence and Knowledge Processing Track - Paper No.: D-8-1; pp.84.

Johnson, C.H. (2001). A survey of current research on online communities of practice. Internet and Higher Education Vol.4, pp.45-60.

Marchionini, Gary (2006). Exploratory Search: From Finding to Understanding, Communications of ACM, Vol.49, No.4, pp.41-46.

MediaFrame (2006). MediaFrame Web site; http://Mediaframe.org/; Last accessed May 2006.

Open-Video (2006). Open Video Project Web site; http://www.open-video.org/; Last accessed May 2006.

PERSONALIZED DIGITAL LIBRARY USING DYNAMIC USER PROFILE AND CYC ONTOLOGY

Khulthida Chumsaktrakul and Md Maruf Hasan

School of ICT, Shinawatra University, Thailand
47312003-8@padawan.shinawatra.ac.th, maruf@shinawatra.ac.th

ABSTRACT

Digital libraries are new tools for making information ubiquitous, processable and dynamic. As digital libraries become commonplace on today's Internet, people expect adaptive and personalized services from them. Digital libraries must change from being passive – from little adaptation to their user to being more proactive in tailoring and offering information to users by using personalization and recommendation techniques. Maintaining a dynamic user profile through usage pattern analysis is also possible by keeping track of user activities. Moreover, external knowledge-bases such as Cyc or other domain ontologies may be used to make intelligent inferences about digital library users and contents.

This paper proposes a personalized digital library system using open source digital library, *Greenstone* and open source ontology, *OpenCyc*. The major contributions of this research are (1) integration of digital library with a knowledge-base, and (2) development of relevant software modules for this purpose. We developed software modules to capture user profiles and analyze usage patterns, and integrated these modules with Greenstone and OpenCyc.

Keywords: digital library, ontology, personalization and recommendation

1 INTRODUCTION

A *digital library* is a collection of digital objects including text, video, and audio, together with methods for access and retrieval, selection, organization and maintenance of the collection [1, 2]. Many easy-to-use tools such as *DL-In-A-Box*, *DSpace* and *Greenstone* help creation of digital libraries with search and browsing interfaces and metadata efficiently. Therefore, many digital collections are built and most of them are made available on the World Wide Web providing library users with knowledge where and when it is needed. These digital library collections do not intend to supersede existing bricks-and-mortar libraries; on the other hand, digital libraries are new ways of dealing with knowledge. We argue that a true digital library should offer more than search and browsing facilities on top of a collection of digital objects.

As users are being overwhelmed both by the quantity of information, and by its variety of form and quality, digital libraries must change to be proactive in tailoring and offering information for their users who expect more personalized services. There are existing solutions to this problem. The widely used *personalization and recommendation* techniques in E-Commerce [3, 4] are readily usable in the context of digital library. Personalization and recommendation attempt to reduce information overload and retain users by learning about a user's needs proactively, and by presenting the user most relevant items based on users' preference and profile. Unlike E-Commerce, several factors may contribute positively to the development of personalized services in digital libraries. For examples, library users are willing to tell the system about themselves more without worrying about privacy. A digital library should unobtrusively keep track of usage patterns of its users and then recommend useful materials to its users based on their interests and profiles. In the digital library, we can make use of both *content-based filtering* (content and metadata analysis) and *collaborative filtering* (usage pattern analysis among users) to offer personalization and recommendation services [4, 5].

Information needs of a user are not limited to the original profile they create at sign-up, but often changes over time. Accordingly, a dynamic profile should be maintained in a digital library for each DL user. We consider the user-defined initial profile as "seed" and attempt to use external knowledge bases (such as *Cyc* ontology) to make intelligent inferences about DL users and contents to maintain *dynamic user profile* in the system.

Cyc ontology is the largest and most complete general knowledge-base and commonsense reasoning engine in the world [6, 7]. Cyc has a rich knowledge-base of concepts and individual instances, and specifically-defined relationships that describe how those concepts are related. Cyc also offers well-defined Java APIs to access it [7]. Therefore, user profile and usage pattern can be intelligently analyzed with the help of Cyc-ontology and inference engine.

Cyc uses its own language, *CycL*, to represent knowledge. *CycLConstants* denote specific individuals (objects) or collections (classes) by using a character string prefixed by #$ *like* #$Dog. *CycLFormula* is a relation applied to some arguments, enclosed in parentheses *like* (#$isa #$GeorgeWBush #$Person). CycL also has logical connectives such as #$and, #$implies, and quantifiers such as #$forAll,

#$thereExists, and uses ?X to denote any variable [7]. Therefore, Cyc can imply rule as follows:

```
(#$implies
      (#$isa ?Person1 #$Person)
      (#$thereExists ?Person2
       (#$and
         (#$isa ?Person2 #$Person)
         (#$loves ?Person2 ?Person1))))
```

which means that *everyone is loved by someone*.

In our study of personalized services for digital library, we use open source digital library system, *Greenstone,* and capture and analyze user's profiles and interests dynamically with the help of *OpenCyc* (open source ontology and inference tool).

2 SYSTEM OVERVIEW

The Personalized Digital Library System explained in this paper consists of *five* main components: web interfaces, user profile, bookshelf database, *Greenstone* digital library and *OpenCyc* knowledge-base. We also used PHP scripts and Java to integrate these five components to work together in maintaining dynamic profile of DL users. Similar to Greenstone, our users are provided with a Web-based interface to interact with the system. In addition, Greenstone's Library Interface, *GLI* is used to develop a digital library collection with appropriate metadata. The dynamic profile management and personalization is done in the backend using OpenCyc.

2.1 Preparation

OpenCyc only includes a small number of assertions and relations. Therefore, we added new assertions and relations and modified existing ones in the context of our digital library collection.

For example, we consider #$Software has the same meaning as #$Program and #$ComputerProgram-CW is a sub-collection of #$Program, thus #$ComputerProgram-CW is also a sub-collection of #$Software. We created this new #$Software collection and then asserted those relations by these formulas:

```
(#$prettyString #$Software #$Program)
(#$genls #$ComputerProgram-CW #$Software)
```

New assertions and relations were stored in OpenCyc using OpenCyc API.

To avoid *cold-start problem*[1], the initial user profile and digital materials are incorporated in our system. We maintain two tables in MySQL database: the *User Table* provides information about user's preferences and the *Bookshelf Table* stores information about user's selected items (along with the metadata).

[1] Recommendations are required for new items or users for whom little or no information has yet been acquired. Poor performance resulting from a cold-start can deter user uptake of a recommender system [5].

We use GLI interface to add metadata and create searching and browsing indexes. Other than full-text search, Greenstone facilitates field-specific search on Title, Author, Publisher and other fields. Greenstone also facilitates hierarchical browsing based on classifiers such as an A-Z List of Titles or a Hierarchy of the File Formats.

2.2 How the Personalized DL System works

The detail operations of our Personalized DL System are shown in figure 1 below. A user must subscribe to use the services with a Web interface and inform the system about his or her interests and preferences which are kept in a MySQL database for further manipulation. (1) Users are required to log into the system; the user ID and password are verified. (2) For a valid user, his or her profile is matched against OpenCyc's assertions and relations to make intelligent inferences; and subsequently, (3) the user profile is also updated. (4) The system then takes this inference result as a new keyword to search for similar items in the digital library and (5) displays a list of recommended items based on user's interest.

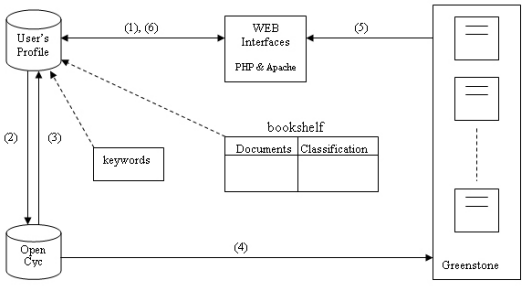

Figure 1: *Architecture of the Personalized Digital Library System*

Concurrently, (3) the new keyword is added or refined into the interested areas in the user profile which increases more specific area of interest. In addition, while the user is searching and opening documents, (6) the system keeps track of what items each user is reading; and optionally putting into their bookshelf.

OpenCyc's inference engine provides one or more rules for each concept; thus, inferring new keyword depends on the concept areas triggered each time (based on the user's current context). Observing user behavior leads to more accurate understanding of user's current information needs. With such a mechanism, a DL user's information space is no longer dependent on the original (static) profile and old history; but it reflects his or her present context.

The entire process of inferring keywords from the Cyc ontology and updating user profile in MySQL database in our system are illustrated in figure 2 using a sequence diagram.

2.3 Integration of Greenstone with OpenCyc

HTML, PHP script and Java API are used to combine all components of the system together. First, integration of

digital library and web interfaces is implemented using PHP and HTML. Second, connection between web interfaces and user profile is done by PHP and MySQL. Finally, the integration of MySQL database and OpenCyc is done by using Java.

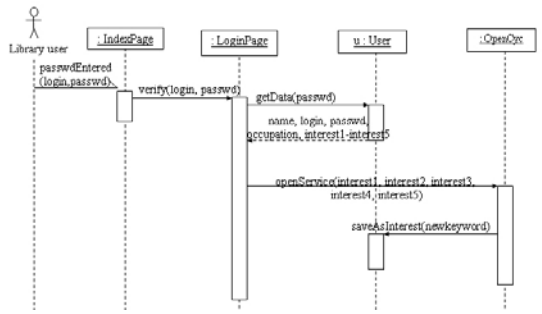

Figure 2: *Sequence diagram of inferring keywords and updating profile in the Personalized Digital Library*

2.4 Specific Features of the Personalized Digital Library System

Following are some specific features of the Personalized DL System we developed.

2.4.1 Inferring new keyword: User profile in the User Database is utilized to make inference with the help of OpenCyc by using Java API. Please note that the rules for inference triggered each time may not be the same in OpenCyc. It depends on user's current context.

Here is an example of how Cyc infers new keyword from the original user profile which may include two interested areas such as *Business* and *Computer*. Before getting the final result, it might take several steps to intelligently make inferences. For instance, #$Business implies #$CommercialActivity; and #$Commercial-Activity is a super-class of #$Account; and therefore, #$Account is a kind of #$Business. In addition to this, #$Account has some relations with #$Computer; implies that, every business needs to do accounting and computers are related with accounting systems. The following rules are triggered in OpenCyc in inferring concepts, such as #$AccountSystem:

```
(#$and
      (#$genls ?X #$Computer)
      (#$equals #$Account ?X))
```

which means that *an account system is a kind of account and is related with computer* (Fig. 3).

2.4.2 Maintaining a dynamic user profile: We start with the *initial* user profile and continuously augment it with the help of Cyc knowledge-base to maintain a dynamic user profile. Meaningful and relevant keywords inferred by OpenCyc are captured and subsequently incorporated with the dynamic user profile. Every time an authenticated user utilizes the personalized service from this system, his or her profile keeps changing based on his or her key area of interest. This key area is assigned as a key to recommend relevant items from the digital library collection. Another alternative way to update dynamic profile is through direct user interaction – an authenticated user may directly update his or her personal profile through Web interface (Fig. 4).

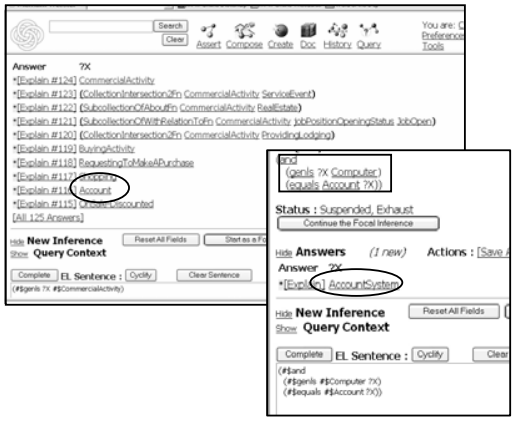

Figure 3: *Cyc infers AccountSystem, a new keyword from two original ones – Business and Computer*

Figure 4: *The user profile is manually updated with new keyword*

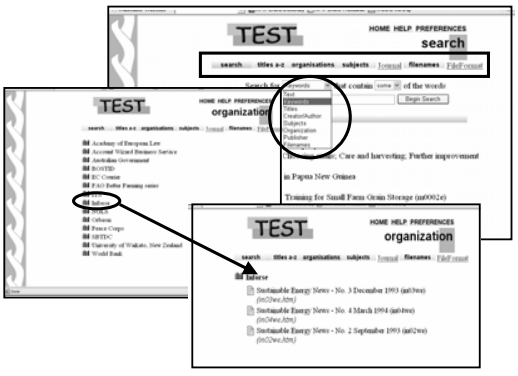

Figure 5: *A user can search by any keywords or browse by hierarchy to access DL contents*

2.4.3 Retrieving documents: The digital library system displays all items similar to user profile after making

inference and updates user profile accordingly. Therefore, users can directly access digital items via Web-browser spontaneously (without having to search). Nevertheless, an authorized user can also browse or search by keywords for relevant information if they prefer to do so (Fig. 5).

2.4.4 Adding document into bookshelf: The system keeps track of user behavior by showing a pop-up window, and by prompting user to add items into his or her bookshelf as he or she uses the service. The selected items are added into his or her bookshelf (Fig. 6). By carefully analyzing the bookshelf for a particular user, we can extract useful information about the user preferences.

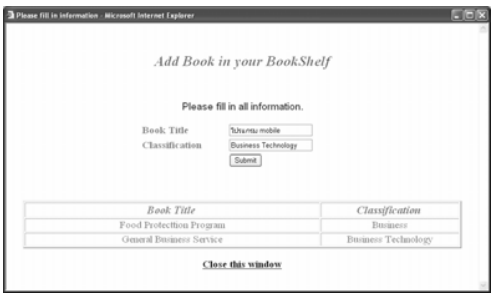

Figure 6: *Bookshelf management in the Personalized DL System*

2.4.5 Displaying and updating bookshelf entries: An authenticated user has full access to his or her selected items in the bookshelf. He or she can add new items, remove old items, and access the items quickly from the personal bookshelf (Fig. 7).

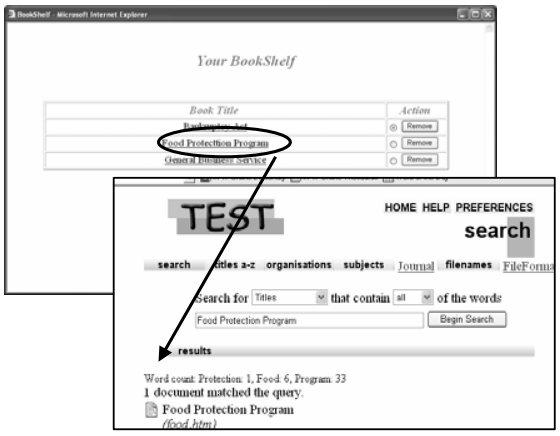

Figure 7: *An authenticated user can easily reach his or her digital items via the bookshelf.*

3 CONCLUSION AND FUTURE WORK

Unlike traditional libraries, digital libraries are unique in a sense that DL contents are mostly available in machine understandable form along with meta-information. Manipulation of such contents along with user's access patterns and profiles obviously enhance our experience in using a digital library. What makes digital libraries superior to its traditional counterpart is its ability to offer intelligent services to users. The Personalized Digital Library System explained in this paper demonstrates an innovative integration of a digital library with a knowledge-base. The system is also capable of maintaining dynamic profiles and making inferences; consequently, DL users are offered with the different digital items at different times based on their current context. We plan to incorporate sophisticated content-based and collaborative filtering algorithms with our system in the near future. Temporal changes in user's information needs are also worth investigating under the proposed personalization framework. We plan to enhance the system to make the user's involvement in profile management minimum (less obtrusive).

Finally, we would like to add that OpenCyc is a restricted version of Cyc knowledge-base. Obviously, ResearchCyc is a better alternative to develop a full-featured system; however, we failed to obtain a license for ResearchCyc due to license restrictions applied to Thailand [5, 8, 9]. We have plans to experiment with alternative knowledge bases.

4 ACKNOWLEDGMENT

This work is supported by *Thailand Research Fund* (TRF Grant: **MRG4880112**) awarded to Dr. Md Maruf Hasan. He wishes to thank his mentor, Dr. Ekawit Nantajeewarawat, and his colleague, Dr. Chutiporn Anutariya for their valuable comments.

REFERENCES

[1] Ian H. Witten and David Bainbridge (2003). How to Build Digital Library. The Morgan Kaufmann Series in Multimedia Information and Systems.

[2] David Bainbridge, Dana McKay, and Ian H. Witten (2004). Greenstone Digital Library Developer's Guide. New Zealand.

[3] Saverio Perugini and Marcos Andre Gonçalves, and Edward A. Fox (2002). Recommendation and Personalization : A Survey. arXiv:cs.IR/0205059 Vol. 1.

[4] Saverio Perugini and Marcos Andre Gonçalves (2004). Recommender Systems Research: A Connection-Centric Survey. Journal of Intelligent Information System Vol. 23 (2): 107-143.

[5] Stuart E Middleton, Nigel R. Shadbolt, and David C. De Roure (2004). Ontological User Profiling in Recommender Systems. ACM Transactions on Information Systems (TOIS) Vol. 22 (1): 54-88.

[6] Nick Siegel, Keith Goolsbey, Robert Kahlert, and Gvin Matthews (2004). The Cyc System. Notes on Architecture. Texas.

[7] Cycorp, Inc (2001). Cyc Tutorial. Cycorp, Inc. http://www.opencyc.org/releases/doc/tut/?expand_all=1 (accessed July 21, 2005).

[8] Douglas B. Lenat (1995). CYC: A Large-Scale Investment in Knowledge Infrastructure. Communications of the ACM Vol. 38 (11): 33-38.

[9] Douglas B. Lenat, George Miller, and Toshio Yokoi (1995). CYC, WordNet, and EDR: Critiques and Responses. Communications of the ACM Vol. 38 (11): 45-48.