Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia Computer Science 10 (2012) 593 – 600
Procedia Computer Science

The 9th International Conference on Mobile Web Information Systems (MobiWIS)

A geo-location context-aware mobile learning system with adaptive correlation computing methods

Nagato Kasaki[a*], Shuichi Kurabayashi[b], Yasushi Kiyoki[b]

[a]*Graduate School of Media and Governance, Keio University 5322, Endo, Fujisawa, Kanagawa, 252-0882 Japan*
[b]*Faculty. of Environment and Information Studies, Keio University 5322, Endo, Fujisawa, Kanagawa, 252-0882 Japan*

Abstract

This paper proposes a context-aware mobile learning system with adaptive correlation computing methods. This system enables users to enhance their knowledge by correlating it with daily experiences. The proposed system contains a hybrid metric vector space to define the correlation between heterogeneous metadata vectors of the user context and learning material. The system integrates heterogeneous metric vector spaces with definitions of the semantic relations between the vector spaces. The significant feature of this system is a hybrid adaptation mechanism for the calculation of correlation. The adaptation mechanism has multidirectional adaptation functions for various learning materials, situations, and learners. We propose a revise-localize-personalize (RLP) adaptation model. In the adaptation mechanism, users only have to improve the metadata or the relations just in their relevant field. The advantage of the system is that the system reduces the time-intensive efforts required for describing direct relations between user contexts and learning materials. This paper presents the feasibility of the context-aware heterogeneous information provision with the hybrid metric vector space, by implementing an actual mobile application system and examining real-world experiments on data provision.

Keywords: mobile learning; context-aware; correlation computing; adaptive computing.

1. Introduction

The development of mobile computing technologies has revolutionized the learning processes of humans in daily life. New methods for acquiring required learning materials at any location and at any

* Corresponding author. Tel.: +81-80-3389-6831.
E-mail address: knagato@sfc.keio.ac.jp.

doi:10.1016/j.procs.2012.06.076

instant of time are evolving [1]. People can access various types and a large number of multimedia learning materials using their mobile devices. The self-paced e-learning product and service market will grow by 9.2% compounded annually from 2010 to 2015 and reach $49.9 billion by 2015 [2]. According to trial research in the m-learning project, people will be attracted by mobile learning materials and will remain motivated for continued learning [3].

However, people encounter difficulties in their search for learning materials corresponding to the requirements in their daily lives. This is a result of the wide semantic gap between the context of the learning materials and the context in the real world. The learning materials are based on certain assumptions about the simple and explicit context where people use the knowledge of the materials, such as "for self-introduction, " "at a station, " and "during shopping." However, the contexts in which the learning materials are actually used are more varied, complicated, and implicit [4]. Owing to this semantic gap, most novices can barely associate appropriate learning materials with their actual contexts. A new context-aware information provision mechanism to bridge the semantic gap is in great demand.

Thus, this paper proposes a context-aware mobile learning system with a hybrid metric vector space to provide learning materials appropriate for daily experiences of users. Fig. 1 shows an overview of the context-aware provision system architecture and user interactions. This system facilitates "experience-connected" learning, which means that people enhance their knowledge by correlating their knowledge with their daily experiences [5]. This system automatically recommends learning materials that fit the daily contexts of learners by dynamically calculating a correlation score between the daily experience and learning materials. This system provides the hybrid metric vector space to define the semantic distance between this heterogeneous information. This vector space differs from the existing vector space models for e-learning [6][7] in that it integrates heterogeneous information of learner experiences and learning materials. It contains metric vector spaces of the heterogeneous information and the correlation matrix of semantic relation values between the dimensions of the metric vector spaces for dynamic association.

The significant feature of this system is a hybrid adaptation mechanism for correlation calculation in the hybrid metric vector space. This adaptation mechanism has multidirectional adaptation functions for the daily experience, the learning material, and their relationship. The adaptation process is executed in an isolated part of the hybrid metric vector space. Hence, in this adaptation mechanism, the actors of the system improve the metadata or the relations only in their relevant field. We propose a revise-localize-personalize (RLP) adaptation model that facilitates a divide-and-conquer approach to improve the dynamic correlation computing process, which is the heterogeneous and complicated problem.

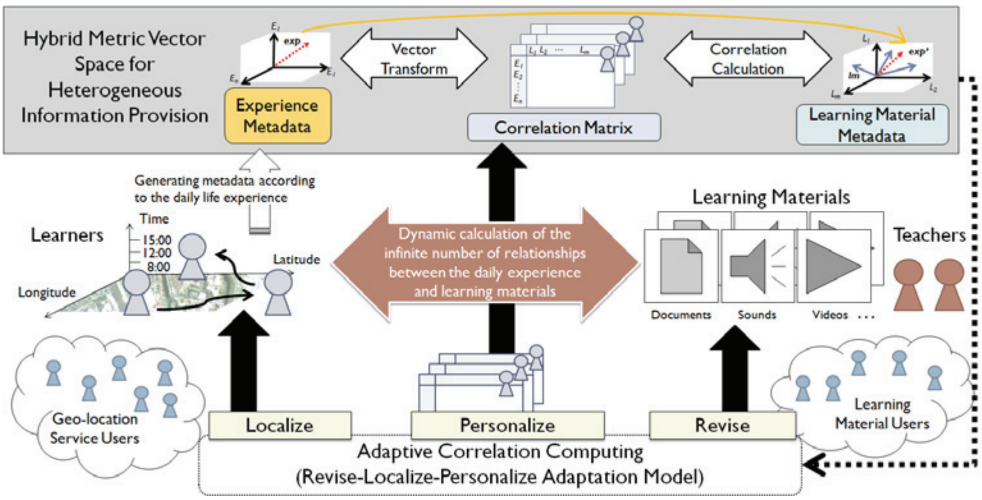

Fig. 1. The model of context-aware mobile learning system with adaptive correlation computing methods

The most notable advantage of this system is that the system reduces the time-intensive efforts required for describing direct relations between daily experiences and learning materials. The vast number of direct relations is difficult to manage. These time-consuming efforts have restricted creation of context-aware learning materials by developers. The proposed system can easily utilize a lot of legacy learning materials for experience-connected learning.

2. Context-Aware Mobile Learning System Architecture

This section presents the basic data models and functions of the context-aware mobile learning system with the hybrid metric vector space.

2.1. Semantic data models for contexts and learning materials

The context-aware mobile learning system describes the semantics of daily experiences and learning materials using metadata vectors. The metadata vector is a feature vector in a metric vector space that is associated with feature terms. The system formulates the feature terms to describe daily experience and learning material features. An experience metadata vector $\boldsymbol{exp}$ and a learning material metadata vector $\boldsymbol{lm}$ are defined as

$$\boldsymbol{exp} \coloneqq (e_1, e_2, \cdots, e_n), \qquad 0 \le e \le 1$$

$$\boldsymbol{lm} \coloneqq (l_1, l_2, \cdots, l_m), \qquad 0 \le l \le 1$$

where e and l are numeric values of relevance to the feature term associated with its dimension, n is the number of experience feature terms, and m is the number of learning material feature terms. The value of e or l is 0 when its object has no relevance to the feature term, and it is 1 when its object has the most strong relevance to the feature term. Vectors $\boldsymbol{exp}$ and $\boldsymbol{lm}$ belong to different metric vector spaces, and hence, they are easy to adapt separately. Vector $\boldsymbol{lm}$ corresponds to the semantic metadata set of a certain learning material such as its category, situation, intended purpose and so forth. Vector $\boldsymbol{exp}$ corresponds to the mixed current experience of the learner in his/her daily life such as "shopping with some friends in the afternoon at an electronics store," *etc*.

The system calculates the correlation between $\boldsymbol{exp}$ and $\boldsymbol{lm}$ by using correlation matrices. The correlation matrices aim to transform the experience metadata vector $\boldsymbol{exp}$ into the dimension of the learning material metadata vector $\boldsymbol{lm}$. The matrices are generated according to not only the target material data-set but also the personal preferences of the learners. Using these matrices, the system can deal with various types of learning materials and personal characteristics. A correlation matrix $\boldsymbol{M}$ is defined as

$$\boldsymbol{M} \coloneqq \begin{bmatrix} r_{1,1} & \cdots & r_{1,n} \\ \vdots & \ddots & \vdots \\ r_{m,1} & \cdots & r_{m,n} \end{bmatrix}, \qquad 0 \le r \le 1$$

where r denotes the semantic relevance between the learning material feature term and the experience feature term. These relevance values describe the definition of the relationship between the experience and the learning material, instead of direct formulation of the individual relationships. The relationship definition enables automatic calculation of the heterogeneous relationship scores on a large scale. The relevance values vary according to the preferences of the learners for provision adaptation. Each value of r is initially defined by manual and will be improved through the adaptation process of the system according to the feedback of learners about the provided learning materials.

2.2. Context-aware learning material provision functions

The system analyzes current experiences of learners in the real world using raw data sensed by their mobile devices in the provision process. The system generates experience metadata $\boldsymbol{exp}$ for each learner experience. Let the raw data vector of mobile device sensors be $s = (s_1, s_2, \cdots, s_x)$. The experience metadata generation function $f_{generate}$ is defined as

$$f_{generate} : \mathrm{s} \to \boldsymbol{exp}$$

An example function of $f_{generate}$ is the following function $g_{location}$ that utilizes learner location and time information. $g_{location}$ initially formulates ideal $\boldsymbol{exp}$ describing a specific learner experience. The ideal $\boldsymbol{exp}$ is linked to the geographical metadata (lat, lng) and the time window metadata from t_{start} to t_{end}. The geographical metadata and time window metadata represent the location and time of the experience in their daily lives of learners. $g_{location}$ generates $\boldsymbol{exp}$ by synthesis and normalization of the k sets of values of ideal $\boldsymbol{exp}$ closely matching the learner contexts. First, the system selects the learning material set for provision by using the time window metadata. The system chooses the learning materials specific to the current time of day. Next, the system sorts the selected ideal metadata in ascending order of the distance between the geographical metadata and the current location of the learners. Finally, the system synthesizes the top k ideal $\boldsymbol{exp}$ in the ranking and applies 2-norm normalization.

The system transforms the generated $\boldsymbol{exp}$ into the dimension of the target $\boldsymbol{lm}$ by using personalized correlation matrix $\boldsymbol{M}$. The vector transform function $f_{transform}$ is defined as

$$f_{transform}(\boldsymbol{exp}, \boldsymbol{M}) := \left(\sum_{k=1}^{n} \boldsymbol{M}[1,k] \cdot \boldsymbol{exp}[k], \sum_{k=1}^{n} \boldsymbol{M}[2,k] \cdot \boldsymbol{exp}[k], \cdots, \sum_{k=1}^{n} \boldsymbol{M}[m,k] \cdot \boldsymbol{exp}[k]\right)$$

This transformation calculation denotes the multiplication of a matrix and a vector, i.e., $\boldsymbol{M} \cdot \boldsymbol{exp}$.

The system determines a learning material ranking ordered by the correlation score between the transformed experience metadata $\boldsymbol{exp}'$ and each learning material $\boldsymbol{lm}$. The correlation calculation function $f_{correlation}$ is defined as

$$f_{correlation}(\boldsymbol{exp}', \boldsymbol{lm}) := \sum_{k=1}^{m} \boldsymbol{exp}'[k] \cdot \boldsymbol{lm}[k]$$

This correlation calculation denotes the inner product of $\boldsymbol{exp}' \cdot \boldsymbol{lm}$. The inner product method is suitable for defining the correlation value because the value of the inner product is high only when both vectors have high values of the elements corresponding to the same dimensions. A high inner product value signifies high relevance of the two metadata vectors. The system finally recommends the appropriate learning materials according to the calculated correlation values of the materials.

2.3. Functions for RLP adaptation model

The system provides the multidirectional adaptation functions to improve the accuracy of correlation computing adapting to various learning materials, situations, and learners. The main adaptation functions are f_{revise} that revises the learning material metadata, $f_{localize}$ that improves the experience metadata generation using localization, and $f_{personalize}$ that personalizes material provision using matrix modification. We call the adaptation mechanism comprising these three functions a revise-localize-personalize (RLP) adaptation model. This model provides comprehensive coverage of the hybrid metric vector space data. These adaptation functions can be executed manually by learners, teachers, and so forth.

We define the automatic algorithm of adaptation functions to support manual adaptations. Let operation $[\![x]\!]$ be a relevance normalization operation; then, the adaptation functions are defined as follows:

$$[\![x]\!] = \begin{cases} 1 & (1 < x) \\ x & (0 \leq x \leq 1) \\ 0 & (x < 0) \end{cases}$$

$$f_{revise} := [\![\boldsymbol{lm}[x] + \mu_r(\boldsymbol{exp}'[x] - \boldsymbol{lm}[x])]\!]$$

$$f_{localize} := \left[\!\!\left[\boldsymbol{exp}[x] + \mu_l \sum_{k=1}^{m} \frac{\boldsymbol{M}[k,x]}{\sum_{l=1}^{n} \boldsymbol{M}[l,x]} (\boldsymbol{lm}[k] - \boldsymbol{exp}'[k])\right]\!\!\right]$$

$$f_{personalize} := [\![\boldsymbol{M}[x,y] + \mu_p(\boldsymbol{lm}[x] - \boldsymbol{exp}'[x]) \cdot \boldsymbol{exp}[y]]\!]$$

The output of each function is a new value of vector or matrix computed from the experience metadata $\boldsymbol{exp}$, correlation matrix $\boldsymbol{M}$, and learning material metadata $\boldsymbol{lm}$. Vector $\boldsymbol{exp}'$ is the current output generated by $f_{transform}$ calculated using $\boldsymbol{exp}$ and $\boldsymbol{M}$. Vector $\boldsymbol{lm}$ in $f_{personalize}$ and $f_{localize}$ denotes the training data, which is usually determined by user feedback or learning material providers. f_{revise} adapts the current $\boldsymbol{lm}$ to $\boldsymbol{exp}'$.

The adaptation functions improve the provision results by increasing the correlation of appropriate materials and by decreasing the correlation of inappropriate materials using parameter μ. Parameters μ_r, μ_l, μ_p are the adaptation rates; the domain of the values is $-1 \leq \mu \leq 1$. The value of μ is nearly 1 when the training data is appropriate for the current experience, and the value is nearly -1 when the training data is inappropriate.

We show an example of applying $f_{localize}$, which is the most unique function in the RLP adaptation model because it improves the heterogeneous metadata of experience from the learning material metadata. Let $\boldsymbol{lm_i}$ be the ideal learning material metadata vector. In the example considered, $\boldsymbol{lm_i}$ contained high values of l_1 among the learning material feature terms. The generated experience metadata vector $\boldsymbol{exp}$ contained high values of e_1. The correlation matrix $\boldsymbol{M}$ contained high relation values of $\mathrm{r}(l_1, e_2)$ and $\mathrm{r}(l_2, e_1)$. The transformed experience metadata vector $\boldsymbol{exp}'$ contained only the value l_2 and showed no correlation with $\boldsymbol{lm_i}$. The system applies $f_{localize}$ to $\boldsymbol{exp}$ in order to improve the correlation with $\boldsymbol{lm_i}$. From $\boldsymbol{lm_i}$ and $\boldsymbol{M}$,$f_{localize}$ assumed that $\boldsymbol{exp}$ should have the values e_1. After applying $f_{localize}$ to $\boldsymbol{exp}$, the new $\boldsymbol{exp}$ contained these feature values and its correlation with $\boldsymbol{lm_i}$ was improved.

3. System Implementation

For evaluation of the context-aware mobile learning system considering practical users in the real world, we implemented an actual working application of the system. The application focuses on language learning because linguistic knowledge is closely associated with real-world experiences. This application system is a web application for mobile web browsers using geo-location API to implement the experience metadata generation function.

3.1. Data structure implementation

The application system implements the experience feature terms using terms from typical situations in daily life that involve language usage. The number of experience feature terms may be extremely large if there are no restrictions. However, most terms that do not have a significant relationship with language use will not influence the correlation after vector transformation. Thus, the experience feature terms must only include the terms that have a close relationship with the situations in which users require the target learning materials. The application system particularly applies the 82 category terms that are adduced as "specific notions" in *threshold level 1990* as the experience feature terms [8]. These categories encompass the basic situations of language usage encountered when living in foreign countries. This is referred to as the "threshold level" or the "A2 level" in the Common European Framework of Reference for Languages (CEFR) [9]. These categories contain terms such as "at work," "schooling," "leisure," and "public transport."

The learning material feature terms should be defined by the authors of these materials or teachers of the curriculum that includes these learning materials, because their definition requires a comprehensive knowledge of the learning materials. The terms may consist of the legacy categories and tags initially found in the learning materials. In this case, the semantic relation value between the learning material and each feature term is easy to define. The implemented system utilized the 73 category terms initially found in the target linguistic learning materials, such as "transit," "restaurant," and "sport" for the learning material feature terms. If the learning material belongs to the corresponding category, the relation value is 0.5; else, its value is 0.

4. Experimentation for adaptive correlation computing methods

This section presents experimental evaluations of the quantitative precision of ranking results with actual experimental areas and multi-scale location evaluations of participants. We proposes the idea of quantitative experiment method with manually prepared relevance data and executed preliminary experiments. We visualized the precision of the results to the experimental area maps in order to acquire a comprehensive view. After calculating the precision values, we applied the adaptation methods and evaluated their effectiveness.

4.1. Evaluation metric

We examined extensive results of 300 locations for 2653 learning materials. We compared ideal ranked materials and actual ranked materials that were determined by the system. Then, we applied the normalized discounted cumulative gains (NDCG) to compare the two rankings. The NDCG is defined as

$$DCG_q = \sum_{i=1}^{q} \frac{rel_i}{log_2 i}, NDCG_q = \frac{DCG_q}{IDCG_q}$$

where rel_i is the relevance value of the i-th material of the target ranking. $IDCG_q$ is the DCG_q of an ideal ordered material ranking. The NDCG metric is suitable for the evaluation of the web search-oriented ranking because the NDCG is affected to a greater extent by relevant materials with higher ranking [10].

Initially, we specified scores for subjective relevance between the learning materials and each experimental area. First, we chose experimental areas and a time for user experiences. We selected the central place of Tokyo, Japan, at noon because this area consists of various buildings such as shops, restaurants, stations, and offices. We divided this area into smaller zones of various scales. Sub-squares of 3 sizes, i.e., having side lengths of 1000 m, 200 m, and 50 m are created, and 10×10 squares of each type of sub-square cover the experimental area. The reason of defining multi-scale squares is that the context of learning material use depends on not only the point location but also the breadth of vision. For example, a certain same location point means "city area," "university," or "computer room" according to the viewpoint of the learners. Finally, we generated the relevance scores between for each learning material and the experience in each square. We computed each relevance score by using the evaluations of 3 persons in order to ensure unbiased scores. In the experiments, 7 volunteers participated in the voting process to indicate whether the learning material in an experimental square was appropriate. We defined the relevance score as the number of participants who voted the material as appropriate. Each square obtained at most 3 votes indicating that the learning material was appropriate. Thus, the relevance score is a natural number from 0 to 3.

4.2. Experimental results

We created visualization maps of the ranking result precision for each experimental square. These visualization maps offer an overview of the geographical features in the learning material provision. We generated rankings for the material by using the application system for each experimental square, and we calculated the $NDCG_{10}$ of the top 10 materials. The 10th NDCG is used because the users of this system are more likely to view a maximum of the top 10 materials on their mobile devices in the daily life. Fig. 2 shows the precision visualization maps using Google Maps overlaid with a translucent color according to the $NDCG_{10}$ value of each square.. The average $NDCG_{10}$ values for each type of experimental area are 0.79, 0.54, 0.66, and 0.67 for 1000 m, 200 m, and 50 m long squares, and the entire area, respectively. Approximately 21 % of squares show an $NDCG_{10}$ value of less than 0.5, which squares we call the problem squares.

In the second experiment, we evaluated the effectiveness of the adaptation functions of the system. We applied the functions to the problem squares and compared the adapted precisions with the original precisions. We determined the function to be applied by detecting the problematic metadata in the material provision process. If the learning material metadata has undesired values, i.e., when the learning material is not provided in most areas, f_{revise} should be applied to the learning material metadata. If the correlation matrix has undesired values, i.e., when there is high variation between the precisions calculated using relevance scores of each user, $f_{personalize}$ should be applied to the correlation matrix. If the experience metadata has undesired values, i.e., when the square area always has a low precision, $f_{localize}$ should be applied to the experience metadata. Fig. 3 shows the improvement in the average precisions after applying the adaptation functions to each experimental area. The average values of $NDCG_{10}$ after the adaptation experiment are 0.86, 0.64, 0.72, and 0.74 for 1000 m, 200 m, and 50 m long squares, and the entire area, respectively.

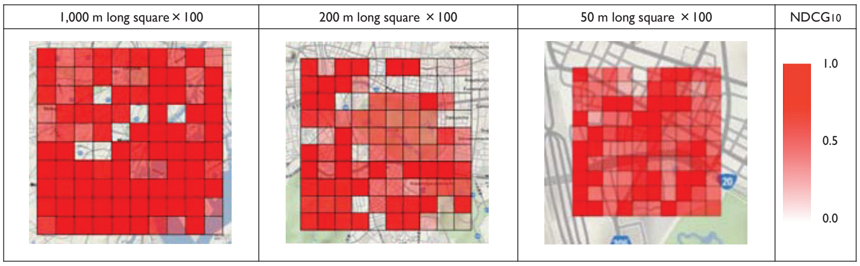

Fig. 2. The precision visualization maps overlaid with translucent color according to the $NDCG_{10}$ value of each square

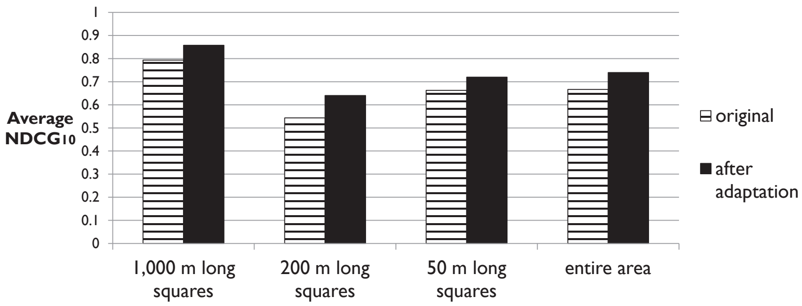

Fig. 3. The improvement in the average precisions after applying the adaptation functions for each experimental area

4.3. Analysis

The experimental results show that the system with the hybrid vector metric space realizes the context-aware heterogeneous information provision in the linguistic mobile learning field, and the RLP adaptation model is effective in improving the precision of the provision results. The provision results show the average value of $NDCG_{10}$ to be 0.72 after the adaptation process. This precision score should increase in order to realize an actual effective provision system. The current adaptation functions are certainly effective but not sufficient owing to the existence of squares whose precision was not significantly improved. The current adaptation functions mainly aim to enhance the potential provision of the ideal materials defined by users. The adaptation functions may need to create an elimination mechanism for inappropriate materials. Further, the novelty of the provided information is another important metric in addition to the precision. The creation of improved adaptation functions considering the novelty metric could be an area of future work.

5. Conclusion

This paper proposes a context-aware mobile learning system with adaptive correlation computing methods. In this system, a hybrid metric vector space defines the relevance between the heterogeneous information of daily experiences and learning materials. The significant feature of the hybrid metric vector space is its precision adaptation mechanism with the RLP adaptation model. This paper demonstrates the feasibility of the heterogeneous information provision with the adaptive correlation computing methods by analyzing the experiments with the actual application system. The system will enable users to enhance their knowledge by correlating their daily experiences and transforming their daily lives into a new learning environment.

References

[1] Keegan, D. (2002). *The future of Learning: From eLearning to mLearning*. Hagen: FernUniversität.

[2] Adkins, S. S. (2011). *The Worldwide Market for Self-paced eLearning Products and Services : 2010-2015 Forecast and Analysis*. Middle East: Ambient Insight Research.

[3] Attewell, J. (2005). *Mobile technologies and learning*. London: Learning and Skills Development Agency.

[4] Dey, Anind K. & Gregory D. Abowd. (2000). Towards a better understanding of context and contextawareness. In the *Workshop on the What, Who, Where, When and How of Context-Awareness, affiliated with the 2000 ACM Conference on Human Factors in Computer Systems (CHI 2000)*, The Hague, Netherlands. April 1-6.

[5] Nagato, K., Shuichi, K., & Yasushi, K. (2011). An Experience-Connected e-Learning System with a Personalization Mechanism for Learners' Situations and Preferences. *IES 2011-Emerging Technology for Better Human Life*, 118-123.

[6] Mangina, E., & Kilbride, J. (2008). Utilizing vector space models for user modeling within e-learning environments. *Computers & Education*, 51(2), 493-505.

[7] Ghauth, K. I. B., & Abdullah, N. A. (2009). Building an E-learning Recommender System Using Vector Space Model and Good Learners Average Rating. *2009 Ninth IEEE International Conference on Advanced Learning Technologies*, 194-196.

[8] Van E., J. A. & Trim, J.L.M. (1990). *Threshold level 1990*. Council of Europe Press.

[9] Council of Europe. Council for Cultural Co-operation. Education Committee. Modern Languages Division (Strasbourg). (2001). *Common European Framework of Reference for Languages: learning, teaching, assessment*. Press Syndicate of the University of Cambridge.

[10] Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern information retrieval (2nd ed.)*. Addison-Wesley Professional.