Unlocking the Potentials of Mobile Learning Object Compilation by Using Random Forest and Semantic Web Based as Tool for Lecturers

Mu'azu Abubakar Ahmed*[a,b], Muhammad Rafie Mohd Arshad[a], Mohd Heikal Husin[a],Hasimah Mohamed[a],Yussuf Yakubu[a,c], Rogers Phillip Bhalaluses[d],

a: School of Computer Sciences, Universiti Sains Malaysia, Pulau Penang, Malaysia.
b: Department of Mathematics, Kaduna State University, Kaduna, Nigeria.
c: Department of Computer Sciences, FCE Zaria, Kaduna State, Nigeria.
d: Open University of Tanzania, Dar es Salaam, Tanzania.
**Corresponding Author Email:abubakarmuaz11@gmail.com*

Abstract

Mobile learning, is an area which make use of moveable devices to get access to learning and its activities. Regrettably, most universities and establishments in developing countries and beyond do not have learning contents that are compatible with the mobile devices. This makes the development of mobile learning contents difficult and therefore the practice of reusing online Learning Objects (LO), is generally employed to make the development of the mobile learning contents much easier. The major issues encountered with semantic web at times of processing RLO, are that first, not all online LO are accessible on mobile devices, subsequently the LO metadata are not readily available. The Mobile Learning Objects Compilation Framework (MLOC) which is a mixture framework of random forest and semantic web is proposed by this study to address these problems. The hybrid framework would include a method that will generate RLO metadata from repositories and use those metadata to evaluate the RLO. The assembled related RLO are made available to these learning contents to other systems through the web services so that mobile apps can access the RLO easily. This research therefore examines the methods to enhance semantic web in the reuse of LO for mobile devices. The research will first introduce a method to generate learning metadata from public search results based on learning theories. Thereafter, establish the semantic methods to evaluate the RLO and assemble RLO into complete learning units in a repository that can be accessed by mobile devices without hitches The projected framework would be able to search and extract the RLOs which are much more effectual compared to RLOs retrieved by other related mobile apps which in turn confirms that MLOC can be used to process RLOs for mobile devices.

Keywords: *Lecturers; MLOC; RLO; Mobile Learning*

1 Introduction

Reusability is a process within product development lifecycle of software engineering which reduces the production time and resources by using an existing asset within a development of another product. One area that is using software engineering to develop its products is learning systems commonly known as electronic learning (e-learning). In e-learning the digital learning contents are products of software engineering and can be reused in the production of other e-learning contents. Mobile learning which uses mobile devices such as smart phones and tablets is a part and parcel of e-learning and can re-use e-learning contents[1].The learning contents in e-learning are composed of a lot of subject related pieces of information known as Learning Objects (LO).When the LO are reused in different learning systems they are known as Reusable Learning Object (RLO).Due to the limitations of financial and Information and Communication Technologies (ICT) resources, many institutions rely on only lecturers to develop the learning contents on the own without much

ISSN: 2005-4238 IJAST
Copyright ⓒ 2020 SERSC

support from instructional designers and Information Technology (IT) experts. Developing e-learning contents from scratch is difficult and many lecturers take the option of using the RLO found freely on the internet and customize them to fit their students' needs. However, many lecturers fail get the effective learning contents to be used in mobile devices (such as lectures, presentation and simulations). In turn the lecturers have to search and assemble specific RLOs from the search engines which is not an easy task. The search tools that are used by lecturers to reuse RLO from internet adopt Semantic Web technologies. Semantic web is the use of artificial intelligence in the web so that the computers can understand the links and the web resources without human interaction. These features currently lack in SCORM compatible systems which in turn renders the SCORM based RLO not to be effective in mobile learning.

2 Problem Statement

The restrictions of mobile devices pose a problem in reusability using developing and accessing the reusable learning objects in Mobile learning even with the help of Semantic Web(Csete, Wong, & Vogel, 2004). There exists a lack of mutual platform used by mobile devices requires diverse IT skills to be utilized when positioning the RLO into Mobile learning environment. Subsequently, the mobile phone screen size necessitates for small and effectively engineered RLO to be used in mobile devices and the unpreserved power supply makes it difficult to use the mobile devices for a long time(Allen, 2016). This necessitates the knowledge contents to be judiciously checked earlier being allowed to be used for mobile learning. By way of online repositories containing a diversity of learning contents, it is significant to evaluate the content prior repossessing to be used for mobile devices. In addition to that the whole procedure of saving the RLO from an online repository contains numerous processes, of which presents hitches to the lecturers. Primarily, the lecturers must hunt for RLO from different packages from the internet by means of search retrieval tools such as Google (Bhalalusesa& Arshad, 2014).Subsequently, since not entirely all RLO can be used in mobile devices due to lack of compatibility, the lecturers have to appraise if the RLO can be functioning in mobile devices, then download the RLO and store it in the local repository. Subsequently that the RLO has to be combined with other RLO based on the learning templates by the lecturer in order to arrange a complete operative learning contents. In the meantime, all these tasks are tough, most of the lecturer's flop to find the appropriate RLO for mobile devices using the search apparatuses. Problem arises by not knowing the category of reusable learning contents (RLO) that can be operative on mobile devices. After knowing the type of operative learning content for the different types of mobile devices another problem is that of downloading the RLO. Meanwhile mobile devices have boundaries especially of sizes and power, it is unproductive and waste to just download all RLOs from the search results. There must be a mechanism to be put in order to evaluate if the RLO can be effective in mobile devices before downloading takes place. Unfortunately, it is hard to evaluate RLO since majority of them are deposited in repositories without the proper metadata. The common practice is to store RLO in Relational Database (RDB) (Polsani, 2003). Nonetheless the RDB cannot be public in semantic web due to the closed nature of the databases and in turn the ontology structure (such as Resource Description Framework (RDF) is essential to simplify the usage of these Databases. The effective way to create the ontologies for RLO is to map their RDB into RDF. Mapping the RDB into RDF presents a huge challenge in institutions (Cardinaels, 2007). All in all the problem statement can be summarized as the "Gathering of on RLO into effective learning contents for mobile devices using semantic web by lecturers requires knowing exactly what types of RLOs are effective for mobile devices, searching for those effective RLO for mobile devices in the internet, evaluating if the RLOs are operative and collecting a set of the identified effective RLOs

Copyright ⓒ 2020 SERSC

based on their metadata into learning unit in a repository system that can be accessed by mobile devices.

3 Research Questions

The study to be undertaken is directed by the following detailed research questions

i. How can RLO be compiled by semantic web and random forest into effective learning contents for mobile devices for lecturer's usage?
ii. Is Classification Algorithm faster than Transitivity Engine in classification of RLO?

4 Representational State Transfer (REST)

Representational State Transfer (REST), is an architecture that is used to describe elements as data used in applications over the web using HTTP as its primary protocol. When Web Services use REST they are called as RESTFUL services. RESTFUL service is an emerging Web Service that is used by the modern media such as Facebook and Twitter (Ong et al., 2015). The core functionality of REST is just like SOAP in a sense that it serves back end up application to interact with any devices and platform. In REST protocol the client and server exchange raw data instead of methods using GET, PUT, POST and DELETE html method (Kennedy et al., 2011). The table 5 gives a comparison between SOAP and REST.

Table 1: Comparison between SOAP and REST

NO	REST	SOAP
1.	Using HTTP to communicate to your service.	Can use many protocols (HTTP preferred) to communicate the service function
2.	URL represents the structure of your service	Function represent the structure of service
3.	Data is transferred in any standard (XML or JavaScript Object Notation (JSON))	Used only XML
4.	The sever should be stateless	The server may not be stateless
5.	Exchange Data	Exchange Functions to pull data
6.	Uses GET	Uses GET and POST

5 Comparison of Classification Algorithms

The Machine learning from the huge datasets can be very difficult if and once the data has many classifiers. Hence, it is hard to distinguish which classifier is the most imperative to determine the solution. Hence, other classifiers may be weak at some point but could be the most vital ones to determine the outcome. Machine learning algorithm generally take a collection of classifiers based on their weights (Quinlan, 2014). The distinct classifiers (decision trees) after which data are learned are called learners and the combinations of classifiers based on the weighted of the learners are called ensembles (collection of decision trees). Thus, preference of algorithm is data dependent and the more efficient the data, the more efficient is the algorithm. Similarly, the large datasets are good to accomplish better classification since data grows over time and consequently the initial data could be small and rise over a period of time. Correspondingly, the good tuning of data can be imperative in order to get better classification and therefore algorithms that provide all these features must be preferred compared to others. Nevertheless, additional factors such as accuracy, speed, usefulness multi-collinearity, outliers are considered in order to upsurge the effectiveness. The table 2 illustrates the summary on how the classification of algorithms perform on the features that govern the selection of algorithms for a problem. Through evaluation shown in table 2, it can be understood that Random Forest performs well like other algorithms in most of the settings but outperforms others when it comes to over-fitting, number of variables and missing variables. Comparable results have also been obtained by other searchers(Ogutu, Piepho, & Schulz-Streeck, 2011) In addition to that, the study concluded that Random Forest

Copyright ⓒ 2020 SERSC

performs pretty well across different types of datasets(Liaw& Wiener, 2002). Significantly, classification of RLO, Random forest can provide optimum classification. It is consequently important to understand how Random Forest functions so that it can be incorporated into semantic web framework to help classify RLO instead of slow transitive rules.

Table 2: Comparisons of Classification Algorithms

	NB	**SVM**	**kNN**	**NN**	**DT**	**RF**
Classification errors	Low	Low	Low	Low	Low	Low
Numberof variables	Low	High	High	High	High	Low and High
Numberof categories	Low	High	High	High	LowandHigh	LowandHigh
Convergence time	Low	High	Low	High	Low	Medium
Over fitting	Yes	Yes	Yes	No	Yes	No
Multicollinearity	Yes	Yes	Yes	Yes	No	No
Easytointerpret	Yes	No	No	No	Yes	Yes
AffectedbyMissing variables	No	Yes	Yes	Yes	Yes	No

5 The web Service Interface

MLOC permits deployment of RLO hooked on different systems and devices. The deployment of RLO is by web services and the RLO stored in MLOC can be accessed by mobile apps as well as desktop computers which makes it is easier for the RLO to be reachable by all type of devices including mobile application operating systems vary but the main dominant ones are Android and Apple Operating System (iOS). Mobile application requires Java or Adobe flash in order to develop the apps that can be served well in Android or iOS. Subsequently, the MLOC uses web services to receive different topics or courses that are to be used to find the RLO from the internet. In turn MLOC uses web services to send back the list of RLO obtained from the search result back to the web application that initiated the request. The web services that are conducted in MLOC are summarized in table 3 below.

Table 3: MLOC Web services

Course Services	Topic Services	RLO Services
1. Create Course 2. Edit Course 3. Delete Course 4. Import Course Template 5. Upload Course manually 6. Automatic generation from RLO repositories	1. Create Topic 2. Edit Topic 3. Delete Topic 4. Import Topic Template 5. Upload Topic manually 6. Automatic generation from RLO repositories	1. Upload RLO 2. Edit RLO Metadata 3. Delete RLO 4. Search RLO 5. Aggregate RLO 6. Evaluate RLO 7. Map RDB to Ontologies

Table 4: MLOC RLO Assembly Test Report

Level Test Plan (LTP)	This testing was set to check if MLOC can assemble RLO as per requirement of MLOC. All the RLO of a course were passed in the MLOC and all possible combinations of RLO were formed. The expected results were a series of combinations of RLO for a given course.
Level Test Design (LTD)	Expected results is a series of combinations of RLO
Level Test Case (LTC)	- Course A was used to test - All the course RLOs were used as inputs.
Level Test Procedure (LTPr)	- Input a course name - Get all the RLO of the course and combine them separately
Anomaly Report (AR)	A course with many topics cannot be computed

Copyright ⓒ 2020 SERSC

Level Interim Test Status Report (LITSR)	RLO assembly can be done by MLOC but the course must not have more than 15 topics.
Level Test Report (LTR)	- Passed - Need to declare the condition that the course must have less than 15 topics

On the basis of the above, in demand to see if MLOC can aggregate RLO the following test was run in the MLOC database. The test report in the table 4 describes the testing done for combination using Powerset. In the testing, RLO of one course was taken and passed on to the MLOC engine using Semantic Rules in RLO assembly. Hence, the expected results were a list of possible combinations where by at most one of them is selected as the best combination. Significantly, the same procedure is repeated for MLOC engine using Random Forest in RLO assembly and the results of the testing show that the MLOC engine can produce and identify the most effective combination of the RLO of the course in the RLO Assembly process.

6 Comparison with Similar Systems

Comparison with other similar system centered on the variables established. The accuracy, which involves of precision, recall, F Measure and MAE was 166 compared to other similar system. Comparisons were done in two folds. The Random Forest was related other similar classification algorithms to check if its performance is at per with other classification algorithms. However, MLOC accuracy obtained from Random Forest was compared with other similar RLO evaluation systems. The choice of algorithm was based on the related work by Caruana and Niculescu whose findings on the performance of Random Forest was based on their own datasets(Caruana &Niculescu-Mizil, 2006). However, they tested on various multiple types of datasets their results cannot be generalized in MLOC setting until a proper comparison is done using the same dataset since the performance of algorithm may vary when depending on dataset. Subsequently the MLOC and Random Forest were set up in Weka then the algorithms to which Random Forest was compared with, were the ones which could be run in Weka. These are Neural Networks (NN), Support Vector Machines (SVM), Decision Tree (DT), Random Forest (RF) and Naive Bayes (NB). Decision trees algorithm similar to Random Forest and therefore it was not included in the comparison. These algorithms were then run using the MLOC dataset. In general, the results of the tests were similar to the finding by Caruana and Niculescu. For NN using voted perception, classification resulted in lower F-measure (0.851) compared to RF (0.986). It also had very low ROC curve which indicated that it has weaker classification compared to Random Forest. BN also had a little bit lower accuracy (0.95) compared to RF (0.986). One key feature was the fact the results 167 obtained from SVM (0.98) were similar to the results obtained in RF. Both SVM and RF had 98.6% accuracy and ROC curve that was above 0.95. This research study however opted for RF since it allows scalability of dataset and therefore allows data to grow in the training dataset with less fluctuation of the precision. Since MLOC started with small dataset (1099) and expects the database to increase in size it will require re-training at some stage in the future and therefore an algorithm that performs well in both small and large dataset is required. In addition to that Random Forest is easy to interpret and easy to use contrary to SVM which require further technical knowledge of the structure of dataset in order to train and fine tune the model (Rodriguez-Galiano, Sanchez-Castillo, Chica-Olmo, & Chica-Rivas, 2015). Therefore, with changes of the ontology in the future the accuracy difference in RF model will be smaller compared to SVM model.

Table 5: Comparison of Classification Algorithms using MLOC Dataset

Classification Algorithm	NN (Voted Perception)	SVM	BN	Random Forest
Precision	0.808	0.986	0.953	0.986
Recall	0.899	0.986	0.948	0.986
F-Measure	0.851	0.986	0.95	0.986

Copyright ⓒ 2020 SERSC

MAE	0.1011	0.0137	0.0524	0.0275
ROC	0.5	0.956	0.99	0.996

Grounded on the assessments made between Random Forest and other classification algorithms it can be understood that the performance is equally accurate and sometimes more accurate than the other algorithms. These results are in line with the 168 previous researchers such as Rodriguez et al who compared various algorithms using mining data and obtain higher accuracies for Random Forest (Rodriguez-Galiano et al., 2015) Caruana and Niculescu-Mizil Furthermore, other findings concluded that Random Forest accuracy was high when he compared different algorithms using different datasets thus, for this reason random forest can genuinely be used to enhance the performance of semantic operations of classifying RLO as effective or not(Caruana &Niculescu-Mizil, 2006). Significantly, preceding studies also have used different algorithms to get the effectiveness of RLO obtained from search results. However very few have been able to evaluate their mechanisms and document the results for public use. The results from this study were therefore compared and the comparison generalized to other similar studies. Therefore, similar study which was compared against MLOC is the research on personalized learning objects which also reported on the MAE alone (Wang, Tsai, Lee, & Chiu, 2007).The detailed comparisons are listed in the table 6.

Research Work	Personalized Learning Objects Recommendation based on the Semantic Aware Discovery and the Learner Preference Pattern (Wang et al., 2007)	Ranking Metrics and Search Guidance for Learning Object Repository (Yen et al., 2010)	MLOC
Precision	Not given	86.90%	98.6%
Recall	Not given	95.71%	98.6%
MAE	The values oscillate between 0.5 and 1	Not given	0.0275
F-Measure	Not given	0.9109	0.986

Based on the data presented in the table 6 it can be practical that the results generated from MLOC are slightly better than those presented in the(Yen, Wu, Cheng, & Huang, 2010) which also suggests that the results from MLOC would be similarly higher in the studies which fall in the same category such as earlier studies. For this motive, we can conclude that MLOC stand to achieve better results compared to other similar mechanisms of evaluating RLO from search results.

7 Overall Framework Evaluation

The overall framework evaluation is the result of individual parts of the framework that have been evaluated in the experiments. In the first experiment the first part of framework which consist of RLO storage was evaluated by proof of concepts and it was shown that ontological mapping to create knowledge base can be well handled in MLOC. The subsequent experiment which found that MLOC can first search and find RLO that are effective for mobile devices and also it can be used to deploy RLO to mobile devices without any problem. The third part which included RLO evaluation Here the results showed that RLO evaluation in MLOC framework performs well like other systems. It was also random forest improves the semantic web speed and it was understood that RLO assembly is performed well in MLOC framework leading to the reuse of learning contents. Since all the parts of the framework were seen to function well it can be generalized that overall the framework functions well in reusing the RLO for mobile devices. In finalizing the evaluation, it was important to compare how the framework performs in comparison to other related systems (models or framework) that produce learning contents in el-learning. On the other hand, the research works (Trifonova and Ronchetti, 2006, Pathmeswaran and Ahmed, 2011) which can find RLO for mobile devices have a short coming in a sense that the RLO that they find are not assembled with other RLO to make complete effective learning contents.

8 Conclusion

The problems pertinent of mobile devices in accessing mobile worthy educational resources has made the mobile learning reaching its full potential. Hence, this necessitated the techniques of using reusable learning objects that are meanable to mobile devices. in doing this the Representational State Transfer (REST) is used. REST is an architecture that is used to describe elements as data used in applications over the web using HTTP as its primary protocol.

The REST protocol enables the client and server exchange raw data instead of methods using GET, PUT, POST and DELETE html method. comparisonswere made with other algorithms and it was found out that, Random Forest that was used performs well like other algorithms and in most of the settings but outperforms other algorithms when it comes to over-fitting, number of variables and missing variables.The MLOC was used as it permits the deployment of RLO hooked on different systems and devices.

When compared with other related works, the MLOC performed better than the others with based on the metrics of 98.6% for Precision, 98.6% for Recall, MAE was given as 0.0275 where as other related works did not measure this metrics and F-Measure of 0.986. Thus, MLOC was proven to have aided the Lecturers in searching and extracting more effective RLOs as compared to RLOs retrieved by other related mobile apps which in turn further confirms that MLOC can be used to process RLOs for mobile devices.

Acknowledgements:

The authors would like to acknowledge the Malaysian Ministry of Higher Education Under the Fundamental Research Grant Scheme (FRGS) 2018, with account number 203.PKOMP.6711580.

References

[1] Allen, M. W. (2016). Michael Allen's guide to e-learning: Building interactive, fun, and effective learning programs for any company: John Wiley & Sons.

[2] Bhalalusesa, R. P., & Arshad, M. R. M. (2014). Using Search Results Metadata to Discover Effective Learning Objects for Mobile Devices. Paper presented at the Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013).

[3] Cardinaels, K. (2007). A dynamic learning object life cycle and its implications for automatic metadata generation.

[4] Caruana, R., &Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. Paper presented at the Proceedings of the 23rd international conference on Machine learning.

[5] Csete, J., Wong, Y.-H., & Vogel, D. (2004). Mobile devices in and out of the classroom. Paper presented at the EdMedia: World Conference on Educational Media and Technology.

[6] Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. R news, 2(3), 18-22.

Copyright ⓒ 2020 SERSC

[7] Ogutu, J. O., Piepho, H.-P., & Schulz-Streeck, T. (2011). A comparison of random forests, boosting and support vector machines for genomic selection. Paper presented at the BMC proceedings.

[8] Ong, S. P., Cholia, S., Jain, A., Brafman, M., Gunter, D., Ceder, G., & Persson, K. A. (2015). The Materials Application Programming Interface (API): A simple, flexible and efficient API for materials data based on REpresentational State Transfer (REST) principles. Computational Materials Science, 97, 209-215.

[9] Polsani, P. R. (2003). Use and abuse of reusable learning objects. Journal of Digital information, 3(4).

[10] Quinlan, J. R. (2014). C4. 5: programs for machine learning: Elsevier.

[11] Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., & Chica-Rivas, M. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geology Reviews, 71, 804-818.

[12] Wang, T. I., Tsai, K. H., Lee, M.-C., & Chiu, T. K. (2007). Personalized learning objects recommendation based on the semantic-aware discovery and the learner preference pattern. Educational technology & society, 10(3), 84-105.

[13] Yen, D. C., Wu, C.-S., Cheng, F.-F., & Huang, Y.-W. (2010). Determinants of users' intention to adopt wireless technology: An empirical study by integrating TTF with TAM. Computers in Human Behavior, 26(5), 906-915.

Copyright ⓒ 2020 SERSC