MIP: Mashup for Industrial Internship Placement

Oi-Mean Foong\textsuperscript{1}, Rozana Kasbon\textsuperscript{1}, Suet-Peng Yong\textsuperscript{1} and Norfarahin Adiba Abd Kadir\textsuperscript{2}

Computer and Information Sciences Department, Universiti Teknologi PETRONAS
Tronoh 31750, Perak, Malaysia
\textsuperscript{1}foongoimean, rozank, yongsuetpeng}@petronas.com.my, \textsuperscript{2}purpleorkeed@gmail.com

Abstract — In this paper, Mashup for Internship Placement (MIP) system for undergraduate students is presented. One of the problems in Student Industrial Internship Unit is the lack of proper channel for information sharing among the internship students during or after the training. Thus, it is a challenge for students to decide which internship offer to accept due to lack of knowledge about the potential company and the unavailability of information about the job offered. The MIP system for undergraduate students is proposed to integrate disparate data from multiple sources using Yahoo Pipes. The objectives of this research are two-fold: (1) To investigate mashup as a new paradigm for information visualization and (2) To develop a prototype system that incorporates features for user to view the location and information about the company that offers internship placement in Kuala Lumpur, Malaysia. Rapid Prototyping methodology and Yahoo Pipes were used in the design of the MIP system. Data from Google Maps and host company’s details were extracted using Fetch Feed widget in Yahoo Pipes and integrated with the RSS feed to display search results on interactive map. Preliminary results show that the MIP prototype system is a useful Web 2.0 application for assisting users in making just-in-time decision.

Index Terms—Mashup, knowledge sharing, internship placement, new paradigm, decision making

I. INTRODUCTION

Mashup is an application that combines data either through APIs or other sources into single integrated web application [1]. It is an emerging trend in Web 2.0 especially in World Wide Web. In the past, knowledge sharing system was not as exciting as these days with lots of interactive applications are used to extract contents and integrate data from multiple web sources to provide unique and innovative services. One of the goals of Web 2.0 is to make the mashup applications easy to create, use, share and reuse the resources on the World Wide Web [2].

One of the problems in Student Industrial Internship Unit is that there is lack of proper channel to share information among the internship students during or after the training. Thus, it is a challenge for students to decide which internship to accept due to lack of knowledge about the potential company and the unavailability of information about the job offered. Hence, in this paper, Mashup for Internship Placement (MIP) encompasses solving multiple problems such as extracting data from multiple sources, pre-process and integrate the data together. Heterogeneous information could be collected from various host companies. For example, location of the company on the maps, address and details about a specific company, statistics of students who had completed their internship at particular companies and activities that can be done by trainee/interns will be gathered and relocated on the map in the MIP system. These information given by the MIP system would help students to decide which company to choose and which location that suits their preferences. The MIP system can also help supervisors/lecturers to locate the company easily during the visits.

In fact, the MIP is built as a pipeline linking all search queries by the user and integrating with Google Maps from its admin database. The search results are then displayed on a map widget using a marker which will highlight the number of companies based on its categories and location on the map. The objectives of this research are two-fold: (1) To investigate mashup as a new paradigm for knowledge sharing and (2) To develop a prototype system that incorporates features to allow user to view the location and information about the company that offers internship placement in Kuala Lumpur.

Existing Mashup tools such as Openkapow, Dapper, IBM Mashup Center, WaveMakerStudio and Yahoo Pipes allow developers and novices to build mashups without tedious code-writing [3]-[7]. However, designing mashup is not always trivial. There might be some inherent problems of aggregating different data from multiple heterogeneous platforms or sources. The issue of data integration may arise if the pipes do not link properly. As a result, it will display a mashup with no visual map or diagram.

There are two problems with the widget approach using Yahoo Pipes. Firstly, the number of widgets is increasing as mashup increases in its functionality. Consequently, locating a widget to complete the mashup function is difficult and time consuming. Secondly, there is little programming required but the challenge lies in utilizing the Yahoo Pipes widget. This MIP application can be done by many different ways, using REST (Representational State Transfer) protocol, JSON (JavaScript Object Notation), RSS (Really Simple Syndication) feed or existing widget. Even though there are some
issues with the integration of data from different multiple sources, Yahoo Pipes empower users in selecting and customizing its development via widgets.

The Mashup for Internship Placement (MIP) will help users to use the information and shared knowledge quickly and efficiently as users can get all information in one application without searching for other websites. All important information needed about the companies will be combined and tagged together and accessed with few clicks.

The remainder of the paper is organized as follows: Section II describes the related work of mashups. Section III presents the proposed Mashup Engine using Yahoo Pipes Widget in the MIP prototype. Section IV includes the results & discussion of the proposed MIP. Section V describes the outcomes of system performance evaluation. Last but not least, section VI states the conclusion and future work.

II. RELATED WORK

The idea of mash up all information from several web sites was initiated by Tim O’Reilly [8]. Yu et al. reiterated that most mashups applications can access data or other web pages using formal Representational State Transfer (REST) Application Programming Interfaces (APIs), or Web feed formats such as ATOM (atom syndication format), JSON (JavaScript object notation), screen scraping, or RSS (really simple syndication) feed [9]. Zang et al. adopted user-centric approach to investigate user perceptions on usability of mashups [10]. Benslimane et al. explored visual service mashups which could graphically compose and execute queries from internet search [11]. Hogan et al. investigated the feasibility of mashups as a new approach to bioinformatics experimentation [12]. Scotch et al. developed a public health mashup application that could integrate animal, human, and temperature data to assess the risk of West Nile Virus (WNV) outbreaks [13]. Biörnstard, and Pautasso explored a novel approach to mashup engine [14].

Tuchinda et al. deployed a new paradigm of constructing mashup application using demonstration technique instead of using widgets [15]. It was in contrast to creating a mashup application where each web user dragged widgets onto a canvas, customized each widget, and connected them to form a workflow processes that produced the output either in the form of XML, RSS feed or a visual map.

Wang et al. developed a semantic mashup recommender system by promoting places of attractions based upon user preferences and behaviours [16]. Other mashup application on crime reporting known as Crimespotting was deployed by Oakland, California Police Department [17] whereas LA TIMES launched Google Maps Crime Mashup in Los Angeles [18]. The latter crime mashup application consists of a map of mashing up Google Maps and homicide report of all victims in the county. Zhang et al. developed a centralized Web 2.0 services mashup portal for university students to access his/her emails box and course management system (E-Learning Moodle) to browse through lecture notes, newly published assignments/homeworks, and even social network sites (YouTube, Facebook, and Flickr) [19]. Similarly, Kulathuramaiyer highlighted the role of mashups in the creation of digital journals [20]. Liu et al. designed an enterprise architecture using reusable mashup components on theme data (housing prices, location) on location maps [21].

In fact, mashups have been deployed in health industry, bioinformatics, crimes prevention, tourism, service industry, education, social networking and so forth. From literature study, this is the first attempt that mashup is applied in undergraduate industrial internship programme [22]. Therefore, in this extended paper, we apply mashup in our university undergraduate industrial internship programme to facilitate information sharing among students, lecturers and host companies in Malaysia.

III. PROPOSED MASHUP ENGINE USING YAHOO PIPES WIDGET

A pre-test survey was conducted online using Google Document. The aim of the survey was to conduct feasibility study to justify for the existence of the proposed Mashup for Internship Placement (MIP). The pre-test survey was distributed to 30 respondents and their feedbacks were shown in Fig. 1.

Based on the pre-test survey conducted by 30 respondents, 14 students strongly agreed that it was difficult to choose the best offer for the internship placement (Q1). 21 students had no idea of what will be their tasks in the host company (Q2). 13 students strongly disagreed whereas 8 students disagreed they know the location of all companies that offered internship placement (Q3). 19 students thought that it would be useful to have a system that can search for the host company that offers industrial internship placement (Q4). Finally, 17 students agreed that the system will help them to make better decision in choosing the host companies that offer internship placement (Q5). Summing up, most students believed that they were in dire need of a mashup system for internship placement.
The Mashup Internship Placement (MIP) was proposed and created using several modules or widgets in Yahoo Pipes. The MIP consists of fetch data feed module, filter module, location extractor module, union module, user text input module and location builder module as shown in Fig. 2. The MIP system was designed using Model-View-Controller (MVC) system architecture which was deployed in all interactive Web 2.0 applications [21].

Most pipes begin with a data source. These modules grab data from the internet and bring them to the pipe for processing. One of the source modules to create Mashup for Internship Placement is the Fetch Feed module [7].

### A. Fetch Feed Module

The Fetch Feed module allows one or more RSS news feeds as input to the Pipe. The module understands feeds in RSS, Atom, and RDF formats. Feeds contain one or more items, when one adds more feed URLs it will get a single feed combining all the items from the individual feeds. URLs can also be “wired in” to the module from other modules (like the URL Builder). Fetch feed module extract data from the RSS feed from Google Maps.

### B. Filter Module

When there are more data needed to build the Mashup, the filter module allows developer to include or exclude items from a feed. Filter will create rules that compare feed elements and values that are specified for the Mashup. For example, in MIP prototype, rules were created such as ‘permit items where the description contains the text input from user Kuala Lumpur’ or permit items where the item description contains the ‘city location’ from user text input. A single Filter module can contain multiple rules for an item to match all the rules or otherwise.

### C. Location Extractor Module

This module examines the input feed, looking for information that indicates a geographic location. If it finds geographic data, the module adds a location
element to the output feed. This element contains several sub-elements, including: latitude, longitude, quality, country, state, city, street and postal code. Some of the sub-elements may not be included in the resulting feed and this depends on the amount of the information that can be derived from the input feed. Location Extractor gleams location data from URLs in maps.google.com. The Pipe Preview feature will alternatively present the feed in an interactive map if location element is found in the pipe output.

D. Union Module

This module merges up to five separate sources of items into a single list of items.

E. User Text Input Module

In the Mashup for Internship Placement (MIP) system, text input module was used to let user input the company name or location of the company. Other modules can accept text for all modules except the Fetch Feed.

F. Location Builder Module

This module converts a description of a place into geographical data. It recognizes addresses, zip codes, airport codes, cities/country names, and U.S.A. cities or states. The module outputs a location structure separate the fields for city, state, country, latitude, and longitude. This location can be wired into any module that accepts location types.

IV. RESULTS AND DISCUSSION

The Mashup for Industrial Placement (MIP) system was designed and developed using Yahoo Pipes for students who will undergo industrial internship programme. The Yahoo Pipes extract the latitude and longitude of each location of host companies that offer internship placement in Kuala Lumpur using Google map. When user clicks the “Run Pipe” button, the Yahoo pipes will extract, integrate the data from RSS feed and display the search results on the map as shown in Fig. 3. All companies which offer industrial internship are indicated by markers as shown in Fig. 4. The users enter the search ALL query in the search button. Consequently, all companies in the database will appear on the map. In Fig. 5, it displays the location of the company and its details on the map based on user input in the search button.

Figure 3. Mashup of companies which offer industrial internship placement in Kuala Lumpur

Figure 4. Mashup showing all companies on the map
V. SYSTEM PERFORMANCE EVALUATION

The MIP system was tested by 10 respondents (4 males and 6 females) who are undergraduate students in the university. Using Likert Scale range from 1 – 5, where 1 indicates Strongly Disagree, 2 indicates Disagree, 3 indicates Neither Agree/Disagree, 4 indicates Agree and last but not least 5 indicates Strongly Agree [23].

The results of the MIP System Performance Evaluation are shown in Fig. 6 whereas the evaluation statistics of the MIP results are tabulated in Table I. From the respondents’ feedback about the level of acceptance, neither of them disagree nor strongly disagree with the proposed MIP. Based on the MIP evaluation statistics, it reveals that the MIP is easy to use with mean score 4.4 (Q1). The same argument goes to the question (Q2) where it exhibits an interactive way of accessing information with mean score of 4.1. The MIP contains relevant information with mean score 3.4 (Q3). It helps students to make better decision with their internship placement with mean score 4.4 (Q4). It produces the expected result with mean score 3.3 (Q5).

<table>
<thead>
<tr>
<th>Questions</th>
<th>Mean Score</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1. Function is easy to use.</td>
<td>4.4</td>
<td>0.0848</td>
</tr>
<tr>
<td>Q2. System interface is attractive.</td>
<td>4.1</td>
<td>0.1273</td>
</tr>
<tr>
<td>Q3. System shares relevant information.</td>
<td>3.4</td>
<td>0.0141</td>
</tr>
<tr>
<td>Q4. The system helps students to decide on internship placement.</td>
<td>4.4</td>
<td>0.1556</td>
</tr>
<tr>
<td>Q5. The system gives the exact result from your search.</td>
<td>3.3</td>
<td>0.1273</td>
</tr>
<tr>
<td>Average</td>
<td>3.92</td>
<td>0.1018</td>
</tr>
</tbody>
</table>

From the user evaluation results and users’ point of view, the proposed MIP system could help the users especially future students to undergo internship training to decide which company to choose and which location that suits their preferences. The system can also ease the supervisors to locate the company during their visits.

All company details including the name, their business type and contact details also can be found on the map which can be easily used by students for communication via email or mobile phone.

However, some students had anticipated that the MIP system could share more relevant information such as problems & challenges encountered and even alternative solutions in the project(s) using existing on-line discussion tools (eg. Skype) among new vs. current intern students, current intern students vs. lecturers, and lecturers vs. supervisors from host companies. This feature is yet to be implemented/built into the MIP system. That explains the poor mean score of 3.4 (Neutral) for question Q3. Some students tried to search for those companies in East Malaysia but in vain because these data are not captured into the MIP system. That justifies the mean score of 3.3 (Neutral) for question Q5.

Overall, the MIP prototype system yields a satisfactory average score of 3.92 based on the 5-point Likert scale with standard deviation of nearly 0.1018. Generally, the highest mean score of 4.4 from questions Q1 and Q4 reflect that the MIP is easy to use and helpful to students in deciding their industrial internship placements.
Other information that will be available on the visual map is the details of the project that has been assigned to previous students who had undergone the internship training at the particular company. Therefore, new students could anticipate the type of project that may be given to them by the same company. Consequently, new students who will be undergoing industrial internship could make better decision since information could be easily obtained by selecting its marker on the MIP map. Hence, students are more prepared and can boost their confidence before starting the internship training.

VI. Conclusion

Mashups are certainly an exciting new genre of interactive Web 2.0 application. The combination of data and its representation into more graphical visualization will attract and help users to use this system quickly and more efficiently. In this paper, creating Mashup using Yahoo Pipes is emphasized as it allows novice users to create a Mashup without having to write code or understanding programming concept. Using Mashup to share knowledge among students and staff is a new paradigm to interactive communication and information sharing. This will benefit the students who are going for industrial internship to make a wise decision in choosing the best company that suit them. The Mashup for Internship Placement (MIP) can be easily embedded into other client application such as web browser, university e-Learning system or student portal. In future, we intend to extend the rapid development of MIP mashup application in mobile devices [24]. Mobile mashups could be applied to other areas especially in crimes watch/prevention.

Acknowledgement

Special thanks and gratitude to Universiti Teknologi PETRONAS for the support in this research work.

References

Oi-Mean Foong received her B.Sc. Computer Science and M.Sc. Applied Mathematics in U.S.A. She is currently working as Senior Lecturer with Computer and Information Sciences Department at Universiti Teknologi PETRONAS. She is a member of IEEE and Malaysian Information Technology Society (MITS). She has published a number of papers and journals in the areas of text mining, decision support systems and swarm intelligence. Her research interest includes computational intelligence and data mining.

Rozana Kasbon is a lecturer in Department of Computer and Information Sciences at Universiti Teknologi PETRONAS. Her research interest includes information systems, databases and green computing.

Suet-Peng Yong received her Ph.D from University of Otago, New Zealand, Master in IT and B.IT (Hons.) in Industrial Computing from National University of Malaysia. She is currently a Senior Lecturer with Computer and Information Sciences Department at Universiti Teknologi PETRONAS, Malaysia. Her research interest includes computational intelligence, machine learning and multimedia computing.