WEBSITE STRUCTURE IMPROVEMENT USING ANT COLONY TECHNIQUE

Wiwik Anggraeni¹, Agyl Ardi Rahmadi¹, Radityo Prasetyo Wibowo¹
¹Information System Department, Faculty of Information Technology, Institut Teknologi Sepuluh Nopember
Gedung Sistem Informasi, Kampus ITS, Surabaya, Indonesia
Email: wiwik@its-sby.edu

ABSTRACT

The development of the Internet and the increasingly widespread use of it make the company more and more into e-business activities. In the e-business industry, a web site is the place where the company services are placed. So that the websites hold important role in the success of an e-business. Creating web sites that effectively and efficiently could increase user satisfaction, because they can find the information they want quickly and easily, or in other words to save time and costs.

One way to increase the effectiveness of a web site is by improving the structures of its links. For site owners, improving structure may mean costs and time which is not small. Therefore, a right and efficient method for restructuring a web site is needed. One method that can be used is by using Quadratic Assignment Problem (QAP) approach. Quadratic Assignment Problem used the graph concept to map the structure of web sites and to define the connectivity degree and relations between web pages. Then the QAP is used to modeling the page positioning problem. To resolve the problem, ant colony technique is used.

Using the QAP method solved by ant colony techniques, new structures can be obtained and analyzed. From analysis we could know whether the new structure is better than old website structure or not from QAP objective function, closeness analysis, out degree analysis, and also best practices analysis point of views.

Keywords: Ant Colony Meta-heuristic, QAP, Web Sites Structure, Websites, e-business, Website structure improvement

1. PREFACE

Building an effective website as the center of information is important for a company. Because it will indirectly improves customer and website visitors’ satisfaction. In addition to e-business industry, nature of a website as the center of information also means that a website should have good level of effectiveness and usability, so that users or information seekers can easily and quickly retrieve desired information.

There are many ways for increasing information retrieval effectivity in a website, namely the site map, search engine, and intelligent navigation aid tools [6]. As information provider, the owner provides a website with good level of usability [4]. Especially in e-business industry, where company use the website to get information about customer, competitor, and business partner and also to provide information about their company.

A web site structure improvement method conducted in research [1] by using Quadratic Assignment Problem modeling. By using QAP modeling, website structure problem could be analogous to QAP problem. This analogy also defines the basic assumptions required for the problem of site structure to be modeled in the QAP.

After modeled into QAP, the problem will be solved by using Ant Colony Optimization technique. By using ant colony, it is hoped a better new website structure can be found. Once the new structure found, the assessment of the new structure is using Key Performance Indicator (KPI) benchmarking which is predefined. Based on KPIs, better or not the new structure is can be defined.

2. BASIC THEORY

2.1 Graph

Graph is a discrete structure that is used to represent discrete objects and relationships between those objects (Rossen, 2003). Visual representations of graphs consisting of nodes or vertices (V) in the form of dots to represent the object. To represent the relationships between objects expressed by lines or the edge or edges (E) / arcs (A) which are connecting vertices. There are several different types of graphs, which are distinguished by the nature of the inter-connect of the connected vertices.

Graph is used to solve problems in many areas. As an example of graphs can be used to study the structure of the Internet or World Wide Web (WWW). Other examples of graphs can be used in the problem of finding the shortest path from one place to another in a city or a region.
Graph G can be defined as a pair of sets (V,E), which:

\[ V = \{ v_1, v_2, \ldots, v_n \} \]
\[ E = \{ e_1, e_2, \ldots, e_n \} \]

Or can be written in brief notation \( G = (V, E) \)

In the definition, V otherwise not be empty, while E may be empty. Since a graph may not have even a single edge, but the vertex must exist, at least one.

Vertex in the graph can be labeled with letters like \( a, b, c, \ldots, z \), with the original numbers 1,2,3,\ldots or any combination thereof. While the edge which is connecting vertex \( v_i \) with vertex \( v_j \) expressed by pairs \( (v_i, v_j) \) or by using symbol \( e_1, e_2, \ldots \). Thus, if \( e \) is and edge connecting \( v_i \) and \( v_j \), \( e \) can be written as

\[ e = (v_i, v_j) \]

Graphs can be grouped into several categories. It depends on the parameter used for grouping. One of grouping that can be used, based on the direction and weight. Under the direction and the weights, the graph can be divided into four, namely:

- Directed weighted graph
- Directed unweighted graph
- Undirected weighted graph
- Undirected unweighted graph

Graphs can be used for mapping the structure of website which will be modeled into QAP. 

2.2 Quadratic Assignment Problem

Quadratic assignment problem (QAP) is one of the fundamental problems in combinatorial optimization problem (COP) which is a branch of operations research, from the category of facility location problems.

When first introduced, QAP is a mathematical model for the location of indivisible economical activities. The goal is to place \( n \) facilities to \( n \) locations at a cost comparable to flow between the facilities, multiplied by the distance between locations, plus the cost to place a facility at each location. These problems can be modeled with three \( n \times n \) matrices:

\[ A = (a_{ij}), \text{ flow from facility } i \text{ to } k \]
\[ B = (b_{ij}), \text{ distance form location } j \text{ ke } l \]
\[ C = (c_{ij}), \text{ cost of placing facility } i \text{ at location } j \]

So that the QAP in Koopmans-Beckmann form can be written as follows,

\[
\min_{\pi \in S_n} \left( \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} b_{\pi(i)\pi(k)} + \sum_{i=1}^{n} c_{i\pi(i)} \right)
\]

(1)

\( S_n \) is the set of all permutations of integers \( \{1,2,\ldots,n\} \).

A directed unweighted graph \( G = (V, E) \) is connected if there is a path from any vertex to any other vertex. A graph cannot be connected if it does not contain an edge between each pair of vertices. The minimum cost of the graph is the transportation costs due to placing the facility \( i \) to location \( \pi(i) \) and facility \( k \) to location \( \pi(k) \).

In addition to the mathematical model (1), can also be modeled without using the matrix C at all \( c_{ij} = 0 \), for all \( 1 \leq i, j \leq n \). So that there are only two \( n \times n \) matrices in existing model, flow matrix \( A = (a_{ij}) \), and distance matrix \( B = (b_{ij}) \). So the form of Koopmans-Beckmann QAP can be also written as follows

\[
\min_{\pi \in S_n} \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} b_{\pi(i)\pi(k)}
\]

(2)

The artificial ant, gradually build a solution by moving on the graph. The Process of solution construction is stochastically, non-deterministic which means the state system is determined both by the movement that has been estimated from the process itself and by some random elements.

In addition to stochastic process, the solution construction is also influenced by pheromone model, which is a set of parameters associated with components of the graph (either vertices or edges) whose value changes during the construction process.

ACO is basically a paradigm or framework for designing algorithm meta-heuristic for combinatorial optimization problems. ACO also
could be considered as a class of algorithms. In ACO, artificial ants build solutions to the COP by traversing the construction graphs that are fully connected defined as follows.

First, each decision variable $X_i$ is called the solution components and denoted by $c_{ij}$. The set of all possible solutions components denoted by $C$. Construction graph $G_c(V, E)$ is defined by associating the components of $C$ with the set of vertices $V$ or with the set of edges $E$.

The value of the pheromone trail $\tau_{ij}$ associated with the component $c_{ij}$. Pheromone values are generally followed the function of the iteration algorithm $t : \tau_{ij} = \tau_{ij}(t)$. Pheromone values allow modeling of probability distribution from different components. Pheromone value is used and modified by the ACO in solution construction process.

<table>
<thead>
<tr>
<th>Procedure ACO_metaheuristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set parameters, initialize pheromone trails</td>
</tr>
<tr>
<td>SCHEDULE_ACTIVITIES</td>
</tr>
<tr>
<td>ConstructAntSolutions</td>
</tr>
<tr>
<td>DaemonActions (optional)</td>
</tr>
<tr>
<td>UpdatePheromones</td>
</tr>
<tr>
<td>END_SCHEDULE_ACTIVITIES</td>
</tr>
<tr>
<td>END_Procedure</td>
</tr>
</tbody>
</table>

Figure 2.1 ACO Algorithm

Ants are traversing from vertex to vertex through the edges of the construction graph, utilizing information from the pheromone values and gradually building a solution. In addition, the ants storing a number of pheromone components both on the vertex and the edge which they traversed.

Amount from $\Delta r$ of pheromone stored might depends on the quality of current solution. Thereafter the ants utilize pheromone information as a guide to build a better solution. Generally, the ACO meta-heuristic can be described in pseudo code in figure 2.1.

ACO meta-heuristic consists of initialization steps and three components of the algorithm which is activated in Schedule_Actions procedure. This procedure is repeated until completion criteria are met. As an example of the criteria is the maximum number of iterations.

Schedule_Actions procedure does not specifically define how the three algorithms is scheduled and synchronized. Even then they should be executed in parallel and independently, or if such synchronization is required. In most cases when applying ACO to the problem NP-hard, three of these algorithms perform an iteration consisting of:

(i) The construction of the solution by all ants

(ii) Improve the quality of the solution by performing local search. This step is optional.

(iii) Update the pheromone value

3. WEBSITE STRUCTURE IMPROVEMENT METHOD

Web site structure improvement method outlined as seen in Figure 3.1. There are three parts, namely input, process and output.

3.1 Input

3.1.1 Website Structure

One of the input data is a current website structure. The structure here is the current link structure of the website. To be noted, the website structure as an input ignores such things as follows:

- Interface design of the website
- back function from web browser software, whose function is to return to previously explored web pages
- Links which are cross-links and links to web page previously explored. Regarding cross-links will be explained.

Those three are ignored to eliminate the assumption that might be developed, such as whether the design of structures affecting the structure. In addition to those three, there are other necessary assumptions, namely website has only one root page / homepage or start page. So that the structure of the web site will be mapped will always starts from one initial page.

To get the website structure, the extraction of these structures can be done in two ways, automatically or manually. The automatic method can be done using web crawler software, it is a software that can browse the website and collect any information from the website.

For the manual method, it is done by opening the web site, view, and the mapping manually. In this research, the extraction of the web site structure and mapping is done manually.

After extraction of the website's structure, it needs to be simplified. Simplifying the structure was carried out to meet the needs of the three issues mentioned previously. It also considered simplifying the computation of the optimization. With a simple graph, the calculation can also be simpler. It based on assumption that improvements in the level of simple structure will also increase the value at more complex structure.

ISSN 2085-1944
3.1.2 Website Server Log File

The log file for a website could be taken from web server directly. After that, the log file processed by using log file analysis software. Example of the software used is SmarterStats 5.1, which is a product of SmarterTools Inc. The log file processed to get report about popular paths in the website.

After get the report, next is do the path analysis. Path analysis basically is analyzes the report generated from log file analysis software. Especially the report about most frequent path used in explores the website. Path analysis could be done by read the report generated from software. From report, usually it could be seen the paths which is frequently used by the visitor and how much that path is used. And also from the report, User Visiting Sessions (UVSs) and User Visiting Pattern (UVP) can be defined.

User visiting sessions or UVSs is a usage pattern from the paths traversed by website visitors. UVS data obtained from processed website’s log file, which is the paths that frequently traversed by website visitor while in a website.

Those paths are listed and sorted according to how often it is crossed. Previously, website pages have also been labeled by using a number or letter, as a marker of each page of the website. From those, one row matrix can be formed which contains the representation of whether a website page visited or not in the path. The number of columns in that matrix is the number of paths on the website.

User visiting pattern or UVPs is a matrix \( V_{MxN} \) where \( M \) is the number of UVSs from a website and \( N \) is number of pages on that website.

Table 1. Example of UVP with 10 UVSs from a website with 14 pages

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

3.2 Process

After getting the website structure data and UVP data from processed website server log file, then next step is modeling QAP from existing website structure problem.

3.2.1 QAP Modeling

If the graph of the popular path structure from a website is G, then

\[ G = (P, L, W) \]

Defined as:

- \( N \) = total number of pages on website
- \( P = \{P_i | i \in [1,N]\} \) is the set of all vertices in G, which in this case is the website pages
- \( L = \{L(i,j) | i \neq j, i, j \in [1,N]\} \) is the set of all edges in G, which in this case is the link between website pages
- \( W = \{W_{ij} | i \neq j, i, j \in [1,N]\} \) is the set of all edges’ weight in G

\( W_{ij} \) is the probability of \( L(i,j) \) selected by website visitors who have accessed \( P_i \) defined as follows

\[ W_{ij} = R_{ij} / \sum_{k=1}^{O(D(i))} R_{idk} \]

(3)

With \( R_{ij} (R_{ij} \in [0,1]) \) is association degree from \( P_i \) to \( P_j \) which defined as follows

\[ R_{ij} = \frac{|P(i, j)|}{|P \cdot i|} \]

(4)

If \( |P \cdot i| \neq 0 \) and \( R_{ij} = 0, i, j \in [1, N] \). Association degree is the level of a website pages associated with the other pages in the paths which website visitors traversed, or it can be considered as the probability of conditions under which the users have visited \( P_j \) via \( P_i \). Also \( D \) defined as set of all destinations from \( P_i \).

Weight matrix for QAP is the Connectivity Degree. According to [2], \( C_{ij} \) is used as connectivity degree between website pages. The greater the \( C_{ij} \), hence more easily the visitors found \( P_j \) through \( P_i \), with \( C_{ij} \) defined as follows

\[ C_{ij} = w_1 + w_2 + \ldots + w_m \]

(5)

With \( m \) is the number of links route from \( P_i \) to \( P_j \), or in other words is how many paths that can be traversed in order to get to \( P_j \) from \( P_i \). In addition there are three basic assumptions that must be made in modeling the website structure improvement problem into QAP. These three assumptions are as seen on table 3.2.

As the objective function for the QAP is,

\[ \text{Objective Function} = \sum_{i=1}^{N} \sum_{j=1}^{N} C_{ij} W_{ij} \]

\[ \longrightarrow \text{Minimize} \]

\[ \text{Subject to} \sum_{j=1}^{N} W_{ij} = 1, \sum_{i=1}^{N} W_{ij} = 1, \forall i, j \in [1, N] \]
Min\( \sum_{i \leq j \leq m} \left[ C_{ij} \times d(a_i, a_j) \right] \) (6)

Having the assumptions defined and objective function is known, the next is how to form the QAP model. Namely the distance matrix and weight matrix in the website structure problem. Both will be used in the calculation when solving QAP model with ant colony technique.

In forming the required distance matrix for QAP model, it is necessary to calculate by performs operation on the graph of the website structure that has been mapped. The graph used is the one which is the mapping of whole website.

<table>
<thead>
<tr>
<th>Param.</th>
<th>QAP Definition</th>
<th>The assumption in the structure of the web site</th>
</tr>
</thead>
<tbody>
<tr>
<td>( d(a_i, a_j) )</td>
<td>The distance between two facilities</td>
<td>The distance between two web pages based on the number of links/steps to reach any target page initiating from an initial web page</td>
</tr>
<tr>
<td>( C_{ij} )</td>
<td>The interaction cost of two facilities per distance unit</td>
<td>The amount of interaction or connectivity between two web pages based on the ( C ) parameter</td>
</tr>
<tr>
<td>( TC(a) )</td>
<td>Total cost of current facility layout</td>
<td>Total cost of current website structure</td>
</tr>
</tbody>
</table>

The generator worked by using cost matrix and number of pages from the QAP model as input. The generator is made using the basic concepts of ACO that has described.

The ACO concept used for the generator is the state transition rule and pheromone update rule of the ant system. State transition rule from [1]:

\[
P_{ij}(t) = \frac{\tau_{ij}(t)^{\alpha} \eta_{ij}}{\sum_{j \in \text{allowed nodes}} \tau_{ij}(t)^{\alpha} \eta_{ij}}
\]

With,
- \( P_{ij}(t) \) is the probability of selecting the page \( j \) from \( i \) at the next step
- \( \tau_{ij}(t) \) is the amount of pheromone on the edge \( i-j \)
- \( \alpha \) is the weight of pheromone in computing probability

The pheromone update rule:

\[
\tau_{ij}(t+1) = (1-\rho)\tau_{ij}(t) + \frac{Q}{\ln Q \sum_{k \in \text{allowed edges}} \frac{1}{\tau_{ik}(t)^{\alpha} \eta_{ik}}}
\]

With,
- \( \tau_{ij}(t) \) is the amount of pheromone on the edge \( i-j \)
- \( \tau_{ij}(t+1) \) is the secondary amount of pheromone on the edge \( i-j \)
- \( \rho \) is the evaporation rate of pheromone
- \( Q \) is the total amount of pheromone on the Ant System

The generator need matrix \( C \) for input. After that the matrix \( C \) is used to get the structure which has a smaller cost value than the previous structure. How the generator work is explained as follows.

Each ant always starts from the vertex 1, which in this case is the initial page of the website. Each ant traverses from the initial vertex to the next vertex. Selection of the next vertex based on the state transition rule. Once selected, the ants moved into the vertex. After that ants choose the next vertex to be visited. How many times do ants move from vertex 1 to another depends on the defined hop parameters.

After an ants move as many as hops, the ant is stopped and the next ant do the same process. It is processed until all the ants had traverse or all the vertices that are available have visited. If the first ant has visited vertex 4, then on the second hop it is not allowed to visit the vertex 4, and so are for the next hops. When the next ants start to do the hops, for the first hop the ant is allowed to visit the vertex 4, but not on the second and so on.
State transition rule

1

2

3

4

... n

1st ant

2nd ant

n ant

Figure 1 Simple illustration how the generator work based on ACO

So a vertex not allowed to be visited again if in the previous hop that vertex has been visited. Different ant can visit it if has the same hop number when will visit that vertex. Suppose that the first ant visiting the vertex 4 in the first hop, then the number-two ants can also visit the vertex 4 in the first hop, but not in the second hop and so on. For each vertex is visited, their status marked as visited. The purpose is to stop the current iteration.

The process is limited by the number of iterations. Each iteration has a number of ants. Each ant has a number of steps (jumps) itself. An iteration is completed when each ant has completed the hop, or if all nodes have been visited. In each iteration, the global pheromone update to change the probability of each point visited. So each ant can be expected to form a graph with no isolated vertices. Figure 1 is the illustration how the generator work.

3.3 Output Analysis

After QAP modeling and the use of ACO algorithm, which in this case a generator for website structure created, implemented and used to solve the existing QAP model, then expected to get the output in the form of optimized new website structure.

An analysis of the new website structure is conducted after get the new structure. The analysis was done with reference to predefined KPI. In this case, the goal of the new structure is how visitors can be more quickly and easily to find the information their need. So that the KPIs established for the purpose are:

- Amount of time when visiting the website
- Number of pages viewed per visit

In the analysis assumed that the number of pages viewed per visit is a maximum for each visit, which means all the pages in the web site visited. So every visitor explored every page in the website. The website visiting time taken from the website’ log file analysis for the average time per page visit. Then every average time of the web pages are summed.

In regards to defined KPIs, if the time required by based on scenario to browse the web site is less than when the scenario is run on old structures, new structures will be considered better than the old structure.

In addition, analysis in terms of graph analysis is conducted from the statistical values generated from the graph analysis. And also in terms of practical analysis, that is if the structure is used in the real world, how the structure affects the factors that are ignored in the calculation such as website interface design and placement of links in a web site.

4. CASE STUDY

4.1 Input, Process, Output

In the case study, the website taken as a case is the English version website of Institut Teknologi Sepuluh Nopember (ITS), which has initial page address in http://www.its.ac.id/en.

Log files from the web site cases taken directly from the web server where the website are hosted. The website is hosted in an Apache web server. The log files is located at the website directory / logs from the main directory of website with the type of log file is the access log which records website access. Log files taken are have recording period from 22 December 2009 to 20 January 2010.

After get the website log files from website and get it prepared, then next is the selection of software to process the log files. In this case study SmarterStats 5.1 software which is a product of SmarterTools Inc. is used.

Software used to process the log files is used to generate statistical reports from the website. Reports which is required, as described, is a report on the popular paths frequently traversed by visitors of website.

In the report, there are paths used by visitors while browsing the website. But due to the Indonesian and English version from the ITS website proved to have joined access record, so that
from the reports there are also paths used by visitors for Bahasa Indonesia version of ITS website.

Because of it, it is necessary to segregate popular path report data that can be done manually or by using the log files processing software used. In this case study, it is manually segregated. So we get the popular paths from the case website.

In this case study, because the small value of the use of some popular paths, then the popular paths have ignored the minimum limits of the number it is used.

Of the two hundred data obtained from the previous segregation, the data suitable for use are selected again because the data are still mixed with the Indonesian version of ITS website data. From the selection, it is found approximately fifty data used for further processing. Furthermore UVP data obtained from analysis of the popular paths.

Next is using the UVP data and structure of the website obtained as input for the software made and coded by using Java language. The software run with the parameter Q = 100, alpha = 0.1, rho = 0.9, and the maximum is one hundred iteration. Then the final result has been obtained.

For this case study, we got a good structure as a new structure of the website case with less total cost value. With the old structure, the case website has a total cost of 3.902, after the software used, obtained the total cost of 3.4509. Figure 5 is a new structure generated.

4.2 Output Analysis

From the output obtained from the software, we then analyzed these results to be assessed whether these outcomes can be judged better than the old structure or not.
From the output results, the total costs were analyzed. From the new structure, the total cost value is 3.451; while the old structures, the total cost value is 3.902. It concluded that the new structure better than the old structure in terms of total cost because the value has met the objective function.

For KPIs analysis, from the reports generated by the web log analysis software, it cannot be found the average visit time for each website pages. So the KPI’s analysis cannot be done.

Because the KPI analysis cannot be done, another analysis needs to be done to evaluate the new structure of the case website. Analysis undertaken is graph analysis for new structure, and also practical analysis. Those analyze are Closeness analysis, Out Degree analysis, and analysis of practicability.

4.2.1 Closeness Analysis

Closeness analysis is an analysis to determine the level of closeness of each vertex in a graph. The smaller the steps required, the higher the value of closeness. For the case of website structure, closeness shows the level of convenience to roam from one page to other pages. In this analysis, closeness analysis to the old and new structures of case website is shown in table 2.

<table>
<thead>
<tr>
<th>Notes</th>
<th>Old Structure</th>
<th>New Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>0.01369863</td>
<td>0.01639344</td>
</tr>
<tr>
<td>Maximum</td>
<td>1.0</td>
<td>0.02439024</td>
</tr>
<tr>
<td>Mean</td>
<td>0.032140605</td>
<td>0.01876567</td>
</tr>
</tbody>
</table>

From the average value of closeness, it can be seen that the old structure has a higher value than the new structure. From the decrease in the average value of closeness from the old structures, it can be concluded that the new structure will be made visitors experience more difficulties when browse the website.

4.2.2 Out Degree Analysis

Out Degree analysis is an analysis to show the level of edges out from each vertex in a graph. The higher the more out degree value for each vertex in a graph. Table 3 shows the comparison value for the out degree analysis of new and old structures for the case website.

<table>
<thead>
<tr>
<th>Notes</th>
<th>Old Structure</th>
<th>New Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>0.023809524</td>
<td>0.45238096</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.4047619</td>
<td>0.97619045</td>
</tr>
<tr>
<td>Mean</td>
<td>0.023255814</td>
<td>0.64451826</td>
</tr>
</tbody>
</table>

From the table we can see that for the minimum and maximum value and the average of the out degree has increased for the new structure, which means that in every page of a new structure there are many new links to other pages. Thus the analysis shows that the new structure requires a considerable change in the problem of placement of new links on each page.

4.2.3 Practicability analysis

Practicability analysis conducted here to see how the new structure can be used in terms of interface design and links layout of the website. From the graph in Figure 4.4 can be seen that from each vertex there are many edges to mutually connect with each other vertex. Thus on every page there is a new link that did not exist.

From the new structure it can be seen that there are pages which is not mutually connected. There are pages that are linked to all other pages, for example, are page 1. But there are also pages that are not connected to all the pages. An example is page 2. If new structures are used, from terms of the web site's interface it can use the rollover menu type to place links. The example of rollover menu is shown in Figure 4.5.
But from new structure is obtained, each page does not always linked to every page. In terms of practicality, it would be very difficult for web designers because it means every page has a number of different links. Yet at every website it is common that each page has a form of the same view (template) with the same number of links. So that the new structure from terms of practicality cannot be implemented.

5. CONCLUSION AND FUTURE DISCUSSION

After doing the case study and analyzing results from the website structure improvement method with ant colony technique, we can conclude the following things for the website structure problem issues:

- Website structure improvement by using ant colony techniques done by modeling the website structure problem into QAP where the solution found using the website structure generator that works with the concept of ant colony.
- Implementation of improved methods of website structures improvement using ant colony has been able to do with making software for modeling QAP and web log analysis software for initial analysis of log files from a website.
- It has been found a web site proposal to improve the structure from English version of ITS website.

The suggestions for future discussion related to the research that has been done can be given as follows:

1. When mapping the structure of the website, it is done manually. For a web site with a vast number of pages can be quite difficult. For the next for mapping the structure of web sites can be made a web crawler that can map and construct graph of website structure automatically.
2. The website structure generator that is used only uses the basic concepts of ACO. The results obtained from structure improvement can be considered bad in terms of closeness and practical analysis. If possible, the algorithm for website structure generator is specially made so the result of improvement will be better.

6. REFERENCES
