An Android Application Using Machine Learning Algorithm for Clique Detection in Issues Related to Transportation

https://doi.org/10.3991/ijim.v16i14.30625

Fairouz Hussein(✉), Subhieh M. El-Salhi, Rajaa Alazazma, Tasneem Abu-Hantash, Haneen Abu-Hantash, Heba Thaher
The Hashemite University, P. O Box 330127, Zarqa 13133, Jordan
fairouzf@hu.edu.jo

Abstract—The school bus routing problem relates to designing the optimum distribution and collection routes for the school buses serving geographically scattered students and has been the focus of many academics for a long time. In order to meet the tremendous distance and the complex geographic structure, we proposed a method to divide the geographic regions into cliques that contain houses in close proximity to each other. This paper presents the development of a new variant mobile application with a local search algorithm that has been designed to detect routes for school transportation. The proposed system has the following advantages: (1) grouping the students' houses into cliques (clusters), as each clique includes the houses that are close to each other, and a bus is assigned to each group; (2) the ability to determine the route for designated locations, resulting in: (3) more effective use of time and fuel; and (4) integration with the Google API to handle access to map display, data downloading, and Google Maps servers. The application is implemented on the Android operating system since Android is currently widely utilized. A survey was conducted to evaluate the proposed application. The findings from the reviews were impressive and substantial. The proposed solution to the school bus routing problem can be applied to solve problems related to the vehicles of institutions and organizations.

Keywords—machine learning, clustering algorithms, maximal clique mining, intelligent transportation systems, android platform

1 Introduction

In recent years, urban life is getting more and more complex with the rise of the population, the increase in population density, and the spread of buildings. These factors lead to an increased need for easy mobility at a time when traditional transportation methods do not guarantee a sustainable future, and they may not be able to adapt to the increasing demand for easy and efficient transportation. For many years, the response to this growing demand was mainly to build more new infrastructure and increase the number of vehicles on the roads. However, the rapid development in information and communication technology (ICT), along with advances in the field of intelligent transportation systems (ITS) in many countries around the world, has encouraged researchers, decision-makers, and entrepreneurs to search for new solutions to mitigate the challenges of providing accessible transportation that meets the needs of humanity. Nowadays, the increasing development of technology directs the innovations to facilitate the work of transportation. The configuration of roads can reflect the economic development in a region or country as well as the level of traffic.

When schools in Jordan open their doors, nearly 2.027 million students will be headed back to the classrooms. And keeping in mind that much consideration is given to what occurs while students are at school, most probably don't think much about how they arrive there. The lack of consideration is astounding, especially given that school transportation accounts for between 20 and 25 percent of rush hour traffic. There are important factors to ensure that students safely arrive at their schools without delay: compliance with traffic laws, driving ethics, the nature of roads, and knowledge of drivers on roads, etc.

In general, there is a limited number of buses in schools. Each bus has a specific capacity of students. Students are distributed in different regions and may be away from school. All students in a given region have at least one bus allocated to their region. The students must arrive at school at a specific time. Therefore, time is of considerable concern in the arrival of students to school.

However, organizing the school's transportation is a very complicated process that consumes time and needs the efforts of all the people in the school. Many algorithms have been used in school transportation systems to generate bus routes and schedule buses. These systems use machine learning clustering techniques to assign the best cluster (route) to the school bus. Examples of these clustering algorithms are the K-means algorithm, the Mean-Shift algorithm, and the Branch and Bound algorithm. The Branch and Bound algorithms are a well-known approach that can be applied to find optimal solutions to many optimization problems, such as the Traveling Salesman Problem (TSP). Inspired by the Branch and Bound paradigm, we have adapted the Bron-Kerbosch algorithm to solve the school bus routing problem.

In this paper, we proposed a new system that selects the route for the buses to pick up the students from their houses. The students are grouped into cliques (clusters) while satisfying bus capacity constraints. The students in a clique (route) are assigned to a bus, which satisfies some constraints in the problem, such as the maximum number of students in a bus and the minimum distance between students' houses. This work presents a new approach to utilizing machine learning algorithms in clustering analysis. With a framework as extensive as this (and one that provides such a vitally important service), it is sensible to expect that it would be a noteworthy piece of transportation approach talk. The proposed method is based on the Traveler Salesman Algorithm. To achieve these goals, we have proposed a system that can optimize the TSP to find an optimal route in a short time. As a result, many of the problems faced by school transportation have been solved in our proposed system, such as: bus distribution, arranging the arrival and departure times of students, and fuel costs.

Due to the importance of development in the educational and academic fields, smart devices and laptops have spread in homes and even schools. It has become imperative to take advantage of these modern technologies to benefit students in utilizing their time and not wasting it on the roads. Therefore, we designed an electronic mobile application that works with the school's management system. The application facilitates communication between the school's administrative staff, students, their parents, and bus drivers. Through the application, the school bus is fully tracked on the mobile phone, through an account for both the administration, the student's parent, and the driver, which makes work and follow-up during the student transportation process easier and more accurate.

Getting one's current location via network positioning or GPS has become one of the most attractive fundamentals in most location-based service applications. The proposed application can offer an optimal route for designated locations using Google Maps, resulting in an efficient use of time and fuel. The application is implemented on the Android operating system since Android is currently broadly utilized. The framework is integrated with the Google API to handle access to map display, data downloading, and Google Maps servers.

In this work, we designed and implemented a mobile application with a machine learning algorithm to solve the Traveling Salesman Problem (TSP). The main contribution of this work is to utilize the Bron-Kerbosch algorithm in presenting an intelligent mobile application for school bus routing and scheduling.

The next sections will present an overview of the Bron-Kerbosch algorithm, specifically related to solving the Travelling Salesman problem, as well as a more in-depth analysis of the proposed methodology. Related research in the field will be described first, with a discussion of what distinguishes it from this research. Explicit outcomes from the current research will be discussed, as well as ideas for extending this research in the future.

2 Related work

The school bus routing problem is one of a wide set of problems recognized as the vehicle routing problem, or VRP. However, it differs qualitatively from the simple vehicle routing problem in that students are not like products that are picked up from one place and sent to another. In the literature, there are many studies regarding school bus routing problems (SBRP). One of the important previous studies in this field was carried out by Newton and Thomas [1]. The School Bus Routing Problem (SBRP) is concerned with designing the distribution and combination of optimal routes for school buses serving geographically dispersed students and has long been a focus of many scientists' attention. This problem (SBRP) comprises in its context small problems that motivated scientists to study and research, such as: selection of bus stops, data preparation, scheduling school bells, generation of bus routes, and bus scheduling. In this research, we were keen to solve two of these problems, which are the problem of bus route generation and the problem of bus scheduling. According to [2] research, the problem of bus route generation with limited routing time and bus capacity constraints is considered an NP-hard problem. Therefore, some researchers resort to heuristic methods [3, 4, and 5].

In recent years, there have been numerous studies examining various strategies and solutions for resolving the school bus routing problem (SBRP). The SBRP can be classified as a single-load or multi-load problem [6, 7], depending on the bus load of students from the same school or from different schools. In this study, we focused on the single-load, due to financial considerations preferred by school administrations on the one hand, and the desire of parents on the other hand.

According to Spada et al. (2005) [3], research on the SBRP problem can be divided into two categories: a home-based approach and a school-based approach. The home-based approach is concerned with finding the best way for students to travel from their home to the school. In the school-based approach, the problem of each school is solved separately by generating routes, and these routes are allocated to buses and organized according to different school times. Furthermore, in the school-based approach, each school allows only its students to use its buses (no mixed travelers). The work presented by [4] preferred the school-based approach, and they proposed a heuristic approach for the problem of school bus routing and scheduling.

Braca et al. (1997) [8] followed the home-based approach. They solved the problem in a greedy procedure by constructing a route between houses and schools that minimizes the length of the route. In one way or another, the solution remains constant, unless there are buses holding a few students due to absence. Unlike the school-based approach, the home-based approach allows travelers from different schools to board the same bus (mixed travelers). Martínez et al. (2011) [9] split the problem down into two parts. They determined the most suitable spots for students in the first stage, and then found the best routes to serve those stops. They attempted to reduce the amount of time pupils spent walking to their designated bus stations. [10] simulated the SBRP and attempted to reduce the total journey time. To overcome their challenge, they offered a hybrid model and a clustering strategy with time windows.

From the modelling perspective, the most common objectives of the routing problem are: 1) minimizing the total tour time [11]; 2) minimizing the number of tours; 3) balancing the bus capacity, and 4) minimizing the distance students walk to get to the bus. A literature review of some of the objectives of routing problems is mentioned in [12]. Although bus tour time, number of tours, bus capacity, and student walk bus have all been studied separately in the literature, all four objectives are highly interrelated. In addition to the time it takes for the bus to finish its longest route, school districts are concerned with minimizing the overall distance traveled. Researchers have used tabu search [13], metaheuristic search [14], vertical algorithms [15], and hybrid algorithms [16, 17, and 18] to produce a variety of solutions for this concern. Some scholars have recently concentrated on optimizing school transportation from many perspectives in the literature, such as the number of buses [19], the number of students [20], supply transportation [21], school transport time [22], demand analysis [23, 24], and etc.

Currently, satellites are used to locate vehicles (GPS is the most widespread). In most location-based service applications, getting one's current location via network positioning or GPS has become one of the most appealing fundamentals [25]. The work offered by [26] involved the development of a mobile application for smart bus transportation utilizing GPS. The work in [27, 28] investigated actions using machine learning techniques to categorize them into distinct groups. The obtained data was used to train machine learning algorithms to classify the actions using mobile devices.

In the literature, the objectives varied between 1) minimizing the tour time; 2) minimizing the number of tours; 3) balancing the bus capacity, and 4) minimizing the student walk. However, there were no studies aimed at reducing the number of buses per round. To fill the gaps in the literature, this study proposes a new approach to the SBRP problem based on a routing approach and simultaneously allocating students to those specified bus routes. Our research differs from previous studies in that it considers two main objectives: (1) minimizing the number of buses required and (2) minimizing total bus travel time. The main constraint is the limited number of buses. Generally, the existing school systems support academic and administrative aspects. In this research, we present a specialized and integrated system that concerns organizing buses, routing and transferring students, with a mobile application for parents, teachers, and drivers. To the best of our knowledge, there are no existing mobile systems in schools that group the students' houses into cliques (clusters). And this is what led to the novelty in presenting this research.

3 An overview of Bron-Kerbosch algorithm for solving the Travelling Salesman problem

This research addressed the following objectives:

3.1 Minimize the bus travel time

This objective is important from the point of view of the parents and the school. The objective is formulated as a traveling salesman problem (TSP). The traveling salesman problem is about a salesman who should make a tour of a number of cities using the shortest path available. The salesman must visit all cities by taking the shortest route, visiting each city only once and returning to the city of departure. Unfortunately, there are n! permutations of paths for n cities, so this problem can be used for very small n, which is O (n!). For example, if it takes one microsecond to calculate 1 permutation, the algorithm takes about 3.6 seconds for n = 10 and about 6 months for n = 15. The Traveling Salesman Problem is an NP-hard computational optimization problem. When looking for an algorithm to handle an optimization problem, we frequently ask ourselves, "Does it locate all the exact optimal solutions, and how can we know that they are optimal?". The TSP can be solved by using one of the available algorithms, such as dynamic programming or heuristic search algorithms.

For n number of cities, the number of paths that must be explored is n! Miller–Tucker–Zemlin (1960) articulated the classic traveling salesman problem formally as follows:

$$Minimize\ Z = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \tag{1}$$

Subject to

$$\sum_{i=1}^{n} x_{ij} = 1 \qquad j = 1, 2, \ldots, n \tag{2}$$

$$\sum_{j=1}^{n} x_{ij} = 1 \qquad i = 1, 2, \ldots, n \tag{3}$$

$$y_i - y_j + n x_{ij} \le n - 1 \qquad \forall i \ne j \tag{4}$$

$$x_{ij} = \{0 \ or \ 1\} \qquad \forall i,j \tag{5}$$

Here, n is the number of cities, and c_{ij} shows the distance between the cities *i* and *j*. Term x_{ij} shows whether it goes from city *i* to city *j*. Equations (2) and (3) are aimed at guaranteeing that each city will be visited only once. In equation (4), the variables y_i are arbitrary real numbers that satisfy equation (4) to prevent sub tours being created. The term x_{ij} in equation (5) is 1 if the salesman travels from city *i* to city *j* and 0 if there was no tour. We used the Euclidean distance metric to compute the distance between houses, which is defined by:

$$\boldsymbol{c_{ij}} = \left[\sum_u (i-j)^2\right]^{1/2} \tag{6}$$

In the Google API system, each house is represented by its latitude and longitude values. Note that the latitude and longitude values are not coordinates in a Cartesian system. Therefore, these values will be converted to values at the Cartesian level. A question arises here: "How to calculate the distance between Latitude and Longitude points?" The answer is one we should know both from Trigonometry and from Algebra. Assume we are given two houses h_1 and h_2 presented by their latitude and longitude values, $h_1(\theta_1, \varphi_1)$ and $h_2(\theta_2, \varphi_2)$, respectively. In three-dimensional Cartesian space, each point has three coordinates. To compute the Euclidean distance c in Cartesian coordinates between $h_1(p_1, q_1, r_1)$ and $h_2(p_2, q_2, r_2)$, use equation (6):

$$c^2 = (p_1 - p_2)^2 + (q_1 - q_2)^2 + (r_1 - r_2)^2 \tag{7}$$

In spherical co-ordinates, we have

$P = R \cos\theta \cos\varphi$
$q = R \cos\theta \sin\varphi$
$r = R \sin\varphi$

and

$R \approx 6.378 \times 10^3$ m

Put the values p, q, and r in equation (7), then we have

$$c^2 = R^2 \, [2 - 2\cos\theta_1 \cos\theta_2 \cos(\varphi_1 - \varphi_2) - 2\sin\theta_1 \sin\theta_2] \tag{8}$$

Where R is the average radius of Earth at the equator which is 6,378 kilometers (3,963 miles). Then, the distance in Cartesian co-ordinate c in (8) is converted to the value C measured along the surface of the earth, which is given by equation (9):

$$C = 2R \sin^{-1}\left(\sqrt{\sin^2((\theta_2 - \theta_1)/2) + \cos\theta_1 \cos\theta_2 \sin^2((\varphi_2 - \varphi_1)/2)}\right) \tag{9}$$

For example, the distance between Al Zarqa School (36.098891 N, 32.026309 W) and a student's house (36.122446 N, 32.064756 W) is 4.822597 km.

3.2 Minimize the number of buses

This objective is important from the school's point of view to avoid the costs associated with purchasing redundant buses. The minimal number of buses M that is required to serve N students can be calculated using the equation below:

$$M = \text{Minimize } k \quad \text{s.t.} \quad \sum_{i=1}^{k} i \le N \tag{10}$$

This objective can be formulated as a clique mining problem (assume one bus for each clique). Inspired by the clustering framework which is based on the branch and bound algorithm, this paper referred to the maximal clique mining approach [29]. Maximal cliques are a systematic technique to describe communities in the community detection problem, also known as graph clustering. The maximal clique problem (MCP) is based on the idea of arranging the set of all routs into smaller sets of routs that have common properties. The MCP is known for its significance in modeling real-world applications and can be found in many areas. For example, in data mining, information retrieval, social networks, bioinformatics, media gateways, service functions, etc.

The objective of this research is to use the maximal clique problem to divide the geographic regions into sub-regions where each sub-region covers nearby houses and a certain bus is allocated to a certain sub-region (clique). Figure 1 shows an illustrative example of how the proposed problem is modeled.

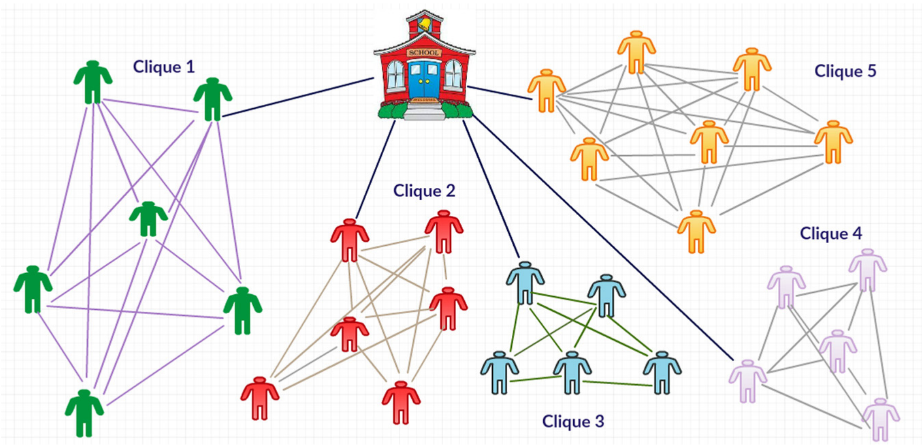

Fig. 1. A motivating example

Let us recall a fundamental definition in graphs about clique.

Definition. Let G = (V, E) be an undirected graph with n = |V| and m = |E|. We define the size of G, denoted by |G|, as |G| = (n+m). A clique C in a graph G is a subset of vertices, where $C \subseteq V$ such that, for any two vertices $v_i, v_j \in C$, there exists an edge $(v_i, v_j) \in E$. Mainly, if there is a subset of **k** vertices that are connected to each other, we say that graph contains a **k**-clique. C is called a maximal clique in G if C cannot be extended by adding any more adjacent vertices.

Finding all the cliques is expensive as the number of cliques grows exponentially with every node added. So we apply the Bron-Kerbosch [30] algorithm, which is a recursive backtracking solution that computes all cliques in linear time. It takes $O(3^{n/3})$ time in the worst case. This is optimal as a function of n, because it can have up to $3^{n/3}$ maximal cliques in an n-vertex graph. The Bron-Kerbosch algorithm is widely used and is denoted as one of the fastest algorithms [31]. Figure 2 shows a graph of 6 vertices after applying the Bron-Kerbosch algorithm.

Algorithm: **BRON-KERBOSCH ALGORITHM**

Input: P is the set of all vertices in Graph G, and two empty sets R and X
Bron-Kerbosch (P, R, X)
if {P = X = Φ} then
report R as a maximal clique
else
for each vertex v in P **do**
Bron-Kerbosch (P ∩ {v}, R ∪ {V} , X \{v})
P = P \ {v}
X = X ∪ {v}
end for
end if

The algorithm keeps track of three distinct sets of vertices: P, R, and X. Here, P is the set of the possible vertices that may be selected, R is the set of vertices that construct the maximal clique, and X is the set of vertices that cannot construct maximal clique. The purpose of X is to prevent including the same maximal clique several times, through the update P = P \ {v}. At the initialization step, P contains all the vertices of the graph, and R and X are empty sets. In each iteration, P and X are distinct sets, and their union consists of the vertices that form cliques when added to R. There are no more elements that can be added to R when P and X are both empty, so R is a maximal clique, and the procedure returns R. Results for this algorithm are listed in Table 2.

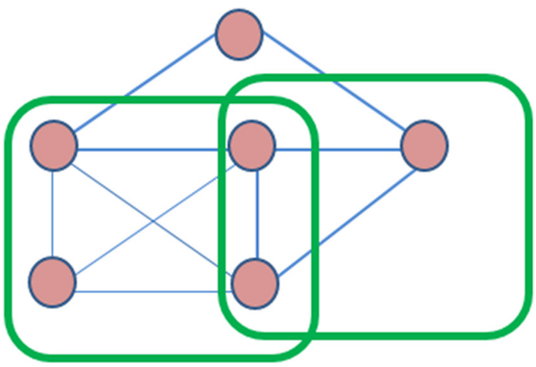

Fig. 2. Applying Bron-Kerbosch algorithm on a region of 6 vertex, we get 4 maximum cliques, 1 triangle, 1 square and 2 edges

4 Proposed system

4.1 System functionality

At the beginning of every new academic year, thousands of students suffering of bus transportation. The process of managing bus traffic in schools takes a lot of time. This process must take into account the limited number of existing buses and the arrival of students to and from schools at a specific time. The key feature of this application is dividing the geographic regions into sub-regions and allocating buses to the appropriate sub-regions.

The following requirements specify the functions of the system and its components:

- The system shall contain the login operations.
- The system shall organize the tour that the bus should take, which should be determined based on the geographical information of the students' houses.
- The system has to choose the shortest routes that the bus can take in a specific region.
- The system shall list the students in each tour.
- The system shall enable the user to view the route and some important information about the students that have been transported.
- The system shall enable the user to delete, add, and update students from tours.

4.2 Analysis and design

In this section, we present the main analysis techniques to characterize the proposed system: (a) the sequence diagram (b) the use case diagram, and (c) the class diagram. The sequence diagram is used to model the flow of operations within the system and to describe how and in what order- the main objects of the system interact together to accomplish a process. Typically, the system serves three users: the school's administrators, the buses' drivers and the student's parents. Users should have accounts by signing up into the application. The administrator enters the student's information and their house locations. Students' houses are located directly on the Google map in the application. In order to ease the burden, parents can enter these information. The system then calculates the cliques, determines the tour for each clique, and allocates buses for each. After that, the bus driver's portal displays tours and some important information about the students on those tours, such as the parents' phone number. The sequence diagram for the school's administrator and the driver is shown in Figure 3.

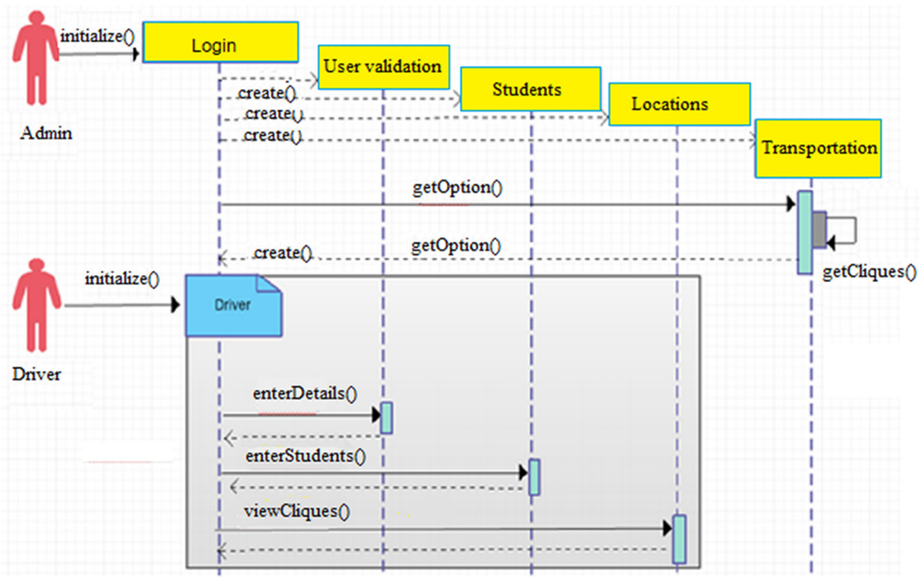

Fig. 3. The sequence diagram for the school's administrator and the driver

The use case diagram is used to describe a set of operations that objects (actors) should do in the system. The use case modeling is a powerful approach for eliciting needs. It depicts the software system's requirements in a graphical format. Figure 4 describes the main tasks that the administrator can perform in the school mobile application. Besides this, we used the class diagram to present the main entities in the system, their properties, their behaviors (methods), and the interactions between them. The class diagram of the proposed system is shown in Figure 5.

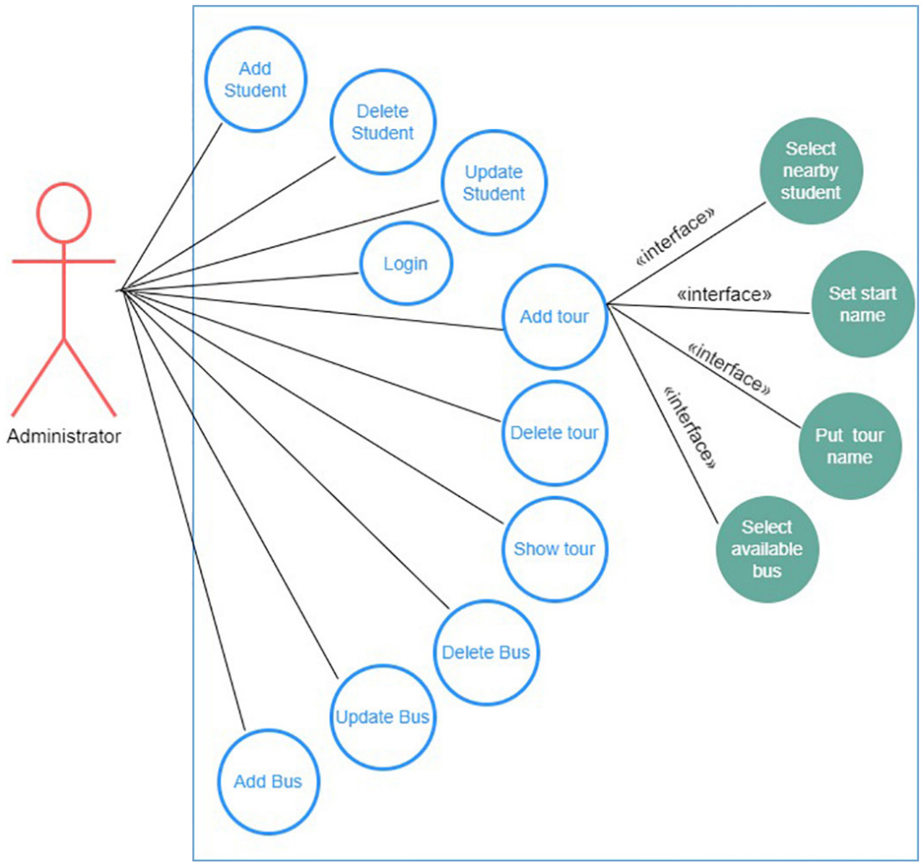

Fig. 4. The use case diagram of the administrator's operations

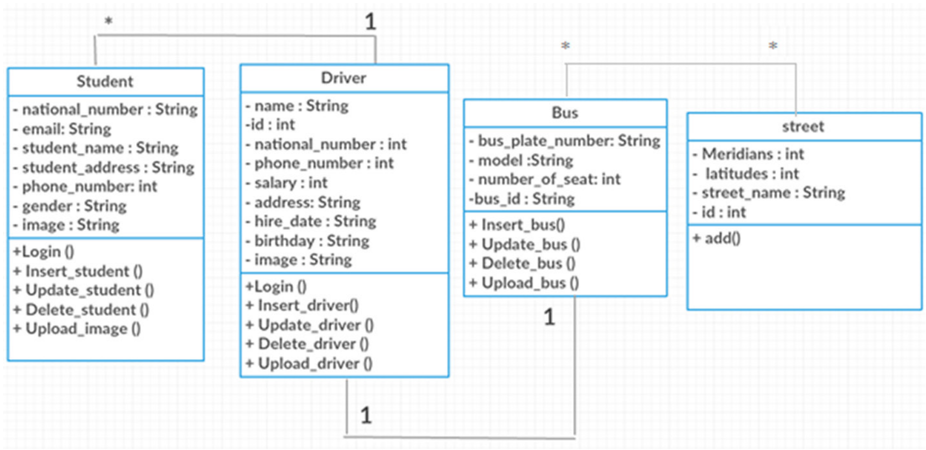

Fig. 5. The class diagram of the system

5 Experimental results

In order to prove the effectiveness and performance of the Bron-Kerbosch algorithm, which is based on the maximal clique mining algorithm, we developed an application for organizing school transportation. The experiments are performed on the following platforms: Intel(R) Core(TM) i7-4790 CPU @ 3.60 GHz, 64-bit operating system of Windows 10. The developed application works on mobile devices that use the Android operating system. We tested it at a school affiliated with Hashemite University. We chose 84 houses located around the school and saved their geographic information. Table 1 shows the latitude and longitude coordinates of the selected quarter of houses and the corresponding values on the Cartesian level axes. Note that the home address appears in latitude and longitude values in the Google API system. However, these values are not coordinates in a Cartesian system. Therefore, the latitude and longitude values are converted to the corresponding values at the Cartesian level by referring to the equations mentioned earlier in section 3.

Table 1. Longitude & latitude values of houses around a school near Hashemite University

House Number	Longitude & Latitude Pair	Cartesian Coordinates (x, y)
45	31.970969, 35.952460	4384647.523, 3180077.764
37	31.959529, 35.963077	4384601.810, 3181284.632
40	31.958737, 35.964182	4384578.085, 3181396.495
43	31.959401, 35.964825	4384510.834, 3181422.81
23	31.991409, 35.967448	4382843.793, 3180519.494
19	31.986013, 35.973521	4382763.209, 3181170.249
46	31.959929, 35.966188	4384410.064, 3181508.908
31	31.967657, 35.951473	4384859.733, 3180116.408
16	31.965991, 35.951956	4384912.108, 3180210.801
30	31.990899, 35.973639	4382524.351, 3181010.658
22	31.988825, 35.973038	4382656.332, 3181036.264
49	31.958846, 35.966134	4384464.517, 3181542.113
42	31.972744, 35.952514	4384560.149, 3180020.701
48	31.970241, 35.951624	4384728.528, 3180038.885
17	31.988752, 35.969251	4382870.046, 3180749.102
26	31.990836, 35.971086	4382669.083, 3180817.552
51	31.968612, 35.953083	4384724.979, 3180206.698

Table 2. Cliques around the school

The Allocated Bus Number	The Maximal Clique
7	{17, 23, 26, 30, 22, 19}
9	{42, 45, 48, 51, 31}
2	{37, 40, 43, 46, 49, 16}

Table 2 shows how the selected houses were grouped into cliques using the Bron-Kerbosch algorithm based on their geographical information and the reservation of a bus for each tour (clique). The aim of using the Bron-Kerbosch algorithm is to group students' houses that are close to each other and mark them as a tour so that this tour is reserved for a particular bus.

The proposed mobile application was developed using the Android Studio platform. Android Studio is Google's official integrated development environment (IDE) that offers the fastest tools for implementing apps on various types of Android devices. The platform enables us to integrate the Google Map API into the mobile application, display any location on the map, as well as different routes, and customize the map according to the desired functions. Those functions comprise a secure and reliable authorization application programming interface (API) for Google accounts.

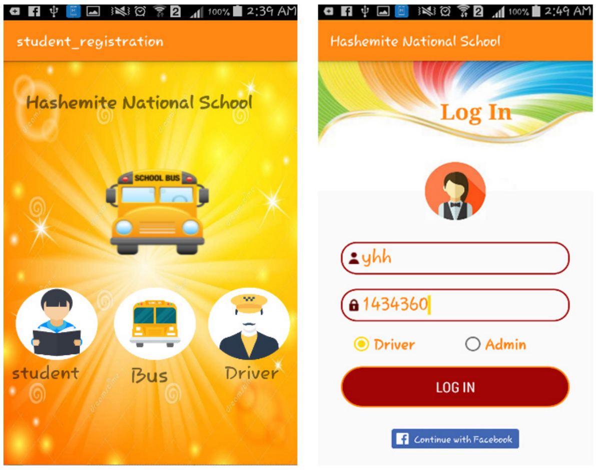

Fig. 6. The home page and the sign-in page

Figure 6 shows the home page and the sign-in page for the users of the application. Mainly, Figure 7 shows some screenshots of the key functions of the application, which are arranged in a tabular form.

Here are some of the main functions included in the application:

- Log in: Provides a consistent and safe way to access accounts and log in to the application.
- Google maps: Shows the location of the school and the student's house on the map.
- Bus route (clique): This function is the core service of the mobile application. It determines the route that the bus should take, which was determined by the system based on the geographical information of the students' houses.
- Bus portal: The school principle can edit important information about buses, such as their number, license, or capacity in terms of the number of students.
- Student portal: After the user registers in the application and is authorized to login, he can update his personal information, view some of the driver's information, and the bus number. The user might be a student or a parent.
- Driver portal: The driver can view the route he has to take and view some important information about the students he transports, such as their phone numbers.

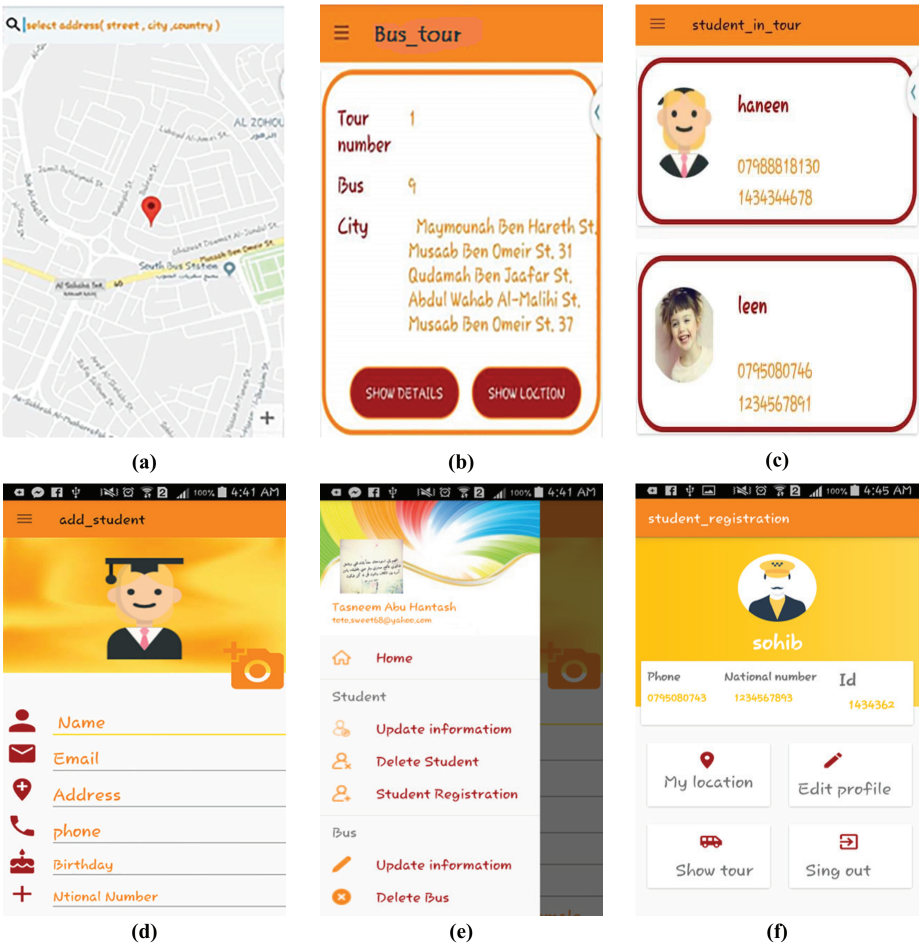

Fig. 7. Here some screenshots of the key functions of the application such as (a) determining the student's house, (b) showing the bus tour, (c) listing the students in a specific tour, (d) adding a new student, (e) updating a student information, and (f) the student registration

To achieve the objectives of the proposed method in terms of usability, we asked real users to download and use the application. Also, to measure the satisfaction and conviction of the users of the application, we conducted a survey and distributed it to the users of the application (the administrators, the drivers, and the parents of the students). The sample size was 88, with 10 school principals, 20 drivers, and the rest were students' parents. The 10-question questionnaires were distributed to the selected sample. The contents of the questionnaires focused on (1) the importance of using the proposed application, (2) the main functions of the proposed system, and (3) whether it achieved the desired objectives in facilitating the transport of students. The answer options varied between "yes" and "no" and "strongly agree," "neither agree nor disagree," "disagree, "strongly disagree."

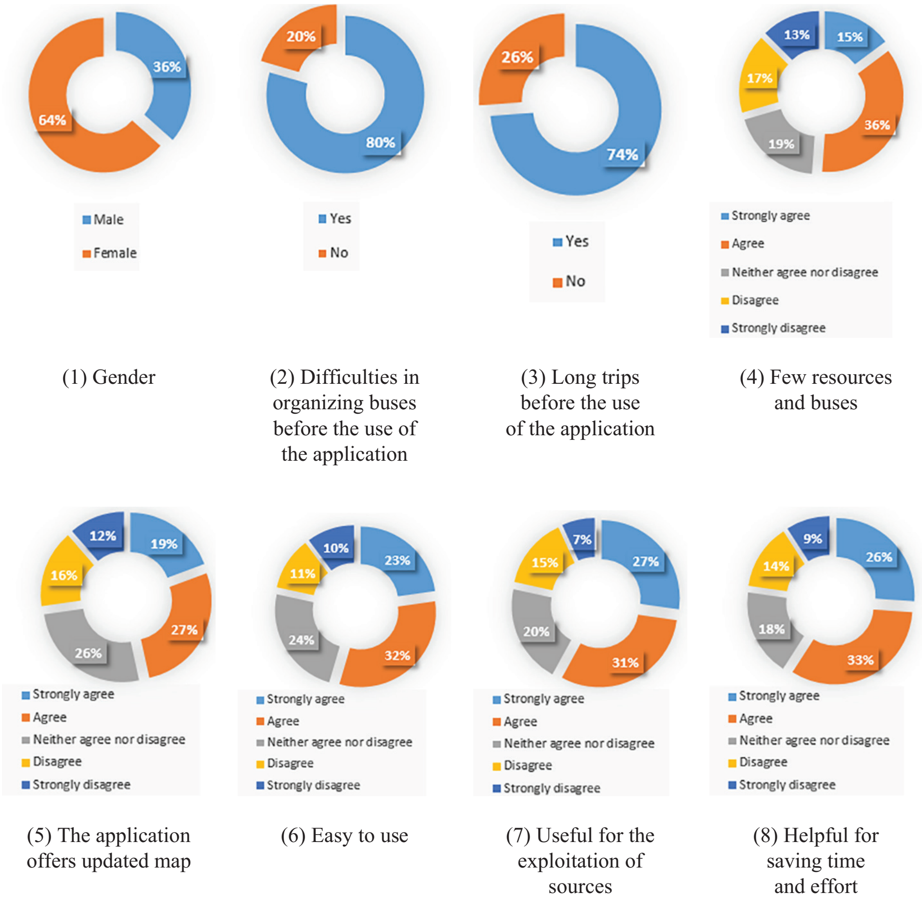

Fig. 8. The questionnaire analysis of the collected data

Figure 8 shows a questionnaire analysis for some questions, such as (1) gender, and (2) are the problems associated with distributing and organizing bus transport at the beginning of each semester becoming less? , and(3) students usually spend a long time getting home from school; (4) the school owns a few buses, which affects the delivery of students to and from school on time; (5) the application provides a modern map of the roads that lead to the school; (6) the application is easy to use and can be used by users of different ages and cultures; (7) the application helps schools make efficient use of their transportation resources; and (8) the application helps in the process of organizing school transportation, which consumes time and effort.

Table 3. Standardized leadings: composite reliability and average variance

Factor	Question	Composite Reliability	Average Variance
System Usefulness	(7), (8)	0.803	0.83
System Reliability	(2), (4)	0.819	0.79
Ease of Use	(5), (6)	0.871	0.81
User Satisfaction	(3)	0.794	0.78

We used a confirmatory factor analysis (CFA) to assess the research framework's reliability and validity. The standardized loadings, composite reliability, and average variance were extracted for important factors, and the results are listed in Table 3. In general, and to ensure appropriate convergence, values of composite reliability of between 0.60 and 0.70 are acceptable in exploratory research, but values higher than 0.70 are required in more advanced stages. As a rule of thumb, an average variance extracted of at least 0.50 is highly recommended. The results from the collected data were encouraging, and all standardized loadings were larger than 0.5. The average variance scales are considered to have good convergent validity. In addition, all composite reliability values are larger than 0.7, showing good reliability. The findings showed that the school stakeholders, drivers, and students' families would like to use the application in their daily lives, as this is of great importance in utilizing time and saving fuel from the school's point of view, and, on the other hand, is a direct way to reassure the parents of their children.

6 Conclusion and future work

Overall, this paper presents a mobile application for school transportation. The proposed application is characterized by the ability to divide the geographic regions into sub-regions where each sub-region covers nearby houses and allocate buses for the resultant sub-regions. The proposed method relies on a graph theoretic technique in the field of machine learning called the Bron-Kerbosch algorithm. The application allows drivers to select the appropriate routes to pick up students, provides a modern map of the roads, makes efficient use of transportation resources, and gives advice to drivers and staff in the process of organizing school transportation, which consumes time and effort. A survey was conducted to assess the proposed application. The reviews' conclusions were impressive and significant.

This work can be extended in the near future through the following topics:

- The Google Maps-based automatic notification system: notifying parents that the bus is on its way reduces their concern for their children.
- The school administration and parents have more assurance when they follow the movements of the school bus, especially when there is a traffic crisis or lane change.
- Develop an interface that can work on mobile devices with iOS operating systems.

All of these efforts will be directed towards giving the application the most prominent advantage in Jordan. This work could easily be extended to application fields of clustering algorithms such as computational biology.

7 Acknowledgment

The authors would like to thank the Hashemite University for its contribution to the development of the mobile application for this research.

8 References

[1] Newton, R.M. and Thomas, W.H. "Bus routing in a multi-school system." Computers & Operations Research 1.2 (1974): 213–222. https://doi.org/10.1016/0305-0548(74)90047-1
[2] Bektaş, Tolga, and Seda Elmastaş. "Solving school bus routing problems through integer programming." Journal of the Operational Research Society 58.12 (2007): 1599–1604. https://doi.org/10.1057/palgrave.jors.2602305
[3] Spada, Michela, Michel Bierlaire, and Th M. Liebling. "Decision-aiding methodology for the school bus routing and scheduling problem." Transportation Science 39.4 (2005): 477–490. https://doi.org/10.1287/trsc.1040.0096
[4] Kara, Imdat, and Tolga Bektas. "Integer linear programming formulations of multiple salesman problems and its variations." European Journal of Operational Research 174.3 (2006): 1449–1458. https://doi.org/10.1016/j.ejor.2005.03.008
[5] Corberán, Angel, et al. "Heuristic solutions to the problem of routing school buses with multiple objectives." Journal of the Operational Research Society 53.4 (2002): 427–435. https://doi.org/10.1057/palgrave.jors.2601324
[6] Parvasi, Seyed Parsa, Mehdi Mahmoodjanloo, and Mostafa Setak. "A bi-level school bus routing problem with bus stops selection and possibility of demand outsourcing." Applied Soft Computing 61 (2017): 222–238. https://doi.org/10.1016/j.asoc.2017.08.018
[7] Miranda, Douglas Moura, et al. "A multi-loading school bus routing problem." Expert Systems with Applications 101 (2018): 228–242. https://doi.org/10.1016/j.eswa.2018.02.014
[8] Braca, Jeffrey, et al. "A computerized approach to the New York Cityschool bus routing problem." IIE Transactions 29.8 (1997): 693–702. https://doi.org/10.1080/07408179708966379
[9] Martínez, Luis M., and José M. Viegas. "Design and deployment of an innovative school bus service in Lisbon." Procedia-Social and Behavioral Sciences 20 (2011): 120–130. https://doi.org/10.1016/j.sbspro.2011.08.017
[10] López Santana, Eduyn Ramiro, and José de Jesús Romero Carvajal. "A hybrid column generation and clustering approach to the school bus routing problem with time windows." Ingeniería 20.1 (2015): 101–117. https://doi.org/10.14483/udistrital.jour.reving.2015.1.a07
[11] Schittekat, P., Kinable, J., Sörensen, K., Sevaux, M., Spieksma, F. and Springael, J. "A metaheuristic for the school bus routing problem with bus stop selection." European Journal of Operational Research 229.2 (2013): 518–528. https://doi.org/10.1016/j.ejor.2013.02.025
[12] Park, Junhyuk, and Byung-In Kim. "The school bus routing problem: A review." European Journal of Operational Research 202.2 (2010): 311–319. https://doi.org/10.1016/j.ejor.2009.05.017
[13] Pacheco, Joaquín, et al. "Bi-objective bus routing: an application to school buses in rural areas." Transportation Science 47.3 (2013): 397–411. https://doi.org/10.1287/trsc.1120.0437
[14] Díaz-Parra, Ocotlán, et al. "A vertical transfer algorithm for the school bus routing problem." 2012 Fourth World Congress on Nature and Biologically Inspired Computing (NaBIC). IEEE, 2012. https://doi.org/10.1109/NaBIC.2012.6402241
[15] Schittekat, Patrick, et al. "A metaheuristic for the school bus routing problem with bus stop selection." European Journal of Operational Research 229.2 (2013): 518–528. https://doi.org/10.1016/j.ejor.2013.02.025

[16] Park, Junhyuk, Hyunchul Tae, and Byung-In Kim. "A post-improvement procedure for the mixed load school bus routing problem." European Journal of Operational Research 217.1 (2012): 204–213. https://doi.org/10.1016/j.ejor.2011.08.022
[17] Euchi, Jalel, and Rafaa Mraihi. "The urban bus routing problem in the Tunisian case by the hybrid artificial ant colony algorithm." Swarm and Evolutionary Computation 2 (2012): 15–24. https://doi.org/10.1016/j.swevo.2011.10.002
[18] Riera-Ledesma, Jorge, and Juan José Salazar-González. "A column generation approach for a school bus routing problem with resource constraints." Computers & Operations Research 40.2 (2013): 566–583. https://doi.org/10.1016/j.cor.2012.08.011
[19] Schlossberg, Marc, et al. "School trips: effects of urban form and distance on travel mode." Journal of the American Planning Association 72.3 (2006): 337–346. https://doi.org/10.1080/01944360608976755
[20] De Witte, Astrid, et al. "The impact of "free" public transport: The case of Brussels." Transportation Research Part A: Policy and Practice 40.8 (2006): 671–689. https://doi.org/10.1016/j.tra.2005.12.008
[21] Florian, Michael, and Marc Gaudry. "A conceptual framework for the supply side in transportation systems." Transportation Research Part B: Methodological 14.1–2 (1980): 1–8. https://doi.org/10.1016/0191-2615(80)90029-6
[22] Ibeas, Angel, Jose L. Moura, and Luigi Dell'Olio. "Planning school transport: Design of routes with flexible school opening times." Transportation Planning and Technology 32.6 (2009): 527–544. https://doi.org/10.1080/03081060903374734
[23] Rodrigue, Jean-Paul. The geography of transport systems. Routledge, 2020. https://doi.org/10.4324/9780429346323
[24] Ortega, Jairo, et al. "Optimization model for school transportation based on supply-demand analyses." Journal of Software Engineering and Applications 12.06 (2019): 215. https://doi.org/10.4236/jsea.2019.126013
[25] Duong, Chinh, et al. "Smartphone-based sensing for intelligent inland waterway transportation." International Journal of Interactive Mobile Technologies (iJIM) 14.18 (2020): 195–203. https://doi.org/10.3991/ijim.v14i18.16449
[26] Kassim, Murizah, and Ahmad Syafiq Aiman A. Bakar. "The design of augmented reality using unity 3D image marker detection for smart bus transportation." International Journal of Interactive Mobile Technologies 15.17 (2021): 33–48. https://doi.org/10.3991/ijim.v15i17.22071
[27] Alqudah, Yazan, et al. "Machine learning to classify driving events using mobile phone sensors data." International Journal of Interactive Mobile Technologies (iJIM) 15.2 (2021): 124–136. https://doi.org/10.3991/ijim.v15i02.18303
[28] Hussein, Fairouz, Sari Awwad, and Massimo Piccardi. "Joint action recognition and summarization by sub-modular inference." 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2016. https://doi.org/10.1109/ICASSP.2016.7472167
[29] Bomze, Immanuel M., et al. "The maximum clique problem." Handbook of combinatorial optimization. Springer, Boston, MA, 1999. 1–74. https://doi.org/10.1007/978-1-4757-3023-4_1
[30] Bron, C and J. Kerbosch: finding All clique of an undirected graph. ACM, 1973. 16(9):p.575–577. https://doi.org/10.1145/362342.362367
[31] Stix, Volker. "Finding all maximal cliques in dynamic graphs." Computational Optimization and applications 27.2 (2004): 173–186. https://doi.org/10.1023/B:COAP.0000008651.28952.b6

9 Authors

Fairouz Hussein is currently an assistant at the Faculty of Prince al Hussein bin Abdullah II for Information Technology, the Hashemite University, Zarqa, Jordan. Her research interests include: machine learning, computer vision, image processing, greedy algorithms and multimedia.

Subhieh M. El-Salhi is currently an associate professor and the head of the department of Computer Information System at the Faculty of Prince al Hussein bin Abdullah II for Information Technology, the Hashemite University, Zarqa, Jordan. Her research interests include 3D surface representation, machine learning and data mining.

Rajaa ALazazma is a senior full stack web developer with a Bachelor degree in Computer science and Application from Faculty of Prince Al-Hussein bin Abdullah II of Information Technology, the Hashemite University, Zarqa, Jordan.

Tasneem Abu-Hantash has a Bachelor degree in software engineering from Faculty of Prince Al-Hussein bin Abdullah II of Information Technology, the Hashemite University, Zarqa, Jordan.

Haneen Abu-Hantash has a Bachelor degree in software engineering from Faculty of Prince Al-Hussein bin Abdullah II of Information Technology, the Hashemite University, Zarqa, Jordan.

Heba Thaher is a web and software developer and a computer educator for all educational levels with a Bachelor degree in software engineering from Faculty of Prince Al-Hussein bin Abdullah II of Information Technology, the Hashemite University, Zarqa, Jordan.

Article submitted 2022-03-03. Resubmitted 2022-04-27. Final acceptance 2022-05-18. Final version published as submitted by the authors.