CHAPTER 1

INTRODUCTION

1.1 Rationale

In the era of globalization, the tourism sector plays a vital role in a country's economy [1]. The increased number of domestic and foreign tourists has marked the development of the tourism sector in Indonesia. The variety of tourist destinations in the area often makes it difficult for tourists to determine which attractions to visit, considering their limited time. In addition, tourists also tend to visit new places and spend time traveling for more than one day. Their limited knowledge of the areas visited makes it difficult for them to determine efficient tourist routes [2–4]. For this reason, it is necessary to build a system that can automatically generate optimal tourist route recommendations for several days (N-days) of visit.

Several previous studies have examined tourist route recommendations for N-days tourist visit using optimization algorithms such as Artificial Bee Colony (ABC) [5], Firefly Algorithm (FA) [6], and Simulated Annealing (SA) [7] to solve the Traveling Salesman Problem (TSP). TSP is an optimization problem that aims to find the shortest distance by visiting a set of points, where each point must be visited exactly once, and the route must end at the starting point [8]. However, these studies have a significant drawback: the tourist route recommendations obtained in N-days are not optimal. Due to the fact that the TSP analogy generates a single path, in order to determine the tourist route for N-days, the path must be cut according to the time window constraints and the number of days specified. The process of cutting paths in the TSP is illustrated in Fig. 1.1.a. TSP generates N-day itineraries by dividing paths. In other words, the route per day is obtained by iterating, where every possible point that could be included in the route on the day N ($N = 1, 2, 3, \dots, x$) is checked individually to ensure that the time window constraints are met without considering the shortest distance between the points. If a point exceeds the time constraint, it and the following points will be checked for inclusion in the route for the next day, while the previous point will be included in the route for the N-th day. The resulting routes fail to emphasize daily optimization by cutting the path that way.

To overcome the routing limitations in the previous studies [5–7], we propose the Vehicle Routing Problem (VRP) analogy to produce more optimal tourist routes. VRP can be understood as a conceptual expansion and generalization of the NP-Hard TSP [9]. The initial concept of VRP was first introduced in 1959 by Dantzig and Ramser [10]. VRP is an optimization problem that aims to determine the most efficient route for several vehicles by considering additional constraints, including capacity, time, and others [11]. In the VRP with capacity constraints, known as a Capacitated Vehicle Routing Problem (CVRP), each

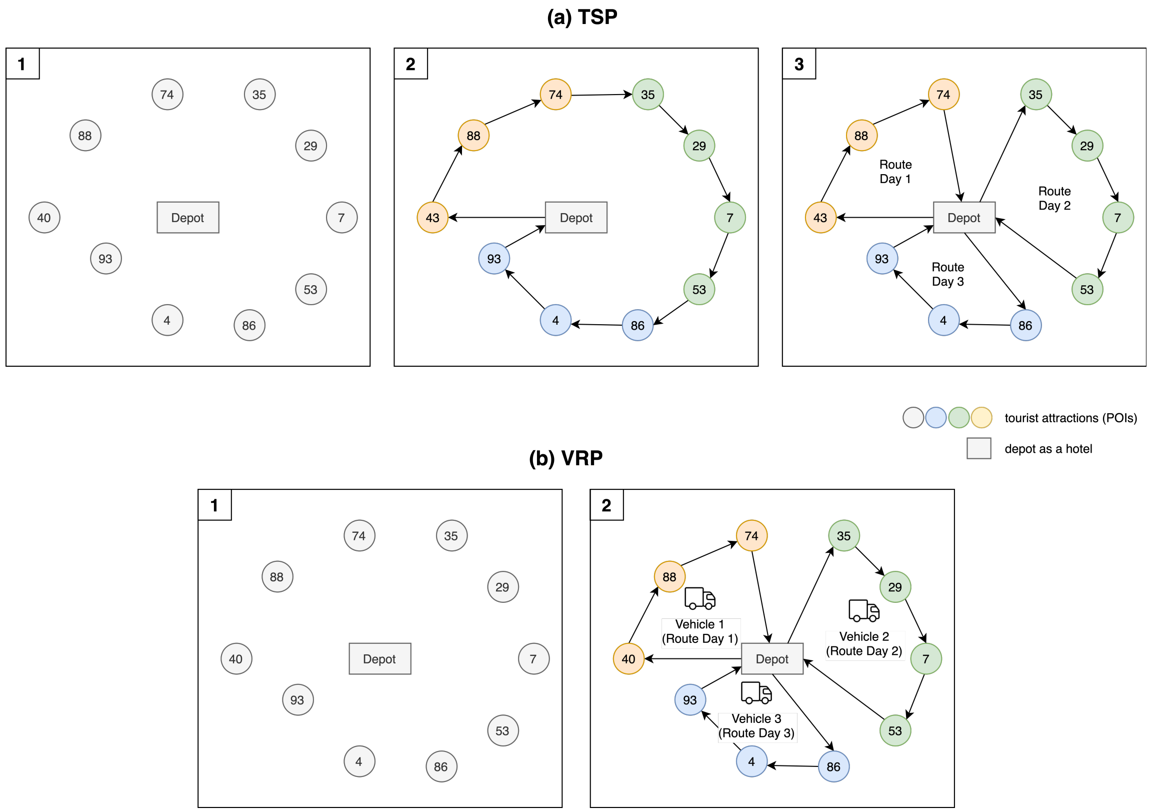

Figure 1.1: The Differences between TSP and VRP.

vehicle is assigned a maximum load capacity that must not be surpassed during travel [12]. On the other hand, the Vehicle Routing Problem with Time Windows (VRPTW), a variant of VRP with time constraints, introduces the time aspect where each vehicle has specific time windows. The vehicle must arrive or leave the location during the visit according to the predetermined time windows [13]. The N-days tourist route problem involves time constraints. The vehicle operating time (time windows of the vehicle), the opening and closing hours of the Points of Interest (POIs) (time windows of the POIs), and the time required to visit each POI (POI visiting duration, which reflects the time spent by the user at each POI) are assumed to be the time constraints. Therefore, this study analogizes the N-days route problem as a VRPTW. The process of finding the N-days route with the VRP analogy is shown in Fig. 1.1.b. In this analogy, the VRP assigns multiple vehicles simultaneously, each with certain specified constraints.

Metaheuristic algorithms are increasingly recognized as superior solutions for complex optimization problems like VRP, especially when conventional methods often fail to manage the vast and unstructured search space [14–16]. One of the most widely used algorithms is the Artificial Bee Colony (ABC), which draws inspiration from the collective intelligence of worker bees in efficiently finding and exploiting food sources [17]. Despite its strengths in simplicity and exploration [18–20], it often faces challenges in fully exploiting the best solutions and tends to get trapped in local optima [21–23], particularly when dealing with more complex VRP optimization challenges. To address these limitations, a hybrid approach known as the Improved Artificial Bee Colony Algorithm based on reverse learning Harris Hawks Optimization (HABC) was developed [24]. This hybridization integrates Harris Hawks Optimization (HHO) into the onlooker bee phase to boost exploration capabilities by mimicking the cooperative hunting strategy of hawks [25], while the Cauchy Reverse Learning strategy is applied in the scout bee phase to help the algorithm escape local optima by introducing a more diverse search pattern [26–28]. These improvements not only boost the algorithm's ability to explore the search space but also accelerate its convergence, allowing HABC to find optimal solutions more quickly and effectively. Although still new, HABC has been tested in one study on benchmark problems and demonstrated superior performance over other algorithms including traditional ABC [24], making it the most promising choice for tackling the complex VRP challenges in our research.

Our research proposes a model to solve the N-days tourist route problem by analogizing it to a VRPTW (VRP analogy). For route distribution per day, we design a Greedy algorithm to optimize route distribution. The route search is performed using HABC. To generate daily tourist routes, the system considers several criteria (multi-criteria) based on user needs: popularity (rating), time, cost, and the number of tourist attractions (POIs) included. Then, the generated tourist route recommendations are evaluated based on the fitness value. The fitness value is calculated based on Multi-Attribute Utility Theory (MAUT) as an objective function in order to estimate user interest based on the four predetermined criteria.

1.2 Theoretical Framework

The theoretical framework in this research is built based on several main theories and concepts that underlie the development of an optimal tourist route recommendation system for N-days of visit. These theories and concepts include recommender systems, the Vehicle Routing Problem (VRP), optimization algorithm, and Multi-Attribute Utility Theory (MAUT).

1. Recommender Systems Theory

   Recommender systems are crucial for providing personalized suggestions based on user preferences and historical data. In the tourism context, recommender systems can suggest attractions, accommodations, and routes tailored to individual users. This research integrates recommender systems with VRP models and optimization algorithms to generate personalized N-days tourist routes. By leveraging user preferences (multi-criteria evaluation), the system can offer tailored recommendations that enhance user satisfaction and travel experience. The integration of recommender systems ensures that the generated routes are not only optimized for efficiency but also personalized to meet the specific interests and preferences of each user.

2. Vehicle Routing Problem (VRP) Theory

   VRP serves as the foundation for understanding how to optimize routes for multiple vehicles under various constraints such as capacity, time windows, and others. Unlike simpler routing problems, VRP allows for the incorporation of multiple days and complex constraints, making it particularly suitable for tourism scenarios where tourists need to visit multiple points of interest (POIs) over several days. In the context of N-days tourist route planning, VRP helps in structuring the problem by considering time constraints for each day, ensuring that the routes are feasible and efficient for daily travel.

3. Optimization Algorithm Theory

   Optimization algorithms are critical in efficiently solving complex routing problems like VRP. One such algorithm is the ABC algorithm, which mimics the foraging behavior of bees to find optimal solutions. Despite its effectiveness, ABC can sometimes be slow to converge and may get stuck in local optima. To address these limitations, the Improved Artificial Bee Colony Algorithm based on reverse learning Harris Hawks Optimization (HABC) is used. HABC combines the strengths of the Cauchy Reverse Learning strategy and the HHO algorithm, enhancing the search process. This hybrid approach increases the accuracy and speed of finding optimal solutions, making it well-suited for planning multi-day tourist routes. By leveraging HABC, the system can more effectively navigate the search space, avoiding local optima and improving overall route optimization.

4. Multi-Attribute Utility Theory (MAUT)

   MAUT is essential for evaluating and ranking alternative tourist routes based on multiple criteria. MAUT allows the proposed system to assess routes considering various factors such as popularity, time, cost, and the number of POIs included. By applying MAUT, the system can estimate user preferences and optimize routes to ensure they align with user needs, resulting in more satisfying and personalized travel experiences. The use of MAUT ensures that the recommended routes provide a balanced approach, taking into account all relevant criteria to maximize user satisfaction.

1.3 Statement of the Problem

In previous research, the N-days route problem is often analogized to the TSP [5–7]. However, the routing concept in TSP that only produces a single path, affects the results of the N-days route to be non-optimal. With this single-path concept, the N-days route is generated by visiting all POIs, and then cutting the path according to the specified time constraint and number of visiting days. This process is executed without reconsidering the distance between each POI. Therefore, this TSP analogy is not suitable to be analogous to the N-days tourist route problem because it does not emphasize daily optimization. To overcome these problems, this research proposes VRPTW (VRP analogy) to solve the N-days route problem.

In addition to solving the N-days tourist route problem by analogizing it as a VRP analogy to address the limitations of the TSP from previous studies [5–7], this research also seeks to overcome the weaknesses of the traditional ABC algorithm. The traditional ABC algorithm is a metaheuristic algorithm that can solve various kinds of problems, one of which is the optimization problem of route determination. The selection of the traditional ABC algorithm is based on three main considerations, namely: 1) ABC is a simple algorithm; 2) it is more efficient than other Swarm Intelligence algorithms; and 3) it is proven to be effective in solving VRP in single-depot cases [18]. Besides these advantages, the traditional ABC algorithm also has weaknesses, namely the slow convergence of solutions and solutions can be trapped into local optimum [21–23]. This happens because the exploration may not be thorough enough and relies too much on random changes, which can limit diversity and lead to premature convergence. Therefore, we use the HABC algorithm to optimize routes in VRP, specifically applied to the tourist route problem in this study. The HABC algorithm, which integrates the advantages of the Cauchy Reverse Learning strategy and the HHO algorithm into the traditional ABC algorithm, helps to enhance search accuracy, escape from local optima, and achieve rapid convergence during the search process [25–28]. This algorithm has been proven to work well in solving benchmark problems [24]. The main question that arises is how does the HABC algorithm performs in generating N-days tourist route recommendations compared to the VRP and TSP analogies, the traditional ABC algorithm, and other algorithms such as Ant Colony System (ACS), Simulated Annealing (SA), and Cuckoo Search Optimization (CSO)?

1.4 Objective and Hypotheses

The objectives of this study are as follows:

1. To optimize tourist routes per day based on tourist needs (multi-criteria) by analogizing it as a Vehicle Routing Problem (VRP) and comparing it to the Traveling Salesman Problem (TSP).

2. Implementing Improved Artificial Bee Colony Algorithm based on reverse learning Harris Hawks Optimization (HABC) in the area of a recommender system, especially in the problem of generating N-days tourist route.

3. To prove that the HABC algorithm outperforms the traditional ABC and other algorithms (ACS, SA, and CSO) in generating optimal N-days tourist route.

The hypothesis in this study is that the N-days tourist route problem, analogous to VRP and optimized using HABC, can produce more optimal routes. In the context of VRP, which extends TSP, a number of vehicles can be assigned according to the number of days and specified time constraints to form effective routes [9, 11, 13]. In addition, optimizing the solution search process is crucial for finding the best routes. This is where the HABC algorithm plays a crucial role by improving the quality of the solution search process in VRP. HABC provides a more effective mechanism for exploring different possible routes and refining them to find better solutions, thereby producing optimal routes [24]. Therefore, the use of the HABC algorithm not only addresses the weaknesses of the traditional ABC algorithm but also optimizes the search process in VRP more effectively than in TSP. Thus, HABC can provide better solutions compared to the traditional ABC and other algorithms such as ACS, SA, and CSO.

1.5 Assumption

Several assumptions were made in this study. First, we assume that users already know the attractions (POIs) they want to visit, the hotel they want to stay in, how long they want to spend on vacation, and the Degree of Interest (DOI) values for the three criteria (rating, time, and cost) that affect their interest. Therefore, the system focuses on recommending N-days tourist route based on the information provided by users. Second, this study assumes that the mode of transportation used is only a car, with each vehicle starting and ending at one hotel. Third, the travel duration used in this study, between POIs and between POIs and hotels, is based on the travel duration obtained from the Google Maps API. Fourth, the cost of travel is the entry fee for each POI. Fifth, each POI is visited exactly once. Finally, this study assumes that the tourist visit time (time window) per day is from 08:00 AM to 09:00 PM.

1.6 Scope and Delimitation

This research assumes that users already have knowledge about their itinerary, such as their selected POIs, hotels, vacation duration, and DOI values for rating, time, and cost. Based on these assumptions, the independent variables in this study are the user's desired POIs and the DOI values for rating, time, and cost, while the DOI of the number of POIs included in the itinerary is considered as a control variable. Additionally, the generated N-days tourist route recommendation and fitness values are treated as dependent variables.

The dataset used in this study consists of data on POIs and hotels in the Yogyakarta area. It does not consider time-dependent types of tourism, such as culinary tourism that requires specific lunch and dinner schedules.

1.7 Significance of the Study

The main contribution of this research is to propose a model for finding the most optimal tourist route for N-days by solving VRP using the HABC algorithm. The HABC algorithm has been proven to solve benchmark problems and can overcome the shortcomings of the traditional ABC algorithm, particularly in terms of convergence speed and avoiding local optima [24]. In addition, we combine the concept of MAUT to consider user preferences. Furthermore, a Greedy strategy is used to decode the N-days tourist route from a vector representation in the VRP, efficiently assigning POIs to daily routes while respecting time window constraints. Thus, the model that we build can generate the most optimal N-days tourist route based on user preferences, ensuring a balance between efficient travel and personalized experience.