Generating Efficient Rebalancing Routes for Bikeshare Programs Using a Genetic Algorithm

James R. Kroes  
*Boise State University*

Andrew S. Manikas  
*University of Louisville*

Thomas F. Gattiker  
*Boise State University*

**Publication Information**


This is an author-produced, peer-reviewed version of this article. © 2020, Elsevier. Licensed under the Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 license. The final, definitive version of this document can be found online at *Journal of Cleaner Production*, doi: 10.1016/j.jclepro.2019.118880
Generating Efficient Rebalancing Routes for Bikeshare Programs Using a Genetic Algorithm

James R. Kroes*
Boise State University
College of Business and Economics
Boise, ID
jimkroes@boisestate.edu

Andrew S. Manikas
University of Louisville
Management Department
College of Business
Louisville, KY
andrew.manikas@louisville.edu

Thomas F. Gattiker
Boise State University
College of Business and Economics
1910 University Drive
Boise, ID
tomgattiker@boisestate.edu

* Corresponding author

1. Introduction

Growth in urban areas often leads to problems such as increased traffic congestion and poor air quality. To help alleviate these issues, shared mobility networks have been launched in hundreds of cities worldwide to provide citizens with alternatives to personal autos and to other less sustainable methods of transport (Fishman, 2016; Zhang et al., 2015). Shared mobility includes carsharing, ridesharing, scooter sharing and bikesharing (SAE, 2018). Bikeshare programs allow users to pick up bicycles (often at hub locations), utilize the bicycle for a journey, and return it to a location within the system (DeMaio, 2009). While bicycle sharing has been in existence for many years in various forms, the advent of modern telecommunications (i.e., cellular technology and the internet) have enabled these programs to proliferate.

While the growth of bikeshare programs over the past decade is impressive, operating these systems is not without challenges. One of the crucial operational issues faced by bikeshare programs involves the rebalancing (also called repositioning) of bicycles throughout the system (Raviv et al., 2013). Rebalancing is required when users sign out bicycles from popular locations and return them to less popular areas. Over time, this results in limited availability (or complete unavailability) of bicycles at popular locales. For example, one press account (Dobkin, 2013) describes a community where the commercial district is at the base of several large hills - downhill from the major residential areas. Commuters tend to use the shared bicycles to commute (downhill) to work in the morning, but they opt for motorized public transportation to go home (uphill) at the end of the day. As a result, without intervention (i.e., rebalancing operations), bicycles would not be available where customers need them the next morning. Rebalancing, which occurs in most systems, requires bikeshare employees to frequently (often daily) pick-up and redistribute bicycles throughout the system (i.e. “rebalance” the inventory). A key challenge is determining rebalancing routes that pick up and drop off the correct number of bicycles using an efficient sequence. For typical bikeshare programs, which may have over one hundred bicycles and several dozen stops, these routing exercises are mathematically complex variations of the traveling salesperson problem (Rainer-Harbach et al., 2015).

The nature of bikeshare and other shared mobility programs heightens the need for efficient rebalancing operations: First, numerous programs operate as non-profit government supported programs with severe budget constraints, which necessitates the effective use of both people and equipment (DeMaio, 2009; Shaheen et al., 2010). Second, shared mobility programs are generally viewed positively from an environmental standpoint (i.e., lower carbon emissions than driving) (Wadud, 2014); however, rebalancing operations pose a challenge to sustainability because rebalancing is routinely accomplished using fossil fuel powered vehicles. For example, a recent study, which has been heavily covered in the popular press, finds that, the majority of the time that they are used, shared bicycles generate more carbon emissions than the alternative and that rebalancing operations are the greatest source of bike share emissions (Luo et al., 2019).
Ingenious solutions to the rebalancing problem have been offered in a number of academic works. However, these solutions have limitations with respect to adaptability to real operating environments, including the use of Euclidean distances rather than actual distances, homogeneous vs. heterogeneous fleets, ease of use for practicing managers, and affordability. Consequently, the objective of this paper is to present a software solution that allows managers to quickly and inexpensively develop efficient routings for rebalancing operations. Like many other solutions described in the literature, our heuristic generates efficient, near-optimal, routes. However, the solution is built using a ubiquitously available spreadsheet program. Furthermore, the system uses actual travel distances and travel times. These are obtained from a no-cost, publicly available source. Additionally, the algorithm allows multiple heterogeneous vehicles to be used for repositioning. Finally, the solution can incorporate real-time traffic conditions and closures when developing a route.

The next section reviews the existing literature in order to demonstrate the motivation for our study. The methodology and data used to develop the rebalancing solution are presented in the third section of the paper. The results from a case study application of the tool and conclusions from this study are discussed in the fourth and fifth sections of the paper.

2. Literature Review

2.1 Financial Constraints

Many bikeshare programs are not-for-profit organizations that rely on corporate sponsorships or local government subsidies since rental fee revenue collected from patrons is frequently insufficient to cover costs (DeMaio & Gifford, 2004; DeMaio, 2009). Many private shared mobility companies are not profitable. In such environments, operational efficiency is paramount. The criticality of efficient operations is evinced through a number of bikeshare program bankruptcies, including recent high-profile cases in Montreal and Washington, D.C. (Lindsey, 2016). This suggests that bikeshare programs need to be efficient in all aspects of operations, including fleet rebalancing. It also suggests that affordability is an imperative when it comes to software and other tools used to increase efficiency. These two imperatives – efficiency achieved at a low cost – are the starting point of the motivation for the present paper.

2.2 Stakeholder Acceptance

The growth of bikeshare and other shared mobility programs has been widespread -- mainly in urban areas and other dense locales, such as university campuses (Ji et al., 2014; Tian et al., 2018). Societal acceptance and individual demand for shared mobility depend on a number of factors including topography (Wygonic et al., 2014), land use patterns (El-Assi et al., 2017), characteristics of proximal destinations (Faghii-Imani & Eluru, 2015), availability of substitute modes (Wang et al., 2018) weather (Gebhart & Noland, 2014), and user attitudes (e.g., Chen & Lu, 2016). Important drivers of user attitudes are perceptions that programs reduce costs (O’Brien et al. 2014), reduce congestion (Kaviti et al., 2018; Kahn, 2012), and reduce carbon emissions (Wadud, 2014; Plazier et al., 2017).

However, researchers are beginning to call into doubt the emissions-related benefits of mobility sharing. Moreover, these questions about the net carbon greenhouse gas (GHG) impact of shared mobility are migrating into the public conversation as evidenced by popular news coverage (e.g., Magill, 2014. Rosalsky, 2019, Temple, 2019; Thill, 2019). Importantly, whether or not the benefits actually accrue depends substantially on the efficiency of rebalancing operations. This is corroborated by two recent life-cycle assessments of shared mobility solutions. The first study (Luo et al., 2019) finds that rebalancing is the greatest source of bike sharing emissions, comprising 36% of total emissions for dock-based systems and 73% for dockless systems. The study concludes that the most effective way to reduce the GHG footprint of bike sharing (dock or dockless) is by reducing motorized vehicle travel associated with rebalancing. The second study (Hollingsworth et al., 2019) finds that the two biggest contributors to carbon greenhouse gas emissions of scooter sharing systems are materials and manufacturing (50% of total emissions) and charging and rebalancing operations (43%). The study identifies reducing driving for rebalancing as one of three effective strategies for reducing GHG emissions — i.e., the research finds that total emissions can be reduced approximately 25% by more efficient redistribution.

Rebalancing-related emissions can potentially be reduced using a number of approaches including optimizing the locations of docking stations, providing users with incentives to return bikes to locations where they are needed, moving bikes or scooters using environmentally friendly (e.g., non-motorized) vehicles and optimizing the vehicle routings used for rebalancing (Luo et al., 2019). The need for efficient (rebalancing) vehicle routing and for
rebalancing systems that can incorporate a mixed fleet of rebalancing vehicles, including non-motorized ones, motivates the present paper. There is an impressive body of extant literature on rebalancing. In the next section, we discuss how our paper differs from and contributes to this existing research.

2.3 Rebalancing

Focusing on operations research-based approaches to improving efficiency, Laporte et al. (2015, 2018) suggest five categories of research based on their review of the shared mobility literature: (1) station location (e.g., Mete et al., 2018; Kabak et al., 2018) (2) fleet dimensioning (e.g., Fricker & Gast, 2016), (3) station inventory (e.g., Nair & Miller-Hooks, 2011; Raviv & Kolka, 2013; Wygonik et al., 2014; Saltzman & Bradford, 2016) (4) rebalancing incentives (e.g., Raimbault, 2015) and (5) vehicle repositioning, which is commonly referred to as the Bike Repositioning Problem, or BRP. (Many studies combine categories three and five - e.g., Schuijbroek et al. (2013); Erdogan et al. (2014); Lu (2016).) The focus of the present paper is on Laporte’s fifth category.

In bikeshare programs, the flow of bicycles from one station to another station is typically unequal to the flow in the opposing direction (Lu, 2016). Thus, effective rebalancing is a critical component of running an efficient bikeshare program. Bikeshare programs need to redetermine the best rebalancing route regularly (e.g., daily) due to large variability in user behaviors (Sarkar et al., 2015) - i.e., the inventories of bicycles at the various hubs within a system are not consistent enough to allow managers to utilize the same rebalancing route day to day. Laporte et al. (2018) observe that the repositioning literature can be subdivided into static and dynamic repositioning (i.e., when demand and inventory levels change “on the fly”). Dynamic repositioning (Brinkmann et al., 2019; Caggiani & Ottomaneili, 2013; Nair & Miller-Hooks, 2011; Regue & Recker, 2014; Labadi et al., 2015; Xu et al., 2019) is relatively less researched. Repositioning studies characterize the problem as one of minimizing the cost of rebalancing (i.e., vehicle miles traveled or travel time sometimes with the addition of loading and unloading time) given a target inventory level that must be achieved (Benchimol et al., 2011; Chemla et al., 2011) or a combination of cost and customer service (e.g., unmet demand, departure from target inventory levels) (e.g., Raviv et al., 2013; Szeto et al., 2016). Existing static rebalancing solutions differ on the inclusion or exclusion of a number of important characteristics, including stochastic demand (Rainer-Harback et al., 2015), the number of vehicles (i.e., one or many) (e.g., Forma et al., 2015), number of allowable visits to a station by a vehicle (i.e., one or many) (Chemla et al., 2013), whether vehicles have homogenous capacity (in terms of cost, as well as available time, distance limits, or the number of bicycles they can carry), whether or not bicycles can be temporarily dropped at a station en route to their destination station (Benchimol et al., 2011; Di Gaspero et al., 2016). Additionally, static rebalancing research has examined whether the rebalancing operation can create a temporary deficit at a station (Cruz et al., 2017), variations in the number of types of bicycles (one or many) (Li et al., 2016), demand intervals for stations (fixed or variable), and inclusion of bicycle loading and unloading times (in addition to vehicle travel times) (e.g., Raviv, Tzur & Forma, 2013).

Four characteristics of the extant rebalancing literature motivate the present paper. First, most of these prior research efforts require substantial technical expertise, and thus they exceed the day-to-day reach of the typical bikeshare program operators. Therefore, there is a need to develop rebalancing tools that managers can use on a regular basis - i.e., containing a user interface enabling day-to-day managerial use without an operations researcher serving as an intermediary.

Second, in most of the existing studies on rebalancing, the algorithms assume that geographic information is readily available, or it is simplified through the use of a Euclidean coordinate system that adopts linear paths between locations. However, actual travel distances can depart significantly from Euclidean distances (Love & Morris, 1972; Shahid et al., 2009).

Third, while the use of actual travel distances can improve the quality of solutions, even when actual driving distances are used, the current traffic conditions cause large departures from estimates that do not use real-time traffic information (Kim et al., 2005). Specifically, studies have shown that driving directions created utilizing real-time traffic data can result in driving time savings averaging 12% during periods of heavy traffic (Ng et al., 2006). These potential improvements result from routes being developed that avoid delays including construction, road closures, and traffic congestion. Without real-time traffic data, routes might be developed that direct the rebalancing vehicle into slow moving traffic congestion or the route may be infeasible due to a road closure not reflected in historical data.
Fourth, extant repositioning studies suppose a single vehicle or a fleet of homogeneous vehicles. However, many bikeshare programs are in the position of having to opportunistically create fleets by acquiring and retrofitting heterogeneous vehicles, which can become further differentiated by additions of trailers, roof racks, etc. Some systems, including the one we studied, also supplement motorized vehicles with human powered carts or bicycle trailers. (Interestingly, in a conference proceeding, Ghosh and Varakantham (2017) propose a system where patrons, as opposed to the system operators, are incentivized to redistribute bicycles via trailers that are attached to their rental bike.)

2.4 Optimizing Rebalancing Routings

Finding the optimal solution with certainty to transportation problems typically requires the full enumeration of all possible solutions – a process which often cannot be completed within a reasonable time frame (Alshamsi & Diabat, 2017). Therefore, heuristic methods that can rapidly generate near optimal solutions are often employed (Mahmoudsoltani et al., 2018; Paydar & Olfati, 2018). As computing power increased over recent decades, optimization heuristics that can be solved on personal computers, such as Genetic Algorithms (GAs) (Holland, 1975), Simulated Annealing (Kirkpatrick et al., 1983), and Tabu searches (Glover, 1986), have been developed and widely applied to a variety of problems. Of these, the GA heuristic, which was originally developed by Holland (1975) and applied to a variety of transportation problems, was deemed to be an effective tool for this application. Multiple factors drove this decision: Specifically, GA’s are relatively simple to create and very effective at rapidly determining near optimal solutions to complex optimization problems (Davis, 1991). Additionally, GAs are inherently flexible which makes them well suited to problems with a wide variety of constraints like the fleet rebalancing problem (Xiao & Konak, 2017).

3. Rebalancing Route Generation Methodology

3.1 Rebalancing Solution Goals

As detailed point by point in Section 2, the solution developed in this effort needs to meet the following criteria:

- Little or no cost
- Increases efficiency (i.e. vehicle miles, GHG emissions, travel time) of rebalancing operations
- Does not require detailed end user understanding of the underlying technical processes
- Utilizes actual travel distances and actual travel times, rather than Euclidean distances
- Utilizes real time traffic data
- Accommodates a combination of a vehicles—including non-motorized vehicles-- with differing capacities

To address the goals of this study, we determined that any solution should utilize assets currently available to potential users with little or no incremental costs. Thus, it was decided to create a tool that utilizes the MS Excel spreadsheet program as the user interface because Excel is ubiquitous in most workplaces. To solve the issue of access to actual travel time data, we employ a mechanism to import actual bicycling and driving travel times and directions data from Google Maps (free of charge) and store it in the spreadsheet tool. The only additional technical requirement is that the user needs to have an internet connection when running the tool.

3.2 User Configuration

The sequence of steps required to operate the tool is detailed in Figure 1. First, the user opens up the Excel spreadsheet file and updates the hub inventory levels (this may be done manually or, if the program has an inventory management software package, by pasting the system generated inventory report directly into one of the tool’s worksheets.) The user then selects if the rebalancing solution will use both the van and bicycle-towed trailer or just the van. Since the bicycle-towed trailer requires substantial manual effort, a feature was added to allow the user to input the maximum time that he or she would like to employ the bicycle when it is used during a rebalancing operation. During
development, the client expressed an interest in utilizing real-time traffic data so that the driving portion of the operation would avoid congested areas. Therefore, the user can also select a real-time traffic option to use that data. The trade-off between using stored historical data and real-time traffic data is that the use of real-time data requires an active internet connection and the additional time for the data to be retrieved.

The remaining steps are fully automated and do not require further user input. The program, which is coded using MS Excel’s Visual Basic for Applications (VBA) functionality, does not require any resources beyond those included with the Excel.

Fig. 1. Flowchart of the rebalancing tool’s operation.
3.3 Determination of Hubs to Be Visited

The program initially checks the inventory levels at all of the hubs to determine which hubs are outside of the allowable stocking range and thus need to be visited (for our test client, the allowable inventory levels at each hub were determined by the program’s Director based on historical usage levels.) The system verifies that the sum of bicycles that need to be added to hubs balances with the number that need to be removed from hubs. A balance between bicycles picked up and dropped off on a rebalancing route is required to prevent either a shortage of bicycles at a hub stop or a return to the office with bicycles still on the van or trailer. If the hub needs are not balanced, the system will achieve a balance by either adding an additional hub stop to pick up bicycles from (a hub with a current inventory level at the high end of the allowable range) or by adding a stop to drop off extra bicycles at a hub with a current inventory at the low end of the allowable range. In both cases, the final inventories at hubs where bicycles are removed or added remain within the user specified allowable ranges.

3.4 Travel Time Data Acquisition

Next, the program uses the Google Maps Directions Application Programming Interface (API) (Google Maps, 2018) to retrieve (via an internet connection) the bicycling and driving travel time data and turn by turn directions between the all possible pairings of the locations that need to be visited – which are stored in a worksheet within the Excel spreadsheet. If the user has selected real-time traffic information, the driving travel times under current conditions are retrieved for all of the hubs to be visited; otherwise, the standard travel times are downloaded for any location pairs not already stored in the workbook. The Google API allows users to retrieve 2,500 records per day free of charge – a limit which was not approached during our testing of our test client’s 37 hub system.

3.5 Generation of Rebalancing Routes Using a Genetic Algorithm

The ease of use and effectiveness a GA made it well suited for this application. Whether for a bicycle or a van, the objective of the GA utilized in this effort is to minimize total travel time. To find a near-optimal solution for a route, the GA assesses numerous possible rebalancing route sequences and improves them through an evolutionary process. As prescribed by Mitchell (1998) the evolutionary process consists of four steps that (i) starts with an initial population of feasible route sequences, (ii) from which the best solutions are selected for inclusion in the next generation’s population, (iii) then, routes within the current population are combined (called crossover) to create the successive population (offspring) of potential route sequences, and (iv) finally, the visit order of two stops for a small percentage of the offspring routes are randomly swapped (analogous to mutation) – which often allows the GA to overcome local optima. This process, described in detail below and depicted in Figure 2, iterates through multiple generations of solutions until a routing solution which meets predetermined criteria is developed.

As mentioned above, the tool can generate a rebalancing solution that utilizes a bicycle-towed trailer and a van or one that uses a van only (note that an option to only use a bicycle-towed trailer is not explicitly offered because a typical bicycle trailer has capacity limit that would preclude it from efficiently moving more than a handful of bicycles at one time; however, if the number of bicycles that need to be moved between hubs is less than the trailer capacity and the rebalancing time is not constrained, the tool will generate a solution that only uses a bicycle-towed trailer.) If the bicycle trailer and van are both selected as vehicles for rebalancing by the user, the tool first finds a routing solution for the bicycle trailer and then proceeds to find a route to service any remaining stops using the van.
3.5.1 Representing Route Sequences for the GA

A GA utilizes “chromosomes” that represent potential rebalancing routing sequences. In our GA, a chromosome consists of all non-origin locations that need to be visited by a route – each of which is called an allele. The origin is the first location always and not encoded in the chromosome as an allele for simplicity. Similarly, since the rebalancing
routes are round trip solutions, the origin is also always the last location and also not encoded in the chromosome. Examples of how a route is sequenced within chromosomes are shown in Figure 3, which depicts the crossover process used in the GA optimization routine.

3.5.2 Creating Initial Populations

Following Mitchell’s (1998) evolutionary steps, the GA begins the optimization process by creating an initial population of possible rebalancing routes. If a bicycle-towed trailer and van are both used for rebalancing, the initial population for the stops served by the bicycle is created first. An important difference between the application of the GA to bicycle-towed trailer versus motorized van driving routes is the method used to create the initial population of route sequences. To create the initial population of routes for rebalancing via bicycle-towed trailer, the system employs a version of the nearest neighbor “greedy” algorithm to select the stops that will be served by the bicycle-towed trailer. This algorithm has been shown to be an effective method for determining reasonable solutions to traveling salesman type problems (Anderluh et al., 2017). To start this process, the program management office is selected as the start and finish point for the route since the bicycle-towed trailer is stored at this location. Next, the system selects the first stop within the route as the hub location closest to the starting point (i.e., the nearest neighbor) that needs to have a number of bicycles picked up which does not exceed the trailer capacity (which is three bicycles in our test case.) The program then surveys the additional stops that need to be visited and selects the nearest stop that (i) either needs some or all of the bicycles on the trailer or (ii) if the trailer still has capacity, the stop that needs to have a number of bicycles picked up that will fit within the remaining unused trailer capacity. This process is repeated until an additional stop cannot be visited within the maximum bicycle route time entered by the user. Once the stops to be visited by the bicycle trailer are identified, 50 randomly generated sequences of those stops are created for the initial population of the GA. A population size of 50 potential routes was chosen based on similar applications of GAs to scheduling problems (Meister et al., 2005).

It is important to note that each randomly generated route sequence is checked to ensure that it will service the pickup and drop off requirements of the stops to be visited before it is included in the population of possible routes to be evaluated. We assess the feasibility of each randomly generated route sequence using the following process: starting at the origin and progressing in order through each stop in the proposed route, we examine if the number of bicycles on the rebalancing vehicle ever drops below zero or if the capacity of the rebalancing vehicle is exceeded (e.g., a route is not feasible if an empty trailer shows up at a stop that needs two bicycles dropped off – which would result in an ending inventory on the trailer of negative two, or if the number of bicycles on the vehicle after a pickup is greater than the vehicle’s capacity.) If either condition occurs, that proposed route is deemed to be infeasible and replaced with a new randomly ordered sequence (which is also checked for feasibility.) This process continues until 50 feasible proposed routes have been created. Once created, these 50 potential routes for the bicycle trailer serve as the initial population that will use the GA’s iterative process (described below in detail) to optimize the order of the stops.

A van serviced route is required if the bicycle trailer route does not visit all of the stops required to rebalance the system or if a stop requires a number of bicycles to be moved that exceeds the towed trailer’s capacity. Thus, the initial population of potential route sequences for the van route is created by randomly reordering the remaining stops not visited by the bicycle trailer (again using the program management office as the origin and finish point of each route) until a population of 50 potential route sequences is created. As with the bicycle routes, the feasibility of each randomly generated route sequence is assessed before a route is included in the population to be evaluated.

3.5.3 Optimizing a Rebalancing Route

Once a population of potential routes exists for either the bicycle-towed trailer rebalancing or van rebalancing, the GA develops routing solutions using the same process. Continuing with the second step in Mitchell’s (1998) recommended procedure, the GA first evaluates each possible route to determine how effectively it is servicing the stops (referred to as the “fitness” of a route.) Fitness is measured as the time required to either bicycle (with trailer) or drive to all of the hub stops in the route and return to the starting location. To assess the fitness of each potential route sequence, we calculate the objective function value for each – which is the total travel time required to complete a route (T). This is calculated as the sum of transit times between each of the stops in a route. Thus, the objective of the GA is to find a route sequence that minimizes the sum of the transit times between the stops within a route. Expressed mathematically, for a route consisting of n stops where mi is the transit time in minutes from the prior stop i-1 to the stop i, the total route time objective T that the GA is optimizing is:
Once the fitness of potential route sequences within a population are evaluated, the process to create a new generation of 50 route sequences begins. To preserve the best existing solutions (and prevent a new generation from underperforming a prior generation), the best solutions from the current generation are cloned into the next generation’s population – in our case we copy the best three route sequences to the next generation. Crossover, the third step in Mitchell’s (1998) evolutionary process, is then used to create the next 42 potential route sequences (Miller and Goldberg, 1995). As shown in Figure 3, a single new potential route is generated using crossover, which in essence, is a hybrid of two routes in the current population that are randomly selected as “parents” that are then “mated” to create this new “child” route. To do this, the GA randomly selects the next unvisited stop (i.e., allele) in one of the two parent routes as the next stop in the child route; this random selection process is replicated until a child route that visits all of the required stops is created. If the child route is found to be feasible, it is then added to the next generation’s population. If it is not feasible, two new parents are randomly selected and the creation of a new child route is initiated. Crossover is complete once the next generation reaches a population size of 45 possible feasible routes (comprised of the 3 best clones from the current generation and the new 42 offspring routes.)

Mutation, the final step suggested by Mitchell (1998), is used to create the final five child routes in a new population generation. The five routes are created by mutating routes randomly selected from the 45 child routes in the new generation. Mutation, which adds diversity to the extant pool of routes, is accomplished by randomly selecting one of the child routes, copying it and then arbitrarily swapping the order of two of the stops within the route’s sequence (Paydar & Olfati, 2018). If the mutated route is found to be feasible, it is added to the population; otherwise, a new child is randomly selected and mutation is restarted. The mutation process repeats until the final 5 child routes are added to the next generation and the total population size reaches 50 feasible route sequences.

![Fig. 3. Crossover within the Genetic Algorithm.](image-url)
The GA routine is set to run for a maximum of 1,000 generations, after which the best route identified is selected. However, the GA will terminate earlier if no improvement in the best route time is found for the prior 100 generations. If the GA is optimizing a bicycle trailer route it will then run again to find a rebalancing schedule that utilizes the van for any remaining stops to be visited. Upon completion of this process, the program creates a set of turn by turn instructions for the van and/or the bicycle-towed trailer (depending on which modes are utilized in the plan.) These instructions also include details on how many bicycles to pick up or deliver at each stop.

4. Testing and Results

4.1 Test Bikeshare Program

We partnered with the Boise GreenBike bikeshare program (https://boise.greenbike.com/) to test the tool. The Boise GreenBike program was launched in 2015 as a sustainable transportation alternative to motor vehicles and other transit modes. As is the case for most modern bikeshare programs, bicycles are equipped with a GPS tracker and user interface. Patrons can also interface with the system (i.e., access station inventory levels and reserve bikes) via a cell phone application or internet browser. Management can access relevant data, including hub inventory levels and usage patterns, via the internet. At the time of this study, the program operated 114 bicycles, available for rent to the general public at 37 “hub” locations (Boise GreenBike, 2016).

Like many bikeshare programs, the Boise GreenBike program faces severe resource and budgetary constraints that limit its ability to effectively execute operations – the program relies largely on grant funding and its operations are supported by a staff comprised of only three full-time employees. Currently, rebalancing is accomplished using a motorized transport van; however, the Program Director is hoping to improve the sustainability of the operation by utilizing a trailer towed behind a bicycle in addition to the van. In discussions, it was determined that bicycle-towed trailer rebalancing would be more readily adopted if the employees have access to a planning tool that quickly develops efficient rebalancing plans. Specifically, more efficient routing will reduce the travel distance needed to redistribute bicycles across the system, which translates into time savings and (perhaps more important to the employees) less pedaling effort.

At the initiation of the study, the staff at Boise GreenBike planned its rebalancing routes using a purely manual process. This process required a staff member to download and open a bicycle hub inventory report. This report was then examined to determine which hubs need to have bicycles added and/or removed. Then, using a map of the system, the staff member would manually create a rebalancing route that visited the hubs needing attention in a sequence that resulted in each of these hubs ending up with the desired number of bicycles. The rebalancing route planning process typically took between 15 to 20 minutes to complete. This manual method did not include real-time traffic, nor were multiple different routes generated in an attempt to find an optimal or near-optimal route.

4.2 Evaluation of the Tool

To test the functionality of the tool, we utilized historical system data from the summer of 2017 and spring of 2018. Spring and summer encompass the busiest times of the year for Boise GreenBike and thus the times with the greatest need for rebalancing. We utilized randomly selected inventory scenarios from days within these two time periods to evaluate the tool’s ability to generate high quality rebalancing schedules. We assessed the performance of the tool in three key areas: 1) Changes in planning time, 2) Route efficiency versus optimal routings, and 3) Route efficiency versus manually scheduled routings.

4.2.1 Improvements in Planning Time

The determination of the near-optimal routings and creation of directions typically requires one to two minutes of processing time. Altogether, the planning of a typical day’s rebalancing routes requires between 3 to 6 minutes – a considerable savings compared to the manual process which generally requires between 15 to 20 minutes. We estimate that the streamlined process could reduce the time spent planning by over 40 hours per year.
4.2.2 Improvements in Route Quality vs. Optimal Routings

To evaluate the effectiveness of the GA in developing routings compared to the optimal routings, we fully enumerated 10 typical rebalancing scenarios. Since the GA separately optimizes the bicycle trailer and van routes, we chose five scenarios in which the bicycle-towed trailer alone could serve the hubs and five scenarios that only required the van. In practice, the trailer routes are typically shorter, subsequently the bicycle scenarios that were selected for full enumeration each required five stops. In contrast, the van scenarios that were tested each required 10 stops. The upper limit of 10 stops was chosen because these cases could be solved in a reasonable time frame, whereas full enumeration of routes incorporating more than 10 stops requires tens of millions of possible routings to be explored. The full enumeration routine required the feasibility of each possible routing to be evaluated, i.e., the route had to visit its stops in an order that permitted the vehicle to pick up and then have an inventory of bicycles onboard to drop off where needed in the system. Consequently, the optimal route for a scenario is not necessarily the shortest duration route, but rather the shortest duration route among the feasible solutions. The results of the full enumeration tests, which are presented in Table 1, show that the GA determined routes that, on average, were within 1.3% of the optimal solution. The GA determined the optimal routing for four of the five shorter bicycle trailer routes and for one of the five longer van routes. The GA also determined its routes much more quickly compared to full enumeration – on average the GA solved each five-stop route in 17 seconds and each ten-stop route in 25 seconds. By comparison, full enumeration required 20 minutes on average for each five-stop route and 160 minutes on average for each ten-stop route. On average, the GA solution time required approximately 98% less time than full enumeration.

Table 1 Genetic algorithm performance versus optimal solutions

<table>
<thead>
<tr>
<th>Test Route</th>
<th>Number of Hubs Visited</th>
<th>Optimal Route Travel Time - Full Enumeration (Minutes)</th>
<th>Genetic Algorithm (GA) Route Travel Time (Minutes)</th>
<th>GA vs. Optimal (% Difference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bicycle Trailer Route A</td>
<td>5</td>
<td>56.4</td>
<td>56.4</td>
<td>0.0%</td>
</tr>
<tr>
<td>Bicycle Trailer Route B</td>
<td>5</td>
<td>93.9</td>
<td>93.9</td>
<td>0.0%</td>
</tr>
<tr>
<td>Bicycle Trailer Route C</td>
<td>5</td>
<td>129.3</td>
<td>129.3</td>
<td>0.0%</td>
</tr>
<tr>
<td>Bicycle Trailer Route D</td>
<td>5</td>
<td>86.2</td>
<td>86.4</td>
<td>0.2%</td>
</tr>
<tr>
<td>Bicycle Trailer Route E</td>
<td>5</td>
<td>74.1</td>
<td>74.1</td>
<td>0.0%</td>
</tr>
<tr>
<td>Van Route A</td>
<td>10</td>
<td>104.1</td>
<td>108.6</td>
<td>4.3%</td>
</tr>
<tr>
<td>Van Route B</td>
<td>10</td>
<td>114.3</td>
<td>114.3</td>
<td>0.0%</td>
</tr>
<tr>
<td>Van Route C</td>
<td>10</td>
<td>105.6</td>
<td>112.8</td>
<td>6.8%</td>
</tr>
<tr>
<td>Van Route D</td>
<td>10</td>
<td>137.1</td>
<td>138.3</td>
<td>0.9%</td>
</tr>
<tr>
<td>Van Route E</td>
<td>10</td>
<td>104.1</td>
<td>104.7</td>
<td>0.6%</td>
</tr>
</tbody>
</table>

GA vs. Optimal (Average % difference) = 1.3%

The performance of the routing algorithm was also gauged by comparing routes generated by the GA against routes created using the existing manual process. Using historical data, routings were manually generated for 15 van only, 15 bicycle and van, and 15 bicycle only rebalancing scenarios. These same 45 scenarios were also routed using the tool. On average, the tool generated route sequences that required 14% less travel time compared with the manually created routes. Explicitly, the average manually generated route required 110 minutes of driving and/or bicycling transit time, while the average routing generated by the tool required 94 minutes of transit time. Further, in all 45 cases, the routes generated by the GA had transit times which never exceeded the transit times of the manually created routings. Figures 4a and 4b illustrate a sample route that was manually scheduled and then improved using the GA tool. In this example, the GA suggested routing reduced the driving time on this route by 5 minutes, representing an 11%-time savings. Given that the routes used in this exercise are representative of typical rebalancing routes at Boise GreenBike, over a year, the optimized routings will save approximately 50 hours of travel time per year. Additionally, we estimate the new routings will eliminate roughly 240 miles of driving for the van serviced routes – which will prevent over 600 pounds of CO₂ from being emitted into the atmosphere (U.S. EIA, 2018).
Fig. 4a. Manually scheduled route (Driving time = 45 mins., Distance = 12.7 mi.)

Fig. 4b. GA optimized route (Driving time = 40 mins., Distance = 12.3 mi.)
5. Contributions, Limitations, and Future Research

We set out to develop a tool that would generate solutions that are as effective as others described in the bikeshare rebalancing literature, but using lower cost software while considering some of the complexities encountered by bikeshare programs. By using MS Excel, which already resides on most desktops, our solution can be delivered at lower cost than algorithms that require higher cost routing packages. Our tool is also distinct from other solutions because it uses actual travel data, including real-time traffic conditions (acquired at no cost to the user). In addition to its applied value, our simple method for acquiring realistic travel time and distance data via an API facilitates more realistic testing of future routing algorithms. Our tool makes a further contribution by accommodating multiple heterogeneous repositioning vehicles. In our test organization, the tool generated solutions that are (on average) within 1.3% percent of the optimal solutions, while demonstrating the potential to save approximately 90 hours of labor annually.

Generalizability, especially with respect to scale, is an issue in any study of this type. Currently, 75% of all bikeshare rides in the US are generated by programs in just four cites (New York City, Chicago, Boston and Washington, DC) with the remainder being dispersed over a large number of programs that are smaller in terms of overall population, size of the bikeshare program or both (NATCO, 2018). This study’s test city, Boise, Idaho, falls into that latter group. Boise has a metropolitan statistical area (MSA) population of 709,000 making it the 80th largest city in the United States (U.S. Census Bureau, 2017). To put this in perspective, the United States has 134 MSA's with populations between 250,000 and 1,000,000. This suggests that many communities could potentially institute bikeshare programs of roughly the same magnitude as the one studied here. That said, the size of our test program should not be viewed as an upper limit for our tool. Although our system was developed to meet the needs of a program that utilized 37 hubs and approximately 120 bikes, the user interface and the underlying GA can be applied to programs with more hubs and bicycles without modification. However, the factor that limits the no-cost use of the tool is the query limit imposed by the Google Maps Directions API. The daily limit of 2,500 queries means that users are prevented from building a travel matrix incorporating more than 50 locations in a single day (because for \( n \) hubs, \( n^2 \) queries are required to collect data between all possible location pairings.) Nevertheless, this limitation can be overcome if the travel matrix is built utilizing standard travel data (i.e., non-real-time) and if the queries are spread across multiple days. Conversely, if the user is choosing to utilize real-time traffic data, the query limit will prevent the full travel matrix from being completed for routes incorporating more than 50 stops. However, this limitation is easily (although not freely) overcome through a paid data subscription.

Several emerging trends in shared mobility systems could be the subject of extensions to this research project. Dockless systems (Ma et al., 2018; van Waes et al., 2018) are replacing many hub systems and this trend is likely to continue. For example, in China (where dock-based systems were already better established than in most other geographies), dockless bike sharing grew tenfold from 2016 to 2017 (Gu et al., 2019). Our tool can be easily applied to dockless systems. The user simply needs to enter the addresses of the bicycles to be picked up (or a token address if there are bicycles at numerous addresses within a small area) and the desired locations where they are to be dropped off. This is an advantage of a system like ours that generates routes using actual travel distances. With that said, an area of future research would be creating optimal clusters of for pickup and drop-off in a dockless system. Since the relative share of the carbon greenhouse gas footprint attributable to rebalancing is much higher in dockless systems than it is in dock-based systems, the benefits from improving rebalancing routings in dockless systems are potentially greater as well (Luo et al., 2019). An additional area for future research is predictive rebalancing. Such a system would contain a forecasting component in order to predict demand at origin locations, and this would be paired with routing optimization. Such a system could guide the user in making trade-offs between rebalancing-related costs and customer service (stockout) levels.
References


Raviv, T., & Kolka, O. (2013). Optimal inventory management of a bike-sharing station. IIE Transactions, 45(10), 1077-1093.


