Determining an optimum inventory route for an areal object: the case of forest inventory in Finland

HENNA ETULA AND HARRI ANTIKAINEN

In recent decades, routing based on Geographic Information Systems (GIS) has become a major branch of technology, which has been used especially in applications related to transport and logistics. However, in terms of the development of methods, routing in a cross-country environment is more difficult, and hence research into it has been relatively scarce. This is particularly true in the context of complex routing problems involving visits to several locations. A typical example of a problem of this kind is field inventory, which is a data collection procedure used in many application areas, particularly those related to environmental research and the management of natural resources. This study presents a problem in which an efficient inventory route is determined for an areal object, such that the area visible from the route meets a prescribed threshold, while maintaining the shortest possible route. Although this problem, referred to here as the Areal Inventory Problem (AIP), is closely related to a multitude of routing and location allocation methods known in the context of GIS, none of them is very well-suited for solving the AIP. This study describes a general solution procedure for the AIP, and introduces an implementation of a heuristic algorithm that can be used to solve a real-world AIP within a reasonable time frame. The proposed approach is demonstrated with actual data related to field inventory practices carried out by the Finnish Forest Centre.

Keywords: Finland, inventory of forest resources, Areal Inventory Problem, route optimization, location allocation, maximum coverage

Henna Etula, Suomen metsäkeskus, Myllykatu 15, FI-65100 Vaasa, Finland. E-mail: henna.etula@metsakeskus.fi
Harri Antikainen, University of Oulu, Department of Geography, Po Box 3000, FI-90014 University of Oulu, Finland. E-mail: harri.antikainen@oulu.fi

Introduction

The determination of optimum routes is a line of technology, which has become an essential part of modern society, mostly used in logistics, fleet management, and private transport. Besides the commercial or public sector applications, it is equally well established also in scientific research, manifested by the vast amount of studies considering accessibility, traffic simulation and site selection problems, which all involve the search for optimum routes.

The general view of routing technology tends to emphasize the recent advances in Geographic In-formation Systems (GIS), positioning technology (mainly Global Positioning System, GPS) and the computational capability of mobile devices, together with the increased availability of digital data products concerning transportation networks, as the main contributors to its broad use today. While the central role of these enabling factors cannot be disputed, it is important to realize that the route search methods and representations used in routing applications date back to at least the 1950s, or even earlier. Indeed, modern-day routing applications are strongly founded on the concepts and representations of graph theory, which is a branch of mathematics focused on the notion of

a graph (Miller & Shaw 2001). A graph is essentially a structure used to model pairwise relations between objects, and it is typically denoted as

$$G = (N, E), \qquad (1)$$

consisting of a set of nodes (or vertices) N and a set of edges (or arcs or links) E. An edge

$$e = (i, j) \in E, \qquad (2)$$

connects nodes i and j, and has an associated cost c_{ij} (Diestel 2000). The cost may represent the impedance of travel, such as geographical distance or time, between i and j. An optimum, least-cost path between any two locations within the graph can be calculated using a graph search algorithm, such as the classic algorithm originally proposed by Dijkstra (1959).

The graph structure lends itself very well to representing the natural topology of transportation networks, allowing optimum transportation routes to be found efficiently. This is because a transportation network is functionally one-dimensional, and therefore it can be directly represented as a graph. However, finding routes through two-dimensional space, such as cross-country terrain, is more difficult because the two-dimensional space with unlimited movement options must first be transformed into a one-dimensional graph structure with a finite set of movement options. In GIS, this can be realized by the cost surface approach. A GIS-based cost surface is essentially a raster representing continuous two-dimensional space. Each cell of the cost surface raster is assigned a "cost", which depicts the impedance of movement across the cell. The raster is transformed into a graph by treating the center of each cell as a node, and the connections between adjacent nodes as edges, weighted according to the underlying cost values. An optimal path between any two locations (nodes) can be found by determining the sequence of moves constituting the least possible accumulative cost between the locations (Bolstad 2002). For this purpose, the same graph search algorithms can be used as in the case of graphs representing transportation networks.

The cost surface method has been used in many real-world applications, such as in determining the optimal alignment of linear constructions, like trails (Xiang 1996), roads (Yu et al. 2003), canals (Collischonn & Pilar 2000), power transmission lines (Bagli et al. 2011), and pipelines (Feldman et al. 1995). The method has also been used for modeling the patterns of human movement in archaeological research (Howey 2007; Anderson 2012), as well as in the management of protected and recreational areas (Theobald et al. 2010; Tomczyk & Ewertowski 2013). Likewise, the method has been applied in ecological research to predict the movement and migration of wildlife (Lundqvist 2007; Parks et al. 2013). These examples depict the versatility of the method: it can be applied in very different domains by simply adjusting the cost parameter used in the method.

One of the main limitations of the least-cost path method based on cost surface analysis is that the path can only be calculated from one location to another location (or a set of other locations). Hence, it is not possible to use the method to solve more complex routing problems, such as finding a route visiting several locations in an optimal order. Certainly, methods for solving problems of this nature exist, and some of them have been implemented in the mainstream GIS software packages. These implementations typically assume that a transportation network (or a similar, existing network) is used as a basis for route optimization. Using a raster graph for this purpose is impractical due to the large number of nodes with many connections. Instead, the least-cost path analysis is sometimes used to first create a graph connecting a predetermined set of locations, and this graph is then used to solve a routing problem involving visits to several locations. For example, Balstrøm (2002) has employed cost surface analysis to calculate least-cost paths between all pairs of rain gauges in a study area located in mountainous terrain, and the network composed of these least-cost paths was used to determine the most optimal visiting sequence of the gauges with a routing method. In another study, Store and Antikainen (2010) adopted a similar approach by calculating least-cost paths based on a cost surface raster between forest stands to be inventoried. However, in this case each forest stand was assigned an importance score, and the task was to use the graph created with the least-cost path analysis to determine the most important stands that could be inventoried within a prescribed amount of time spent in the field.

A basic assumption used in the aforementioned studies is that the locations to be visited comprise

a finite set of discrete locations, which can be represented as a graph. However, in some routing problems, there may not be any predefined discrete locations that could serve as a basis for the graph representation. This study is motivated by a real-world routing problem related to the field inventory of areal objects, which in this case are forest stands. This problem asks for the shortest possible route to be found inside an areal object such that a proximity buffer drawn around the route, representing visibility from the route, covers at least the predefined proportion of the areal object. In the following, the background of the problem and the motivation for the study are described in more detail, and a literature review of related problem types is presented. This is followed by the description of a formal solution procedure for the problem, and the implementation of the solution is presented along with a sample problem and solution examples.

Field inventories

Field inventory is an essential but labor-intensive way to gather and produce information in many application areas, especially those pertaining to environmental research and the management of natural resources. A prime example of this are forests, for which there is a worldwide demand for information (Tomppo et al. 2008). Some of the need for information is related to the questions of biodiversity and ecosystems, while some are obviously related to economic aspects, involving a variety of resource modeling procedures and management plans. The information is typically collected using different methods and is carried out by many organizations. For example, the Finnish Forest Centre, a governmental forestry organization, collects information about privately owned forests throughout Finland. The information is collected both by the means of remote sensing (RS), including airborne laser scanning (ALS) and aerial images, and field inventory (Maltamo et al. 2011). While the RS techniques, combined with reference plots examined in the field, allow valid data to be produced for most areas of interest, for a subset of areas (such as seedling stands and temporarily unstocked regeneration areas), traditional field inventory remains the only appropriate option for data collection.

There are two types of inventory units used in the field inventory practices of the Finnish Forest

Centre: points and polygons. Points are associated with the measurement of reference plots, in which the center point of the plot is located with the help of a GPS device. The location is used as a focal point around which the trees are measured according to the given data collection instructions. Polygons are associated with the so called stand-wise inventory. A forest stand is a basic unit in forestry, defined as an area which is more or less homogeneous with regard to growing stock and forest type. The size of a stand typically ranges from a half hectare to five hectares. The stand-wise inventory is not based on specific measuring points, as in the measurement of reference plots, but rather it is performed by making overall observations and measurements along an inventory route passing through the stand. The accuracy of the inventory may not always be the same for all stands (for example, mature stands are measured more carefully while data from seedling stands is often collected only by visual observations); however the same field inventory procedure is applied for all stands, requiring that an adequate level of information has to be acquired from the stand. The only way to accomplish this is to visit and inventory the stand in the field.

The efficiency of field inventory is highly important. For example, the inventories carried out by the Finnish Forest Centre cover about 1.5 million hectares of forestland every year, constituting a major data collection and maintenance effort. While around 75% of data can be obtained by using modern RS techniques, one fourth of the inventory area still requires a stand-wise field inventory. Due to the high cost of field inventories, special emphasis must be devoted to the efficiency of the inventory.

There are two principal ways of improving field inventory, and thereby keeping the costs within acceptable limits. The more traditional way is the careful planning of the used inventory method and the data to be collected. It is important to gather all required data during a single field visit, without a need to revisit the area later for complementary data collection. It is also useful to concentrate on properties that can be measured easily, and then use them for further calculations to derive other properties. For example, in forest inventory, the volume of the growing stock is difficult to estimate correctly in the forest, but it can be calculated based on the average diameter and height of the trees, which can be measured easily.

Forestry field inventory methods and practices have been developed and fine-tuned over decades, and the current procedures can be considered to be efficient already, leaving little room for further improvement. Therefore, instead of the actual data collection methods, there is growing interest towards seeking improvement by means of route optimization, which constitutes an additional strategy for making field inventory more efficient. The potential advantages of optimizing inventory routes seem obvious, as it provides a way to design routes that avoid unnecessary and inefficient movement in the inventoried area. By this means, a maximum amount of information can be gathered with a minimum effort or cost. This study is specifically concerned with the task of designing the route in such a way that an adequate level of coverage is reached, while the route itself is kept as short as possible. The coverage of the route means the surrounding area that can be observed or otherwise reached from the route without stepping out of the route. In this article, this problem is referred to as the *Areal Inventory Problem* (AIP).

Related optimization problems

In addition to the least-cost path problem, the Areal Inventory Problem (AIP) is closely associated with many location allocation problems. Perhaps the best-known and most intensively studied location allocation problem is the *p*-median problem (Hakimi 1964), which seeks to find optimum locations for any number *p* of facilities such that the sum of distances between each (weighted) demand point and the closest facility is minimized (Longley et al. 2005). The solution of a *p*-median problem is a set of *p* facilities, such as health centers or schools, located as centrally as possible with regard to the demand points (representing, e.g., population at a certain level of aggregation).

While minimizing the overall distance needed to access the facilities may be applicable to most situations, for certain kinds of facilities, such as fire stations and other emergency services, the facilities should cover as much demand as possible within a prescribed time or distance. This kind of problem, which is commonly referred to as the Set Covering Location Problem (SCLP) (ReVelle & Toregas 1972), seeks to allocate the demand in a particular area to a

minimum number of facilities such that the distance between any facility and a demand point allocated to it does not exceed a pre-specified threshold. Of course, due to the scarcity of resources, in many real-world applications it is not feasible to attempt to allocate facilities to cover all demand within a certain distance threshold; rather it is more useful to allocate facilities such that as many demand points as possible are located within the distance threshold. This variation of the problem is known as the Maximal Covering Location Problem (MCLP) (Church & ReVelle 1974). The MCLP is typically used to solve commercial allocation problems, such as the location of new retail centers.

The AIP can be thought of as an instance of the MCLP where the "facilities" are the vertices of the route and the distance threshold is the degree of visibility from the vertices. The length of the route can be minimized in the AIP by applying methods used to solve the Traveling Salesman Problem (TSP) (Applegate et al. 2006). The TSP is a classic problem that seeks to find a shortest possible tour through a set of locations such that the tour visits each location exactly once.

The TSP and the MCLP can be combined to constitute a specific kind of problem, namely the Covering Salesman Problem (CSP) (Current & Schilling 1989). In the CSP, the task is to find the shortest possible tour visiting a certain subset of locations, while making sure that all locations excluded from the tour are within a prescribed distance from the closest location included in the tour. Several variations of the CSP have been proposed in the literature. In the Covering Tour Problem (CTP) (Gendreau et al. 1997), the locations visited by the tour can be predetermined. Unless locations are pre-assigned to the tour, the problem is reduced to the general CSP. On the other extreme, if all locations must be included into the tour, the problem is equivalent to the basic TSP.

In the Generalized Covering Salesman Problem (GCSP) (Golden et al. 2012), visits to the locations can also be varied: it is possible to "stay overnight" at a location or visit it more than once, if this is necessary to cover the demand. The CSP has also been refined by the original authors of the problem in a study where they define two specific kinds of problems: the Median Tour Problem (MTP) and the Maximal Covering Tour Problem (MCTP) (Current & Schil-

ling 1994). In the MTP, the objective is to minimize the total (weighted) distance of locations not on the tour to the nearest location included in the tour. In the MCTP, on the other hand, the task is to minimize the total demand associated with the locations which are not within a certain pre-specified maximal distance from a location on the tour.

Apart from the formulations and solution options of the allocation problems, there has been an increasing interest towards assessing the problems with respect to how the demand is covered in allocation problems. A traditional approach is to assume that the demand is divided into cells of equal size, and for simplicity, each such cell has been discretized into a point located at the centroid of the cell. This zero-dimensional location thus represents all of the demand contained by the cell, such as population or area, for example. If the centroid is covered in the solution of the allocation problem, the entire demand of the cell is considered to be covered. This may give rise to misleading and suboptimal results. Obviously, a more realistic approach would be to represent the demand by polygons rather than points.

Alexandris and Giannikos (2010) have assessed coverage gaps by comparing the coverage calculated according to grid cells, and the coverage calculated according to spheres drawn around point locations. In their own experiment, the authors demonstrate that a better coverage can be achieved with the same number of facilities as in the traditional approach. At the same time, a partial coverage is taken into account. This signifies that a majority of the demand contained by the cell may become covered, even if the centroid of the cell is not covered (Murray 2005). In that case, the calculation of the coverage is not particularly sensitive to changes in cell size, and it is also less susceptible to the Modifiable Area Unit Problem (MAUP).

Overall, the AIP is related to multiple location allocation and routing problems well known in the literature, the MCTP in particular. However, in the AIP, the coverage is assessed in terms of areal coverage, instead of the coverage of discrete locations. As a result, it is essential to take account of the suggestions made by Alexandris and Giannikos (2010) regarding the calculation of the actual coverage when implementing a solution procedure for the AIP.

Solving the Areal Inventory Problem

Problem formulation

The inventory of an areal object (polygon) is carried out by visiting a set of observation points, which are locations where the inventory technician stops to carry out measurements and make observations about the surrounding environment. There is no supply variation, signifying that all observation points are assumed to produce the same amount of information. In addition, there are no preconditions to the location of the observation points, other than that they are supposed to be located inside the inventoried polygon. However, the number of observation points inside the polygon, n, is determined before the inventory, based on the existing knowledge of a similar inventory. The inventory route goes through these points, such that the transfer between two consecutive points takes place along the shortest path between them.

As the route should be as efficient as possible, the length of the route must be minimized. The solution algorithm proposed here is thus related to the MCTP, signifying that the aim is to produce the shortest possible route that seeks to reach an adequate level of coverage attainable by n points. The observation points are not determined in continuous space, instead the polygon is represented as a finite set of candidate locations among which the n observation points are selected. However, the *coverage* of the inventory route is determined in continuous space, which is done by defining a buffer around the points where observations are made. Since the inventory technician is also expected to make observations along the route when moving from one observation point to the next, the entire route is buffered, instead of the actual observation points only. The width of the buffer, *Buf,* which in this case represents the length of visibility, can be a constant for all stands, or a variable depending on the properties of the forest (e.g., the length or density of growing stock) or the ruggedness of terrain. The visible area outside the boundaries of the inventoried polygon is discarded as irrelevant. The calculated route is valid provided that its coverage C (which is the visible portion of the inventoried polygon divided by the entire surface area of the polygon), is equal to or greater than a predetermined threshold C_enough. Again, this parameter can be a constant or a stand-specific variable depending on the growing stock.

The definition of the problem is as follows:

Minimize the length of the inventory route subject to

> (1) the route goes through n points,
> (2) the coverage C of the route is equal to or greater than C_enough,

where
> (3) the visible area of the polygon to be inventoried (*P*) is created by buffering the inventory route and the observation points with the buffer width *Buf* and, resulting in a buffer polygon *B*,
> (4) the buffer polygon *B* is intersected by polygon *P*, in order to produce the visible area (*BI_area*) inside the polygon and
> (5) the coverage *C* of the route is calculated by dividing *BI_area* by the surface area of *P*.

Solution procedure

This section describes an algorithm that can be used to solve the AIP. The procedure of the algorithm is presented in the form of pseudocode, but an illustration of the algorithm is also provided as a flowchart (Fig. 1).

Input parameters:
P = the areal object to be inventoried (polygon),
e = entrance (point),
x = exit (point).

Constants:
A = the surface area of *P*,
n = the number of the observation points,
Buf = a buffer width and
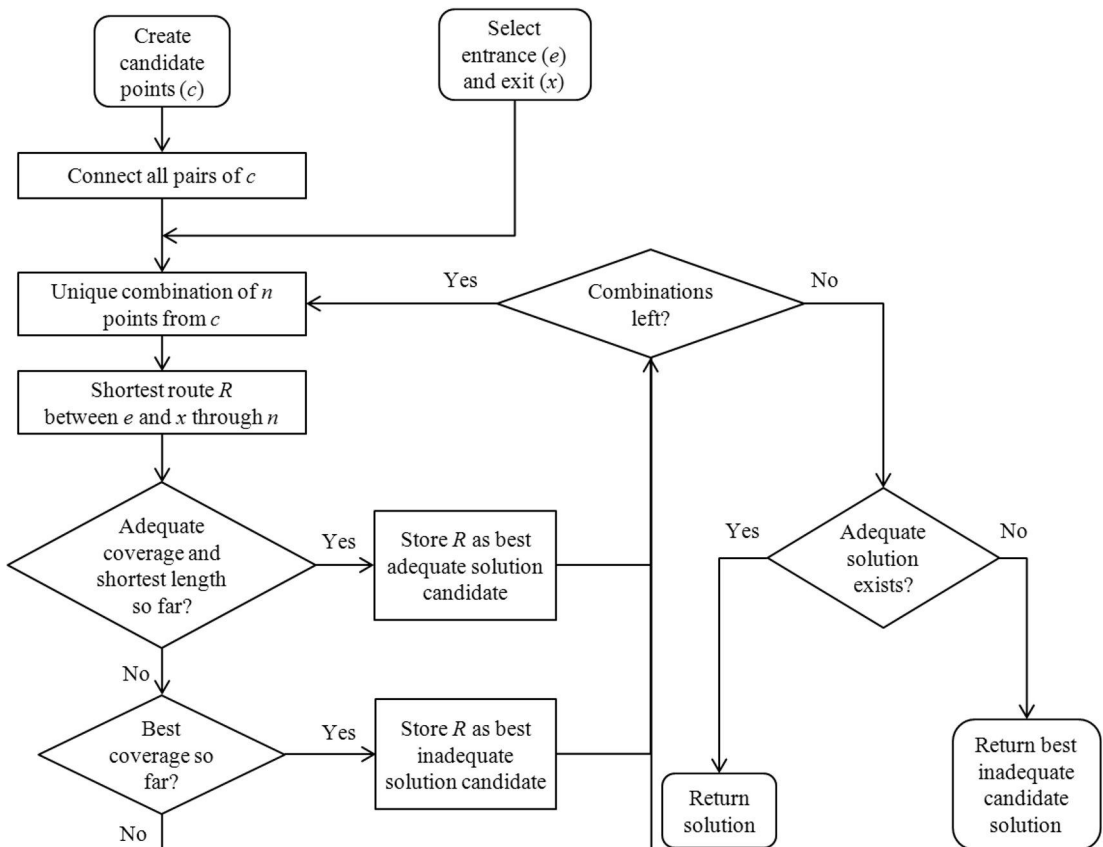*C_enough* = an adequate coverage for inventory.

Fig. 1. The procedure for solving the areal inventory problem.

Variables:

R_temp = a temporary variable for the route in process (polyline),
L_temp = a length of R_temp,
B = a buffer around the route R_temp (polygon) with buffer width Buf,
BI = B intersected by P (polygon),
BI_area = the area of BI and
C_temp = coverage associated with route R_temp.

Output parameters:
R = the best route (polyline),
L = the length of the best route and
C = coverage of the best route,
or alternatively
C_max = the maximum coverage if the coverage of any route does not reach C_enough
$R2$ = the output route if C_max is used (polyline) and
$L2$ = the length of the route $R2$.

The algorithm:

Step 1: Create a set of candidate points with regular intervals over P. The interval depends on A. An appropriate point interval may be, for example, 20 meters.

Step 2: Set variables:

R	= null,
L	= ∞,
C	= 0,
$R2$	= null,
$L2$	= 0 and
C_max	= 0.

Step 3: Select points which are inside of P.

Step 4: **For** all possible sets of n points **do**
  a) Set variables:

R_temp	= null,
L_temp	= 0,
B	= null,
BI	= null,
BI_area	= 0 and
C_temp	= 0.

  b) Choose n points.
  c) Produce the shortest possible route between e and x through n points = R_temp.
  d) Calculate the length of R_temp = L_temp.
  e) Buffer the route with buffer width Buf = B.
  f) Intersect B by P = BI.
  g) Calculate the area of BI = BI_area.
  h) Calculate the coverage C_temp = BI_area / A.
  i) **If** $C\_temp \geq C\_enough$ **and** $L\_temp < L$ **then**
      R = R_temp
      L = L_temp
      C = C_temp

else if $C\_temp > C\_max$ **then**
  $R2$ = R_temp
  $L2$ = L_temp
  C_max = C_temp

Step 5: Print out the output parameters R, L and C. If none of the routes has reached the coverage C_enough, print out $R2$, $L2$ and C_max.

Sample problem

This sample problem illustrates the solving of the AIP for a single forest stand with an area of 2.13 hectares. It is assumed here that a model has been constructed to estimate the parameter n for stands of different shape and size, and according to the model, n is six for this particular stand. As the stand is covered by commercial thinning forest of low density, the visibility in the area is estimated to be 30 m. Due to the small size of the area, topography is not assumed to have any effect on visibility. It is also determined that the entire area does not have to be observed: instead, a 60% coverage is expected to be adequate for data collection purposes.

A set of candidate points at the regular interval of 20 m is created over the stand. The selected interval is a trade-off between accuracy and computation time. 20 m was estimated to be suitable for a medium-sized forest stand, such as the one considered here. The placement of the candidate locations at this interval in the stand results in 54 points. As the inventory route is expected to visit six points, this would amount to $\binom{54}{6} = 25\,827\,165$ different combinations of points. Four of these combinations, and their associated routes, are illustrated in figure 2. Each one of the routes is buffered, resulting in a buffer polygon B, which is then intersected with the boundaries of the stand (resulting in a clipped buffer polygon BI). The variable BI_area is used to denote the observed area of the stand, and C is the coverage score of the route (i.e. the ratio of BI_area to the entire area of the stand). The route length, L, represents the traversed distance between the start and end points of the route.

The coverage of the routes presented in figure 2 varies clearly. Route 3 does not exceed C_enough because all n points are concentrated in the western part of the stand. In Route 4, n points are located near the boundary of the stand. The attained coverage is high but this comes at the cost of route
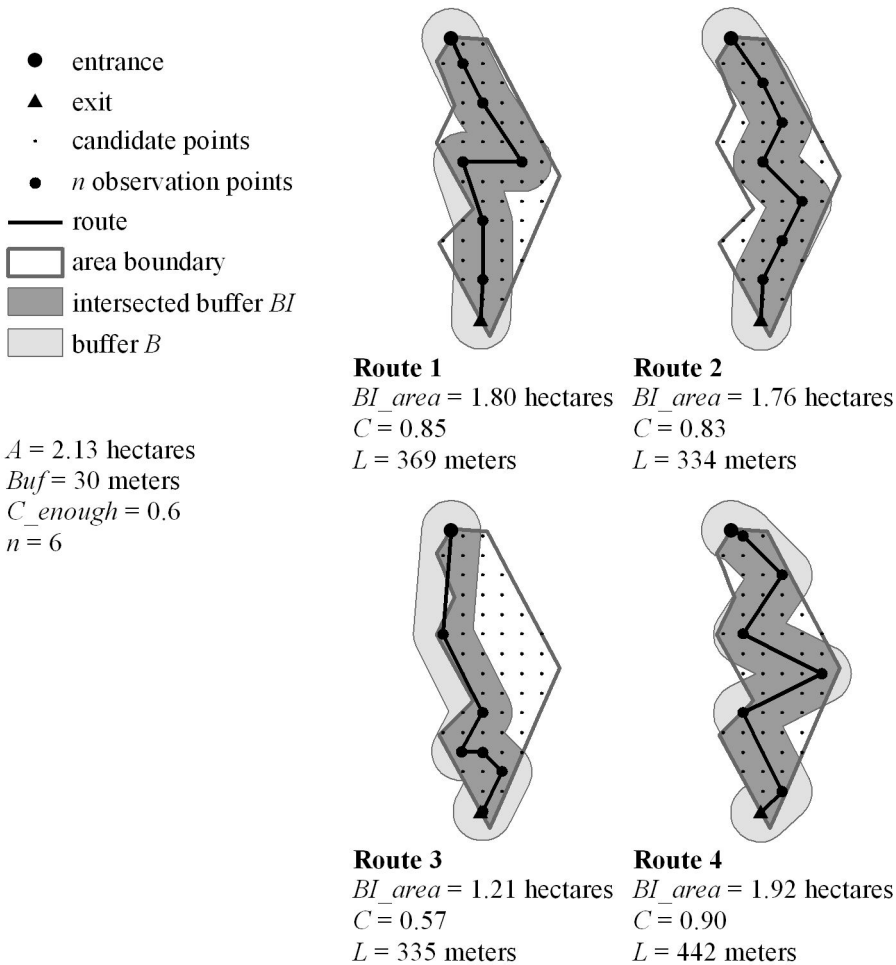
- ● entrance
- ▲ exit
- · candidate points
- ● n observation points
- —— route
- area boundary
- intersected buffer BI
- buffer B

A = 2.13 hectares
Buf = 30 meters
C_enough = 0.6
n = 6

Route 1
BI_area = 1.80 hectares
C = 0.85
L = 369 meters

Route 2
BI_area = 1.76 hectares
C = 0.83
L = 334 meters

Route 3
BI_area = 1.21 hectares
C = 0.57
L = 335 meters

Route 4
BI_area = 1.92 hectares
C = 0.90
L = 442 meters

Fig. 2. Four alternative inventory routes produced for one forest stand. Route 2 is the most efficient because it produces enough observed area with the shortest route length.

length. In addition, a great deal of B is located outside the boundaries of the stand, which is not an ideal situation. Routes 1 and 2 are located in the middle of the stand and the coverage becomes high because B is almost completely located within the stand. Route 3 is the shortest but its coverage is below the required threshold. The three other alternative routes have an adequate coverage, among which Route 2 is the best because it is the shortest of the presented solutions.

It is important to realize, though, that only a small set of alternative solutions are presented for this sample problem, leaving 25 827 161 other solutions without consideration. This signifies that

instead of an exhaustive examination of all alternatives, the actual implementation of the algorithm must employ a heuristic strategy in order to keep computation time within reasonable limits. Heuristic algorithms are used to quickly find a reasonably good solution to a problem in situations where it is considered too time-consuming to search for the absolutely optimal solution. Indeed, due to the NP-hardness of many allocations problems (denoting that their computational complexity increases exponentially with the size of the problem), their solution strategies are commonly based on heuristic algorithms (Church & Sorensen 1994).
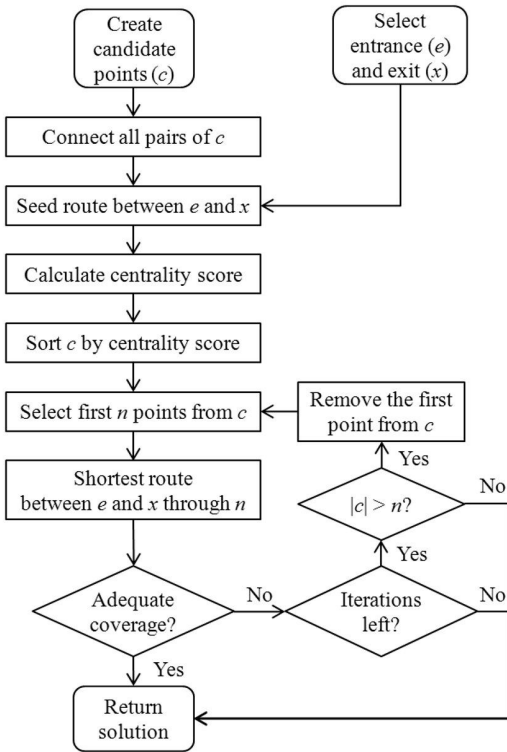
Fig. 3. The heuristic procedure for solving the areal inventory problem.

Heuristic solution procedure

Implementation

A software tool was implemented by Esri Finland to solve the AIP, and it was constructed as a geoprocessing model to be run on the ArcGIS software. The AIP tool utilizes the built-in functions of ArcGIS 10.1, including the Network Analyst and Spatial Analyst extensions. Instead of the general solution algorithm presented in this study, the tool employs a heuristic strategy to quickly find a good solution to the AIP. The main steps of the implementation are illustrated in figure 3.

The implemented AIP tool consists of two principal stages. In the first stage, a set of candidate points at a regular 20-m interval is created inside a polygon, and then a mesh of straight-line paths is constructed between the candidate points. The points, along with the mesh, constitute a discretized representation of the movement options inside the polygon. Since forest stands are, by definition, internally homogeneous, a straight line between any two points can be safely assumed to be the shortest path connecting the points. The mesh contains all possible paths between all the candidate points, with the exception that candidate points located less than 10 m from the boundary of the area are excluded, as it is unlikely that they could belong to the optimal inventory route.

The other stage of the AIP tool produces the inventory route inside the polygon. The tool has six parameters: a point where the inventory route starts (entrance point), a point where the inventory route ends (exit point), a number of the observation points (n), a buffer width (Buf), an adequate coverage for inventory (C_enough), and the maximum number of iterations. The entrance and exit points are user-defined locations that can be positioned anywhere on the boundary of the polygon. The route is calculated between these points, visiting any combination of n observation points that provides an adequate level of coverage, defined in C_enough, with the buffer width Buf.

Instead of enumerating all possible route point combinations to find the absolute best solution, the tool employs a heuristic strategy to quickly find a reasonably good route. First, the tool calculates a seed route, which is the shortest possible path between the entrance and exit points. After that, it calculates a centrality score for each candidate point, which is the combined distance of the point from the seed route and from the geographic center of all of the candidate points (Fig. 4). The centrality score is used to "guide" the route calculation process to search for routes that are close to the shortest possible path between the entrance and exit points, and, at the same time, centrally located, thereby being potentially good route candidates in terms of coverage. The points are sorted according to the centrality score, placing the best one first. Then, the route is sought in an iterative manner. At each iteration, the first n points of the list are included in the route as intermediate points between the fixed entrance and exit points. The shortest route through this combination of points is calculated, by allowing the visiting sequence of the intermediate points to
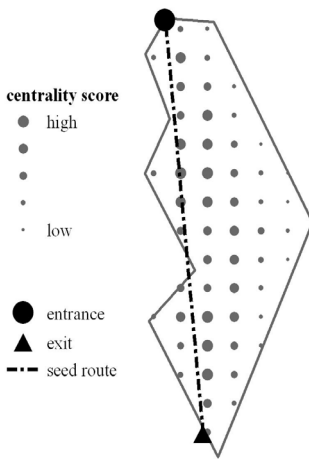
Fig. 4. The heuristic procedure begins with creating a seed route and calculating the centrality score.

vary in order to find the shortest route. As the route is calculated, the tool examines whether the route reached the value *C_enough*. Unless *C_enough* is reached, the first point in the ordered list is removed and the tool moves into the next iteration; otherwise, the tool will stop. The tool will also terminate if all possible combinations of candidate points have already been examined, or if the maximum allowed number of iterations has been reached.

The final result of the tool will be the first of the routes that reaches the threshold. However, if the tool stops due to the maximum iteration condition without reaching *C_enough*, the result will be the best of the examined routes. If the polygon is very small, that is, if *n* equals one, the tool will only calculate a route that visits the center point of the polygon.

Using the software tool

A selection of routes calculated with the AIP tool are depicted in figure 5. The routes are calculated for a forest stand with different entrance and exit points. The area of the stand is 2.2 ha. The candidate points are located at 20-m intervals, the buffer width is 25 m, the adequate coverage is 60%, and the number of the observation points is five. With these parameter values, the tool runs quickly.

The size of the polygon is a significant factor determining for the computational performance of the calculation. Figure 6 shows a relatively large forest stand, having a surface area of 8.0 ha. The parameters are the same as above except for the number of the observation points, which in this case is nine. The AIP tool was run with 10, 25 and 50 iterations, failing to reach the adequate coverage with any one of these options. It can easily be seen in figure 3 that if the polygon in question is large, containing plenty of candidate points, the number of iterations greatly affects the quality of the resultant route. While better solutions can be achieved by allowing more iterations, this has a negative effect on computation time. For example, in the case of this particular polygon, the overall computation time more than doubles between 10 and 50 iterations.

When using the tool it is important to estimate the appropriate parameters and computation time, which are critical when calculating complete inventory routes that visit many forest stands. The computation time can be controlled by adjusting the maximum allowed number of iterations, and by creating a more sparsely spaced set of points for choosing *n* points. Unfortunately, the solution quality will deteriorate as the point set becomes more scattered. On the other hand, a very dense set of points may not help produce a better solution at all; instead, it just increases the solving time considerably. As described above, the number of points can be reduced by ignoring points close to the boundary of the polygon, since it is unlikely that such points can contribute to the extent of the visible area (the buffer outside the polygon will be cut away anyway). Although this helps decrease computation time, its positive effect quickly becomes insignificant as the size of the polygon increases.

Conclusions

This study has presented the Areal Inventory Problem (AIP), which attempts to find a route for an areal object, such that the area visible from the route meets a prescribed threshold, while maintaining the shortest possible route. Although the AIP is related to several classic routing and allocation problems, this particular problem has not received attention in the literature so far. In this study, an algorithm that
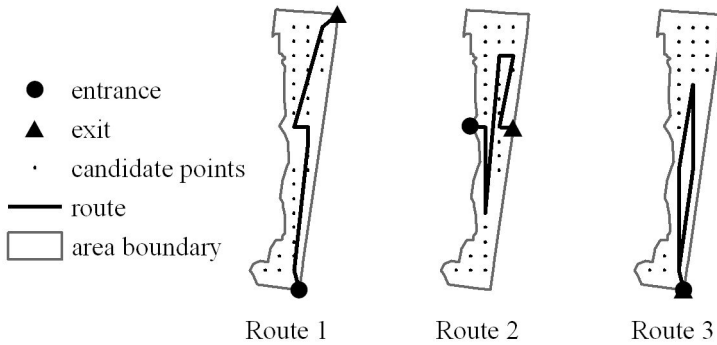
Fig. 5. Three different inventory routes produced for one forest stand with different entrance and exit points.
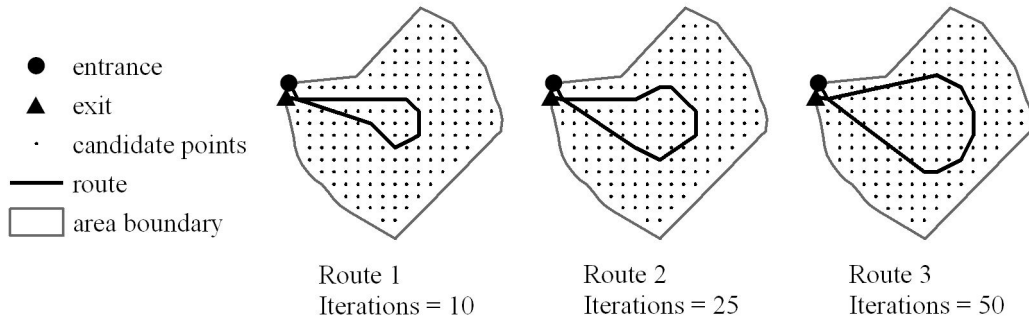

Fig. 6. Three different inventory routes produced for one forest stand with different numbers of iterations.

can be used to solve the AIP in the GIS environment is proposed. This has been demonstrated with a software tool implemented in GIS, which has been tested with actual data related to forest field inventory.

Despite its deceptive simplicity, the AIP is a complicated problem. This requires making certain presuppositions that help reduce the complexity of the problem. The presuppositions employed in this study include the appropriate number of observation points (i.e., the fixed number of points visited by the route), and the interval of candidate points used to represent the polygon and to construct the route. Even with these simplifying presuppositions in effect, the computation time of the AIP increases exponentially with the size of the problem. Due to this property, it is only possible to find a valid solution to very small problems without a heuristic strat-

egy. The heuristic method devised for this study is efficiently guided to quickly find good solutions. Nevertheless, the number of iterations needed to find a solution that meets the given threshold may be high, and an adequately good solution may not be found at all within the limits of the prescribed number of iterations.

Although the proposed method for solving the AIP is intended to enable the automation of a route finding procedure, and to make it possible to solve routing problems that may be too difficult to solve manually, it is necessary to realize that the use of the method still relies strongly on the GIS specialist designing the representation of the AIP. The processing of GIS-based data elements often involves a considerable computational overhead which makes the calculation of routes, already very complicated by itself, even more challenging. The task of a GIS

specialist is therefore to set up the representation of the problem and define its parameters such that the computational effort of running the AIP tool is in a reasonable relation to the potential benefit acquired in the field work.

With regard to the representation of the AIP in the proposed solution procedure, there are certain limitations related to the way that polygons are represented by discrete points, and how the route is only allowed to visit these points, instead of searching for the route in continuous space. This effectively means that the route is always constrained to the predetermined set of candidate points and the network of connectivity defined between these points. Nevertheless, the discretization scheme and the associated sacrifice in the alignment options of the route is necessary in order to devise a reasonable solution algorithm. The AIP does, however, take into account continuous space by determining the covered area by means of a buffer zone drawn around the route, instead of allocating discrete candidate points to the vertices of the route. This procedure allows the discovery of better and more accurate solutions compared to the conventional MCTP.

The AIP seeks to solve a routing problem inside a single polygon, which, of course, constitutes only a part of the larger inventory route planning problem. The complete inventory route typically visits a number of targets (forest stands) in a certain order. The implemented AIP tool was designed in such a way that it can be integrated as part of a GIS workflow used to determine inventory routes. In the case of forest inventory planning, the AIP tool will calculate routes inside individual forest stands. When those routes are merged with the optimal routes calculated between forest stands (e.g., using the procedure proposed by Store & Antikainen 2010), complete routes can be produced for actual forest inventory purposes. This involves accounting for several factors, including topography, vegetation and the traversability of the terrain (Etula & Antikainen 2012).

There are many practical aspects of routing that call for further development of the AIP tool. For example, when calculating complete inventory routes (visiting a number of areas), it is possible that the areas to be inventoried are located immediately adjacent to each other. In such a situation, it might be useful to assess the coverage of the route simultaneously for both areas, instead of assessing the coverage separately. Particularly in a situation where a small area is located next to a larger area, it is possible that the smaller area will actually be entirely visible from a route calculated for its large neighbor, without a need to calculate a route for the small area at all. In addi-

tion, depending on the location and shape of the areas, the most optimal strategy might involve visiting certain areas several times. For example, the shortest route might be accomplished by inventorying part of a certain area first, then moving on to inventory a different area, and then returning to the first area to complete its inventory. The ultimate objective of inventory route optimization is, after all, to minimize the overall length of a complete route, and this requires that the mutual position of the inventory targets is taken into consideration. This aspect also needs to be accounted for when selecting the entrance and exit points of consecutive targets.

Although the AIP algorithm is formulated for the purposes of forest inventory, it is not limited to this particular application, as analogous applications can be seen in other fields as well. The value of the visibility parameter can be determined by using viewshed analysis to make the method suitable for different landscapes, and it is even possible to replace the visibility parameter with any other factor representing distance or accessibility. Polygons of any shape and size can be used in the method as well, as the performance of the method can be guaranteed by adjusting the density of the candidate points to match the scale. In addition, there is no need for the polygons to be internally homogeneous, as the method can be extended to calculate the internal network of connectivity using the cost surface technique. This allows the variable traversability conditions typical to larger geographical areas to be taken into account, generating even more application possibilities.

REFERENCES

Alexandris G & Giannikos I 2010. A new model for maximal coverage exploiting GIS capabilities. *European Journal of Operational Research* 202: 2, 328–338. http://dx.doi.org/10.1016/j.ejor.2009.05.037.

Anderson DG 2012. Least cost pathway analysis in archaeological research. In White DA & Surface-Evans SL (eds). *Least Cost Analysis of Social Landscapes*, 239–257. The University of Utah Press, Salt Lake City.

Applegate DL, Bixby RE, Chvatal V & Cook WJ 2006. *The traveling salesman problem. A computational study*. Princeton University Press, Princeton.

Bagli S, Geneletti D & Orsi F 2011. Routeing of power lines through least-cost path analysis and multicriteria evaluation to minimise environmental impacts. *Environmental Impact Assessment Review* 31: 3, 234–239. http://dx.doi.org/10.1016/j.eiar.2010.10.003.

Balstrøm T 2002. On identifying the most time-saving walking route in a trackless mountainous terrain. *Geografisk Tidsskrift* 102: 1, 51–58.

Bolstad P 2002. *GIS Fundamentals*. Eider Press, White Bear Lake, Minnesota.

Church RL & ReVelle CS 1974. The maximal covering location problem. *Papers in regional science* 32: 1, 101–118.
http://dx.doi.org/10.1007/BF01942293.

Church RL & Sorensen P 1994. Integrating normative location models into GIS: Problems and prospects with the p-median model. National Center for Geographic Information and Analysis. *Technical report 94–5*.

Collischonn W & Pilar JV 2000. A direction dependent least-cost-path algorithm for roads and canals. *International Journal of Geographical Information Science* 14: 4, 397–406.
http://dx.doi.org/10.1080/13658810050024304.

Current JR & Schilling DA 1989. The covering salesman problem. *Transportation Science* 25: 3, 208–213.
http://dx.doi.org/10.1287/trsc.23.3.208.

Current JR & Schilling DA 1994. The median tour and maximal covering tour problems: formulations and heuristics. *European Journal of Operational Research* 73: 1, 114–126.
http://dx.doi.org/10.1016/0377-2217(94)90149-X.

Diestel R 2000. *Graph Theory*. 2nd ed. Springer, New York.

Dijkstra EW 1959. A note on two problems in connexion with graphs. *Numerische Mathematik* 1: 1, 269–271.

Etula H & Antikainen H 2012. Maaston kulkukelpoisuuden mallintaminen metsäsuunnittelijan näkökulmasta reitinoptimointia varten. *Terra* 124: 1, 29–43.

Feldman SC, Pelletier RE, Walser E, Smoot JC & Ahl D 1995. A prototype for pipeline routing using remotely sensed data and geographic information system analysis. *Remote Sensing of Environment* 53: 2, 123–131.
http://dx.doi.org/10.1016/0034-4257(95)00047-5.

Gendreau M, Laporte G & Semet F 1997. The covering tour problem. *Operations Research* 45: 4, 568–576.
http://dx.doi.org/10.1287/opre.45.4.568.

Golden B, Naji-Azimi Z, Raghavan S, Salari M & Toth P 2012. The generalized covering salesman problem. *INFORMS Journal on Computing* 24: 4, 534–553. http://dx.doi.org/10.1287/ijoc.1110.0480.

Hakimi S 1964. Optimum location of switching centers and the absolute centers and medians of a graph. *Operations Research* 12: 3, 450–459.
http://dx.doi.org/10.1287/opre.12.3.450.

Howey MCL 2007. Using multi-criteria cost surface analysis to explore past regional landscapes: a case study of ritual activity and social interaction in Michigan, AD 1200–1600. *Journal of Archaeological Science* 34: 11, 1830–1846. http://dx.doi.org/10.1016/j.jas.2007.01.002.

Longley PA, Goodchild M, Maguire DJ & Rhind DW 2005. *Geographic information systems and science*. 2nd ed. John Wiley & Sons Ltd, Chichester.

Lundqvist H 2007. Ecological cost-benefit modelling of herbivore habitat quality degradation due to range fragmentation. *Transactions in GIS* 11: 5, 745–63.
http://dx.doi.org/10.1111/j.1467-9671.2007.01070.x.

Maltamo M, Packalén P, Kallio E, Kangas J, Uuttera J & Heikkilä J 2011. *Airborne laser scanning based stand level management inventory in Finland*. Proceedings of SilviLaser 2011, 11th International Conference on LiDAR Applications for Assessing Forest Ecosystems, University of Tasmania, Australia, 16–20 October 2011, 1–10.

Miller H & Shaw SL 2001. *Geographic Information Systems for Transportation*. Oxford University Press, New York.

Murray AT 2005. Geography in coverage modeling: Exploiting spatial structure to address complementary partial service of areas. *Annals of the Association of American Geographers* 95: 4, 761–772.
http://dx.doi.org/10.1111/j.1467-8306.2005.00485.x.

Parks SA, McKelvey KS & Schwartz MK 2013. Effects of weighting schemes on the identification of wildlife corridors generated with least-cost methods. *Conservation Biology* 27: 1, 145–154. http://dx.doi.org/10.1111/j.1523-1739.2012.01929.x.

ReVelle CS & Toregas C 1972. Optimal location under time or distance constraints. *Papers of the Regional Science Association* 28: 1, 131–143.

Store R & Antikainen H 2010. Using GIS-based multicriteria evaluation and path optimization for effective forest field inventory. *Computers, Environment and Urban Systems* 34: 2, 153–161. http://dx.doi.org/10.1016/j.compenvurbsys.2009.12.003.

Theobald DM, Norman JB & Newman P 2010. Estimating visitor use of protected areas by modeling accessibility: A case study in Rocky Mountain National Park, Colorado. *Journal of Conservation Planning* 6, 1–20.

Tomczyk AM & Ewertowski M 2013. Planning of recreational trails in protected areas: Application of regression tree analysis and geographic information systems. *Applied Geography* 40, 129–139. http://dx.doi.org/10.1016/j.apgeog.2013.02.004.

Tomppo E, Olsson H, Ståhl G, Nilsson M, Hagner O & Katila M 2008. Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sensing of Environment* 112: 5, 1982–1999. http://dx.doi.org/10.1016/j.rse.2007.03.032.

Xiang WN 1996. A GIS based method for trail alignment planning. *Landscape and Urban Planning* 35: 1, 11–23. http://dx.doi.org/10.1016/0169-2046(96)00303-9.

Yu C, Lee J & Munro-Stasiuk MJ 2003. Extensions to least-cost path algorithms for roadway planning. *International Journal of Geographical Information Science* 17: 4, 361–376. http://dx.doi.org/10.1080/1365881031000072645.