Vehicle routing: less "artificial", more "intelligence"

J. W. Joubert
Department of Industrial & Systems Engineering,
University of Pretoria, South Africa

Abstract

The integration of multiple constraints of the Vehicle Routing Problem (VRP) variants is computationally expensive. Although vehicle routing problems have been well researched, variants are typically treated in isolation, whereas industry requires integrated solutions. Solution algorithms are also tested using benchmark data that are questionable, and that do not represent typical applications. The paper proposes an approach that solves a problem by analyzing its environment through cluster analysis, chooses an appropriate solution strategy, and tests the results in an attempt to learn for the purposes of improved future decisions.
Keywords: vehicle routing, VRP, heuristics, metaheuristics, artificial intelligence, learning.

1 Introduction

Vehicle routing and scheduling problems are well-researched in the field of Operations Research. The main objective of these types of problems are to minimize the distribution costs for individual carriers. Given the complexity of the type of problem, extensive research has been conducted to develop exact and heuristic solution techniques for urban distribution problems. The Vehicle Routing Problem (VRP) can be described as the problem of assigning optimal delivery or collection routes from a depot to a number of geographically distributed customers, subject to side constraints. In its basic form, the VRP can be defined with $G = (V, E)$ being a directed graph where $V = \{v_0, v_1, \ldots, v_n\}$ is a set of vertices representing customers, and with v_0 representing the depot where m identical vehicles, each with capacity Q, are located [1]. $E = \{(v_i, v_j)|v_i, v_j \in V, i \neq j\}$ is the edge set connecting the vertices. Each vertex i, except for the depot $(V \backslash v_o)$,

WIT Transactions on The Built Environment, Vol 89,© 2006 WIT Press
www.witpress.com, ISSN 1743-3509 (on-line)
doi:10.2495/UT060251

has a non-negative demand q_i and a non-negative service time s_i. A distance matrix $C = \{c_{ij}\}$ is defined on E. In some contexts, c_{ij} can be interpreted as travel cost or travel distance. Hence, the terms distance, travel cost, and travel time are often used interchangeably. The VRP consists of designing a set of m vehicle routes having a minimum total length such that

- each route starts and ends at the depot,
- each remaining vertex $(V \backslash v_o)$ is visited exactly once by one vehicle,
- the total demand of a route does not exceed Q, and
- the total duration (including service and travel time) of a route does not exceed a preset limit L

The VRP is a hard combinatorial optimization problem for which Laporte [2] has indicated several exact and approximate solution algorithms. An np-hard problem implies that the solution space will increase at an exponential or factorial rate (nonpolynomial) as the number of customers/vertices increases. Early researchers such as Clarke and Wright [3] realized that exact algorithms can only solve relatively small problems, but a number of heuristic (near-optimal) algorithms have proved very satisfactory. Laporte and Semet [4] present a number of classical heuristics, while Gendreau et al. [5] cover an array of metaheuristics, including Simulated Annealing, Genetic Algorithm, Granular Tabu Search, Ant Colony Optimization, and Swarm Intelligence.

2 A case for intelligence

Freight carriers are sharing the road network with various modes of public transport while the use of private vehicles have rapidly increased as well. Carriers are continuously expected to provide higher levels of service at lower rates, and therefor try to minimize their logistic costs, and maximize their profit. Sharing the road infrastructure with other vehicles such as private cars and public transport forces carriers to plan their freight routes more carefully. Enhanced vehicle routing and scheduling takes the congestion constraint into account and attempts to improve the vehicular utility through shorter routes and higher load factors. Software applications often do not provide adequate functionality by not being able to address complex business requirements such as companies having a fleet of vehicles that differ in capacity and/or running costs, and double scheduling where vehicles are allowed to complete a trip, return to the depot to renew it's capacity, i.e. offload goods collected, or loading goods to be delivered. The reason for software deficiencies are related to the extreme computational complexity when solving routing models. Human intervention is required to, for instance, split the fleet into vehicle categories that represent similar or the same capacity and/or costs. Each category is then solved independently, adjusting demand as customers are serviced by other categories. Human operators can also intervene by evaluating vehicular routes, and identifying vehicles that may be used for a second trip, and then schedule such vehicles accordingly. Although such interventions are mechanistic in nature, they require the time and effort of experienced individuals having a thorough understanding of vehicle routing so as to intervene wisely.

We refer to ourselves (in a more formal way) as *homo sapiens* — man the wise — and value our mental abilities to think and reason to assist us in improving our surroundings. We require our thought processes and intelligence to make decisions that will maximize the utility that we obtain from logistics — moving goods from points of manufacture to points of consumption that are geographically dispersed.

> "What is mind? What is the relationship between mind and the brain? What is thought? What are the mechanisms that give rise to imagination? What is perception and how is it related to the object perceived? What are emotions and why do we have them? What is will and how do we choose what we intend to do? How do we convert intentions into action? How do we plan and how do we know what to expect from the future?" — Albus [6]

It seems clear from the quote by Albus that before one toss terms such as thinking and planning around, one should carefully consider how such actions take place, and how one intends to employ such actions to improve, within the scope of this paper, urban freight congestion.

2.1 Intelligence

In their leading text, Russell and Norvig [7] introduce Artificial Intelligence (AI) as not only understanding the human intellect, but also building entities (or agents) that are intelligent. Although it encompasses a huge variety of subfields of study, with many varying definitions, the authors have categorized AI approaches in a two-dimensional framework represented in Figure 1.

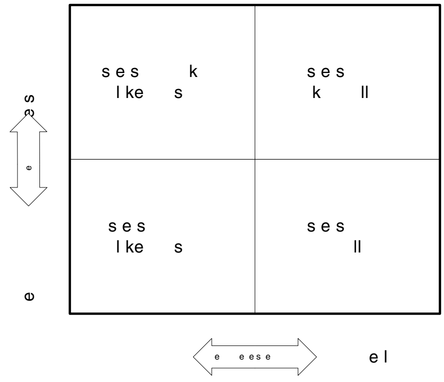


Figure 1: Categories of artificial intelligence. (Adapted from [7].)

The top half of the framework is concerned with thought processes and reasoning, as opposed to the lower half that is concerned with the behavioral element of intelligence. The left side of the framework measures the success of an agent's intelligence against the fidelity of human performance. The right half establishes an *ideal* concept of intelligence as a benchmark, referred to as *rationality*. This is analogous to effectiveness — *doing the right* things. However, the *right* within *rationality* is only relative to what is known at the time of the doing.

An *agent* is something that acts. This paper is concerned with the development of a computer agent that could intelligently intervene in the routing and scheduling of distribution vehicles. But how is it to be distinguished from mere programming? It should be able to operate autonomously, perceive the environment, persist over a period of time, and be able to adopt the goals and objectives of another entity. As an improvement on a basic agent, the author of this paper propose a rational agent that has a strategy to achieve the best possible outcome for a given objective, either known, or the expected outcome, should some of the parameters be uncertain. The focus of this paper is therefor not on understanding the human thought processes, but on creating a system that can think, and act rationally.

2.2 Complexity

Perfect rationality in modeling is often too difficult to attain due to too high computational demands when looking for exact solutions. Problems such as the routing and scheduling of vehicles can often not be solved exactly, and require the use of solution algorithms that provided approximate solutions where the optimality of the solution can neither be proved in advance, nor confirmed once a solution is found. The different opinions with regards to either finding an exact optimal solution versus settling for a *good enough* solution given a specific environment have led to the split that occurred between *Decision Theory* and AI in the latter half of the twentieth century.

Decision Theory is the field of study where probability theory and utility theory are combined to present a formal framework for decision making under uncertainty. The field of operations research addresses complex management decisions rationally. The intention of the pure branch of decision theory is to obtain a rational decision, or a *global* optimum. On the contrary, the complexity in finding a single optimum value led the pioneers of AI such as Herbert Simon (1916-2001) to prove that being able to find a *good enough* answer describes human behavior more accurately — and earned him the Nobel prize in economics in 1978. And although the computational ability of computers have increased dramatically over the past decade, the intention is still to assist mere mortal logistics decision makers to improve their ability to manage distribution fleets.

3 An intelligent vehicle routing agent

The primary research question that should be answered is *whether it is feasible to develop a rational and intelligent agent to schedule a predefined variant of the VRP*. In order to answer the question, a number of secondary questions are stated in terms of the concept of an intelligent agent.

In his paper on the engineering of mind, Albus [6] identifies four functional elements of an intelligent system.

Sensory perception — accepting input data from both outside and from within the system. The data is then transformed through classification and clustering into meaningful representations of the real world. The first secondary research question addresses the analysis of input data and is stated as follows:

> *How should customer parameters be clustered so that meaningful classification can be done prior to executing the solution process?*

Behavior generation — planning and controlling actions so that goals are achieved. An intelligent agent accepts task with goals, objects and priorities. The tasks are then broken up into jobs and, along with resources, are assigned to agents. Hypothetical plans are created and simulated to predict the outcome of the plans. The simulated results are evaluated, and the agent selects the best expected hypothesized plan. In terms of this paper an agent refers to computational elements that plan and control the execution of a routing algorithm, correcting for errors and perturbations along the way. The planning processes of the agent are heuristics and metaheuristics that attempt to converge to optimal vehicle routes and schedules. This lead to another secondary research question:

> *How can heuristics and metaheuristics be used to establish vehicle routes and schedules in a complex and constrained environment?*

Value judgement — the computation of a predefined set of costs, risks, benefits, and or penalties related to the vehicle routes. In operations research terms these computational expressions are referred to as the objective function(s). The third secondary research question is derived from value judgement:

> *What should constitute the objective function of the model so that the real problem is adequately represented?*

World modeling — an overall strategy that uses input parameters and variables to update a knowledge database. Data is used to query the behavior generation of plans regarding current routes and schedules. The strategy further simulates possible results of future plans after analyzing the current plans. Simulated results are evaluated, using the value judgement, so the best expected plan for execution can be selected. After execution, the strategy allows for sensory expectations to be created regarding future actions — analogous to bumping your feet against an obstacle in the dark. After stumbling, and reacting to the pain, you lift your feet unnaturally high so as to avoid the next obstacle. The fourth and fifth, probably the most challenging secondary research questions addresses the agents ability to learn from the past and improve in future:

> *What critical parameters influence the agent's learning, and should therefor be included in creating future expectations? How are future expectations created from the past performance?*

4 Unsupervised clustering

The idea behind learning is not so that an agent can act, but rather to improve an agent's ability to act in future. I the context of vehicle routing the *agent* is a routing system. The acting is the routing of vehicles, given the demand inputs, using some metaheuristic with its associated parameter settings. For a routing system to learn, it must perceive certain characteristics of the inputs, for example the geographical dispersion of customers or the width of time windows provided by customers, and choose an appropriate metaheuristic, and know what parameter values to suggest in order to obtain the best route in the shortest possible time. The execution of the metaheuristic makes up the *performance* element of the agent. Deciding which metaheuristic to use forms the *learning* element of the agent.

The concepts of *representation* of an agents knowledge and its reasoning processes that brings that knowledge to life are central to the entire field of AI. In this paper the concept of pattern identification on input data is investigated. The design of a learning element is affected by three distinctive components:

- Which components of the performance element are to be learned?
- What feedback is available to learn these components?
- What representation is used for the components?

What is peculiar about the benchmark problem sets proposed by both Solomon [8] and Homberger and Gehring [9] are the fact that they are preempting specific theoretical characteristics, unlike problems found in real applications. This is clearly illustrated when the assignment of time windows is discussed. For the problem sets $R1$, $R2$, $RC1$, and $RC2$ a percentage of customers are selected to receive time windows, say $0 < f \leq 1$. Next n random numbers from the random uniform distributions are generated on the interval $(0,1)$, and sorted. Customers $i_l, i_2, \ldots, i_n$ are then assigned time windows, where the number of customers requiring time windows is approximated by $n_l \approx f.n$. The center of the time window for customer $i_j \in \{i_1, i_2, \ldots, i_n\}$ is a uniformly distributed, randomly generated number on the interval $\left(e_0 + t_{0i_j}, l_0 - t_{i_j0}\right)$, where e_0 and l_0 denotes the opening and closing times of the depot, respectively, and t_{0i_j} and t_{i_j0} denotes the travel distance from the depot to customer i_j, and back, respectively.

For clustered problem sets $C1$ and $C2$ the process becomes questionable. Customers in each cluster are first routed using a $3-opt$ routine. An orientation is chosen for the route, and time windows are then assigned with the center being the arrival time at the customer. The width and density are derived in a similar fashion as for random and semi-clustered data. Although Solomon [8] states that "this approach permits the identification of a very good, possibly optimal, cluster-by-cluster solution which, in turn, provides an additional means of evaluating heuristic performance", it does not provide a credible means to evaluate real life problems where customers do not negotiate their sequence prior to stating a preferred time window.

Literature provides good references to what type of metaheuristics, or metaheuristic configurations provide good answers to which of the six benchmark problems. When given a real data set from industry, however, one is not provided with the classification of "this a $C1$ problem set". To therefore determine which solution algorithm to use, and which parameter configuration, the routing agent first needs to classify the input data. The components of the performance element that the agent should learn from input data provided are the geographical distribution of customers; the relation between customer demand and vehicle capacity, and time window characteristics. In order to determine the nature of learning for the agent, the type of feedback available to the agent is extremely important. Russell and Norvig [7] distinguishes between three types of feedback:

Supervised learning Learning takes place by providing both input and output examples. For instance, if an agent is provided with many pictures that he is told contain buses, the agent learns to recognize a bus. Both the input and the output is provided.

Unsupervised learning Patterns are learned by providing input, but in the absence of specific outputs. When commuting from home to work, a person might be able to distinguish between "good traffic days" and "bad traffic days", without ever being given examples of either of the two. A purely unsupervised agent cannot learn as it has no information as to what constitute a desirable state, or a correct action.

Reinforcement learning The most general of the three types of feedback. Without being told by a supervisor what to do, a reinforcement learning agent must learn through reinforcement, for example an action that is not followed by a tip or any confirmation is interpreted as an undesirable state.

The routing agent is typically given a data set without knowing whether it is clustered, randomly distributed, or whether the time windows are tight. As a supervisor also do not know whether it is clustered, or not, it would also not be possible to reinforce a correct action taken, as the evaluation of correctness would be flawed. The routing agent would hence have to learn unsupervised. Knowledge and reasoning are both required for problem solving agents to perform well in complex environments. The concept of knowledge representation is important as an agent would require some structure in which to put the information that it has learnt, so as to be able to revisit its knowledge base in future when decision are made. All in an attempt to improve future decision making. The central component of a knowledge-based agent is its knowledge base, expressed as sentences in a knowledge representation language. Each sentence asserts something about the agent's world. There are ways to add new sentences to the knowledge base, and ways to query what is already known. In AI these two actions are standardly referred to as Tell and Ask. Being a logical agent, when 'Ask'ed a question, the answer would be related to what the knowledge base has been 'Tell'ed previously. Also, the two tasks may involve inference where new sentences are derived from old ones.

The clustering problem is defined as partitioning a given data set into groups, or clusters, such that data points in a cluster are more similar to each other than the other points in different clusters. According to Gath and Geva [10] and Xie and Beni [11] the criteria for the definition of optimal partition of the data into subgroups is based on three requirements:

- Clear separation between the resulting clusters.
- Minimal volume of the clusters.
- Maximal number of data points concentrated in the vicinity of the cluster centroid, i.e. maximum cohesion.

Thus, although the environment is fuzzy, the aim of the classification is generation of well-defined subgroups. To solve the clustering problem, a number of clustering algorithms have been proposed. One of the most important families of clustering techniques are partitioning clustering, with the most commonly used algorithm in this family being the k-means clustering algorithm and its numerous variants [12]. A main problem of the k-means clustering variants is that the algorithms require the number of clusters, c, as an input so that a data set can be clustered into c partitions. Unsupervised clustering is the problem of discerning multiple categories in a collection of objects. The categories referred to are the components of the input data that the agent should learn, while objects refer to the input data points, i.e. the customers in the network. The learning process is unsupervised as the agent does not know whether the input data is randomly distributed, clustered, or a combination of both. So if the number of clusters, c, is not known when learning should occur, the agent can perform a number of clustering attempts, each using a different values for c. In such a way the most appropriate value for c can be determined. Such an approach is defined as cluster validation. In this chapter, the behavior of a number of validation indices will be tested on benchmark data sets for the Vehicle Routing Problem with Time Windows (VRPTW). The objective is to establish trends that can be used to Tell the routing agent how to identify input data as belonging to either the Rl, $R2$, Cl, $C2$, RCl, or $RC2$ group of problems. The most appropriate metaheuristic can then me identified, along with its most appropriate parameter settings.

4.1 Fuzzy *c*-means clustering

One of the variants of the k-means clustering algorithm, fuzzy c-means (FCM) clustering, attempts to find the most characteristic point in each cluster $v_i \in \mathbf{V} = \{v_1, \ldots, v_c\}$, which is considered the *center* of cluster i and then grade the membership for each node $x_j \in \mathbf{X} = \{x_l, \ldots, x_n\}$ in cluster i. The member allocation is achieved by minimizing the commonly used *membership weighted* *with-in cluster error* objective function defined in (1).

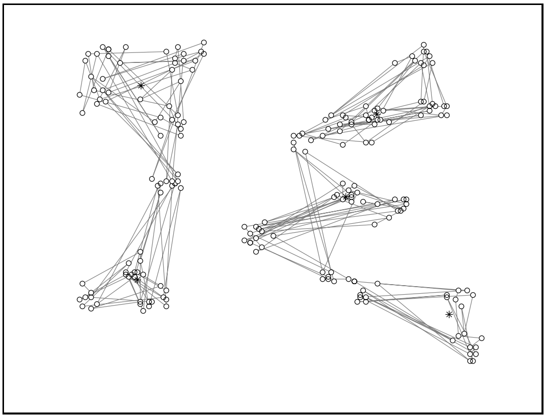
$$J_e(\mathbf{U}, \mathbf{V}) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m d_{ij}^2 \qquad (1)$$

where d_{ij} is the Euclidean distance between object j and the i^{th} center, and u_{ij} is the fuzzy membership of object j belonging to the i^{th} cluster. The FCM requires the number of classes, a fuzzy factor m and a convergence threshold as input. The centers matrix $\mathbf{V}$ is initialized using a random selection of c nodes from the node set $\{i, \ldots, n\}$. The iteration count is zeroed before the membership matrix $\mathbf{U}^k$ is calculated. A new centers matrix is calculated, before the convergence of the objective function is tested. Xu and Brereton [12] notes that when the fuzzy factor m approaches 1, the FCM is similar to the standard k-means clustering. When m approaches infinity, however, the clustering of the FCM is at its fuzziest: each node is assigned equally to each cluster. The authors also note that the FCM is but a local search algorithm, and at best will find a local minimum, and is therefore sensitive for the random initial guess for v_0. Figure 2 illustrates the clustering of one of the Cl problem sets provided by Gehring and Homberger [13], $C1\_2\_1$, the first of their problem sets with 200 customers.

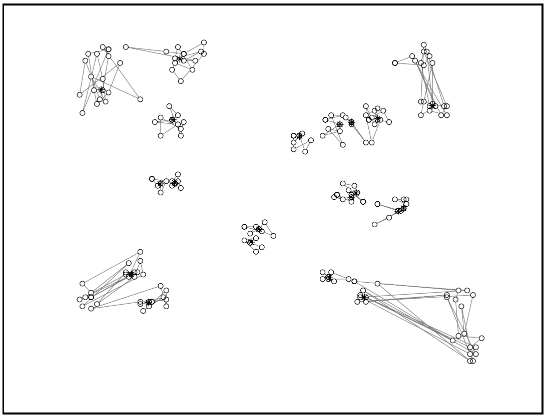
The small circles indicate the customer nodes, while asterisks indicate the center of the cluster. All nodes clustered together are linked with gray lines. In establishing the clusters, a fuzzy factor of $m = 3$, convergence threshold of $c = 1.0 \times 10^5$, and an iteration limit of $k^{\max} = 1000$ is used. A number of validation indices are subsequently considered to evaluate the clustering.

4.2 Validation indices

is a single real value that describes the quality of a cluster partition. Some of the validation indices are only concerned with the membership value of the final clustering partition. The Partition Coefficient, V_{PC}, and the Partition Entropy, V_{PE}, have been introduced by Bezdek [14, 15]. The disadvantages of these two indices are the lack of direct connection to a geometrical property, and the monotonic decreasing tendency with c. The next indices involve not only the membership value, by also the actual data set. Xie and Beni [11] introduced an index that give weight to both compactness, and separation. First the fuzzy deviation of node j from cluster i, denoted by d_{ij} is determined as the Euclidean distance between node j and cluster i, weighted by the fuzzy membership of node j belonging to cluster i. The sum of the squares of the fuzzy deviations of each node j is referred to as the variance of cluster i, denoted by σ_i. The total variation of the data set with respect to the given fuzzy c-partition is referred to as σ. The compactness of the partition is the ratio between the total variation of the data set to the size of the data set, expressed as $\frac{\sigma}{n}$. The centers between all cluster center combinations $i, r \in \{1, \ldots, c\}, i \neq r$ is calculated, and the minimum inter-center distance is denoted by $d_{\min}$. The separation of clusters is then determined by $s = d_{\min}^2$. A high value of s indicates well-separated clusters. The index becomes finding the minimum value for $\frac{\sigma}{n \cdot s}$.



(a) 5 Clusters



(b) 20 Clusters

Figure 2: Clustering the C1-2-1 problem set.

Pal and Bezdek [16] extend the Xie-Beni index for cases where the fuzzy factor $m \neq 2$, denoted by V_{XB}^{+}. Kwon [17] also investigates the Xie-Beni index, and proposes and index, V_K, that eliminates the monotonically decreasing tendency as the number of clusters increases and approaches n, the number of nodes in the data set. The index is defined in (2).

$$V_K = \frac{\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{m} \left\| x_j - v_i \right\|^2 + \frac{1}{c}\sum_{i=1}^{c} \left\| v_i - \overline{v} \right\|}{n\left(\min_{i,r\in\{1,\ldots,c\}, i\neq r}\left\{\left\| v_i - v_r \right\|^2\right\}\right)} \tag{2}$$

The second term in the numerator is an *ad hoc* punishing function used to eliminate the decreasing tendency when c becomes large and close to n.

The *Compose Within and Between Scattering* index was introduced by Rezaee *et al.* [18] and is defined by (3).

$$V_{CWB} = \alpha Scat(c) + Dis(c) \tag{3}$$

The interested reader can refer to [18] for a detailed formulation of each component. The V_{CWB} index tends to find an optimum between the compactness and separation. $Scat(c)$ denotes the average scattering (compactness) for the c clusters, while $Dis(c)$ denotes the distance between cluster centers (separation). With $Scat(c)$ taking on much smaller values than $Dis(c)$, a scaling factor α is introduced to balance the two terms' opposite trends. The authors perform the validation over cluster partitions with values $2 \leq c \leq c_{\max}$. In an application where a cluster is considered to be more than p nodes, a $c_{\max}$ of $\frac{n}{p}$ is used.

5 Conclusion

In this paper the author challenges current approaches in the development of vehicle routing optimization algorithms. With the introduction of intelligence into routing, an algorithm is required not only to solve pre-classified problem sets, but also be able to *classify* an unknown problem into a specific set. To achieve this characteristic of an intelligent agent, the author proposes fuzzy c-means cluster validation. Concepts are merely introduced in the paper, while cluster validation results are included in the conference presentation. The classification of results is used in an extended research project that investigates the *learning* of an algorithm over a period of time.

Future research could include an investigation into different clustering configurations that include time-window characteristics. Such a research approach will typically have to challenge the heart of the problem sets as proposed by earlier research contributions.

References

[1] Van Breedam, A., Comparing descent heuristics and metaheuristics for the vehicle routing problem. *Computers and Operations Research*, **28(4)**, pp. 289–315, 2001.

[2] Laporte, G., The vehicle routing problem: An overview of exact and approximate algorithms. *European Journal of Operational Research*, **59(3)**, pp. 345–358, 1992.

[3] Clarke, G. & Wright, J., Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research*, **12**, pp. 568–581, 1964.

[4] Laporte, G. & Semet, F., Classical heuristics for the capacitated VRP. *The Vehicle Routing Problem*, eds. P. Toth & D. Vigo, Society for Industrial and

Applied Mathematics (SIAM): Philadelphia, SIAM Monographs on Discrete Mathematics and Applications, chapter 5, pp. 109–128, 2002.

[5] Gendreau, M., Laporte, G. & Potvin, J.Y., Metaheuristics for the capacitated VRP. *The Vehicle Routing Problem*, eds. P. Toth & D. Vigo, Society for Industrial and Applied Mathematics (SIAM): Philadelphia, SIAM Monographs on Discrete Mathematics and Applications, chapter 6, pp. 129–154, 2002.

[6] Albus, J.S., The engineering of mind. *Information Sciences*, **117(1)**, pp. 1–18, 1999.

[7] Russell, S.J. & Norvig, P., *Artificial Intelligence: A Modern Approach*. Prentice Hall Series in Artificial Intelligence, Prentice Hall: Upper Saddle River, New Jersey, 2nd edition, 2003.

[8] Solomon, M., Algorithms for the vehicle routing and scheduling problems with time windows. *Operations Research*, **35(2)**, pp. 254–265, 1987.

[9] Homberger, J. & Gehring, H., Two evolutionary metaheuristics for the vehicle routing problem with time windows. *INFOR*, **37(3)**, pp. 297–318, 1999.

[10] Gath, I. & Geva, A.B., Unsupervised optimal fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11(7)**, pp. 773–781, 1989.

[11] Xie, X.L. & Beni, G., A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **13(8)**, pp. 841–847, 1991.

[12] Xu, Y. & Brereton, R.G., A comparative study of cluster validation indices applied to genotyping data. *Chemometrics and intelligent laboratory systems*, **78(1–2)**, pp. 30–40, 2005.

[13] Gehring, H. & Homberger, J., A paralel hybrid evolutionary metaheuristic for the vehicle routing problem with time windows. *Proceedings of EUROGEN99 — Short Course on Evolutionary Algorithms in Engineering and Computer Science*, eds. K. Miettinen & M.M. Mäkelä, Finland, volume No. A 2/1999 of *Reports of the Department of Mathematical Information Technology*, pp. 57–64, 1999.

[14] Bezdek, J.C., Numerical taxonomy with fuzzy sets. *Journal of Mathematical Biology*, **1**, pp. 57–71, 1974.

[15] Bezdek, J.C., Cluster validity with fuzzy sets. *Journal of Cybernetics*, **3(3)**, pp. 58–73, 1974.

[16] Pal, N.R. & Bezdek, J.C., On cluster validity for the fuzzy c-means model. *IEEE Transactions on Fuzzy Systems*, **3(3)**, pp. 370–379, 1995.

[17] Kwon, S., Cluster validity index for fuzzy clustering. *Electronic Letters*, **34(22)**, pp. 2176–2177, 1998.

[18] Rezaee, M.R., Lelieveldt, B.P.F. & Reiber, J.H.C., A new cluster validity index for the fuzzy c-mean. *Pattern Recognition Letters*, **19(3–4)**, pp. 237–246, 1998.