Presentation of Location Problems

Alexander Tsenov¹

Abstract: The paper addresses the set of possible presentations of different types of location – allocation problems. This is an important stage of the whole network planning process, consisting the definition of the network type, the main network properties and the tools used to find an optimal solution for the network topology.

Depending of a network type it is important to define an optimal presentation of the network design problem. Each one presentation must correspond to a known optimization algorithm or a new algorithm must be developed.

The paper consists of a survey of known presentations and algorithms and a new presentation of location allocation problem named "set based presentation". The idea for developing the new presentation is to apprise the usability of Genetic Algorithms for solving network design problems.

Keywords: Network design, Optimization, Genetic Algorithms

I. INTRODUCTION

The problem of optimally designing a network in order to meet a given set of specifications (such as prescribed traffic requirements, achieving a desired level of reliability, respecting a given maximum transit time), while minimizing total cost, arises in a wide variety of contexts: computer networks, telecommunications networks, transportation networks, distribution systems.

Network design algorithms draw an increasing amount of attention nowadays. Considering the complexity, high cost factor and fast deployment times of today's communications systems (such as IP and ATM backbones, optical networks, numerous types of access structures etc.), network operators can benefit a lot from the use of network design tools. These tools can help speeding up and 'automating' the design process, ensuring superior quality (i.e. lower cost and/or better Quality of Service) and more justifiable solutions. Network design tools typically incorporate a wide range of functionality, such a geographical database handling, traffic estimation, link dimensioning, cost calculation, equipment configuration databases etc. The real benefit of using these tools, however, comes from the possibility of using the algorithmic network optimization approaches. In this way, there arises a possibility for finding solutions of better quality in much shorter time, as compared to the manual network design.

The topology optimization problems can be characterized as follows:

Given:

- a set of access nodes with given geographical locations;
- traffic demand between each node pair.

Objective:

- find the number and locations of the nodes;
- find links connecting the nodes;

¹Alexander Tsenov is with Telecom Department at Technical University of Sofia, "Kliment Ohridsky Blvd 8, 1756 Sofia, Bulgaria, E-mail: <u>akz@vmei.acad.bg</u> This paper introduces representation strategies that are appropriate to genetic algorithms for access network planning and applies them to some simple network optimization problems. The first section describes a number of genetic representations in the context of a general location-allocation problem. These representations are presented along with their appropriate operators. Also a novel set-based representation and operator is developed. The idea is to find most applicable presentation of a location – allocation problems witch can be easily manipulated by the basic genetic operators such a crossover and mutation. The required representation has to help the network planer to find the most optimal solution of planing and optimizing problems concerning the network topology.

II. LOCATION-ALLOCATION PROBLEMS

The general problem involves the allocation of customers to a number of (supplier) sites. They can be broadly divided into two types: site generation problems and site selection problems [1]. Site generation problems require that the optimization choose a location for the sites from a continuous space. For site selection problems, the sites are chosen from a finite set of candidate locations. It is the second of these groups of problems that is concentrated on here.

The objective function for a solution to a locationallocation problem may be defined in many ways but is usually the total cost, distance or time for supplying all the customers, or some combination of these three. Although for some applications such as the siting of a fire station, it may be minimization of the maximum distance between location and customer.



Fig. 1: Example location-allocation problem; there are four possible sites (A-D) to which the 10 customers (a-j) must be allocated. A line connecting a site and customer shows that this site supplies the customer.

An example problem is illustrated in Figure 1, where there are four possible sites and ten customers to be supplied. An example solution to the problem is illustrated; sites A, C and D have been chosen to supply the ten customers.

The following section describes representations of location – allocation problems, oriented to application of genetic algorithms for optimizing these types of problems.

III. REPRESENTATIONS

There are many possible representations available but it is important to develop the one that is both suited to the problem and that works efficiently with a crossover operator.

Preferably, the representation and operators can be used unchanged for harder problems. For example, it is desirable that a general-purpose representation and optimizer be found that can be used on a multi-period access network-planning problem. That is, the aim is to produce a general algorithm that covers this class of problems.

An ideal representation for a genetic algorithm has a number of features:

• There should be no redundancy – each point in the problem search space should map to a unique genetic representation [2]. Redundancy in the representation increases the size of the genetic search space and means that many different genomes representing a single point in the problem space will have the same fitness.

• It should avoid representing parts of the problem space that break constraints. Clearly if a constraint programming problem is being considered this is impractical for all of the constraints.

Simple Strategy

A simple representation for a location-allocation problem is a bit string of length equal to the number of candidate sites. If a bit is equal to one then the corresponding location is present



Fig. 2: Simple genetic representation for a location-allocation problem. For each site a gene represents whether that site is in the solution using binary valued alleles. Each customer is then assigned to the site, which can supply it most cheaply. The values in the final line show how the example in Figure 1 would be represented, that is a solution where sites A, C and D are present.

in the solution, if zero it is not. The customers can then simply be assigned to the nearest (or lowest cost connection) site. Figure 2 illustrates this representation and shows the possible allele values for each gene and the allele values necessary to represent the solution shown in Figure 1. This representation allows standard bit-string crossover operators to be used, such as uniform crossover.

This approach is used by Kratica [4][5] for the warehouse location problem and by Chardaire et al. [3] for optimizing the topology of concentrator networks. This representation is discussed no further due to its inflexibility for representing harder capacity constrained and dynamic problems.

Customer Specifies Location

A more general representation for the problem is one where for each customer there is a set of alleles which represent all the possible sites that can supply the customer, the actual allele chosen represents the site which supplies the customer. This representation was used by Routen [3] for a concentratorlocation problem. It is good as there is no redundancy – one genome maps to a unique instance of a location - allocation.

A diagram illustrating this representation is shown in Figure 3.



Fig. 3: A genetic representation where each customer has a set of alleles that represent which sites could supply the customer. The values in the final line show how the example in Figure 1 would be represented.

This representation allows manipulation of the genome using standard crossover operators such as n-point crossover and uniform crossover. Uniform crossover is the most appropriate, as there is no information contained in the ordering or position of the genes within the genome.

IV. SET-BASED REPRESENTATION

A natural representation for a location-allocation problem is one in which sets (or clusters) of customers are formed along with an associated location. The set of customers is allocated to this associated location. The objective of the optimization is then to form good clusters of customers and to find good locations as centers for these clusters. The problem can be decomposed so that first a cluster of customers is found and then a location is selected on which to center them.

The problem with representing sets as lists is that it is inefficient. A list representing a set contains much redundant information. For example if customer A and customer B share a set then it is irrelevant whether customer A comes before or after customer B in the list. This redundancy causes problems for the GA as it means that there are many points in the search space with equal fitness values.

The above problems led to the development of a novel representation for sets. The new representation is described using forma theory developed by Radcliffe [7].

Starting from the idea that customers sharing a node belong to the same set, the requirement for the optimization is that it optimizes the set membership of all the customers. This promotes the idea of finding good clusters of customers to the fore and makes finding good locations a secondary aim.

This is because, as will be described later, there are good techniques for finding an optimal location given a cluster of customers. Using the terminology of forma theory, the equivalence relation used is therefore Ψ_{ab} = "customer *a* shares a set with customer *b*"

So, if there are three customers a, b and c, the following equivalence classes (or forma) can be induced:

$$\Xi_{\Psi_{ac}} = \left\{ \xi_{ac}, \xi_{\overline{ac}} \right\}$$
(1)

$$\Xi_{\Psi_{ab}} = \left\{ \xi_{ab}, \xi_{\overline{ab}} \right\}$$
(2)

$$\Xi_{\Psi_{bc}} = \left\{ \xi_{bc}, \xi_{\overline{bc}} \right\}$$
(3)

where for example:

 $\{\xi_{ab}\}$ means that *a* shares a set with *b*, and

 $\left\{\xi_{\frac{ab}{ab}}\right\}$ means that *a* does not share a set with *b*.

A basis for customers a, b and c which is capable of representing all possible sets of customers can be obtained using the following equivalence relations:

$$\mathbf{E} = \left\{ \Psi_{ab}, \Psi_{ac}, \Psi_{bc} \right\}$$
(4)

Given E, a genetic representation can be formed by

· using the equivalence relations as genes

 \cdot and the set of formal that the equivalence relationship induces as the alleles.

Implementation

The above representation can be implemented as a bit string but is more logically represented as an array. The representation that is used to implement the forma is actually the upper half of an array. Each position in the array *A* is set to a binary value so that:

if A(x, y) = 1, then
$$\xi_{xy}$$

if A(x, y) = 0, then
$$\xi_{\frac{1}{y}}$$

As $\xi_{ab} \equiv \xi_{ba}$, also A(x,y) \equiv A(y,x).. Therefore only half an array is needed; arbitrarily the upper half is used. Figure 4 demonstrates the process of populating an array and the dependencies that exist between values. Figure 4 also shows the many dependencies between values in the array (as the forma are not orthogonal). Despite these dependencies, it was considered that the representation provided a general-purpose approach to modeling the clustering of customers around primary and secondary nodes. The representation is such that the application of a traditional operator is difficult; therefore, a new crossover operator is to be developed.

So far, this section has only described the method for representing the partitioning of a set of customers; no mention has been made of representing the associated site. The next section describes how this site can be represented. *Representing the Associated Site*

A simple method has been chosen for representing the site that is associated with each set of customers. The association



Fig. 4: The set based representation uses a half array for its implementation. This diagram illustrates the population of an array and the dependencies that exist between values. The values on the left are specified and the values that result from these specified values are shown on the right. In this example the customers a, b, c, d, e are partitioned into the two sets illustrated in the bottom right of the figure.

is made through the customers; each customer has an associated gene that represents their preferred supplying site. In fact, this is the same as the representation introduced on Figure 3. However, when a particular location-allocation is instantiated, the first customer to be assigned to a set has its associated location used to supply the whole set of customers. Once all the sets have been instantiated then the genome representing the customers' associated sites is updated to reflect the current configuration.

For example, if the set genome specifies the following groupings of six customers:

$$\{\{1, 2\}, \{3, 4, 5\}, \{6\}\}$$

and the customers are associated with the following sites (Table I).

TABLE I INITIAL ALLOCATION

Customer	1	2	3	4	5	6
Supplying Site	С	D	В	Е	А	А

Then, when the genome is converted to a locationallocation instance, $\{1, 2\}$ are supplied by site C, and $\{3, 4, 5\}$ are supplied by site B, and $\{6\}$ by site A. This information is then passed back to the customer genome, and the resulting genome is (Table II).

TABLE II END ALLOCATION

Customer	1	2	3	4	5	6
Supplying Site	С	С	В	В	В	A

An alternative to the above strategy would be to search through all the sites associated with a particular set to find the best of the collection. For large problems or computationally intensive objective functions this approach would be very inefficient and has therefore not been pursued in the current work.

V. CONCLUSION

In this work two different types of location-allocation problems were considered. The benefits and the disadvantages of these representations are discussed. A new representation named set-based representation is introduced. The new representation is strongly oriented to a set of genetic operators such a crossover and mutation. The operation of these genetic operators is being still tested over many test problems including known location problems. The firs results are acceptable and demonstrate the applicability of the propounded representation.

REFERENCES

- Minoux, M.: Network Synthesis and Optimum Network Design Problems: Models, Solutions Methods and Applications, *Networks*, vol. 19, 1989, pp. 313 – 360
- [2] Hoang, H.H.:A Computational Approach to the selection of an Optimal Network, *Management Science*, vol. 19, 1973, pp. 488 - 498
- [3] Routen T., Genetic Algorithms and Neural Network Approaches to Local Access Network Design, Proceedings of the 2nd International Workshop on Modeling, Analysis and Simulation of Computer and Telecommunications Systems, 1994
- [4] Kratica J., V. Filipovitch, V. Ljeljum, D. Toljijg Solving of the Uncapacitated Warehouse Location Problem Using a Simple Genetic Algorithm, *Proceedings of the XIV International Conference on Material handling and warehousing*, 3.33-3.37, Belgrade, 1996
- [5] Kratica J., Improvement of Simple Genetic Algorithm for Solving the Uncapacitated Warehouse Location Problem, Proceedings of the 3rd On-line World Conference on Soft Computing (WSC3), 1998
- [6] P. Chardaire, A. Kapsalis, J.W. Mann, V.J. Rayward-Smith and G.D. Smith, Applications of Genetic Algorithms in Telecommunications. In J. Alspector, R. Goodman, T.X. Brown (Eds.), *Proceedings of the 2nd International Workshop on Applications of Neural Networks to Telecommunications*, 1995, pp. 290-299
- [7] N.J. Radcliffe The Algebra of Genetic Algorithms, Annals of Mathematics and Artificial Intelligence, 10, 1994, pp. 339-384.