Optimising the deployment of fibre optics using Guided Local Search

Sam Cramer
University of Kent
Email: S.A.Cramer@kent.ac.uk

Michael Kampouridis
University of Kent
Email: M.Kampouridis@kent.ac.uk

Abstract—The deployment of fibre optics poses a huge investment risk, thus telecommunication companies are skeptical about replacing copper given the high cost of doing so. Over recent times, the usage of the internet has changed and led to a need for fibre optics. The decision on whether to deploy or not is made through the use of complex models. However, the problem being that deployment plans are manually predefined based on previous knowledge, this process does not guarantee that the plans are optimal. This paper demonstrates that the deployment of fibre optics can be optimised by using intelligent algorithms. We implemented a metaheuristic (Guided Local Search) to the problem to demonstrate the effectiveness and benefit of looking for an optimal deployment plan. Results indicate that Guided Local Search lead to a significant increase in the profit and can address the problem of finding an optimal deployment plan.

I. INTRODUCTION

The demand for high speed internet has increased dramatically over the years due to changes in internet usage. Ideally, telecommunication companies would like to provide fibre optics for all consumers. The problem arising is the substantial cost associated with replacing the copper wires. This poses a huge investment risk for companies [1], given the cost per connection differs from different concentrations of customers willing to switch to fibre optics in areas. Different options exist, such as fibre to the home (FTTH) and fibre to the cabinet (FTTC). However, this paper uses the general term called fibre to the x (FTTx).

In order for telecommunication companies to make a decision on whether they should deploy fibre optics, a business plan based on a techno-economic model is employed. Techno-economic models are tools that can take into account both technical (e.g. bandwidth requirements and length of cable) and financial fields (e.g. number of customers and cost). Significant research exists for the economical benefit by applying techno-economic models in telecommunications such as [1]. Following this, work has been done in providing a foundation for benefits derived from undertaking techno-economic analysis. [2]-[7] all address the issues surrounding the economic side of the deployment of fibre optics with different focuses.

[4] introduces a novel approach by including technical aspects using Geographic Information Systems to map real life areas. The paper also applies an algorithm in an attempt to optimise deployment, to gain more accurate results. Similarly, [8] looks to provide an optimised layout of a network. Despite all the research that has gone into techno-economic

models for telecommunications, the goal for companies is to maximise profit. Both in industry and literature, there are only a handful of plans that are tested based on past knowledge and experience [9]. Once the plans are chosen, which are then tested for their profitability. Finally, a decision is made based on which plan yields the highest profit. However, if the information provided is incorrect or outdated, the plans may be suboptimal. In this paper we help to address the problem, that decisions to deploy fibre optics are not necessarily optimal.

Consider the following scenario. If telecommunication companies aim to deploy fibre optics over 4 years in a city, made up of 10 areas. The companies would be faced with a total of 9,765,625 (5^{10}) different possibilities. It is unlikely that taking three or five different plans and evaluating them against one another will result in an optimum deployment. The impact from this gap in the literature, is that telecommunication companies could be missing out on more profitable plans.

An approach to find an optimal deployment plan without past knowledge and experience, can be to use metaheuristics. [10] is the first to tackle the problem of finding an optimum deployment plan by applying Genetic Algorithm (GA). Results of applying GA lead to an increase of about 3% in profit, around $17m. Moreover, a more recent extension [11] included three more algorithms.

The aim of this paper is to extend the work by [10][11] to optimise the deployment of fibre optics. The extension will involve implementing a new metaheuristic to the problem, Guided Local Search (GLS). GLS is chosen over similar optimisation algorithms, because of the effectiveness shown [12]-[15]. The main benefit of GLS is its ability to overcome being stuck in a local optima. In order for a comparison to be drawn, the best performing algorithms from previous work is used, Simple Hill Climbing (SHC) and Simple Hill Climbing with Move Swaps (SHC-MS).

The remainder of the paper is as follows: Section II outlines the economic model, for our algorithms to work on. Section III presents the algorithms used in this paper. Section IV presents the experimental design. Section V presents and analyses the results. Section VI concludes the findings.

II. ECONOMIC MODEL

As we are extending the work previously done by [10][11], it falls outside the scope of this paper to present a detailed techno-economic model for two reasons. Numerous number of

models already exist in the literature and economic decisions mainly influence the deployment of FTTx.

A. Overview

Given that the deployment of FTTx is across a city, we define that a city can be divided into a number of areas. The model used in this paper defines areas differently by identifying three aspects. Firstly, all areas have a different population size. Secondly, each area is identified by its own social category. The categories can either be a business, high-income, medium-income and low-income. Finally, each area is part of a town or is a town itself. The three aspects that constitute an area have been randomly assigned and do not represent a real city.

The economic model used in this paper consists of different parameters [10][11], the parameters are the inputs we give the model to generate a profit figure. The following parameters are used:

1) *Rollout customers*. The population of the area.
2) *Rental service customers*. Customers that are currently using the rental services.
3) *Rental service tariffs*. The cost for the customer using rental services.
4) *Pay as you go (PAYG) service customers*. Customers that are currently using the PAYG services.
5) *PAYG tariffs*. The cost for the customer using PAYG services.
6) *Service installation*. A one time cost for the customer when he/she takes up FTTx.
7) *Innovators*. Customers that take up FTTx as soon as the service is made available.
8) *Imitators*. Customers that wait and see before taking up FTTx.
9) *Lost customers*. Customers that leave FTTx for any reason, such as finding a cheaper deal.
10) *Interest rate*. The cost of borrowing money for the investment of deploying FTTx.
11) *Budget*. The maximum investment allowed per year.
12) *Study period*. The length of time the model is run.
13) *Rollout period*. The number of years to deploy FTTx.
14) *No income period*. The number of years taken to bring the service up for customers to use.

An additional input to the model is the *deployment plan*, where we specify for an area the year in which FTTx should be deployed to. Once a deployment plan has been generated, the above parameters are used in the following steps for each year of the study period:

1) Calculate innovator customers
2) Calculate imitator customers
3) Calculate total customers
4) Calculate lost customers
5) Subtract lost customers from total customers
6) Re-calculate total customers
7) Calculate annual revenue
8) Calculate annual cost
9) Calculate profit
10) Calculate Present Value
11) Calculate Net Present Value

The previous steps are explained in more detail below.

B. Innovators

Innovators are those whom adopt FTTx as soon as it becomes available. The percentage of innovators that take up FTTx also depend upon the social category of the area. So in order to calculate the total innovators for an area, we take the product of the percentage of innovators (different percentage depending on social category) and the rollout customers. It is important to note that innovators are only calculated once per area, and the calculation is from the year following the no income period until the end of the rollout period. The year in which the innovators is calculated is dependent upon when the area will be deployed to. An example showing how the innovators are calculated is given in Section II-D.

C. Imitators

Imitators are those whom wait and see before joining FTTx. Similar to innovators, they are based on a percentage that differs depending on the social category they belong to. Like the innovators, the total imitators for an area is the product of the percentage of imitators and the rollout customers. However, because these have a *'wait and see'* attitude, imitators can only be calculated starting the year after the area has been deployed to. In other words, they initially follow the innovators the following year, in subsequent years the imitators will imitate the imitators. An example showing how the imitators are calculated is given in Section II-D.

D. Total customers

Once the innovators and imitators have been calculated, the sum of the two for a given year gives us the total customers. An illustration is shown below for how to calculate the total customers.

TABLE I.
CALCULATING THE TOTAL CUSTOMERS

Customers	Year 0	Year 1	Year 2	Year 3
Rollout customers	100,000	50,000	0	0
Innovators	0	25,000	12,500	0
Imitators	0	0	7,500	10,500
Total new customers	0	25,000	20,000	10,500

Table I shows in Year 0, we rollout to 100,000 potential customers and in Year 1, we rollout to a further 50,000 potential customers. After Year 1, we do not deploy to any new customers. For this particular example, we assume that there only exists one social category. The innovators percentage is 25%, the imitators percentage is 10%, the no income period is one year and that the percentages remain unchanged each year.

Since the no income period lasts for a single year, we have no customers for Year 0 as the services are not available for customers to take-up. Therefore, for Year 1, there are 25,000 (25% * 100,000) innovators. Given that the imitators require a year to 'wait and see' no imitators are calculated for Year 1. For Year 2, there are 12,500 more innovators, which are the new customers from the previous year (25% * 50,000). We also have the imitators from the remaining customers that didn't take it up in Year 1, in this example there are

7,500. It is calculated by taking the difference between the Year 0 customers (100,000) and the innovators who joined in Year 1 (25,000), then multiplying the difference of 75,000 by 10% (imitator percentage). Thus, total customers for Year 2 is 20,000 (7,500 + 12,500).

In Year 3, as there are no new customers to roll out to in Year 2, hence there are no innovators, only imitators. The imitators are calculated in the following way. Firstly, take the difference between the remaining customers from Year 0 (75,000) and the imitators who joined in Year 2 (7,500). Secondly, multiply the difference of 67,500 by 10%, giving 6,750 new customers from the rollout in Year 0. To calculate the imitators from the customers rolled out to in Year 1, take the difference of Year 1 customers (50,000) and the innovators calculated in Year 2 (12,500). Then, multiply the difference (37,500) by 10% to get 3,750 imitators. It gives us a grand total of 10,500 imitators (6,750 + 3,750) for Year 3.

This method for calculating the total customers is carried out in the same way for the remainder of the study period.

E. Lost customers

To make the model more realistic, it is common for those who are using FTTx to leave the service. Whether they are switching to a competitors product or simply do not require the service any more. Similar to imitators and innovators, different social categories have different percentages of lost customers, because a low income family may no longer be able to afford using FTTx. The product of the lost customer percentage (social category dependent) and the total customers returns the lost customers.

F. Total customers

Recalculate the total number of customers by removing those who left the service. The lost customers are then re-added back into the rollout customers.

G. Annual revenue

There are three sources of revenue for the company: rental services, PAYG services and service installation.

The rental services and PAYG services are both calculated in the same way. The number of customers on each plan is expressed by the product of the percentage of each service and the total customers. The calculation of revenue generated per month is the product of the number of customers on each service plan and the tariff for each plan.

In order to work out the annual revenue we need to multiply the revenue per month by 12. As the total number of customers are spread out over the year, an average annual revenue figure needs to be calculated. To give a reasonable estimate under the assumption that customers join linearly, dividing the annual revenue by 2, a rough average can be calculated.

The service installation revenue is the product of the new customers for the year and the service installation charge. Then, take the sum of annual average revenue and the service installation to calculate the total revenue for a year.

H. Annual Cost

The annual cost is calculated by the sum of capital expenditure (CAPEX) and operational expenditure (OPEX). CAPEX is a one time cost per area when it is deployed to and OPEX is a reoccurring annual cost from deployment till the end of the study period. CAPEX refers to the infrastructure cost, i.e. the cost of tunnelling and laying the cables. OPEX refers to the on going costs, such as maintenance charges and utilities.

I. Profit

The difference between the annual revenue and cost is defined as the profit or loss for the given year.

J. Present Value

The Present Value (PV) is the future amount of money, discounted to reflect the time value of money. The PV of profit P with a given interest rate i, t time periods in the future is given by:

$$P_t = \frac{P}{(1+i)^t}. \tag{1}$$

K. Net Present Value

Net Present Value (NPV) is the sum of all PVs over the study period. The NPV is given by:

$$\sum_{t=0}^{n} \frac{P}{(1+i)^t}, \tag{2}$$

where n is the study period.

III. ALGORITHMS

The economic model presented only takes a deployment plan as a single input. This limits us of only being able to evaluate a single manually generated plan. To overcome this limitation, we need to generate and evaluate multiple plans in order to find the most profitable one. To draw a comparison for the effectiveness among of our algorithms, there must be some consistency in the selected algorithms and how they interact with the economic model. As mentioned earlier, we are comparing GLS to the best performing algorithms from [11] SHC and SHC-MS.

A. Representation

The deployment plan mentioned in Section II-A is treated as an input and is represented by a list of numbers. Firstly, the length of the list indicates the number of areas we are interested in rolling out to. Secondly, each position within that list represents the area that the deployment plan relates to. In other words, the first element of the list relates to 'Area 1', the second element of the list relates to 'Area 2' and so on. Finally, each element within the list requires the year in which each area deploys to, consisting of a numerical value between 0 and the rollout period.

To give a visual example of how we represent our deployment plan, see Figure 1 below.

From Figure 1 we can see the years that each area will deploy to. Note that you can have many areas deploying in

Fig. 1. 10 areas representing our deployment plan
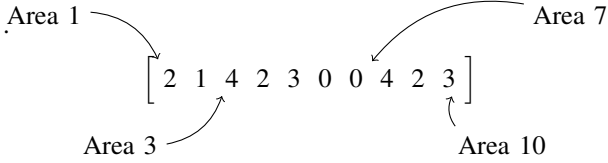
the same year. For example, areas 1 and 9 are both deploying in the second year, you can deploy to as many areas so long as it does not violate the annual budget. Additionally, you can have areas where no deployment occurs (areas 6 and 7).

B. Fitness function

A fitness function needs to be assigned to compare deployment plans. For this paper the following definition will be used:

$$\text{Fitness} = \text{NPV}. \tag{3}$$

Any output could be used depending on what is hoped to achieve through the model. For example, if the aim is to maximise customers, the fitness function is equal to total customers. As our aim is to maximise profit, then assigning the fitness function as such is the most appropriate. Thus, the most 'fit' deployment plan will be the most profitable.

Next we present the three algorithms that are built around the deployment plan and fitness function. First we will discuss the first two local search techniques before moving on to GLS.

C. SHC

SHC is a metaheurisitc, which is part of the local search family and is used for optimising computationally complex problems. Local search algorithms have been very successful in tackling large scale problems, such as the Travelling Salesperson Problem [16]. SHC initiates by randomly generating a deployment plan, similar to that as shown in Figure 1. SHC operates by making iterative improvements to the deployment plan in order to arrive at a new neighbourhood i.e. a new deployment plan, whilst converging towards the optimum position. The process begins by randomly visiting each area, until all elements within the deployment plan are visited but are not necessarily in sequential order. Whenever SHC lands on a particular area, it reduces the value of that area and tests the new fitness of the deployment plan. If the fitness level is higher than the previous deployment plan, then SHC moves to another area within the deployment plan. Otherwise, it continues to reduce the value of the element until either a higher fitness is returned, or no improvement can be made. When all areas within the deployment plan have been visited, the search continues in the same way until there are no neighbouring plans that results in an improvement in fitness. Therefore, an optimal deployment plan has been found.

D. SHC-MS

SHC-MS is a combination of two algorithms, SHC and Move Swaps (MS). Essentially, the algorithm is the same as SHC, but SHC-MS has an additional way of modifying a deployment plan using MS. MS allows for two areas to swap with each other. To determine whether a move making a move either decreases the value of the area by 1 or performs a swap with another area. A probability is assigned to the MS and is calculated when SHC lands on an area.

E. Guided Local Search

GLS is a metaheuristic that is used to assist local search techniques to escape from the problem of being stuck in local optimas, by the use of penalties, i.e. GLS sits on top of a local search algorithm. GLS penalises costly features present within a solution and so deters the local search from visiting such solutions again. In order to apply GLS to optimise our deployment plan, solution features must be defined for the problem at hand. A feature is a property that either exists or not in the given solution. For the purpose of this problem the solution features are the years that deployment can take place. For each feature present in the solution, there is a cost associated with it. The cost is mathematically calculated according to the feature present, so in the case of our deployment plan the cost of our features would be the total operating and capital expenditure per area over the study period.

Whether a feature f_{ij} (i refers to area and j refers to deployment year) is present in the current solution or not is given by:

$$I_{ij}(s) = \begin{cases} 1 & \text{if solution } s \text{ has property } ij \\ 0 & \text{otherwise} \end{cases}. \tag{4}$$

GLS makes use of these features by combining the penalties with our fitness function. This combination means that we are effectively making our deployment plan more expensive and decreasing our fitness function. The result of this combination is an *augmented fitness function* (denoted as $h(s)$, where s is a deployment plan) as defined below:

$$h(s) = g(s) - \lambda \cdot \sum_{\substack{i=1 \\ j=0}}^{M} p_{ij} \cdot I_{ij}(s), \tag{5}$$

where $g(s)$ is the fitness function for deployment plan s, M is the total number of features, p_{ij} is the penalty value according to f_{ij}, and λ is the *regularization parameter*. λ is necessary as it allows us to control how much the penalties influence the augmented fitness function. If λ is set at 0, then GLS would never be able to escape the local maximum. On the other hand, if λ is set to 1, then this completely removes the penalised feature, which is ideal. However, if information is incorrect, it could potentially remove useful features. The target is to now work out an effective value for λ [17]. In order to do so, we use the formula from [17]. A new parameter α is introduced shown by the following formula:

$$\lambda = \alpha \cdot g(s)/(F_s), \tag{6}$$

where $g(s)$ is the original fitness function, F_s is the number of features present in the solution and α is our lambda coefficient, which is dependent upon the problem at hand.

For every local search that gets stuck in a local maxima, modifications are made to the augmented fitness function.

Local search then resumes from the last deployment plan. However, the fitness function is less than it previously is, because the fitness of that plan has been penalised allowing the algorithm to move out of its local maxima. GLS works by incrementing the penalty value of those features whose *utility* (denoted as $u(s, f_{ij})$) is maximised, the utility of each feature is given as:

$$\arg\max_{s, f_{ij}} u(s, f_{ij}) = I_{ij}(s) \cdot \frac{c_{ij}}{1 + p_{ij}}, \qquad (7)$$

where c_{ij} is the cost of the feature present in the deployment plan s and p_{ij} is the penalty of the feature present in deployment plan s.

To avoid continuously penalising areas with high cost, the penalty parameter p_{ij} is included so that it avoids bias and allows for penalties to be allocated more diversely. When local search gets stuck in the next local maximum then the process repeats itself.

IV. EXPERIMENTAL DESIGN

This section establishs the initial setup of our model in order to run experiments using our algorithms. Different simulations need to be created in order to test the effectiveness. The different simulations looks to maximise our fitness function (as specified in Section III-B) as long as it fits within the constraints presented next.

A. Constraints

This first constraint limits the number of areas deployed to within a certain year, as long as the cost of deployment for the year does not exceed the budget level. Having a budget constraint is a very typical problem for any large scale investment project as companies simply do not have the resources to carry out such a large scale activity. The reasons could be that the telecommunication companies are restricted by money, labour, capital or time.

The second constraint has a maximum number of areas per town to be deployed to in a certain year, given a project with activities such as digging up roads to lay new cable may be disruptive.

The third constraint forbids us from deploying to too many areas in close proximity to each other, again given the potential disruption caused.

Finally, the fourth constraint is to force the deployment of FTTx to an area within a specified time frame. This is common in real life as it may be necessary to go through certain areas first in order to reach others.

B. Experiments

In order to evaluate the algorithms, three experiments are conducted for 100 areas and 500 areas using the constraints outlined in Section IV-A.

Experiment one will introduce constraint one. The budget that will be used for the experiments given in Section IV-D.

Experiment two will introduce constraints two and three. Constraint two limits the number of areas within a town, dependent upon the size of the town e.g. if a town consists of 10 areas then only a maximum of 3 areas per year can be deployed to. Constraint three will be that the deployment plan cannot have 3 or more areas in a row having the same deployment year.

Experiment three introduces constraint four. The areas chosen to be forced are assigned randomly per town, at least one per town.

C. Experimental setup

The values we use for the parameters are artificially made up specifically for the model. Ideally, the use of real data would be useful so that this approach can be applied to a real city. However, the experimental nature of this paper allows us to deal with the problem in a more controlled environment.

D. Parameter settings

Initially, the population per area is based on a random number that follows a Normal distribution from 0 to 10,000 customers, where a customer may be either a household or a business. Each area is assigned a social category as is mentioned in Section II-A, business (b), high income (h), medium (m) and low income (l). The total number of each social category is chosen and is shown in Table II.

TABLE II.
THE NUMBER OF AREAS PER SOCIAL CATEGORY FOR TWO DIFFERENT AREA SCENARIOS

Social category	Number per 100 areas	Number per 500 areas
Business	3	15
High income	10	50
Medium income	33	165
Low income	54	270

The different social categories are important, because those from a higher income area would be more available to take up FTTx. Alternatively, it might be too expensive for those of low income. This is reflected by the different percentages of uptake among innovators and imitators, as well as the remaining parameters values as shown in Table III.

TABLE III.
THE PARAMETERS FOR OUR ECONOMIC MODEL

Name of parameter	Input values
Number of areas	100, 500
Social categories	4
Rental services customer percentage	100%
Rental services tariffs	£80 per month
PAYG services percentage	0%
PAYG tariffs	£15
Service installation	£150
Innovators percentage (b, h, m, l)	25%, 20%, 15%, 10%
Imitators percentage (b, h, m, l)	20%, 15%, 10%, 5%
Lost customer percentage (b, h, m, l)	5%, 5%, 5%, 5%
Interest rate	10% per annum
Budget 100 areas	£60,000,000 per year
Budget 500 areas	£300,000,000 per year
Study period	20 years
Rollout period	4 years
No income period	1 year

From this table, we assume that all customers use rental services. This is purely for simplicity and has no effect on the results of our experiments. The value of £80 for the rental service charge is what exists within the literature [18]. We also identify different uptake rates for innovators and imitators given their social category. All the values are kept the same throughout the study period and all experiments use the same values. 100 and 500 areas are used throughout the experiments had have a budget of £60,000,000 and £300,000,000 respectively.

The CAPEX and OPEX per area are £2,500,000 and £40,000 respectively, these values are consistent with what appears in the literature [19]. As is mentioned in Section II-H, CAPEX only applies to the rollout period and OPEX applies across the whole study period. Due to these costs and our budget, only a certain number of areas can be deployed to in a given year. Therefore, it is up to the algorithms to determine which areas are the least profitable to leave out.

Additionally, areas within a town may share common resources such as service stations or connection boxes. This shared cost only gets calculated once for an area when deployment occurs for a certain town.

E. Algorithm settings

For both SHC and SHC-MS, we continue running until a local optimum position has been located, causing the loop to stop. For SHC-MS, in order to determine whether a swap will occur requires a probability. In this paper, the probability is chosen at 0.5, a 50% chance of swapping, based on the effectiveness from [11].

For GLS, the algorithm makes a call to the local search 100 times before returning the best plan found, based on initial tests of different calls. Regarding the tuning of λ mentioned in Section III-E, tests are conducted to calculate an effective α value and concludes that 0.001 is the most appropriate.

For statistical purposes, we run each algorithm for 50 individual runs. In order for a fairer comparison to be made among algorithms, the initial deployment plan used are randomly generated. For every run, a new randomly generated plan is used. Therefore, there are a total of 50 different initial deployment plans for our algorithms to work through.

V. RESULTS

In this section, we discuss the results from our three experiments under both 100 and 500 areas. In order to compare how well the algorithms performed against each other, they are ranked accordingly using the Friedman test [20]. The results are then checked at a given significance level, to see if there is a significant difference among algorithms. In other words, determining whether we accept or reject our null hypothesis. Next, we determine which algorithms are significantly different, by using the Holm post-hoc test. Since GLS sits on top of SHC and SHC-MS, we refer to these as GLS-SHC and GLS-SHC-MS respectively.

Looking at Tables IV and V, one thing to observe is that SHC struggled in all experiments under 100 areas and finished last as can be shown in Table VII. The values in bold represent the best performing algorithm(s).

TABLE IV.
THE RESULTS FROM RUNNING THE EXPERIMENTS UNDER 100 AREAS

Measure	SHC	SHC-MS	GLS-SHC	GLS-SHC-MS
Mean	3.36E+08	**3.79E+08**	3.45E+08	**3.79E+08**
SD	1.08E+07	1.91E+05	4.61E+06	**0.00E+00**
Max	3.64E+08	**3.79E+08**	3.64E+08	**3.79E+08**
Min	3.14E+08	3.78E+08	3.36E+08	**3.79E+08**

(a) Experiment one

Measure	SHC	SHC-MS	GLS-SHC	GLS-SHC-MS
Mean	3.34E+08	3.45E+08	3.42E+08	**3.65E+08**
SD	9.94E+06	1.09E+07	4.66E+06	**1.85E+06**
Max	3.60E+08	3.67E+08	3.53E+08	**3.69E+08**
Min	3.14E+08	3.08E+08	3.32E+08	**3.61E+08**

(b) Experiment two

Measure	SHC	SHC-MS	GLS-SHC	GLS-SHC-MS
Mean	3.33E+08	3.42E+08	3.41E+08	**3.64E+08**
SD	9.53E+06	1.31E+07	7.33E+06	**1.69E+06**
Max	3.56E+08	3.62E+08	3.60E+08	**3.69E+08**
Min	3.15E+08	3.05E+08	3.25E+08	**3.62E+08**

(c) Experiment three

TABLE V.
THE RESULTS FROM RUNNING THE EXPERIMENTS UNDER 500 AREAS

Measure	SHC	SHC-MS	GLS-SHC	GLS-SHC-MS
Mean	1.71E+09	**1.91E+09**	1.72E+09	**1.91E+09**
SD	1.44E+07	6.08E+04	1.19E+07	**3.83E+03**
Max	1.74E+09	**1.91E+09**	1.75E+09	**1.91E+09**
Min	1.67E+09	**1.91E+09**	1.69E+09	**1.91E+09**

(a) Experiment one

Measure	SHC	SHC-MS	GLS-SHC	GLS-SHC-MS
Mean	1.69E+09	1.66E+09	1.70E+09	**1.81E+09**
SD	1.63E+07	5.60E+07	1.70E+07	**7.06E+06**
Max	1.73E+09	1.76E+09	1.74E+09	**1.82E+09**
Min	1.66E+09	1.53E+09	1.66E+09	**1.79E+09**

(b) Experiment two

Measure	SHC	SHC-MS	GLS-SHC	GLS-SHC-MS
Mean	1.69E+09	1.64E+09	1.70E+09	**1.80E+09**
SD	1.62E+07	7.39E+07	3.56E+07	**6.60E+06**
Max	1.72E+09	1.76E+09	**1.82E+09**	1.82E+09
Min	1.66E+09	1.46E+09	1.66E+09	**1.79E+09**

(c) Experiment three

SHC-MS is able to hit high NPV as shown by Tables IV and V. However, SHC-MS may be too random at times and thus loses consistency, shown by the larger standard deviation (SD) in Tables IV and V.

Across the experiments using 100 areas GLS-SHC gives on average around a 2.56% increase in NPV to SHC as shown in Table VI, which is impressive given how limited SHC is. Additional, GLS helps lower the SD, as shown in Table IV, where the SD of GLS-SHC compared to SHC

is 4.61E+06 and 1.08E+07 respectively for experiment one, 4.66E+06 and 9.94E+06 respectively for experiment two and 7.33E+06 and 9.53E+06 respectively for experiment three. It is more inconsistent under 500 areas as the SD compared to SHC is greater in experiment two (1.70E+07 and 1.63E+07 respectively) and experiment three (3.56E+07 and 1.62E+07 respectively), as shown in Table V.

TABLE VI.
THE PERCENTAGE GAIN IN MEAN NPV FROM APPLYING GLS TO OUR LOCAL SEARCH ACROSS ALL EXPERIMENTS

Algorithm	Area	Exp. one	Exp. two	Exp. three
GLS-SHC vs SHC	100	2.54%	2.54%	2.61%
GLS-SHC-MS vs SHC-MS	100	0.04%	6.07%	6.60%
GLS-SHC vs SHC	500	0.87%	0.51%	0.70%
GLS-SHC-MS vs SHC-MS	500	0.01%	8.49%	10.14%

By applying GLS on top of SHC-MS, the performance gain is highly significant, where it manages to outperform every algorithm on every experiment. As shown in Tables VII and VIII, where GLS-SHC-MS ranks first in every experiment. What is most impressive is the consistency that it shows in all experiments, especially when faced with 500 areas, as shown in Table V. The SD of GLS-SHC-MS is the lowest in every experiment compared to all algorithms not just SHC-MS. The performance increases dramatically when compared against its respective local search. On average the performance gain in mean NPV is approximately 4.24% over all experiments under 100 areas and is approximately 6.21% under 500 areas, as is shown in table VI.

GLS-SHC-MS returns the highest maximum NPV for every experiment we run, as shown in Tables IV and V. Under these experiments, we can see that GLS-SHC-MS is able to create very large differences when we compare against the highest NPV from other algorithms. In experiment three using 100 areas, we see that GLS-SHC-MS outperforms its nearest rival by 1.87%. This relates to a monetary gain of £7m as shown in Table IV. When we compare that of 500 areas using experiment three, GLS-SHC-MS outperforms the second highest maximum NPV of GLS-SHC by £1.5m with the third highest maximum NPV of SHC-MS being £60m.

When we apply the Friedman test to the experiments, we observe the following results as shown in Tables VII and VIII. The values highlighted in bold represent that there is a significant difference against the best performing algorithm.

Given the results shown in Tables VII and VIII, GLS-SHC-MS achieves the highest performance in all experiments in both 100 and 500 areas. The above results are very consistent, and show that GLS-SHC-MS is significantly better than SHC and GLS-SHC in all six instances. As the p-values of SHC and GLS-SHC are less than the Holm's critical value. Additionally, in four instances GLS-SHC-MS is significantly better than SHC-MS, as the p-values of SHC-MS are less than the Holm's critical value. Therefore, from the results presented we can conclude that GLS-SHC-MS was the best performing algorithm.

GLS-SHC is able to find smaller increases in mean NPV compared to SHC, because the performance of SHC is fairly limited. This still shows that GLS is an effective algorithm as it could make improvements with significant limiting factor.

TABLE VII.
THE RESULTS FROM THE FRIEDMAN TEST AT 5% SIGNIFICANCE LEVEL FOR 100 AREAS

Algorithm	Ranking	p-value	Holm
GLS-SHC-MS (best)	1.43	-	-
SHC-MS	1.57	0.59	0.05
GLS-SHC	**3.15**	**2.71E-11**	**0.025**
SHC	**3.85**	**7.08E-21**	**0.0167**

(a) Experiment one

Algorithm	Ranking	p-value	Holm
GLS-SHC-MS (best)	1.00	-	-
SHC-MS	**2.52**	**3.93E-9**	**0.05**
GLS-SHC	**2.75**	**1.22E-11**	**0.025**
SHC	**3.73**	**3.97E-26**	**0.0167**

(b) Experiment two

Algorithm	Ranking	p-value	Holm
GLS-SHC-MS (best)	1.00	-	-
SHC-MS	**2.68**	**7.69E-11**	**0.05**
GLS-SHC	**2.70**	**4.58E-11**	**0.025**
SHC	**3.62**	**3.41E-24**	**0.0167**

(c) Experiment three

TABLE VIII.
THE RESULTS FROM THE FRIEDMAN TEST AT 5% SIGNIFICANCE LEVEL FOR 500 AREAS

Algorithm	Ranking	p-value	Holm
GLS-SHC-MS (best)	1.50	-	-
SHC-MS	1.50	1.0	0.05
GLS-SHC	**3.10**	**5.76E-10**	**0.025**
SHC	**3.90**	**1.47E-20**	**0.0167**

(a) Experiment one

Algorithm	Ranking	p-value	Holm
GLS-SHC-MS (best)	1.00	-	-
GLS-SHC	**2.67**	**9.94E-11**	**0.05**
SHC	**3.03**	**3.78E-15**	**0.025**
SHC-MS	**3.30**	**5.20E-19**	**0.0167**

(b) Experiment two

Algorithm	Ranking	p-value	Holm
GLS-SHC-MS (best)	1.04	-	-
GLS-SHC	**2.59**	**1.94E-9**	**0.05**
SHC	**2.87**	**1.36E-12**	**0.025**
SHC-MS	**3.50**	**1.61E-21**	**0.0167**

(c) Experiment three

On the other hand, the randomness of SHC-MS proves to be very effective as SHC-MS is able to achieve high maximum NPV and so GLS is able to learn and adapt more effectively. Therefore, it appears that GLS is only as effective as the hill climber it sits on.

Overall, the results from our experiments are of great significance and are showing a sign that computational intelli-

gence can be used to find an optimal deployment plan, especially when using GLS. The use of computational intelligence can help decision making and give companies more confidence in undergoing potentially risky investment decisions compared to existing methods.

VI. CONCLUSION

This paper considers a new metaheuristic to determine the optimal deployment plan of fibre optics. In particular, this approach addresses the problem that telecommunication companies use a manual approach for deploying fibre optics to a city. As companies may not necessarily be using the most optimal deployment plan from the use of past experience may include outdated information. In this paper, we introduced the metaheuristic GLS to extend the work by [10][11].

GLS was our chosen metaheuristic as it had been shown through the work of [12]-[15] to be very effective at optimisation problems and outperforms the state of the art algorithms like Genetic Algorithm. We compared GLS with two different local search algorithms, as the previous work by [11] showed that they were effective for the problem at hand.

To compare the effectiveness of GLS, we outlined six different experiments and presented the findings. The results from the experiments were of great significance to the literature. From the experiments we were able to show that GLS was an effective metaheuristic, as GLS was able to significantly improve its local search. The use of GLS was able to provide a substantial monetary gain when compared against other algorithms e.g. under 500 areas GLS-SHC-MS found £150m more in mean NPV compared to the others. We also noticed that GLS-SHC-MS achieved the highest maximum NPV in every experiment. We found that GLS-SHC-MS was significantly better than the other algorithms and showed that the application of GLS was very beneficial to the problem.

Future work could be to test other metaheuristics to the problem, other types of optimisation techniques or to test different local search techniques with GLS. Alternatively, GLS could be further tested on a techno-economic model rather than the economic model put forward in this paper to be build a more complete model that is common practice in industry.

REFERENCES

[1] L. A. Ims, D. Myhre, B. T. Olsen, "Economics of Residential Broadband Access: Network Technologies and Strategies," *Network, IEEE*, Vol. 11, Issue 1, pp. 51-57, 1997.

[2] B. T. Olsen, D. Katsianis, D. Varoutas, K. Stordahl, J. Harno, N. K. Elnegaard, I. Welling, F. Loizillon, T. Monath, P. Cadro, "Technoeconomic evaluation of the major telecommunication investment options for European players," *Network, IEEE*, Vol. 20, Issue 4, pp. 6-15, 2006.

[3] T. Rokkas, D. Katsianis, D. Varoutas, "Techno-economic Evaluation of FTTC/VDSL and FTTH Roll-Out Scenarios: Discounted Cash Flows and Real Option Valuation," *Optical Communications and Networking, IEEE/OSA Journal of*, Vol. 2, Issue 9, 2010.

[4] B. Van Den Bossche, R. Meersman, J. Vanhaverbeke, A. Schoutteet, "Maximizing the Return on Investment for FTTX-rollout Through the Sse of GIS Street Maps and Geomarketing Data," *Telecommunications Internet and Media Techno Economics (CTTE)*, 9th Conference, 1,6, and 7-9, 2010.

[5] M. Tahon, J. Van Ooteghem, K. Casier, L. Deschuttere, S. Verbrugge, D. Colle, M. Pickavet, P. Demeester, "Cost allocation model for a synergetic cooperation in the rollout of telecom and utility networks," *Telecommunication, Media and Internet Techno-Economics (CTTE), 10th Conference of*, 2011.

[6] I. Tomkos, "Techno-Economic Comparison of Next Generation Optical Access Network Architectures," *50th FITCE Congress*, 2011.

[7] K. Walczyk and A. Gravey, "Techno-Economic Comparison of Next-Generation Access Networks for the French Market," *Information and Communication Technologies, Lecture Notes in Computer Science*, Vol. 7479, pp. 136-147, 2012.

[8] A. R. Shaghaghi, T. Glover, M. Kampouridis, E. Tsang, "Guided Local Search for Optimal GPON/FTTP Network Design," *In Chaki, N., Meghanathan, N., Nagamalai, D. (Eds.): Proceedings of the Fourth International Conference on Networks & Communications, Lecture Notes in Electrical Engineering*, Vol. 131, Springer, 2013.

[9] Y. Kawai, Y. Horita, "Fibre Optic Communities in the U.S. Their Deployment, Application, and Spatial Planning Strategies for the Suburban and Rural America," *Journal of Housing Research Foundation*, No. 35, pp. 1-12, 2009.

[10] M. Kampouridis, T. Glover, A. R. Shaghaghi, "Using a Genetic Algorithm as a Decision Support Tool for the Deployment of Fibre Optic Networks," *Proceedings of the IEEE World Congress on Computational Intelligence*, Brisbane, Australia, 2012.

[11] M. Kampouridis, T. Glover, A. R. Shaghaghi, "Deciding the optimal roll-out plan for the deployment of fiber optic networks," *Journal of Heuristics*, (Under Review), 2014.

[12] P. Kilby, P. Prosser, P. Shaw, "Guided Local Search for The Vehicle Routing Problem," *In Proceedings of the 2nd International Conference on Metaheuristics*, pp. 473-486, 1997.

[13] C. Voudouris, "Guided local search for combinatorial optimisation problems," *PhD Thesis, Department of Computer Science, University of Essex, Colchester, UK, July*, 1997.

[14] C. Voudouris and E. P. K Tsang, "Guided Local Search and its application to the Travelling Salesman Problem," *European Journal of Operational Research*, Vol. 113, Issue 2, pp. 469-499, 1999.

[15] Y. Zhong, M. H. Cole, "A Vehicle Routing Problem with Backhauls and time windows: a guided local search solution," *Transportation Research Part E: Logistics and Transportation Review*, Vol. 41, Issue 2, pp. 131-144, 2005.

[16] L. McGeoch and D. Johnson, "Experimental analysis of heuristics for the STSP," *In G. Gutin, and A. Punnen (Eds.), The traveling salesman problem and its variations*, Boston: Kluwer Academic Publishers, pp. 369487, 2002.

[17] A. Alsheddy and M. Kampouridis, "Off-line Parameter Tuning for Guided Local Search Using Genetic Programming," *Proceedings of the IEEE World Congress on Computational Intelligence*, Brisbane, Australia, 2012.

[18] K. Casier, S. Verbrugge, R. Meersman, D. Colle, M. Pickavet, P. Demeester, "Techno-economic evaluations of FTTH roll-out scenarios," *Proceedings of NOC2008, the 13th European Conference on Networks and Optical Communications*, pp. 113-120, 2008.

[19] K. Casier, "Techno-economic evaluation of a next generation access network deployment in a competitive setting," *Ph.D. dissertation, Ghent*, 2009.

[20] J. Demsar, "Statistical comparisons of classifiers over multiple data sets," *Journal of Machine Learning Research*, Vol. 7, pp. 1-30,2006.