Journal of the Eastern Asia Society for Transportation Studies, Vol. 6, pp. 624 - 639, 2005

OPTIMAL IMPLEMENTATION-PATH OF ROAD PRICING SCHEMES WITH TIME-DEPENDENT MODEL

Agachai Sumalee
Senior Research Fellow
Institute for Transport Studies,
University of Leeds
38 University road, Woodhouse Lane
LS2 9JT, LEEDS, UK
Fax: +44-113-343-5334
Email:asumalee@its.leeds.ac.uk

Abstract: This paper proposes an optimisation algorithm for designing an optimal implementation path for a charging cordon scheme. The main incentive for time-dependent design is public acceptability issue. A simple equilibrium based time-dependent model is defined. Different patterns of the implementation path are investigated including (i) evolution of a single cordon, (ii) single cordon and additional screenlines, and (ii) inner cordon and gradual introduction of outer cordon. The optimisation algorithm developed is based on the idea of Genetics Algorithm (GA). The algorithm is tested with a network of the Edinburgh city in UK.

Key words: Road Pricing, Bilevel Optimization, Genetics Algorithm, Time-dependent model

1. INTRODUCTION & MOTIVATION

Economists claim that road users do not pay the true price of their road use which leads to inefficiency and congestion in the transport system (Pigou 1920; Vickrey 1969). Particularly, there exist externalities (in the form of congestion, pollution, and accidents) that the car users do not perceive and are not charged for. A fiscal tool to rectify the failure in transport market and resolve the congestion problem, thus, has been proposed which is the idea of congestion charging or road pricing. Congestion charging raises the social welfare of the system, in theory, by imposing appropriate charges for the usage of roads.

Although this economic concept sounds very promising, the actual implementation and design of a road pricing scheme is much more complex in the real world. Sumalee (2001) and May *et al* (2002) investigated practical design criteria for a road pricing scheme as considered by the practitioners in the UK. The main conclusions from their survey are that:

(i) the local authorities consider a simple charging design (i.e. cordon format) with simple charging structure (uniform toll) to be most appropriate;
(ii) issues related to public and political acceptability are probably the main criteria for the design of the initial scheme;
(iii) related to (ii) it is widely accepted that the initial scheme will not be an optimal design but there exists possibility to evolve the scheme overtime to achieve better performance.

A number of algorithms for the second-best optimal toll problem have been proposed (See Clegg et al. 2001; Verhoef 2002; Lawphongpanich and Hearn 2004; Meng et al. 2001; Shepherd and Sumalee 2004). However, none of these works have considered the practical aspects mentioned above. Several authors have attempted to address some of the practical aspects in the design of a road pricing scheme. For instance, some attempts have been made to integrate the requirement of a charging cordon formation into the optimal design (Sumalee 2004; Zhang and Yang 2004; Hyman and Mayhew 2002; Mun et al. 2001). Sumalee (2003) extended the methodology to incorporate the practical constraint on the scheme design (e.g. equity impact) into consideration. These researches represent the shift from the normative economics (i.e. study of the optimal toll policy in theory) toward the positive economics paradigm (study of how to successfully implement a road pricing scheme). Nevertheless, despite these recent efforts, there has been little attention on the issue of evolution design of road pricing scheme over time.

The aim of this paper is to develop an optimisation algorithm for designing an optimal implementation-path of a charging cordon scheme over the time horizon. This implementation strategy is expected to help raising the public acceptability during the initial stage of the scheme as well as enhancing the performance of the scheme over a longer time period. Interestingly most of the real world road pricing schemes has been implemented following this idea. The Singapore Area Licensing Scheme (ALS) was initially implemented with a very simple cordon based design in 1970s (Holland and Watson 1978). Then, the scheme has been gradually modified to achieve better performance by introducing additional screenlines, modifying the complexity of charges, adding additional cordons, etc. Although the main reason for evolution of the scheme in Singapore is not entirely due to public acceptability, it is an evidence of the benefit of a gradual scheme implementation to adjust to the change of the available charging technology and traffic condition.

The congestion charging scheme in London was implemented in 2003. Similar to the case of the scheme in Singapore, the initial design of the London congestion charging scheme is also very simple. Recently there has been an initial discussion on the possible extension of the scheme to the west part of the central London to increase the performance of the existing scheme after gaining some public support from the initial scheme. Schade and Schlag (2003) produced a research result supporting this idea by differentiating the difference between ‘acceptability’ (before the implementation) and ‘acceptance’ (after the implementation) of the scheme.

The paper is structured into further four sections. The next section will describe the problem formulation of the optimal implementation-path of a charging cordon scheme. Some attention will be given to the definition of the time-dependent traffic model and possible patterns of cordon scheme modification. Then, Section 3 explains the development of the optimisation algorithm for optimal implementation-path based on Genetics Algorithm (GA). Section 4 presents the numerical results from the tests with the Edinburgh network. Finally, Section 5 concludes the paper and suggests further research issues.

2. PROBLEM FORMULATIONS

2.1 Standard optimal charging cordon design problem

If a toll is imposed on one of the links in a two-link network, it is likely that the existing drivers on the tolled route will switch to the untolled route (route choice) or they may give up their trip given the increase in the travel cost (elastic demand condition). In analysing the optimal toll design, one needs to include the demand response into the problem. In this paper, Wardrop's equilibrium condition (Wardrop 1952) will be adopted as the representation of users' behaviours in the network. The discussion of the plausibility of different modelling assumptions is beyond the scope of this paper.

In the optimal charging cordon design problem, the main aim is to choose a charging cordon location with its uniform toll so as to maximise an objective function (after the travellers adapt their behaviours). Let ε_j be a dummy variable taking the value of 1 if link j is to be tolled and 0 otherwise (ε is a vector of ε_j), $\mathbf{v}$ and $\mathbf{d}$ be a vector of link flow and demand level respectively, and t be the uniform toll level of the cordon. In addition, let Θ be the set of all possible combinations of ε_j that will form a charging cordon. A number of constraints on the feasible cordon location (e.g. cordon should not cover a particular area) can also be imposed in which Ω represents the set of feasible cordons. **g** represent vector functions of other constraints on the design (e.g. equity impact)The optimal charging cordon design can then be defined as follows:

$$\begin{aligned}
&\max_{(\mathbf{v},\mathbf{d},\varepsilon,t)} F(\mathbf{v},\mathbf{d},\varepsilon,t)\\
&s.t.\\
&(\mathbf{v},\mathbf{d}) \rightarrow \text{Wardrop's equilibrium solution given } (\varepsilon,t)\\
&\mathbf{g}(\mathbf{v},\mathbf{d},\varepsilon,t) \leq \mathbf{0}\\
&\varepsilon \in \Omega \cap \Theta
\end{aligned} \tag{1}$$

Note that formulation (1) above is a very informal mathematical form of the problem and is presented here only for the demonstration purpose. In fact, the optimisation algorithm developed later in the next section does not use any of the mathematical property of (1).

The optimal toll design problem as shown in (1) can be categorised as Mathematical Program with Equilibrium Constraints (MPEC). MPEC is well known for its undesirable properties for optimisation problem (i.e. non-convex feasible region and potentially non-smooth objective function) (Luo et al. 1996). Thus, an optimisation method based on Genetics Algorithms (GA) will be developed instead in this paper. In addition, GA can handle the complexity of the constraints on the location of the tolled point which will be described later in Section 2.3.

The objective function in (1) is defined as the social welfare. This can be defined as: consumer surplus + operator surplus. Using the rule of half, the consumer surplus can be approximated as:

$$CS = \frac{1}{2} \cdot \sum_{rs} \left[\left(d_{rs}^0 + d_{rs} \right) \cdot \left(\mu_{rs}^0 - \mu_{rs} \right) \right]$$

, where d_{rs}^0 and d_{rs} are the demand levels in the do-nothing and do-something scenarios for OD pair *rs*, and μ_{rs}^0 and μ_{rs} are the minimum travel cost between OD pair *rs* in the do-nothing and do-something scenarios. With the road pricing policy, the operator surplus is the net of the revenue generated by the scheme and the cost of the scheme.

A number of possible outcome constraints are available. For instance, there could be a minimum level of revenue required from the scheme or a maximum level of the equity impact (see Sumalee, 2003 for the quantification of the equity impact using the *Gini* coefficient).

2.2 Simple time-dependent model and optimal implementation-path problem

There are many possible frameworks for time-dependent model available (References). These frameworks can all be integrated into the optimisation algorithm developed in Section 3. It should be noted that a time-dependent model developed in this paper is a very simplified version.

The model adopted in this paper covers a 15-year planning horizon. It is an equilibrium based model assuming that the traffic state is stable within the period of five year interval. Thus, over the whole 15 years there will be three asynchronous Wardrop's equilibrium states each of which covers a period of five year. Figure 1 demonstrates this model. The planner in this problem is allowed to implement an initial cordon in the first year, and then modify the scheme twice in year 6 and year 11. It also assumed that there exists some exogenous demand growth.

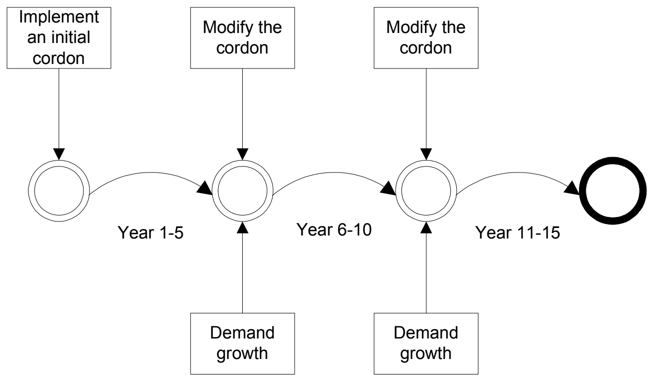


Figure 1: Framework of time-dependent model and implementation path

The linkage between each state of equilibrium condition is established through the demand function. The demand function adopted in this model is defined as:

$$d_{rs}^{state\ i} = d_{rs}^{state\ i-1} \cdot \left(\frac{\mu_{rs}^{state\ i}}{\mu_{rs}^{state\ i-1}} \right)^{\beta_{state\ i}} \tag{2}$$

From (2), the demand level at the equilibrium state i depends explicitly on the existing demand level for the same OD pair at the previous equilibrium state (i -1), the proportion between the travel costs at the equilibrium state i ($\mu_{rs}^{state\ i}$)and state i-1 ($\mu_{rs}^{state\ i-1}$), and the elasticity for that state ($\beta_{state\ i}$). Although this linkage can not be claimed as a plausible representation of the effect of traffic demand by land use change, it should be a simple approach to capture some of these land use effects.

During each time interval different constraints on the cordon design can be defined. For instance, the complexity of a charging scheme can be increased over time by allowing additional cordons or toll points in some new areas of the network. Similarly, the level of constraint on different outcomes may be changed over time. This is indeed the benefit of allowing the adjustment of charging cordon overtime to reflect the condition in each time period. The objective functions of the optimal implementation-path of a charging cordon scheme will be the net present value of the social welfare over the whole 15 years:

$$F = \sum_{y=1}^{15} \frac{SW_y}{(1+r)^{y-1}}$$

, where SW_y is the social welfare in year y and r is the interest rate. Note that the costs of the scheme comprise of (i) cost of implementing a new toll point; (ii) cost of operating a toll point; (iii) cost of removing a toll point.

2.3 Possible patterns of scheme modification

There are many possible patterns of cordon implementation. In this paper, three possible implementation patterns are considered:

(i) evolution of a single cordon

A single charging cordon will be implemented initially in year 1. Then the cordon structure and toll level can be modified in years 6 and 11. Several constraints can be imposed on the cordon structure in each time period, e.g. possible coverage area or distributional impact.

(ii) a single cordon + implementation of additional screenlines inside the cordon

This implementation pattern represents the possible increase in complexity of the charging scheme over time by introducing additional screenlines inside the cordon. The exact definition of a screenline will be given later in the next section. Under this implementation path, a single charging cordon will be implemented in year 6. The cordon will not be changed during the whole time-horizons. Two additional screenlines will be added to the scheme sequentially in years 6 and 11. The uniform tolls on the cordon and two screenlines will also be optimised for each time frame.

(iii) inner cordon + gradual introduction of an outer cordon

Similarly to the previous patter, this implementation path introduces more complexity to the charging scheme by implementing additional cordon. The first cordon implemented in year 1 will be referred to as an inner cordon. The second cordon which will be gradually implemented in years 6 and 11 will be called the outer cordon. The second cordon will surround the inner cordon. The outer cordon will be implemented in two stages. In the first stage of the implementation, an incomplete cordon will be implemented in year 6. At this stage, the outer cordon is not a complete cordon yet. Then in the second stage, the rest of the tolled links of the outer cordon will be implemented to complete the cordon in year 11.

3. EVOLUTIONARY OPTIMISATION ALGORITHM

Apart from the complexity of the equilibrium constraint of the problem stated in (1), the constraint on the closed cordon formation of the combination of tolled links adds further complexity to the problem. To author's knowledge, there exists no derivative based optimisation algorithm that is capable of solving the problem in (1). Sumalee (2004) and Zhang & Yang (2004) independently proposed heuristic optimisation algorithms for solving (1) using Genetics Algorithm (GA).

The foundations of GA described by Holland (1975) establish a great breakthrough in the application of GA in optimisation. The basic idea is that GA will evolve chromosomes over a number of generations as inspired by the evolution process described earlier. Initially, GA produces a set of *'chromosomes'*, which is the representation of the solution, for the current generation. The algorithm then evaluates each chromosome against a specified objective function to determine the *'fitness'* of that chromosome. The *'probabilistic selection process'* mimics the natural selection or natural survival process. The selection process ensures that all chromosomes have a possibility of being selected while the stronger chromosomes have higher probability to be selected. After selecting the survival chromosomes, the genes of a pair of survival chromosomes are combined to reproduce the offspring for the next generation (this is called the *'crossover process'*). This process represents the breeding and genetic inheritance transformation in nature. Then, the natural stochastic process is introduced via mutations that randomly *'mutate'* the value of the genes. The optimisation algorithm developed for tackling optimal implementation path problem is based on the idea of GA. Figure 2 depicts the overall process of the algorithm.

The main difficulty in developing the GA based algorithm for the optimal implementation path problem is the design of the chromosome structure representing the cordon design (and its implementation path). Before introducing the implementation step into the chromosome structure, we will firstly deal with the chromosome design of a single and double cordon.

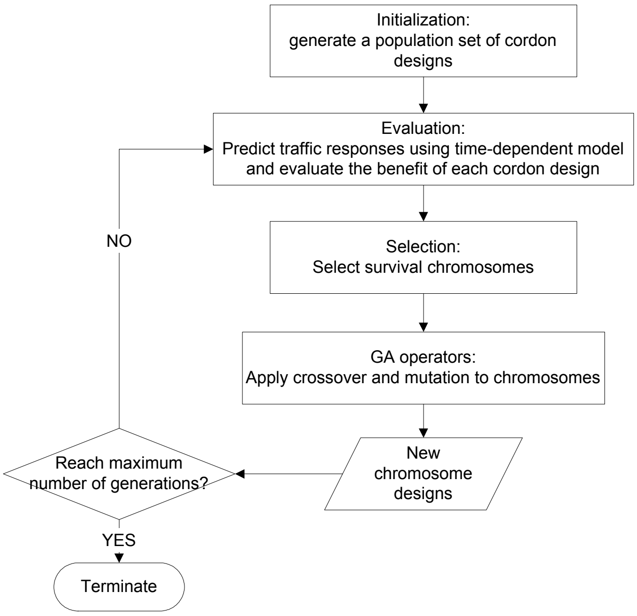


Figure 2. Process of the GA based optimisation algorithm

3.1 Standard chromosome structure for a single cordon and double cordon

A good chromosome design for a charging cordon should preserve the closed cordon format even after applying the crossover and mutation operators. This property is required to ensure that the searching process of GA is kept within the feasible region of the problem in (1). Sumalee (2004) developed a branch-tree framework for representing a charging cordon that ensures the cordon formation even after applying the crossover or mutation operations. This structure will be employed as the foundation of the development of the chromosome representation of the cordon implementation path in this paper. Some brief description of the branch-tree structure will be explained here. Those who are interested in the full detail of this method should consult Sumalee (2004).

A closed charging cordon, in the context of graph theory, is a cordon where all paths from all zones outside it connecting to nodes inside it are charged at least once on a link related to those paths. In other words, all car users driving into or passing through a designated area (charging area) are charged. This initial definition leads to a simple but rigorous application of graph theory to formulate a structure of a charging cordon as a tree. This idea will be exemplified using an example shown in Figure 3.

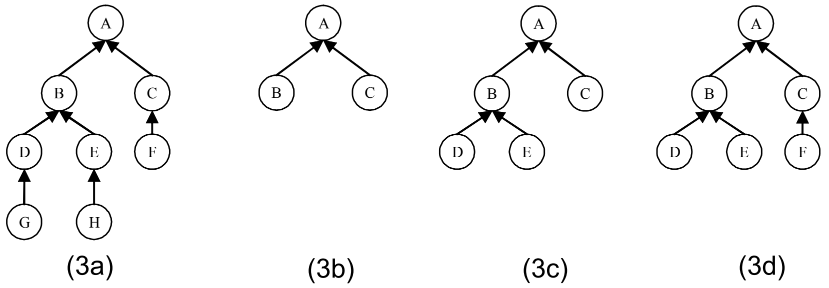


Figure 3: An example of branch-tree concept

Figure 3a above shows a full road network considered in this example. Assume that we are interested in constructing a charging cordon around node A (i.e. all paths accessing node A must be tolled). Figure 3b shows a sub-set of the full network in 3a. The tree in 3b was constructed by including all links entering node A into the tree. Tolling both links B-A and C-A ensures that all paths accessing node A are tolled. Thus, the tree shown in 3b can represent a charging cordon. Figure 3c shows another tree created from the tree in Figure 3b by including all links entering node B into the tree. Similarly, if links D-B, D-E, and C-A are all tolled then all paths accessing node A are also tolled. Thus, the tree in Figure 3c also represents a charging cordon. Figure 3d shows a tree adapted from the tree in Figure 3c by adding all links entering node C to the tree. Again, if links C-D, E-B, and C-F are tolled the tree in Figure 3d represents a closed cordon. This structure can actually be formalised. With an initial tree representing a charging cordon, one can build a new cordon based on this tree by including all links entering any ‘leaf nodes’ (nodes at the bottom of the tree) into the existing tree. Then, imposing the tolls on all links originated from the leaf nodes will create a charging cordon system.

In GA, a chromosome must be represented in a form of a string of numbers (or alphabets). The structure of the branch-tree explained above can be encapsulated into two strings of numbers. The first string is the node string which contains information about the nodes included in the tree. The second string is the degree string which contains the number of children nodes of that node in the tree (e.g. nodes D and E are children nodes of node B in Figure 3c). The structure of both string chromosomes also contains crucial information about the shape of the tree. Figure 4 shows an example of the chromosome representing a tree.

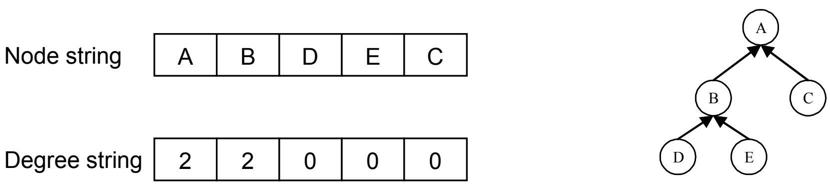
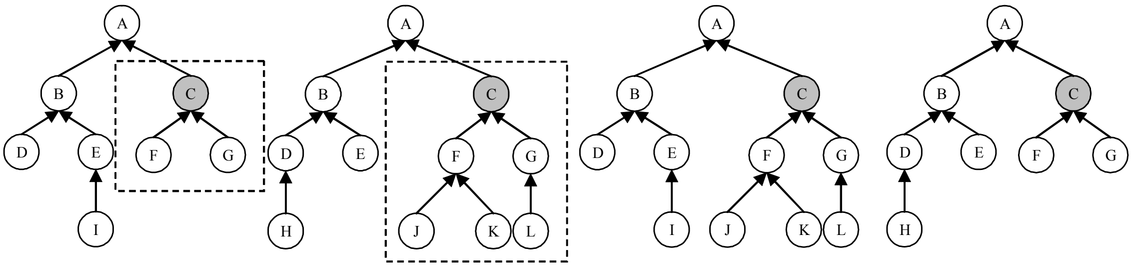


Figure 4: An example of chromosome strings

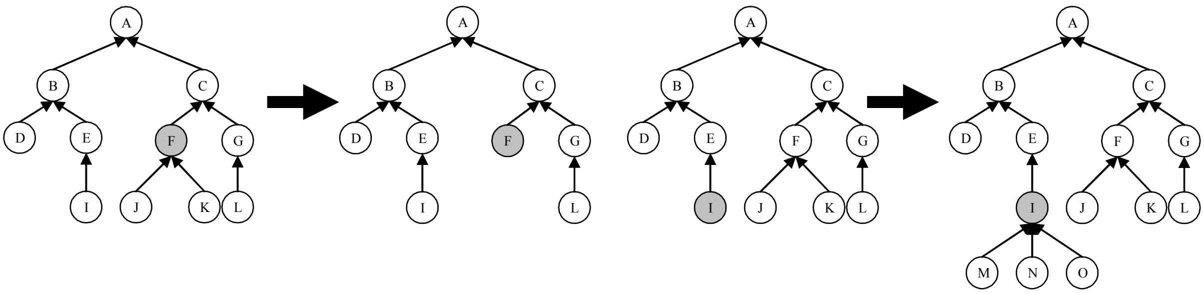
The crossover process will be applied to the chromosomes by (i) listing the common nodes in both mated trees, (ii) randomly choose a node for crossover, and (iii) swap the sub-tree rooted from that chosen node down to the leaf nodes between the two chromosomes. This crossing process with the branch-tree ensures the formation of a charging cordon of the newly created cordon. Figure 5 illustrates this process. In this example, the trees representing charging cordons shown in Figure 5a and 5b are mated for the crossover process. Presume that node C is randomly chosen as the crossover node. As explained above, the crossing parts from both tress are the sub-tree rooted from node C down to the leaf nodes. In this case, the sub-trees for the crossover are those parts in the boxes (dash-line). Then, the sub-trees from both chromosomes are exchanged producing two new chromosomes as shown in Figures 5c and 5d. Notice that the tree in Figure 5c is the tree in Figure 5a with the sub-tree under node C from Figure 5b. Similarly, the tree in Figure 5d is the tree in Figure 5b with the sub-tree under node C from Figure 5a.



(5a) (5b) (5c) (5d)

Figure 5: An example of crossover process

For the mutation process, a chromosome will be mutated by either expanding or shrinking the tree at a randomly selected node. The mutation process will operate as follow: (i) randomly check if the chromosome will be mutated; (ii) if so, randomly choose the node inside the tree to be mutated; (iii) if the selected node is a leaf node that node will be expanded, otherwise that node will be shrunk. Figure 6 exemplifies the mutation process. The tree in Figure 6a is selected to be mutated and assume that node F is randomly selected for the mutation. Since node F is not a leaf node, it will be shrunk during the mutation process. The shrinking process then removes all the children nodes and corresponding branches underneath node F from the original tree making node F the leaf node (with degree 0). The resulting tree from the shrinking process is shown in Figure 6b. On the other hand, assume that node I in the original tree (Figure 6c) is instead randomly chosen. In this case, as node I is the leaf node, it will be expanded. The resulting tree from the expanding process is shown in Figure 6d where all preceding nodes of node I in the original traffic network are included into the tree.



(6a) (6b) (6c) (6d)

Figure 6: Illustration of the mutation process (shrinking and expanding operations)

As mentioned in Section 2.3, the other possible scheme design includes the introduction of the outer cordon. The scheme with two cordons (inner and outer cordons) can be represented directly using the branch-tree structure with the pre-defined boundary of both cordons. This is illustrated in Figure 7 below. With this example network two cordon boundaries are defined (boundary 1 and 2). The area within the boundary 1 is the valid area of implementing the inner cordon whereas the area between the boundaries 1 and 2 is the area of the outer cordon. Two trees are required to represent the scheme with the inner and outer cordons. The structure of each of these trees follows the explanation above. The mutation and crossover will be applied separately to each cordon. The reason for restricting the possible areas of the inner and outer cordons is to ensure that the outer cordon always covers the inner cordon (even after the crossover and mutation).

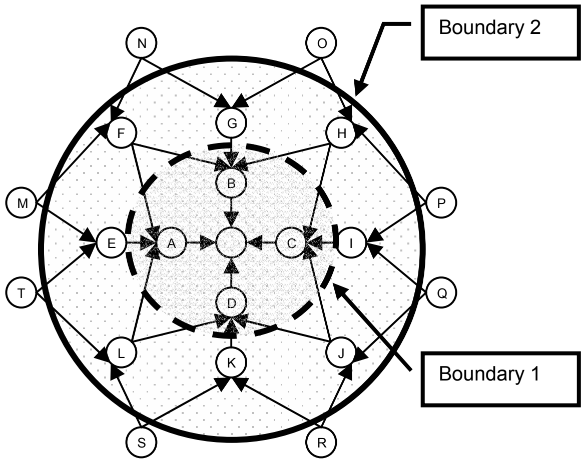
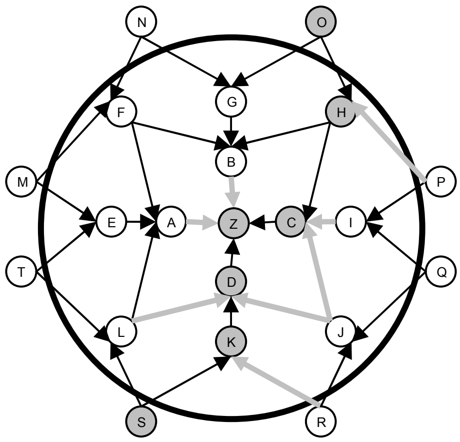


Figure 7: An example of the boundary definition for the scheme design with inner and outer cordons

3.2 Chromosome representation for screenline

This section presents a possible framework for representing a screenline as a chromosome. A screenline is a set of tolled point defining a boundary such that any trip passing it must be tolled. Normally, a screenline is delineated from one side of the borders of a charging cordon to another. An idea adopted here is to define a screenline as a series of connecting nodes (in either direction). The tolled links are then those links entering the nodes included in the series but not connecting between two nodes inside the series. Figure 8 is used to explain the idea.



S	K	D	Z	C	H	O

Figure 8: An example of the structure of a screenline and its chromosome representation

From this figure, a cordon is already in place and a screenline is defined by a series of the nodes in grey (nodes S, K, D, Z, C, H, and O). Notice that the series of nodes representing a screenline starts from and end at border nodes of the charging cordon. The grey links are those links entering one of the nodes in the series and do not connect between any two of the serial nodes. These are the tolled links of the screenline. Obviously, any trips traversing from one side of the screenline (defined by the series of the nodes) to the others will be tolled since all links entering these nodes are tolled. The chromosome representation of these serial nodes as a screenline is simply a string of node labels as shown in Figure 8.

The crossover process for two screenlines can be conducted by (i) generating the list of common nodes in both chromosomes; (ii) generate a set of feasible crossover partitions of the two mated chromosomes; these are node strings between two of the common nodes which do not have any of the common nodes in between them; (iii) if there are more than one feasible crossover partition then randomly selects one from the list; (iv) crossover the feasible partitions between two chromosomes.

For the mutation process, if the chromosome is selected for the mutation, the first stage involves randomly selecting the mutation node. Then, at the selected node a new sequence of the nodes representing the screenline will be randomly generated until the node chosen is one of the nodes in the original chromosome. For instance, from the example shown in Figure 8 assumes that node D is selected as the mutation node. At node D assume that node L is chosen as the next node in the series. Then, at node L node A is chosen, and at node A node Z is chosen. Since node Z is an existing node in the original chromosome, the mutation process terminates. The mutation process will then include all the nodes in the original chromosome after node Z into the new chromosome. Thus, the new chromosome will be S-K-D-L-A-Z-C-H-D. Notice that the part of the chromosome after node D until node Z is a new part generated by the mutation process (D-L-A-Z), and the parts S-K-D and Z-C-H-D is the part from the original chromosome.

3.3 Chromosome representation of time-dependent charging scheme design

In Section 3.1 and 3.2, the chromosome representations of the three main components of a charging scheme are described (inner cordon, outer cordon, and screenline). These components will be associated with its implementation stage (year). A chromosome representing the total design of the scheme is simply a stack of these chromosomes. A stack of three chromosomes will be used for representing the implementation of a single cordon over time (see Figure 9a). Notice that two strings are required for representing one cordon (node and degree strings). For instance, the first two strings compose the chromosome for the inner cordon implemented in year 1.

For the implementation of a single cordon with additional screenlines a stack of three chromosomes will also be employed but the second and third chromosomes represent two screenlines to be implemented in year 6 and 11 respectively (rows 3 and 4 in Figure 9b). Note that only one string is required for representing a screenline. The third pattern of implementation path is the introduction of the inner cordon and gradual introduction of the outer cordon. Figure 9c shows the stack chromosome structure for this design. Again, the first two rows contain information about the inner cordon (implemented in year 1). The outer cordon is then represented in the third and fourth rows (in grey) which have the same structure as the inner cordon (see Section 3.1). The last row contains the information on the implementation stage of each toll link of the outer cordon. The value of 1 indicates that all tolled links originated from the node in the corresponding column of the outer cordon chromosome will be implemented in year 6. Otherwise, they will be implemented in year 11.

The stack chromosomes which involve only the outer and inner cordons (Figure 9a and 9c) can be applied directly with the crossover and mutation techniques described earlier.

However, the crossover and mutation processes for the stack chromosome with screenlines need some modification to ensure the consistency of the scheme (e.g. the screenline does not start outside the cordon or does not start from and end at one of the leaf nodes). To resolve the possible conflicts, after the crossover over and/or mutation is applied to the chromosomes, the process for checking the consistency of the screenlines will be conducted.

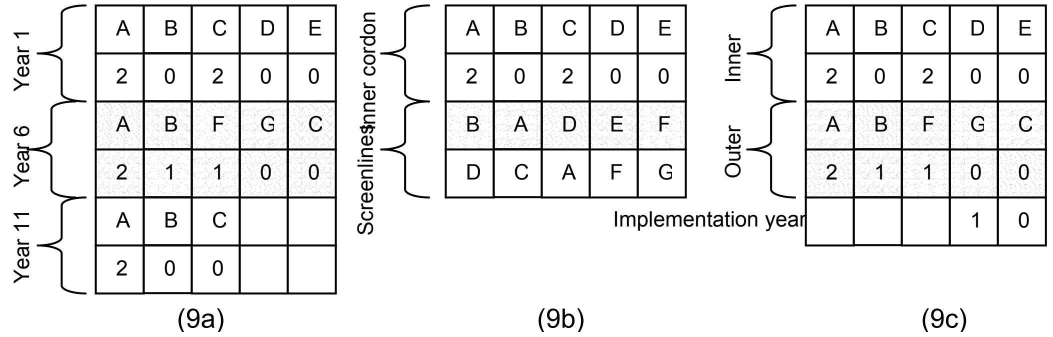


Figure 9: Examples of the stack chromosome structure for the scheme implementation paths

The process will divide the screenline into different parts and check if that part is (i) feasible part, (ii) infeasible part, or (iii) fixable part. The process will divide the screenline so that there is no border node in the middle of the node string. The feasible part is simply a series of nodes started from a boundary and ended at a boundary node and is inside the cordon. A part of a screenline is infeasible when it is outside the cordon. The fixable part is the part of a screenline that is inside a cordon but does not started and/or ended at one of the border nodes. Figure 10a shows example of different types of a part of screenline. Figure 10b shows an example of a screenline that comprises of different parts. The consistency checking process will identify four different parts of this screenline: Q-J-C-Z-D-K (feasible part), K-S-L-T-E (infeasible part), E-A-F (feasible part), and F-B (fixable part). The checking process will then randomly select one part from the set of feasible and fixable parts. If a fixable part is chosen, the process will then complete the screenline by randomly generate a sequence of nodes until reaching one of the border node.

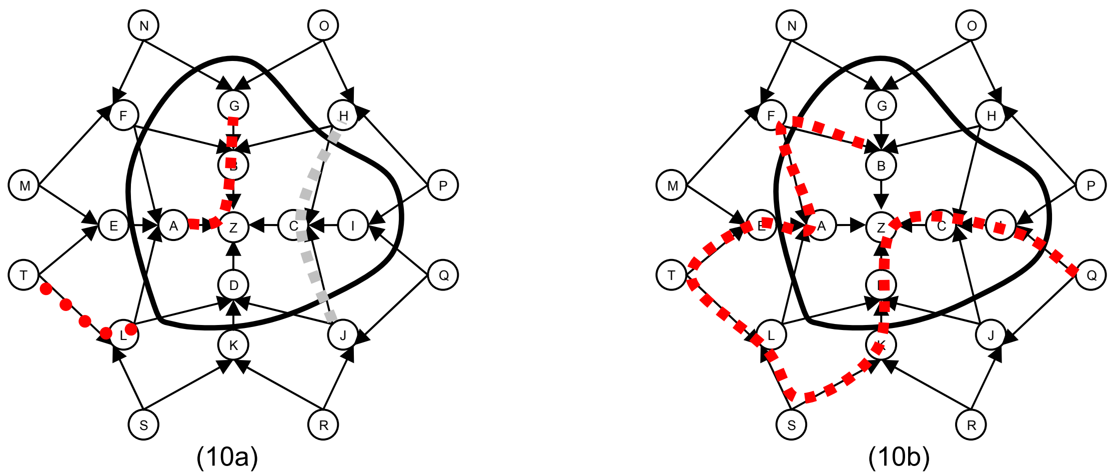


Figure 10: Inconsistency problems with screenline

4. NUMERICAL EXPERIMENTS

This section presents some test results of the algorithm developed in the previous section. The tests are conducted with the network of the Edinburgh city (see Figure 11). The elasticity, value of time, and vehicle running cost are assumed to be -0.57, 7.63 p/min, and 5.27 p/km respectively throughout the planning period. The exogenous growth of traffic is assumed to be around 5% in years 6 and 11. A single AM peak traffic model is used to represent each equilibrium state of the traffic in each year. In order to take account of the discounting effect of the costs and benefits of the scheme, the benefit of the road pricing scheme will be converted to a benefit per annual using an assumption that the scheme will gain the same level of benefit for 4 hours a day and 260 days a year. The discount rate is assumed to be 6%. The costs of the scheme operation and implementation are assumed to follow the recent report of Catling (2001) which estimated the operation and implementation cost per toll point of a cordon scheme to be £m 12.1 and £m 5.4 per year respectively. The cost of removal a toll point was not reported in this document. Instead, we assume this to be equal to the labour costs for the implementation which is around £50k per toll point.

Table 1 shows the test results where each pattern of implementation path is optimised. In addition, the table also includes the performances of the schemes created by the "naïve" method which simply chooses the best design in each year for each pattern from the optimisation results. The first, second, and third columns in Table 1 show the benefits of each scheme incurred in years 1, 6, and 11 respectively. The last column shows the net present value (NPV) of the benefit of the scheme over 15 years. Note that additional constraint is imposed on the design of the inner + outer cordon in which only half of the tolled points of the outer cordon can be implemented in year 6. The possible charging area of the inner cordon is defined roughly as the area inside the outer ring road for the evolving cordon and inner cordon + screenlines schemes. On the other hand, for the inner + outer cordons scheme the chargeable area of the inner cordon is slightly smaller and the outer cordon can be placed anywhere outside the chargeable area of the inner cordon.

Table 1: Tests results with the Edinburgh network

Scheme	Benefit Y1 (£m)	Benefit Y6 (£m)	Benefit Y11 (£m)	NPV of benefit (£m)
Evolving inner cordon	7.5	8.2	8.5	77.4
Inner cordon + screenlines	6.2	11.7	16.3	101.3
Inner + outer cordon	3.9	16.8	14.9	104.3
Naïve evolving inner cordon	7.5	7.2	6.4	69.3
Naïve inner cordon + screenlines	7.5	8.1	9.6	79.7
Naïve inner + outer cordon	3.9	14.5	11.2	88.4

From Table 1, the best implementation pattern is the inner + gradual outer cordon producing the NPV benefit of around £m 104.3 which is about 31% higher than the NPV benefit of the evolving inner cordon scheme. Obviously, this is mainly due to the larger chargeable area of the inner + outer cordons scheme compared to those of the other two patterns. The inner cordon + screenlines scheme generates around £m 101.3. Again, this is higher than the benefit from the evolving inner cordon which is mainly due to the higher degree of freedom of the design based on the inner cordon and screenlines (higher number of tolled links).

Figure 11 shows the locations of the inner cordons from the evolving inner cordon scheme. The initial cordon implemented in year 1 is actually the optimal cordon scheme for the problem considered in Sumalee (2004b). In year 6, the cordon is extended to cover a wider area of the network in the southeast area of the city. This is a result of the diversion of the traffic. In year 1, the main congested area is the west part of the city; hence the cordon focuses on that area. In year 6, the congestion (and traffic demand) migrates from the west to the south and east parts of the city. This effect is then exacerbated by the exogenous demand growth causing a significantly higher level of congestion in the south and east part of the city. Thus, the cordon is moved to deter the congestion (and externality) in that area in year 6. The cordon is then slightly modified in year 11 to be more compact.

Figure 12 shows the optimal design for the inner cordon + screenlines scheme. The inner cordon is much larger than the inner cordon in Figure 11. This may be due to the possibility of the introduction of additional screenlines in which a wider cordon allows a longer screenline. Then, the first screenline added to the scheme is located to trap the traffic in the west part of the city. Figure 13 shows the inner + outer cordons scheme. Again most of the tolled points implemented in year 6 are on the west part of the city. One in particular is on the Forth Bridge.

When the inner + outer cordon scheme is constructed by the naïve method, the performance of the scheme decreases by 18% compared to that of the optimised design. Similarly, the benefits of the evolving inner cordon and inner cordon + screenlines schemes also decrease by 10.4% and 27% respectively. This clearly demonstrates the importance of the time-dependent based design.

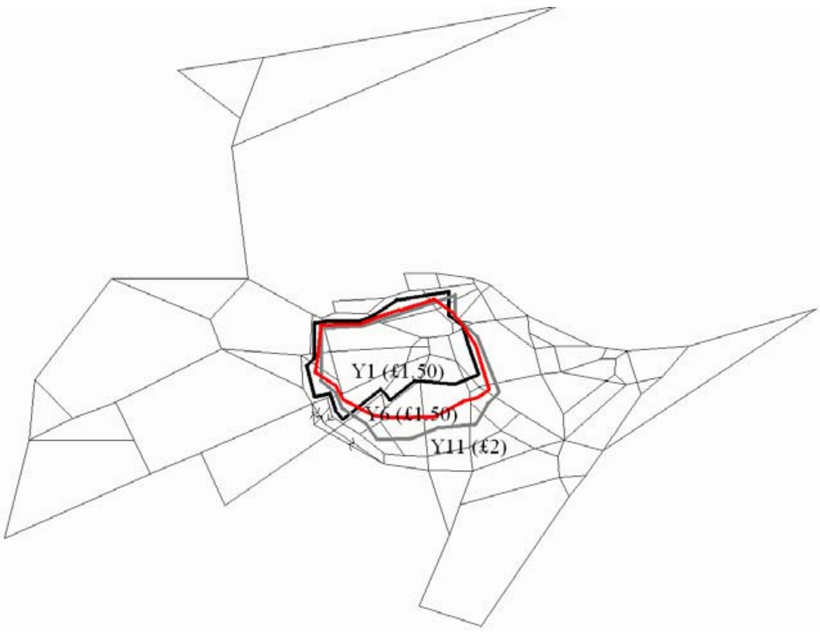


Figure 11: Optimal design of evolving inner cordon scheme

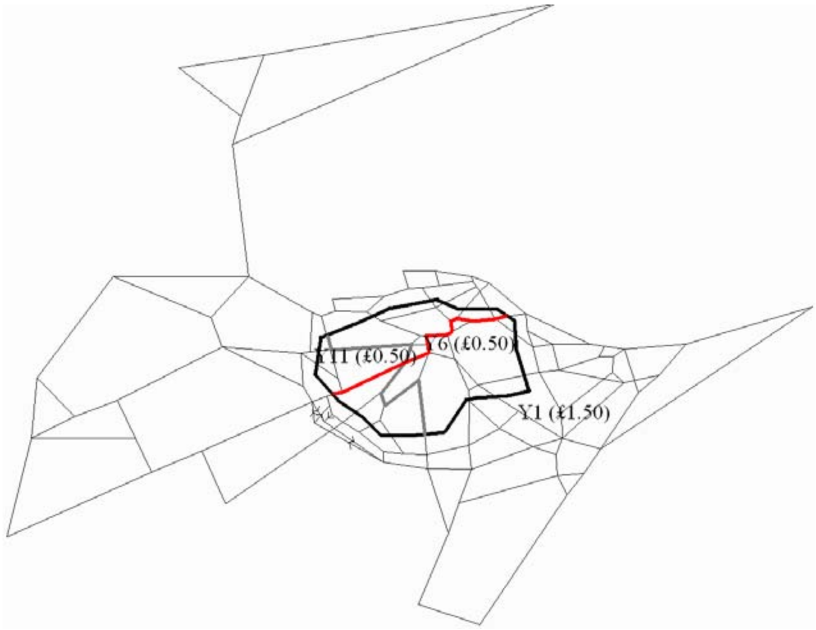


Figure 12: Inner cordon + screenlines

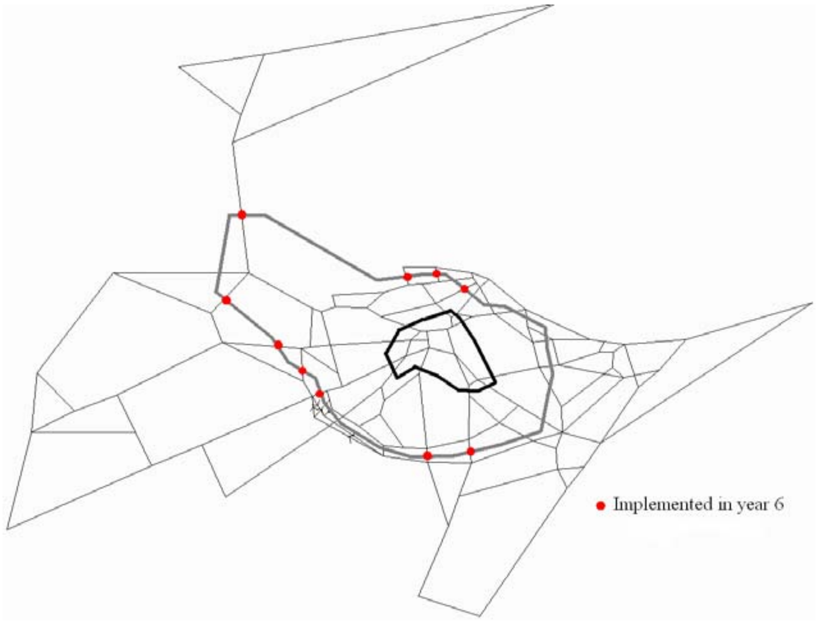


Figure 13: Inner + outer cordons

5. CONCLUSIONS

The main motivation of this research is due to the past experience and current development of a number of successful real-world road pricing schemes (e.g. Singapore and London schemes). Initially, these schemes were implemented in a simple format to ensure the users' friendliness and public acceptability. Then, the schemes have been evolved over the time adding more complex charging structure to adjust the schemes to more optimal designs. The paper proposes an optimisation algorithm for defining an optimal implementation path for a road user charging scheme. Three different implementation patters are considered including (i) evolution of a charging cordon, (ii) inner cordon + screenlines, and (iii) inner + outer cordons. All of these patterns have some specific features for the time-dependent implementation. A simple time-dependent model based on the equilibrium concept is adopted. The algorithms are tested with the city network of Edinburgh in the UK. With the setting of the model and other assumptions in this test, the results show that the best performance implementation path is the inner cordon and gradual introduction of the outer cordon. However, it should be noted that this implementation pattern is allowed to cover a wider area of the network compared to the other two. Three different implementation paths are also developed using the naïve method which simply combines the best design in each time period without considering the time-dependent factor. The results shows that ignoring the time-dependent factor in the design process may cause a substantial loss of the scheme benefit (up to 27%).

The work in this paper represents an initial step into the evolution design of a charging cordon scheme. Several future research issues are envisaged to be particularly useful. Firstly, the model adopted in this paper is a simple one and a more complex land use transport model is needed. The second issue is the consideration of the uncertainty of the future scenario into the present decision following the concept of real option theory. Rather than relying on some specific setting of modelling parameters, under this framework one can define distribution of these parameters and their variation over time. Thirdly, additional outcome constraints of the pricing scheme should be considered during the time-dependent design. Fourthly, a variety of implementation patterns should be investigated. Finally, one should consider a way to introduce the perception of the scheme complexity into the model consideration (e.g. constraint on the maximum number of toll crossing points between each OD movement) in order to ensure the users' friendliness of the scheme design.

REFERENCES

Catling, I. (2001), **Initial assessment of RUC scheme design options for Edinburgh**, Edinburgh City Council, Edinburgh.

Clegg, J., Smith, M., Xiang, Y. and Yarrow, R. (2001) Bilevel programming applied to optimising urban transportation, **Transportation Research Part B: Methodological***,* vol. 35, no. 1, pp. 41-70.

Holland, E. P. and Watson, P. L. (1978) Traffic restraint in Singapore: measuring the impacts of area license scheme, **Traffic Engineering And Control***,* vol. 19, pp. 14-17.

Holland, J. H. (1975) **Adaptation in Natural and Artificial Systems**, MIT Press, Ann Harbor, MI.

Hyman, G. and Mayhew, L. (2002) Optimizing the benefits of urban road user charging, **Transport Policy***,* vol. 9, pp. 189-207.

Lawphongpanich, S. and Hearn, D. W. (2004) An MPEC Approach to Second-Best Toll Pricing, **Mathematical Programming B***,* vol. 101, no. 1, pp. 33-55.

Luo, Z. Q., Pang, J. S. and Ralph, D. (1996) **Mathematical programs with equilibrium constraints**, Cambridge University Press, New York.

May, A. D., Liu, R., Shepherd, S. P. and Sumalee, A. (2002) The impact of cordon design on the performance of road pricing schemes, **Transport Policy***,* vol. 9, pp. 209-220.

Meng, Q., Yang, H. and Bell, M. G. H. (2001) An equivalent continuously differentiable model and a locally convergent algorithm for the continuous network design problem, **Transportation Research Part B**, vol. 35, no. 1, pp. 83-105.

Mun, S., Konishi, K. and Yoshikawa, K. (2001), 'Optimal Cordon Pricing', in **The Ninth World Conference on Transport Research**, Seoul, Korea.

Pigou, A. C. (1920) **Wealth and Welfare**, Macmillan, London.

Schade, J. and Schlag, B. (2003) Acceptability of urban transport pricing strategies, **Transportation Research Part F: Traffic Psychology and Behaviour**, vol. 6, no. 1, pp. 45-61.

Shepherd, S. P. and Sumalee, A. (2004) A Genetic Algorithm Based Approach to Optimal Toll Level and Location Problems, **Networks and Spatial Economics**, vol. 4, pp. 161-179.

Sumalee, A. (2001), **Analysing the Design Criteria of Charging Cordons**, Institute for Transport Studies, University of Leeds, Leeds, ITS working paper no 560.

Sumalee, A. (2003) Optimal toll ring design with spatial equity impact constraint: an evolutionary approach, **Journal of Eastern Asia Society for Transportation Studies**, vol. 5, pp. 1813-1828.

Sumalee, A. (2004) Optimal Road User Charging Cordon Design: A Heuristic Optimisation Approach, **Computer-Aided Civil and Infrastructure Engineering**, vol. 19, pp. 377-392.

Verhoef, E. T. (2002) Second-best congestion pricing in general networks: Heuristic algorithms for finding second-best optimal toll levels and toll points, **Transportation Research Part B**, vol. 36, no. 8, pp. 707-729.

Vickrey, W. S. (1969) Congestion theory and transport investment, **American Economic Review**, vol. 59, no. 2, pp. 251-260.

Wardrop, J. (1952) Some theoretical aspects of road traffic research, **Proc. of the Institute of Civil Engineers**, vol. 1, no. 2, pp. 325-362.

Zhang, X. and Yang, H. (2004) The optimal cordon-based network congestion pricing problem, **Transportation Research Part B**, vol. 38, no. 6, pp. 517-537.