Abstract: This paper addresses the optimal design problem of toll ring in a general urban traffic network. This problem is categorised as a Mathematical Programming Problem with Equilibrium Constraints (MPEC) due to the implicit constraint of user equilibrium condition. In addition, the constraint on the topology of the combination of tolled links is imposed where the combination of tolled links must form a toll ring. A Genetic Algorithms (GA) approach used to tackle this topological constraint is based on Sumalee (2003). In this study, the approach is extended to take account of the potential spatial equity impact. The Gini coefficient, invented for the analysis of income distribution, is adopted as the measure of the equity impact. The toll ring design must ensure that the Gini coefficient is lower than the pre-specified criteria. The modified GA-based method is tested with a network of the city of Edinburgh in UK.

Key Words: road pricing, optimal toll location, equity, genetic algorithms, constraint handling

1.INTRODUCTION

The concept of road pricing is a long-established fiscal mechanism aimed to control the demand of road usage (Pigou, 1920 and Knight, 1924). The idea is proposed that by alleviating the appropriate tolls the traffic level and distribution will organise itself in an optimal way. The “optimal way” is regularly referred to the traffic pattern and demand level that maximise the social welfare. The problem of finding the optimal tolls when only a subset of links is to be tolled (called the second-best optimal toll problem) is a very complex problem. Since the emergence of the concept of the second-best optimal toll problem, many economists have illustrated the approach to quantify the optimal toll level. (Levy-Lambert, 1968; Marchand, 1968; Liu and McDonald, 1999; and Verhoef, 2002). However, it should be noted that most of the studies only concentrate on the theoretical construction of the second-best marginal cost toll with a very simplified traffic network.

In reality, the design and implementation of the scheme usually gives more weight on the issues of acceptability and practicality rather than the economic benefit (May et al 2002; Sumalee 2001). Sumalee (2001) identified some characteristics of the scheme design that are in favour of the public and practitioner. The main concept is to implement a simple system, i.e. charging cordon system with a uniform toll. This closed cordon format is believed to be a user-friendly format where people can understand easily and also response to the tolls more efficiently. As the shadow of the belief in the potential progressive influence of road user charging scheme to the transport system, the potential equity problem, both in vertical and horizontal dimensions across different groups of population in the society, is always an opposing question toward the justification of the scheme. The “vertical dimension” of the equity issue can be seen as an unequal impact from the scheme across the different groups of the population segregated by income level, sex, available alternative to car, age, or even race. For instance, one may argue that
the implementation of road user charging system will benefit the rich whilst worse-off and exclude the poor (or lower income level group). However, the model adopted in this paper does not permit us to analyse this issue yet due to its assumption on the homogeneous users. Nevertheless, “the horizontal dimension” of the equity impact is also a very important issue and will be considered in this study. The horizontal equity impact is referred as the “spatial equity impact”. The spatial equity impact can be described as the distribution of the benefits and costs of the scheme across the population from different areas in the network. If the scheme benefits only a small group of people from some areas, but the rest of the population experience the decline in the social welfare, the scheme can be argued as inequitable policy. Thus, the design of a toll ring must ensure that the spatial equity impact is carefully considered and kept minimum.

Recently, there have been an increasing number of the attempts to develop analytical approaches to the optimal toll location problem. May, et al (2002), Verhoef (2002) and Shepherd and Sumalee (2002) tackled the problem of the optimal toll location and toll level problem for a detail network under user equilibrium condition (but the toll point may not form a toll ring). Mun, et al (2001) and Hyman and Meyhew (2002), in contrast, developed approaches to define the optimal toll ring but with a very aggregate representation of the traffic network. Yang, et al (2002) and Sumalee (2003) independently developed evolutionary based optimisation algorithms for computing an optimal toll ring with detail traffic networks. Nevertheless, very few studies have considered the equity impact in designing optimal toll rings. Meng and Yang (2002) recently proposed a framework to considering the spatial distribution impact of the change in a traffic network structure (i.e. network design problem). They proposed a measure which is the ratio of the origin-destination (O-D) travel cost in the “before” and “after” scenarios. Also, Yang and Zhang (2002) use the same indicator to study in the toll design problem with multiple user classes. This measure cannot capture the lost in consumer surplus from the depressed demand. Thus, it is not an appropriate measure for the policy expected to depress the travel demand significantly, like road toll.

In this study, the measure for evaluating the spatial equity impact is the Gini coefficient which is a long established measure in evaluating the income distribution in the population. The detail of this measure will be discussed later on. The aim of this paper is to develop an optimisation method for defining an (heuristic) optimal location of the toll ring while the equity constraints is satisfied. This paper extends the Genetic Algorithms (GA) based method proposed in Sumalee (2003) to handle the constraints during the optimisation process. The detail of the constraint handling method will be discussed in the detail in §5. The paper is structured into seven sections. The next section explains the mathematical formulation of the optimal toll ring problem. The third section describes the detail of the adopted equity measure, Gini coefficient. The fourth section, then, presents an innovative approach to define a set of toll rings based on the branch-tree concept following Sumalee (2003). The fifth section introduces the concept of genetic algorithms (GA) and explains the optimisation algorithm developed for finding a constrained optimal toll ring. Section six presents the numerical results for the test of the algorithm with the Edinburgh network. The last section concludes the paper.

2. OPTIMAL ROAD PRICING DESIGN PROBLEM

The problem discussed in this paper is to find the optimal location of tolled links forming a toll ring with their optimal toll levels so as to optimise the social welfare of society and satisfying the equity constraint. This problem is, thus, to define three aspects of toll rings including i) optimal number of tolled points, ii) locations of optimal toll points and iii) optimal uniform toll levels on
selected toll points. This problem can be formulated as following (see Appendix A for the notations):

$$\max_{(\tau, \epsilon)} W(\mathbf{T}, \mathbf{F}, \tau, \epsilon)$$

s.t. \((\tau, \epsilon) \in \Psi\)

\((\mathbf{T}, \mathbf{F}) \in X_{ad} \subseteq \mathbb{R}^I \times \mathbb{R}^p\)

\(h(\mathbf{T}, \mathbf{F}, \tau, \epsilon) \leq 0\)

(1)

where the objective function of this optimisation program \(W\) is to maximise the net benefit which is the social welfare function (total benefit) following Marshallian’s rule measure deducted the costs of the road pricing system (i.e. implementation and operation costs):

$$W(\mathbf{T}, \mathbf{F}, \tau) = \sum_{i} \int_{0}^{y} D_i(x) dx - \sum_{j} \sum_{p} \delta_{ip} F_p \cdot c_j - \sum_{j} \epsilon_j s_j$$

(2)

\(X_{ad}\) denotes the feasible space for path flows which is assumed to be the user-equilibrium path flows satisfying flow conservation constraint. The rational road users’ responses to the tolls imposed is assumed to follow Wardrop’s user equilibrium (Wardrop, 1952) which can be defined as a variational inequality (VI) (cf Smith, 1979). Let \(K\) denotes the set of flows satisfying the flow conservation constraint. Therefore, \((\mathbf{T}, \mathbf{F}) \in X_{ad} \iff (\mathbf{T}, \mathbf{F}) \rightarrow sol(VI)\) and \(\mathbf{F} \in K, \mathbf{T} \in M\) where \(K\) and \(M\) is the set of feasible path and O-D flows. In the other words, the path flows will be a feasible solution of the optimisation problem iff it solves the VI and satisfies the flow conservation constraints. The VI defined by two pair of vectors, \(\mathbf{C}\) and \(\mathbf{F}\), and \(\mathbf{D}\) and \(\mathbf{T}\) can be formulated as follows:

$$\mathbf{C}(\mathbf{F})^{\top} \cdot (\mathbf{F}' - \mathbf{F}) - \mathbf{D}(\mathbf{T})^{\top} \cdot (\mathbf{T}' - \mathbf{T}) \geq 0 \quad \forall \mathbf{F}' \in K; \forall \mathbf{T}' \in M$$

(3)

This UE condition is imposed as constraints for the optimal road pricing design problem. In this paper, the single user class and separable cost function is assumed. In addition, the monotone condition for the cost function is assumed to guarantee the uniqueness of the link flow equilibrium solution. As mentioned, the main objective of the paper is to solve the optimal toll ring. Let \(A\) and \(B\) denote the feasible set of toll levels and a set of toll rings. Therefore, \(\psi = A \cap B\).

This means that the feasible combination of toll point \((\tau, \epsilon)\) must be the toll point forming a toll ring with feasible uniform toll. The last set of constraints for the optimisation program is the equity impact constraints defined by \(h(\mathbf{T}, \mathbf{F}, \tau, \epsilon) \leq 0\) which will be discussed in the detail in the next section. The main barriers in solving this optimisation problem are the existence of VI and the complex structure of the topological constraint of a toll ring. To author’s knowledge, there does not exist any derivative based optimisation approach to the problem stated in formula (1).

In this paper, the derivative free approach is utilised to solve the problem. The idea to tackle this problem is to use the flexibility of Genetic Algorithms (GA) in dealing with the constraint of closed cordon. Sumalee (2003) proposed the GA-AS method which can be used to solve the optimal toll ring problem based on the branch-tree framework of the graph in the network. The description of this method will be discussed in §4 and §5. The improvement of the method presented in this paper is the additional mechanism to deal with the equity impact constraint which will be discussed later in §5.4. Therefore, the algorithm proposed in this paper is not only able to find an optimal toll ring maximising the economic benefit, but also generate a toll ring that satisfies the constraint on the level of equity impact.
3. MEASURE FOR EVALUATING THE SPATIAL EQUITY IMPACT

Equity, like the related concepts of justice, fairness and right, is not a simple thing. Different people have different concepts of equity, but also, which of the aspects of equity that seems important will depend very much on the particular context and circumstances (Langmyhr 1997).

In this paper, the spatial distribution of the benefits and costs (in the form of social welfare improvement) of the toll ring scheme is considered. There exists the economic framework that can be utilised in analysing the spatial equity effect which is the indicator for the income inequality. The Gini coefficient is the most commonly used income inequality measure. In fact, Fridström, et al (2000) used the same measure to analyse the distribution effect of various transport policies. It can be explained with reference to figure 1 below. On the horizontal axis, a population is ordered by income from the lowest to the highest. On the vertical axis is the cumulative share of total income. The ideal situation or the most equitable case is that everybody had the same income and this will produce the straight line. In reality, the cumulative income distribution will not pose a straight line. This is shown by the “Empirical distribution” curve or called “Lorentz curve”. The area between the two curves represents the skew of the actual income distribution from the ideal equitable distribution. Thus, this area can be used as an indicator of income inequality, ranging from 0 for perfectly equal distributions to 0.5 for distributions where one person earns all income (which is the area of the triangle under the equitable income distribution line). The Gini coefficient is twice this area to get a measure of inequality between 0 and 1. Recall that the aggregate benefit measure of the toll ring design is the Marshallian measure:

$$\tilde{W}(F, \tau) = \sum_i \int_0^T D_i(x, \tau) dx - \sum_f \sum_p \delta_{jp} \cdot F_p \cdot c_j$$

The change in social welfare for each O-D pair is used as the measure for the distribution impact. Thus, the disaggregate form of the aggregate Marshallian measure (see Formula 5) will be used as the measure of the benefits and impacts distribution for traveller group from that O-D pair where the superscript 0 and 1 denote the no-toll and toll scenarios:

$$\tilde{W}_1(F, \tau) - \tilde{W}_0(F, \tau) = \left( \int_0^T D_i(x, \tau) dx - \sum_f \sum_p \Delta_{jp} \cdot F^{11}_p \cdot c_j \right) - \left( \int_0^T D_i(x, \tau) dx - \sum_f \sum_p \Delta_{jp} \cdot F^{00}_p \cdot c_j \right)$$

The change in social welfare for each O-D pair is used as the measure for the distribution impact.
Here we have assumed the average social welfare improvement is $\bar{W}$. There are totally $I$ O-D pairs with social welfare improvement $W = (\bar{W}_1, ..., \bar{W}_I)$, $T_i$ members of group $i$ is the total number of travel demand (i.e. number of trips) in the un-tolled case, $i = 1, ..., I$ and $\sum_i T_i = T$ is the total travel demand in the whole network. However, the social welfare improvement can be both negative and positive. This contrasts to the sign of the income (used in the calculation of the original Gini coefficient) which can be only positive. Thus, the raw value of the social welfare improvement for each O-D pair will need to be re-scaled before calculating the Gini coefficient following formula 6. Let $\bar{W}_k$ be the lowest value of the social welfare improvement (from O-D pair $k$) from all values (which is normally negative). Then, the social welfare improvement for the other O-D pairs can be re-scaled as follows:

$$\bar{W}_i^* = \bar{W}_i + |\bar{W}_k| + \kappa \quad \forall i$$

where $\kappa$ is the given constant. Also, the average social welfare improvement will be replaced by the average of the re-scaled social welfare improvement. It is easy to verify that this re-scaling will not change the scale of the Gini coefficient since the main purpose is to measure the area under the equitable curve. As mentioned, the value of the Gini coefficient is in between “0-1”. The lower the Gini coefficient, the more equitable the toll pricing scheme. The constraint of the maximum value of the Gini coefficient will be used in order to constrain the spatial distribution impact (i.e. $Gini \leq \phi$; where $\phi$ represents the acceptable equity distribution value). It should be noted that using Gini coefficient implicitly assumes the normative view of the planner on what the equitable system should be.

4. BRANCH-TREE FRAMEWORK FOR TOLL RING FORMULATION

Firstly, the definition of toll ring must be clarified. A toll ring is a set of tolled points surrounding a designated area so that all travellers passing the area will be tolled. The other explanation of the toll ring in the context of graph theory is that all paths from all zones outside the toll ring passing through one of the nodes inside the toll ring must be tolled at least once on a link related to those paths. The latter concept of the toll ring formulation based on graph theory will be used next to relate the branch-tree concept to closed cordon formulation.

4.1 Notation of the branch tree concept

Let $G(a,n)$ be a directed graph representing an urban traffic network where $a$ and $n$ is a set of links and nodes respectively in the graph. A link is defined by two nodes, $i$ and $j$ where $i \neq j; i, j \in n$. The direction of a link is from $i$ to $j$. $i$ is termed “the preceding node” of $j$. The number of preceding nodes for each node $a$ can be more than one. A set $\Xi_j = \{|i| i \text{ is the preceding node of } j \}$ is defined as a set of “all-preceding nodes” of node $j$ where $|\Xi_j|$ is the size of set $\Xi_j$ (the total number of preceding nodes of node $j$). Let $\beta_j = \{\{n,d\}\}$ be a branch-tree where $j$ is the root node of this branch and $d$ is a degree of node $n$ in the branch. The degree of a node is the number of children or preceding nodes of that node in the branch. Node with the degree of zero is termed “leaf node”. Figure 2 shows an example of a branch. The branch in Figure 1 consists of 5 nodes. The root node of this branch is node A. Node B and C are children
or descendants of node A. Similarly, node D and E are children of node B. Node A and B has degree of two. Node C, D, and E has no children (zero degree) hence they are ‘leaf’ nodes of the branch.

![Figure 2: Example of a branch tree](image)

Next, the branching process is introduced. The branching process is defined as a process to expand the current branch from leaf nodes. At node $n$ which is one of the leaf nodes, the branch can be expanded by changing the degree of node $n$ from zero to $|\Xi|$ and include all nodes in $\Xi$ into the branch. Figure 3 illustrates the branching process. Figure 3a shows the complete graph $G(a,n)$, i.e. a road network. In this example, the branch in Figure 3b is expanded at node C. By observing the full network in Figure 3a, nodes F and G are the preceding nodes of node C. Thus, these two nodes are added to the expanded branch after node C. Figure 3c shows the branch expanded where the expanded node is in gray (node C). The last notation that will be introduced is the sub-branch. Inside a branch tree, a number of sub-branches can be defined. Recall that $\beta = \{(n,d)\}$ is a branch tree. Given $i$ which is one of the node in $\beta$, $\tilde{\beta} \subset \beta$ is a sub-branch of $\beta$. Figure 4 shows $\tilde{\beta}_B$ which is a sub-branch of the branch in Figure 3b rooted from node B.

![Figure 3: Example of the branching process](image)

![Figure 4: Example of a sub-branch](image)

### 4.2 Relationship between branch tree and toll ring

After introducing the necessary notations used for the branch tree concept, this section explains the connection between the branch tree concept and the closed cordon formulation. For a given $\beta = \{(n,d)\}$, the tolled links are defined by a set of leaf nodes and their preceeding nodes in a branch. These tolled links will form a toll ring around the root note $j$ if and only if all nodes in $\beta = \{(n,d)\}$ have either $d = 0$ or $d = |\Xi|$. This proposition can be easily verified by considering the definition of a toll ring mentioned earlier. By tolling on all leaf links of a branch, all paths entering the root node are definitely tolled, since all paths entering the root node must have at least one of the leaf links as the relevant link. This condition satisfies the definition of a toll ring. In a network, a set of links forming a closed cordon around the tolled area must be predefined.
Figure 5a shows the hypothetical network used for exemplification. The grey node is assumed to be the city centre which is the tolled area. Cordon 1 is defined as the initial cordon. From this initial cordon, a virtual root node (name “C1”) is defined for the branch and the first level nodes in the branch are all preceding node of the links forming Cordon 1. Figure 5b shows the branch C1. The original branch in (5b) is then expanded at node E and G creating the new branches in (5c) and (5d). These new branches form the Cordon 2 and Cordon 3 shown in (5a). Note that nodes E-L, predefined by the user, are referred as “target nodes”. This notation will be used in the algorithm in the next section. For brevity, those who are interested in the further detail of the branch-tree framework is referred to Sumalee (2003).

Figure 5: Demonstration of the relationship between branch trees and closed cordons

5. APPLYING GA TO SOLVE THE OPTIMAL CLOSED CORDON

To solve the optimal toll ring design problem, GA is used to produce and evolve a set of toll rings encapsulated in the form of chromosomes (using the branch tree formulation). An optimal uniform toll for each chromosome (toll ring) will be found by testing the set of tolled links with different predefined toll levels. As stated in formula (1), an optimal toll ring location and associated uniform toll must be defined under the equilibrium condition stated in formula (3). In the framework of GA, each of the toll rings and uniform tolls will be evaluated its fitness (or net social welfare improvement following formula 2) under the user equilibrium condition. Any traffic modelling software can be used in this stage. In this paper, SATURN (Van Vliet, 1982) is used to find the effect and benefit of each toll ring design. The process will be terminated by the predefined number of generations (given by user’s input). This algorithm, as proposed by Sumalee (2003), is termed GA-AS. Figure 6 depicts the overall process of GA-AS.

5.1 Chromosome design

It is crucial to design a chromosome structure that is compatible with the structure of branch-tree explained in the previous section. More importantly, the chromosome structure should be able to maintain the feasibility of the solution (in this case, the closed cordon format) even after
applying crossover and mutation operators. As mentioned in the previous section, the members of a branch set have two key characteristics, i.e. their node numbers and degrees. Thus, two strings will be used to represent a chromosome. The first string, termed node string, will contain the node numbers of a branch. The second string, termed degree string, will contain the degree of each node in the corresponding column of node string. Figure 7 shows an example of this chromosome structure.

![Figure 7: Chromosome structure of a branch tree](image)

The node and degree strings tell us that this branch comprises of nodes A, B, D, E, and C with the degrees of 2, 2, 0, 0, and 0 in that order. This chromosome represents the branch shown in figure 2. However, from the node and degree strings, the information provided is only which nodes are in this branch and the degrees of each node. They do not provide information about the connectivity of the nodes in the branch. For example, how one may know that node B is the preceding node of node A without looking at figure 2. The answer is in the order of the nodes in the node string stated in the algorithm A1 in Appendix B.

5.2 Initialisation

The encoding method explained is used to create the set of closed cordons in this section. The generation of a closed cordon is designed as a random process. The variable Prop is defined as the probability of a node to be expanded (recall the definition of branching process explained...
earlier). Also, the user must define the initial closed cordon by defining a set of links. The preceding nodes of these links will become a set of target nodes (see Figure 5).

5.3 Evaluation process

Once the chromosome is generated, the next task is to evaluate its fitness. The fitness is measured according to the objective function stated in formula (2). In this paper the simple uniform toll regime is assumed. This assumption is consistent with the judgmental design criteria mentioned earlier. In the numerical test presented later in §6, eight possible toll levels are predefined: £0.50, £0.75, £1.00, £1.25, £1.50, £2, £3 and £4. Each chromosome (representing a charging cordon) will be evaluated with each toll level. If the equity constraint is imposed, the algorithm will also calculate the Gini coefficient for each toll level. The toll level producing the highest objective function with the acceptable value of Gini (if the constraint is activated) will be chosen as an optimal uniform toll of that cordon. The objective function associated with that optimal toll is then used as the fitness of that chromosome.

5.4 Constraint handling

The particular property of GA is its ability to solve an arbitrary problem. GA does not require any information about the gradient or even the explicit functional form of the objectives or constraints in the optimisation program. As explained earlier, this advantage has already been exploited in this paper in dealing with the topological constraint for the set of tolled points. In this section, the constraint on the equity impact will be included in the optimisation process. There are various methods in dealing with the constraints in GA. The discussion of the integrity of each method is beyond the context of this paper. The main mechanism used to handle the constraints in this paper is to impose the penalty to the fitness value. The algorithm will generate potential solutions without considering the constraints and then penalise those violating the constraints by decreasing the “fitness” of the chromosomes. In this way, a constrained problem is transformed to an unconstrained problem by associating a scalar penalty with all constraint violations. Let \( f(x) \) be the objective function of the original optimisation problem (without lost of the generality, assume maximisation program) with a set of constraints; this constrained optimisation problem can be transformed by changing the objective function to:

\[
f(x) - \alpha \sum_{i=1}^{l} \omega_i \tag{8}
\]

where \( l \) is the number of constraints, \( \alpha \) is a penalty coefficient, and \( \omega_i \) is a penalty related to the constraint \( i \). In using this penalty-based method, one must carefully consider the choice of \( \omega_i \). If one chooses a too low value of \( \omega_i \), one may run the risk of spending too much time to evaluate illegal chromosomes in the GA process. On the other hand, if the chosen value of \( \omega_i \) is too high, one may lose some useful information about the unfeasible area which can lead to a better solution. In this paper, the method of “dynamic self adaptive penalty” proposed by Richardson, et al (1989) is adopted. The modified objective function is as follows:

\[
f(x) - \left[ k \cdot \left( \frac{G}{g} \right)^\rho \cdot \bar{f}_{g-1} \cdot \sum_i v_i \right] \tag{9}
\]

where \( G \) is the total number of generations tested, \( g \) is the current generation number, \( \rho \) and \( k \) are parameters, and \( v_i \) returns “the degree of constraint violation” for constraint \( i \). Thus from formula (2) the modified objective function becomes:
\[ \tilde{W} = W - \left( k \cdot \frac{g}{G} \cdot f_{g-1} \cdot v \right) \]

where

\[ v = \begin{cases} 0 & \text{if } \phi > Gini \\ \phi - Gini & \text{if } \phi < Gini \end{cases} \]

### 5.5 Selection process

The selection process used in GA-AS is based on “stochastic universal sampling” which uses a single wheel spin (Michalewicz, 1992). The so called “roulette wheel” is constructed where each slot represents a chromosome. In the original form of roulette wheel, the slots are sized according to the fitness of each chromosome which represents the probability of a chromosome to be selected. However, the fitness value (total benefits net of costs) can be negative which causes a problem for allocating the space for each chromosome on the wheel. The linear ranking approach proposed by Whitley (1989) is adopted to solve this problem. The slots in the roulette wheel are sized according to the chromosome at rank \( i \), where the first is the best chromosome (based on the fitness value), by the following equation:

\[ P_i = \frac{1}{|P|} \left( 2 - c \cdot (2c - 2) \cdot \left( \frac{|P| - i}{|P| - 1} \right) \right) \]

\( |P| \) is the size of the population (set \( P \), and \( 1 \leq c \leq 2 \) is “the selection bias”: higher values of \( c \) cause the system to focus more on selecting only the better individuals. The best individual in the population is thus selected with the probability \( \frac{c}{|P|} \); the worst individual is selected with the probability \( \frac{2 - c}{|P|} \). After each chromosome is assigned its probability to be selected, the next step is to calculate a cumulative probability \( (q_i) \) for each chromosome:

\[ q_i = \sum_{j=1}^{i} p_j \]  

Then, each time a string chromosome is selected for the new generation by generating a random number \( r \) from the range [0..1]. If \( r < q_i \) then select the first chromosome; otherwise select the \( i \)-th chromosome such that \( q_{i-1} < r < q_i \). Indeed, some chromosomes would be selected more than once according to the selection probability of each chromosome. As part of the selection process, the idea of “elitism” is also adopted to ensure that the best chromosome in the current generation will be included in the population of the next generation.

### 5.6 Crossover process

The chromosome structure in GA-AS is in fact very similar to the chromosome used in Genetic Programming (GP) which is a branch. Thus, it is natural to adopt the crossover process currently used in GP type of algorithm. The normal process of crossover in GP is to cross sub-branches below chosen nodes in two mated chromosomes. The complication involved in crossing the chromosomes in GA-AS is the strict structure of branch tree. The process has to start by identifying identical nodes in two mated branches. Then, the crossing node is randomly chosen from the set of identical nodes. The part of node and degree strings representing the sub-branches in two mated branches rooted from the chosen node will be defined. Then, the two sub-branches are crossed over to produce two new chromosomes.
Figure 8 illustrates this process. From figure 8, there are two original mated branches (in 8a and 8b). Node C is one of the common nodes in these two branches which are selected for crossover. The sub-branches in the dash line boxes in (8a) and (8b) are the parts of branches that will be crossed over. Figures 8c and 8d show the two new branches after crossing over. The existence of dummy node requires the crossover and mutation processes to conduct the checking process after the operation. After applying GA operators, an algorithm for detecting the new dummy nodes as the results of the crossover operation must be applied to the new chromosomes.

5.7 Mutation process

The second GA operator is the mutation process aimed to preserve the diversity amongst the population and to represent the stochastic evolution process in the nature. The chromosome structure in GA-AS is not consistent with the traditional binary string chromosome. Thus, a new approach to mutate the chromosome is developed. The mutation process in GA-AS involves the branching process at a node (including both branching in and out). For a leaf node, if the node is to be mutated, the node will be expanded following the branching process explained earlier (this is branching out process). On the other hand, if the node to be mutated is not a leaf node, the branch will be branched in at that node, i.e. converting that node to a leaf node and remove all nodes in a sub-branch below the mutated node. Figure 9 gives some idea on how the mutation works.

The branch in Figure 8d is used as the original branch this time. The two branches in the left hand side of Figure 9 exhibit the mutation as branching in where node F is contracted. The two branches in the right hand side demonstrate the branching out as mutation process where node I is expanded.

6. NUMERICAL RESULTS WITH THE EDINBURGH NETWORK

This section presents the results from the test for the GA with the network of Edinburgh city in UK. This network is also used by May, et al (2002). The GA-AS will be used with and without
the equity constraint to find the best two toll rings. In order to compare the performance of this algorithm, the toll rings found will be compared with three judgmental cordons which are predefined including the inner cordon 1, inner cordon 2, and outer cordon 1. The Edinburgh network has 350 links and 25 zones. Figure 10 shows the network and three judgmental cordon charging designs resulting from applying the judgmental approach. The hexagon and triangular nodes represent junctions and zones. All of the cordons are tested with eight predefined toll levels, which are £0.50, £0.75, £1.00, £1.25, £1.50, £2, £3 and £4, and the optimal uniform toll for each toll ring can be identified.

The inner cordon 1 is used as the initial cordon for defining the target nodes for the GA. The boundary of the cordon is also limited by the outer ring road to ensure the availability of the diversion route. Table 1 shows the test results in detail. Note that the equipment and operational costs have been estimated to be £183,400 per toll point and £85,300 per toll point per annum respectively, (Oscar Faber, 2001). If we assume a 30 year life, a discount rate of 6% and 1000 peak hours per year then this is approximately equivalent to £100 per toll point per peak-hour for implementation and operating costs. Table 2 shows the parameter used in the GA for the optimisation process. Note that the parameters used for the penalty in the case of constrained optimisation is $k = 10$ and $\rho = 0.95$ (see §5.4).

Figure 11a and 11b below show the first and second best cordons as the results of applying the GA without the equity constraint, named OPC1 and OPC2 respectively. The uniform tolls for both cordons are £1.50. The direction of the tolls imposed is inward to the centre of the cordons. The OPC1 and OPC2 cordons produce the total benefit of around £7,640 and £6,850 per peak hour which is about 20.5% and 18.4% of the first-best benefit respectively. The net benefits of OPC1 and OPC2 are around £6,340 and £5,550 per peak hour respectively. The improvement in terms of the net benefit of the OPC1 and OPC2 over the inner 2 cordon is around 160% and 140% respectively. The difference in terms of the net benefit that the local authority can get between the OPC1 and inner 2 cordons is about £m 2.3 per year.
Table 1: Test result with the Edinburgh model (all figures are per peak hour)

<table>
<thead>
<tr>
<th>Toll rings</th>
<th>Optimal toll</th>
<th>No. of toll points</th>
<th>Cost (£k)</th>
<th>Total benefit (£k)</th>
<th>% of Total benefit to FB</th>
<th>Net benefit (£k)</th>
<th>Gini coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inner cordon 1</td>
<td>£0.50</td>
<td>9</td>
<td>0.90</td>
<td>3.00</td>
<td>8.1%</td>
<td>2.10</td>
<td>0.326</td>
</tr>
<tr>
<td>Inner cordon 2</td>
<td>£0.75</td>
<td>7</td>
<td>0.70</td>
<td>4.69</td>
<td>12.6%</td>
<td>3.99</td>
<td>0.295</td>
</tr>
<tr>
<td>Outer cordon</td>
<td>£0.75</td>
<td>20</td>
<td>2.00</td>
<td>3.96</td>
<td>10.6%</td>
<td>1.96</td>
<td>0.221</td>
</tr>
<tr>
<td>OPC1</td>
<td>£1.50</td>
<td>13</td>
<td>1.30</td>
<td>7.64</td>
<td>20.5%</td>
<td>6.34</td>
<td>0.487</td>
</tr>
<tr>
<td>OPC2</td>
<td>£1.50</td>
<td>13</td>
<td>1.10</td>
<td>6.85</td>
<td>18.4%</td>
<td>5.55</td>
<td>0.368</td>
</tr>
<tr>
<td>OPC3</td>
<td>£0.75</td>
<td>11</td>
<td>1.00</td>
<td>5.70</td>
<td>15.3%</td>
<td>4.60</td>
<td>0.246</td>
</tr>
<tr>
<td>OPC4</td>
<td>£0.75</td>
<td>10</td>
<td>1.00</td>
<td>5.00</td>
<td>13.4%</td>
<td>4.00</td>
<td>0.243</td>
</tr>
<tr>
<td>First-best condition (FB)</td>
<td>Varied</td>
<td>35.00</td>
<td>37.19</td>
<td>100.0%</td>
<td>2.19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 11: Two best charging cordons found by applying the GA-AS method without equity constraints

Table 2: A set of the GA-AS parameters used in the test with the Edinburgh network

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Generation number</th>
<th>Population number</th>
<th>Prob. of crossover</th>
<th>Prob. of mutation</th>
<th>No. of elitism</th>
<th>Prop parameter (see §5.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value used</td>
<td>20</td>
<td>40</td>
<td>0.35</td>
<td>0.15</td>
<td>1</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Next, the constraint on the Gini coefficient is imposed. The upper bound of the Gini coefficient is set at 0.25. Note that all judgmental toll rings except Outer cordon have the Gini coefficient higher than the criteria set. Also, the optimal toll rings found by unconstrained GA-AS, OPC1 and OPC2, all fail to meet the equity constraint (the Gini coefficients of OPC1 and OPC2 are 0.487 and 0.368 respectively). After applying the GA-AS with the equity constraint, the toll...
rings OPC3 and OPC4 are found as the best toll rings satisfying the equity constraint. Figure 12a and 12b show these two toll rings. The benefit of OPC3 is, in fact, substantially lower than those of OPC1 and OPC2. This implies the trade-off between the aggregate economic efficiency in the whole network and the equity of the system. Nevertheless, the design of both OPC3 and OPC4 can produce higher benefits compared to the judgmental cordons despite the lower Gini coefficient.

Figure 12: Two best charging cordons found by applying the GA-AS method with equity constraints ($Gini < 0.25$)

7. CONCLUDING REMARKS

The GA based approach for solving the optimal toll ring design with the equity constraint is developed and tested with the network of Edinburgh. The method is the extended version of GA-AS proposed in Sumalee (2003). The algorithm utilises the Gini coefficient, developed originally for the study of income distribution, to evaluate the spatial equity impact (across different zones) of the toll ring designs. The self-adaptive penalty based approach is used to include the equity constraint in the GA-AS process. The tests with the example network show the significant trade-off between the objective of maximising the social welfare improvement over the whole network and the constraint of keeping the distribution impact at a certain level. The optimal toll rings with and without the equity constrain produce the benefits around 15.3% and 20.5% of the first best condition respectively. On the other hand, the unconstrained design generates a potential equity problem in which the Gini coefficient for this design is about 0.49 compared to the value of 0.25 from the constrained toll design.

Comparing to the judgmental approach, the algorithm successfully produces the toll ring designs outperforming all the judgmental toll rings both in terms of the value of social welfare improvement and the spatial equity impact in the network. Further study will look at the alternative formulation of the toll ring design problem with the equity constraint as Multi-objectives optimisation aiming to maximise the social welfare function whilst minimise the spatial equity impact. Also, the extension of model to be able to analyse the vertical equity (multiple user classes) is also considered. The concept of Gini coefficient, which is very normative, will also be re-considered its plausibility and different equity measures will be developed and compared with the Gini coefficient.
REFERENCES


Knight, F.H. (1924) Some fallacies in the interpretation of social cost, Quarterly Journal of Economics 38, 582-606.


Wardrop, J.G. (1952) Some theoretical aspects of road traffic research, Proc. of the Institute of Civil Engineers 1(2).


### APPENDIX A: NOTATIONS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>the set of OD-pairs, denoted i = 1,…,I</td>
</tr>
<tr>
<td>T_i</td>
<td>the continuous number of users (or OD-flow) for OD pair i, with T_i ≥ 0</td>
</tr>
<tr>
<td>D'(T_i)</td>
<td>the inverse demand function for trips for OD-pair i, with D_i' ≤ 0</td>
</tr>
<tr>
<td>J</td>
<td>the set of directed links in the network, denoted j=1,…,J</td>
</tr>
<tr>
<td>V_j</td>
<td>the continuous number of users (or link flow) on link j, with V_j≥ 0</td>
</tr>
<tr>
<td>c_j(V_j)</td>
<td>the average cost function for the use of link j, with c_j≥ 0</td>
</tr>
<tr>
<td>C_p</td>
<td>the average path costs</td>
</tr>
<tr>
<td>F_p</td>
<td>the continuous number of users (or path-Flow) for path p, with F_p≥ 0</td>
</tr>
<tr>
<td>δ_jp</td>
<td>A dummy that takes on the value of 1 if link j belongs to path p, and a value of 0 otherwise</td>
</tr>
<tr>
<td>ε_j</td>
<td>A dummy that takes on the value of 1 if a toll can be charged on link j, and a value of 0 otherwise</td>
</tr>
<tr>
<td>s_j</td>
<td>the cost of implementation and operation of a tolled link</td>
</tr>
<tr>
<td>τ_j</td>
<td>the level of the toll on link j if ε_j=1</td>
</tr>
<tr>
<td>I</td>
<td>index for OD pairs</td>
</tr>
<tr>
<td>J</td>
<td>index for links</td>
</tr>
<tr>
<td>p</td>
<td>index for paths</td>
</tr>
</tbody>
</table>

### APPENDIX B: ALGORITHMS

**Algorithm A1:** (The programming approach adopted is the “recursive” program.)

- **Step 1:** Set k to be the first digit in node and degree string. Set j equals to 1. Put the root node and its degree into column k.
- **Step 2:** Set P_j = a set of all preceding nodes of the node in k. Set P_j = ∅.
- **Step 3:** Check if P_j = P_j. If so, set j = j – 1 and then go to step 7. Otherwise go to Step 4.
- **Step 4:** Set k = k+1.
- **Step 5:** Set n_j to be a node in P_j but not in P_j. Put n_j into the k column of node string.
- **Step 6:** Put the degree of n_j into the k column of degree string. Add n_j to P_j. If degree of n_j is equal to zero, go to Step 3. Otherwise set j = j+1 and go to Step 2.
- **Step 7:** Check if j = 0. If so, stop. Otherwise, go to Step 3.