Review on Urban Transportation Network Design Problems

Reza Zanjirani Farahani [1*], Elnaz Miandoabchi [2], W.Y. Szeto [3], Hannaneh Rashidi [4]
(1) Department of Informatics and Operations Management, Kingston Business School, Kingston University, UK
(2) Department of Industrial Engineering, Amirkabir University of Technology, Tehran, Iran
(3) Department of Civil Engineering, The University of Hong Kong, Hong Kong, China
(4) Department of Industrial Engineering, Mazandaran University of Science and Technology, Iran

Abstract

This paper presents a parsimonious review on the definitions, classifications, objectives, constraints, network topology decision variables, and solution methods of the Urban Transportation Network Design Problem (UTNDP), which includes both the *Road Network Design Problem* (RNDP) and the *Public Transit Network Design Problem* (PTNDP). The current trend and gap in each class of the problem are discussed and future directions in terms of both modeling and solution approaches are given. This review intends to give a bigger picture of transportation network design problems, allow comparisons of formulation approaches and solution methods of different problems in various classes of UTNDP, and encourage cross-fertilization between the RNDP and PTNDP research.

Keywords: Transportation; urban transportation network design problem; road network design; transit network design and frequency setting problem; multi modal network design problem

1. Introduction

Transportation is important in the sense that it allows people to take part in human activities. With the increasing population, the demand for transportation is increasing. More and more traffic is on roads, which in turn creates more and more mobility related problems such as congestion, air pollution, noise pollution, and accidents, especially in city centers where the level of human activities is high. Governments need to plan transport networks properly and control the urban traffic movements to ensure mobility and mitigate the mobility related problems simultaneously. The higher population also leads to more expensive land especially in the city centre and hence more people living in new towns or suburbs, thereby requiring new transportation infrastructures for serving new towns or improving existing transportation structures to cope with the increasing population in the suburbs. These planning, design and management issues are traditionally addressed in UTNDP. This problem can actually include the design problems in suburban areas in additional to those in urban areas because the methodology involved is basically the same. Moreover, this problem can involve transit networks in addition to road networks since the transportation include both public and private transport.

UTNDP has been continuously studied during the last 5 decades, and the number of related publications is growing over time probably because the problem is highly complicated, theoretically interesting, practically important, and multidisciplinary. A number of reviews has also been published by Boyce (1984), Magnanti and Wong (1984), Friesz (1985), Migdalas (1995), Yang and Bell (1998a), Desaulniers and Hickman (2007), Guihaire and Hao (2008), and recently by Kepaptsoglou and Karlaftis (2009). Some of these reviews deal with general network design problems but some focus specifically on urban network design or on one part of urban transportation networks. For example, the first five reviews only focus on RNDP while the last three reviews only focus on PTNDP. As a result, the similarities and differences of the formulation approaches and solution methods between the RNDP and PTNDP cannot be addressed in these reviews, which cannot encourage the cross-fertilization of the two research areas. Moreover, the problem that considers the interaction between road and public transit network designs has been ignored in these reviews.

This paper attempts to provide a holistic view to UTNDP and its classifications by uniting the decisions for

* Corresponding. Fax: +44 020 8417 7026; Tel: +44 020 8517 5098; Email: zanjiranireza@gmail.com

improving transportation networks. With regard to this, the current paper covers both problem categories, and presents a third problem category for the joint decisions in road and public transit networks with at least two modes, which considers the interactions of these modes. This paper also contains an updated literature for both fields until early 2011. This contrasts to the last review paper on RNDP which published in late 1990s, and the previous reviews of PTNDP which cover the literature until 2007. The main aim of this paper is to cover problems related to urban transportation network topology and its configuration. In this regards, only the **strategic** level and a number of **tactical** level decisions related to **network topology** are covered in this paper; those papers not related to network topology decisions, such as operational level decisions and tactical decisions that are not related to network topology, will not be covered unless they are considered together with network topology decisions. The problems such as the traffic signal setting, parking pricing, toll setting, and the public transit ticket pricing are important sub-problems of UTNDP and even some of these has long history with lots of important features and development. These sub-problems are deserved to be examined and reviewed in a comprehensive manner in future review papers and hence excluded in this paper. Traffic signal setting has the strongest relevancy with network configuration decisions, as the network topology directly affects the flow pattern and the conflict points at intersections. This subproblem has been studied extensively and it has a relatively large body of literature (e.g. Cantarella et al., 1991; Meneguzzer, 1995; Wong and Yang, 1999; Wey, 2000; Cascetta et al., 2006).

Moreover, this review focuses on deterministic transport networks and deterministic travel demand. That is, we focus on papers that assume no supply and demand uncertainty. For example, there is no randomness in travel demand and road capacity considered in the reviewed papers. Nevertheless, we can still identify current trends and gaps as shown later, and bring out new research directions in this field as shown in the last section.

Other than reviewing UTNDP, we also review the solution methods. This allows comparisons of solution methods of different problems in various classes of UTNDP and proposes new algorithmic research directions. This algorithmic review and the new directions are particularly important given that these methods are required to solve practical design problems, and their problem sizes become larger and larger. Real case scenarios are also reviewed to give insight about the size of the networks for each problem catalogues currently considered and give some hints on the future requirement of the solution methods for practical problems.

The rest of the paper is organized as follows. Section 2 explains the key definitions, classifications, and general formulations of UTNDP. Section 3 reviews the specific problem studied in the literature. Section 4 depicts the solution methods used in the literature. Section 5 describes the application to real case scenario. Section 6 presents an overall of the research development of UTNDP. Finally, the summary and further research directions are presented in Section 7.

2. Definitions, general formulations and classifications of UTNDP

2.1. Definitions of UTNDP

There are at least three different definitions of UTNDP in the literature:

1. UTNDP is concerned with building new streets or expanding the capacity of existing streets (Dantzig et al., 1979). This definition is quite common in the literature, but most studies use other names for this problem catalog such as the Road Network Design Problem (RNDP), the transportation network design problem, and the network design problem.
2. UTNDP is to determine the optimal locations of facilities to be added into a transportation network, or determine the optimal capacity enhancements of existing facilities in the network (Friesz, 1985). In this definition, the facilities may be represented by either nodes or links. Therefore, this definition is wider than the

first one.

3. UTNDP deals with a complete hierarchy of decision making processes in transportation planning, and includes strategic, tactical and operational decisions (Magnanti and Wong, 1984). *Strategic decisions* are long-term decisions related to the infrastructures of transportation networks including both transit and road networks; *tactical decisions* are those for the effective utilization of infrastructures and resources of existing urban transportation networks, and *operational decisions* are short-term decisions, which are mostly related to traffic flow control, demand management or scheduling problems. Figure 1 gives examples of strategic, tactical and operational decisions in UTNDP. The shaded items are related to network topology. This definition has the widest scope among the three since it includes the management aspect on the tactical and operational levels in addition to the planning aspect on the strategic level.

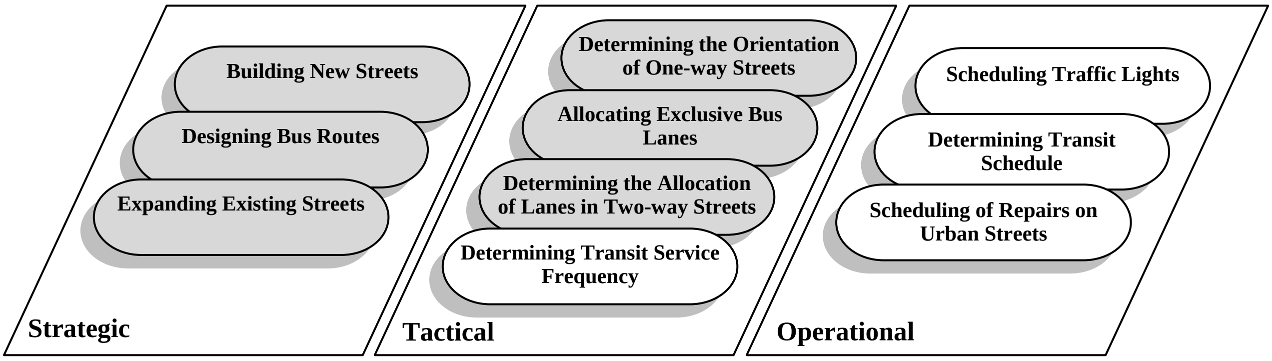

Figure 1. Examples of Decisions in UTNDP

In order to create a comprehensive and integrated collection of classifications under a single umbrella, this paper adopts the third definition to define UTNDP, but mainly limits itself to the decisions specified in the second definition – the decision related to network topology. Actually, we believe that this definition encompasses both the *Road Network Design Problem* (RNDP) and the *Public Transit Network Design and Scheduling Problem* (PTNDSP) in which the term public will be omitted in the rest of paper for the sake of brevity, that determines the optimal transit routes, frequencies, and time-tables, because transport networks include both transit and road networks. In addition, we believe that this definition includes two big classes of UTNDP: 1) the problem of developing a new network via adding links, which is related to strategic decisions and 2) the problem of improving or managing the current network, which is related to the tactical and operation decisions.

2.2. General Framework of UTNDP

UTNDP differs from network design problems in other disciplines such as telecommunication because the reaction of travelers has to be taken into account when designing a transportation network. Moreover, designing the network is associated with certain transport policy. Regarding to these two facts, analyzing and modeling in UTNDP involves two issues: (1), policy making for the network improvement, and (2) predicting the network user behaviors in response to the formulated design policies. The rest of this section will discuss the problem from the mentioned aspects.

2.2.1. General Mathematical Model for UTNDP

The problem is usually formulated as a bi-level problem or a leader-follower problem. The upper level problem is the leader's problem, the design problem, or the problem of the decision maker, (e.g., the government), who plans or manages the transport network. This upper level problem is related to the policy discussion in practice and includes the measurable goal (e.g., reducing total travel time), restrictions (e.g., political, physical, and environmental constraints) and the design decisions to be made (e.g., new roads to be built). This upper level problem assumes that the leader can predict the behavior of the travelers. The lower level problem is the followers' problem or the problem of travelers who decide whether to travel, and if so, their travel modes and routes. The bi-level structure allows the decision maker to consider the reaction of the travelers and improve the network to influence the travel choice of travelers but has no direct control on their choice. This structure does not allow the travelers to predict the decision of the leader, but only allows them to determine their choice after knowing the decision of the leader. Mathematically, the problem can be represented as follows:

$$\text{(U0)} \quad \min_{u} \ F(u,v(u)) \tag{1}$$

$$\text{s.t.} \quad G(u,v(u)) \leq 0 \tag{2}$$

where $v(u)$ is implicitly determined by:

$$\text{(L0)} \quad \min_{v} \ f(u,v) \tag{3}$$

$$\text{s.t.} \quad g(u,v) \leq 0 \tag{4}$$

F and u are, respectively, the objective function and (network design) decision variable vector of the upper level problem (U0), and G is a vector function in the upper level constraint. f and v are, respectively, the objective function and decision variable (flow) vector of the lower level problem (L0), and g is a vector function in the lower level constraint.

$v(u)$ is called the reaction or response function, which depicts the user reaction in terms of a flow pattern for each network design u. Since $v(u)$ is an implicit function that cannot be shown explicitly, it is depicted by L0. In each network design problem, $v(u)$ is an optimal solution of L0. In fact, the objective of the bi-level network design problem is to find an optimal decision vector $u*$ to optimize the objective function F subject to the network design constraint (2) as well as the user reaction constraints (3)-(4).

The above lower-level problem can be expressed as a variational inequality; in this case the bi-level network design problem can be formulated as a *mathematical program with equilibrium constraints*. UTNDP then becomes single-level but conceptually, it is bi-level with two different types of games involved, namely the leader-follower game and the non-cooperative Nash game.

Solving a bi-level network design problem using exact solution methods is very difficult because the problem is NP-hard. Ben-Ayed et al. (1988) studied bi-level problems and concluded that even a simple bi-level problem with both linear upper-level and lower-level problems is also NP-hard. Another reason is the non-convexity of bi-level network design problems. Even if both the upper and lower level problems be convex, the convexity of the bi-level problem cannot be guaranteed (Luo et al., 1996).

The presented mathematical model mostly corresponds to RNDPs, while due to the complexity of TNDSPs only in some cases formulating the TNDSP as a bi-level network design problem is possible. Most of the TNDP are single level problems where the reactions of users are simplified. As mentioned by Chakroborty (2003) it is difficult to formulate TNDSP as a mathematical problem, since it is inherently discrete and concepts such as *transfers* and *route continuity* are hard to represent. Finally, Baaj and Mahmassani (1991) discussed the complexity of this problem arised from its combinatorial, nonlinearity, non-convexity and multi-objective nature. They also depicted the difficulties in formulating as a mathematical model, and in defining acceptable spatial route layouts.

2.2.2. Classifications of Upper Level Problems or Design Problems in UTNDP

UTNDP can be classified into problems which arise from variety of possible network design policies and decisions. Traditionally, UTNDP is separately considered in two main catalogues, namely RNDP and TNDSP. RNDP mainly considers street networks and does not distinguish the flow of public transit vehicles and other private vehicles. RNDP usually supposes all traffic flows as homogenous. Based on the nature of the decisions considered, RNDP can be further classified in three groups: (1) *Discrete Network Design Problem* (DNDP), which only deals with *discrete design decisions* such as constructing new roads, adding new lanes, determining the directions of one-way streets, and determining the turning restrictions at intersections, (2) *Continuous Network Design Problem* (CNDP), which is only concerned with *continuous design decisions* such as expanding capacity of streets, scheduling traffic lights, and determining tolls for some specific streets, and (3) *Mixed Network Design Problem* (MNDP), which contains a combination of continuous and discrete decisions. If the demand, land-use, or network changes over the planning horizon are also considered, RNDP can also be further classified into DNDP-T, CNDP-T, and MNDP-T, which are respectively the time-dependent extensions of DNDP, CNDP and MNDP.

TNDSP mainly considers topology of the public transit networks, as well as service frequency and time-table. The problem can be classified five types based on the decisions addressed: (1) The *Transit Network Design Problem* (TNDP) exclusively attends to design routes of transit lines including the origins and destinations of the transit routes and the sequence of links visited. (2) The *Transit Network Design and Frequency Setting Problem* (TNDFSP) determines the service frequency of each bus line in addition to route design. (3) The *transit network frequencies setting problem* purely deals with frequency setting given the route structure. (4) The *transit network timetabling problem* deals with the time-table issues given the service frequency and routes. (5) The *transit network scheduling problem* considers both the frequency and time-table decisions given the route structure.

Nearly all RNDP studies deal with improvements in the existing road networks, as most decision variables are related to the manipulation of the existing network. Moreover, new city constructions rarely occur in the real world. In contrast, TNDSP studies (especially those with topological decisions) mostly deal with new transit network configurations, or complete reconfiguration of the existing transit networks. A common characteristic of RNDP and TNDSP is that they consider only single mode (i.e., private or public transit mode), and also are often limited to the network for the mode studied (i.e., focus on single tier or level network).

In reality, there are multiple modes and their demands are interrelated. The *Multi-Modal Network Design Problem* (MMNDP) can be a suitable title for another category, which encompasses at least two different modes for UTNDP. Traffic flow in MMNDP can cover automobiles, taxis, vans, buses, bicycles, motorcycles, metro, etc. The special case of MMNDP is the *Bi-Modal Network Design Problem* (BMNDP), which considers only two modes. Decisions considered in MMNDP can be decisions of a single mode (i.e. road, transit, etc.), or combinations of various decision for the modes under focus. In fact, the multimodal problem arises when at least two modes are considered and simulated even if design decisions are related only one of the modes.. The multi-modality in urban transportation networks is captured in several forms in the literature:

- *No interactions between flows of different modes*: In this case, the networks of different modes are not related to each other, and thus the flows of one mode do not have any effect on the flows of the other modes. For example, in an automobile-metro problem or in an automobile-bus problem where buses move in exclusive lanes, transit flows are physically separated from automobile flows.
- *Interactions between flows of different modes*: When buses share the same roads with automobiles, the flows of bus and automobile affect each other. In such problems road and bus networks are related to each other in terms of both nodes and arcs.
- *Interrelations in flows and in decisions*: Most multi-modal problems only consider flow interactions, while in problems with mode related topographic decisions, the effects of decisions of one mode on the other are

addressed. For example, in an automobile-bus mode problem, converting a two-way street into a one-way street will affect the bus route on that street.

MMNDPs usually consider multi-level or multi-tier networks in which the road network and different public transit networks are considered to be sub-networks of the urban transportation networks. Another dimension of MMNDPs is the number of modes involved in travelling between two points. There are two cases in this dimension. First, travelers can only choose a one mode, and second, travelers can choose a combination of modes to finish their trips and the number of modes involved is greater than one. In the latter, the trips are called *combined-mode trips*. One example is that a traveler drives to the metro station and then takes a metro to his/her destination.

2.2.3. Classification of Lower Level Problems in UTNDP

The lower level problem has different forms, depending on the choice dimensions considered and the assumptions for the travelers used. The lower level problems in the road network category are different from those of transit network category. The road network problems deal with vehicles as flow units, but transit network problems consider passengers as flow units. The former allocates the vehicle flows to road networks under a specific transport policy scenario, while the latter allocates the passengers to the public transit vehicles after considering a fixed flow of public transit vehicles resulted from the specific transit design scenario.

The most prevalent form of the lower level problem considers the route choice of the users, which is called the *trip assignment problem*. This common form of problem occurs when single mode is under focus in pure road or transit network design problem. The problem determines the flow pattern of the transportation network for a network design scenario. Since there are two types of networks, namely road networks and transit networks, there are two types of trip assignment, namely *traffic assignment* and *transit assignment*, for determining the flow patterns in road and transit networks respectively. Trip assignment usually assumes some sort of behavior principles to depict the route choice behavior. The traditional one is Wardrop's (1952) principle, which states that the traveler selects the shortest route. This principle assumes that all travelers are non-cooperative and know the actual travel time for each route. This principle can be easily extended to consider toll, parking cost and excess time cost. The resulting flow pattern is called the *user-equilibrium* (UE) pattern, which is the pattern such that no travelers have incentive to switch to other routes. If they switched routes, their travel time would be higher. This assignment is called UE assignment. The extension of UE assignment is the *stochastic user equilibrium* (SUE) assignment (Daganzo and Sheffi, 1977), which assumes that travelers select routes based on perceived travel time. This equilibrium assumes that travelers may not know the travel time precisely. Opposite to the UE assumption, if the travelers are assumed to behave cooperatively, the resulting assignment is called system optimal (SO) assignment.

In all the above traffic assignment problems, congestion effect is normally considered (i.e. the travel costs do depend on the traffic flows) but in some cases, the congestion effect is ignored. This leads to congested and uncongested assignment respectively. In particular, if the congestion effect is ignored and UE is assumed, the conventional UE traffic assignment becomes *all-or-nothing* (AON) assignment in which all the demand is assigned to the shortest route. Moreover, if SUE is assumed instead, the stochastic assignment problem, which is called *stochastic uncongested assignment*.

The transit assignment problem can also be divided into categories of *congested* and *uncongested transit assignment* problems. In uncongested transit assignment, no passenger capacity is considered for transit vehicles. The congested transit assignment considers the transit passenger capacity restrictions. In both problems, different criteria and different passenger behaviors are assumed for this problem, which leads to various approaches of allocation of passenger demands to transit paths. Similar to trip assignment, stochastic version exists for both categories of transit assignment.

Traditional trip assignment problems treat all classes of traffic flow as a single unified traffic flow. Another important extension to traffic assignment is to distinguish between different classes of travelers or vehicles. The *multiclass trip assignment* considers different demands and travel cost functions for each user class. Travelers may be classified by their vehicle ownership, trip purposes, socioeconomic attributes such as income, etc. Vehicles can be classified into different vehicle types such as private cars, taxis, trucks, motorcycles, etc.

On the other hand, the lower level problem can allow the users to have more than one travel choice other than route choice. In such case, if both destination and route choices are considered, it becomes the *trip distribution-assignment problem*. In multi-modal problems, the lower level problem has to tackle with more than one travel mode. If both mode choice and route choice are considered, the problem becomes the *modal split-traffic assignment* problem. If all destination, mode, and route choices as well as the choice of whether to travel or not are incorporated, the lower level problem becomes the *combined travel choice problem*. Depending on the model assumption, users may select different travel choices sequentially or simultaneously.

In the above problems, demand (for each mode) between each origin-destination (OD) pair can be fixed or elastic, based on the assumptions of the main model. Basically, it is called *fixed demand* when the demand for the mode considered is unchanged. When the demand for travel is dependent on various factors such as travel time of the mode considered or other system performance attributes, it is called *elastic demand*. Elastic demand can be found in two forms; first when the demand between each OD pair is variable (i.e., simple elasticity is considered), and second when the total demand between each OD pair is fixed but the share of each mode for the total demand is variable. The first case is defined for circumstances that travelers decide to give up their travel when the travel cost is too high or they change their destination, but the second case is for circumstances that travelers only decide to change their mode of travel. Moreover, the second case is often found in MMNDP where the total demand is divided between various modes of travel. Nevertheless, these lower level problems implicitly consider more than one travel choice.

3. Specific Urban Transportation Network Design Problems (UTNDP)

Tables 1-6 summarize the problems studied and solution methods used in each publication. In the following sections, the specific problem being studied in each class of UTNDP is reviewed. Section 3.1 focuses on RNDP. Section 3.2 focuses on TNDP and TNDFSP. Section 3.3 focuses on MMNDP. Section 4 focuses on the solution method. For each class of the problems, the literature is presented in tables which summarize the problem attributes discussed in the current section, and the solution methods addressed in Section 4.

3.1. Studies of the Road Network Design Problem (RNDP)

Generally, the inputs of RNDP are as follows: (1) The network topology, (2) The travel demand between each origin-destination pair for a specific time interval in terms of a matrix or a function, (3) The characteristics of streets such as the capacity, the number of lanes, the free flow travel time, and the specifications of the travel time function, (4) The upper and lower bounds of decision variables for handling physical or political considerations such as the maximum allowable increase in capacity and the maximum toll level, (5) The set of candidate projects for network improvement, (6) The available budget, and (7) The cost of each candidate project. For some problems such as scheduling traffic light and determining road tolls, inputs (4)-(7) are not necessary.

The upper level constraints of RNDP may include the following: (1) The simple upper and lower bound constraints for the decision variables, (2) The budget constraint, and (3) The definitional constraints such as capacity constraint and cycle time constraint: The capacity constraint is included in the upper level problem for preventing the street flow to exceed the street capacity. This constraint is not necessary when the traffic assignment problem has considered capacity constraints. The cycle time constraint is used to ensure that the sum of green times and all loss times for all approaches equals cycle time.

3.1.1. Studies of Continuous Network Design Problem (*CNDP*)

The studies on CNDP are summarized in Table 1. They are categorized based on objectives used, number of objectives, travel demand (fixed or elastic), traffic assignments, decisions, and solution methods involved. As reflected in Table 1, a significant body of RNDP research has focused on CNDP. One reason for this may be the continuousness of variables, which allows many different modeling approaches and solution methodologies as shown in Table 1.

Table 1. A Summary on the Studies of Continuous Network Design Problems.

Reference	Objective						Single/multiple optimization	Traffic Assignment	Demand	Decision			Solution Method
	Max. Consumer Surplus	Max. Reserve Capacity	Min. Total Vehicle Miles	Min. Travel + Construction Cost	Min. Construction Cost	Min. Total Travel Time/Cost				Determining Road Tolls	Scheduling Traffic Light	Street Capacity Expansion	
Steenbrink (1974a)				●			S	SO	F			●	Iterative-Optimization-Assignment Algorithm
Abdulaal and LeBlanc (1979)				●			S	DUE	F			●	Powell's Method, Hooke and Jeeves Method
Dantzig et al. (1979)				●			S	SO	F			●	Iterative Decomposition, Lagrangian Multiplier
Marcotte (1983)				●			S	DUE	F			●	Exact Algorithm Based on Constraint Accumulation, Heuristic
						●					●		
LeBlanc and Boyce (1986)				●			S	DUE	F			●	Bard's Algorithm
Marcotte (1986)				●			S	DUE	F			●	Four Heuristics based on Iterative Optimization Assignment Algorithms
Ben-Ayed et al. (1988)				●			S	SO	F			●	The authors only discuss on the possible solution approaches
Suh and Kim (1992)				●			S	DUE	F			●	Bi-level Descent Algorithm
Friesz et al. (1992)				●			S	DUE	F			●	Simulated Annealing + Bertsekas-Gafni’s projection method
Marcotte and Marquis (1992)				●			S	DUE	F			●	Two Heuristics based on Iterative Optimization Assignment Algorithms
Friesz et al. (1993)			●		●	●	M+W	DUE	F			●	Simulated Annealing
Davis (1994)				●			S	SUE	F			●	Generalized Reduced Gradient Method, Sequential Quadratic Programming
Yang (1997)	●					●	S	DUE	E	●		●	Sensitivity Analysis Based Method
Yang and Bell (1998b)		●					S	DUE	F			●	Nil
Meng et al. (2001)				●			S	DUE	F			●	Augmented Lagrangian Method
Meng and Yang (2002)						●	S	DUE	F			●	Simulated Annealing
Yang and Wang (2002)		●				●	M+W	DUE	F			●	Simulated Annealing
Ziyou and Yifan (2002)		●					S	DUE	F		●	●	Sensitivity Analysis Based Method
Chiou (2005)				●			S	DUE	F			●	Four Gradient-based Methods
Gao et al. (2007)						●	S	DUE	F			●	Gradient-based Method
Chiou (2008)		●					S	DUE	F		●		Hybrid Approach for Simultaneously Solving the Two Problems
						●	S	DUE	F			●	
Xu et al. (2009)				●			S	DUE	F			●	Genetic Algorithm, Simulated Annealing
Mathew and Sharma (2009)						●	S	DUE	F			●	Genetic Algorithm

Wang and Lo (2010)				●			S	DUE	F			●	Transformation into a Mixed Integer Linear Program

In terms of objectives and decisions, travel time related objectives and street capacity expansion have been received most attention. Capacity decisions are supposed to be continuous. Some studies also consider both the link capacity decision with the decisions on road tolls (e.g., Yang, 1997) and traffic lights (e.g., Ziyou and Yifan, 2002 and Chiou, 2008) simultaneously.

Table 1 also shows that most studies on CNDP assume fixed demand (denoted by F). However, some studies (e.g., Yang, 1997) consider elastic demand (denoted by E). Moreover, most studies only have single objective (denoted by S) and limited studies consider multi-objective (denoted by M). Multi-objective problems are all handled by the weighted sum approach (denoted by W). Furthermore, most studies focus on DUE traffic assignment and total travel time.

3.1.2. Studies of Discrete and Mixed Network Design Problems (*DNDP and MNDP*)

A summary of DNDP studies is given in Table 2. The body of DNDP literature is somehow smaller than that of CNDP, probably because of the complexity resulted from the presence of discrete variables. These DNDP studies have been investigated in three distinct groups. The first group of studies is similar to the classical literature of CNDP in the sense that they focus on strategic decisions such as constructing new streets, or increasing the capacity of existing streets as the yes-no type decisions. The second group of studies considers the tactical decisions such as determining the orientation of one-way streets, allocating lanes in two-way streets, and changing some two-way streets into one-way streets. The last group considers both strategic and tactical decisions.

Compared with CNDP, additional objectives such as minimizing the sum of total network cost and total flow entropy (where the entropy of the link flow provides a measure of flow balance of the link: the larger the flow entropy is, the more unsymmetrical flow is), and minimizing total travel distance have been considered. However, maximizing consumer surplus, which is used in the elastic demand case, has not been considered. Moreover, the problem considered so far has only one objective with fixed demand. Some work can be done for the multi-objective or elastic demand case.

Table 2. A Summary on the Studies of Discrete Network Design Problems.

Reference	Objective						Single/multiple optimization	Traffic Assignment	Demand	Decision					Solution Method
	Max. Reserve Capacity	Min. Network Travel cost + Flow Entropy	Min. Total Travel Distances	Min. Total Societal Cost	Min. Travel + Construction Cost	Min. Total Travel Time/Cost				Turning Restrictions at Intersections	Lane Allocation in Two-way Streets	Making Some Streets One-Way	New Street Constructing	Street Capacity Expansion	
DNDP with Strategic Decisions															
Steenbrink (1974b)				●			S	SO	F				●	●	Iterative Decomposition Algorithm
Leblanc (1975)						●	S	DUE	F					●	Branch and Bound
Poorzahedy and Turnquist (1982)						●	S	DUE	F				●	●	Branch and Backtrack Heuristics
Chen and Alfa (1991)					●		S	SUE	F				●		Branch and Bound+ Stochastic Increment Assignment
Yang and Bell (1998b)	●						S	DUE	F					●	Nil

Gao et al. (2005)						●	S	DUE	F				●		Generalized Benders Decomposition Method with Support Function
Poorzahedy and Abulghasemi (2005)						●	S	DUE	F				●	●	Ant Systems
Poorzahedy and Rouhani (2007)						●	S	DUE	F				●	●	Hybrids of Ant Systems, Ant Colony with Genetic Algorithm, Simulated Annealing, Tabu Search
DNDP with Tactical Decisions															
Lee and Yang (1994)						●	S	DUE	F			●			Simulated Annealing
Drezner and Wesolowsky (1997)			●				S	AON	F			●			Branch and Bound, Heuristics, Simulated Annealing
Drezner and Salhi (2000)			●				S	AON	F			●			Tabu Search
Drezner and Salhi (2002)			●				S	AON	F			●			Genetic Algorithm, Simulated Annealing
Zhang and Gao (2007)						●	S	SUE	F		●				Particle Swarm Optimization
Wu et al. (2009)		●					S	SUE	F		●				Chaotic Optimization Algorithm
Long et al. (2010)						●	S	SUE	F	●					Branch and Bound with Sensitivity Analysis Based
DNDP with Both Strategic and Tactical Decisions															
Drezner and Wesolowsky (2003)			●				S	AON	F			●	●		Descent Algorithm, Simulated Annealing, Tabu Search, Genetic Algorithm
Miandoabchi and Farahani (2010)	●						S	DUE	F		●	●		●	Hybrid of Simulated Annealing and Genetic Algorithm, Evolutionary Simulated Annealing

A review of the studies of MNDP is shown in Table 3. As shown from this table, these studies involve decisions that are combinations of those in CNDP and DNDP. More importantly, only a few studies in this field have been accomplished in the recent decade. It seems that a lot of work can be done in this area (for example, considering lane allocations in two-way streets or considering multiple objectives with environmental concerns).

Table 3. A Summary on the Studies of Mixed Network Design Problems.

Reference	**Objective**	**Traffic Assignment**	**Demand**	**Decision**							**Solution Method**
				Discrete				Continuous			
				Making Some Streets One-Way	Street Capacity Expansion	Constructing New Streets	Orienting Sequences of Streets	Traffic Light Setting	Street Capacity Expansion	Road Tolls Setting	
Yang and Bell (1998a)	General Weighted Sum Multi-Objective	DUE	F			●			●		Enumeration Scheme with Other Methods
Cantarella et al. (2006)	Min. Total Travel Time	DUE	F				●	●			Hill Climbing, Simulated Annealing, Tabu Search, Genetic Algorithm, Hybrids of Tabu Search
Dimitriou et al. (2008a)	Max. Profit	SUE	E		●					●	Genetic Algorithm
Zhang and Gao (2009)	Min. Total Travel Cost + Construction Cost	DUE	F			●			●		Gradient Based Method with Penalty Function
Gallo et al. (2010)	Min. Total Travel Time	SUE	F	●				●			Scatter Search

3.1.3. Time-dependent Studies of Road Network Design Problem (*RNDP*)

Because demand and land use are changing over time, the consideration on the planning horizon is important to be considered in the strategic decision making. However, there are only a few time-dependent studies of the RNDP, which can be classified into two groups, namely the time-dependent *Continuous and Discrete Network Design Problems* (referred to as CNDP-T and DNDP-T respectively) as shown in Table 4.

Table 4. A Summary on Time-Dependent Network Design Studies.

Reference	Objective	Traffic Assignment	Demand	Decision				Solution Method
				Discrete		Continuous		
				Constructing New Streets	Street Capacity Expansion	Street Capacity Expansion	Road Tolls Setting	
CNDP-T								
Lo and Szeto (2004)	Max. Consumer Surplus	DUE	E			•		Generalized Reduced Gradient Method
Szeto and Lo (2005, 2008)	Max. Discounted Social Surplus	DUE	E			•	•	Generalized Reduced Gradient Method
Szeto and Lo (2006)	Max. Discounted Social Surplus Max. Intergeneration Equity	DUE	E			•	•	Generalized Reduced Gradient Method
Lo and Szeto (2009)	Max. Discounted Consumer Surplus	DUE	E			•	•	Generalized Reduced Gradient Method
DNDP-T								
O'Brien and Szeto (2007)	Max. Discounted Consumer Surplus	DUE	E		•		•	Generalized Reduced Gradient and Branch and Bound

All these groups deal with determining the sequence of the decisions to be made within the modeling horizon. Their bi-level models can be viewed as an extension of the bi-level network design model (1)-(4) and time indices appear in the resulting model as shown below:

$$\text{(U1)} \quad \min_{u} \; F(u_1,\ldots,u_T, v_1(u_1),\ldots,v_T(u_T)) \tag{5}$$

$$\text{s.t.} \quad G(u_1,\ldots,u_T, v_1(u_1),\ldots,v_T(u_T)) \le 0 \tag{6}$$

where $v_t(u_t)$ is implicitly determined by:

$$\text{(L1)} \quad \min_{v} \; f(u_1,\ldots,u_T, v_1,\ldots,v_T) \tag{7}$$

$$\text{s.t.} \quad g(u_1,\ldots,u_T, v_1,\ldots,v_T) \le 0 \tag{8}$$

where $u = [u_t]$ and $v = [v_t]$, and the modeling horizon is [1, *T*]. If v_t is not equal to v_{t-1}, then a project should start at time *t*. In this way, the project phasing is captured.

3.2. Studies of Transit Network Design Problems and Transit Network Design and Frequency Setting Problems (TNDP and TNDFSP)

Usually, the main inputs to TNDP include (1) the network of available infrastructure, which might consist of the urban road network (with or without dedicated facilities such as bus stops), rail structures (e.g. tramways) and (2) the estimated travel demand. Depending to the type of problem, TNDPs may include various combinations of inputs such as (3) the available budget, (4) capacity of buses, (5) maximum or desired number of bus lines, (6) set of predefined possible bus lines, (7) maximum and minimum allowable round trip time of bus lines, (8) frequency of bus lines, (9) maximum number of bus stops, (10) minimum coverage area by the bus network as the percentage of demand that can be served, (11) maximum number of transfers, etc. For TNDFSP, additional information is required such as fleet size, and the minimum service frequency, etc.

As stated in Guihaire and Hao (2008), the typical considerations for TNDP are normally the following: (1) dependency on the existing routes, (2) area coverage, (3) route and trip directness, (4) demand satisfaction, (5) the number of lines or total length of routes, and (6) the overall shape of the network. For TNDSFP, additional considerations such as the number of runs and frequency bounds for each line are usually included. Depending on the decision maker's policy for transit networks, these considerations can be reflected in constraints or objective functions by setting a bound for the measure or optimizing the measure respectively. For example, route length can appear in the objective function (e.g., Barra et al., 2007) and route constraints (e.g., Szeto and Wu, 2011). For the former, the total route length is minimized whereas for the latter, the route length of each transit line cannot be exceeded by a user-predetermined value.

Regarding to numerous studies in TNDFSP and TNDP, a summary are presented in Table 5. The studies in these two fields are very diverse and different in terms of objectives, the constraints, and the solution methods used. The problems are usually approached by multiple criteria to consider the public sector's, line operators' and the users' interests but normally a weighted sum objective function is eventually used. Also, most of them assume a very simplified passenger behavior that leads to a single-level optimization problem.

In most of the existing bus network design problems, any flows on the road network other than buses such as cars, taxis, and bicycles were generally ignored, and hence the interactions between these flows and bus flows were not taken into account.

Table 5. A Summary on TNDP and TNDFSP studies.

Reference	Constraints	Objective(s)	Solution Method
TNDP Studies			
Patz (1925)	- Bus Capacity. - Demand.	Min. Number of Empty Seats.	Heuristic
Sonntag (1977)	Restricted Set of Possible Lines.	- Min. Average Travel Time. - Min. Number of Transfers.	Heuristic
Mandl (1980)	- Constant Frequency. - Area Coverage.	- Min. Travel Time. - Max. Route Directness.	Heuristic
Pape et al. (1992)	Area Coverage.	- Min. Number of Lines. - Max. Number of Direct Passengers.	Heuristic
Xiong and Schneider (1992)	Nil	- Min. Total Travel Time. - Min. Construction Cost.	Cumulative Genetic Algorithm
Chakroborty and Dwivedi (2002)	Route feasibility.	- Min. Total Travel Time. - Min. Unsatisfied Demand. - Max. Number of Passengers, Who Travel with at Most Two Transfers.	Genetic Algorithm
Zhao and Gan (2003)	- Predefined Routes and Areas - The Number of Lines and Stops. - Route and Network Directness. - Deviation from Main Routes. - Route Length	- Min. Number of Transfers. - Max. Route Directness. - Max. Area Coverage.	Greedy Search, Hill Climbing, Hybrid of Tabu Search, Simulated Annealing, and Greedy Search
Zhao and Ubaka (2004)			Greedy Search, Fast Hill Climb Search
Yu et al. (2005)	- The Length of Lines. - Route Directness.	- Min. Total Number of transfers. - Min. Maximum Passenger Flow in Each Route	Parallel Ant Colony
Guan et al. (2006)	- The Link Capacity. - The Length of Line. - Number of Transfers.	- Min. Total Length of Routes. - Min. Total Numbers of Routes Taken by Passengers. - Min. Total Distances Traveled by Passengers.	Mathematical Method
Zhao and Zeng (2006)	- Route Directness. - Route Feasibility. - Number of Routes. - Route Length. - Budget.	- Min. Average Number of Transfers - Max Service Coverage	Hybrid Metaheuristics
Barra et al. (2007)	- Travel Demand Satisfaction. - Budget. - Service Level.	Min. Total Length of Routes.	Mathematical Method
Yang et al. (2007)	The Length of Routes.	- Max. Number of Direct Travelers Per Unit Length.	Parallel Ant Colony
Mauttone and Urquhart (2009)	Travel Demand Satisfaction.	- Min. Number of Routes. - Min. Total Travel Time.	Heuristic
Fan and Mumford (2010)	The Number of Lines.	- Min. Total Travel Time. - Min. Total Number of Transfers.	Hill-Climbing, Simulated Annealing

TNDFSP Studies			
Lampkin and Saalmans (1967)	Fleet Size.	- Max. Number of Direct Passengers. - Min. Total Travel Time.	Heuristic
Silman et al. (1974)	Budget.	- Min. Fleet Size. - Min. Journey Time. - Min. Overcrowding.	Heuristic
Dubois et al. (1979)	Budget.	Min. Total Travel Time.	Heuristic
Hasselström (1979; 1981)	Budget.	- Min. Number of Transfers. - Max. Number of Passengers.	Mathematical Method
Ceder and Wilson (1986)	- Minimum Frequency. - Fleet Size. - Route Length.	- Min. Excess Travel Time, Transfer and Waiting Time. - Min. Vehicle Costs.	Heuristic
Van Nes et al. (1988)	Fleet Size.	- Max. Demand Satisfaction. - Max. Number of Direct Trips.	Mathematical Method+ heuristic
Shih and Mahmassani (1994); Shih et al. (1998)	Nil	- Min. Travel Time. - Max. Demand Satisfaction. - Min. Fleet Size.	Heuristic
Baaj and Mahmassani (1995)	- Headway. - Fleet Size. - Capacity.	- Max. Number of Direct Trips. - Min. Waiting Time, Transfer Time.	Heuristic
Bussieck (1998)	- Number of Transit Vehicles. - Frequency. - Line and Vehicle Capacity.	- Max. Number of Direct Passengers. - Min. Operator Costs.	Mathematical Method
Pattnaik et al. (1998)	- Headway. - Load Factor	- Min. Operator Costs. - Min. Passengers' Travel Time.	Genetic Algorithm
Carrese and Gori (2002)	- Demand Satisfaction. - Routes Length. - Number of Transfers. - Total Travel Time. - Fleet Size.	- Min. Users Waiting Time and Excess Time Compared to the Minimum Path. - Min. Operator Costs.	Heuristic
Bielli et al. (2002)	Using Predetermined Lines.	Max. 24 Criteria of Network Performance.	Genetic Algorithm
Fusco et al. (2002)	- Level of Service. - Satisfied Demand. - Lines Configuration. - Frequency.	Min. Overall Cost.	Heuristics
Ceder (2003)	- Route Length. - Deviation from Shortest Path.	- Min. Operator and Users Costs. - Min. Fleet Size	Heuristic
Ngamchai and Lovell (2003)	Area Coverage.	- Min. Cost of Waiting and Traveling. - Min. Fleet Size or Fleet Cost.	Genetic Algorithm
Tom and Mohan (2003)	Nil	- Min. Operation Costs. - Min. Total Travel Time.	Genetic Algorithm
Wan and Lo (2003)	- Service Frequencies of Lines. - Bus Capacity.	Min. Operation Costs.	MIP Solver in CPLEX
Agarwal and Mathew (2004)	- Frequency. - Load Factor.	- Min. Operator and Users Cost	Heuristic, Parallel Genetic Algorithm
Fan and Machemehl (2004)	Route Length.	- Min. Time of Waiting, Traveling, and Walking - Min. Fleet Size. - Min. Cost of Unsatisfied Demands.	Genetic Algorithm
Hu et al. (2005)	- Route Length. - Average Transfer Times, Station Stopping Times and Headways.	- Max. the Nonstop Passenger Flow - Min. Sum of Passengers' and Operator Costs	Genetic Algorithm, Ant Colony
Zhao (2006)	Route Directness.	- Min. Total No. of Transfers. - Max. Demand Coverage.	Simulated Annealing
Zhao and Zeng (2007)	- Headway Bound. - Fleet Size. - Route Length. - Load Factor.	Min. Weighted Sum of Users' and Operator Cost.	Heuristics, Simulated Annealing
Borndörfer et al. (2008)	Demand Satisfaction	- Min. Operation Cost. - Min. Total Travel Time.	Column Generation
Pacheco et al. (2009)	- Number of Routes. - Fleet Size. - Location of the Bus Stops.	Min. Total Time of Waiting and Traveling.	Local Search, Tabu Search
Cipriani et al. (2010)	- Bus Capacity. - Frequency. - Route Length.	Min. Sum of Operator and User Costs.	Heuristic, Parallel Genetic Algorithm
Szeto and Wu (2011)	- Fleet Size. - Number of Transit Stops. - Frequency. - Route Length.	Min. Weighted Sum of the Number of Transfers and Network Travel Time.	Genetic Algorithm

3.3. Studies of the Multi-Modal Network Design Problem (MMNDP)

A summary of the studies of MMNDP is mentioned in Table 6. The lower-level problem of these studies is formulated as the modal split-traffic assignment problem to capture the interaction of the demand between modes. As we can see, these studies can be divided into two groups. The first group covers the studies in allocating exclusive lanes to specific transportation modes. There are few works in this group, in which most of them analyze some restricted scenarios as decision making options, and there is just one case dealing with mathematical modeling. The second group encompasses the studies which only deal with transit network design decisions. The third group covers studies that simultaneously consider various strategic and tactical decisions of UTNDP. Again, limited works are in this group. Moreover, these two groups assume that travelers can reach their destinations by using one mode and combined mode trips such as bus-metro trips and *Park'n Ride* trips are not considered. Indeed, there is a third group that further considers the operational decisions of RNDP and TNDP. We do not cite them here, as they fall outside the scope of this paper.

Table 6. A Summary on MMNDP studies.

Reference	Decision(s)	Lower-Level Problem	Modes	Flow Inter-action	Approach	Objective(s)	Solution Method
First Group Problems							
Seo et al. (2005)	Allocating Exclusive Bus Lanes	Nil (Software)	Car and Bus	Yes	Scenario Comparison	-	-
Elshafei (2006)	Allocating Exclusive Bus or Bicycle Lanes	Nil (Software)	Car, Bus, Bicycle	Yes	Scenario Comparison	-	-
Mesbah et al. (2008)	Allocating Exclusive Lanes for Buses	Logit Model, DUE, Uncongested Transit Assignment	Car and Bus	Yes	Mathematical Model	Min. Total Travel Time	Enumeration Method
Li and Ju (2009)	Allocating Exclusive Bus Lanes	Combined Mode-Split Dynamic DUE	Car and Bus	Yes	Scenario Comparison	-	-
Second Group Problems							
Lee and Vuchic (2005)	- Determining Bus Routes - Determining Bus Line Frequencies	Logit Model, Congested Transit Assignment	Car and Bus	No	(No mathematical model)	Min. Total Travel Time.	Heuristic
Cipriani et al. (2006)	- Determining Bus Routes - Determining Bus Line Frequencies	Logit Model, DUE, Congested Transit Assignment	Car and Bus	No	Mathematical Model	Min. Sum of Operator, Users and External Costs.	Heuristic
Fan and Machemehl (2006a)	- Determining Bus Routes - Determining Bus Line Frequencies	Logit Model, Congested Transit Assignment	Car and Bus	No	Mathematical Model	Min. Total Travel Time.	Genetic Algorithm
Fan and Machemehl (2006b)	- Determining Bus Routes - Determining Bus Line Frequencies	Logit Model, Congested Transit Assignment	Car and Bus	No	Mathematical Model	Min. Total Travel Time.	Simulated Annealing
Fan and Machemehl (2008)	- Determining Bus Routes - Determining Bus Line Frequencies	Logit Model, Congested Transit Assignment	Car and Bus	No	Mathematical Model	Min. Sum of Operator Cost, User Cost, and Unsatisfied Demand Costs	Tabu Search
Beltran et al. (2009)	- Determining Bus Routes - Determining Bus Line Frequencies	Logit Model, DUE, Congested Transit Assignment,	Car and Bus	No	Mathematical Model	Min. Sum of Operator, Users and External Costs.	Heuristic, Genetic Algorithm
Gallo et al. (2011)	Determining Bus Line Frequencies	Combined Mode-Split /Assignment	Car, Bus, Metro	Yes	Mathematical Model	Min. Sum of Operator, Users and External Costs.	Heuristic, Scatter Search, Genetic Algorithm
Third Group Problems							

Reference	Decision(s)	Lower-Level Problem	Modes	Flow Inter-action	Approach	Objective(s)	Solution Method
Cantarella and Vitetta (2006)	- Determining Signal Setting - Determining Parking Spaces - Lane Allocation in Two-Way Streets - Determining One-Way Street Directions	Logit Model, DUE, Uncongested Transit Assignment	Car and Bus	No	Mathematical Model	- Min. Total Travel Time of Cars and Buses - Min. Total Pedestrian Travel Time - Min. CO Emission - Min. Number of Vehicles that Park Outside Desired Destinations - Max. Number of Users that shifts to Bus	Genetic Algorithm
Szeto et al. (2010)	- Lane additions over time - Toll setting over time	Combined Mode-Split /Assignment	Car and Bus	No	Mathematical Model	Max. Social Surplus	Generalized Reduced Gradient Method and Branch and Bound
Miandoabchi et al. (2011a)	- Determining One-Way Street Directions - Lane Allocation in Two-way Streets - Constructing New Streets - Lane additions - Designing Bus Routes	Combined Mode-Split /Assignment	Car and Bus	Yes	Mathematical Model	- Max. Total User Benefit - Max. Bus Demand Share - Max. Bus Demand Coverage -Min. Average Generalized Cost of Bus Trips	Simulated Annealing Hybrids with: - Genetic Algorithm - Clonal Selection Algorithm
Miandoabchi et al. (2011b)	Same as Miandoabchi et al. (2011a) Except Allocating Exclusive Bus Lanes Instead of Designing Bus Routes	Combined Mode-Split /Assignment	Car and Bus	Yes	Mathematical Model	- Max. Total User Benefit - Max. Bus Demand Share	SA Hybrids with: - Genetic Algorithm - Particle Swarm Optimization - Harmony Search

4. Review of Solution Techniques to Urban Transportation Network Design Problem (UTNDP)

The solution methods mentioned in Tables 1-6 can be classified into three categories: (1) exact or mathematical methods, (2) heuristics, and (3) metaheuristics. Exact methods such as the branch and bound method, branch-backtrack based algorithms or mathematical programming techniques rely on some mathematical properties to solve the problem to at least local optimality. Although some of them such as mathematical programming techniques for solving linear and nonlinear problems can be used to solve for realistic, large networks efficiently, others for integer programming problems such as the branch and bound method is inapplicable for medium or large sized networks because of computationally inefficiency.

Heuristics are usually developed from the insight of the problem but they may not be convergence. However, they are always more efficient than the branch and bound method and can be applied to large networks.

Metaheuristics such as Simulated Annealing (SA) and Genetic Algorithm (GA) were proposed based on the analogy on the physical, chemical, or biological process. They do not require any mathematical property of the problem to solve for solutions and can be used for obtaining nearly global optimal solutions. The computation speed of these methods is much faster than exact methods for integer programming problems in general at the expense of the solution accuracy. Figure 2 summarizes the applications of some metaheuristics. This figure shows that GA and SA have been mostly used to solve UTNDP, probably because they are classical metaheuristics. Some other common methods such as Tabu Search (TS), Scatter Search (SS), Ant Colony (AC), Ant Systems (AS), and Particle Swarm Optimization (PSO) have been used in a few studies. Also, some studies have developed hybrid (H) metaheuristics to achieve a better solution quality than non-hybrid metaheuristics. This figure also reveals that metaheuristics have been mostly applied to TNDFSP, DNDP, and MMNDP, mainly because of the presence of discrete decision variables in these problems and their non-convexity which causes computational burden if exact methods were used.

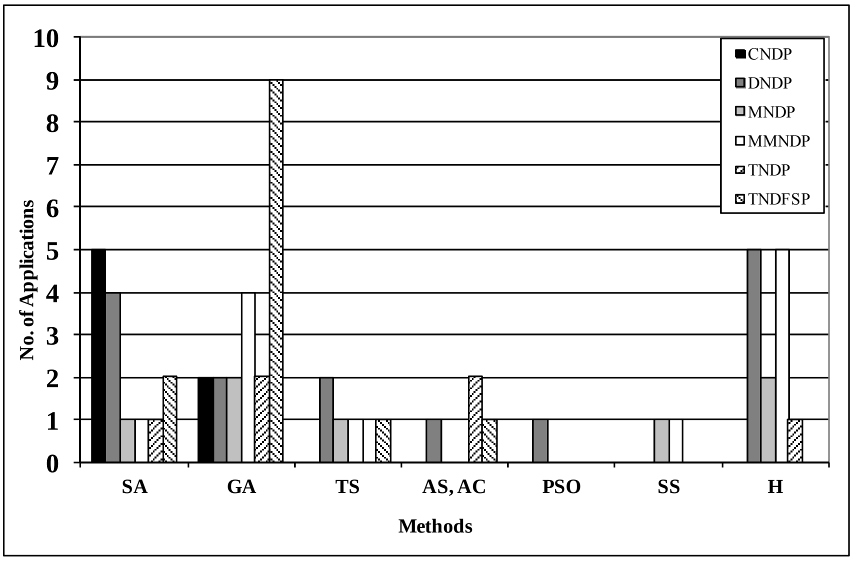

Figure 2. A Summary of metaheuristics used for UTNDP.

In the following, more discussion of solution techniques to RNDP, TNDP and TNDFSP, and MMNDP will be given in Sections 4.1-4.3 respectively.

4.1. Solution Techniques to the Road Network Design Problem (RNDP)

The literature of the Continuous Network Design Problem (CNDP) encompasses a variety of solution techniques, given the differentiability of the continuous variables. This diversity is very noticeable in 1970 and 1980 works. For Discrete and Mixed Network Design Problems (DNDP and MNDP), discrete variables make these problems NP-hard and non-convex so the exact methods such as the Branch and Bound (BB) method cannot solve these problems efficiently. We have noted that due to the intrinsic complexity of these problems, diversity of solution algorithms is somehow limited. A large amount of solution algorithms belong to the metaheuristics category and the few remaining algorithms are of exact and heuristic type methods. For time-dependent problems, very limited solution algorithms have been developed as the problems are relatively new.

As expected, the solution techniques to RNDP can be categorized into exact methods, heuristic methods and metaheuristics, since RNDP belongs to UTNDP. Exact methods such as branch and bound can be found mostly in DNDP. Examples of using such methods include LeBlanc (1975), Chen and Alfa (1991), Drezner and Wesolowsky (1997), and Long et al. (2010). They developed the BB algorithm to directly solve their (upper) problems. LeBlanc (1975) and Long et al. (2010) solved the 24-nodes and 76-links Sioux Falls network, and Chen and Alfa (1991) solved four networks, the largest of them being the 41-nodes and 73-links Winnipeg network. Drezner and Wesolowsky (1997) solved small and medium sized networks, the largest being a 40-nodes and 99-links randomly made network.

Some studies transformed RNDP to a single level problem and used exact methods to solve the resultant problem. For example, Gao et al. (2005) transformed DNDP to a nonlinear programming problem using support function concept and solve the nonlinear problem by existing nonlinear programming techniques. Gao et al. (2007) developed a gradient-based method for CNDP using lower-level problem’s optimal-value function. Zhang and Gao (2009) reformulated MNDP as CNDP, and solved the bi-level model by the use of optimal-value function of the lower-level model in a gradient-based method.

Other studies directly formulated RNDP into a mathematical program with equilibrium constraints, and then exact methods to solve the resultant problem. Not many studies adopted this approach. For CNDP, Lo and Szeto

(2004, 2006, 2007, 2009) and Szeto and Lo (2005, 2009) adopted the generalized reduced gradient based method to solve the resultant nonlinear program while for MMNDP, Szeto et al. (2010) used the branch-and-bound algorithm to the mixed-integer problem and the generalized reduced gradient based method was employed to solve the relaxed problem.

The second category is heuristics. There exists a wide variety of heuristics for RNDP and especially for CNDP. Before applying heuristics, a number of studies of RNDP use the indirect approach, in which the proposed bi-level model is first transformed to a single level optimization model or a mathematical program with equilibrium constraints, and then the resultant model is solved. For instance, Steenbrink (1974b) approximated user optimal flows with system optimal flows and solved DNDP using iterative decomposition method. Abdulaal and LeBlanc (1979) reformulated their problem as an unconstrained optimization and solved it by a direct search method. Marcotte (1983) transformed CNDP into a single level equivalent differentiable optimization problem and solved it using a constraint accumulation algorithm. Poorzahedy and Turnquist (1982) approximated the bi-level problem to a single level problem, using the objective function of the lower-level problem as the objective function, and all constraints of upper and lower-level problems as the constraints set, and then used two branch-backtrack based algorithms to solve the Sioux Falls network. Davis (1984) transformed CNDP into a single level problem by including constraints to represent the SUE condition and solved it using the sequential quadratic programming and generalized reduced gradient methods. LeBlanc and Boyce (1986) reformulated CNDP into a single level linear model with the combined lower and upper-level objective functions and solved it using Bard's method. Marcotte (1986) and Marcotte and Marquis (1992) both transformed the UE conditions into variational inequalities and developed heuristics for CNDP based on iterative optimization assignment method. Meng et al. (2001) developed augmented Lagrangian method in which the DUE conditions are represented by a single constraint in terms of the marginal function. Finally, Wang and Lo (2010) transformed CNDP into a mixed integer linear program and solved it using CPLEX.

The other group of RNDP studies has developed heuristics to directly solve the bi-level model of the problem. They mostly used descent search methods and exploit the derivative information of the implicit response function of $v(u)$ that can be obtained from the lower-level problem. Most of these studies are of CNDP type. For example, Suh and Kim (1992) used a bi-level descent algorithm based on variational inequality sensitivity analysis. Yang (1997) applied sensitivity analysis based method to set road tolls. Ziyou and Yifan (2002) used the sensitivity analysis based method to determine signal setting and link expansion. To solve CNDP, Chiou (2005) employed four gradient-based methods, namely Rosen’s gradient projection method, conjugate gradient projection method, quasi-Newton projection method, and Rosen’s gradient projection method with the techniques of parallel tangents (PARTAN), with the use of derivative information. Chiou (2008) proposed a hybrid approach to iteratively solve the two bi-level models (one for reserved capacity maximization and one for delay-minimization) and used projected quasi-Newton in part of it. Again, the sensitivity analysis based method was also incorporated to find the gradient information.

There are few studies from 1970s that developed single level models because of using system optimization for the traffic assignment, and directly solved them by using heuristic methods. For instance, Dantzig et al. (1979) developed an iterative decomposition method and Steenbrink (1974a) used iterative-optimization-assignment.

The third category of solution methods belongs to metaheuristics. In almost all of the proposed metaheuristics, the lower level problem is solved for each newly generated solution to calculate the objective function value. The application of this type of solution methods is prevalently found in DNDP and MNDP as mentioned earlier.

Therefore, many works available in the literature (as shown in Tables 2-3) rely on metaheuristics or their hybrids.

CNDP has the fewest applications of metaheuristics with few applications of SA and GA. The budget limit as the main constraint of CNDP has been considered as the penalty for the objective function by Meng and Yang (2002) and Yang and Wang (2002) in SA and by Mathew and Sharma (2009) in GA.

In contrast, DNDP has a variety of metaheuristics and their hybrids and these solution techniques usually handle constraints directly within the algorithmic steps. For DNDP with strategic decisions, AS and its hybrids have been applied. For DNDP with tactical decisions, classical metaheuristics such as SA, GA, TS have been proposed for orientation of one-way streets and PSO has been used to solve the lane allocation problem. In DNDP with strategic and tactical decisions, a number of classical metaheuristics and two hybrid metaheuristics of GA have been developed. The construction budget in DNDP is tackled inside the algorithmic steps rather than by penalties in the objective function. For instance, Poorzahedy and Abolghasemi (2005) and Poorzahedy and Rouhani (2007) considered budget feasibility when new solutions were generated by the ants. In Miandoabchi and Farahani (2010), budget feasibility was maintained when perturbation was performed or it was repaired when generating new solutions (e.g. by crossover operator in GA). In DNDP including tactical decisions, constraints some are related to street orientations (e.g. connectivity between OD pairs) or lanes configurations. These are either considered in the move generation phase (e.g. in SA), or are checked after generating new solutions (e.g. in GA) and the infeasible solutions are discarded. Only in lane allocation problem by Zhang and Zhao (2007), the out of bound lane allocations resulting from solution update by PSO were repaired.

For MNDP, various metaheuristics such as SA, TS, GA and SS have been used to solve the problems. Signal timing constraints are sometimes imposed in problems with signal setting decisions (e.g. Cantarella et al., 2006; Gallo et al., 2010). Signal timings are obtained before solving the lower level model for each network design scenario.

4.2. Solution Techniques to the Transit Network Design Problem and the Transit Network Design and Frequency Setting Problem (TNDP and TNDFSP)

As in DNDP and MNDP, the NP-hard nature of TNDFSP and TNDP requires developing innovative solution methods to get nearly optimal solutions efficiently for practical sized problems. For this reason, heuristics and metaheuristics were always used and sometimes parallel computing strategies were incorporated (e.g. Agarwal and Mathew, 2004; Yu et al., 2005; Yang et al., 2007; Cipriani et al., 2010), although not many studies were found to apply parallel computing strategies. They usually include methods to generate candidate routes using heuristic algorithms and to define the configuration of routes using heuristics or metaheuristics (e.g. Chakroborty and Dwivedi, 2002; Zhao and Zheng, 2006; Pattnaik et al., 1998; Bielli et al., 2002; Tom and Mohan, 2003; Agarwal and Mathew, 2004; Lee and Vuchic, 2005; Fan and Muchemehl, 2004; Zhao and Zeng, 2007). Other methods construct a set of routes and then try to improve them usually by metaheuristics (e.g. Hu et al., 2005; Yang et al., 2007; Mauttone and Urquhart, 2009). Generation of candidate routes is usually done using shortest-path-based algorithms which add nodes or links to the route, until some sort of defined constraints for route length, travel time or etc. are violated. The configuration of routes is usually defined by selecting and improving the generated routes which is accompanied by line frequency determination where required. Assignment of passenger demand values to the network is performed at this stage. According to Kepaptsoglou and Karlaftis (2009), the procedure for route configuration can be categorized based on the algorithms used (such as mathematical programming, iterative heuristic procedures, local search heuristics, or metaheuristic procedures), and based on whether frequencies and routes are determined sequentially or simultaneously.

Although a large collection of metaheuristic applications to these problems can be found in the literature, the applications are limited to very few numbers of classical and non-classical metaheuristics. Classical metaheuristics such as GA and SA are most prevalent among the other classical and non-classical methods such as TS and AC. Among these four methods, GA has the most number of applications especially in TNDFSP. Further research on testing the efficiency and solution quality of metaheuristics that have not been applied to TNDP and TNDFSP can be done.

There are limited studies that employed mathematical methods for obtaining solutions. For example, Bussieck (1998) applied a relaxation, branch-and-bound method in combination with commercial solvers to obtain solutions. Wan and Lo (2003) used MIP solver in CPLEX to solve their mixed integer problem. Guan et al. (2006) employed a standard branch and bound method to solve their problem. Barra et al. (2007) applied the constraint programming technique for their problem. Although these solution methods are exact, they can normally obtain solutions for small networks.

4.3. Solution Techniques to the Multi-Modal Network Design Problem (MMNDP)

Given the relatively small body of the literature in MMNDP, the solution techniques proposed in this field are few. In fact, to our best knowledge, only Mesbah et al. (2008) obtained exact solutions using simple enumeration for the exclusive lane allocation problem, which belong to the first group problems. Almost all of the studies in the second and third group of MMNDP, which includes various combinations of UTNDP decisions, developed metaheuristics to solve the problem. This is because the high complexity arising from the problem structure. The only exception is Szeto et al. (2010) which applied a generalized reduced gradient method together with the branch-and-bound algorithm.

A wide variety of metaheuristics such as SA, TS, GA, and non-classical hybrid metaheuristics such as hybrids of SA with PSO, clonal selection algorithm and harmony search have been proposed to solve the problems. For example, Miandoabchi et al. (2011a; 2011b) applied SA as an improvement heuristic to the new solutions generated by PSO, clonal selection algorithm and harmony search. In other multi-modal network design problems, the values of non-topological decision variables such as signal timings or bus line frequencies are determined after a new solution is generated by the metaheuristics. In such problems, the values of these variables are obtained after solving the lower level problem to calculate the value of the objective function (e.g. Cantarella and Vitetta, 2006; Beltran et al., 2009). However, parallel computing strategies have not been implemented.

4.4. Solution Techniques to Lower-Level Problems

The lower-level problems in RNDP are always in the form of pure traffic assignment problem. The most prevalent forms of traffic assignment in these problems are the DUE assignment with SUE and AON assignments as the other forms. These lower level problems are often solved many times in some solution methods such as metaheuristics, and hence the computation burden of the overall solution method can be due to solving these problems too often.

The most common method to solve DUE assignment in the literature is the convex-combination method (the Frank-Wolfe (1956) method) and its variants. Few different solution methods can be found such as the method of successive averages based on the flow averaging proposed by Powell and Sheffi (1982) or cost averaging proposed by Cantarella et al. (2006). The most common algorithm applied in the literature to solve the SUE problem is the Method of Successive Averages (MSA) (e.g. Chen and Alfa, 1991; Zhang and Gao, 2007; Gallo et al., 2010; Long et al., 2010) based on flow averaging.

Transit assignment in TNDP and TNDFSP is mostly based on uncongested transit assignment which is considered as constraints of mathematical models (if any) or in computation steps of the algorithms. In such assignment problems, concepts of shortest path assignment are used. Transit assignment models with the hyperpath concept of Spiess and Florian (1989) are absent in TNDP and can scarcely be found in TNDFSP. For instance, Cipriani et al. (2010) solved a TNDFSP considering the concept of hyperpath. Although they did not mention how to solve the transit assignment model, it is believed that they adopted Spiess and Florian’ solution strategy.

In MMNDP, the modal-split and trip assignment problem is solved using variant approaches. Some used simulation software packages such as NETSIM (Seo et al., 2005) and VISSIM (Elshafei et al., 2006) to solve the lower-level problem. The others considered mode split and assignment in separate steps. Cantarella and Vitetta (2006) solved the mode split problem, then equilibrium traffic assignment and considered fixed travel times for transit network. Mesbah et al. (2009) solved the mode split problem, traffic assignment and transit assignment in iterative steps until convergence was met, and Beltran et al. (2009) solved the mode split and equilibrium traffic and transit assignment problems in two iterations. Transit assignment was solved using the hyperpath approach. Moreover, some studies used logit formula for mode split, and then solved the assignment problems for one or all modes. The others considered the joint mode split and traffic assignment problem. Li and Ju (2009) proposed a heuristic algorithm, and Miandoabchi et al. (2011) used the diagonalization algorithm to solve their joint problem.

4.5. Conclusions for Solution Techniques to UTNDP

From the investigation of solution techniques, many conclusions on metaheuristics can be drawn. For example, it can be concluded that most metaheuristic techniques developed for UTNDP belong to the category of classic metaheuristics. Although a number of studies attempted to use some newer techniques, none of them have explored the application of the more recent metaheuristics in UTNDP. Furthermore, even though some versions of hybrid metaheuristics have been proposed for many other problems, no such applications are present in CNDP, and only one such study can be found in TNDFSP. Another point is that, most proposed metaheuristics in UTNDP are of conventional type, and very few of them have used parallelizing strategies to improve the computational efficiency. Similarly, multi-objective type metaheuristics are rarely used in UTNDP and most studies have used the conventional weighted-sum approach to obtain a single objective function and then applied single objective type metaheuristics to obtain solutions. Another observation from the literature is that, the branch and bound algorithm and metaheuristics have been used separately from each other and no attempt has been made to combine them. Finally, the above discussions suggest that most efforts to solve UNTDP are directed towards metaheuristic methods and little attention has been devoted to non-metaheuristics.

In addition, two points can be addressed about solution techniques for the lower-level problem. One point is that the studies that used SUE or transit assignment mostly rely on method of successive averages which is not fast enough to solve such problems. Another point is that nearly all studies have used the conventional algorithm to solve the lower level problem, which is the most important source of computational effort. No study has attempted to speed up the solution process of lower level problems using other approaches, other than solving every single lower level problem.

Lastly, we find that that descent search methods, such as SAB which are prevalent in RNDP, have not been used in TNDSP and MMNDP. This observation together with conclusions for metaheuristics and solution methods for the lower-level problem give us hints on the future research directions.

5. Applications in real-world case studies

In this section, some real-world case studies in the literature of UTNDP are reviewed. The aim is to provide information about real world examples solved in the studies and give insight about the size of the networks for each problem catalogues to be handled in the future. As UTNDP is a practical problem, some studies have applied their developed models to design realistic city networks. Table 7 summarizes the real world networks studied in RNDP publications. The city names, countries, and numbers of nodes and links of the networks are reported if they can be found in the publications. Tables 8 and 9 summarize real-world case studied in TNDP, TNDFSP, and MMNDP, including city names, countries, and attributes of the basic network of infrastructures for designing public transit networks. As we can see, the size of the network varies substantially from study to study even under the same problem catalogue, probably because the scope of the study varies from one to another. Some requires more nodes to cover more locations or larger areas but some do not.

When we compared the number of papers with case studies with those without as shown in Tables 1-6, we can conclude that most of UTNDP papers do not include a case study. 18 out of 47 studies for transit networks reviewed in this paper have solved real world cases. The case studies in RNDP are fewer than that in TNDP and TNDFSP and only 6 out of 51 studies cited here, have reported case studies. CNDP with 29 cited studies constitutes a major part of RNDP literature, but only 1 recent case study can be found in this category. This is a good indication of non-practicality of continuous street expansion decisions. For DNDP, 3 out of 17 studies have used real world problems and this is limited to DNDP with only strategic decisions. For MNDP, 2 out of 5 studies have cited real problems, in which the decisions are traffic light settings and street orientations. Again, continuous street expansion decisions with case study applications are absent in MNDP. Finally for MMNDP, 4 out of 9 studies have reported case studies. We believe that most of the studies do not include a case study because their focuses may be on developing efficient solution methods for practical problems and the authors have no real data to show the applications of the method.

Table 7. A Summary of Case Studies in RNDP.

Problem	References	City, Country	No. of Nodes	No. of Links	No. of OD pairs
CNDP (Continuous)	Mathew and Sharma (2009)	Pune, India	273	1131	4083
DNDP (Discrete)	Steenbrink (1974b)	Part of the Netherlands Road Network	not reported		
	Chen and Alfa (1991)	Winnipeg, Canada	41	73	23
	Poorzahedy and Rouhani (2007)	Mashhad, I.R. Iran	1298	1726	not reported
MNDP (Mixed)	Cantarella et al. (2006)	Barcellona Pozzo di Gotto, Italy	176	510	650
		Villa San Giovanni, Italy	88	225	380
		Melito Porto Salvo, Italy	96	189	342
	Gallo et al. (2010)	Benevento, Italy	104	175	1260

Table 8. A Summary of Case Studies in TNDP and TNDFSP.

Problem	References	City, Country	Size of the Basic Network
TNDP (Transit Network)	Mandl (1980)	Swiss Network	14 nodes and 20 links
	Zhao and Zeng (2006)	Miami-Dade County, USA	4,300 street segments, 2,804 nodes and 120,000 OD pairs
	Yu et al. (2005); Yang et al. (2007)	Dalian, China	3200 links, 2300 nodes and 1500 bus stops
	Guan et al. (2006)	The Hong Kong Mass Transit Railway, China	9 nodes and 36 OD pairs
	Mauttone and Urquhart (2009)	Rivera, Uruguay	84 nodes and 143 links
TNDFSP (Transit Network and Frequency)	Dubois et al. (1979)	Ten towns in France	For town of Niece: 700 nodes, 250 zones and 5400 links. No data provided for the all towns
	Shih et al. (1998); Baaj and Mahmassani (1995)	Austin, USA	177 bus stops
	Carrese and Gori (2002); Cipriani et al. (2010)	Rome, Italy	1300 nodes and 7000 unidirectional links
	Pattnaik et al. (1998)	A part of Madras, India	25 nodes and 35 links
	Bielli et al. (2002)	Parma, Italy	459 bus stops + A network of 99 centroids and 685 pedestrian nodes
	Agarwal and Mathew (2004)	New Delhi, India	1,332 bus stops with 4,076 links
	Hu et al. (2005)	Changchun, China	No data provided for number of bus stops or railway stations

The Hong Kong Mass Transit

Problem	References	City, Country	Size of the Basic Network
		Railway, China	
	Zhao (2006)	Miami-Dade County, USA	4,300 street segments, and 2,804 nodes 120,000 OD pairs
	Pacheco et al. (2009)	Burgos, Spain	382 bus stops
	Szeto and Wu (2011)	Tin Shui Wai, Hong Kong, China	23 nodes (zones), and 28 bus stops

Table 9. A Summary on Case Studies in MMNDP.

References	City, Country	Network Attributes
Seo et al. (2005)	Two arterial roads in Seoul, South Korea	-
Elshafei (2006)	A single roadway in Washington, USA	-
Cantarella and Vitetta (2006)	Crotone, Italy	254 nodes, 449 links, 21 origins and 21 destinations
Beltran et al. (2009)	Rome, Italy	1300 nodes and 7000 unidirectional links
Gallo et al. (2011)	Campania, Italy	262 road nodes, 43 rail nodes, 764 road links, 98 rail links, and 91 centroids

6. An Overview to the Literature of UTNDP

The literature of UTNDP can be dated back to the second decade of 20th century, when modern cities started to emerge and their design issues and the related complexities gradually grew over time. RNDP proposed in 1970 and 80s focused on both continuous and discrete decisions for street expansion or construction. The complexity of the bi-level model led to the use of SO concept in the design model, resulting in a single level model, and the resultant model was solved by branch and bound methods and heuristics that didn't rely on derivatives. In the 70s both DNDP and CNDP were researched, but later CNDP became more common since it was easier to solve. In the 1980s, following the trend in using derivative based heuristic methods and their applicability to bi-level models, the use of SO concept in CNDP and DNDP became nearly abandoned and the concept of DUE was incorporated into CNDP and DNDP. In 1990s orientation of one-way streets was proposed which further increased the complexity of DNDP. This led to the use of newly emerged metaheuristic methods at that time such as simulated annealing. On the other hand, the studies of CNDP commonly used descent search methods and simulated annealing to obtain solutions.

The availability of various metaheuristics and the development of computing technology in the 2000s resulted in the increased number of DNDP with variant decisions which was solved by metaheuristics. In the late 2000s, lane allocation in two-way streets and turning restrictions at intersections were proposed as new decisions. Heuristic methods were not applied for solving DNDP in 2000s, except for a new street construction problem which possessed specific attributes that made it possible to be solved by such a method. Also, a branch and bound method was used for solving the turning restriction design problem, which belongs to DNDP. Inclusion of movement direction decisions such as street orientation leads to new problems and applications of metaheuristics. SUE traffic assignment was used in the lower level problem of DNDP in limited number of studies since it increased the computation complexity of the bi-level problem by introducing more paths than DUE traffic assignment. SUE traffic assignment has been considered in 3 studies in the late 2000s in which used metaheuristics to solve for solutions.

A separate line of RNDP was proposed in 2000s based on the decision making over a time horizon. This line does not restrict the choice of design variables but increases the complexity of the problem since the user equilibrium problem in each time period have to be considered simultaneously.

The first study of MNDP was proposed in 90s, while it was not studied until the mid 2000s. These studies used various combinations of decisions and nearly all of them were solved by metaheuristics because of their high degree of complexity.

Unlike RNDP, TNDP and TNDFSP have such a high complexity that exact methods or local search heuristics have rarely been applied on them. These problems have been studied since 1920s. Researchers seem to be more interested in TNDFSP than TNDP, probably because transit network configuration with frequency setting is more realistic. The number of TNDFSP studied in the 2000s is almost 1.6 times of that of TNDP. The development of computing technology in the recent years have facilitated the application of advanced computing strategies such as parallelized computing in metaheuristics, and a number of recent studies of TNDP and TNDFSP have deployed them to develop more efficient solution methods.

MMNDP was studied after mid 2000s and all of them used some form of modal-split traffic assignment in the lower level model. There are three main groups of studies. The first group is related exclusive bus lanes. Although exclusive bus lanes have been used since 1930s, this problem had not been studied thoroughly until 2000s. All exclusive bus lane allocation problems were studied as MMNDP. The researchers have used scenario comparison to determine solutions or developed a model that was the enumeration method.

The second group encompasses the studies that only focus on transit network design decisions. All of the studies in this group, use the total travel time, or total costs as the single objective function of their problems. The third group covers studies that simultaneously consider various strategic and tactical decisions of UTNDP. The inclusion of various combinations of road and public transit network decisions led to the consideration of various criteria for the users of both networks and for the community in the second group. For this reason, all of MMNDPs in the third group are multi-objective. Also, due to the high complexity of the above problems, only metaheuristic methods have been applied.

In conclusion, the development trends in UTNDP are related to the development of computer technology, the cumulative knowledge on the solution methods and travel behavior, and the hot policies topic at that time. In the past, the computer technology was limited so big problems could not be solved quickly or exactly by the computer. Therefore, to solve for solutions at that time, highly simplified design problems were considered. Very often, linear functions and continuous variables were used as an approximation of the nonlinear and discrete variables and a single-level transit network design problem was studied with the consideration of a simplified passenger choice behavior. With development of new solution methods such as branch-and-cut method and metaheuristics and the advanced of computer technology, more design problems includes integer decision variables could be solved efficiently for at least nearly optimal solutions, and the approximation of continuous variables by discrete variables may not be necessarily. Moreover, the problem to be solved became bigger and bigger and more design decisions can be simultaneously incorporated in the design problem.

As time goes, more understanding on the travel behavior is obtained and more realistic travel behavior was included in the lower level problem. Hence, the design problems became more complicated but could still be solved efficiently due to the improvement of computer technology and solution methods. Furthermore, the hot policy discussion in the past leads to new design problems or at least new objective functions and constraints. From the above discussion, it seems that the existing network design studies try to find a balance between the computer technology available, cumulative knowledge on the solution methods and travel behavior, and the hot policy topic at that time.

7. Outlook

After reviewing major publications of UTNDP, we realized that there are some gaps in the literature, which need more studies. This section proposes future directions that are believed to be able to lead to essential improvements over the existing literature because the suggested directions are related to two major needs: First, the need to define problems that are more realistic in terms of policy (or design) requirement and travelers' behavior. Second, the need for more efficient solution methods after taking into account the development and the limitations of computing technology, and the applicability of solution methods in practice.

In the following, we have categorized our future suggestions for each problem type and for solution methods separately. This categorization reflects the three axes of UTNDP, namely policy discussions in the upper level problems (problem approach, decision types, objective functions, etc.), lower level problems and solution methods (for the whole problem and for lower level problems). The future research suggestions are summarized in Table 10 and elaborated in Sections 7.1-7.4.

Table 10. Summary of Future Research Directions for the UTNDP.

Problem/Solution Techniques	Subject category	Possible Extension
Road Network Design Problem (RNDP)	Objective Functions and Constraints	Using environmental related objective functions such as pollutions with traditional objective functions (for all RNDPs)
		Using objective functions that consider equity (for all RNDPs)
		Including delays incurred in intersections in the travel times (for Continuous and Discrete NDPs)
		Using reliability and robustness concepts in objective functions or in the constraints (for all RNDPs)
	Decisions	Considering operational decisions such as signal setting together with network topology decisions
	Demand	Considering elastic demand (for Discrete NDPs)
	Lower-Level Problem	Using the SUE assignment problem (for Continuous NDPs)
		Using multiclass user equilibrium assignment problems
		Using dynamic traffic assignment problems (for all RNDPs)
Transit Network Design Problem (TNDP) and Transit Network Design and Frequency Setting Problem (TNDFSP)	Objective Functions	Using environmental related objective functions such as pollutions with traditional objective functions (for all problems)
	Decisions	Using the time dependent design approach (for both problems)
	Lower-Level Problem	Adopting more realistic passenger behaviors that consider the effects of advanced passenger information system, real time transit information, and seat availability
Multi-Modal Network Design Problem (MMNDP)	Decisions	Using the time-dependent design approach (for problems with strategic decisions)
	Multi-Modality	Including fixed public transit networks with no effects on private vehicle flows, considering elastic demands and the related objective functions for each mode (for RNDPs)
		Including a fixed bus network on the existing road network that share the roads with private vehicle, considering the restrictions on road network improvement decisions and using elastic demands and the related objective functions for each mode (for RNDPs)
		Defining new combinations of multi-modal decisions to define MMNDPs
	Inter-Modal Connectivity	Considering inter-modal connection issue in combined mode trips (e.g. bus + metro networks, road + metro networks)
	Lower-Level Problem	Developing behavioral principles that can capture transfer penalty
		Develop methods to model non-additive transfer penalty
		Develop approaches to determine waiting time at transfer locations over time while taking both public and private modes into account
Solution Techniques	Metaheuristics	Using recently proposed metaheuristics and comparing their performance with genetic algorithm.
		Using various hybrids of recent and conventional metaheuristics
		Using parallelizing and distributed computing to enhance the efficiency of metaheuristics
		For multi-objective problems, developing metaheuristics that use Pareto frontier in the solution process rather than those relying on the weighted sum objective functions
	Hybrid Methods and Non-Metaheuristics	Incorporating SAB heuristic into TNDSP and MMNDP
		Using hybrids of BB method and metaheuristics
		Developing heuristic methods for problems with discrete decisions
	Lower-Level Problem Solution Methods	Using self-regulated method to solve SUE or transit assignment problems
		Using approximation schemes such as neural networks to solve traffic assignment problems

7.1. Future Research for Road Network Design Problem (RNDP)

Objective functions and constraints

- Most studies of CNDP and DNDP have only one objective and very often the objective is related to travel time. This implies that much emphasis is on congestion. Moreover, there is only one research in MNDP (Cantarella and Vitetta, 2006) that considers CO emissions minimization as one of the problem objective functions. Environmental objectives should be considered in these problems since there are usually tradeoffs between different objectives and the environmental impact is one of the hot topics nowadays. Optimizing one objective is at the expense of the others. Moreover, the impacts of emissions on health and global warming have been received attention recently. Hence, more studies of RNDP with environmental concerns can be proposed and studied in the future.
- Minimization of total travel time/cost is the most prevalent form of objective function in RNDPs. This objective accounts for the aggregate network congestion, and this may lead to unbalanced congestion levels throughout the network and hence some OD pairs can be benefited more than the others in terms of travel time. Using equity measures in the objective function can ensure more equitable benefit to be introduced to each OD pair.
- Most studies in RNDP only consider travel times in street segments, and ignore the delays incurred at intersections (e.g. signalized or priority related intersections). Intersection delays may significantly contribute to the trip time. Inclusion of intersection delays in the objective function can lead to more realistic travel times. However, there are only few related studies in DNDP and MNDP. For example, Lee and Yang (1994) considered intersection delay for the street orientation network design problem; Cantarella et al. (2006) and Gallo et al. (2010) included signal setting in the network design decisions. Intersection delays can be included in other CNDPs and DNDPs.
- Nearly all RNDPs adopted the deterministic modeling. However, travel demands or link capacities are stochastic in nature. Robustness and reliability in transportation networks thus drew attention. Nevertheless, up to now, only few studies incorporated these dimensions in UTNDPs. For example, Ukkusuri et al. (2007) considered robustness in the objective function, and Dimitriou et al. (2008b) developed CNDP with travel time reliability constraints. These two concepts can be adopted for other problems in RNDP.

Decisions

- As shown in this review, network topology decisions are seldom considered with operational decisions such as signal setting simultaneously. Indeed, traffic signal setting has the strongest relevancy with network configuration decisions, as the network topology directly affects the flow pattern and the conflict points at intersections. Although traffic signal setting problem has been studied extensively and it has a relatively large body of literature (e.g. Cantarella et al., 1991; Meneguzzer, 1995; Wong and Yang, 1999; Wey, 2000; Cascetta et al., 2006), not many network design papers considered both signal setting and network topology decisions. Hence, more works can be done in this direction in the future. Similar conclusions applied to other operational decisions such as road pricing and scheduling of repairs of urban streets.

Demand

- For DNDP, elastic demand can be studied in the future given that no studies can be found on this so far.

Lower-level problems

- Although the Stochastic User Equilibrium (SUE) problem has been used in several DNDPs, its application in CNDPs is limited to only very few studies (e.g., Davis, 1994; Long et al., 2010). Its application in CNDPs can provide more realistic behavior of the users in these problems.

- In almost all user equilibrium problems are considered with single class users only. While multiclass-user equilibrium problems have been discussed in many studies, they have been rarely applied in RNDPs. The inclusion of multiple user classes is a possible extension.
- All RNDPs in this review considered traffic assignment or its time-dependent extension in the lower level problem. These traffic assignment problems assume that the traffic condition is in a steady state, in which the traffic condition does not vary with time. Hence, they are not suitable to analyze the traffic congestion effect at the fine-grained temporal level. To examine this effect, it is more appropriate to adopt Dynamic Traffic Assignment (DTA). This problem can provide a higher degree of realism and capture changing traffic conditions. It is also essential for designing dynamic tolls, since they are changing with traffic conditions. However, most of the existing UTNDPs rarely capture DTA in the lower level problem and very few studies (Ukkusuri and Waller, 2007) did capture it. Hence, more research can be done in this direction.

7.2. Future Research for Transit Network and Transit Network Design and Frequency Setting Problem (TNDP and TNDFSP)

Objective functions

- Similar to RNDP, environmental objectives are largely ignored in these problems. More focus can be put on addressing the environmental cost in this field.

Decisions

- The time-dependent approach mentioned in Section 3.1.3 has not been considered in TNDSP, which also involve long-term, strategic decisions. Therefore, one future direction can be to develop time-dependent models for TNDP and TNDFSP. However, this adds to the complexity of the problem and needs fast and efficient solution procedures to handle this.

Lower-level problems

- Most existing studies of TNDP or TNDFSP assume simplified passengers' line choice behavior when designing transit routes. This can lead to a single level TNDSP, which is easier to solve but not too realistic. For those studies with the consideration of transit passengers' behavior, they usually adopt classic assumptions for boarding and alighting, and ignore the effect of advanced passenger information system, real time transit information, and seat availability on their choice. More realistic passenger's behavior should be captured in this field in the future.

7.3. Future Research for Multi-Modal Network Design Problem (MMNDP)

Decisions

- The time-dependent approach for MMNDP has not been considered and can be a future extension for this field. However, it adds to the complexity of the problem.

Multi-modality

- As previously mentioned, the issue of multi-modality has been largely ignored in road and transit network design problems. In fact, the literature of MMNDP is very limited, while a number of future directions are possible in this field by extending the single mode problems to include various forms of multi-modality.
- Multimodality with no interaction between flows of different modes: In RNDP, public transit modes such as buses in exclusive lanes, trams, or metro can be considered such that they have no effects on vehicle flows on roads. The demand shares will vary as road network configuration changes and consequently disutility levels of modes compared to each other will change. The existence of private and public transit modes in the problem

can imply the multi-objective nature of the problem because of the existence of different objectives that can be related to individual modes, all travelers, community, etc. The only two existing studies in this field are Cantarella and Vitetta (2006) for multi-objective signal setting and one-way street orientation problem with a fixed travel time bus network, and Szeto et al. (2010) for lane additions and toll setting. New problems can be proposed by considering other road network improvement decisions, in the presence of a tram or a metro network. Such problems may have multiple objective functions (e.g. total user benefit, demand share of metro or tram, air pollution, etc.).

- Multimodality with interaction between flows of different modes: In an existing fixed bus network that shares the same roads with other vehicles, the road network improvement decisions must take into account the restrictions caused by the bus network, the changes in mode demand shares, and both bus and non-bus travel times. Different objective functions and various constraints are required in these problems to consider multiple performance criteria such as the demand share of each mode, total user benefit, travel cost by bus mode, etc.
- Multimodality with interrelations between flows of different modes and in decisions: Although this issue has been recently studied by some researchers (e.g., Miandoabchi et al., 2011a, b), more new problems can be proposed along this direction. For example, the turn-restriction problem (Long et al., 2009) has been proposed but the impact of turn-restrictions on transit routes has been ignored. One new problem can be to simultaneously design the turn restrictions and transit routes. Besides, the previous studies only consider changes in the routes of existing bus lines between terminal stations. Another extension can be designing the whole bus network in combination with various RNDP and MMNDP decisions.

Inter-modal Connectivity

- MMNDP traditionally assumes single mode trips and ignore the possibility of combined mode trips such as *Park'n Ride*. Combined mode trips (i.e. multi-modal trips) indeed arise the inter-modal connection issue. This has been considered in the literature for some operational level decisions such as pricing, but no network topology design problem has yet considered this. In such problems, an important issue for travelers is to be able to transfer conveniently from one mode to another to finish their trips. This means that the network of each transit mode must be connected in the best way to provide an integrated and connected public transit network. This integration can only become possible under the condition that when designing the transit network of one mode, the locations of transit stops of other modes are explicitly considered. Another extension would be an RNDP that considers an existing metro network (for car-metro travels) and accounts for the allocation of parking spaces in the proximity of metro stations, such that travelers can access the metro network by their cars in the most convenient way.

Lower-level problems

- The adoption of combined trip concept for inter-modal connectivity problems also sets new requirements for the lower level problem to be used. One requirement is to capture transfer penalty in the behavior principle used within the lower level problem, since some travelers weigh transfer penalty heavily and they do not just select a combination of modes based on the out-of-pocket cost. Whether a transfer involving long walking time, long waiting time and walking uphill or going up stairs are important considerations from the travelers' perspective. The transfer penalty may not be additive because having two transfers can be weighed more than double than having one transfer. Hence, developing a realistic travel choice principle that can capture transfer penalty and how to model transfer penalty in the lower level network are two future research directions.
- Another requirement is to model the waiting time accurately since travelers often need to wait at transit stops

when they transfer from one mode to another. In reality, the waiting time varies from time to time due to different service frequency of various public transit modes and time-varying road congestion. However, in the literature, most of them only capture expected waiting time in the lower level problem since the micro-grained temporal level is not considered. Hence, one future research direction is to develop a methodology to determine waiting time at transfer locations over time while taking into account both public and private modes.

7.4. Future Research for Solution Techniques

Metaheuristics

- The metaheuristics shown in Figure 2 are actually those classic or well-known metaheuristics that have also been applied to other problems. It is observed that genetic algorithm and simulated annealing are the most prevalent metaheuristics among the others. Furthermore, applications of a number of recently developed methods such as artificial bee colony, swarm particle optimization, harmonic search, and clonal selection algorithm can be found in some studies. Although genetic algorithm has been proved to give better results than other algorithms in several studies, it is not clear whether other recently developed metaheuristics can outperform it. Instances of such algorithms are differential evolution method, firefly algorithm, cuckoo search, artificial immune systems, and generalized evolutionary walk algorithm (Gendreau and Potivn, 2010; Koziel and Yang, 2011; Yang, 2011). Therefore, one future direction can be to compare the performance of recently developed heuristics with genetic algorithm and the most commonly used algorithms to solve UTNDP.
- Each metaheuristic has its own strength and each problem has its own properties. It is therefore important to identify and analyze the problem properties, and develop tailored solution methods for the problem. Solution methods can be developed for new and existing network design problems based on hybrids of different metaheuristics. Most of hybridization research in UTNDP has been done with simulated annealing and tabu search with genetic algorithm and a number of metaheuristics. While other possibilities still remain to hybridize recently developed metaheuristics or to incorporate the strength of various well-known metaheuristics to solve UTNDP.
- Computational complexity of many classes of UTNDP is a major barrier of efficiency of conventional metaheuristics. Parallelizing strategies and distributed computing can improve the efficiency but there are few applications in TNDP and TNDFSP (e.g. Agarwal and Mathew, 2004; Yu et al., 2005; Yang et al., 2007; Cipriani et al., 2010) and have never considered in RNDP and MMNDP to the best of our knowledge. Hence the incorporation of these considerations into solution algorithms is definitely one future direction.
- Most studies of TNDSP and most multi-objective analyses of RNDP consider more than one objective but the resultant problems are usually formulated as a single objective problem using a weighted sum objective function. This approach can actually simplify the problem but we need to predetermine the weight for each objective function. Consequently, this method precludes some Pareto solutions. Alternatively, we can avoid presetting the weight and solve the multi-objective problem directly using existing multi-objective metaheuristics to determine the Pareto frontier and the modeler can then determine the best solution in the frontier. This alternative is rarely considered in TNDSP and should be considered in the future.

Hybrid methods and non-metaheuristics

- For UTNDP with discrete variables, it was usually solved by metaheuristics or the branch and bound method. Metaheuristics can give nearly optimal solutions quickly whereas the branch and bound method can be global optimal solutions at the expensive of computation time. To make a good tradeoff between solution speed and quality, one can be developed a hybrid of the branch and bound method and metaheuristics, where the metaheuristics are used to find good upper and lower bounds whereas the branch and bound method is used to check for global optimality.

- SAB heuristic has been used for finding descent directions of RNDP but the heuristic has rarely extended and applied to solve TNDSP or MMNDP with a bi-level structure, probably because there are few TNDSPs or MMNDPs with a bi-level structure. Incorporating SAB heuristic into TNDSP and MMNDP can surely speed up the computation process and is one potential research direction. Other speeding up strategies should also be considered.
- As reflected in Section 4.1, most of studies of RNDP are directed towards finding more efficient non-metaheuristic solution methods for continuous variable problems. There are few works which has attempted to develop efficient non-metaheuristic solution procedures to solve discrete variable problems. It seems that more effort can be put on developing faster and better heuristics for RNDP.

Solution procedures for lower level problems

- When solving the SUE problem or the transit assignment problem in UTNDP, current methods rely on the method of successive averages, which has a slow convergence rate. To speed up the computation process, one can employ newly developed self-regulated method proposed by Liu et al. (2009). This method can be integrated with existing solution methods for solving bi-level UTNDP.
- Solving the lower-level problem is often the major source of the computational burden in UTNDP. This can be done by devising faster solution methods, or by incorporating approximation methods such as neural networks (as proposed by Cantarella et al. (2006)) to obtain the travel time and flow values.

Acknowledgements

The research was jointly supported by a grant (200902172003) from the Hui Oi Chow Trust Fund and two grants (201001159008 and 201011159026) from the University Research Committee of the University of Hong Kong. We are grateful to the comments from two anonymous referees.

References

Abdulaal, M., & LeBlanc, L. J. (1979). Continuous equilibrium network design models. *Transportation Research Part B, 13*(1), 19-32.

Agrawal, J., & Mathew, T. V. (2004). Transit route design using parallel genetic algorithm. *Journal of Computing in Civil Engineering, 18*(3), 248-256.

Baaj, M. H., & Mahmassani, H. S. (1991). An AI-based approach for transit route system planning and design. *Journal of Advanced Transportation*, 25(2), 187–210.

Baaj, M. H., & Mahmassani, H. S. (1995). Hybrid route generation heuristic algorithm for the design of transit networks. *Transportation Research Part C, 3*(1), 31-50.

Barra, A., Carvalho, L., Teypaz, N., Cung, V. D., & Balassiano, R. (2007). Solving the transit network design problem with constraint programming. Paper presented at the *Proceeding of 11th World Conference on Transport Research,* 24-28.

Beltran, B., Carrese, S., Cipriani, E., & Petrelli, M. (2009). Transit network design with allocation of green vehicles: A genetic algorithm approach. *Transportation Research Part C, 17*(5), 475-483.

Ben-Ayed, O., Boyce, D. E., & Blair III, C. E. (1988). A general bilevel linear programming formulation of the network design problem. *Transportation Research Part B, 22*(4), 311-318.

Bielli, M., Caramia, M., & Carotenuto, P. (2002). Genetic algorithms in bus network optimization. *Transportation Research Part C, 10*(1), 19-34.

Borndörfer, R., Grötschel, M., & Pfetsch, M. E. (2008). Models for line planning in public transport. In M. Hickman, P. Mirchandani & S. Voß (Eds.), *Computer-Aided Systems in Public Transport* (pp. 363-378) Springer Berlin/ Heidelberg.

Boyce, D. E. (1984). Urban transportation network-equilibrium and design models: Recent achievements and future prospects. *Environment and Planning A, 16*(1), 1445-1474.

Bussieck, M. R. (1998). *Optimal lines in public rail transport.* Braunschweig University of Technology, Germany.

Cantarella, G. E., Improta, G., & Sforza, A. (1991). Road network signal setting: equilibrium conditions. In M. Papageorgiou

(Ed.), *Concise Encyclopedia of Traffic and Transportation Systems* (pp. 366-371) Pergamon Press.

Cantarella, G. E., Pavone, G., & Vitetta, A. (2006). Heuristics for urban road network design: Lane layout and signal settings. *European Journal of Operational Research, 175*(3), 1682-1695.

Cantarella, G., & Vitetta, A. (2006). The multi-criteria road network design problem in an urban area. *Transportation, 33*(6), 567-588.

Carrese, S., & Gori, S. (2004). Chapter 11: An urban bus network design procedure. In M. Patriksson, & M. Labbé (Eds.), *Transportation planning: State of the art (applied optimization) 64* (pp. 177-196) Springer.

Cascetta E., Gallo, M., & Montella, B. (2006). Models and Algorithms for the Optimization of Signal Settings on Urban Networks with Stochastic Assignment. *Annals of Operations Research*, *144*(1), 301-328.

Ceder, A. (2003). Chapter 3: Designing public transport network and routes. In W. Lam, & M. Bell (Eds.), *Advanced Modeling for Transit Operations and Service Planning* (pp. 59-91) Pergamon Imprint, Elsevier Science Ltd Pub.

Ceder, A., & Wilson, N. H. M. (1986). Bus network design. *Transportation Research Part B, 20*(4), 331-344.

Chakroborty, P. (2003). Genetic algorithms for optimal urban transit network design. *Journal of Computer Aided Civil and Infrastructure Engineering*, 18, 184-200.

Chakroborty, P., & Dwivedi, T. (2002). Optimal route network design for transit systems using genetic algorithms. *Engineering Optimization, 34*(1), 83-100.

Chen, M., & Alfa, A. S. (1991). A network design algorithm using a stochastic incremental traffic assignment approach. *Transportation Science, 25*(3), 215-224.

Chiou, S. (2005). Bilevel programming for the continuous transport network design problem. *Transportation Research Part B, 39*(4), 361-383.

Chiou, S. (2008). A hybrid approach for optimal design of signalized road network. *Applied Mathematical Modelling, 32*(2), 195-207.

Cipriani, E., Gori, S., & Petrelli, M. (2010). Transit network design: A procedure and an application to a large urban area. *Transportation Research Part C, In Press.*

Cipriani, E., Fusco, G., & Petrelli, M. (2006). A multimodal transit network design procedure for urban areas. *Journal of Advances in Transportation Studies*, *Section A, 10*, 5-20.

Daganzo, C. F., & Sheffi, Y. (1977). On stochastic models of traffic assignment. *Transportation Science, 11*(3), 253-274.

Dantzig, G. B., Harvey, R. P., Lansdowne, Z. F., Robinson, D. W., & Maier, S. F. (1979). Formulating and solving the network design problem by decomposition. *Transportation Research Part B, 13*(1), 5-17.

Davis, G. A. (1994). Exact local solution of the continuous network design problem via stochastic user equilibrium assignment. *Transportation Research Part B, 28*(1), 61-75.

Desaulniers, G., & Hickman, M. D. (2007). Chapter 2: Public transit. In C. Barnhart, & G. Laporte (Eds.), *Handbooks in operations research and management science* (pp. 69-127) Elsevier.

Dimitriou, L., Tsekeris, T., & Stathopoulos, A. (2008a). Genetic computation of road network design and pricing stackelberg games with multi-class users. In M. Giacobini et al. (Eds.), *Applications of evolutionary computing* (pp. 669-678) Springer Berlin/ Heidelberg.

Dimitriou, L., Stathopoulos, A., & Tsekeris, L. (2008b). Reliable stochastic design of road network systems. *International Journal of Industrial and Systems Engineering, 3*(5), 549-574

Drezner, Z., & Salhi, S. (2000). Selecting a good configuration of one-way and two-way routes using tabu search. *Control and Cybernetics, 29*(3), 725-740.

Drezner, Z., & Salhi, S. (2002). Using hybrid metaheuristics for the one-way and two-way network design problem. *Naval Research Logistics (NRL), 49*(5), 449-463.

Drezner, Z., & Wesolowsky, G. O. (1997). Selecting an optimum configuration of one-way and two-way routes. *Transportation Science, 31*(4), 386-394.

Drezner, Z., & Wesolowsky, G. O. (2003). Network design: Selection and design of links and facility location. *Transportation Research Part A, 37*(3), 241-256.

Dubois, D., Bel, G., & Llibre, M. (1979). A set of methods in transportation network synthesis and analysis. *The Journal of the Operational Research Society, 30*(9), 797-808.

Elshafei, E. H. (2006). *Decision-making for roadway lane designation among variable modes.* University of Maryland, USA.

Fan, W., & Machemehl, R. B. (2004). *Optimal transit route network design problem: Algorithms, implementations, and numerical results.* No. SWUTC/04/ 167244-1. Austin, Texas: Center for Transportation Research, University of Texas.

Fan, W., & Machemehl, R. B. (2006a). Optimal transit route network design problem with variable transit demand: Genetic algorithm approach. *Journal of Transportation Engineering, 132*(1), 40-51.

Fan, W., & Machemehl, R. B. (2006b). Using a simulated annealing algorithm to solve the transit route network design problem. *Journal of Transportation Engineering, 132*(2), 122-132.

Fan, W., & Machemehl, R. B. (2008). Tabu search strategies for the public transportation network optimisations with variable transit demand. *Computer-Aided Civil and Infrastructure Engineering*, *23*(7), 502-520.

Fan, L., & Mumford, C. (2010). A metaheuristic approach to the urban transit routing problem. *Journal of Heuristics, 16*(3), 353-372.

Frank, M., & Wolfe, P. (1956). An algorithm for quadratic programming. *Naval Research Logistics Quarterly, 3*(1-2), 95-110.

Friesz, T. L. (1985). Transportation network equilibrium, design and aggregation: Key developments and research opportunities. *Transportation Research Part A, 19*(5-6), 413-427.

Friesz, T. L., Anandalingam, G., Mehta, N. J., Nam, K., Shah, S. J., & Tobin, R. L. (1993). The multiobjective equilibrium network design problem revisited: A simulated annealing approach. *European Journal of Operational Research, 65*(1), 44-57.

Friesz, T. L., Cho, H., Mehta, N. J., Tobin, R. L., & Anandalingam, G. (1992). A simulated annealing approach to the network design problem with variational inequality constraints. *Transportation Science, 26*(1), 18-26.

Fusco, G., Gori, S., & Petrelli, M. (2002). A heuristic transit network design algorithm for medium size towns. *Proceedings of the 13th Mini-EURO Conference.*

Gallo, M., D'Acierno, L., & Montella, B. (2010). A meta-heuristic approach for solving the urban network design problem. *European Journal of Operational Research, 201*(1), 144-157.

Gallo, M., Montella, B., & D'Acierno, L. (2011). The transit network design problem with elastic demand and internalisation of external costs: An application to rail frequency optimization. *Transportation Research Part C: Emerging Technologies*, *19*(6), 1276-1305.

Gao, Z., Sun, H., & Zhang, H. (2007). A globally convergent algorithm for transportation continuous network design problem. *Optimization and Engineering, 8*(3), 241-257.

Gao, Z., Wu, J., & Sun, H. (2005). Solution algorithm for the bi-level discrete network design problem. *Transportation Research Part B, 39*(6), 479-495.

Gendreau, M., & Potvin, J.-Y. (2010) *Handbook of Metaheuristics*. Second Edition. Springer-Verlag, New York, USA.

Guan, J. F., Yang, H., & Wirasinghe, S. C. (2006). Simultaneous optimization of transit line configuration and passenger line assignment. *Transportation Research Part B, 40*(10), 885-902.

Guihaire, V., & Hao, J. (2008). Transit network design and scheduling: A global review. *Transportation Research Part A, 42*(10), 1251-1273.

Hasselström, D. (1979). *A method for optimization of urban bus route networks.* Göteborg, Sweden: Volvo Bus Corporation.

Hasselström, D. (1981). *Public transportation planning - a mathematical programming approach.* Göteborg, Sweden.

Hu, J., Shi, X., Song, J., & Xu, Y. (2005). Optimal design for urban mass transit network based on evolutionary algorithms. In L. Wang, K. Chen & Y. Ong (Eds.), *Advances in natural computation* (pp. 429-429) Springer Berlin/ Heidelberg.

Kepaptsoglou, K., & Karlaftis, M. (2009). Transit route network design problem: Review. *Journal of Transportation Engineering, 135*(8), 491-505.

Koziel, S., & Yang, X. S. (2011). *Computational Optimization, Methods and Algorithms*. Springer Berlin/Heidelberg.

Lampkin, W., & Saalmans, P. D. (1967). The design of routes, service frequencies, and schedules for a municipal bus undertaking: A case study. *Operational Research Quarterly, 18*(4), 375-397.

Leblanc, L. J. (1975). An algorithm for the discrete network design problem. *Transportation Science, 9*(3), 183-199.

LeBlanc, L. J., & Boyce, D. E. (1986). A bilevel programming algorithm for exact solution of the network design problem with user-optimal flows. *Transportation Research Part B, 20*(3), 259-265.

Lee, C., & Yang, K. (1994). Network design of one-way streets with simulated annealing. *Papers in Regional Science, 73*(2), 119-134.

Lee, Y., & Vuchic, V. R. (2005). Transit network design with variable demand. *Journal of Transportation Engineering, 131*(1), 1-10.

Li, S. G., & Ju, Y. F. (2009). Evaluation of bus-exclusive lanes. *IEEE Transactions on Intelligent Transportation Systems, 10*(2), 236-245.

Liu, H. X., He, X. Z. & He, B. S. (2009). Method of successive weighted averages (MSWA) and self-regulated averaging schemes for solving stochastic user equilibrium problem. *Networks and Spatial Economics*, *9*(4), 485-503.

Lo, H. K., & Szeto, W. Y. (2004). Chapter 9: Planning transport network improvements over time. In D. Boyce, & D. H. Lee (Eds.), *Urban and Regional Transportation Modeling: Essays in Honor of David Boyce.* (pp. 157-176). Northampton (Mass.): E. Elgar, cop.: Cheltenham (G.B.).

Lo, H. K., & Szeto, W. Y. (2009). Time-dependent transport network design under cost-recovery. *Transportation Research Part B, 43*(1), 142-158.

Long, J., Gao, Z., Zhang, H., & Szeto, W. Y. (2010). A turning restriction design problem in urban road networks. *European Journal of Operational Research, 206*(3), 569-578.

Luo, Z., Pang, J., & Ralph, D. C. (1996). *Mathematical programs with equilibrium constraints*. Cambridge: Cambridge Univ. Press.

Magnanti, T. L., & Wong, R. T. (1984). Network design and transportation planning: Models and algorithms. *Transportation Science, 18*(1), 1-55.

Mandl, C. E. (1980). Evaluation and optimization of urban public transportation networks. *European Journal of Operational Research, 5*(6), 396-404.
Marcotte, P. (1986). Network design problem with congestion effects: A case of bilevel programming. *Mathematical Programming, 34*(2), 142-162.
Marcotte, P. (1983). Network optimization with continuous control parameters. *Transportation Science, 17*(2), 181-197.
Marcotte, P., & Marquis, G. (1992). Efficient implementation of heuristics for the continuous network design problem. *Annals of Operations Research, 34*(1), 163-176.
Mathew, T. V., & Shrama, S. (2009). Capacity expansion problem for large urban transportation networks. *Journal of Transportation Engineering, 135*(7), 406-415.
Mauttone, A., & Urquhart, M. E. (2009). A route set construction algorithm for the transit network design problem. *Computers & Operations Research, 36*(8), 2440-2449.
Meneguzzer, C. (1995). An equilibrium route choice model with explicit treatment of the effect of intersections. *Transportation Research Part B*, *29*(5), 329-356.
Meng, Q., Yang, H., & 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, 35*(1), 83-105.
Meng, Q., & Yang, H. (2002). Benefit distribution and equity in road network design. *Transportation Research Part B, 36*(1), 19-35.
Mesbah, M., Sarvi, M., & Currie, G. (2008). New methodology for optimizing transit priority at the network level. *Transportation Research Record, 2089*, 93-100.
Miandoabchi, E., & Farahani, R. Z. (2010). Optimizing reserve capacity of urban road networks in a discrete network design problem. *Advances in Engineering Software*, *42*(12), 1041-1050.
Miandoabchi, E., Farahani, R. Z., Dullaert, W., & Szeto, W. Y. (2011a). Hybrid evolutionary metaheuristics for concurrent multi-objective design of urban road and public transit networks. *Networks and Spatial Economics,* 1-40. doi:10.1007/s11067-011-9163-x.
Miandoabchi, E., Farahani, R. Z., & Szeto, W. Y. (2011b). Bi-objective bimodal urban road network design using hybrid metaheuristics. *Central European Journal of Operations Research*, 1-39. doi:10.1007/s10100-011-0189-4.
Migdalas, A. (1995). Bilevel programming in traffic planning: Models, methods and challenge. *Journal of Global Optimization, 7*(4), 381-405.
Ngamchai, S., & Lovell, D. J. (2003). Optimal time transfer in bus transit route network design using a genetic algorithm. *Journal of Transportation Engineering, 129*(5), 510-521.
O'Brien, L., & Szeto, W. Y. (2007). The discrete network design problem over time. *HKIE Transactions, 14*(4), 47-55.
Pacheco, J., Alvarez, A., Casado, S., & González-Velarde, J. L. (2009). A tabu search approach to an urban transport problem in northern Spain. *Computers & Operations Research, 36*(3), 967-979.
Pape, U., Reinecke, E., & Reinecke, Y. (1992). Entwurf und implementierung eines linienplanungssystems fr den busverkehr im pnv unter einer objektorientierten grafischen entwicklungsumgebung. *Gruppendiplomarbeit,*
Pattnaik, S. B., Mohan, S., & Tom, V. M. (1998). Urban bus transit route network design using genetic algorithm. *Journal of Transportation Engineering, 124*(4), 368-375.
Patz, A. (1925). Die richtige auswahl von verkehrslinien bei großen Straßenbahnnetzen.*50/51*
Poorzahedy, H., & Abulghasemi, F. (2005). Application of ant system to network design problem. *Transportation, 32*(3), 251-273.
Poorzahedy, H., & Rouhani, O. M. (2007). Hybrid meta-heuristic algorithms for solving network design problem. *European Journal of Operational Research, 182*(2), 578-596.
Poorzahedy, H., & Turnquist, M. A. (1982). Approximate algorithms for the discrete network design problem. *Transportation Research Part B, 16*(1), 45-55.
Powell, W.B., & Sheffi, Y., 1982. The convergence of equilibrium algorithms with predetermined step sizes. *Transportation Science 16*(1), 45–55.
Seo, Y. U., Park, J. H., Jang, H., & Lee, Y. I. (2005). A study on setting-up methodology and criterion of exclusive bus lane in urban area. *Proceeding of the Eastern Asia Society for Transportation Studies, Volume 5,* 325-338.
Shih, M., & Mahmassani, H. S. (1994). *A design methodology for bus transit networks with coordinated operations.* No. SWUTC/94/60016-1. University of Texas, Austin: Center for Transportation Research.
Shih, M., Mahmassani, H. S., & Baaj, M. H. (1998). Planning and design model for transit route networks with coordinated operations. *Transportation Research Record, 1623*, 16-23.
Silman, L. A., Barzily, Z., & Passy, U. (1974). Planning the route system for urban buses. *Computers & Operations Research, 1*(2), 201-211.
Sonntag, H. (1977). *Linienplanung im öffentlichen personennahverkehr.* Technische Universitt, Berlin.
Spiess, H., & Florian, M. (1989). Optimal strategies: A new assignment model for transit networks. *Transportation Research Part B, 23*(2), 83-102.

Steenbrink, P. A. (1974a). *Optimization of transport network*. New York, USA: Wiley.

Steenbrink, P. A. (1974b). Transport network optimization in the Dutch integral transportation study. *Transportation Research, 8*(1), 11-27.

Suh, S., & Kim, T. (1992). Solving nonlinear bilevel programming models of the equilibrium network design problem: A comparative review. *Annals of Operations Research, 34*(1), 203-218.

Szeto, W. Y., Jaber, X., & O'Mahony, M. (2010). Time-dependent discrete network design frameworks considering land use. *Computer-Aided Civil and Infrastructure Engineering, 25*(6), 411-426.

Szeto, W. Y., & Lo, H. K. (2005). Strategies for road network design over time: Robustness under uncertainty. *Transportmetrica, 1*(1), 47-63.

Szeto, W. Y., & Lo, H. K. (2006). Transportation network improvement and tolling strategies: The issue of intergeneration equity. *Transportation Research Part A, 40*(3), 227-243.

Szeto, W. Y., & Lo, H. K. (2008). Time-dependent transport network improvement and tolling strategies. *Transportation Research Part A, 42*(2), 376-391.

Szeto, W. Y., & Wu, Y. (2011). A simultaneous bus route design and frequency setting problem for Tin Shui Wai, Hong Kong. *European Journal of Operational Research, 209*(2), 141-155.

Tom, V. M., & Mohan, S. (2003). Transit route network design using frequency coded genetic algorithm. *Journal of Transportation Engineering, 129*(2), 186-195.

Ukkusuri, S. V., Mathew T. V., & Waller, S. T. (2007). Robust transportation network design under demand uncertainty. *Computer-Aided Civil and Infrastructure Engineering, 22*(1), 6-18

Ukkusuri, S. V., Waller S. T. (2007). Linear programming models for the user and system optimal dynamic network design problem: formulations, comparisons and extensions. *Networks and Spatial Economics*, *8*(4), 383-406.

van Nes, R., Hamerslag, R., & Immers, B. H. (1988). Design of public transport networks. *Transportation Research Record, 1202*, 74-83.

Wan, Q., & Lo, H. K. (2003). A mixed integer formulation for multiple-route transit network design. *Journal of Mathematical Modelling and Algorithms, 2*(4), 299-308.

Wang, D. Z. W., & Lo, H. K. (2010). Global optimum of the linearized network design problem with equilibrium flows. *Transportation Research Part B, 44*(4), 482-492.

Wardrop, J. (1952). Some theoretical aspects of road traffic research. *Proceedings of the Institute of Civil Engineers, Part II,* 325-378.

Wey, W. M. (2000). Model formulation and solution algorithm of traffic signal control in an urban network. *Computers Environment and Urban Systems*, *24*(4), 355-377.

Wong, S. C., & Yang, C. (1999). An iterative group-based signal optimization scheme for traffic equilibrium networks. *Journal of Advanced Transportation*, *33*(2), 201-217.

Wu, J. J., Sun, H. J., Gao, Z. Y., & Zhang, H. Z. (2009). Reversible lane-based traffic network optimization with an advanced traveller information system. *Engineering Optimization, 41*, 87-97.

Xiong, Y., & Schneider, J. B. (1992). Transportation network design using a cumulative algorithm and neural network. *Transportation Research Record,* 1364, 37-44.

Xu, T., Wei, H., & Hu, G. (2009). Study on continuous network design problem using simulated annealing and genetic algorithm. *Expert Systems with Applications, 36*(2, Part 1), 1322-1328.

Yang, H., & Wang, J. Y. T. (2002). Travel time minimization versus reserve capacity maximization in the network design problem. *Transportation Research Record, 1783*, 17-26.

Yang, Z., Yu, B., & Cheng, C. (2007). A parallel ant colony algorithm for bus network optimization. *Computer Aided Civil and Environmental Engineering, 22*(1), 44-55.

Yang, H. (1997). Sensitivity analysis for the elastic-demand network equilibrium problem with applications. *Transportation Research Part B, 31*(1), 55-70.

Yang, H., & Bell, M. G. H. (1998a). Models and algorithms for road network design: A review and some new developments. *Transport Reviews, 18*(3), 257-278.

Yang, H., & Bell, M. G. H. (1998b). A capacity paradox in network design and how to avoid it. *Transportation Research Part A, 32*(7), 539-545.

Yang, X. S. (2011). *Nature inspired metaheuristic algorithms*. Luniver Press, London, UK.

Yu, B., Yang, Z., Cheng, C., & Liu, C. (2005). Optimizing bus transit network with parallel ant colony algorithm. *Proceedings of the Eastern Asia Society for Transportation Studies, Volume 5,* 374-389.

Zhang, H., & Gao, Z. (2007). Two-way road network design problem with variable lanes. *Journal of Systems Science and Systems Engineering, 16*(1), 50-61.

Zhang, H., & Gao, Z. (2009). Bilevel programming model and solution method for mixed transportation network design problem. *Journal of Systems Science and Complexity, 22*, 446-459.

Zhao, F., & Zeng, X. (2007). Optimization of user and operator cost for large scale transit networks. *Journal of Transportation*

Engineering, 133(4), 240-251.

Zhao, F. (2006). Large-scale transit network optimization by minimizing user cost and transfers. *Journal of Public Transportation, 9*(9), 107-129.

Zhao, F., & Gan, A. (2003). *Optimization of transit network to minimize transfers.* No. BD-015-02). Miami, Florida: Research Center of Department of Transportation, Florida International University.

Zhao, F., & Zeng, X. (2006). Simulated annealing-genetic algorithm for transit network optimization. *Journal of Computing in Civil Engineering, 20*(1), 57-68.

Zhao, F., & Ubaka, I. (2004). Transit network optimization – minimizing transfers and optimizing route directness. *Journal of Public Transportation, 7*(1), 67-82.

Ziyou, G., & Yifan, S. (2002). A reserve capacity model of optimal signal control with user-equilibrium route choice. *Transportation Research Part B, 36*(4), 313-323.