# A METAHEURISTIC METHODOLOGY FOR FEEDER BUS NETWORK DESIGN PROBLEM

*F. Ciaffi - Università degli Studi Roma Tre - francesco.ciaffi@uniroma3.it*

*E. Cipriani - Università degli Studi Roma Tre - ernesto.cipriani@uniroma3.it*

*M. Petrelli - Università degli Studi Roma Tre - marco.petrelli@uniroma3.it*

## ABSTRACT

The present paper deals with the transit network design problem related to the feeder bus lines defined as transit services for connecting suburban areas, where the demand has to be gathered, with the stops of the mass transit network, usually railway or underground stations. The objective of the research is the development of a procedure that simultaneously generates routes and frequencies of the feeder bus network in a actual-size large urban area. The solving procedure is articulated in 3 phases: in the first one, the study area is divided in several different basins, one for each main transit stop, on the base of some criteria like urban framework, main skeleton of the road network and pedestrian accessibility. In the second one, a heuristic algorithm generates two different and complementary sets of feasible routes, in order to provide a good balance between the maximization of the service coverage area and the minimization of the overall travel time. First set is composed by circular routes, generated solving a travelling salesman problem (TSP) connecting the highest demand node pairs in the area with the stop of main transit network. The second feasible set aims at developing feeder routes more direct than the others using the k-shortest path algorithm. The set of all feasible routes, generated taking into account only the main skeleton of the road network, is then the input data for the third phase where a GA is utilized for finding a sub-optimal set of routes with the associated frequencies.

The proposed procedure has been implemented on a real-life size network, Rome, in order to compare its effectiveness with the performances of the existing transit network. The results of the application of the design procedure show that the feeder routes imply a more integrated transit network with a reduction of the total travel time, despite an increase of the number of transfers, in a more efficient way for the reduction of the operating costs and for an increase of the average load factor.

*Keywords: feeder bus network design, genetic algorithms, heuristic route generation*

# 1  INTRODUCTION

In the last decades we are witnessing to a huge developments of policies and strategies for managing and planning urban transit networks, mainly due to the suburban sprawl that can be found all over the world and the resulting overuse of private cars. Differently from past studies and experiences, nowadays most of researchers have to take into account main issues of sustainable mobility, introduced from 80's just for facing increases of air and noise pollution, energy consumption, traffic congestion and car accidents. Designing a reliable and functional transit network represents one of the cheapest and most effective tools to achieve final goals of sustainability.

Based on such issues, this paper presents a methodology for solving the "Feeder Bus Network Design Problem" ("FBNDP") whose solution seems very useful to improve integration between main urban transit services, like rail and bus networks. FBNDP represents one of those new policies developed in last decade for managing and  planning urban transit networks since it allows to design those specific bus lines suitable to cover residential areas gathering transit demand and feeding main transit system in specific transfer points. Therefore each step of the solving model proposed in this work has been completed taking into account all the typical FBNDP goals; the sub-optimal set of lines, routes and frequencies was designed aiming at the balance between service coverage areas and service effectiveness and efficiency, and looking for the improvement of the integration between rail and bus networks in urban area of Rome.

Even if the solving model is inspired to a previous model by Cipriani et al. (2009), dealing with the mass transit network design, main novelty of this paper is represented by the adjustment of the procedure to solve the feeder bus network design problem. First of all, a basin definition procedure and a zone aggregation procedure are introduced in order to define mass transit stations to be fed and to properly apply the HRGA, making the solving procedure suitable to different study areas. Moreover an innovative structure of HRGA allows to design two sets of preliminary feasible routes according to two different design criteria, both inspired to the typical goals of feeder bus services just described above. The robustness of this new solving procedure make possible the application of innovative algorithms for optimal solution calculation, in order to compare their efficiency and effectiveness with Genetics. This evolution represents one the further developments of this model, still now studied to reduce computational times and to improve the quality of the solutions provided.

The structure of this paper is based on the procedure framework which represents the different steps of the work: firstly it is shown a wide review of other works about feeder bus networks design; next chapters, instead, deal with the problem formulation and his mathematical pattern and, most importantly, the solving procedure framework, based on three different phases. Finally is shown the application of the model to two different real-life size networks: the city of Winnipeg, in order to test the procedure quality, and the urban area of Rome, so as to compare its effectiveness with the performances of the existing transit networks.

# 2 STATE OF THE ART

The transit network design is a complex non convex problem (Newell 1979, Baaj and Mahmassani, 1991). It is usually formulated as a non linear optimization problem with both discrete and continuous variables and constraints. The best and most efficient solution methods are based on heuristic procedures but their applications are mainly limited to test cases or real-life networks of small size. A global review about route design, frequency setting, timetabling of transit lines, and their combination is proposed by Carrese and Gori (2002), Desaulniers and Hickman (2007), Guihaire and Hao (2008), Kepaptsoglou and Karlaftis (2009).

Among the most remarkable works on this matter, we should mention Baaj and Mahmassani (1995), whose work proposes an Artificial-Intelligent heuristic algorithm for route generation. This algorithm selects a given number of high-demand node pairs and builds an initial network skeleton by connecting these node pairs through the shortest paths. The skeleton is then progressively expanded to routes according to a node selection strategy that reflects different trade-offs between performance measures and users' and operators' costs.

Over the last years, the evolution of operational research and computer technology has produced great and renewed attention for the transit network design problem; also feeder bus network planning has been taken into account. Most of these remarkable studies about feeder bus network design mainly deals with heuristic procedure for routes generation rather than the solving procedure techniques; among them, one of the most interesting is proposed by Jerby and Ceder (2006) with their threefold research: create a method for estimating potential demand for a shuttle bus service, elaborate a model focusing on optimal route automatic design and, finally, define a heuristic algorithm designed to take into account road networks of all sizes. All the research is based on a modular approach allowing the entire problem to be partitioned in a chain of sub-problems; first four stages are to estimate the potential demand, creating a base network of a defined service area using inputs and constraints like average travel speeds, maximum travel time and walking distances after the scan of urban entire network and the discard of all the links not used by transit service. Fifth and sixth steps show the model formulation proposed for this particular feeder route design problem, based on a decision variable aiming at maximizing potential demand and minimizing walking distances. Last two stages of the threefold research are about the heuristic algorithm framework, introduced to define circular routes in urban context, both starting and finishing in the same node with a total travel time lower than a fixed constraint.

Feeder bus networks differs from usual bus networks since they are located not only on main direct-through streets but also on secondary roads in order to reduce passengers' access impedance. Therefore tortuous routes could be considered although they might increase in-vehicle time for users and total cost for operators. Chien and Yang (2000) proposed in their work an alternative methodology for solving feeder bus route design problem in a typical urban irregularly shaped service area, based on a model for finding just the optimal bus route

location and its operating headway. All this research starts from the assumption that demand is uniformly distributed within each zone of the service area but differs among zones and that it is no sensitive to service quality. An algorithm allows to design not the entire feeder bus network but a single bus service optimal route, based on a many-to-one travel pattern, aiming to the maximization of service coverage and demand collection in the service area, having a line-haul distance of 10 km from the transfer station to be fed.

New approaches based on metaheuristic techniques (Genetic Algorithm –GA–, Simulated Annealing or Tabu Search) have been frequently applied to solve optimization problems. Due to the discrete nature of several variables of the transit network design problem as well as the nonlinearity and the non-convexity of its objective function, probabilistic optimization techniques such as GAs seem to be appropriate. Among all these researches about FBNDP and based on metaheuristic techniques, one of the most interesting is the methodology proposed by Shrivastava e O'Mahony (2006) and its application to real life case of DART, the rapid transit system of Dublin. In this paper, most appealing innovation is represented by the application of a Genetic Algorithm (GA) in order to obtain the sub-optimal set of feeder bus lines and simultaneously the associated frequencies leading to a schedule coordination with main transit system (DART). Hence we can consider at the same time this solving procedure useful for both routing problems and scheduling problems.

Differently from this research, our paper adopts a design approach similar to the one proposed by Cipriani et al. (2009). Specifically, methodology proposed in the present work is based on a route generation procedure specific for feeder bus network in order to design a basin of different and complementary lines; then, a genetic algorithm combines the candidate lines in order to find a sub-optimal network of feeder services.

## 3. PROBLEM FORMULATION

The feeder bus network design is formulated as an optimization problem consisting in the minimization of all resources and costs related to the public transport system with fixed demand. The optimization problem is subject to route choice model on transit network and a set of feasibility constraints on route length and line frequency.
The optimization problem can be formally defined as follows:

$$\left(\hat{\mathbf{r}},\hat{\mathbf{f}}\right) = \arg\min \; z\left(\mathbf{r},\mathbf{f},\mathbf{q}_t^*\right) \quad \text{(1)}$$

subject to hyperpath assignment on transit network

$$\left(\mathbf{q}_t^*\right) = \Lambda\left[\mathbf{C}_t\left(\mathbf{r},\mathbf{f}\right)\right] \quad \text{(2)}$$

and a set of feasibility constraints that define both minimum and maximal values for route length and bus frequency:

$$L_{\min} \leq L_i \leq L_{\max} \qquad (3)$$

$$f_{\min} \leq f_i \leq f_{\max} \qquad (4)$$

where the following notations have been introduced:

- $z$ is the objective function;
- $\mathbf{r}$ is the vector of routes;
- $\hat{\mathbf{r}}$ is the vector of optimal routes;
- $\mathbf{f}$ is the vector of lines frequencies;
- $\hat{\mathbf{f}}$ is the vector of optimal frequencies;
- $\mathbf{q}_t^*$ is the equilibrium vector of segment flows on the transit network;
- $\Lambda$ is the user route choice model function;
- $\mathbf{C}_t$ is the vector of path generalized costs on the transit network;
- $q_{hk,i}$ is the number of transit passengers on segment ($hk,i$) of line i;
- $fc_{max}$ is the maximum load factor;
- $Cv$ is the vehicle capacity;
- $f_i$ is the frequency of line i and $f_{min}$ and $f_{max}$ are its minimum and maximum value;
- $L_i$ is the length of line i and $L_{min}$ and $L_{max}$ are its minimum and maximum value.

The road network is represented with the definition of a unidirectional graph G= (N,E), where N is the set of nodes and E is the set of links representing connections between nodes. A route is a sequence of adjacent nodes in G while a line is specified as a pair (**r**,**f**).

The public transport supply is represented as a frequency based service. Equation (2) represents the demand-supply consistency constraints (assignment constraints). Such constraint corresponds to a hyperpath approach for the simulation of user choice behaviour on transit (see Spiess and Florian 1989). Transit capacity constraints are not considered, as they are included in the heuristic design procedure.

The feasibility constraints for route length (3) and line frequency (4) have been introduced. The required frequency of service on the resulting route does not exceed the maximum operationally implementable value because it is impractical to maintain as well as the length does not exceed a maximum allowable value because schedules are too difficult to maintain. Analogously, both frequency and length are not lower than a minimum value because it is not possible to serve a very close OD demand pair with a bus (walking would be a better option and no operator would maintain such a line) or to offer a very low frequency service in an urban context (it would be perceived by users as no service availability).

The objective function z is defined as the sum of operator's costs z1 and users' costs z2 plus an additional penalty related to the level of unsatisfied demand z3:

$$z\left(\mathbf{r},\mathbf{f},\mathbf{q}_t^*\right) = z_1\left(\mathbf{r},\mathbf{f}\right) + z_2\left(\mathbf{r},\mathbf{f},\mathbf{q}_t^*\right) + z_3\left(\mathbf{r},\mathbf{f},\mathbf{q}_t^*\right) \qquad (5)$$

The transit users' costs are a weighted sum of in-vehicle travel time, access time, waiting time and a transfer penalty. Transit operator's costs are computed as a combination of total bus travel distance and total bus travel time. To provide transit services to as many transit users as possible, another additional component is included in the objective function z. This supplementary term represents a penalty that is proportional to the unsatisfied transit demand of the design network. The third term reflects the need to reject the banal solution of minimum cost ("zero users and zero service"). Solutions characterized by large increase of the unsatisfied transit demand are also discarded. Thus objective function formulation, similar to the one proposed by Cipriani et al. (2009) for main transit network design, is developed to properly weight "$z_1$", "$z_2$" and "$z_3$" terms, in order to represent specific needs of feeder bus networks.

The input data are the public transport demand matrix, the characteristics of the road network and the rapid rail transit system, the operating and users' unit costs. Outputs are bus routes and their frequency as well as the total costs and the vector of flows on the public transport network.

# 4. SOLUTION APPROACH

The proposed solution framework consists of three main stages:

1. a basin definition and zone aggregation procedure, in order to identify and select areas and nodes to be served;
2. a heuristic route generation algorithm (HRGA) that generates a large and rational set of feasible routes, by applying different design criteria and practical rules;
3. a genetic algorithm (GA) that finds the optimal network of routes and their frequencies.

Of course, given the well-known non-convexity of the problem and the heuristic nature of the method, there is no guarantee that the solution found, indicated as [$\hat{\mathbf{r}}$, $\hat{\mathbf{f}}$] in equation (1), will be optimal. In other words, the outcome of the heuristic procedure corresponds to a (known) minimum that is local respect to the (unknown) global one.

In the first phase (Stage 1) two different procedures are carried out to identify areas and nodes to be served. A basin definition procedure is applied for each mass transit station to be fed, chosen according to analyses on the "feeder" transit demand (part of the transit demand exceeding maximum allowable walking distance) the urban framework and the resulting walking accessibility; then traffic zones are associated to the closest major stations in terms of travel distances. A zone aggregation procedure is applied according to the "feeder" transit demand level for each traffic zone and is carried out to identify traffic zones to be served, defining the main skeleton of the road network for the application of the HRGA in the following stage.

In the second phase of the solution procedure (Stage 2), a heuristic algorithm generates two different and complementary sets of rational and realistic routes (K-shortest path and TSP

type routes). This provides a large set of feasible routes that are nevertheless quite diversified among them, because they are built according to typical feeder bus network design criteria. Therefore balance between effectiveness and efficiency (user or operator point of view) and maximization of service coverage in the area and improvement of integration between rail and bus networks represents the remarks which led us in choosing main elements of HRGA. The "K-shortest path" type routes are composed by direct routes connecting main system stations (transfer points) with any centroids laying in our service area. The TSP-type routes come from the application of "Travelling Salesman Problem algorithm" and connect all centroids to railway or subway stations with a single path providing a widespread demand collection aiming at maximizing service coverage.

The resulting set of feasible routes is the basin from which the GA select routes to build a network (Stage 3). The design variables are transit routes and the GA is implemented in the Matlab language while the fitness evaluation requires computing, for each solution generated, the three terms of the objective function by simulating the public transport network with the EMME software (EMME User's manual, 2008).

As the performance of the transit system depends on the service frequencies, which should be optimized depending on the passenger volumes, an iterative assignment and frequency setting procedure, first introduced by Baaj and Mahmassani (1990), is applied.

## 4.1 Basin definition and zone aggregation procedures

The first stage of the solution framework is the application of two different procedures in order to define areas and traffic zones to be served. Such procedures allow to identify: mass transit stations to be fed (hereafter called "major" stations), an influence area for each major station (hereafter called "basin") and which traffic zones are associated to each basin.

These results are provided according to several considerations and evaluations. Firstly for the definition of major stations, the "feeder" transit demand has to be computed; "feeder" demand is the part of the whole transit demand exceeding a maximum allowable constraint in terms of walking distance run viz all the passengers not properly served by mass transit system in the access/egress phase. In our application to the actual-size network of Rome, this allowable walking constraint has been chosen equal to 500 metres due to the large size of the study area. An "All or nothing" assignment allows to define main pedestrian paths of "feeder" transit demand and then major stations to be fed are identified. Specifically, for each district of the city only a part of all mass transit stations are chosen, on the base of analyses about the size of these pedestrian paths flows and where most of them meets in.

Once defined major stations to be fed, it is possible to identify an influence area for each one, hereafter called "basin", by applying a basin definition procedure. This procedure is carried out taking into account urban framework, road infrastructure framework and resulting walking accessibility for all the traffic zones in the study areas. Specifically in this stage, all the traffic zones are associated to the closest major station in terms of walking distances so that each mass transit station can be considered fed by only a part of all traffic zones of the

study area. This basin definition procedure allows an easier application of the HRGA since defines a priori which traffic zones have to be connected with each major station through the set of routes built according the two different design criteria.

Finally, in order to reduce computational times and the route design complexity for both the algorithms used in following HRGA stage, a zone aggregation procedure is carried out for each basin. Sure enough that a lower number of traffic zones to be served allows to reduce the deviousness of preliminary TSP and K-shortest paths and then the overall computational times. These procedure is carried out on the base of transit demand level of the traffic zones laying in each basin and is divided in 4 steps:

1. for any basin, all traffic zones are listed according to the demand level;
2. for the first zone, an influence area is considered;
3. covered surrounding zones are aggregated and removed from the list;
4. the procedure is repeated until the list is empty

Worth noting that two different zones are aggregated with when they are less than 300 metres far.

Once defined the number and the list of major stations to be fed, their basins, which traffic zones are associated with each basin and which traffic zones are aggregated together, the stage 1 can be considered completed and the procedure ready to apply the algorithms for designing preliminary sets of feasible routes.

## 4.2 Heuristic route generation algorithm

The second component of the solution framework is the HRGA. Two complementary sets of candidate routes are generated by HRGA applying different design criteria and practical rules.
The HRGA is divided into the following five sequential steps:

- Step 1: Generation of the K-shortest type routes;
- Step 2: Generation of the TSP-type routes;
- Step 3: Storage of the K-shortest and TSP-type routes in the overall set of routes;
- Step 4: Check of the feasibility constraints (maximum and minimum allowable route length) for all routes stored in the basin;
- Step 5: Set feasible routes as input data for the GA.

The first set of feasible routes is called "K-shortest type". The route generation criterion reflects the users' point of view and is addressed to develop direct lines seeking routes effectiveness; it is composed by "k" the shortest paths connecting the main transit network stations with any single centroid laying in the service area. The order "k" of the paths represents a positive integer chosen in order to increase line influence in covering more zones of study area. Solving procedure is implemented in Matlab and is based on Dijkstra algorithm, because of its capability of finding paths and trees of lowest cost with high

computational speed in each kind of graph. Therefore two of the inputs required are source node and destination node and specifically, in our application, first one was a centroid in the area while second one was the railway station to be fed. Application of Dijkstra algorithm needs another important input as the total travel distance matrix among all the regular nodes of road network. This matrix is built having the "adjacency matrix" of the graph; specifically this matrix contains in each cell "ij" not the usual binary code but a value equal to the length of the edge connecting nodes "i" and "j" or, otherwise, a code "Inf" if nodes "i" and "j" are not connected by a direct edge.

Last but not less important input for K-shortest-type routes is the parameter "k"; its value has to be chosen in order to design a wide range of "k" alternative and feasible routes between each centroid and the railway station to be fed. Therefore, in our application, this choice was made taking into account location of traffic zones and railway stations in our service area and road network features.

The second set of feasible routes is based on a completely different approach, which aims at reducing passengers' access impedance allowing a widespread demand collection in service area. The design strategy, based on Travelling Salesman Problem Algorithm, seeks to provide service coverage in the area since also circular but rational routes are accepted even if they might increase total travel costs for both users and operators.

From the user's point of view, the obvious increase of in-vehicle time with respect to the K shortest-type routes should be balanced by some positive expected aspects, as lower auxiliary access time to the network. From the operator's point of view, the network composed by this family of routes seems less efficient than that composed by all K shortest-type routes due to the deviousness of the lines but it allows a service coverage improvement in the service area, typical goal of feeder bus network design problem.

As it is well known "TSP" is a minimization problem of connecting all vertex of a graph with the lowest cost path, starting and finishing at a specified vertex after having visited each other vertex exactly once. In this work this family of routes derives from a TSP algorithm based on a Genetic Algorithm (GA).

For this reason two inputs required for generation of TSP routes family with this solving procedure are number of iterations of GA and number of individuals of the population to be evaluated. Two other parameters have to be taken into account to apply this particular TSP algorithm. First one is the total distance matrix between all centroids of service area; each cell "ij" of this matrix shows the length of the shortest path between centroid "i" and centroid "j". Coordinates "X" and "Y" of any centroid are, instead, last input needed to implement this TSP-algorithm; inflecting number of GA iterations and number of individuals of GA population is possible to obtain a wide range of alternative and feasible TSP routes, starting and finishing at the railway station to be fed after having visited each centroid of the service area exactly once. As we've already described, generation of this kind of routes follows the main goal of providing the service coverage improvement in the area, the demand collection

maximization and, from users' point of view, the minimization of auxiliary access time to the transit network.

At the end of K-shortest type and TSP-type routes generation, all routes generated are checked to verify if feasibility route constraints are satisfied. In particular, the constraints concerning the maximum and minimum allowable route lengths are investigated. If the constraints are satisfied the routes are stored in the set of feasible routes as input data for the last phase of the solution framework.

## 4.3 Optimal bus network calculation

The third stage of the solution framework is characterized by the use of the genetic algorithm (GA) to find the optimal sub-set of routes and their frequencies.

As we showed shortly in the paragraph before, Genetic Algorithms are stochastic optimization algorithms founded on the applications of concepts of natural selection and natural genetics (Goldberg, 1989) and have been used in recent years to solve many optimization problems. The transition scheme of GA simulates the natural evolution of a population and investigates the solution space by applying a probabilistic search process to all the individuals representing a population of solutions simultaneously. In general, the "best" individuals of any population of solutions tend to reproduce themselves and survive to the next generation, thus improving successive generations. GAs explore all regions of the solution space by applying genetic operators (mutation, crossover, selection and elitism) that simulate the reproduction of individuals in the population. Crossover takes a pair of individuals and generates two new individuals (their offspring) by combining their chromosomes' sets. Mutation provides a probabilistic modification of the chromosomes that may alter the reproduction process. Elitism is used to save the few best individuals that tend to reproduce and survive to the next iteration, thus improving successive generations.

In our application (Section 5), resulting from the HRGA procedure are about 80 "K-shortest", and "TSP" type routes. The set containing these routes is the basin from which the GA picks to build up a network; for instead, in case of a population composed by 50 individuals, any composed by 10 chromosomes, any generation presents 50 networks, any composed by 10 lines, randomly picked from the basin containing the 81 lines. Any route is identified by a code (line number); any network is represented as a string (in this example, 10 characters long). For any individual (network) the objective function value is computed. Then, a linear fitness scaling is performed in order to convert the raw objective function scores to values ranging in an interval suitable for the roulette wheel selection that has been implemented in the present algorithm. Once 2 parents have been selected, the crossover operator randomly selects half of the chromosomes from the first parent and half from the second one, and combines the selected chromosomes to generate the offspring. Mutation operator replaces, with a mutation probability equal to 1.5%, each chromosome of the individual with a new chromosome (line) chosen randomly from the basin. Reproduction options that have been utilized to create the next generation are: elitism fraction equal to 10%; crossover fraction, other than elite fraction, equal to 85%; mutation fraction equal to 15%.

The GA has been implemented in the Matlab language too; the Emme scripting language is used to perform transit assignments required for the evaluation of the objective function. The fitness evaluation requires computing, for each solution generated, the three components of the objective function by simulating the public transport network. As the performance of the transit system depends on the service frequencies, which should be optimized depending on the passenger volumes, an iterative assignment and frequency setting procedure, first introduced by Baaj and Mahmassani (1990), is applied. The procedure consists of an iterative process between the transit demand assignment and the route frequency setting equation:

$$f_i = \frac{q_{hk,i,\max}}{fc_{\max} \cdot C_V}$$

(10)

where *qhk,i,max* is the maximum segment volume of line i as resulting from the assignment.

The procedure stops when the maximum difference among route frequencies in two consecutive iterations is lower than a given threshold. The convergence of the iterative frequency setting procedure is not guaranteed, but all computational tests performed converged in few iterations.

# 5. REAL SIZE NETWORK APPLICATION

The proposed procedure has been implemented on a medium sized real-life network in order to compare its effectiveness versus the performance of the existing transit network. As explained in the introduction of this paper, before the application of our model on the Rome transit network, the methodology has been tested on the Winnipeg transit network in order to assess OF weights and the feasibility of final results.

## 5.1 Winnipeg test network

These preliminary tests have been carried out on the Winnipeg network because its size is very close to the size of suburban study area of Rome network chosen for the application of the model. This experiment has allowed to evaluate:

1. the sensibility of the solution search process to the weights adopted in the objective function;
2. the best network configuration varying the number of bus lines (from 10 to 20).

In the HRGA stage a set of 80 feasible routes has been identified; this is the basin for the application of the second step of the procedure: finding the optimal solution among a population 30 of individuals (networks), each one having 20 chromosomes (lines).
Each term of the objective function has been computed according to different combinations of weights in order to: (1) make the objective function terms homogeneous; (2) reflect the

trade-offs between different subjects involved (users and operators). Calibration has taken into account mainly the comparison between access time weight and transfers weight; these have been considered the best parameters to represent the typical FBNDP goals since auxiliary access time weight stands for maximization of service coverage aim while transfer weight represents the objective of improving integration between bus network and rail network. Three different tests have been carried out with three different ratios between access time weight and boardings weight: $\alpha/\beta = 4$ (Test A), $\alpha/\beta = 1/4$ (Test B) and $\alpha/\beta = 1/10$ (Test C), where:

- $\alpha$ =boardings weight;
- $\beta$ =auxiliary access/egress time weight

Such weights have been calibrated applying a detailed and exhaustive sensitivity analysis that has allowed to underline some important aspects:

- high sensitivity of identified solutions with respect to the changes of access time weights (value adopted equal to 10);
- low sensitivity of identified solutions when using high values for the transfers' penalty weight (value adopted equal to 1);
- low variance of the waiting and in-vehicle travel times with respect to the change of access time weight and boarding weight (values adopted equal to 0.04 and 0.02, respectively).

Specifically, the comparison among tests has shown that a global decrease of auxiliary access time can be detected with test "C"; any test is characterized by an uniform value of unsatisfied demand, that is the sum of all passengers not served by the transit network within 75 minutes. At the same time very low increases of waiting time, in-vehicle time, and operative costs (veh-h and veh-km) show that it is possible to achieve the maximization of demand collection with an optimization of current available resources, as it's illustrated in "Table I":

Table I: Objective function terms for the different test networks

| TEST | OF value | Veh-h (n) | Veh-km (n) | Transfers (n) | In-veh time (min) | Access time (min) | Waiting time (min) | Uns. demand (n) |
|---|---|---|---|---|---|---|---|---|
| A | 120,517 | 625 | 9,922 | 4,984 | 998,671 | 318,356 | 112,226 | 3,458 |
| B | 214,290 | 625 | 9,928 | 5,072 | 999,521 | 317,578 | 112,490 | 3,459 |
| C | 404,749 | 628 | 9,976 | 5,019 | 999,880 | 317,224 | 113,403 | 3,461 |

Then, three additional tests have been carried out on Winnipeg network in order to assess a suitable number of lines composing the sub-optimal feeder network. Specifically 10, 15 and 20 lines networks have been compared as it's shown in "Table II" with corresponding objective function terms:

Table II: Objective function terms for the different networks

| N° of lines | OF value | Veh-h (n) | Veh-km (n) | Transfers (n) | In-veh time (min) | Access time (min) | Waiting time (min) | Uns. demand (n) |
|---|---|---|---|---|---|---|---|---|
| 10 | 404,781 | 619 | 9,828 | 5,060 | 998,671 | 317,968 | 113,383 | 3,464 |
| 15 | 404,484 | 625 | 9,920 | 5,099 | 1,000,430 | 317,201 | 113,244 | 3,462 |
| 20 | 404,749 | 628 | 9,976 | 5,019 | 999,880 | 317,224 | 113,403 | 3,461 |

On the base of such results, a feeder network made of 10 lines has been considered as the suitable for a study area of medium size like Winnipeg and Rome one; 15-lines network shows higher values of transfers, in-vehicle time and boardings while 20-lines network presents too low values of average and maximum passenger volumes (halved than 10-lines). Small increase of access time showed by 10-lines feeder network has been considered acceptable taking into account the lower employ of resources needed to provide service in the area.

Since methodology tests on Winnipeg network gave the expected results, model has been considered calibrated and ready to be applied on Rome real-life network.

## 5.2 Rome real-life network application

A FBND methodology has been applied to twonorth-western suburbs of the city Aurelio and Primavalle, located in the 18th and 19th districts,represent the study area, about 20 km2wideand with a population of about 125.000 inhabitants.

Main transit network is composed of two lines: one subway line, Metro A, and one urban railway service, FR-3 Roma-Cesano; transfer points chosen to be fed lay just on this two lines: Battistini, end of the Metro A line, and Gemelli, stop of FR-3 urban railway. The existing bus transit network is, instead, composed of 15 bus lines, with many overlapping routes and not very high frequencies (average lines headway is about 15 minutes). In order to offer a regular service in the study area, respecting planned headways on each line route, in the morning peak hour operator needs about 120 vehicles. Despite of this huge number of required buses, average and maximum volumes don't exceed 35% and 76% respectively. Transit demand in the morning peak hour amounts to about 13000 trips with almost 8000 users (the 60%) exiting from the service area.

Once defined its geographical and network features, the study area has been divided in two basins, one for each "major" station; therefore each "major" station has been considered fed by only a part of all traffic zones, whose demand has been considered attracted only by the basin reference station. Both the partition and aggregation have been carried out according to the basin definition and the zone aggregation procedure, already described in the paragraph 4.1. The HRGA has allowed the identification of a set of feasible routes composed by 81 elements (66 K-shortest type routes, and 15 TSP-type routes), showed in "Figure 1", which represents the basin for the application of GA for finding optimal solution.
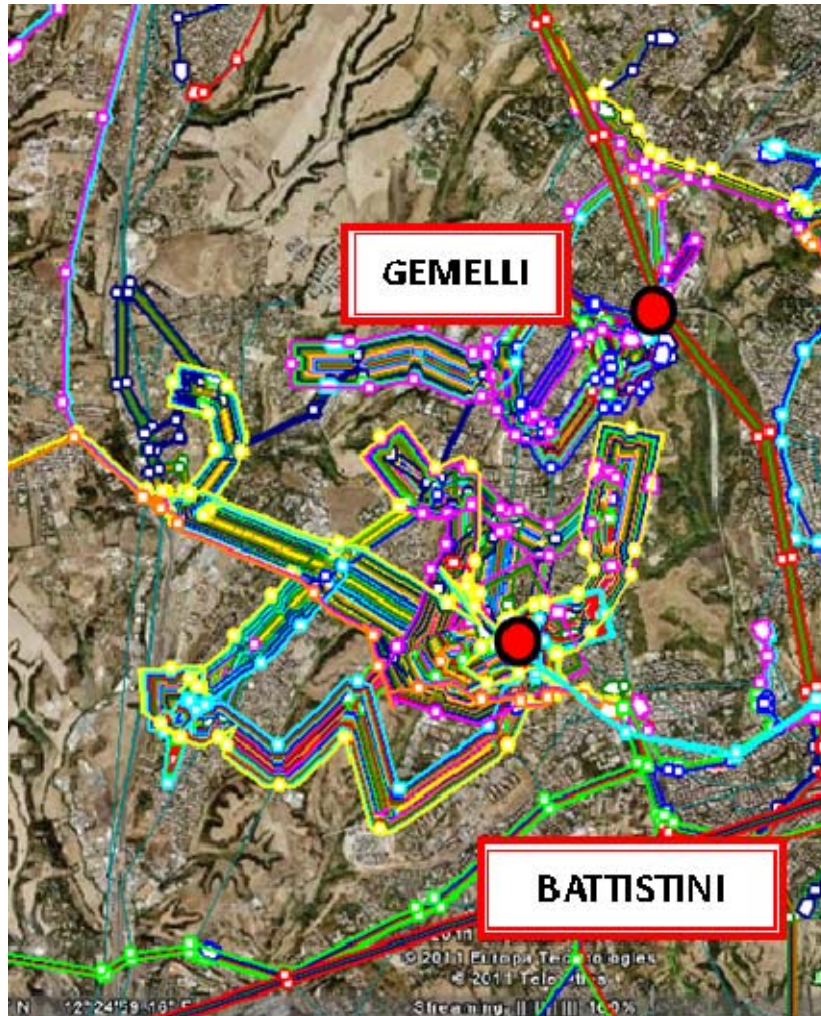
Figure 1: Basin of 81 elements with location of railways station to be fed

According to preliminary results, previously shown, individuals composed by 10 chromosomes (i.e. 10 lines bus network) have been adopted. Differently, a varying size of the population has been considered creating five scenarios, each composed by a different number of individuals (10, 20, 30, 40 and 50).

GA was run on a PC Pentium 4 with a 1.86 Ghz processor and 1 GB of RAM for a computation time of about 15 hours for 100 iterations. Such computation time depends on the number of iterations needed for the GA to obtain a significant reduction of the objective function.

Tables III and Table IV summarize the results obtained by applying the GA. Table III shows the values of the different terms of the objective function; this has been evaluated by assigning all transit demand on entire network and taking into account only the results of our study area; the rest of entire network has been considered invariant in time and space.
Table IV, instead, provides the comparison among these solutions in terms of percentage differences calculated with respect to the existing network of 15 bus lines in the study area.

Table III: Performance comparison between existing and designed networks of 10 lines (#)

| Scenario | Individuals. (n) | OF value (n) | Transfers (n) | In-veh time (min) | Access time (min) | Waiting time (min) | Unsatisfied demand (n) |
|---|---|---|---|---|---|---|---|
| Existing | - | 183,545 | 16,520 | 305,591 | 166,530 | 73,172 | 1,582 |
| 1 | 30 | 140,691 | 24,279 | 280,043 | 125,920 | 73,691 | 948 |
| 2 | 40 | 139,670 | 24,649 | 284,456 | 124,502 | 75,479 | 1,000 |
| 3 | 50 | 140,245 | 24,474 | 284,133 | 125,237 | 73,188 | 987 |
| 4 | 60 | 139,277 | 24,742 | 281,567 | 124,206 | 76,733 | 974 |
| 5 | 70 | 139,121 | 24,331 | 286,404 | 124,171 | 73,971 | 958 |

Table IV: Performance comparison between existing and designed networks of 10 lines (%)

| Scenario | Individuals. (n) | Δ(%) OF value | Δ (%) Transfers | Δ (%) In-veh time | Δ (%) Access time | Δ (%) Waiting time | Δ (%) Unsatisfied demand |
|---|---|---|---|---|---|---|---|
| 1 | 30 | -23 | 47 | -8 | -24 | 1 | -40 |
| 2 | 40 | -24 | 49 | -7 | -25 | 3 | -37 |
| 3 | 50 | -24 | 48 | -7 | -25 | 0 | -38 |
| 4 | 60 | -24 | 50 | -8 | -25 | 5 | -38 |
| 5 | 70 | -24 | 47 | -6 | -25 | 1 | -39 |

As it can be seen above, all scenarios gave the results we expected; Table 3 shows that changes of the components vary in a very limited range among all solutions but have a deep gap with respect to existing scenario; Table 4 illustrates that we have a global decrease of OF value, of unsatisfied demand and access time; at the same time boarding increase represents, instead, the improvement of integration between rail and bus network, third initial aim of this application.

Scenario 3, composed by 6 K-shortest type routes and 4 TSP-type and obtained with a GA population of 50 individuals, has been considered as the best network due to its balance between transfers increase and the decrease of all others OF components.

Actually the comparison between the existing network and the proposed one shows that larger amount of transit demand in the study area can be served more effectively (reduction of 25% of the access time and decrease of 38% of unsatisfied demand) by a bus fleet composed by a lower number of lines (reduction of 33%) in a more efficient way (reduction of 7% of in-vehicle time), while still guaranteeing the same waiting time as the current network. At the same time, boarding increase represents, instead, a natural consequence of a route design procedure based on the initial goal of improving integration between rail and bus network (one transfer per passenger is due to the boarding from bus to railway).

Table V shows a summary of best design network; as it can be seen average headway, number of vehicles needed, average line length and time also decreased:

Table V: Best network features compared with existing network ones:

|  | *Number of vehicles (n)* | *Average Headway (min)* | *Average length (km)* | *Average running time (min)* |
|---|---|---|---|---|
| Existing (A) | 119 | 15,6 | 18 | 85 |
| Design (B) | 60 | 11,4 | 9 | 36 |

Detailed comparison between the existing and the "10 lines" designed network in terms of line frequency is reported in Table VI. It is possible to see, as expected, that feeder network is mostly composed by high frequency lines (headway equal to 4 min or less), differently from the existing network.

Table 6: Classification of existing and design network in terms of headway:

| *Network* | *Number of lines (n)* | *Headway ≤ 4 min (% lines)* | *5 min < Headway ≤ 15 min (% lines)* | *15 min < Headway ≤ 30 min (% lines)* |
|---|---|---|---|---|
| Existing | 15 | 0 (0%) | 8 (53%) | 7 (47%) |
| Scenario | 10 | 3 (30%) | 3 (30%) | 4 (40%) |

At the end, in order to have a more homogeneous comparison between our optimal solution (Scenario 3) and the current network, we optimized the frequencies of the current network since, as exposed in chapter 4, our solving procedure optimizes all the frequencies of each line on the base of transit volumes. The comparison shows an increase of "current" operative costs (veh-h and veh-km), both equal to 3%, due to the maximization of service coverage and the increase of associated lines frequencies.

The analysis of auxiliary volume distributions shows that the design network has allowed an improvement of service coverage in the study area. Analogously, as it's shown in Figure 2, the passengers volume distribution highlights an increase of almost 3000 passengers on the FR-3 and of 800 passengers on the first links of Metro A. This consideration is confirmed by analyzing the distribution of transit volumes on the entire railway network, showing an increase of users on entire route of FR-3 between our optimal scenario and current network.

Moreover the analysis on the entire transit network shows a global decrease of total travel times and distances on the bus network (-2% in pax-h and pax-km) while, instead, a global increase of the same parameters is registered on the urban-rail network (+3% in pax-h and pax-km). These results represent the achievement of improving integration between bus and rail modes since passengers use mainly the rail network to complete their trips reaching their destinations, after an initial stretch of route on optimal feeder network given by the model.
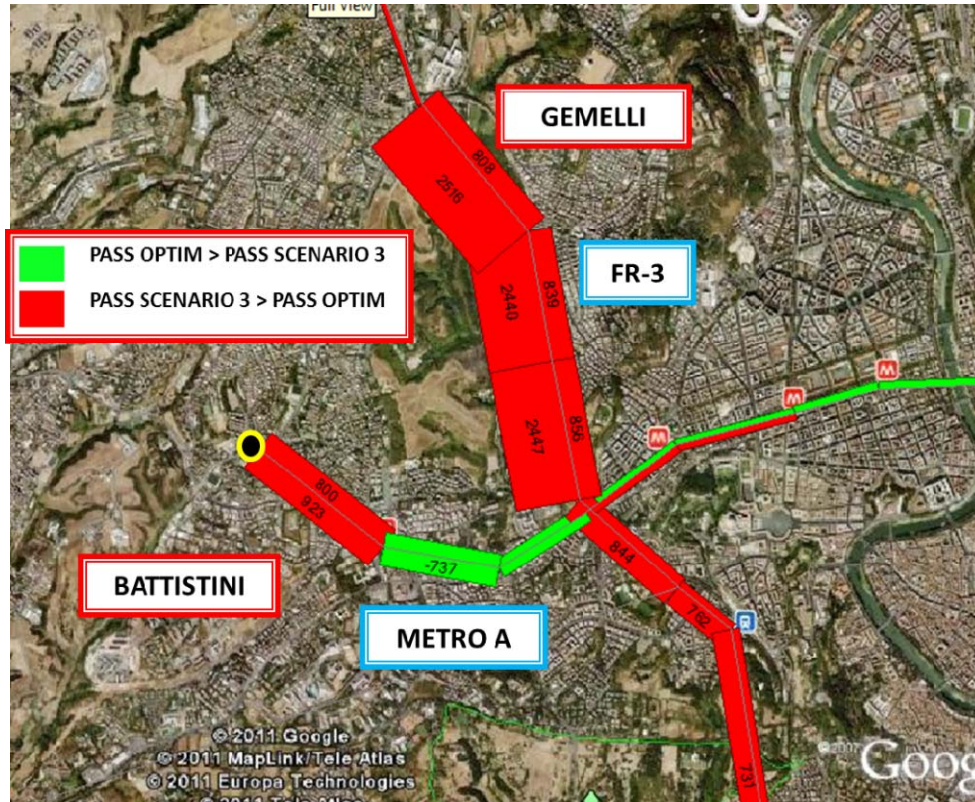
Figure 2: Comparison of transit volumes on rail network between current network with optimized frequencies and optimal Scenario 3

# 6. CONCLUSIONS AND FURTHER DEVELOPMENTS

In this paper, authors propose a procedure for solving the feeder bus network design problem in a urban area characterized by a multimodal transit system . The solving procedure consists of a set of heuristics, which includes a first routine for route generation based on two different algorithms (TSP and K-shortest path), in order to achieve all the typical goals of feeder bus network like improving integration between rail and bus transit systems; in the second phase a genetic algorithm is implemented for finding an optimal or near-optimal network of routes with the associated frequencies. Main novelties introduced by this paper are the adoption of a new heuristic procedure for route generation process and the application of the transit network design methodology, suited for feeder bus networks, to a real-life urban area (the city of Rome).

The application of various generation criteria in the HRGA has led to a consistent, diversified and exhaustive set of feasible routes. The implemented genetic algorithm has proved to be robust and effective in producing reasonable solutions. Numerical experiments carried out on network of the city of Rome highlighted that study area can be served witha more extensive demand collection in respect of our initial goals of maximization of service coverage(reduction of 25% of the access time and decrease of 38% of unsatisfied demand) by a bus fleet composed by a lower number of lines (reduction of 33%) in a more efficient way (reduction of 7% of in-vehicle time), while still guaranteeing the same waiting time as the current system.

Further developments will be focused on supplementary analysis for the definition of feasible routes generated only on the main skeleton of road network in order to reduce its complexity. Additional effort has to be spent in the specification of the objective function in terms of components and weights to improve the effectiveness of the procedure by the transportation point of view. Additional refinements and improvements of GA efficiency or the use of other metaheuristics techniques (like PSO) are necessary for the reduction of computational times.

# REFERENCES

Baaj M.H. and Mahmassani H.S. (1990). "TRUST: A Lisp program for the analysis of transit route configurations", Transportation Research Record, 1283, pp. 125-135.

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

Baaj M. H. and Mahmassani H.S. (1995). "Hybrid route generation heuristic algorithm for the design of transit networks", Transportation Research, Vol.3C, pp. 31-50.

Carrese, S. and Gori, S. (2002). "An urban bus network design procedure". Applied Optimization 64, pp. 177-196.

Chien S. and Yang Z.. (2000). "Optimal feeder bus routes on irregular street network", Journal of Advanced Transportation, Vol.34 n.2, pp. 213-248.

Cipriani E., Gori S., Petrelli M. (2009). "Transit Network Design: a procedure and an application to a large urban area". "Transportation Research – Part C", Elsevier;

Desaulniers, G., Hickman, M., (2007). Public transit. Handbooks in Operation Research and Management Science, pp. 69-120.

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

Jerby S. and Ceder A. (2006). "Optimal routing design for shuttle bus service", Transportation Research Record, 1971, pp. 14-22.

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

Newell G.F. (1979). "Some issues relating to the optimal design of bus routes", Transportation Sci., 13(1), pp. 20-35.

Shrivastava P. and O'Mahony M. (2006). "A model for development of optimized feeder,routes and coordinated – A genetic algorithm approach", Transport policy 13, pp. 413-425.

Spiess H. and Florian M. (1989). "Optimal strategies: A new assignment model for transit networks", Transportation Research, vol. 23B,  pp.83-102.