Computers, Environment and Urban Systems 36 (2012) 445–455

Contents lists available at SciVerse ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/compenvurbsys

# Identifying bus stop redundancy: A gis-based spatial optimization approach

Eric M. Delmelle [a,*], Shuping Li [b], Alan T. Murray [c]

[a] Department of Geography and Earth Sciences, University of North Carolina at Charlotte, Charlotte, NC 28223, USA
[b] Environmental Research System Institute (ESRI), Redlands, CA, USA
[c] GeoDa Center for Geospatial Analysis and Computation, School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ 85287-5302, USA

ARTICLE INFO

*Article history:*
Received 6 July 2010
Received in revised form 21 January 2012
Accepted 22 January 2012
Available online 25 February 2012

*Keywords:*
GIS
Interaction
Location modeling
Public transit
Simulated annealing

ABSTRACT

Transit planners are often faced with a tradeoff between improving accessibility through the addition of stops while simultaneously increasing efficiency so that destinations can be reached in a reasonable amount of time. In this paper, we propose the development of an optimization framework integrated within a Geographical Information System (GIS) for addressing this specific problem. Our proposed modeling framework departs from well-known facility location coverage models by considering both the impact of walking distance from an individual residential location to a stop and the transit facility attractiveness (ease to cross, number of destinations served). Integration within a GIS environment is accomplished using a simulated annealing heuristic. An example on an inbound urban bus route illustrates the utility of the approach for transit planning, using model parameters developed in collaboration with local transit agencies.

Published by Elsevier Ltd.

## 1. Introduction

Promoting alternative modes of travel is a struggle in today's highly competitive but strongly automobile-dominated urban travel markets (Taaffe, Gauthier, & O'Kelly, 1996). Problems associated with automobiles, such as congestion and pollution, have exerted considerable pressure on transportation infrastructure, especially in urban regions. Socially, people should have access to public transportation, keeping in mind that public transportation needs to best serve the population (Bullard, 2003). In recent years there has been renewed interest in public transit and it continues to be viewed as an important component of the overall transportation planning and management of urban regions (Murray, 2003). The most distinguished advantages of public transit derive from its unique ability to thrive on high volumes of travel demand concentrated in space and time (Pucher, 2004). Another positive characteristic of transit is its potential for serving passenger trips at lower energy consumption cost and less pollutant emissions per passenger mile than a private automobile (Garrison & Levinson, 2006; Giuliano, 2004).

Despite the attractive and sustainable attributes of public transportation, ridership continues to decline as a proportion of overall trips. In the USA, approximately 64.37% of all commuting trips were made by private automobile in 1980 (19.73% carpool, 6.22% transit), and increased to around 75.20% in 2000 (13.36% carpool, 4.58% transit) with transit utilization at about 5% (Taaffe et al., 1996; TRB, 2006). Following the annual American Community Survey (US Census Bureau, 2009), car utilization further increased to 76.1% in 2009 (with a drop to 10% in carpool and 5% in public transportation). The situation is not as severe in large, dense cities with a well connected transit system. Newman and Kenworthy (1999) point out that transit use in large cities in the USA was 9.0% in 1990 in comparison to 38.8% in European cities. Buehler (2010) highlights that European transport policies are usually integrated with stringent land use development measures, keeping settlements compact. Transit policies and high gasoline prices have encouraged the use of alternate modes of transportation in Europe, as public transit is four to six times higher (Pucher & Lefevre, 1996).[1] Loose land-use development policies have resulted in a separation of employment and housing in the USA, known as spatial mismatch (Blumenberg & Manville, 2004; Preston & McLafferty, 1999).

One way to address spatial mismatch is by expanding the coverage of an existing system (adding more stops and lines), which may also increase transit utilization since access to the system is improved. Adding transit stops along a route will increase coverage of potential riders, but slows down travel speeds, reducing the number of reachable destinations within a travel budget (Murray & Wu, 2003). Greater ridership may actually be achieved when the system is made more efficient, given that faster travel speeds are preferred (Blumenberg & Manville, 2004; Newman & Kenworthy, 1999) and uncertainty due to traffic congestion and multiple stops along the way is at a minimum (Murray, Davis, Stimson, &

* Corresponding author. Tel: +1 704 687 5991.
*E-mail address:* Eric.Delmelle@uncc.edu (E.M. Delmelle).

[1] Americans are more generally sensitive to switching modes of transportation when gasoline prices are increasing (Buehler, 2010; Currie & Phung, 2008).

0198-9715/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.compenvurbsys.2012.01.002

Ferreira, 1998). Transit efficiency reflects how quickly a system can bring people from their transit stop to their destination. Accessibility of the public transit network is getting individuals from their system entry point to their system exit location in a reasonable amount of time (Murray & Wu, 2003; Murray et al., 1998). One strategy to this conundrum (decreasing travel time and augmenting public transit ridership) is to remove redundant stops, while retaining those stops offering higher levels of service, or make a route less circuitous (Blumenberg & Manville, 2004). Strategic transit planning models should be flexible enough to support transit attractiveness while simultaneously reducing redundancy.

There is no clear consensus on what constitutes a well-located transit stop, however, the FTA (1996) provides general planning guidelines that have been echoed by the literature. The FTA (1996) distinguishes between macro and micro-level criteria in determining stop placement. For example, macro-level considerations may place an emphasis on particular segments of the population who are completely reliant upon transit access such as lower-income groups (Wells & Thill, in press), or those with mobility constraints such as the elderly or disabled (Casas, 2007; Church & Marston, 2003). These considerations deal with the overall equitable distribution of public transit. Micro-level or so-called street side factors include, but are not necessarily limited to, the spacing between stops, distance to the nearest intersections, presence of sidewalks, adjacent land use, pedestrian access, and safety concerns (i.e. crosswalks, street lights, or street parking). Such street side factors are not detached from overall equity concerns as stops are often designed and placed with the needs of the mobility constrained population in mind; walking distances between a stop and other transit lines should be minimized and the placement of mid-block stops is discouraged as it forces pedestrians to walk greater distances to the nearest intersection, thereby placing an extra burden on deprived riders (Hess, in press). Furthermore, stops located far from an intersection will encourage jaywalking, increasing the risks and severity of pedestrian/vehicle crashes (FTA, 1996; Hess, in press; Pulugartha & Vanapalli, 2008; Truong & Somenhalli, 2011). Finally, a stop can also be strategically located when it is served by multiple routes (O'Sullivan & Morrall, 1996); it is more likely to be used if it serves multiple lines and multiple destinations. This paper is concerned with these latter, micro-level placement objectives, as stops along a single route are evaluated in addressing the redundancy/efficiency tradeoff. Emphasis is placed on preserving the most desirable stops according to street level objectives specified by the local transit agency.

Specifically, in this paper, we integrate the notion of population ridership (demand), physical accessibility (nearness of a crossing) and connectivity (number of destinations from a stop). We use a Spatial Interaction Coverage (SIC) model which explicitly maximizes covered demand weighted by the physical attraction of the transit facilities demand. The attraction of a facility is either reflected by its physical access, the number of destinations it serves or both. This paper contributes to the literature as follows: (1) it proposes a framework to address transit accessibility and efficiency by explicitly modeling facility attraction, distance decay and the importance of destinations served on a single route; (2) its parameters are estimated in consultation with transportation planners; and, (3) it integrates a non-linear model through simulated annealing, developed within a geographical information system (GIS) framework.

Section 2 reviews the importance of optimization methods when applied to transit planning. Section 3 addresses the spatial interaction coverage model, which explicitly accounts for facility attraction and distance decay. To solve this non-linear model, a simulated annealing heuristic is implemented in GIS. This integrated tool facilitates running different transit scenarios. Section 4 describes the transit planning context for the City of Charlotte (CATS), followed by an application to an inbound transit route. Concluding remarks and future developments are discussed in Section 6.

## 2. Background

In transit planning, two important system factors are transit access and system efficiency. For access, distance has a significant attenuating effect on spatial interaction: when the separation from an individual to a transit point of access decreases, the likelihood of using that public transportation service increases (Griffith & Jones, 1980). If the distances or barriers to access a public transportation service are too great at either the trip origin or destination, however, then it is unlikely to be utilized as a mode of travel. As stressed by O'Sullivan and Morrall (1996), knowing how far transit riders are willing to walk for service has serious implications for planners and developers in determining a catchment area for each station. The access coverage distance is affected by various factors, such as the walking environment to the transit point of access (Agrawal, Schlossberg, & Irvin, 2008), the structure of the age group in the originating demand areas (Neilson & Fowler, 1972), and the reliability of the public transit service. In recent research, Daniels and Mulley (2011) state that modes of transportation (train or bus) have a stronger impact on walking distance to public transit. For instance, research in Sydney points to a 573 m average distance to public transit, but this distance almost doubles for transit by train (805 m versus 461 m). A 5 min or 400 m walking distance is considered a reasonable access standard for bus stop transit in urban areas (Ammons, 2001; Demetsky & Lin, 1982; FTA, 1996; Levinson, 1983; Schobel, 2005). For system efficiency, fewer stops along a route will increase travel speeds and lengthen travel distance possible over a fixed period of time (FTA, 1996; Furth & Rahbee, 2000; Saka, 2001; Wirasinghe & Ghoneim, 1981). However, decreasing the number of stops will also decrease the access to transit facilities for customers (Foda & Osman, 2010; Murray, 2003). A trade-off exists between the number of transit facilities and service access, and striking a balance between these two components is a critical consideration in transit planning (Ibeas, dellOlio, Alonso, & Sainz, 2010; Murray & Wu, 2003).

The main difference among modeling approaches to public transportation stems from how effectiveness is defined and measured. A coverage modeling approach can maximize the number of individuals with suitable access to public transportation, while limiting the number of stops enhances system efficiency. Coverage is treated as a binary variable, making non-linear decline in demand with distance difficult to account for (Farhan & Murray, 2006). A distance based approach, such as the $p$-median, minimizes the weighted average travel time for individuals to the nearest access point, also keeping the number of stops to a predefined number (Hakimi, 1964). Although the $p$-median problem can account for distance decay indirectly, every individual is assigned to its closest transit access point even if it is not with a suitable distance.

Combining coverage and distance decay approaches is attractive, but there remains modeling hurdles. Coverage and distance decay approaches ignore that facility attraction influences ridership. Potential transit riders are likely to travel to the nearest facility when those facilities are equally attractive (Farhan & Murray, 2006). However, they are willing to walk longer distances for more attractive facilities (Daniels & Mulley, 2011). A large population concentrated around a transit stop does not automatically translate into ridership, especially when the destinations served by a transit facility are not useful for that population. Research has suggested that transit riders are willing to travel larger distances for faster or more reliable services, such as express routes (O'Sullivan & Morrall, 1996).

A number of well-known options exist for implementing models. It is possible to use commercial solvers, such as Lingo, Gurobi or CPLEX, but also through programming in a language such as C++ for large or non-linear problems. Increasingly, the benefits of linking GIS with optimization solvers are being realized. GIS has a long contributing history to location science (Murray, 2010). By means of network analysis, GIS facilitates the modeling of distance and delineating service areas (Gutierrez & Garca-Palomares, 2008). Also, GIS can help identify which segments of the population remain underserved once transportation infrastructure has been modified. Implementing the formulation of a location model in a GIS environment has been made easier by using supported programming languages (e.g. Python or Visual Basic for Application in ArcGIS). Two distinct avenues exist to capitalize on the flexibility of GIS for decision making in general, and location modeling in particular (Delmelle, Delmelle, Casas, & Barto, 2011; Ghosh, 2008; Malczewski, 1999). One option is a loose coupling between a regular solver and GIS. Recent examples of this include Tong and Murray (2009) associated with the MCLP, accounting for joint service when a demand node is served by multiple facilities, Murawski and Church (2009) who optimize an existing transportation network to improve health care accessibility, and Alexandris and Giannikos (2010) exploiting GIS capabilities to model complementary partial coverage. Such approaches have limitations, however. For instance, input and output files are exchanged manually and have different data formats. Loose coupling has the disadvantage of being less efficient, reducing performance, and there is a need for more than one software program to solve a problem. Alternatively, file exchange can be automated; in ArcGIS for example, external solvers or other programs can automatically be called through the use of ArcObjects (Delmelle et al., 2011). An instance of a graphical user interface (GUI) in the area of location modeling is LOLA by Hamacher, Hennes, Kalcsics, and Nickel (2003), using a flexible C++ programming platform and integrated in ArcView (Bender, Hennes, Kalcsics, Melo, & Nickel, 2004). Heuristics are often necessary when the model to be solved is non-linear. Integrating heuristics within a GIS can be accomplished through a tightly coupled system, such as the application proposed in this paper, or by means of a standalone system.

## 3. Modeling spatial interaction

Distance decay and coverage models each have shortcomings. There is a lack of flexibility in structuring distance decay and no explicit representation of facility attraction is found in coverage models. The Spatial Interaction Coverage (SIC) model addresses these issues by capitalizing on the interaction between a demand node and a facility. The demand can be split across several facilities within the coverage distance: a closer facility may not be as attractive as a facility further away with more amenities. The likelihood of an individual using a facility will increase with the attraction to that facility (Huff, 1963). In public transit, different characteristics reflect attraction, such as the number of transit lines and their frequencies, the number of non-stop destinations and also the quality of the pedestrian environment in the vicinity of a stop. While existing facility location models typically do not consider physical access of stops and importance of stops in terms of destinations served, the ability of the spatial interaction coverage to incorporate these attributes and better model distance decay makes it an appealing option (Alam, Thompson, & Brown, 2010).

### 3.1. SIC formulation

The formulation of the SIC utilizes the following notation:

| | |
|---|---|
| $i$ | index for demand nodes |
| $j$ | index for candidate facility locations |
| $I$ | set of demand nodes |
| $J$ | set of all candidate facility locations |
| $d_{ij}$ | shortest distance or travel time between demand node $i$ and candidate location $j$ |
| $R$ | service access distance standard |
| $w_j$ | attraction weight for candidate facility location $j$ |
| $a_i$ | demand at location $i$ |
| $\alpha$ | exponent controlling attraction weight |
| $\beta$ | exponent controlling distance $d_{ij}$ |
| $N_i$ | $\{j \mid d_{ij} < R\}$; set of candidate facility locations within a threshold distance $R$ of demand node $i$ |

**Decision variables**:

$$X_j = \begin{cases} 1 & \text{if existing candidate facility location } j \text{ is} \\ & \text{included in the system} \\ 0 & \text{otherwise} \end{cases}$$

$S_{ij}$ — interaction between demand node $i$ and candidate facility location $j$.

The spatial interaction coverage model is as follows:

$$\text{Maximize } Z = \sum_{i \in I} \sum_{j \in N_i} S_{ij} \tag{1}$$

$$\text{Subject to } S_{ij} = \left[ \frac{a_i w_j^{\alpha} d_{ij}^{-\beta}}{\sum_{k \in N_i} w_k^{\alpha} d_{ik}^{-\beta} X_k} \right] X_j \quad \forall i \in I,\ \forall j \in J \tag{2}$$

$$\sum_{j \in J} S_{ij} \leqslant a_i \quad \forall i \in I \tag{3}$$

$$\sum_{j \in J} X_j = p \tag{4}$$

$$S_{ij} \geqslant 0 \quad \forall i \in I,\ \forall j \in J \tag{5}$$

$$X_j \in 0,1 \quad \forall j \in J \tag{6}$$

The objective (1) of the spatial interaction coverage model maximizes the interaction between all demand nodes and facilities. The decision variable $S_{ij}$ summarizes the interaction between a demand node $i$ and facility $j$ as is a function of the magnitude of the demand at $i$, the separating distance $d_{ij}$, and the attractiveness of facility $j$, denoted $w_j$. This decision variable is zero when there is no attraction between a demand node and a facility (scenarios include no open facilities in the vicinity of demand node $i$, no demand at $i$, or the selected facility attraction $w_j$ is 0). The SIC contrasts with the $p$-median and MCLP model in that demand can partially be assigned to more than one facility, and facility attraction is explicitly integrated. Exponents $\alpha$ and $\beta$ control the importance given to facility attraction and distance deterrence. A large value for $\beta$ (e.g. $\beta = 5$) limits the interaction to nearby facilities, reflective of difficulty for a certain age group to reach a facility, while the distance constraint vanishes when $\beta = 0$. The denominator accounts for additional facilities that are open within the neighborhood of $j$, acting as competitors. The interaction between a demand node $i$ and candidate facility $j$ decreases as more facilities open in the vicinity of $j$. Note that the sum of the interactions originating from $i$ must be smaller or equal to the demand, (3). Constraint (4) specifies that $p$ facilities must be selected. When facility costs are constant, the constraint is equivalent to a budget constraint (ReVelle & Swain, 1970). The interaction between a demand node and a facility is required to be positive in constraint (5). Finally, constraint (6) imposes integer restrictions on the decision variables.

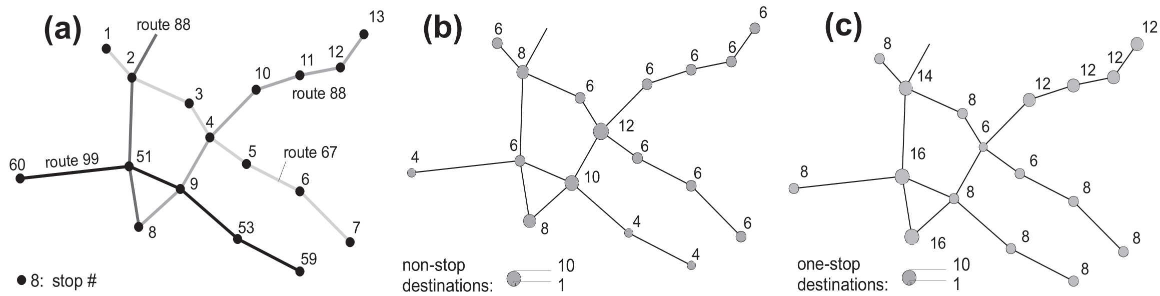


**Fig. 1.** Illustration of the the non-stop connectivity and 1 stop connectivity for a hypothetical network of 4 routes (route 67, 88, 88, and 99).

Various challenges exist in modeling bus stop attractiveness. One is a composite weight combining various facility characteristics: whether the stop is on an express route or if many other destinations can be reached without a transfer. In this paper, we quantify facility attraction by estimating how connected that facility is to the rest of the network (see Fig. 1). The following notation is introduced:

| | |
|---|---|
| $\mathfrak{L}$ | set of bus routes in the transit system |
| $e$ | index of bus stops (facility) on a route $E \in \mathfrak{L}, e \in J$ |
| $l$ | index of bus stops (facility) on a route $L \in \mathfrak{L}, l \in J, e \neq l, E \neq L$. |
| $k$ | number of connections allowed. |

$$C^k(l,e) = \begin{cases} 1 & \text{if stop } e \text{ can be reached in } k \\ & \text{connections from } j \\ 0 & \text{otherwise} \end{cases}$$

The weight associated with a candidate stop is scaled from 0 to 1:

$$w_l^k = \frac{\sum_{e \in E} C^k(l,e)}{\underbrace{max}_{j} \sum_{e \in E} C^k(l,e)} \tag{7}$$

when $k$ is zero, the weight $w_l^k$ of a candidate stop reflects how many other stops can be reached without having to connect. Transfer stations tend to receive a greater weight while stops located on shorter routes with a limited number of destinations receive a lower weight. Due to the presence of decision variables in the numerator and denominator, the spatial interaction coverage model is a non-linear optimization problem and a solution cannot be obtained by a commercial solver. Farhan and Murray (2006) have transformed the model to a Maximal/Minimal Covering-Distance Decay Problem (MCDDP) in which all facilities are characterized by the same attraction value and separation distance is linear ($\alpha = \beta = 1$). Although a solution can be found to the MCDDP by commercial solvers, it has limitations in the choice of exponent parameters, explicitly modeling the importance of attraction and distance decay. An alternative is to use a heuristic to solve the spatial interaction coverage model.

*3.2. Simulated annealing*

Since Eq. (1) is non-linear, a total enumeration of all candidate facilities is not feasible, due to combinatorial explosion (Grötschel & Lovàsz, 1995). The search for an approximate solution is conducted using a suitable heuristic method $H$, which helps to identify an optimal set of facilities $J^*$ (or near optimal $J^+$) $\subset J$. Simulated Annealing (SA) is a method by which a metal cools and freezes into a minimum energy crystalline structure. The heuristic algorithm was originally proposed as a means of finding the equilibrium configuration of a collection of atoms at a given temperature. Kirkpatrick, Gelatt, and Vecchi (1983) made the connection between the cooling technique and the mathematical optimization. The major advantage of SA is its ability to avoid becoming trapped at a local optimum. The algorithm employs a random search that accepts changes improving the objective function, but also non-improving moves. The latter is accepted with probability $\delta_T$, where $T$ is the current temperature. $T$ cools down as the algorithm progresses, and so does the probability of accepting non-improving solution. In location modeling, simulated annealing has been applied to various problems, such as the $p$-median problem (Murray & Church, 1996) and non-linear problems like as cell tower allocation (Akella, Delmelle, Batta, Rogerson, & Blatt, 2010), sampling selection (Delmelle & Goovaerts, 2009) and multi-site land use allocation problems (Aerts & Heuvelink, 2002). Some parameters are first defined, followed by the simulated annealing algorithm in Fig. 2, adapted to the SIC.

**Parameters**

| | |
|---|---|
| $J(k)$ | solution at iteration $k$, $J(k) = \{X_j = 1, \forall j = 1, \ldots, J\}$ |
| $J(k+1)$ | solution at iteration $k + 1$. |
| $Z[J(k)]$ | objective function value at $J(k)$. |
| $\triangle Z$ | difference in objective function value between two successive iterations |
| $T$ | current temperature. |
| $T_{st}$ | starting temperature. |
| $T_{fin}$ | stopping temperature (tolerance). |
| $T_{max}$ | number of iterations per temperature level. |
| $\kappa$ | temperature decreasing factor (cooling schedule). |
| $\delta_T$ | probability to accept a move as a function of temperature $T$. |
| $J^\diamond$ | incumbent solution |
| $J^+$ | best solution |
| $J^*$ | optimal solution |

The optimization starts with a randomly selected scheme $J(k)$ of $p$ candidate locations ($|J(k+1)| = \sum_{j \in J} X_j = p$), the objective function $Z[J(k)]$ (Eq. 1) is evaluated and called the incumbent solution. Since the solution spqce has not been fully explored yet, the value of $Z[J(k)]$ is kept in memory as the best solution $J(k) = J^+$ found so far. $J(k)$ becomes $J(k+1)$, by swapping two elements of $J(k)$, but making sure that the number of open facilities remain equal to $p$ (see Table 1). $J(k+1)$ becomes the new, current solution if it has a better objective function value than $J(k)$. When $Z[J(k+1)] < Z[J(k)]$, $J(k+1)$ is accepted with probability $\delta_T$ (see Eq. 9).

**Step 1**: Set $k = 1$, $T = T_{\mathrm{st}}$ Generate initial solution $J(k)$ and its objective function value $Z\big[J(k))\big]$, with $|J(k)| = \sum_{j\in J} X_j = p$. Make $J^{\diamond} = J^{+} = J(k)$ and $Z\big[J^{+}\big] = Z\big[J^{\diamond}\big] = Z\big[J(k)\big]$.

**Step 2**: While $T > T_{\mathrm{fin}}$

- **Step 2.1**: Initialize the number of iterations per temperature $T_{\mathrm{it}} = 1$.
- **Step 2.2**: While $T_{\mathrm{it}} < T_{\max}$:
  - Increment $k$, $k = k + 1$. Define $J(k+1)$ as a modified set, where two elements $m, n \in J(k), m \neq n$, are swapped, and $|J(k+1)| = \sum_{j\in J} X_j = p$ (see Table 1) Report its objective function value difference $\triangle Z = Z\big[J(k+1))\big] - Z\big[J(k))\big]$.
    - If $\triangle Z < 0$ accept with probability $\delta_T$. Generate random number $\alpha$
      - If $\alpha \leq \delta_T$, accept solution. Make $J(k+1) = J(k)$ and update incumbent $J^{\diamond} = J(k+1)$.
      - If $\alpha > \delta_T$ do not accept the move.
    - If $\triangle Z \geq 0$, make $J(k+1) = J(k)$ and update incumbent $J^{\diamond} = J(k+1)$.
  - If $J^{\diamond} \geq J^{+}$, make $J^{+} = J^{\diamond}$.
- **Step 2.3**: Make $T_{\mathrm{it}} \leftarrow T_{\mathrm{it}} + 1$. If $T_{\mathrm{it}} < T_{\max}$, go back to **Step 2.2**.

**Step 3**: Reduce the temperature by a factor $\kappa$, i.e. $T = T \times \kappa$. Go back to **Step 2**.

**Fig. 2.** The Simulated Annealing algorithm (spatial interaction coverage model).

$$\delta_T\{J(k) \to J(k+1)\} = 1 \qquad \text{if } Z[J(k+1)] \geqslant Z[J(k)] \tag{8}$$

$$\delta_T\{J(k) \to J(k+1)\} = \frac{1}{1 + e^{\left(\frac{\triangle Z}{T}\right)}} \qquad \text{if } Z[J(k+1)] < Z[J(k)] \tag{9}$$

The process continues in a similar fashion until it reaches a cooled state. In order to find a solution $J^{+}$ close to the optimal $J^{*}$, a high starting temperature $T_{\mathrm{st}}$ and a cooling factor $\kappa$ close to 1 are necessary. This allows the algorithm to escape from a local maximum.

A tight GIS coupled system integrates the SIC model in ArcGIS[2] by means of its programming language, Visual Basic for Applications (VBA). Demand nodes and facilities serve as input in the form of points. An interface for the simulated annealing heuristic was built (see Fig. 3), allowing the user to input various parameters, such as the weight $w_j$ of each facility -corresponding to a table field-, the importance of that weight $\alpha$, the distance deterrence factor $\beta$, the number of facilities and the coverage radius (demand nodes located beyond a certain distance will not be covered). Various simulated annealing parameters also need to be specified and calibrated. The sum of the objective is computed according to which facilities $X_j$ are open, $\sum_{j\in J} X_j$. The procedure continues in a similar fashion following the structure of the simulated annealing heuristic. Once the model is solved, the final optimal solution is automatically displayed as a spider map showing selected facility locations and covered demand areas.

## 4. System evaluation

In this section, we apply our methodological framework to an urban route of the transit system in Charlotte, North Carolina. The interface was presented during two meetings (December 2010 and May 2011) with one of the authors and three transit planners of the Charlotte Area Transit System (CATS). CATS operates 56 local routes and 19 express routes. Express routes link nearby suburbs at the periphery of the city to its main center. There are 3662 bus stops scattered over Mecklenburg county, with a few stops located in adjacent counties serving as terminal stations for express routes (Fig. 4). Overall boardings have increased in the last ten years, but at a much smaller pace than the CATS operating budget. The structure of the transit network radiates out from the center along major roads, forcing commuters to make a stopover in the center of the city. Another discouraging factor is the redundancy of bus stops (small inter-separation distance), decreasing overall system efficiency. Systematic evaluation is an important step in the development of sustainable transportation (Banister, 2008; Bender et al., 2004). During our first meeting, the spatial interaction coverage model interface was demonstrated on a small dataset. A questionnaire of six open-ended questions was given to obtain feedback regarding bus stop placement considerations, implementation and usefulness of the developed system for modeling bus transit.

**Table 1**
Simulated annealing algorithm: 5 iterations. Swapped elements are indicated in bold.

| $k$ | $k+1$ | $k+2$ | $k+3$ | $k+4$ |
|---|---|---|---|---|
| 0 | **1** | 1 | 1 | **0** |
| 1 | **0** | 0 | 0 | **1** |
| 1 | 1 | **0** | 0 | 0 |
| 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 |
| 0 | 0 | **1** | 1 | 1 |
| 1 | 1 | 1 | **0** | 0 |
| 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | **1** | 1 |

CATS uses a 400 m ($\frac{1}{4}$-mile) coverage radius for their stop catchment area, but this distance can vary to 1080 m ($\frac{2}{3}$-mile) when some individuals are heavily dependent on public transportation. It was agreed that 400 m ($\frac{1}{4}$-mile) would provide a conservative estimate. Connectivity, surrounding land-use and residential density, distance to the nearest crosswalk, and ease of crossing are important components when modeling facility attraction. Surrounding land-use and residential density is generally reflected on the demand side of the model. The presence of large retail opportunities was regarded as critical for keeping stops on a route. Finally, CATS seeks feedback from the public through town-hall like meetings. An example was given on how boardings on particular segments of the airport route increased since a stop was recently placed closer to a large retail store due to public request, which eventually increased ease of crossing and boardings. CATS suggested applying the model to a single

[2] Interface available at http://geoearth.uncc.edu/people/edelmelle/gis/SICSA.html Last accessed January 21 2012.

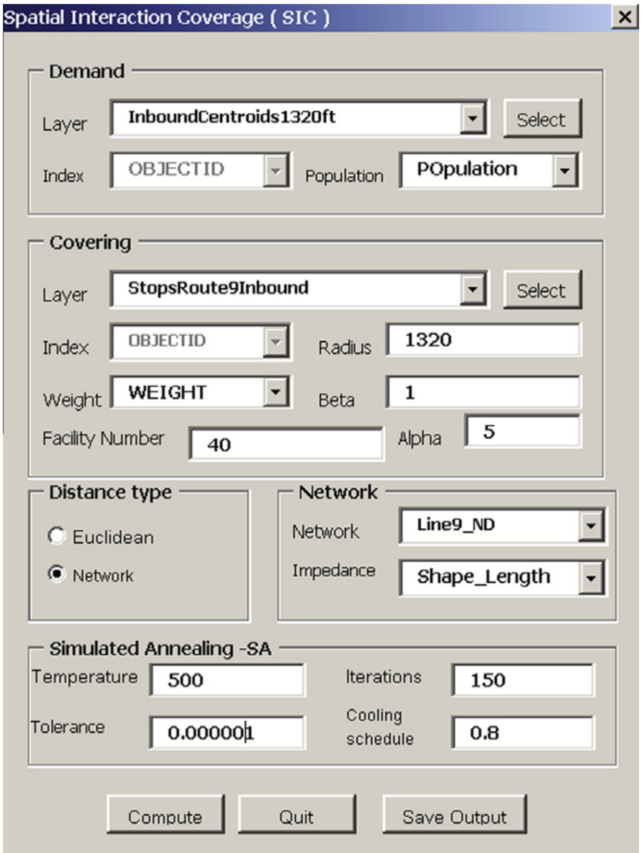


**Fig. 3.** The GIS-based interface for the SIC simulated annealing heuristic.

inbound urban route rather than optimizing an entire system, recognizing that each route is specific to its commuters. Transit planners suggested that the results from this analysis could provide valuable insights in their decision to remove some stops separated by very small distances.

## 5. Application

The spatial interaction coverage was applied to route 9 in the city of Charlotte, North Carolina. Route 9 in thick, black in Figs. 4 and 5a runs along central Avenue, connecting the southeastern end of Charlotte to the Charlotte Transportation Center in the downtown area, and serving the Eastland Mall. The line is 16-km (10-miles) long, with an average frequency of five buses in peak hours, taking a little over 30 min to connect the two ends of the line. Route 9 is characterized by 46 inbound and 47 outbound stops. On the inbound route alone (Fig. 5), some stops are only separated by one block, indicating possible redundancy.

**Connectivity** at each bus stop in CATS is computed using Eq. 7. Fig. 4 illustrates the spatial distribution of the probability to reach all other bus-stops, either with no transfers ($k = 0$) or one connection ($k = 1$) in the city of Charlotte. In the non-stop scenario, two to three major stations exhibit superior connectivity. One of them in the center of Charlotte is the Transportation Center, well connected to many routes, while to the Southwest is the second highly connected area around the SouthPark Community Transit Center, close to the SouthPark mall. The area North of the city is less connected, given that a very limited number of express routes run along the North–South corridor to the periphery, and these routes have few bus stops. The center of the City is well connected as are areas to the west corresponding to the University.

The non-stop connectivity of bus stops for Route 9 is summarized in Fig. 5b, keeping in mind that this number will increase when there are multiple bus routes running in parallel at a given stop. Not surprisingly, inbound stops closer to the terminal stations (where different routes converge) have higher connectivity. For the same reason, stops close to the beginning of the line have greater non-stop connectivity (since the Albermale Express −40×, route 221 and route 222 are also parallel on some segments of route 9).

To assess the **walking accessibility** of each transit stop, we used 2008 aerial photos. For each stop, the network distance to the nearest crosswalk was estimated in GIS. This approach was suggested during our first meeting with CATS as an important aspect of bus stop accessibility, and also recently echoed in Daniels and Mulley (2011). In Fig. 6, we illustrate this criterion for the stops at the beginning of the inbound route. The map indicates that stops nearby a crosswalk were more likely to receive a higher crossing score.

**Demand** is generally modeled using population information from the census; in the United States, this data is made available at different levels (tract, block group or block). In this paper, we take a more disaggregated approach using parcel data for the year 2008. This vector data for Mecklenburg County, North Carolina contains, for each parcel, descriptive information on land use (commercial, offices, government institutions and residential) and square footage. We pool these different types of data together into one single demand field, since only considering residential

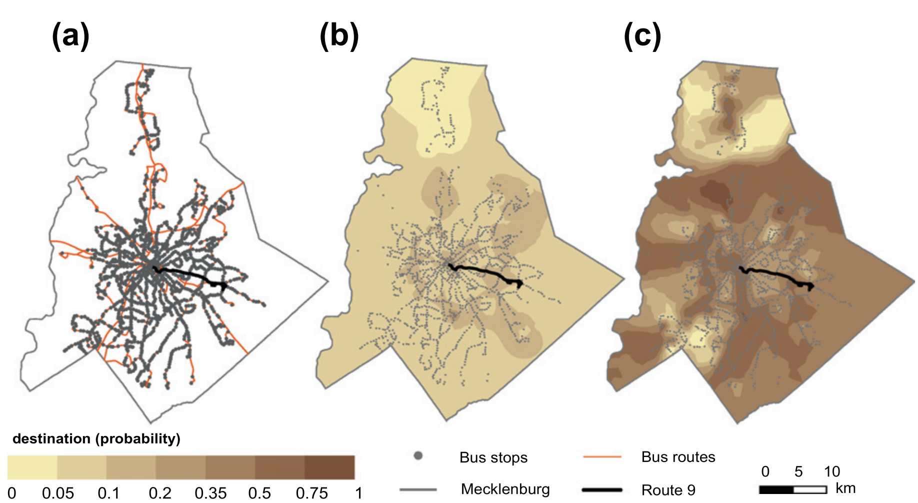


**Fig. 4.** Bus transit system for Mecklenburg County, NC. The interpolated maps in (b) and (c) represent the spatial distribution of the probability to reach all destinations with no transfers ($k = 0$) and one transfer ($k = 1$), respectively.

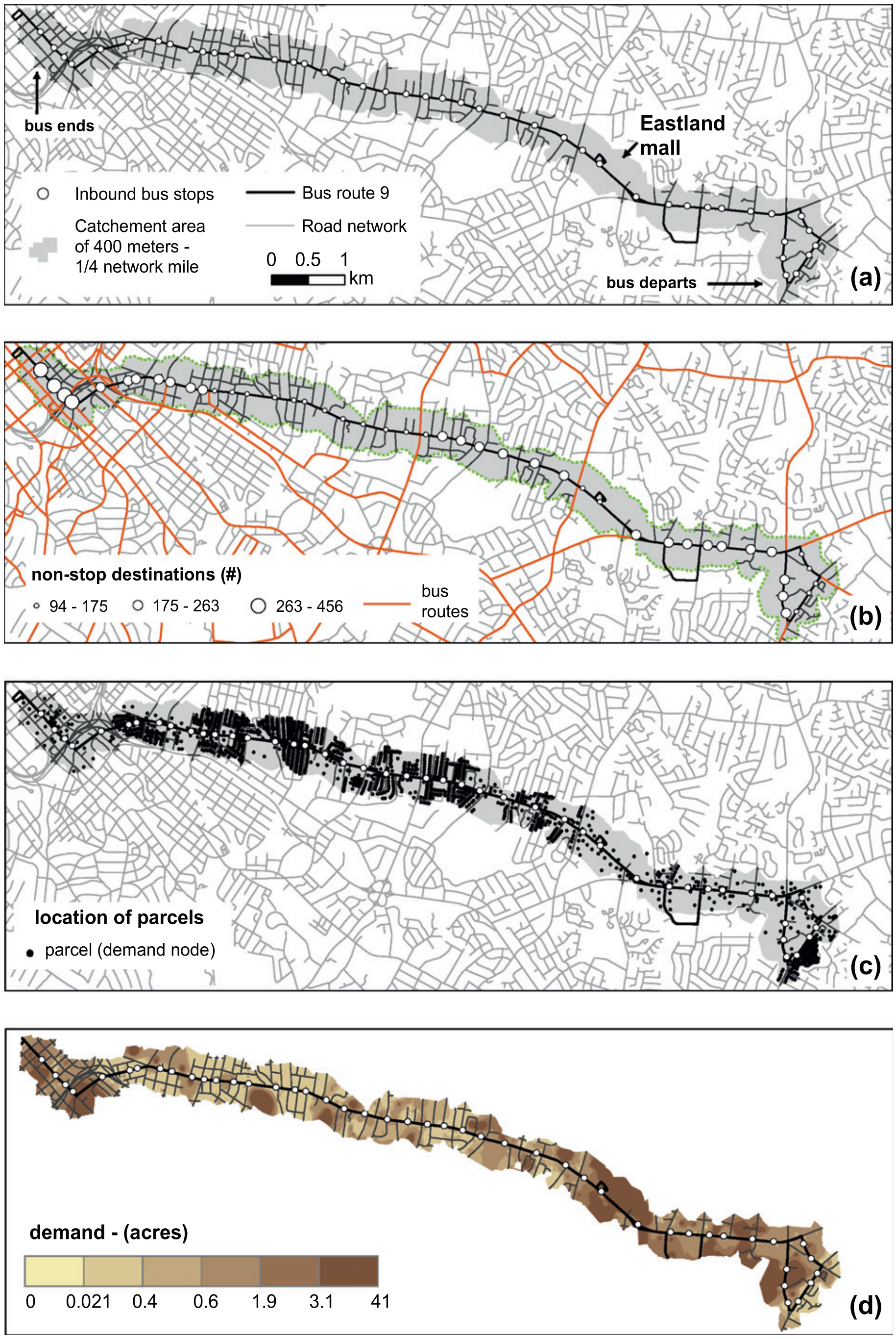


**Fig. 5.** CATS bus route 9, with inbound stops (direction to downtown). (*b*) indicates the non-stop connectivity on the line, with a minimum number of 93 stops (along route 9). (*c*) displays the location of parcels (demand nodes), while (*d*) is the interpolated demand.

parcels would underestimate transit usage. Fig. 5d illustrates the variation in demand along the route, with increasing values (in gray) close to the Eastern Mall and in the downtown area. Residential areas occur more so in the second part of the route.

We applied the SIC on two different scenarios, keeping a similar access distance of $R = 400$ m (or 1320 feet; $\frac{1}{4}$-mile), but changing how many transit stops are maintained. In scenario 1, we modeled facility attraction solely based the non-stop connectivity at each stop, using parameters $\alpha = 2$ and $\beta = 2$. That is, we ignore accessibility attributes surrounding each stop. In the second scenario, bus stop attraction was modeled as a function of walking access and ignored non-stop connectivity along a route. That is, a facility was made more attractive with decreasing distance to the nearest intersection ($\alpha = 2$ and $\beta = 2$). Fig. 7 illustrates that a greater amount of interaction can be achieved with a larger number of stops, but up to a critical number of stops ($p = 25$). Although not displayed here, when the distance impedance was raised significantly ($\beta = 2 \rightarrow \beta = 5$), the interaction increases up to 30 stops, indicating that a greater number of stops was necessary and that the same amount of interaction could not be reached when using $\beta = 2$. Although this result is intuitive, it has some critical implications when modeling access for groups with limited mobility, since distance is experienced much differently. The shape of the interaction function was quite similar when only optimizing walking access in scenario 2. Adding stops beyond $p = 30$ in scenario 2 did not increase the objective function. Fig. 8 shows the solution for scenarios 1 and 2 with $p = 15$ and $R = 400$ m. Selected stops are green-colored, and so are their catchment area for cartographic

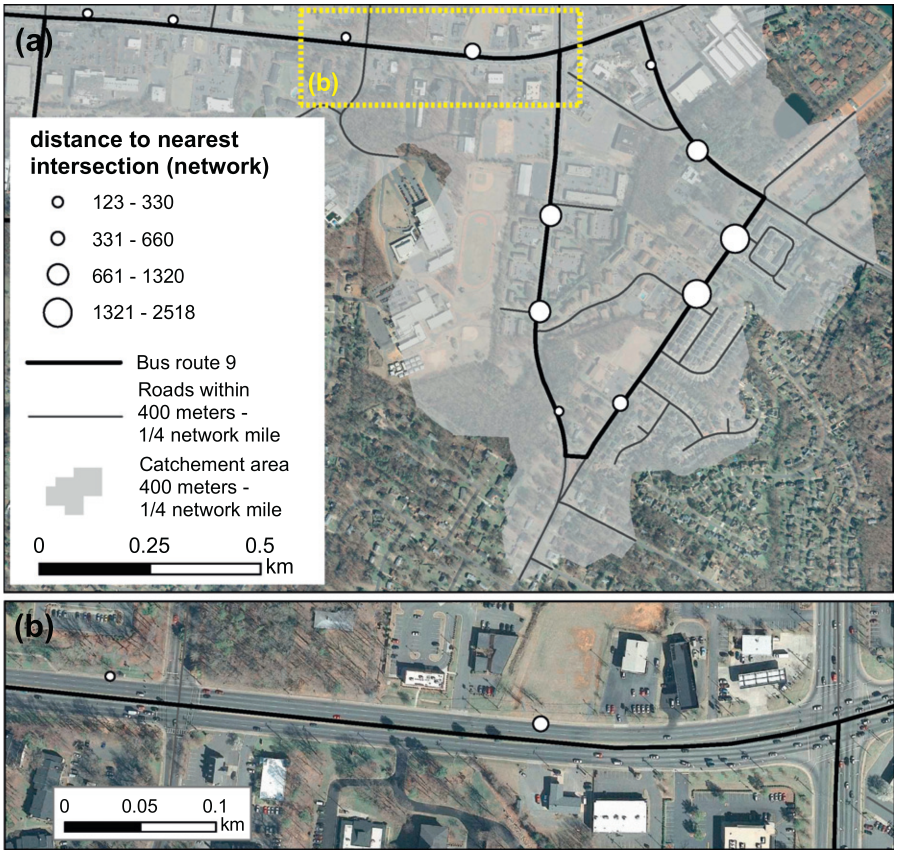


**Fig. 6.** Distance to the nearest crosswalk for each inbound stop on route 9.

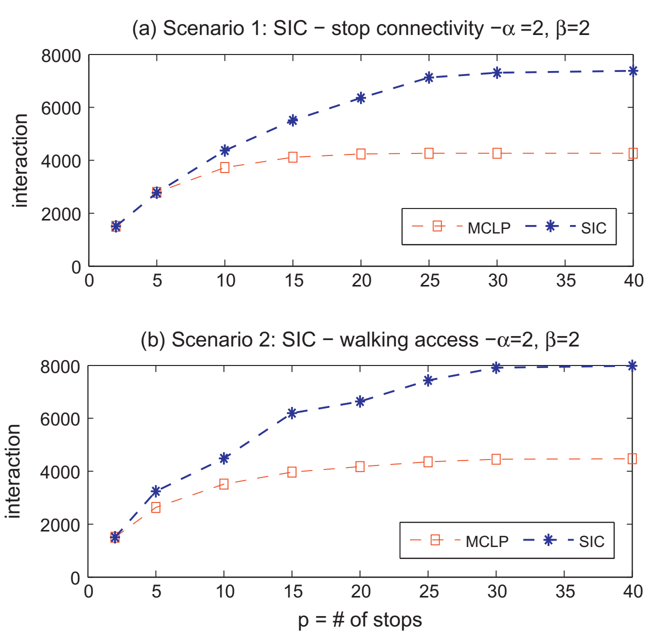


**Fig. 7.** Variation in the interaction as a function of the number of transit stops $p$ with varying model parameters (($a$): scenario 1 and ($b$): scenario 2).

association. Reducing the number of stops increases transit efficiency, resulting in a shorter transit time to connect origin and destination of the route. In **scenario** 1, the model keeps those stops in the close vicinity of large demand areas (mall) and close to the Charlotte Transit Center. Most significantly, the model selects stops with high attractiveness, that is stops with high number of non-stop connectivity. The last four stops close to the destination are selected in the optimization process, since all of them are characterized by a relatively large demand and very high connectivity. There is a long stretch of residential area in the second section of the route which is not covered at all. This contrasts with the results for **scenario 2** (Fig. 8b), where the model ignores connectivity and solely attempts to keep these stops with good physical access -closer to crosswalks- and where demand is large. In this case the model also identifies a stretch of 10 stops in the second residential area which are not deemed vital. When **scenario** 2 with Fig. 6, five consecutive stops characterized by small distances to the nearest intersection are selected. Although providing good crossing scores, the stops in the lower end where the bus starts its route are not selected due to very low demand.

Of interest is whether results found for a Maximal Covering Location Problem (MCLP) differ significantly from the ones optimizing spatial interaction coverage. To obtain a solution to the maximal covering location problem, a VBA script was developed within the interface, exporting the associated integer program to a text file and solved using Lingo, an optimization solver. We compare the solutions obtained for the MCLP for various values of $p$ with a radius $R = 400$ m (1320 ft. or $\frac{1}{4}$-mile) to the SIC for Route 9. The MCLP solution for $p = 15$ is displayed in Fig. 8c. Although not shown here, the MCLP objective function increases up to $p = 15$ facilities and flattens out to $p = 20$, suggesting that adding more facilities beyond $p = 15$ will not increase covered demand. As expected, MCLP chooses bus stops that can be utilized by the greatest number of individuals, regardless of their attractiveness, but tends to spread them more geometrically. In the MCLP, more stops are located in the residential area than in the SIC model. The latter can be a consequence of simultaneously modeling facility attraction

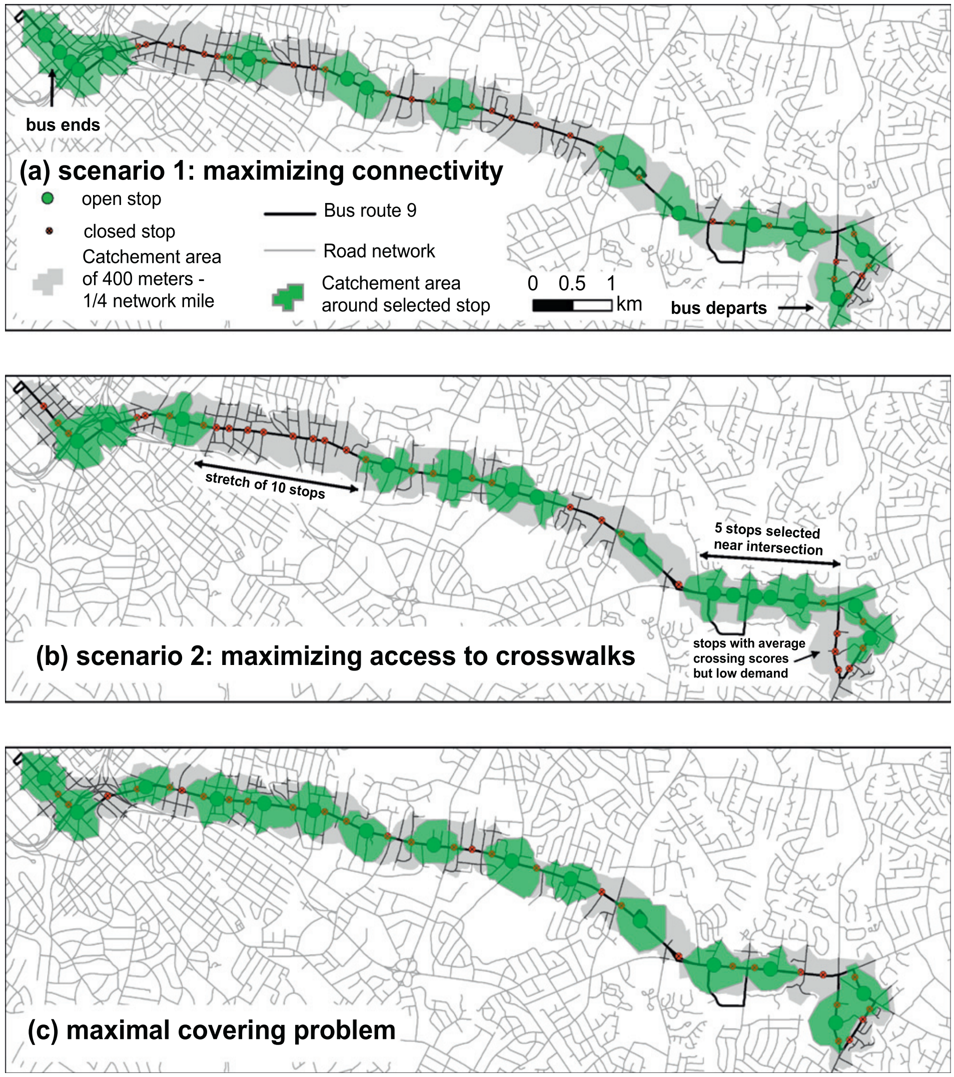


**Fig. 8.** Solution to scenarios 1 and 2, and the maximal covering location problem, with $p = 15$ and $R = 400$ m (1320 ft. or $\frac{1}{4}$-mile).

multiplied by demand and distance decay. The red curve in Fig. 7 summarizes the MCLP solutions evaluated as a SIC objective. We take the optimal MCLP solution for each $p$-value and compute what the SIC objective would be for that solution. We observe that the MCLP solution is non-inferior (below the pareto curve), indicating that the MCLP cannot encapsulate the goals inherent to the SIC.

## 6. Conclusions

The Spatial Interaction Coverage (SIC) model has been developed within a GIS, and applied to a transit planning problem. The model represents the interaction between demand nodes and facilities by explicitly accounting for facility attraction, distance deterrence and competition among nearby, open facilities. Because of its non-linearity, a simulated annealing heuristic was developed to solve the model. A GIS-based interface was built to integrate both candidate locations and demand nodes as point data, but also to manage model parameters. It gives the transit agencies decision making tools to decide on which redundant stops to eliminate and which strategic stops to maintain. The proposed model was further refined through system evaluation with local transit planners, and applied to a inbound bus route within the Charlotte area. Three different scenarios have been discussed reflecting different approaches to model facility attraction.

The spatial interaction coverage model provides an alternate approach to model public transit efficiency and redundancy. In comparison to the MCLP, the SIC model possesses undeniable benefits as it accounts for the magnitude of facility attraction as well the importance of distance decay. The unfortunate drawback is that a solution to the model cannot be found by commercial solvers unless exponents are reduced to 1. Although simulated annealing is relatively easy to implement and its algorithmic structure is straightforward, its success strongly depends on: (1) the generated initial solution, (2) the starting temperature, and (3) the number of iterations per temperature level. More research is needed to understand the behavior of the algorithm with problem instances of different sizes. Ideally, superior and faster search algorithms, fuzzy classification or Tabu search will tend to return optimal (or near-optimal) solutions in a much faster time frame.

The methods used to calibrate facility attraction weights and willingness-to-walk distance are debatable, and further investigation through commuter surveys is necessary. Positive attributes can further be modeled such as handicap access, bike racks, shelters and public phones, while negative attributes such as crime, the presence of vacant buildings, graffiti or litter could impact the attractiveness of a transit stop. It is relatively straightforward to combine these different indicators as a composite index. Ultimately, understanding which factors may impact transit boardings at each stop would improve the estimation of a facility attraction. The choice of an appropriate value for distance and attraction parameters $\alpha$ and $\beta$ is a challenge. As Griffith and Jones (1980) point out, there may be different expressions for distance decay as spatial structure of origins and destinations may vary, as the propensity of origins to emit demand and the attraction of a destination may change with varying urban morphology. Some individuals may be willing to walk a longer distance in suburban areas to reach a transit stop (Daniels & Mulley, 2011; O'Sullivan & Morrall, 1996), and low-income groups with no car ownership and deprived segments of the population may have no other alternative than using public transit. This could partially be addressed by modifying the coverage radius to a location-specific distance $R_i$. Finally, as pointed out in the larger case study, bus routes connect to one another when they share a common bus stop. In reality, connectivity may be increased if closeby stops are merged into one single facility.

Other areas which merit investigation are: (1) how the type of the network (euclidean, network) and distribution of distances affect the quality of the solution (Peeters & Thomas, 1995, 2000; Schilling, Rosing, & ReVelle, 2000), and (2) the possibilities to disseminate this model in UrbanSim (Waddell, 2002), for instance, to project what bus stop attractiveness would be in the future.

## Acknowledgements

The authors would like to thank Jason Lawrence and Ramond Robinson from the Charlotte Area Transit System for the feedback on the SIC interface and guidance in the selection of interaction parameters.

## References

Aerts, J., & Heuvelink, G. B. M. (2002). Using simulated annealing for resource allocation. *International Journal of Geographical Information Science, 16*, 571–587.
Agrawal, A. W., Schlossberg, M., & Irvin, K. (2008). How far, by which route and why? A spatial analysis of pedestrian preference. *Journal of Urban Design, 13*, 81–98.
Akella, M., Delmelle, E., Batta, R., Rogerson, P., & Blatt, A. (2010). Optimal adaptive facility location using geostatistics. *Geographical Analysis, 42*, 227–244.
Alam, B., Thompson, G., & Brown, J. (2010). Estimating transit accessibility with an alternative method: Evidence from Broward County, Florida. *Transportation Research Record: Journal of the Transportation Research Board, 2144*, 62–71.
Alexandris, G., & Giannikos, I. (2010). A new model for maximal coverage exploiting GIS capabilities. *European Journal of Operational Research, 202*, 328–338.
Ammons, D. N. (2001). *Municipal benchmarks: Assessing local performance and establishing community standards* (2nd ed.). Thousand Oaks, CA: Sage.
Banister, D. (2008). The sustainable mobility paradigm. *Transportation Policy, 15*, 73–80.
Bender, T., Hennes, H., Kalcsics, J., Melo, T., & Nickel, S. (2004). Location software and interface with GIS and supply chain management. In H. Hamacher & Z. Drezner (Eds.), *Facility location: Applications and theory* (pp. 233–274). Heidelberg, Germany: Springer-Verlag.
Blumenberg, E., & Manville, M. (2004). Beyond the spatial mismatch: Welfare recipients and transportation policy. *Journal of Planning Literature, 19*, 182–205.
Buehler, R. (2010). Transport policies, automobile use, and sustainable transport: A comparison of Germany and the United States. *Journal of Planning Education and Research, 30*, 76–93.
Bullard, R. (2003). Addressing urban transportation equity in the United States. *Fordham Urban Law Journal, 31*, 1183.
Casas, I. (2007). Social exclusion and the disabled: An accessibility approach. *The Professional Geographer, 59*, 467–477.
Church, R. L., & Marston, J. R. (2003). Measuring accessibility for people with a disability. *Geographical Analysis, 35*, 83–96.
Currie, G., & Phung, J. (2008). Understanding links between transit ridership and gasoline prices: Evidence from the United States and Australia. *Transportation Research Record: Journal of the Transportation Research Board, 2063*, 133–142.
Daniels, R., & Mulley, C. (2011). Explaining walking distance to public transport: the dominance of public transport supply. In *World Symposium on Transport and Land Use Research*. Whistler Canada, 28–30 July.
Delmelle, E., & Goovaerts, P. (2009). Second-phase sampling designs for non-stationary spatial variables. *Geoderma, 153*, 205–216.
Delmelle, E., Delmelle, E. C., Casas, I., & Barto, T. (2011). H.E.L.P: A GIS-based health exploratory analysis tool for practitioners. *Applied Spatial Analysis and Policy, 4*, 113–137.
Demetsky, M. J., & Lin, B. (1982). Bus stop location and design. *Transportation Engineering Journal of ASCE, 108*, 313–327.
Farhan, B., & Murray, A. T. (2006). Distance decay and coverage in facility location planning. *Annals of Regional Science, 40*, 279–295.
Foda, M., & Osman, A. (2010). Using GIS for measuring transit stop accessibility considering actual pedestrian road network. *Journal of Public Transportation, 13*, 23–40.
Federal Transit Administration (1996). Guidelines for the location and design of bus stops. *Transit cooperative research program, report 19*. Washington, DC: National Academy Press.
Furth, P. G., & Rahbee, A. B. (2000). Optimal bus stop spacing trough dynamic programming and geographic modeling. *Transportation Research Record: Journal of the Transportation Research Board, 1713*, 15–22.
Garrison, W. L., & Levinson, D. M. (2006). *The Transportation Experience Policy, Planning, and Deployment*. Oxford: Oxford University Press.
Ghosh, D. (2008). A loose coupling technique for integrating GIS and multi-criteria decision making. *Transactions in GIS, 12*, 365–375.
Giuliano, G. (2004). The urban transport system. In *World conference on transport research society, & institute for transport policy studies, transport and the environment: An international perspective* (pp. 37–98). Brighton, UK: Elsevier Press.
Griffith, D., & Jones, K. (1980). Explorations into the relationship between spatial structure and spatial interaction. *Environment and Planning A, 12*, 187–201.
Grötschel, M., & Lovàsz, L. (1995). Combinatorial optimization. In R. L. Graham, M. Grötschel, & L. Lovàsz (Eds.), *Handbook of combinatorics* (pp. 1541–1579). Amsterdam, The Netherlands: Elsevier.
Gutierrez, J., & Garca-Palomares, J. C. (2008). Distance-measure impacts on the calculation of transport service areas using GIS. *Environment and Planning B: Planning and Design, 35*, 480–503.
Hakimi, S. L. (1964). Optimum locations of switching centers and absolute centers and medians of a graph. *Operations Research, 19*, 1363–1406.
Hamacher, H.W., Hennes, H., Kalcsics, J., & Nickel, N. (2003). LoLA: Library of location algorithms, Version 2.1, Fraunhofer ITWM, <www.mathematik.uni-kl.de/lola/> Accessed: 01.11.11.
Hess, D.B. (in press). Walking to the bus: Perceived versus actual walking distance to bus stops for older adults. Transportation. doi:10.1007/s11116-011-9341-1.
Huff, D. L. (1963). A probabilistic analysis of shopping center trade areas. *Land Economics, 39*, 81–90.
Ibeas, A., dellOlio, L., Alonso, B., & Sainz, O. (2010). Optimizing bus stop spacing in urban areas. *Transportation Research Part E: Logistics and Transportation Review, 46*, 446–458.
Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science, 220*, 671–680.
Levinson, H. S. (1983). Analyzing transit travel time performance. *Transportation Research Record: Journal of the Transportation Research Board, 915*, 1–6.
Malczewski, J. (1999). *GIS and multicriteria decision analysis*. New York: John Wiley & Sons.
Murawski, L., & Church, R. (2009). Improving accessibility to rural health services: The maximal covering network improvement problem. *Socio-Economic Planning Sciences, 43*, 102–110.
Murray, A. T. (2003). A coverage model for improving public transit system accessibility and expanding access. *Annals of Operations Research, 123*, 143–156.
Murray, A. T. (2010). Advances in location modeling: GIS linkages and contributions. *Journal of Geographical Systems, 12*, 335–354.
Murray, A. T., & Church, R. (1996). Applying simulated annealing to location-planning models. *Journal of Heuristics, 2*, 31–53.
Murray, A. T., & Wu, X. (2003). Accessibility tradeoffs in public transit planning. *Journal of Geographical Systems, 5*, 93–107.
Murray, A. T., Davis, R., Stimson, R., & Ferreira, L. (1998). Public transportation access. *Transportation Research Part D: Transport and Environment, 3*, 319–328.
Neilson, G., & Fowler, W. (1972). Relation between transit ridership and walking distances in a low density Florida retirement area. *Highway Research Record, 403*, 26–34.
Newman, P., & Kenworthy, J. (1999). *Sustainability and cities: Overcoming automobile dependence*. Island Press: Washington, DC.
O'Sullivan, S., & Morrall, J. (1996). Walking distances to and from light-rail transit stations. *Transportation Research Record: Journal of the Transportation Research Board, 1538*, 19–26.
Peeters, D., & Thomas, I. (1995). The effect of spatial structure on p-median results. *Transportation Science, 29*, 366–373.
Peeters, D., & Thomas, I. (2000). Choosing a measure of distance for applied location–allocation models: Simulations based on the LP-distance and the k-median model. *Geographical Systems, 2*, 167–184.
Preston, V., & McLafferty, S. (1999). Spatial mismatch research in the 1990s: Progress and potential. *Papers in Regional Science, 78*, 387–402.
Pucher, J. (2004). Public transportation. In S. Hanson & G. Giuliano (Eds.), *The geography of urban transportation*. New York: Guilford Press.
Pucher, J., & Lefevre, C. (1996). *The urban transport crisis in Europe and North America*. London: Macmillan.
Pulugartha, S., & Vanapalli, V. (2008). Hazardous bus stop identification: An illustration using GIS. *Journal of Public Transportation, 11*, 65–83.
ReVelle, C., & Swain, R. (1970). Central facilities location. *Geographical Analysis, 2*, 30–42.

Saka, A. A. (2001). Model for determining optimum bus-stop spacing in urban areas. *Journal of Transportation Engineering, 127*, 195–196.
Schilling, D. A., Rosing, K. E., & ReVelle, C. S. (2000). Network distance characteristics that affect computational effort in p-median location problems. *European Journal of Operational Research, 127*, 525–536.
Schobel, A. (2005). Locating stops along bus or railway lines – a bicriteria problem. *Annals of Operations Research, 136*, 211–227.
Taaffe, E. J., Gauthier, H. L., & O'Kelly, M. E. (1996). *Geography of transportation* (2nd ed.). Prentice Hall: New Jersey.
Tong, D., & Murray, A.T. (2009). Maximizing coverage of spatial demand for service. *Papers in Regional Science, 88*, 85–97.
Transportation Research Board (2006). Commuting in America III. In A. Pisarski (Ed.), *The Third National Report on Commuting Patterns and Trends*. Washington DC.
Truong, L. T., & Somenhalli, S. (2011). Using GIS to identify pedestrian–vehicle crash hotspots and unsafe bus stops. *Journal of Public Transportation, 14*, 99–114.
US Census Bureau (2009). Commuting in the United States: 2009 (edited by B. McKenzie and M. Rapino). <http://www.census.gov/prod/2011pubs/acs-15.pdf> Accessed 01.11.11.
Waddell, P. (2002). Modeling urban development for land use, transportation and environmental planning. *Journal of the American Planning Association, 68*, 27–314.
Wells, K., & Thill, J.-C. (2011). Do transit-dependent neighborhoods receive inferior bus access? A neighborhood analysis in four US cities. *Journal of Urban Affairs, 34*, 43–63.
Wirasinghe, S. C., & Ghoneim, N. S. (1981). Spacing of bus-stops for many to many travel demand. *Transportation Science, 15*, 210–221.