European Journal of Operational Research 224 (2013) 132–140

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

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

Decision Support

Shortest path games

Edward C. Rosenthal *

Department of Marketing and Supply Chain Management, Fox School of Business, Temple University, Philadelphia, PA 19122, USA

ARTICLE INFO

Article history:
Received 3 November 2011
Accepted 27 June 2012
Available online 20 July 2012

Keywords:
Game theory
Shortest paths
Network design

ABSTRACT

We study cooperative games that arise from the problem of finding shortest paths from a specified source to all other nodes in a network. Such networks model, among other things, efficient development of a commuter rail system for a growing metropolitan area. We motivate and define these games and provide reasonable conditions for the corresponding rail application. We show that the core of a shortest path game is nonempty and satisfies the given conditions, but that the Shapley value for these games may lie outside the core. However, we show that the shortest path game is convex for the special case of tree networks, and we provide a simple, polynomial time formula for the Shapley value in this case. In addition, we extend our tree results to the case where users of the network travel to nodes other than the source. Finally, we provide a necessary and sufficient condition for shortest paths to remain optimal in dynamic shortest path games, where nodes are added to the network sequentially over time.

1. Introduction

The problem of finding the shortest paths in a network from a particular source to all other nodes is a well-known problem in operations research, computer science, and discrete mathematics. A polynomial time solution for the problem was given by Dijkstra (1959). Among the many applications of this technique, the one we focus on in this paper is the development of a rail system in a metropolitan area that will bring passengers in outlying stations to or from a central terminus as fast as possible.

We consider a growing metropolitan area with a single hub, or terminus, and a number of surrounding communities whose residents wish to travel to or from the terminus. To facilitate travel, the various municipalities in the region will work together to design and build a passenger rail system that connects the outlying communities to the terminus. The model and solutions we propose may also be employed for existing rail networks that are to be upgraded (for example, to install swifter service) or expanded (to reach outlying communities). The most natural objective for the rail system is to minimize the total travel time across all users. A network of shortest paths, as we show, is the network that achieves this objective.

An important associated problem is how to fairly allocate the cost of the rail network among the participating communities. We propose to solve this associated problem through the use of cooperative game theory. The contribution of this paper is to show how and under what conditions the cost of a shortest path transportation system can be allocated in a manner that is acceptable to all communities.

Despite the extensive literature on shortest path networks, there have been few attempts to study problems akin to the one we propose here. Laporte et al. (2011) study metropolitan area planning and develop networks with various objectives, such as to minimize cost subject to minimally sufficient population coverage, or to maximize a population-to-cost ratio. Laporte et al. (2010) consider the design of a railway with possible link failures and an alternative mode of transport to compensate for failed links. Schöbel and Schwarze (2006) treat network planning in which each line is a player that wishes to minimize cost, which is a function of the traffic on its links. Voorneveld and Grahn (2002) and Grahn (2001) consider cooperative games with a specified source and a sink, where players own sets of arcs and each subset of players S wants to maximize its profit, which is derived from transporting its goods via a shortest source-sink path owned by S. Fragnelli et al. (2000) treat cooperative shortest path games with multiple sources and sinks in which profit derived from transporting goods from a source to a sink is offset by the cost of the path used.

In related research, Nebel (2010) considers computational complexity of a model related to those of Voorneveld–Grahn and Fragnelli et al., while Bachrach and Rosenschein (2009) study power indices and the core of games in which winning coalitions succeed in sending a sufficient flow from a source to a sink. Our model does not study sharing the profit or success derived from transporting goods across a given network, but rather studies sharing of the cost of an efficiently designed network among its users. Finally, we point out that network design and cost allocation problems studied in the related telecommunications literature (Bird, 1976; Granot and Huberman, 1981; Bergantiños and Vidal-Puga, 2010) differ

* Tel.: +1 215 204 8177; fax: +1 215 204 8029.
E-mail address: edward.rosenthal@temple.edu

0377-2217/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2012.06.047

from the present work in the following respect: the objective in the telecommunications problems – to design a least-cost network to connect all users – requires the use of minimum cost spanning trees or Steiner networks, while the objective in the present work instead requires shortest path networks. Thus, despite passing similarities in this literature, none of the works cited study the problem that we do, which involves first, the creation of a network to provide the fastest possible transport for all network members to a central hub, and secondly, the issue of whether it is possible to share the costs of the network in a way that is acceptable to the participating communities.

In Section 2 we introduce the shortest path game and conditions that a solution ought to satisfy. In Section 3 we show that the core of the shortest path game is nonempty, meaning that no subset of players will have any incentive to deviate from the proposed cost-sharing solution. Then in Section 4 we restrict shortest path games to tree networks, provide a simple formula for the Shapley value in this case, and investigate its relation to the core. In Section 5 we extend the Shapley value result to the more general problem in which any node may act as a terminus. Finally in Section 6 we provide conditions under which a dynamic version of the shortest path game can be shown to guarantee the same outcomes as the static version, and in Section 7 we conclude the paper with suggestions for future research.

2. Shortest path games

First we reiterate our application more precisely. We assume that there is a metropolitan area M that includes a set $N = \{1,\ldots,n\}$ of distinct communities. Each community $i \in N$ has a population of *users* that reside within the geographical boundaries of that community. There is a single site s in M that we call the *source* or *terminus*. Site s will be the location of a train station, typically placed in M's central business district. Additionally, we assume that each community i will locate a single station where all of i's users will depart or arrive on their trips between i and s. There is a fixed cost for each community to build a station. While the users in any community may reside at different distances from their designated station i, we only consider their distances from s in units of time measured from i. Finally, we assume that all communities in M have the ability to build a link to s. Such links might go through land that lies outside of M, or else, with permission, cut across land owned by other communities in M. Note that in the network model we will develop, we will require the existence of an underlying complete graph for technical reasons, although it is understood that in real applications certain links are not possible to build (or would be extraordinarily expensive).

Each user will commute to s on a frequent basis. We assume that each user's utility function is based solely on his or her travel time between their local station and s. Since utility increases as travel time decreases, the users in M will each desire a commuter network G^* that minimizes their individual travel times to s.

The following is crucial to the understanding of the development of the model and the results that follow. Our shortest path solution will focus on travel time and not actual construction cost because we assume (as is reasonable to do) that the time saved for thousands of commuters over many years of travel will easily compensate for almost any construction cost premium. (For example, the cost of an expensive tunnel is more than compensated by the time it saves commuters.) Therefore, while the communities desire to reduce the individual (as well as total) travel times, as those travel times *decrease*, the net benefit to the individual communities, as measured by their utility of travel time minus their share of the construction cost, will *increase*, since it is assumed to be inversely proportional to the travel times.

To simplify the model that follows, we will only specify arc "lengths," which represent travel times. To minimize these lengths (or times) is to maximize benefit. In what follows, we will carry out two types of calculations. One is to generate shortest paths in the ordinary sense of minimizing travel times for the users of the network. The other type of calculation – to provide a fair allocation across users – requires an interpretation where the arc lengths are "costs" to be minimized. These latter quantities are not actual costs but rather are inversely related to user *benefits*. For ease of notation, then, when we allocate costs in the network, we will use the same arc quantities that represent distance, but the implicit understanding is that the users of the network aim to *minimize* these quantities, in order to *maximize* their net benefits.

Now we provide a set of conditions that are sensible from the communities' point of view.

Condition 1. Utility for each user is inversely related to their travel time from s.
Condition 2. No subset of communities can connect to s through stations owned by other communities.
Condition 3. The optimal network for a subset of communities should minimize the weighted travel times of their joint user population from the source.
Condition 4. No subset of communities should pay more, in the final cost allocation, than they would on their own.

Condition 1 is consistent with the present focus on transport time. Condition 2 preserves sovereign rights for each community to control the rail traffic through its station. However, the model will provide the opportunity for one community to obtain permission to build a link through another community's land, although not through its rail station site. Condition 3 focuses on the group objective, but below we will show that in the optimal solution the group objective is aligned with the separate users' objectives. Condition 4 calls for no cross-subsidization; more precisely, Condition 4 requires that no coalition in the shortest path game pay more, as part of the larger community, than they would have when building their own network.

To build a mathematical model for this application, we define a network $G = (N \cup \{s\}, A)$ as follows: set N is a set of *nodes* that represents the community set; A is the undirected complete graph of *arcs*, or *links*, between all pairs (i,j) of nodes with $i,j \in N \cup \{s\}$ such that $i \neq j$. The *source*, or *terminus* s may also represent an additional contiguous downtown community, but the residents of that community will not use the rail system. Each arc $(i,j) \in A$ has a *cost* or *distance* $a(i,j) \geqslant 0$ that represents the travel time between i and j (and, as discussed above, is inversely related to the benefit that the travel time confers on those who use this arc). Each node $i \in N \cup \{s\}$ has a cost $f_i \geqslant 0$ that represents the *fixed cost* of building a station at i. In addition, each node $i \in N$ has an associated nonempty set U_i of commuters or *users* present in that community; the cardinality of U_i is given by the *weight* $w_i > 0$ for all $i \in N$. We let $U = \cup(U_i)$. These users are assumed to be identical except for their locations. We define $G^*(N)$ (or just G^*) to be a network that minimizes the travel distance to s for M as a whole. We show in Lemma 1 that the network G' of shortest paths as found by Dijkstra (1959) provides the group-optimal solution G^*.

Lemma 1. *The shortest path network G^* is identical to G'.*

Proof. Dijkstra (1959) finds the network G' of shortest paths starting from a particular source (such as s) to all other nodes in a network, and shows that G' is a tree, meaning that each such shortest path is unique. The proof of Lemma 1 is now immediate: since all users reside at the nodes of G, and since G' minimizes each node's travel distance to s, the sum of these distances, for any non-negative weights w_i, is necessarily minimized as well, since no user has any shorter path to s outside of G'. Therefore, $G^* = G'$. □

It is well known that Dijkstra's algorithm runs in polynomial time, so designing the shortest path network for M is an easy task. We extend the definition of G^* in order to define it on subsets of G as follows: let network $G^*(N')$ for any $N' \subseteq N$ with a well-defined arc set represent the *shortest path network* on $N' \cup \{s\}$. Note that the optimal solution G^* satisfies Conditions 1–3. Now we turn to the problem of allocating the cost of the optimal network among M's communities. A *cooperative game* is defined by a set $\boldsymbol{P} = \{1,\ldots,n\}$ of *players* together with a *characteristic function* c: $2^N \to \mathbf{R}$ such that

$$c(\emptyset) = 0. \tag{1}$$

The characteristic function represents the minimum cost for any subset, or *coalition* $S \subseteq \boldsymbol{P}$, acting on its own, will need to pay in order to participate in the situation being modeled. Typically, there is an additional stipulation on the characteristic function, i.e.,

$$c(S \cup T) \leqslant c(S) + c(T) \quad \text{for } S,T \subseteq \boldsymbol{P} \text{ such that } S \cap T = \emptyset. \tag{2}$$

Expression (2), called *subadditivity*, means that when disjoint coalitions join together, they are able to do at least as well for themselves as they would have separately. (To be consistent with the cooperative game literature, in the cost allocation context we will continue to call the arc quantities $a(i,j)$ *costs*, but we remind the reader that the underlying motivation for coalitions in the game is to reduce the cost shares allocated to them because these costs are inversely related to the players' ultimate benefits.)

We define a cooperative game for our network application by letting the set U of users in G be the set of players. For any subset $S \subseteq U$, let $N(S)$ be the set of nodes in N such that a node is an element of $N(S)$ if and only if it has at least one associated user from S. We define the characteristic function $c(S)$, for any $S \subseteq U$, as the cost of the shortest path network $G^*(N(S))$ restricted to arcs (i,j) such that $i,j \in N(S) \cup \{s\}$, plus the cost of the stations in $N(S) \cup \{s\}$. Once again, we will abuse notation by using the same quantities $a(i,j)$ that represented travel times on arcs to represent costs incurred when establish or maintaining these links. The resulting cooperative game is called the *shortest path game*. We now check that the shortest path game meets (1) and (2). Checking (1) is immediate. To see that (2) holds, consider two disjoint coalitions S and T. Coalition S has a shortest path network $G^*(N(S))$ while T has shortest path network $G^*(N(T))$. Coalition $S \cup T$ can do no worse than the union $G^*(N(S)) \cup G^*(N(T))$ of disjoint trees, but it possible that at least one node $i \in N(S) \cup N(T)$ will acquire a shorter path to s in the expanded network, since $N(S) \cup N(T)$ contains additional arcs (i,j) such that $i \in N(S)$ and $j \in N(T)$.

We now mention the two most important solution concepts for cooperative games. For player set U we define a vector $x = \{x_i{:}i \in U\}$ to be a *cost allocation* (or just *allocation*) provided that $\Sigma_{i\in U} x_i = c(U)$. We define the *core*, denoted $core(U;c)$, of a cooperative game as the set of all allocations $x = \{x_i\}$ such that for all $S \subseteq U$, $\Sigma_{i\in S} x_i \leqslant c(S)$. The core of a cooperative game may be empty in general (Owen, 1982). Another solution concept is called the *Shapley value* (Shapley, 1953). The Shapley value φ is the unique allocation that satisfies symmetry, carrier, and additivity axioms (Shapley, 1953). It is given by the following formula for all $i \in U$:

$$\varphi_i = \Sigma_{T \subseteq U\setminus\{i\}} [t!(u-t-1)!/u!][c(T \cup (i)) - c(T)], \tag{3}$$

where t and u are the cardinalities of T and U, respectively. Expression (3) can be interpreted as allocating to each player its marginal contribution over all the coalitions that it can join.

3. The core of a shortest path game

Proposition 1 below shows that it is indeed possible to ensure that the cost of the shortest path network G^* can be allocated to the users in such a way that no coalition will pay more than it would have when building its own network to connect to terminus s.

Proposition 1. *The shortest path game has a nonempty core.*

Proof. Let G^* be a shortest path network on $G = (N \cup \{s\},A)$. G^* itself is the union of disjoint subtrees of G. (If G^* is not unique, the following will still hold for each of the different solutions.)

We first treat the arc costs for G^*. Consider the following two-stage allocation rule $R(a,x)$: first, for all $i \in N$, let a_i equal the cost of the arc incident to i on the unique s,i-path in G^*. Second, for users u_{im}, $m = 1,\ldots,w_i$ at node i, let $x(u_{im},a_i)$ be a cost allocation such that $\Sigma_m x(u_{im},a_i) = a_i$, for all $i \in N$. We claim that $R(a,x)$ is a core allocation rule. Now suppose that the claim that the allocation given by $R(a,x)$ is in the core is false, i.e., that there exists a coalition $S \subseteq U$ such that $c(S) < \Sigma_{i\in N(S)} a_i$. That is, S has by itself a less costly way to connect to s. Let P be the minimal collection of subtrees of G^* that contains S. Partition P into node sets $N(S)$ and $P\setminus N(S)$. Identify each node $i \in P$ with a particular arc: the incident arc on the unique s,i-path. Now let $a(P) = \Sigma_{i\in P} a_i = \Sigma_{i\in N(S)} a_i + \Sigma_{i\in P\setminus N(S)} a_i$. We observe that $a(P) > c(S) + \Sigma_{i\in P\setminus N(S)} a_i$, by hypothesis. But since $a(P)$ equals the optimal objective value for the shortest path problem on $P \cup \{s\}$, we have a contradiction.

Now we can account for the positive fixed costs f_i separately. We observe that a positive cost to build the station at s would add $f_s > 0$ to the characteristic function for all coalitions but not change the nature of the game. More generally, now assume station costs $f_i > 0$ at all nodes $i \neq s$. Since only players located at node i use station i (other users may pass through, but do not use, station i), the only coalitions S that account for f_i in their characteristic functions $c(S)$ are S such that $N(S) \cap \{i\} \neq \emptyset$; in particular each singleton $\{i\}$ includes f_i in $c(\{i\})$. Therefore, with the f_i added to the game, there exist core points that allocate each quantity f_i only among players located at i. This concludes the proof. □

Note that there exist core solutions that satisfy Condition 4 (non-subsidization), because in the proof the users at each node i pay the entire f_i for their own station and for arcs leading from s to i only. Later on we treat a more general case, in which users may travel to nodes other than s, and then we consider the problem of sharing the station costs as a separate cooperative game.

Finally, note that the node weights w_i played no role in the proof. This is because the argument hinges on ensuring that no coalition's shortest path network deviates from the shortest path network for the entire group. This argument stands regardless of the number of users at the different nodes since the arc costs that coalitions must bear are independent of how many users are located at each node.

One property revealed in the proof of Proposition 1 is that it is constructive, i.e., offers a specific core allocation rule. In general, though, core points are not unique, and the unique allocation given by the Shapley value provides a competing solution concept. Certain games, where the characteristic function is submodular in a cost setting, are called *convex*, and for these games, the Shapley value is always a core solution (Shapley, 1971). This property eliminates the conflict between the two solution concepts. Below, we show that in general the shortest path game is not convex.

Proposition 2. *Shortest path games are not convex in general.*

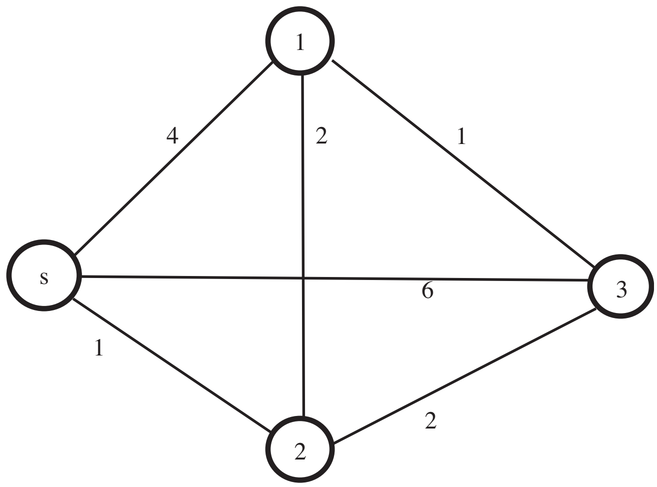

Fig. 1. A network for which the shortest path game is not convex.

Proof. Consider the network G in Fig. 1, where, for simplicity, we assign exactly one user per node, and let the fixed costs $f_i = 0$ for all $i \in N$.

The characteristic function of the game is given by $c(\emptyset) = 0$, $c(\{1\}) = 4$, $c(\{2\}) = 1$, $c(\{3\}) = 6$, $c(\{1,2\}) = 3$, $c(\{1,3\}) = 5$, $c(\{2,3\}) = 3$, and $c(N) = 5$. The core of this game is the set of points satisfying $x_1 \leqslant 4$, $x_2 \leqslant 1$, $x_3 \leqslant 6$, $x_1 + x_2 \leqslant 3$, $x_1 + x_3 \leqslant 5$, $x_2 + x_3 \leqslant 3$, and $x_1 + x_2 + x_3 = 5$. These inequalities can be reduced to the following set:

$$x_2 \leqslant 1,\ x_2 + x_3 \leqslant 3, \quad \text{and} \quad 4 \leqslant x_1 + x_3 \leqslant 5. \tag{4}$$

The Shapley value, however, is given by $\varphi[c] = (13/6, -2/6, 19/6)$. Player 2 is allocated a negative cost since, in two of the six permutations involved in determining φ, Player 2's joining a coalition results in shorter paths and associated cost decreases. In this case the Shapley value fails to satisfy the conditions in (4) and thus is not a core point. Therefore, since the Shapley value is the center of gravity of the extreme points of the core of a convex game (Shapley, 1971), shortest path games in general are not convex. □

While shortest path games are in general not convex, it turns out that convexity holds for an important subclass of these games, as demonstrated in the next section.

4. Shortest path games for tree networks

In tree networks, because each (s,i)-path is unique, we need to modify the definition of the characteristic function in defining the shortest path game. In this case, there is no longer an underlying complete graph but rather, the initial network is based on a particular tree $T = (N \cup \{s\}, A)$. Now we must define the characteristic function $c(S)$ to be equal to the minimal cost in T to connect all nodes in $N(S)$ to s, for all $S \subseteq U$. (It is no longer true that the nodes have alternative ways to connect to s other than the unique paths through T.) This cost function is well-defined: every node $i \in N$ has a single path to s and no other arcs will be included, since these costs are nonnegative. In all other respects, however, the shortest path game for tree networks is defined as in the general case.

Note that the structure of the characteristic function for shortest path games for tree networks violates Condition 2, since now there exist coalitions S whose shortest path networks go through stations not in $N(S)$. We point out that in practical applications this violation of Condition 2 is not a serious drawback, since in many cases an existing transportation network already forms a tree network and an upgrade to the network does not involve alternative paths from existing stations to s.

An important property of convex games (Whitney, 1935; Shapley, 1971) is that for $S_1 \subseteq S_2$ and $i \notin S_2$, we have

$$c(S_2 \cup \{i\}) - c(S_2) \leqslant c(S_1 \cup \{i\}) - c(S_1). \tag{5}$$

It turns out that this "snowballing" property (equivalent to submodularity), where the marginal cost for a node to join with a coalition does not increase as the coalition adds new members, is an important feature for shortest path games on trees, as seen in the following result.

Theorem 1. *Shortest path games defined on tree networks are convex.*

Proof. Given is a tree network $T = (N \cup \{s\}, A)$ with source s. We now show that (5) holds for any T with $S_1 \subseteq S_2 \subseteq N$, and $i \notin S_2$. For any S_i, let $T(S_i)$ be the shortest path tree connecting all users in S_i to s. Now consider $T(S_2 \cup \{i\})$. To ease notation we let i be the node associated with player i. If the shortest (s,i)-path is disjoint from $T(S_2)$, then it must be disjoint from $T(S_1)$ as well since $S_1 \subseteq S_2$. In this case equality in (5) would hold.

Now suppose that the shortest path in $T(S_2 \cup \{i\})$ is found through adding an i,j_2-path such that node j_2 is the first node on $T(S_2)$. Let $P(i,j_2)$ be this i,j_2-path and let $a(P_{ij2})$ be its cost. Now consider $T(S_1)$ with $S_1 \subseteq S_2$. Similarly construct $T(S_1)$ and $T(S_1 \cup \{i\})$, with path $P(i,j_1)$, where j_1 is the first node on $T(S_1)$, and cost $a(P_{ij1})$. Suppose that there exists i, S_1, and S_2 such that the shortest path game on the tree network is not convex, i.e., that (5) does not hold. Then

$$c(S_2 \cup \{i\}) - c(S_2) = a(P_{ij2}) > c(S_1 \cup \{i\}) - c(S_1) = a(P_{ij1}) \tag{6}$$

would hold. But if (6) is true, then $P(i,j_2)$ would not be minimal, because $T(S_2)$ contains $T(S_1)$. This, however, is a contradiction. The argument holds for any set $\{f_i\}$ of fixed costs, since they cancel out on each side of expression (5). Therefore shortest path games on tree networks are convex. □

Theorem 1 is important for practical applications like rail construction for two reasons. One is that when rail systems are updated, the existing system is typically a tree network and while the network design problem disappears, the problem of fairly allocating the costs of the system upgrades still needs to be solved. Another reason that the tree case is important is, even when a new rail system is designed from scratch, the shortest path solution will be a tree network and it makes sense for the participating communities to consider the tree network, and not a hypothetical complete graph, as the baseline from which to plan cost allocations.

Since convexity guarantees that the Shapley value for shortest path games on tree networks will lie in the core of such games, it arguably provides an uncontested solution. Normally, though, for games with more than a handful of players, the Shapley value becomes difficult to compute. We now show that the special structure of tree networks gives rise to an efficient formula for the Shapley value for shortest path games on these networks. Theorem 2 below generalizes the result in Littlechild and Owen (1973) on the Shapley value for the Airport Game.

Before we can state Theorem 2, we need some notation. Let $T = (N \cup \{s\}, A)$ be a tree network. Let i represent any of the w_i users (all identical) located at node i, for all $i \in N$. For all $i \in N$, let p_{si} be the unique s,i-path in T. For each $i \in N$, on p_{si} starting at s and, stopping at i, label the arcs as i_k, $k = 1, \ldots, m_i$ while the arc costs are $a(i_k), k = 1, \ldots, m_i$. Each arc i_k, $k = 1, \ldots, m_i$, on p_{si} is a member of the shortest s,j-paths over all $j \in N$, for a certain number $n(i_k)$ of nodes and corresponding number $n'(i_k)$ of users. That is, there are exactly $n(i_k)$ nodes $j \in N$ whose s,j-paths contain arc i_k. Further, $n'(i_k) = \sum_{j \in N} w_j$ for j such that the s,j-path contains i_k. In Theorem 2, we provide the Shapley value for arc costs for all users U in the network. The formula in Theorem 2 generalizes the Shapley value computation found in Littlechild and Owen's (1973) Airport game.

Theorem 2. *The Shapley value* $\varphi = (\varphi_i)$ *for arc costs in shortest path games defined on tree networks is given by*

$$\varphi_i = \sum_{k=1}^{m_i} a(i_k)/n'(i_k), \text{ for all } i \in U.$$

Proof. See Appendix A. □

We end this section with an example of a shortest path game and its solutions for a tree network. In Fig. 2 below, the fixed (station) costs are indicated next to each node, while arc costs are indicated along each arc. For simplicity, we assume one user at each node.

Example 1. A shortest path game on a tree network.

Since the network in Fig. 2 is a tree network, each unique s,i-path is the shortest path from s to i. The total arc cost is 16, so any feasible allocation x must satisfy $\Sigma_{i=1,\ldots,5} x_i = 16$ for the game with player set $U = 1,\ldots,5$ and arc cost function a. There are $2^5 = 32$ coalitions, and formally, the characteristic function would need to be defined for all of them, which is straightforward. The allocation of the station costs are additive and can be treated separately.

Coalition $A = \{1,2,3\}$ comprises one subtree and coalition $B = \{4,5\}$ comprises another. From this structure we obtain the following constraints: $x_1 + x_2 + x_3 \leqslant 9$ and $x_4 + x_5 \leqslant 7$. Together with $\Sigma_{i=1,\ldots,5} x_i = 16$, we find that neither A nor B can subsidize the other. The following become the conditions defining the core of the game:

$$\text{core}(U;c) = \{x_i, i = 1,\ldots,5 : \ x_1 + x_2 \leqslant 5, \ x_2 \geqslant 2, \ x_3 \geqslant 4, \ x_1 + x_2 + x_3 = 9; \ x_4 \leqslant 2, \ x_4 + x_5 = 7\}.$$

Using Theorem 2 for the arc cost component, we obtain the Shapley value $\varphi = (1,3,5,1,6) \in$ core $(U;c)$. In the Shapley value allocation, Player 3, for example, pays all four units for link $(1,3)$ while link $(s,1)$ is shared equally among players 1, 2, and 3. Thus Player 3's total arc cost share is 4 + 1 = 5, and the other players' allocations are computed similarly.

The station costs are added on as follows. The Shapley value will allocate the station cost f_s equally among all users. Thus, all five

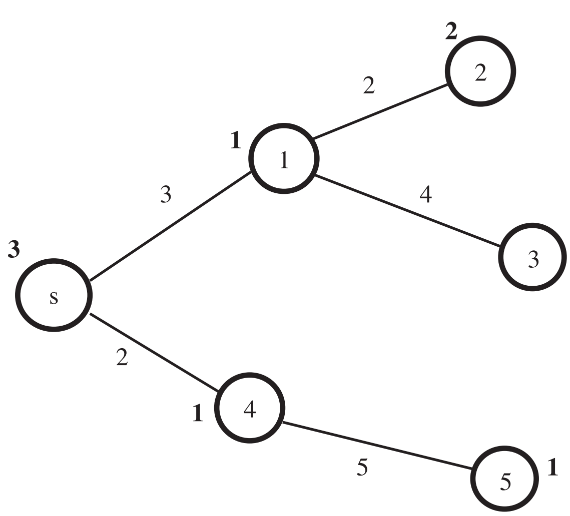

Fig. 2. A shortest path game on a tree network.

users are allocated a cost of 3/5 for f_s. Then, since user i is the sole user of station i, for all 1 = 1,...,5, the Shapley value must allocate the entire cost of station i to user i. Finally, the Shapley value is found to be

$$(1+(3/5)+1, 3+(3/5)+2, 5+(3/5)+1, 1+(3/5)+1, 6+(3/5)+1) = \left(2\frac{3}{5}, 5\frac{3}{5}, 6\frac{3}{5}, 2\frac{3}{5}, 7\frac{3}{5}\right).$$

5. The Shapley value for the general case

Up until now our application was to rail systems where outlying users travel between their "home" location and s. Now we generalize the shortest path game by allowing some number of users $w_{ij} \geqslant 0$ located at node i to travel between nodes i and j, where j may not equal s. We call this game the *generalized shortest path game*. We restrict our analysis to the case where the underlying network is a tree T, as in the previous section, and compute the Shapley value.

The importance of the general case is that rail users normally may travel to destinations other than s, and one needs to properly allocate link costs and station costs in this more comprehensive environment. It is useful to conceptualize the cost allocation problem as the composition of shortest path games where any node $j \in N \cup \{s\}$ can be a terminus. In addition, because the users in general are visiting stations other than their home station and the terminus, the cost allocation problem for the fixed costs f_i is no longer trivial, as it was in the previous case. Further, we note that since the link costs and the station costs are separate, and because the Shapley value is additive (Shapley, 1953), the Shapley value for the combined game involving the sum of these costs is equal to the sum of the Shapley values on the separate link-cost and station-cost games.

To generalize the tree case, let node s be denoted as 0, and introduce w_0 users located at 0. Let $U = \cup U_i, i = 0,1,\ldots,n$ be the player set, where tree $T = (N',A)$ with $N' = N \cup \{0\}$. Let $W = \Sigma_{i \in N'} w_i$. Since users can now travel between pairs of nodes (i,j) for all $i,j \in N'$, we partition each set U_i according to which terminus the users visit, and let $w_{ij} \geqslant 0$ represent the number of users who travel from node i to terminus $j, i \neq j$. We require that all users travel somewhere, i.e., $\Sigma_{j \in N'} w_{ij} = w_i$ for all $i \in N'$. For any coalition S of users, let T_S represent the set of all termini in T that are destinations for users in S. Let $T_S^* \subseteq T$ be the minimal subgraph necessary to connect all users in S to the destination set T_S. Now define the characteristic function $c(S)$ as the total cost of all stations and arcs in T_S^*. Let Φ_i represent the Shapley value for user i in the generalized shortest path game on a tree network. Since the Shapley value is additive, and the arc and station costs are separate, we decompose Φ_i as follows: $\Phi_i = \varphi_{i(A)} + \varphi_{i(F)}$, where $\varphi_{i(A)}$ denotes the Shapley value for allocating the arc costs and $\varphi_{i(F)}$ denotes the Shapley value for allocating the fixed costs at the stations in the two respective shortest path games.

Now observe that we can also decompose $(U;c)$ into separate games in which each node, one at a time, is considered as the terminus. In each constituent game, however, only a subset of the players in U travel to any one source j; the proportion of such players $p_j = \Sigma_{i \in N'} w_{ij}/W$, for each $j \in N'$. All $p_j \geqslant 0$ and $\Sigma_{j \in N'} p_j = 1$. Now consider the problem of allocating the arc costs. To develop a Shapley value allocation for this game, we capitalize on the linearity of the Shapley value (Myerson, 1991) as follows.

The linearity condition states that, loosely speaking, for two or more games, the Shapley value of a convex combination of the games is equal to the convex combination of the Shapley values of the individual games. More precisely (Myerson, 1991), linearity states that for a set of players, for two games v and w on this player set, and $0 \leqslant \delta \leqslant 1$, the Shapley value for player i is the following:

$$\varphi_i(\delta v + (1-\delta)w) = \delta\varphi_i(v) + (1-\delta)\varphi_i(w). \quad (7)$$

We define $\varphi_{i(A)}[j]$, for all $j \in N'$, to be the Shapley value (using arc costs) for user i when node j is treated as the single terminus, using the formula in Theorem 2.

Lemma 2. *The Shapley value, using arc costs only, for user i in the generalized shortest path game on a tree network is* $\varphi_{i(A)} = \Sigma_{j\in N'} p_j \varphi_{i(A)}[j]$.

Proof. We maintain the same player set for each possible terminus j; although, for any j, there may exist users i that do not travel to j, the Shapley value for such i is the null contribution $\varphi_{i(A)}[j] = 0$, since these players add no value. Then Lemma 2 follows directly from (7). □

Next we consider the problem of allocating the costs f_i of the stations in T. Once again we decompose the overall cost allocation problem, where any $j \in N'$ is a terminus, into a set of shortest path games, each with a single terminus. Consider a shortest path game on tree T with terminus j. For each $i \in N'$, there are w_{ij} users that travel between i and j. Consider the characteristic function $c(S)$ for any nonempty coalition $S \subseteq U$ of users. Cost f_j is included in the characteristic function for all $S \subseteq U$ (even singletons), while for all other nodes $i, i \neq j$, cost f_i is included in $c(S)$ only for those coalitions that include users located at i. Since f_j is common to $c(S)$ for all coalitions S, and since the users are symmetric in this regard, it is easy to show that the Shapley value will divide f_j equally among all users who visit j, which total $\Sigma_{i\in N'} w_{ij}$. Moreover, for nodes $i, i \neq j$, since the only coalitions S that bear cost f_i are such that $S \cap U_i \neq \emptyset$, the Shapley value will divide f_i equally among those w_{ij} users in the particular tree game with terminus j. These two observations yield Lemma 3, while Lemma 4 follows immediately from linearity.

Lemma 3. *The Shapley value, using station costs only, for user i in the shortest path game on a tree network with a single terminus j is*

$\varphi_{i(F)}[j] = f_j/(\Sigma_{i\in N'} w_{ij}) + f_i/w_{ij}$ for each user i that travels to j.

Lemma 4. *The Shapley value, using station costs only, for user i in the generalized shortest path game on a tree network is*

$\varphi_{i(F)} = \Sigma_{j\in N'} p_j \varphi_{i(F)}[j]$.

Finally, using the property of additivity together with the results of Lemmas 2–4, *we obtain*

Theorem 3. *The Shapley value for user i in a generalized shortest path game on a tree network is* $\Phi_i = \varphi_{i(A)} + \varphi_{i(F)}$.

Theorem 3 *is important for rail applications because it means that, if the metropolitan area M tracked the movements of its passengers, the Shapley value (which is easy to compute) would provide a fair way for the communities to charge users for their trips, based on the combination of the link costs and the station costs incurred.*

6. Dynamic shortest path games

In the network design problem that motivates this paper, it is natural to consider the expansion of the transport system over time. As new communities connect to the rail system and new nodes are added in its network representation, it is sensible to ask whether the expanded network will accommodate the newcomers effectively, and whether the additional costs involved can be accounted for in an equitable fashion. There are numerous current examples of such network expansions; one such major project is the extension of the Washington, D.C. metro system to serve Dulles Airport (The Washington Post, 2011). For such capital projects that may be extended at future dates, it is important to determine under what circumstances the addition of new players would preserve the cost allocations of current players. Examination of solution properties under the addition of new players to cooperative games was introduced by Rosenthal (1990a,b) and Sprumont (1990).

In the present setting, we want to ask the following two questions: one, whether it is possible that when new players are added to a shortest path network, the network of shortest paths is preserved for all players; and two, whether solutions like the core and Shapley value are likewise preserved when new players are added to the network. It is easy to see that for shortest path games in general, the addition of new players may destroy the optimality of the existing paths.

In Fig. 3, Player 1 (at node 1) is the first to initiate a rail connection to node s. The only arc available to Player 1 is $(s,1)$ with a cost of 4 units. When Player 2 joins, 2's direct link to s provides the shortest s,2-path with cost 1. However, when node 2 and arc $(s,2)$ are present, node 1 would have a shorter s,1-path: $\{(s,2),(2,1)\}$, with a cost of 3. Therefore, in general, when additional players join a network, paths that were previously optimal may no longer be best. The implication for real rail systems is either that current commuters would not benefit from the addition of new communities and stations, or that as new communities and stations are added, the network would constantly be in a state of reorganization (and would accrue unnecessary cost) as the shortest paths from existing communities are updated. In what follows below, we will assume that such network reorganization will not take place as new nodes are added.

The principal question is to discover under what conditions in a network G will the shortest paths from all $i \in N$ to s remain permanently optimal as new nodes are added to N. In Theorem 4 below we state these conditions.

To state Theorem 4 we need some preliminaries. First consider any nodes $i,j,k \in N$ and the circuit of arcs (i,j), (j,k), and $(k,i) \in A$ with respective distances $a(i,j)$, $a(j,k)$, and $a(k,i)$. The *triangle inequality* for undirected graphs states that the sum of the distances of any two arcs of this circuit is no less than the distance

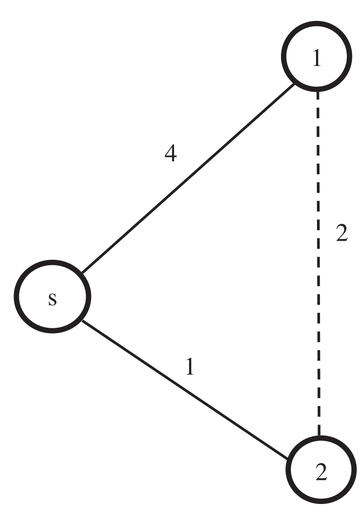

Fig. 3. A network for which the addition of a new player destroys existing optimality. *Note*: solid lines indicate permanent arcs, while the dashed line indicates an underlying arc.

of the third arc. (This inequality is strict if the sum of the distances of any two arcs is greater than the distance of the third arc.)

Second, we develop a notion of *stages* in G. For any two nodes, let a *minimal path* be a path that connects them with the smallest possible number of arcs. Let the set of all nodes that have a minimal path of one arc linking them to source s be called *Stage 1* nodes; all nodes with a minimal path to s consisting of two arcs are called *Stage 2* nodes, and so on. Let m be the maximum number of stages in G. To grow the network dynamically, starting with s, we say that the nodes in N are *added in stages* if no node in Stage $k+1$ can join the network before any node in Stage k, for $k = 1,\ldots,m-1$. Finally, we define a *dynamic shortest path game* on $G = (N \cup \{s\},A)$ to be a sequence of shortest path games G_π where the node set begins with $\{s\}$ and nodes $i \in N$ are added one at a time according to a permutation $\pi:N \to N$.

Theorem 4. *Given is a dynamic shortest path game G_π on G. When node i is added to the network, the shortest s,i-paths for all $i \in A$ remain optimal if and only if*

(1) *the triangle inequality strictly holds for G, and*
(2) *π is such that the nodes are added in stages.*

Proof. See Appendix A. □

Theorem 4 shows us under what conditions the player set can be expanded and still ensure that the shortest s,i-paths, for all $i \in N$, remain permanent once they are discovered. This result addresses an important network design question but not the cost allocation issue. Rosenthal (1990b) shows that the Shapley value is monotonic for convex games in general. That is, when the player set is expanded for such games, no player's Shapley value allocation is inferior to that in the smaller game. The implication of this result for shortest path games is as follows. Since shortest path games in general have a nonempty core but are not convex, one cannot guarantee in general that the Shapley value allocation in a larger game will be no worse than that in a smaller game. However, as we have seen in Theorem 1, shortest path games defined on tree networks are convex. Therefore, as new players are added to the game, no one is worse off when the Shapley value is used as an allocation scheme. And, since the Shapley value lies in the core for these games, the corollary is that there exist core solutions that are monotonic as well.

Regarding the triangle inequality as applied to actual rail construction, it would appear that the triangle inequality holds for some situations, for example, when the rail links are built in relatively flat terrain, with fairly uniform land costs. However, in other instances, topographical features or the ability (or inability) of local governments to claim land rights through eminent domain will lead to certain links being either extraordinarily swift or slow, whereby the triangle inequality will be violated. It is difficult to provide a more general characterization.

7. Extensions and conclusions

Thus far we have not touched on the budgetary nature of rail projects such as we have discussed. It would be quite possible to extend our model by proposing an overall budget of B_N that the community can raise. Briefly, if the quantity B_N were to become binding, it is easy to scale down the project accordingly. It would be straightforward, for example, to develop actual costs for the various links, and then to list the k shortest paths in descending order (Eppstein, 1999; Sedeño-Noda and González-Martín, 2010). In employing such a list, one may find the shortest tree network (that in addition reaches the maximal number of users, for example) within the budget constraint.

The initial contribution of this paper was to define shortest path games in the context of designing commuter rail lines for metropolitan areas. Such design will focus not only on building a network that minimizes travel times but also on allocating capital costs in a fair manner. For such network design and cost allocation problems, we first proposed a set of conditions that serve to provide rapid travel for network users while seeking stable and fair cost allocations. We showed that shortest path games have nonempty cores. This result guarantees that different communities can join together to design a shortest path network in such a way that no subset of communities can do better on its own. Core solutions also satisfy the given conditions.

While in general, the Shapley value allocation is not an element of the core – meaning that these two solution concepts can conflict with one another – we found that the Shapley value indeed lies in the core for shortest path games defined on tree networks. Furthermore, we provided a simple and efficient formula for the Shapley value in these instances and extended it to the more general case where users may travel to stations other than at the source. These results are important because many rail networks will have already been laid out as tree networks from the outset. Finally, recognizing that such commuter rail networks will grow over time, we developed the notion of dynamic shortest path games and characterized which networks share the property that their growth will not only preserve the shortest path properties for existing users but will ensure that cost allocations for those users will not increase as new communities are added. The triangle inequality and the staged node condition would seem to hold for some real-world rail planning scenarios but not others.

One limitation of this paper is that our emphasis on travel time minimization has led us to exclude actual construction costs from the model. An important consideration for future research would be to incorporate construction costs into a similar type of model. Another direction for further work is to consider cost allocation among multiple users of the network, for example if the links are shared by both passenger and commercial rail cars. Finally, it is a challenging problem to study networks with cycles, and therefore multiple source–destination paths, and allocate the costs of these multiple paths to their users.

Acknowledgments

I am very grateful to two referees for their remarkably thorough, detailed and insightful comments, which have improved this paper immeasurably.

Appendix A

Proof of Theorem 2. As before, let $T = (N \cup \{s\},A)$ be a tree network. User i is a generic user located at node i. For $i \in N$, let p_{si} be the unique s,i-path in T. For each $i \in N$, on p_{si} we label the arcs, starting at s and stopping at i, as $i_k, k = 1,\ldots,m_i$, while the arc costs are $a(i_k)$, $k = 1,\ldots,m_i$. Each arc i_k on p_{si} is a member of the shortest s,j-paths over all $j \in N$ for a certain number $n(i_k)$ of nodes and a corresponding number $n'(i_k)$ of users, where $n'(i_k) = \Sigma_{j\in N} w_j$ such that arc i_k lies on the shortest s,j-path in T. Theorem 2 states that the Shapley value allocation for arc costs for user i, for all $i \in U$, is $\varphi_i = \sum_{k=1}^{m_i} a(i_k)/n'(i_k)$.

To prove Theorem 2, first consider arc i_1. Arc i_1 is a member of $n(i_1)$ unique paths from nodes $j \in N$ to s. There are $n'(i_1)$ users and therefore $n'(i_1)!$ permutations of these different users in the Shapley value computation. The key insight is that in each such permutation, the user i that appears first will pay for the entire cost of its path p_{si} that includes arc i_1. This happens because at that point in the computation, the network transitions from an empty one to p_{si}. And, since the cost $a(i_1)$ is paid entirely by the user appearing first in any permutation, then all other users in that permutation contribute zero to it.

Any individual user i appears first in $(n'(i_1)-1)!$ out of the $n'(i_1)!$ orderings. Therefore user i's cost share for arc i_1 is $a(i_1)[(n'(i_1)-1)!/n'(i_1)!] = a(i_1)/n'(i_1)$, since we consider all w_i users at any single node i as equivalent, and their symmetry, as considered by the Shapley value, will allocate the same cost to them all.

Now consider arc i_2. This arc is a member of $n(i_2)$ unique (s,j)-paths. Just as with i_1, there are $n'(i_2)!$ permutations of the different players that connect to s via i_2. But any individual user i appears first in $(n'(i_2)-1)!$ out of the $n'(i_2)!$ permutations and when doing so pays the entire cost $a(i_2)$, while in the remaining permutations user u_i pays nothing. So for i_2, user i's cost contribution is $a(i_2)/n'(i_2)$ and indeed, the same argument holds for all arcs i_k on p_{si}. User i will never contribute toward the cost of any arc not on p_{si}. Therefore user i's cost share as given by the Shapley value is summed over all arcs on p_{si}, yielding $\varphi_i = \sum_{k=1}^{m_i} a(i_k)/n'(i_k)$, for all $i \in U$. □

Proof of Theorem 4. Theorem 4 states that in a dynamic shortest path game G_π, that the addition of new node i to the network will preserve optimality of all shortest paths for all nodes added before i, for all $i \in N$ if and only if

(1.) *The triangle inequality strictly holds for G and*
(2.) *π is such that the nodes are added in stages.*

We prove the necessity of Theorem 4 by induction. The base case consists of the node set $\{s\}$ and is trivially true. Now suppose conditions (1.) and (2.) hold for a subset $N' \subseteq N$ of cardinality n', and consider adding an $(n'+1)$ st node, i. Without loss of generality, select i such that i has the shortest s,i-path out of all the nodes in $N \backslash N'$. Just before i is added, suppose the previously added node $(i-1)$ is a member of stage k; that is, the shortest s, $(i-1)$-path consists of k arcs. Since the triangle inequality holds, we know that no shortest s,i-path exists of length $(k-1)$ or fewer arcs. (For example, if a direct link from s to i existed, $a(s,i)$ would have been optimal by the triangle inequality and i would have been selected as a Stage 1 node; the same argument is true for stages $2,\ldots,(k-1)$.) Therefore i must be a Stage k or Stage $(k+1)$ node. Suppose i is a Stage k node. The only feasible link (i,j) to create a shortest s,i-path must be to a Stage $(k-1)$ node j, because if i's shortest path to s has fewer than k links, we have a contradiction since i would have been selected already.

When node i is added to N', we must determine whether any other node i' in N' will thereby obtain a path shorter than its current s,i'-path. The answer is no, since such a circumstance would mean that node i' is in fact a stage $(k+1)$ node, violating the hypothesis that i' was in Stage k or less.

Finally, if node i is a Stage $(k+1)$ node, a similar argument will apply. Thus we see that the addition of any node $i \in N$ establishes a shortest s,i-path that preserves the shortest paths for all nodes added up to that point, proving necessity of (1.) and (2.).

To prove sufficiency of the two conditions, first consider the triangle inequality. Suppose that the triangle inequality holds but that the nodes are not added in stages. See Fig. 4.

Suppose nodes j and k have already been added to the dynamic shortest path network and now node i is to be added. Suppose further (and without loss of generality) that $a(j,i) > a(k,i)$ but that $a(p_{sj}) + a(j,i) < a(p_{sk}) + a(k,i)$. In this instance, we select arc (cost) $a(j,i)$ since that yields the shorter s,i-path. But now suppose that node k were added to the network before node j. If j is not present when node i is added, the shortest path at that point in time is via node k. However, when node j is subsequently added to the

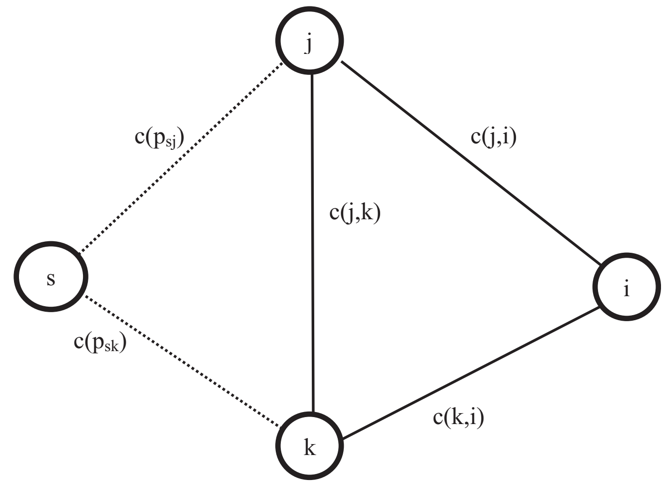

Fig. 4. A dynamic shortest path game. *Note*: $c(p_s)$ is the cost of the unique shortest path from s to another node (dotted lines), while $c(j,i)$ is the cost of arc (j,i), etc. (solid lines).

network, the s,i-path through k is no longer optimal. Therefore, for the shortest paths to remain permanent, the nodes must be added in stages.

Now suppose that nodes are added in stages but that the triangle inequality does not necessarily hold. The previous discussion involving Fig. 3 shows that a network can have two nodes (nodes 1 and 2 in Fig. 3) that are in the same stage, but that if the triangle inequality does not hold, then depending on the order in which the nodes are added within the same stage, permanence of the shortest paths may be violated. This demonstrates the sufficiency of the pair of conditions, and concludes the proof of Theorem 4. □

References

Bachrach, Y., Rosenschein, J.S., 2009. Power in threshold network flow games. Autonomous Agents and Multi-Agent Systems 18, 106–132.

Bergantiños, G., Vidal-Puga, J., 2010. Realizing fair outcomes in minimum cost spanning tree problems through non-cooperative mechanisms. European Journal of Operational Research 201, 811–820.

Bird, C.G., 1976. On cost allocation for a spanning tree: a game theoretic approach. Networks 6, 335–350.

Dijkstra, E.W., 1959. A note on two problems in connexion with graphs. Numerische Mathematik 1, 269–271.

Eppstein, D., 1999. Finding the K shortest paths. SIAM Journal on Computing 28, 653–674.

Fragnelli, V., García-Jurado, I., Méndez-Naya, L., 2000. On shortest path games. Mathematical Methods of Operations Research 52, 251–264.

Grahn, S., 2001. Core and Bargaining Set of Shortest Path Games. Mimeo.

Granot, D., Huberman, G., 1981. Minimum cost spanning tree games. Mathematical Programming 21, 1–18.

Laporte, G., Mesa, J.A., Perea, F., 2010. A game theoretic framework for the robust railway transit network design problem. Transportation Research Part B: Methodological 44, 447–459.

Laporte, G., Mesa, J.A., Ortega, F.A., Perea, F., 2011. Planning rapid transit networks. Socio-Economic Planning Sciences 45, 95–104.

Littlechild, S.C., Owen, G., 1973. A simple expression for the Shapley value in a special case. Management Science 20, 370–372.

Myerson, R.B., 1991. Game Theory. Harvard University Press, Cambridge, MA.

Nebel, F., 2010. Shortest Path Games: Computational Complexity of Solution Concepts. MSc Thesis, University of Amsterdam.

Owen, G., 1982. Game Theory, second ed. Academic Press, New York.

Rosenthal, E.C., 1990a. Monotonicity of solutions in certain dynamic cooperative games. Economics Letters 34, 221–226.

Rosenthal, E.C., 1990b. Monotonicity of the core and value in dynamic cooperative games. International Journal of Game Theory 19, 45–57.

Schöbel, A., Schwarze, S., 2006. A game-theoretic approach to line planning. In: Jacob, R., Müller-Hannemann, M. (Eds.), ATMOS 2006 6th Workshop on Algorithmic Methods and Models for Optimization of Railways. Internationales Begegnungs-und Forschungszentrum für Informatik (IBFI), Schlass, Dagstuhl, Germany.

Sedeño-Noda, A., González-Martín, C., 2010. On the k shortest path trees problem. European Journal of Operational Research 202, 628–635.
Shapley, L.S., 1953. A value for n-person games. In: Kuhn, H.W., Tucker, A.W. (Eds.), Contributions to the Theory of Games, vol. 2. Princeton University Press, Princeton, NJ.
Shapley, L.S., 1971. Cores of convex games. International Journal of Game Theory 1, 11–26.
Sprumont, Y., 1990. Population monotonic allocation schemes for cooperative games with transferable utility. Games and Economic Behavior 2, 378–394.
Voorneveld, M., Grahn, S., 2002. Cost allocation in shortest path games. Mathematical Methods of Operations Research 56, 323–340.
The Washington Post, 2011. How Metro Expansion Might Make Sense. <http://www.washingtonpost.com/opinions/how-metro-expansion-might-make-sense/2011/06/23/AGhRXdjH_story.html>.
Whitney, H., 1935. On the abstract properties of linear dependence. American Journal of Mathematics 57, 509–533.