Reconciling the Overlay and Underlay Tussle

Jin Xiao and Raouf Boutaba, *Fellow, IEEE*

***Abstract*—In the presence of multiple overlays and underlays, the emerging global network behavior is the result of interactions of self-serving overlay routing decisions and independent underlay management actions. It is crucial for network operators, service, and content providers to have a good grasp of the underlying principles in order to better design and manage current and future networks and services. In this paper, we describe special game scenarios wherein the interaction of noncooperative overlays and underlays in multidomain networks can result in an operable global configuration in linear time and the overall convergence is polynomial in the unweighed case. For weighted games, we find that weighted Shapley potential can achieve linear time convergence to an operable state. Furthermore, we analyze the interaction of overlays and underlays as a two-stage congestion game and recommend simple operational guidelines to ensure global stability. We further explore the use of Shapley value as an enabler of mutual cooperation in an otherwise competitive environment. Our simulation results confirm our findings and demonstrate its effectiveness in general networks.**

***Index Terms*—Congestion game, network stability, Shapley value.**

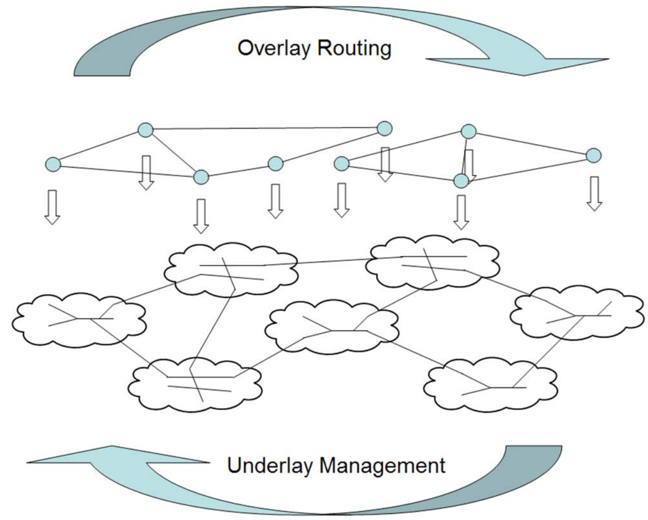

Fig. 1. Overlay and underlay interactions in multidomain networks.

I. Introduction

OVER the past decade, we have seen a gradual evolution of the Internet from data-centric to content/service-centric. Amidst this evolution is the rapid proliferation of application-level networking technology, in particular myriads of overlay networks that span multiple network domains. Their growth is expected to continue in the coming years. Whereas traditional network management and route optimization were conducted exclusively at the underlays to ensure global network performance and stability, the presence of overlays deviates traffic from their underlay routes to achieve their application-level requirements (Fig. 1). Thus, the global network behavior is the result of interactions among the overlay routing decisions and the underlay management actions. What emerges is the combined effects of the self-serving operations of the overlay interactions, the independent operations of the underlay domains, and the lack of communication between the overlay (end-to-end view) and the underlay (local connection and routing view), all of which have been subjects of separate research. Presently, it is difficult to assess the stability and optimality of the networks, especially in multidomain environments where many overlays and underlay providers influence the network operations based on their own (sometimes divergent) objectives. Yet, it is crucial for network operators, service providers, and network architects to have a good grasp over the underlying principles in order to better design and manage current and future networks and services. To date, a number of studies on overlay interactions have been conducted. A common approach is to model the overlay interactions as a game and establish the existence of pure Nash equilibrium (NE) (sometimes unique) in such a game (e.g., [1] and [2]). Congestion games have been frequently explored due to the observation that when a congestion game exhibits a global potential, it guarantees the existence of a pure Nash equilibrium, and in some cases convergence is the natural result of selfish game plays. In the past, analysis has been conducted to establish the existence of pure Nash equilibrium in congestion games (e.g. [3]–[5]). Thus far, we know that there is no guarantee that a pure Nash equilibrium exists in all congestion games [3], and when it does, the convergence of asymmetric games with polynomial cost function can be exponential [4]. Therefore, systemwise stability and convergence is not readily obtainable. On the subject of underlay and overlay interactions, existing works show that such interactions can easily lead to instability in the networks [6], and the self-serving behavior results in a performance tussle [7], [8] between the overlay layer and the underlay layer.

In this paper, we describe special-case scenarios wherein the interaction of noncooperative overlays and underlays in multidomain networks can result in stable global configurations

Manuscript received September 09, 2011; revised October 10, 2012 and May 13, 2013; accepted August 10, 2013; approved by IEEE/ACM Transactions on Networking Editor V. Misra. Date of publication September 30, 2013; date of current version October 13, 2014.

J. Xiao is with the IBM T. J. Watson Research Center, Yorktown Heights, NY 10598 USA (e-mail: jinoaix@us.ibm.com).

R. Boutaba is with the School of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: rboutaba@uwaterloo.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNET.2013.2281276

with convergence in polynomial and linear time. We investigate the degree of optimality in such cases and outline simple guidelines for network design and operations. In particular, we focus on two important operational characteristics: operable states and convergence. We consider operable state as a state of the system wherein there exists no congestive links. It is operable in the sense that a noncongestive network is well behaved and offers bounded performance to the supported services. If a network finds itself in a congestive state, it is important to quickly return to a stable and noncongestive state, thus fast convergence is also necessary.

First, we construct an unweighed congestion avoidance game model of overlay interactions. Through game transformation, we bound the convergence of the game to NE in polynomial time, and to an operable state in linear time. We also derive the resulting price of anarchy (PoA). We then extend our investigation to weighted congestion games and show that such games can have operable pure NE under Shapley utility function, and its convergence to an operable state can be obtained also in linear time. The applicability of Shapley value in noncooperative congestion game allows us to reason about the relation between competition and cooperation in overlays.

We place strong emphasis on operable states in this paper because of their importance to both network operators and service/conent providers, and to resolving the underlay–overlay tussle. Our findings suggest that congestion avoidance brings about significant benefits to noncooperative and coalition-based overlay networks in terms of stability and fast convergence. Our analysis on Shapley-based weighted game also affords us a better understanding of a complex overlay environment, where competition and cooperative coalitions coexist, in particular about the group dynamics affecting stability, operable states, and convergence. Furthermore, we model the interaction of overlays and underlays as a dual game consisting of a short-term overlay game and a long-term overlay–underlay game. This is motivated by the fact that underlay network management actions occur at much longer timescale compared to overlay routing decisions.

First, we abstract the short-term game as a single aggregate overlay player that plays against an aggregate underlay player. Both games are modeled as staged congestion games, and we show that stability can be obtained when the following conditions are met: 1) the overlays can stabilize in a noncongestive configuration within the time interval of underlay management actions; 2) both the overlays and the underlays have strong desire for congestion avoidance despite their otherwise selfish behaviors; 3) the overlays have control over domain traversal. We discuss the feasibility of these guidelines in practice and how the overlays and the underlays should operate to achieve system wise stability. Second, we explore the application of Shapley value as the basis for mutual cooperation among underlays, and we model the general overlay–underlay interactions as a two-stage Stackelberg game. Our analysis shows that stable and operable global states can be obtained for this general environment where three types of interactions coexist: overlay–overlay competition and cooperation; underlay–underlay competition and cooperation; overlay–underlay interaction without explicit information exchange.

The rest of the paper is organized as follows. Section II presents related works. Section III analyzes the overlay interactions and discusses implementation techniques. Section IV shows the long-term overlay–underlay interactions. Simulation study is presented in Section V, and we conclude in Section VI with a summary of results and future work.

II. Related Work

On the topic of overlay network routing, Liu *et al.* [6] have studied the interaction between overlay routing and underlay traffic engineering. They modeled the selfish optimization of the overlay routing and the load balancing objective of the underlying traffic engineering as a two-player noncooperative nonzero sum game. They found that although equilibrium exists in simple networks, oscillations and inefficiency arise due to the divergent objectives in general topologies. Seetharaman *et al.* [7], [8] further expanded along this direction in their study of the routing performance of underlay and overlay layers using a Stackelberg approach. They were able to show that performance gain could be obtained when one layer takes the leader role while the other player takes a follower strategy. They further noted that the degree of performance gain for each layer is a win–lose struggle determined by the selfishness of a layer. In our work, we consider operable states as a way of promoting common welfare among the layers. Using realistic topologies and traffic demands, Qiu *et al.* [9] presented through simulation that contrary to theoretical worst case, selfish routing in the overlay achieves close to optimal average latency. They also noted that some links in the network will have significantly increased congestion. Keralapura *et al.* [10] showed that racing conditions exist between noncooperative overlays due to inadvertent synchronization. Their result confirmed the observation that asymmetric congestion games in general cannot allow simultaneous moves. Zhang *et al.* [11] showed that when the underlay topology is rich, the overlay can compensate underlay routing inefficiencies. To date, few works have considered the combined behavior of interacting overlays as well as overlay–underlay interactions in multidomain setting. Peering-based settlement has been the *de facto* basis of underlay cooperation among autonomous systems (ASs) owned by different ISPs. In their insightful work, Ma *et al.* [12] have investigated the issue of corrosion among ISP bilateral settlement-free peering. They proposed a new multilateral settlement peering strategy based on Shapley value that can achieve efficient, fair, and optimal routing in the underlay. We leverage their results in our overlay–underlay interaction scenario and examine the resulting system optimality.

Game-theoretical analysis has been conducted in many fields of network research in the past (e.g., pricing, flow control, efficiency of wireless networks, etc.). Some works have examined the uniqueness of Nash equilibrium in noncooperative user-based routing environment [1], [2]. Orda *et al.* [1] have shown that in two-node multilink network topology, a unique Nash equilibrium exists. Altman *et al.* [2] studied noncooperative routing games under a general topology network with polynomial cost function. They have shown the uniqueness of Nash equilibrium under bounded cost. Yaiche *et al.* [13] modeled the bandwidth allocation problem as a Nash bargaining solution and devised a distributed optimization algorithm that is Pareto-optimal. They are among the few earliest works that addressed game implementation and explicitly used social fairness as a performance criteria.

Bottleneck routing game was investigated by Banner and Orda [14], in which the user attempts to minimize the load of its bottleneck link rather than to minimize the end-to-end cost. In a nutshell, the objective is a discontinuous MINMAX function. They have shown that in unsplittable bottleneck game, the worst-case convergence bound is exponential $O(2^{|U|^2}|E|^U)$ and the price of anarchy is polynomial $O(|E|^p)$, where $|U|$ is the number of users, $|E|$ is number of links in the network, and p is a constant. In our formulation, we use an exponential objective function that is continuous and strictly convex, which exhibits a number of interesting properties compared to a strict MINMAX function. Furthermore, the formulation of our problem considers both the interactions among overlays and the overlays–underlays game. Recently, Kollias and Roughgarden have shown a novel method of finding pure Nash equilibrium for weighted congestion games under Shapley value [15]. We extend their results to show that our model fits under the same general category and thus can serve as the basis of mutual cooperation among the overlays as well as between the overlay and the underlays.

Congestion games were first introduced by Rosenthal [16] and later formalized by Monderer and Shapley [5]. Fabrikant *et al.* [4] have shown that the complexity of finding a pure Nash equilibrium in asymmetric congestion games is PLS-complete. Milchtaich [3] has shown that polynomial time convergence exists for players with varied payoff functions. Goldberg [17] bounded their convergence to polynomial time, and Even-Dar and Mansour [18] considered the case where all players can move simultaneously according to a Nash rerouting policy and bounded its convergence to polynomial time. Because of complexity, the study on convergence in general congestion games has been mainly focused on convergence to approximate solutions. Christodoulou *et al.* [19] bounded the solution after one round of best-response walk by all players to $\Theta(\mathrm{n})$-approximate in the general case. Chien and Sinclair [20] showed that when the increase in cost of adding a player is bounded, convergence to ϵ-Nash occurs in polynomial time. The congestion game model we study here concerns multicommodity asymmetric games with polynomial cost function. It has a worst-case exponential convergence bound [14], [21]. Therefore, we have established fast convergence bound on the specific games we establish in this paper.

In bounding the optimality of an equilibrium, PoA is frequently used [22]. Some analysis (e.g., [23] and [24]) has been conducted on bounding the PoA ratio in congestion games, and in some cases tight bounds are found. Roughgarden and Tardos [25] have shown that for nonatomic (i.e., each player controls negligible portion of the network) and splittable flows, the price of anarchy is exactly 4/3. This result is extended to atomic unsplittable flow case in their subsequent work [26]. Awerbuch *et al.* [24] have shown that for atomic unsplittable flows, the price of anarchy is exactly 2.5 for a pure strategy game over unweighed players with linear latency cost function. In the above game, if the cost function is a general polynomial cost function of degree d, the price of anarchy bound is at least $\Omega(d^{d/2})$ and at most $O(2^d d^{d+1})$.

III. Interactions Among Overlays

In this section, we analyze the overlay interactions. We will analyze the interaction of overlay–underlay game in Section IV, which builds on many of the properties we establish in this section. First, we model the overlay interactions as a congestion game with a congestion avoidance cost function. We define operable states as states wherein no network links are congested and examine the game under which the Nash equilibria are operable states. Our interest in congestion avoidance and operable states is motivated by the following: 1) network congestion is the major source of delay and transient network failures; 2) the application-oriented nature of overlays means it is often sufficient to satisfy a delay constraint rather than minimizing delay; 3) a congestion-minimization objective at the overlay layer is a necessary precondition to guarantee overlay–underlay network stability as we will show in Section IV. By performing game transformation using binary factoring on the resource, we bound the convergence span of our game model. We also derive the price of anarchy.

Second, we examine the general case of weighted congestion game in which the players may have varied traffic demands. It is known that weighted congestion games do not have pure Nash equilibrium in general. However, a recent study [15] shows that by constructing the utility function based on Shapley value, pure Nash equilibrium can be restored. We extend this finding to our case and show that with the proper choice of a λ parameter, the Shapley induced equilibrium is guaranteed to be an operable state as well. Furthermore, we bound the convergence of such a weighted game to an operable state in linear time. We discuss the implication of Shapley utility function as a natural inducer of cooperation in an otherwise competitive environment. Finally, we relax the sequential move constraint of our game to allow partial simultaneous moves, which further reduces the convergence span.

A. Interaction of Overlays as a Congestion Game

For an overlay, we consider each source–destination pair in the overlay traffic matrix to have a set of candidate overlay routes. Each overlay route traverses a consecutive list of underlay domains. We represent each domain as one or more virtual resource links t. Details of this mapping and its implication on underlay network management are discussed in Section IV. We define a unit traffic κ as a discrete volume of traffic and consider each overlay source–destination pair is split into finite number of such unit traffic. We define the capacity of a resource link t_j as the amount of traffic it can serve and normalize its value according to κ. Thus, a resource link abstractly represents an underlay domain's capacity to host overlay traffic and is considered in operable state unless the capacity is exceeded by overlay demand. Also, each unit traffic can be routed independently of the others and is considered a player in our game model. Accordingly, we define the overlay routing game as follows.

Let $\Gamma_D = \langle N, \{Y_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$ be a game in strategic form. N is the finite set of players $\{1,\ldots,n\}$, and Y_i is the finite set of strategies available to player i and $u_i : Y \rightarrow \Re_+$ where $Y = Y_1 \times Y_2 \ldots \times Y_n$ is the cost function of player i. Given a finite set of resources $T = \{t_1, \ldots, t_m\}$, define $Y_i \subset 2^T$. Let $A_i \in Y_i$ be a strategy of player i, $A \in Y$ be a strategy profile, c_j be the cost function of resource t_j, and l_j be the normalized serving capacity of t_j, then

$$u_i(A) = \sum_{j \in A_i} c_j(A)$$
$$c_j(A) = \left(\frac{x_j(A)}{l_j}\right)^{\lambda} \qquad \lambda \gg 1$$
$$x_j(A) = |\{i \in N : t_j \in A_i\}|.$$

Γ_D is a multicommodity asymmetric unweighed game. The cost function c_j of resource t_j is a strictly increasing function of the number of players using t_j. The game is multicommodity since a strategy includes more than one resource, and the game is asymmetrical since each player may have different strategy sets (e.g., different overlays have different routes). The game is unweighed in that each player has unit load κ. Γ_D has an exact potential $\phi = \sum_{j=1}^{T}\sum_{k=0}^{x_j(A)}(k/l_j)^{\lambda}$ and therefore a pure Nash equilibrium.

A congestion game with an exact potential not only has a pure Nash equilibrium, but also has the finite improvement property (FIP). Hence, from an arbitrary state, following the best-reply path, the game is guaranteed to converge to the pure Nash equilibrium over time, albeit exponential in worst case. We hereby define the operable state of a system as follows.

Definition 3.1: A strategy profile A is an operable state of Γ iff the following condition is true:

$$\frac{x_j(A)}{l_j} \leq 1 \qquad \forall t_j \in T. \tag{1}$$

In a nutshell, we are interested in games wherein NE states are also operable states. Furthermore, by defining operable states, we can examine not only convergence to equilibrium, but also convergence to operable states. One philosophy of ensuring NE states to be operable states is to consider cooperative game play in which players collaborate to achieve a desirable equilibrium. However, as Mahajan *et al.* [27] rightly argue, such collaboration often gives rise to implementation difficulties. Issues such as information sharing, privacy, identity, etc., are essential to cooperative game playing, but difficult or expensive to realize in distributed setting. Furthermore, the players in our scenario do not have strong incentives to faithfully cooperate with each other as their objective functions are largely self-fulfilling rather than system agonistic. These concerns led to the exploration of another philosophy: to examine noncooperative congestion game Γ_D whose NE states are also operable states. This property can be satisfied by defining a cost function that heavily penalizes overloaded resources. Accordingly, we define the polynomial cost function $c_j(A)$. With an appropriate λ value, the cost function ensures an overloaded resource will have value far greater than 1 while any nonoverloaded resource will have value far less than 1. Thus, this cost function reflects the mentality of a congestion-avoiding player by exhibiting a MINMAX property. A cost function modeled based on M/M/1 queuing is also considered, but without a tunable parameter such as λ, it is difficult to guarantee that the MINMAX property always holds. The exact MINMAX function as defined by bottleneck congestion game [14] is also considered. However, as we will show, our particular form of objective function permits a price of anarchy of 1 rather than a polynomial PoA of the MINMAX.

Theorem 3.1: The NE states of Γ_D are operable states if there exist operable states in the system.

Proof by Contradiction: Assume this is not the case. Let the equilibrium be A^* and select an arbitrary operable state A'. By our assumption, A^* contains at least one overloaded resource, while A' does not. As the equilibrium state A^* corresponds to the minimization of the potential function, then $\phi(A^*) \leq \phi(A)$, $\forall A \in Y$, $A \neq A^*$. Since A' contains no resource with load over the threshold, $\phi(A') = c$. where c is some small constant. Since A^* is not an operable state and hence contains at least one resource with load exceeding the threshold, thus $\phi(A^*) \gg c$. Therefore, $\phi(A^*) > \phi(A')$, $A^* \neq A'$. We arrive at a contradiction. □

Corollary 3.1.2: After finite number of moves, Γ_D converges to an NE state that is operable.

The system potential $\phi = \sum_{j=1}^{T}\sum_{k=0}^{x_j(A)}(k/l_j)^{\lambda}$ is strictly convex in the positive domain, thus permits a unique global minimal that is reachable through finite improvement [3]. By Theorem 3.1, it follows that this equilibrium is also an operable state. This state, however, is not guaranteed to be unique. For instance, consider two players with identical strategy set in an NE state. It follows that if the two players are to swap their strategy choices, we arrive at a new NE state. However, this set of NEs all exhibits the same global minimum. □

Corollary 3.1.3: Γ_D has price of anarchy of 1.

This result directly follows from Corollary 3.1.2. The convergence to the global minimum yields a PoA of 1. However, the convergence span may be very long. We address this aspect in Section III-B. □

It is important to set a proper value for λ. λ should be large enough to guarantee that our cost function behaves like MINMAX (without its nasty discreteness). This is because a small $\lambda > 1$ value may not provide MINMAX guarantee as the sum of costs of multiple non-bottleneck links may overcome the cost of the bottleneck link. At the same time, we want the λ value to be as low as possible, because it directly affects the convergence span to both operable state and the NE, as well as the PoA. Therefore, we derive the lower bound of λ and find it to be $O(\log(p)/\log(1 + (1/L)))$, where L is the maximum link capacity and p is the maximum path length. We detail the proof in the Appendix.

B. Convergence to NE and Operable States

Bounding the convergence of Γ_D is challenging due to the nature of the cost function and the variability in resource capacities. More specifically, to bound the convergence span of a potential game, one needs to bound the minimal potential drop due to a player's move, which is the difference between the benefit obtained from reducing congestion on some links and the cost incurred from adding congestion on some other links. With varied resource capacity and a polynomial cost function, this difference can be arbitrarily small. For weighted max-congestion games, it is known that the complexity of pure Nash equilibria is PLS-complete [21], and for Bottleneck congestion games, of which our game model is a more general form, the convergence bound is known to be exponential $O(2^{|U|^2}|E|^U)$ with respect to the number of users $|U|$ and the number of links $|E|$ [14]. Thus, we seek an approximate solution. We first transform the game Γ_D into a binary factor form Γ_D^T, which has the interesting property that the difference in potential due to a player move is always a multiple of a common factor, and obtain a polynomial time convergence bound of $O(L^\lambda Cn)$. We show that Γ_D and Γ_D^T share the same operable states and therefore an NE state obtained via Γ_D^T corresponds to a stable operable state in Γ_D, so we can thus bound the price of anarchy.

We now define the transformed game Γ_D^T as a congestion game in strategic form identical to Γ_D except for a transformed resource collection and a transformed strategy set. Let $L = [\{l_j\}_{j\in T}]_{\mathrm{MAX}}$, construct the binary factor set $B = \{2^0, 2^1, \ldots, 2^{\lfloor \log_2 L\rfloor}\}$. For each resource t_j in T, associate a resource set $t_j^T \subset B$ in Γ_D^T, such that $\sum_{k\in t_j^T} l_k = l_j$. Hence, the set of resources in Γ_D^T is a binary factoring of the resources in Γ_D. For each strategy A_i of player $i \in N$, associate to Y_i^T the set of strategies $\prod_{t_j\in A_i} t_j^T$. Thus, each A_i of i in Γ_D is expanded to a set over the binary factoring of the resources in A_i.

Theorem 3.4: Convergence in the transformed game Γ_D^T is bounded by $O(L^\lambda Cn)$.

It is clear that Γ_D^T has a pure Nash equilibrium induced by a system potential. When a player makes a move to a new state, it must be that the new state has a lower cost than the old state. Since Γ_D^T has an exact system potential, by its exact potential property, a drop in a player's cost must cause an equal drop in system potential. We claim that in Γ_D^T, the smallest drop in potential when a player makes a move is $2^{-\lambda\lfloor \log_2 L\rfloor}$. Suppose player i makes a move, let A be the state before the move and A' be the state after the move. Using the system potential definition from Theorem 3.1, the drop in potential can be expressed as $\Delta\phi = \sum_{j\in(A_i-A_i')}(x_j(A)/l_j)^\lambda - \sum_{k\in(A_i'-A_i)}((x_k(A)+1)/l_k)^\lambda$. Since all l_j are factors of 2, there must exist a sequence of constants $a_1, a_2, \ldots, a_m$ such that for all l_j, $a_j l_j = 2^{\lfloor \log_2 L\rfloor}$. It then follows that

$$\begin{aligned}\Delta\phi &= \sum_{j\in\left(A_i-A_i'\right)}\left(\frac{a_j x_j(A)}{a_j l_j}\right)^\lambda - \sum_{k\in\left(A_i'-A_i\right)}\left(\frac{a_k(x_k(A)+1)}{a_k l_k}\right)^\lambda\\ &= \sum_{j\in\left(A_i-A_i'\right)}\frac{(a_j x_j(A))^\lambda}{2^{\lambda\lfloor \log_2 L\rfloor}} - \sum_{k\in\left(A_i'-A_i\right)}\frac{(a_k\,(x_k(A)+1))^\lambda}{2^{\lambda\lfloor \log_2 L\rfloor}}.\end{aligned}$$

Since all the terms in the above equation have a common denominator, the result of the arithmetic operations is guaranteed to be some multiple of $2^{-\lambda\lfloor \log_2 L\rfloor}$. Therefore, the smallest potential drop from a move in Γ_α^T is bounded by $2^{-\lambda\lfloor \log_2 L\rfloor}$. Let $\phi_{\max}$ and $\phi_{\min}$ be the upper and lower bounds on the potential values respectively, and C be the upper bound on the cost of any player, then the maximum number of steps to convergence is bounded by

$$\frac{\phi_{\max}-\phi_{\min}}{2^{-\lambda\lfloor \log_2 L\rfloor}} \le \frac{nC}{2^{-\lambda\lfloor \log_2 L\rfloor}} = O(L^\lambda Cn).$$

□

Theorem 3.5: An operable equilibrium exists in Γ_D iff an operable equilibrium exists in Γ_D^T.

Given an operable equilibrium A^* in Γ_D, it must be the case that for all $t_j \in T$, $x_j(A^*) \le l_j$. By transforming t_j to its factored set t_j^T, the total capacity does not change: $\sum_{k\in t_j^T} l_k = l_j$. Therefore, $x_j(A^*)$ can easily be distributed among the factored resources in t_j^T such that none of the resources exceed threshold. In fact, there is an optimal distribution under the best-reply dynamic. Therefore, the transformed state A^{*T} is an operable state in Γ_D^T. By Theorem 3.1, it is then clear that Γ_D^T must have an operable equilibrium.

Given an operable equilibrium A^{*T} in Γ_D^T, it must be the case that for all $t_k \in T$, $x_k(A^{*T}) \le l_k$. With similar reasoning as above, one can assign $\sum_{k\in t_j^T} x_k(A^{*T})$ to the corresponding resource t_j in Γ_D, and it is guaranteed that the number of users of resource t_j will not exceed threshold. Hence, the transformed state A^* is an operable state in Γ_D. By Theorem 3.1, Γ_D then must have an operable equilibrium. □

Proposition 3.2.1: For every operable state in Γ_D, there is a corresponding operable state in Γ_D^T, and vice versa.

This is a stronger claim that follows from the proof of Theorem 3.5. Given an operable state in Γ_D, by applying the transformation technique, one will obtain an operable state in Γ_D^T, and vice versa. □

Hence, an equilibrium state reached in Γ_D^T is an operable state in Γ_D. In other words, a finite set of players participating in a game Γ, can in fact obtain convergence to stability by playing a simpler game Γ_D^T and still arrive at an operable state in Γ. This is in effect a tradeoff between the quality of the game equilibrium and the length of the convergence span, with the added benefit of an operationally useful upper bound (i.e., the guarantee of an operable state). Since Γ_D^T is transformed from Γ_D and results in a different operable equilibrium, it is important to bound the price of anarchy. We find that the loss of optimality due to transforming Γ_D to Γ_D^T is 2 (proof presented in the Appendix).

The minimization of system potential in Γ_D naturally leads to an optimal configuration of balanced load in the system. Thus, the definition of price of anarchy can be given as follows:

$$\frac{\Gamma_D}{\left[\left(\frac{x_j(A)}{l_j}\right)_{j\in T}\right]_{\max}}.$$

From Corollary 3.3 we know the price of anarchy for Γ_D is 1, then it follows that the price of anarchy for Γ_D^T is 2.

This game transformation technique is recursive. Γ_D^T could be further transformed by breaking $2^{\lfloor \log_2 L\rfloor}$ into two resources of $2^{\lfloor \log_2 L\rfloor-1}$, and yields a convergence bound of $O((L/2)^\lambda Cn)$ with price of anarchy 4. More generally, letting α be the number of such recursions, the convergence bound is $O((L/2^{\alpha-1})^\lambda Cn)$ with price of anarchy 2^α.

From an operations point of view, it is often sufficient to guarantee that the system operates under an operable state (i.e., a noncongestive state) rather than absolute convergence to the NE. This motivates an analysis on the convergence of our game to an operable state. For ease of analysis, let us categorize a move made in the congestion game Γ_D as either a *big move*—the game state before a move is a congestive state—or a *small move*—the game state before a move is a noncongestive state.

Lemma 3.2.2: The leading moves of congestion game Γ_D must all be big moves under best-response.

This is rather intuitive to see. The change in potential of a big move involving a congestive link is an order of magnitude higher (in terms of λ) than the changes in potential of a small move. Hence, under the best-response dynamics, the big moves are always made first. □

Lemma 3.2.3: The trailing moves of congestion game Γ_D must all be small moves under best-response.

Suppose the converse is true. Then, there exists some big move k that occurs in between small moves. This must imply that there exists a small move before k that has changed the system state from noncongestive to congestive. However, such a move necessitates an increase in potential and thus will never occur. We arrive at a contradiction. □

Theorem 3.7: The convergence of congestion game Γ_D to an operable state is upper-bounded by $(L/\lambda)Cm$.

From Lemmas 3.2.2 and 3.2.3, it follows that the system will arrive at an operable state after all of the big moves are exhausted. Thus, we can obtain the bound on convergence to an operable state by bounding the minimum drop in potential of a big move. Let L be the capacity of a resource, then the minimum drop in potential can be bounded by

$$\left(1+\frac{1}{L}\right)^{\lambda}-1>\lambda\left(\frac{1}{L}\right).$$

The above formulation is based on the observation that the drop in potential is at its minimum when a resource is just above the congestion threshold value 1 before the move and to the threshold value 1 after the move. Let C be the maximum congestion on any resource, and m be the number of resources, then we have the following bound:

$$\frac{(C-1)m}{\lambda\left(\frac{1}{L}\right)}>\frac{L}{\lambda}Cm.$$

□

The above analysis shows that convergence to an operable state can be achieved in linear time rather than convergence to the NE in polynomial time. This result is directly applicable to Γ_D. Therefore, under the objective of congestion avoidance, the selfish behaviors of uncoordinated overlays can produce an operable state quite efficiently. Another useful outcome is that the convergence bound is not affected by the number of overlays in the system, but rather on how many resources are accessible in the system and how congested a resource can be. Thus, it fits with our intuition that a well-engineered network with high excess capacity will achieve an operable state quicker.

C. Pure Nash Equilibrium and Convergence of Weighted Game

Γ_D is an unweighed game wherein the players all have the same traffic demand. In practice, overlay flows may have varied traffic demand due to the different applications they support. Furthermore, a group of overlay players may form coalitions as to jointly optimize their flows to achieve better performance. Both of these scenarios fall under the category of weighted congestion games wherein players have varied traffic demands. It is known that weighted congestion games do not generally have pure NE. However, it has been recently shown [15] that instead of directly constructing the player utility function based on a player's perceived payoff/cost, Shapley value [28] can be used as an alternative. The application of Shapley value to our problem context brings two benefits: 1) Shapley cost guarantees the existence of a pure NE in weighed games; 2) Shapley potential allows us to examine the interplay between selfish play and coalitions that are largely intermixed in today's overlay networks.

Accordingly, we can replace the cost function of Γ_D with the Shapley value of the overlay player i as follows:

$$\mathrm{SV}_{i,t}(S_t)=E\left[\left(\frac{X_{i,t}+w_i}{l_t}\right)^{\lambda}-\left(\frac{X_{i,t}}{l_t}\right)^{\lambda}\right].$$

$S_t\subseteq N$ is the set of players using resource t. Given a uniform random ordering of S_t, $X_{i,t}$ is a random variable with value equal to the total weight of those players that appear before i in t, and w_i is the weight of i. Let Γ_{SV} denote such a weighted congestion game with Shapley utility, the potential of Γ_{SV} can be defined by summing the aggregate Shapley values of the resources in the system as

$$\phi_{\mathrm{SV}}=\sum_{t\in T}\sum_{i\in S_t}\mathrm{SV}_{i,t}\left(S_t(\pi,\pi_i)\right).$$

This is an exact potential of the game [15]. The price of anarchy for this Shapley potential game is $(O(\lambda))^{\lambda+1}$. The player's strategy set is the same as in Γ_D to choose an optimal set of resources to support its end-to-end flow. However, the objective is now to minimize its Shapley value cost rather than the aggregate congestion level of a resource. The operations significance of this objective is that rather than evaluating the cost of purely selfish operations, the players are also concerned with the marginal cost they bring to feasible coalitions. With the addition of monetary incentive (Section IV-C), we can better understand the decision dynamics of overlay players who evaluate whether to form coalitions with others or whether coalition is more beneficial compared to selfish play.

The Shapley cost function is still congestion-avoidance-driven, which leads to stable operable states, as we will show later in this section. However, it is more socially conscious because it considers the weighed mix of the other overlays sharing the same resource and its own expected marginal cost on the resource. In doing so, Γ_{SV} not only allows us to restore pure NE to weighed congestion game, but also allows us to bound the operable states and convergence of Γ_{SV}. In this way, we obtain an analytical framework to understand a complex environment in which selfish play coexists with cooperative coalitions and how such group dynamics affect stability, operable states, and convergence.

As we have done for the unweighed game, we now bound the value of λ that is necessary to guarantee an operable pure NE.

Lemma 3.3.1: A player i will always prefer a transition from a congested state to an operable state when λ is at least $\log(p)/\log(1+(w_i/L))$

First, we observe that the ordering π does not matter. This is the key insight used to prove that $\phi_{\rm SV}$ is an exact potential of $\Gamma_{\rm SV}$. Thus, we can fix an ordering of the players on each resource t such that player i always comes last. This arrangement allows us to apply the proof of Lemma A1.1 and Theorem A1.2 in the Appendix on the worst-case boundary condition for λ, yielding

$$\lambda \geq \frac{\log(p)}{\log\left(1+\frac{w_i}{L}\right)}.$$

□

Letting i be the player with the least weight $w_{\min}$, we can obtain the lower bound on λ.

Theorem 3.8: The NE of $\Gamma_{\rm SV}$ is operable (if the game permits one) when λ is at least $\log(p)/\log(1+(w_{\min}/L))$.

By Lemma 3.3.1, the players of $\Gamma_{\rm SV}$ prefer a move from congested state to operable state, as long as the current game state contains an alternative strategy for the player to shift to (i.e., a strategy arrangement of the player i that would permit an operable state). Since we can fix an arbitrary ordering of players or a resource, the above assertion always holds. Thus, a game state with congestive links cannot be the NE unless all of the players' strategies in the system yield at least one congestive link. In this latter case, there does not exist any operable state in the game.□

Although we cannot bound the convergence of $\Gamma_{\rm SV}$ to polynomial time due to the complexity of this class of games, we can bound its convergence to an operable state. Lemmas 3.2.2 and 3.2.3 are applicable to $\Gamma_{\rm SV}$. Thus, similar to the proof of Theorem 3.7, we can bound the minimum drop in potential by

$$\frac{(C-1)m}{\lambda\left(\frac{w_{\min}}{L}\right)} > \frac{L}{\lambda w_{\min}}Cm.$$

By constructing a Shapley utility of the appropriate degree, we have shown that we can obtain operable pure NE of weighted congestion game whose convergence to an operable state also occurs in linear time. This is again independent of the number of overlays in the system.

Another important property of the Shapley utility is its fairness in distributing payoff of the shared resource usage, congestion cost in this case. By computing the Shapley value $\mathrm{SV}_{i,e}$, we can establish a natural bridge between cooperation and competition among the players. The Shapley value potential yields a system wherein the players have incentive to cooperate with each other because the noncooperative alternative yields a congestion cost that is at least as large as the Shapley value. This key property holds important implications in different overlay environments: 1) in systems where large populations of the overlays do not jointly optimize their routing (e.g., pure selfish routing), $\Gamma_{\rm SV}$ provides an incentive compatible mechanism for indirect cooperation; 2) in systems where some portions of overlays do jointly optimize in coalitions but competitions predominates among coalitions (e.g., different content distribution networks), the potential reachable in the system is lower than the purely noncooperative Shapley NE potential. The degree of reduction is dependent on the relative size of the coalition compared to the player population size. Nevertheless, its worst-case bound is the Shapley NE potential despite competition among coalitions. Our simulation results further demonstrate this property (Section V).

D. Simultaneous Moves Among Overlays

Throughout our analysis, we have considered a system where only one player moves at a time. This property must hold to guarantee convergence under the FIP. Without guarantee of sequential moves, even a game with pure Nash equilibrium can produce perpetual oscillations [10]. We now consider a method to enable partial simultaneous moves. Assuming each player's strategy set is a small subset of the common resource collection, let T_i be the set of resources used by player i's strategy set Y_i ($T_i = \{t_j : t_j \in A_i\}$). Then, we define the *neighborhood* of player i as

$$\mathrm{NB}_i = \{k : T_k \cap T_i \neq \varnothing\}_{k \in N}.$$

We observe that if a player k is not in the neighborhood of player i or vice versa, then players i and k may move simultaneously in the system, and the resulting potential change is as if they have moved in sequence. We can therefore construct a neighborhood graph $G = \langle N, E\rangle$. G is undirected, and an edge exists between two players iff they are in the neighborhood of each other. It follows that all nodes not directly connected on the graph can move simultaneously. Therefore, the minimum number of iterations it takes for all players to have a chance to move is equal to the minimum number of colors needed to color this graph. Furthermore, if a resource t_k is present in all of a player's candidate strategy, we do not include this player when considering the neighborhood of t_k. This is because no matter what move this player makes, it will not change its congestion contribution to t_k. This is a useful reduction since Internet topology is hierarchical. At the top levels, the number of transit domains are rather limited, and hence overlays transiting between similar source and/or destination domains will likely use the same set of transit domains in all of their strategy sets.

IV. Interaction Between Overlays and Underlays

In this section, we model the long-term interaction between overlays and underlays. By abstracting the underlay domains as virtual resource links and mapping the overlay and underlay traffic onto these links, we show that the resulting scenario can be modeled as a two-player congestion game. The congestion avoidance objective of the overlays agrees with the resource optimization objective of the underlays. Hence, we arrive at a stable and sometimes unique game solution. We discuss the game's preconditions and argue that they are achievable in practice. Then, we link the Shapley-based overlay potential with Shapley-based underlay settlement strategies [12] and show system stability and convergence. This approach offers an alternative economic scenario for the overlay operators and underlay providers.

A. Interaction Between Overlays and Underlays as a Congestion Game

We consider that domain-wise management actions are conducted periodically by underlay providers independently

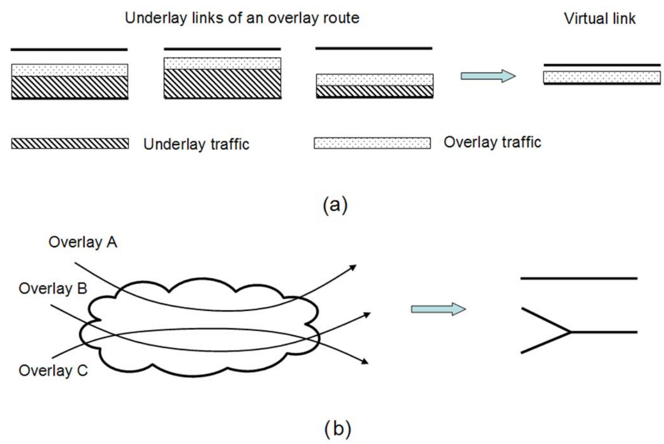

Fig. 2. Underlay routing of overlay routes is mapped to virtual links. (a) Virtual link capacity. (b) Mapping overlay routes crossing a domain to virtual links.

to optimize resource usage (i.e., load balancing) of both the overlay and underlay traffic. We also consider that, overlays aside, the rest of the underlay traffic in a domain is stable in semi-long term, such as weekly. Therefore, we model the underlay traffic matrix for nonoverlay traffic based on fixed volume and source–destination. We consider that dynamics in overlay traffic induce underlay management actions.

Conversely, an underlay's intradomain routing decisions directly control how an overlay traffic traverses its domain. Since overlays are not concerned with the specific underlay links and are affected only by the bottleneck link capacity and other overlays that share the same underlay link, it is then sensible to abstract the resource occupancy of underlay routes as *virtual links* as follows: A virtual link represents the bottleneck underlay link occupancy experienced by an overlay across a network domain (Fig. 2). In addition, two mutually exclusive cases can be observed regarding the interactions of the overlays in an underlay domain: 1) the overlays do not share any underlay links; 2) they do. In case 1), the traffic of the overlays do not affect each other and hence are mapped to disjoint virtual links; otherwise, the overlays are mapped to a disjoint virtual link and a shared virtual link. Fig. 2 illustrates the mapping. Fig. 2(a) shows how the underlay links of an overlay route are mapped to a virtual link and the bottleneck link capacity is taken as the virtual link capacity. Fig. 2(b) shows the mapping of multiple overlay routes in a domain to virtual links. The shared underlay links of Overlays B and C are mapped to a shared virtual link, and the remaining disjoint underlay links are mapped to a disjoint virtual link. The result of this mapping produces a graph of virtual links on which each strategy of the overlays is realized by an adjoining sequence of virtual links. These virtual links are not physically present in the network management operations, but are rather a modeling artifact to relate the resource view of the overlay games to the physical network view of the underlays.

For this mapping to be valid end-to-end across multiple domains, it is important for the networks to be in a congestion-free configuration, such that the workload of an overlay is experienced by all of its virtual links. The overlay must also have control over the network domains it traverses. The former condition is satisfied in Γ_D, and the latter condition will be discussed in Section IV-B. When an overlay route changes, it generally induces a corresponding change in its underlay route mapping, as well as a reroute of the nonoverlay portion of the underlay traffic, resulting in a new set of virtual links. This in turn affects the routing of the overlays. Hence, their interaction affects the network stability.

Let Γ_L be a two player game where Player 1 is the aggregate overlay player and Player 2 is the aggregate underlay player. In this game, Player 1 always moves first, then followed by Player 2. The strategy set of Player 1 consists of the set of all possible overlay traffic distributions Y. The strategy set of Player 2 consists of all possible mappings of $H(Y,Z) : Y,Z \rightarrow T$, where $T = \{t_i\}$ is the set of virtual links and Z is the underlay traffic matrix. More specifically, the overlay player's (Player 1) move produces a demand on the distribution of overlay traffic across specific set of underlay domains; the underlay player's (Player 2) move is then to route the overlay traffic matrix Y and the underlay traffic matrix Z across the underlay network links, thus producing a mapping of the traffic demand set $\{Y,Z\}$ to a set of virtual links $\{t_i\} \in T$. The capacity of each virtual link l_j as viewed by the overlays is then the residual bottleneck underlay link capacity (i.e., $l_j = \mathrm{MIN}\{L_j\} - Z_j$), where L_j is the set of underlay link capacities of domain j carrying the overlay traffic and Z_j is the amount of underlay traffic routed through this bottleneck underlay link. The cost function is defined as

$$\begin{aligned} u(A,Z) &= \sum_{j \in A} c_j(A,Z) \\ c_j(A,Z) &= \left(\frac{x_j(A)}{l_j}\right)^{\lambda}, \qquad \lambda \gg 1 \\ l_j &= \mathrm{MIN}\{L_j\} - Z_j \\ x_j(A) &= \left|\{i \in N : t_j \in A_i\}\right|. \end{aligned}$$

The cost function is similar to the overlay game except that the cost is defined over all utilized virtual links rather than the links used by an individual overlay player. The virtual link capacity is defined based on the mapping function $H(Y,Z)$. The objectives of Players 1 and 2 are aligned in that they both minimize the same aggregate cost function (through the mapping of $H(Y,Z)$), with Player 1 having control over A and Player 2 having control over T and l. In other words, whereas the strategies of an overlay player range over a collection of virtual link resources, the strategies of the underlay player range over a collection of physical resources it employs to support overlay traffic. Since there exists an exact mapping of the two resources $H(Y,Z)$, we obtain a consistent and common cost function. With this basis, we can establish the following: the collective behaviors of the individual overlays can be represented by an aggregate overlay player; the collective behaviors of the individual underlay domain providers can be represented by an aggregate underlay player; Γ_L is a potential game with $u(A,Z)$ being the system potential; and Γ_L has a minimal potential NE when the overlay game minimizes its potential and the network management operations are domain-wise optimal. The proof detailing these assertions can be found in the Appendix.

B. Establishing the Preconditions

The three preconditions that must be satisfied for the long-term overlay–underlay game to be stable are as follows.

1) The overlays and underlays objectives must align.
2) The domain traversal of an overlay is determined by the overlay.
3) The short-term overlay game must stabilize before a move in the long-term overlay underlay game.

Precondition 1 partly motivated our examination on overlay interactions from a congestion avoidance point of view. In practice, congestion is the primary cause of network delay, hence congestion avoidance can be a common objective alignment among the overlays and the underlays. We have thus devised an exponential cost function for overlays that is congestion-avoiding and ensures an operable system state.

Precondition 2 is motivated by the observation that system stability cannot be guaranteed when underlay independently decides on how interdomain routing should occur. Most of the interdomain routing issues today (e.g., route inconsistency, suboptimal performance, route conflict, etc.) can be attributed to this. As a direct result, recent network architecture proposals (e.g., [29] and [30]) advocate the underlays to relinquish control over interdomain routing. At the same time, underlay providers should be in charge of interdomain routing policies. This means that each underlay domain has jurisdiction over what domains it is willing to exchange traffic with, just not control exactly which domain an overlay must traverse. Moreover, each underlay domain need not disclose information about its intradomain operations; only reporting on domain performance aggregate (e.g., congestion level) is necessary. Thus, precondition 2 is feasible.

Precondition 3 relates to the convergence of overlay games. This is a strict requirement on the overlay game design to ensure timely convergence. Design aside, we have also shown implementation techniques such as partial simultaneous moves that can hasten the convergence span. The concept of operable states is another important factor as it can be obtained more readily than convergence to NE.

C. Shapley Based Approach to Underlay–Overlay Interactions

So far, we have considered overlay–underlay interaction where the transit and peering relations among underlay domains are stable and the domains largely adopt bilateral settlement-free interconnection. Although this behavior categorizes the traditional Internet interconnections, it has been shown [12] that new Internet economics are leaning toward more complex and fluid interactions among ASs with complex peering and customer–provider interconnections. Specifically, domains not only make intradomain routing decisions, but may also peer with other domains (i.e., coalitions) as to better improve their own revenue and reduce cost. The game we have analyzed in the previous Section IV-B assumes a static AS topology where each domain is completely selfish. Here, we would like to examine the more complex environment in which domains form coalitions in the form of peering and coalitions of domains compete with each other, under fixed underlay traffic demands and dynamic overlay traffic demands. In Section III-C, Shapley-value-based cost function provided a basis for our analysis. In thissection, we show that Shapley-value-based approach can also help to reason about these complex underlay–underlay interactions and can serve as the basis for joint overlay–underlay operations. Accordingly, our game model is modified from the two-player aggregate game into a Stackelberg game in which the aggregate overlay player (Player 1) is the leader, and the domains (formerly Player 2) are now a set of individual followers. We observe that the establishment of stability and convergence of the short-term overlay game (Section III) is crucial in allowing us to now abstract the overlays' complex behaviors as a single aggregate player, and hence establish the current Stackelberg game formulation.

The cost function of the followers is now defined based on the following Shapley value: Let $\mathcal{S}$ be the set of domains supporting the overlay demand matrix Y, and $S = |\mathcal{S}|$. We can define the Shapley value based cost function of a domain j as

$$\mathrm{SV}_j(\mathcal{S}, v) = \frac{1}{S!} \sum_{\pi \in \Pi} \Delta_j \left(v, \overline{\mathcal{S}}(\pi, j)\right).$$

Π is the set of all orderings of $\mathcal{S}$, and $\overline{\mathcal{S}}(\pi, j)$ is the set of domains preceding j in the ordering π. In another word, similar to the Shapley value definition of Section III-C, $\mathrm{SV}_j(\mathcal{S}, v)$ is the expected marginal contribution $\Delta_i(v, \overline{\mathcal{S}})$ assuming a uniformly random distribution of domain ordering. The value/worth function $v(\cdot) = v_{\mathrm{p}}(\cdot) - v_{\mathrm{c}}(\cdot)$ is directly tied to both the operations of the overlays using the domain (overlay produce profit value $v_{\mathrm{p}}(\cdot)$) and the routing cost $v_{\mathrm{c}}(\cdot) = c_j(A, Z)$ of the underlay domain. As we can see from the definition of the worth function, the term cost function is misleading here, as the underlay domains attempt to maximize $\mathrm{SV}_j(\mathcal{S}, v)$. This leads to a nice definition of the Shapley system potential $\sum_{j \in \mathcal{S}} \mathrm{SV}_j(\mathcal{S}, v)$.

As shown in [12], the above Shapley value has nice meaning at the domain level, as it allows for value decomposition in the following form:

$$\mathrm{SV}_j(\mathcal{S}, v) = \mathrm{SV}_j(\mathcal{S}, v_{\mathrm{p}}, E_{\mathcal{S}}) - \mathrm{SV}_j(\mathcal{S}, v_{\mathrm{c}}, R).$$

The term $\mathrm{SV}_j(\mathcal{S}, v_{\mathrm{p}}, E_{\mathcal{S}})$ is referred to as the Myerson value. $E_{\mathcal{S}}$ is the coalition of domains that domain j peers with as to support and therefore divide the profit v_{p} gained from supporting overlays. $\mathrm{SV}_j(\mathcal{S}, v_{\mathrm{c}}, R)$ is the routing cost of the domain where R is the strategy set of feasible routes traversing the domain. Therefore, the strategy set of a domain j is $\{E_{\mathcal{S}}, R\}$, meaning which domains to form coalition with as to increase its marginal contribution (profit) and what route to take intradomain as to minimize cost. The routing choice R ties in directly to the mapping of $H(Y, Z)$, while $E_{\mathcal{S}}$ deals with the formation of coalitions of underlay domains. Given a fixed domain interconnection topology, a reduction in the cost v_{c} necessarily increases the Shapley value of domain j and is thus always preferred. The NE of such a Shapley value maximization systemwise leads to global congestion reduction.

Lemma 4.2.1: The overlay game moves are incentive-compatible with the underlay Shapley value $\mathrm{SV}(\mathcal{S}, v)$.

Since the domain interconnectivity topology is fixed during the short-term overlay game, the game potential minimization on the cost function of the overlay directly results in the reduction of $v_{\rm c}$, and thus yields an increase in the Shapley value $\mathrm{SV}(\mathcal{S}, v)$. □

Lemma 4.2.2: Given a fixed topology, a Shapley improvement in underlay routing strategy necessitates a drop of system potential in the overlay game.

Since the Myerson value $\mathrm{SV}_j(\mathcal{S}, v_{\rm p}, E_{\mathcal{S}})$ can be easily maximized given a fixed/stable overlay demand matrix, the underlays have incentive to also reduce their Shapley cost term $v_{\rm c}$ by altering their traffic routing strategy as to minimize intradomain congestion. This explicitly results in a potential drop in the overlay game. □

Theorem 4.3: The interactions of the overlay–underlay under Shapley value mechanism results in a stable state.

Lemmas 4.2.1 and 4.2.2 jointly guarantee that the effect the short-term overlay game has on traffic routing is compatible with the effect of the long-term overlay–underlay game. Thus, their interplay converges to a globally stable routing state. Furthermore, the Myerson profit maximization mechanism of [12] guarantees that the settlement relations among underlay domains are also globally stable because mutually profitable and fair settlements emerge over time. □

Given that the respective Shapley values are used for both the overlay routing interactions and the underlay traffic routing and settlement interactions, we obtain an economically *fair* and *efficient* overlay–underlay Shapley solution: The underlays generate profit by cooperating with each other to jointly support the overlay traffics. The profit each underlay domain j receives is proportional to its Shapley marginal contribution to other resources by forming interconnection and reducing routing cost. At the overlay level, the profit function $v_{\rm p}$ translates to a fair monetary compensation an overlay can be charged for using a collection of resources, and the amount charged is proportional to its Shapley marginal congestion cost it brings to the used resources. More specifically, the cost to an overlay for using an underlay resource not only depends on the amount of resources the overlay uses relative to other overlays sharing the same resource, but also on the collaborative efforts the owning ISP of the resource must spend to cooperate with other ISPs in providing end-to-end connectivity. The use of Shapley values is efficient and economically fair to all participants involved, both the overlays and the underlays, and by the inherent property of Shapley value, we obtain a zero-sum transfer of utility among charging and paying players.

V. Simulation Studies

To study the interaction of overlays and underlays in multidomain networks, we simulated a connected network of 210 domains. One hundred randomly connected domains form the *source region*, and 100 other domains form the *destination region*. The two regions are connected by a *transit region* of 10 domains to form a two-tier hierarchy. Fifty source–destination pairs are randomly selected to represent overlays with the source nodes linked to domains in the source region and the destination nodes linked to domains in the destination region. Thus, each overlay route has a minimum span of three domains. Each overlay generates 4 units of traffic that can be routed independently (i.e., 200 overlay players in total) over four disjoint shortest paths. Because all the overlays must travel through the transit region that has limited number of domains, sufficient interaction among the overlays are ensured, and we can control the overall congestion level of the system by modifying the capacities of the transit domains. Each domain is modeled with a simple underlay topology as shown in Fig. 3. This simple topology allows for creating different traffic engineering profiles for the underlay–overlay interaction and is small enough to be computation light. Each domain is randomly assigned a fixed underlay traffic of 6–10 units. A domain planner computes the optimal distribution of traffic (overlay and underlay) in the domains and creates the virtual links for the overlay game. The underlay link capacities are shown in Fig. 3. Links in transit domains have higher capacity ({10, 20, 10}).

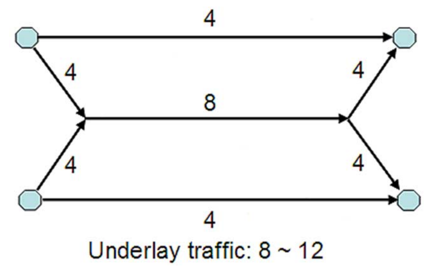

Fig. 3. Simulation setup for intradomain underlay topology.

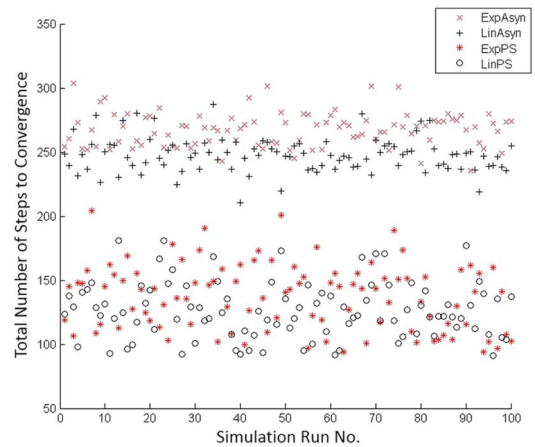

Fig. 4. Overlays convergence time.

Two sets of simulation studies are performed. In the first set, we study the interaction of the overlays when the underlay paths do not change. In the second set, we study the interaction of the overlays and the underlays. One hundreed simulation runs are conducted for each experiment. Time is broken into discrete steps, where a *step* is the time it takes for an overlay player to change its route. A *round* is the time it takes for all players to make a move if required.

Fig. 4 shows the total number of steps it takes for the overlays to stabilize. We study both the exponential cost function (Exp) of our work and the traditional linear cost function (Lin). Furthermore, we want to see the effect of allowing partial simultaneous moves (PS) compared to asynchronous moves (Asyn). For each run, a set of random overlay and underlay traffic is generated, and these varied techniques are applied on the same input. We observe that exponential cost

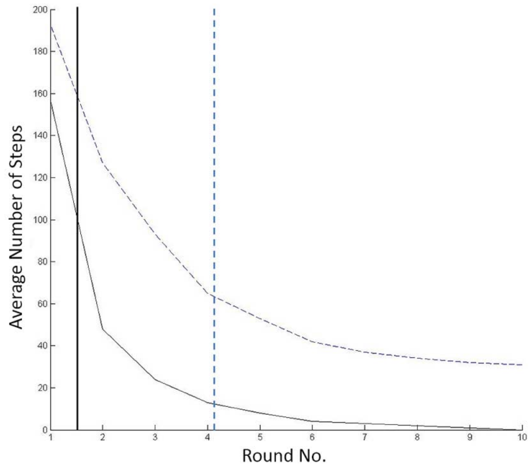

Fig. 5. Number of rounds to convergence.

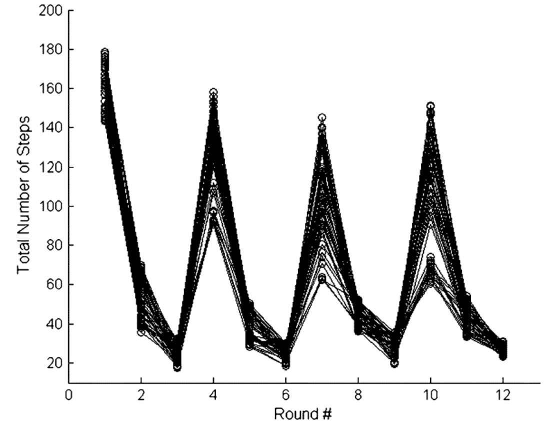

Fig. 6. Overlays and underlays interaction in short cycles.

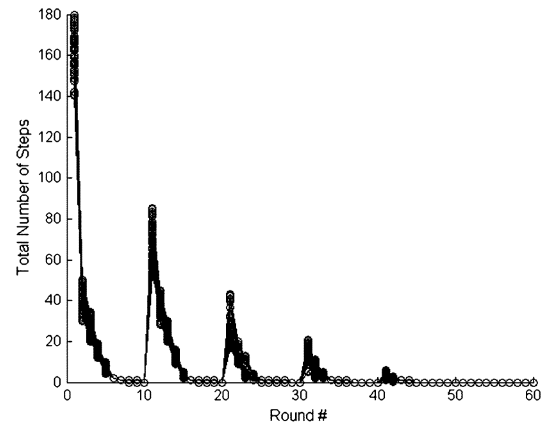

Fig. 7. Overlays and underlays interaction in long cycles.

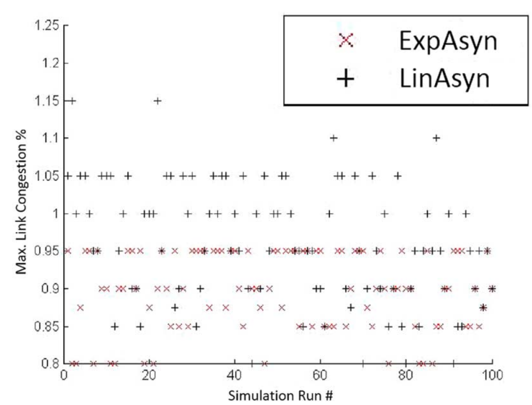

Fig. 8. Maximum link congestion level.

function results in slower convergence (higher total number of steps) than linear cost function, but in general the increase is not as significant as the theoretical worst case. Partial simultaneous moves achieve faster convergence, but the percentage of improvement varies greatly depending on the overlay topologies as expected. In fact, the total number of player moves in PS is the same as the asynchronous case under the same cost function. The difference is that a step in PS allows multiple overlays to move at the same time, while a step in the asynchronous case only allows one overlay to move. The length of each round varies depending on the number of steps in the round. Therefore, as the overlay game converges to equilibrium, the number of steps in a round should also diminish. This is shown in Fig. 5. The average number of steps (over 100 runs) in each round is plotted on the graph. The solid line depicts the game progression of unweighed game using exponential cost function, and the broken line depicts the game progression of weighted game using Shapley utility (traffic volume varying between 3–8 units for each overlay). In each case, the vertical line shows the average number of rounds for the system to reach an operable state. We confirm that it is much smaller than the game's natural convergence span, especially in the case of weighted game.

Figs. 6 and 7 show our experimental results in studying the long-term overlay–underlay interaction. The underlays perform domain-wise optimization at fixed time intervals, and the overlays attempt to optimize their routes in between these intervals. Again, 100 runs of randomized traffic setting are used. We only show the results of the first 50 runs so as not to obscure the graphs. In Fig. 6, the underlay management actions are performed in intervals of 3 rounds, which is not sufficient for the overlays to stabilize. The resulting oscillation of the overlays is evident. We see a spike in overlay rerouting activities immediately following an underlay optimization. The experiment in Fig. 7 shows the result of a 10-round underlay management interval. As the overlays are allowed to reach equilibrium before each underlay management activity, we see a clear convergence over the long term among overlays and underlays. The graph illustrates pictorially the dual-game nature of overlay and underlay interactions.

To show the usefulness of operable states in practice, we study the level of virtual link congestion after the overlay equilibrium has been reached. Fig. 8 shows that our exponential cost function results in noncongestive system configurations and in general the maximum link congestion level is much lower than that of the selfish (linear) utility function. The results appear to be layered due to the discrete workload and link capacity assignment for the simulation setup.

Finally, we investigate the nature of cooperation and competition among overlays. Fig. 9 shows two groups of simulation

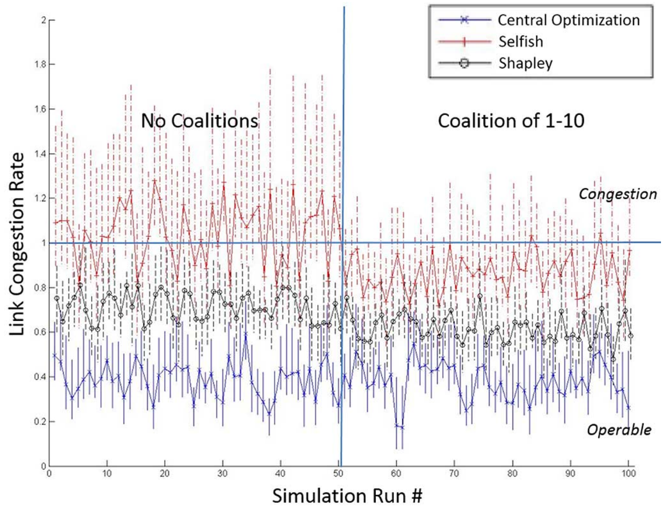

Fig. 9. Link congestion level under varied mode of cooperation.

runs (50 runs each). In the first group, no players form coalitions. The plot shows the average link congestion level as well as the ceiling and floor link congestion levels with 80% confidence. The baseline of comparison is a centralized optimizer that computes the optimal resource allocation. The effect of applying the Shapley utility function is pronounced as we can observe a significant difference in expected link congestion level among independent players when they use selfish utility as opposed to Shapley utility. The second group studies coalition scenarios where players form coalitions of random sizes in the range [1, 10]. In doing so, we aim to simulate certain overlay application scenarios such as those involving content distribution networks. They optimize their own traffic, but behave selfishly among other coalitions. Again, we observe that the expected link congestion level of Shapley utility is lower than that of selfish utility. The congestion level of selfish utility is lower than its counterpart in the no-coalition case as we expect. We also observe that the outcome produced by our Shapley utility function follows the operable state condition exactly. These results agree with our analysis and demonstrate the benefit of Shapley utility in weighted games. Moreover, the contrast of Groups 1 and 2 illustrates the nature of competition and cooperation under selfish utility and Shapley utility.

VI. Conclusion

In this paper, we have modeled the interaction between overlays and underlays in multidomain networks as a dual congestion game. We showed that stability is possible when: 1) underlay management intervals are long enough to permit overlay convergence to an operable stable state; 2) overlays have control over domain traversal; 3) overlay and underlay have aligned objectives of congestion avoidance. With care, these conditions can be satisfied in practice. Furthermore, we showed that Shapley value can be utilized to establish systemwise stability among overlays and underlays and is an economically fair and efficient mechanism for promoting cooperation in an otherwise competitive environment. As future work, we plan to design overlay routing protocols that are underlay-management-friendly, which could also be useful in emerging network virtualization environments, and to conduct further experimentations accordingly.

Appendix

Lemma A1.1: λ is lower-bounded by $O(\log(n)/\log(1+(1/L)))$ in the general case.

The bound can be obtained by investigating the worst case. Consider a set of resources $\{l_1, l_2, \dots l_t, \dots l_n\}$. According to the MINMAX property, given a strategy profile A in which $C_t(A)$ is the bottleneck link, then $u(A)$ must be strictly higher than any other strategy profile of $u(A_i)$ where the bottleneck link is of the condition: $C_i < C_t,\ \forall i$. This implies that $C_t(A)$ must be the dominating term in $u(A)$: $C_t(A) \geqslant \sum C_i(A)$ where $i \in n$ and $i \neq t$. The worst case occurs at boundary condition $C_i(A) = 1$, where $1^\lambda = 1$ for arbitrary λ value. Thus, we obtain the following worst case assuming l_t has the largest capacity L in the network and all of the other links are at capacity $l_i = 1$:

$$\left(\frac{L+1}{L}\right)^\lambda > \sum C_i > n.$$

It follows that unless the above equality holds, there exists an alternative strategy profile A' such that arbitrarily moving a player from a link in C_i to C_t in fact reduces the system potential, and thus the MINMAX property is not guaranteed. Solving the above yields a lower bound on λ

$$\lambda > \frac{\log(n)}{\log\left(1+\frac{1}{L}\right)}.$$

□

Theorem A1.2: λ is lower-bounded by $O(\log(p)/\log(1+(1/L)))$ where p is the maximum path length.

The above general case analysis assumes a player's strategy choice is the entire network. In practice, an overlay has limited number of paths with finite length. It is then sufficient to reestablish the proof of Lemma A1.1 over the set of candidate paths an overlay is interested in.

Proof by Contradiction: Assume Theorem A1.2 is not true. Then, let $\log(p)/\log(1+(1/L)) < \lambda < \log(n)/\log(1+(1/L))$, and take a boundary case where all of $C_i = 1$ and $C_t > C_i,\ \forall i$. Assigning a player k from C_i to C_t will reduce system potential under given λ. However, this move cannot occur since the player k's evaluation of $C_t(A_k)$ must then yield a higher potential than any other alternative as long as $\log(p_k)/\log(1+(1/L)) < \lambda$, where p_k is the player k's path length. Thus, we arrive at a contradiction. Therefore, $\lambda > \log(p)/\log(1+(1/L))$ is sufficient to guarantee that systemwise MINMAX property always holds. □

Theorem A2.1: The game transformation from Γ_D and Γ_D^T incurs a loss of optimality by at most 2.

The difference in the equilibriums of Γ_D and Γ_D^T is caused by the binary factoring of resources in Γ_D. Because of the min-max nature of the cost function, the minimal system potential (i.e., the game equilibrium) is in essence the minimization of the maximum resource congestion in the game space. Each strategy A_i of a player i in Γ_D is expanded into a set of strategies in Γ_D^T over the binary factoring of the resources in A_i. This expansion can potentially cause a rise in the system potential of Γ_D^T compared to Γ_D. To illustrate this effect, consider the following simple example: A resource t has capacity 5, and three players are using this resource, yielding a congestion value of 3/5 on the resource for all the players. in Γ_D^T, t is transformed into a resource set of $\{t_1, t_2\}$ with capacities 1 and 4, respectively. Allocating three players according to min-max property results in congestion levels 3/4 and 0/1. Thus, one arrives at two different system potentials. To bound this loss of optimality at equilibrium, it is sufficient to bound the loss of optimality for all mappings of resources between Γ_D and Γ_D^T. Given any strategy profile A in Γ_D, we examine the transformation process of any resource t_j in Γ_D.

Case 1: $x_j(A)/l_j < 1/2$. t_j is transformed into its binary factors in Γ_D^T. Mapping players of resource t_j to t_j^T in Γ_D^T follows the min-max property, and the resulting congestion on any member of t_j^T is upper-bounded by 1/2. Furthermore, given any congestion level of $t_j < 1/2$, there exists a satisfying mapping among the transformed resources with congestion level at most $2(x_j(A)/l_j)$ on any of the binary factors. This is due to the binary nature of the factors; whenever mapping a player to a binary resource would exceed $2(x_j(A)/l_j)$, the player is mapped to the resource of the next binary grade. For example, a congestion level of 3/7 in Γ_D is mapped to {2/4, 1/2, 0/1} in Γ_D^T. This assertion always holds as long as $x_j(A)/l_j < 1/2$.

Case 2: $1 \geq x_j(A)/l_j \geq 1/2$. This is the simple case. By the definition of operable states, mapping players of resource t_j to t_j^T in Γ_D^T is guaranteed to be bounded by 1 on any of the transformed resources. Thus, the difference in congestion level of t_j^T in Γ_D^T compared to t_j in Γ_D is at most $2(x_j(A)/l_j)$.

The above analysis is conducted assuming there exists operable states in the system. The lack of operable states indicates a resource-starved network and is outside the context of the study. Since the above cases are exhaustive, the loss of optimality is upper-bounded by $\Gamma_D^T/\Gamma_D = 2$. □

Lemma A4.1: Each move by the aggregate overlay player results in a drop in $u(A, Z)$.

The role of the aggregate overlay player is taken on by the short-term overlay game as a whole. Thus, a move in the long-term game is in fact the equilibrium configuration of the short-term overlay game. Due to existence of the system potential in the short-term overlay game, the convergence of the short-term overlay game necessitates a decrease in $u(A, Z)$ given fixed T and Z. In fact, when the minimum potential of the short-term overlay game is obtained, it is a best-response move by the aggregate overlay player. □

Lemma A4.2: Each move by the aggregate underlay player results in a drop in $u(A, Z)$.

The aggregate underlay player represents the systemwise effect of intradomain optimizations conducted by underlays. Since the nonoverlay portion of the underlay traffic is stable, the traffic matrix of each underlay domain is stable if the overlay traffic routing also is stable. As long as this precondition holds, there exists an optimal $H(Y, Z)^*$ that minimizes $u(A, Z)$. Since the interdomain traversal of an overlay is not controlled by the underlay and the overlay configurations result in operable states, the optimal $H(Y, Z)^*$ can be obtained through optimal $H(Y, Z)$ mapping in each domain. This is a best-response move by the aggregate underlay player. □

Theorem A4.3: Γ_L has a pure Nash equilibrium.

From Lemmas A4.1 and A4.2, it follows that Γ_L is a potential game, and thus the minimization of the potential $u(A, Z)$ brings the game to an equilibrium under the FIP. □

Theorem A4.4: Γ_L has a minimal potential equilibrium when the overlay game minimizes its potential and the network management solutions are domain-wise optimal.

Per Lemma A4.1, when the overlay game minimizes its potential, the aggregate overlay player effectively conducts a best-response move. Similarly in Lemma A4.2, when each network domain minimizes its congestion, it corresponds to a best-response move by the aggregate underlay player. Thus, in a two-player potential game, following the best-response moves naturally leads to a game solution that minimizes the potential $u(A, Z)$ [3]. This is the optimal system configuration that minimizes the congestion level of the networks. □

References

[1] A. Orda, R. Rom, and N. Shimkin, "Competitive routing in multiuser communication networks," *IEEE/ACM Trans. Netw.*, vol. 1, no. 5, pp. 510–522, Oct. 1993.

[2] E. Altman, T. Basar, T. Jimenez, and N. Shimkin, "Competitive routing in networks with polynomial cost," in *Proc. IEEE INFOCOM*, 2000, pp. 1586–1593.

[3] I. Milchtaich, "Congestion games with player-specific payoff functions," *Games Econ. Behav.*, vol. 13, pp. 111–124, 1996.

[4] A. Fabrikant, C. Papadimitriou, and K. Talwar, "The complexity of pure Nash equilibria," in *Proc. 36th ACM STOC*, 2004, pp. 604–612.

[5] D. Monderer and L. Shapley, "Potential games," *Games Econ. Behav.*, vol. 14, pp. 124–143, 1996.

[6] Y. Liu, H. Zhang, W. Gong, and D. Towsley, "On the interaction between overlay routing and underlay routing," in *Proc. IEEE INFOCOM*, 2005, vol. 4, pp. 2543–2553.

[7] S. Seetharaman, V. Hilt, M. Hofmann, and M. Ammar, "Preemptive strategies to improve routing performance of native and overlay layers," in *Proc. IEEE INFOCOM*, 2007, pp. 463–471.

[8] S. Seetharaman, V. Hilt, M. Hofmann, and M. Ammar, "Resolving cross-layer conflict between overlay routing and traffic engineering," *IEEE/ACM Trans. Netw.*, vol. 17, no. 6, pp. 1964–1977, Dec. 2009.

[9] L. Qiu, Y. R. Yang, Y. Zhang, and S. Shenker, "On selfish routing in internet-like environments," *IEEE/ACM Trans. Netw.*, vol. 14, no. 4, pp. 725–739, Aug. 2006.

[10] R. Keralapura, C.-N. Chuah, N. Taft, and G. Iannaccone, "Race conditions in coexisting overlay networks," *IEEE Trans. Netw.*, vol. 16, no. 1, pp. 1–14, Feb. 2008.

[11] H. Zhang, J. Kurose, and D. Towsley, "Can an overlay compensate for a careless underlay?," in *Proc. IEEE INFOCOM*, Apr. 2006, pp. 1–12.

[12] R. T. B. Ma, D. Chiu, J. C. S. Lui, V. Misra, and D. Rubenstein, "Internet economics: The use of Shapley value for ISP settlement," *IEEE/ACM Trans. Netw.*, vol. 18, no. 3, pp. 775–787, Jun. 2010.

[13] H. Yaiche, R. R. Mazumdar, and C. Rosenberg, "A game theoretic framework for bandwidth allocation and pricing in broadband networks," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 667–678, Oct. 2000.

[14] R. Banner and A. Orda, "Bottleneck routing games in communication networks," in *Proc. IEEE INFOCOM*, Apr. 2006, pp. 1–12.

[15] K. Kollias and T. Roughgarden, "Restoring pure equilibria to weighted congestion games," in *Proc. ICALP*, Jul. 2011, pp. 539–551.

[16] R. Rosenthal, "A class of games possessing pure-strategy Nash equilibria," *Int. J. Game Theory*, vol. 2, pp. 65–67, 1973.

[17] P. Goldberg, "Bounds for the convergence rate of randomized local search in a multiplayer load-balancing game," in *Proc. 25th ACM PODC*, 2004, pp. 131–140.

[18] E. Even-Dar and Y. Mansour, "Fast convergence of selfish rerouting," in *Proc. 16th ACM-SIAM SODA*, 2005, pp. 772–781.

[19] G. Christodoulou, V. Mirrokni, and A. Sidiropoulos, "Convergence and approximation in potential games," in *Proc. STACS*, 2006, pp. 349–360.

[20] S. Chien and A. Sinclair, "Convergence to approximate Nash equilibria in congestion games," in *Proc. ACM-SIAM SODA*, 2007, pp. 169–178.

[21] C. Alvarez and G. Frances, "The complexity of pure Nash equilibria in weighted max-congestion games," Universitat Politecnica de Catalunya, Barcelona, Spain, Tech. Rep., 2008 [Online]. Available: http://hdl.handle.net/2099.1/5589
[22] C. Papadimitriou, "Algorithms, games and the Internet," in *Proc. 33rd ACM STOC*, 2001, pp. 749–753.
[23] T. Roughgarden and E. Tardos, "Bounding the inefficiency of equilibria in nonatomic congestion games," *Games Econ. Behav.*, vol. 47, pp. 389–403, 2004.
[24] B. Awerbuch, Y. Azar, and A. Epstein, "The price of routing unsplittable flow," in *Proc. 37th ACM STOC*, 2005, pp. 57–66.
[25] T. Roughgarden and E. Tardos, "How bad is selfish routing?," in *Proc. 41st Annu. Symp. Found. Comput. Sci.*, Nov. 2000, pp. 93–102.
[26] T. Roughgarden, "Selfish routing with atomic players," in *Proc. ACM-SIAM SODA*, 2005, pp. 1184–1185.
[27] R. Mahajan, M. Rodrig, D. Wetherall, and J. Zahorjan, "Experiences applying game theory to system design," in *Proc. ACM SIGCOMM Workshops*, 2004, pp. 183–190.
[28] L. S. Shapley, "Additive and non-additive set functions," PhD. dissertation, Dept. Math., Princeton Univ., Princeton, NJ, USA, 1953.
[29] X. Yang, D. Clark, and A. Berger, "NIRA: A new inter-domain routing architecture," *IEEE Trans. Netw.*, vol. 15, no. 4, pp. 775–787, Aug. 2007.
[30] K. Lakshminarayanan, I. Stoica, and S. Shenker, "Routing as a service," UC Berkeley, Berkeley, CA, USA, Tech. Rep. UCB/CSD-04-1327, 2004.

Jin Xiao received the Ph.D. degree in computer science from the University of Waterloo, Waterloo. ON, Canada, in 2010.

He is currently a member of research staff with the IBM T. J. Watson Research Center, Yorktown Heights, NY, USA. His research interests include network and service management, service automation, network economics, smart grid, smart home, and smartphone energy management.

Raouf Boutaba (M'93–SM'01–F'12) received the M.Sc. and Ph.D. degrees in computer science from the University Pierre and Marie Curie, Paris, France, in 1990 and 1994, respectively.

He is currently a Professor of computer science with the University of Waterloo, Waterloo, ON, Canada. His research interests include control and management of networks and distributed systems.