Research Statement

Jacomo Corbo
School of Engineering and Applied Sciences
Harvard University
http://www.eecs.harvard.edu/~jacomo
jacomo@eecs.harvard.edu

Networks permeate our social and economic lives. Consumers benefit from the research of friends and family into new products. In medicine and other technical fields, professional networks shape research patterns. In industry, it has been long posited that research findings spill over to other firms. Peer-to-peer file-sharing networks provide online markets for media content that rival traditional media sources for total share of consumption. The Internet itself consists of a loose federation of independent service providers. In all these settings, the well-being of participants depends on social, geographic, or trading relationships. The countless ways in which network structures affect our well-being make it critical to understand: (i) how network structures impact behavior, (ii) how the rules or processes by which network structures are formed affect the structures that result, (iii) what can be done, in the way of design by stakeholders or policy, to improve systemic outcomes. This area of study, broadly called network economics, in the emerging field of network sciences, is at the heart of my research interests.

Network economics is a fast growing area of study, receiving attention not only within the economics and business school communities, but also in relation to algorithmic game theory, distributed computing, and computational mechanism design within computer science. Partly, this reflects a growing body of empirical work in a multitude of settings suggesting definite structure to the pattern of externalities between agents. Partly, this is due to new opportunities for community formation and connectivity afforded by the Internet. In large part, growing interest in network analysis of social, organizational, and economic settings stems from a desire to influence behavior, provide robustness, or improve efficiency. For example, firms want mechanisms that facilitate the propagation of best practice or knowledge among employees, while the explosion of social networking sites has spurred the development of viral marketing practices. The large and very diverse class of network-based problems makes this a rich of area of study, while also requiring a flexible approach to modeling and a familiarity with techniques from a variety of fields, from social and random network theory to mechanism design to algorithmic game theory.

My own research in this area has spanned a broad range of problems. For example, in my dissertation I propose a generative model of the Internet inter-domain routing graph driven by best-response dynamics of service providers (ISPs) that is the first economically-principled model to yield graphs that share many of the structural features of the actual Internet graph. I also study the interactions between production effort and the creation of synergies that are at the heart of technological collaboration agreements. In particular, the model proposed provides an explanation for the large increase in agreements of collaboration in R&D in the recent past, while also explaining the documented decline in social capital. Altogether, my dissertation focuses on mico-economic models of local interaction problems and the derivation of mechanisms and policies to improve system-wide efficiency.

For the last two years I have also held the position of Chief Race Strategist with the ING Renault F1 Team Ltd. In my capacity, I have led the design and implementation of the Company's predictive race strategy software tools. As part of this work, I have developed strategy optimization algorithms predicated on a game-theoretic description of Formula One (F1) race, using historical data both to inform the equilibrium prescription and to calibrate the utility functions of opponent teams. In a similar vein, I have developed a game-theoretic algorithm to solve an R&D portfolio selection problem for The Boeing Company's advanced research and development unit, Phantom Works. Again, the work involved extensive analysis of empirical data. Notably, the algorithm outperforms Boeing's own in-house algorithms under appropriate market conditions. This work, and my work at Renault F1, have both fostered an appreciation for—and reinforced the imperative of—good empirical analysis, and increased my knowledge of how to broach data-driven projects.

In summary, network economics, and more broadly the network paradigm, have done much in recent years to uncover the mechanisms at play in many phenomena seen in and between firms, organizations, and in social communities. The general goal of my research is to better understand the structure and dynamics of social and economic interaction, with a focus on the Internet and public good provisioning, collaboration networks, and knowledge diffusion in organizations. To that end, I am interested in a number of related topics, including the structure and evolution of networks, the impact of incentives and policy on the efficiency of network structures and processes, and indirect mechanisms to help improve the efficiency of interactions.

I am excited to pursue a research agenda in these areas. Such lines of inquiry are increasingly indispensable as both firms' and individuals' interactions become increasingly complex, numerous, and short-lived, and firms and markets become ever larger, less hierarchical, and more decentralized. This research promises to revolutionize how groups behave, how firms collaborate on joint ventures, how individuals communicate within an organization, how companies are organizationally structured, and even how technological collaborative aids are designed.

Summary of Current Work

1. Towards a Theory of the Internet Inter-Domain Routing Graph

End-to-end packet delivery in the Internet is achieved through a system of interconnections between independent entities called Autonomous Systems (ASes). As of March 2007, the Internet consisted of over 26,000 ASes. While most ASes are Internet service providers (ISPs), the term also refers to enterprises, governmental or educational institutions, and increasingly large content providers with mostly outbound traffic such as Google, Yahoo, and YouTube as well as overlay content distribution networks such as Akamai and Limelight. Each AS controls or administers its own domain of addresses but ASes must physically interconnect to provide end-to-end connectivity across the Internet. Interconnection is not only important from a reachability perspective but also from quality and performance perspectives, because how ASes interconnect, both physically and contractually, determines how packets are routed and impacts the quality and choice of services that may be supported.

The initial pattern of AS interconnection in the Internet was relatively simple, involving mainly ISPs with a balanced mixture of inbound and outbound traffic. In the last 10 years however, the pattern of AS interconnections has become much more diverse and complex, as has the associated contracts governing linkages. The network graph today possesses many complex features like a small average diameter and a power law degree distribution that are critical to its good performance, and yet are poorly explained by most models of AS graph formation. Indeed, there is a strong disconnect in the literature between utility-driven network formation models and random graph generative models: the former capture the incentives of stakeholders in the network, but fail to reproduce realistic graphs; the latter purport to reproduce graphs with aggregate statistics similar to those found in the actual AS graph but then have no explanatory power for the effect of market conditions on the resultant topology (since these models do not capture incentives).

Joint work with Jain, Mitzenmacher, and Parkes [3] has yielded the first economically-principled model to generate graphs with the many of the optimizing features of the actual AS graph, like high throughput and low congestion properties, as well as aggregate statistics like low diameter, high clustering, and a power law degree distribution. The formulation of AS utility includes revenue from an AS's own generated demand for traffic, congestion and routing costs, as well as transfers to and from provider and customer ASes, respectively. The model has the following features: it uses an empirically-motivated model of traffic demand which considers the variation in demand with ASes' business models and the graph of business relationships; it allows for nodes to revise their connections over time; a node's utility explicitly models many of the major economic and technological issues at play. Notably, the model accurately predicts the location of peering (transfer-free) links in the graph, and shows that peering is susceptible to asymmetric flow, a vulnerability that has been corroborated by empirical studies. I completed an earlier investigation on the optimal location of peering points between ASes, in collaboration with Petermann [2], which showed the same sensitivity.

The generative model in [3] also suggests that there is a relative paucity of peering points in the AS graph, in relation to the total welfare-maximizing solution. The under-provisioning of peering links stems from the requirement of bilateral consent from endpoint ASes to establish the link. I showed in an earlier game-theoretic model of AS graph formation [1] that network formation under bilateral consent can lead to topologies that are strictly worse than would otherwise form if links are laid down unilaterally. This paper is significant for being among the first to reconcile and compare different network formation procedures in the literature with different equilibrium outcomes and quantifying the impact of these procedures on the efficiency of resulting network structures.

2. Spillovers, Production and Synergies in Networks

External effects (spillovers) pervade economies and societies in general. Both inter- and intra-industry cross-interaction between firms have been the object of many studies. I have studied productive efforts in settings where the external effects are both exogenous and endogenous.

When the pattern of externalities between agents is fixed, I have examined the incentive to provide goods that exhibit strategic complementarities, where efforts to contribute are mutually reinforcing, or substitutabilities, where they offset each other [4]. The model admits linear quadratic utilities with very general interdependencies and I show how to formulate a network game from these. Specifically, such utility functions features in models of how job opportunity information travels along friendship links, models for the provisioning of content in a peer-to-peer file-sharing network (which is the primary working example in this work), and innovation diffusion models. I provide a characterization of equilibrium contribution levels in terms of a measure of social importance in a network, called the Katz-Bonacich index, and discuss the viability of various equilibria by showing which are stable to a perturbed best-response dynamic. In particular, I show the persistence of free-riding behavior in a game with strong strategic substitutes and show that its likelihood is tied to an actor's position in the network.

Still in the context of linear quadratic utilities, I consider the case when participants can manipulate both productive efforts and spillovers [5]. Here the interdependencies between actors vary with the strength of the synergistic linkage across different pairs of agents. The main innovation of this study is, precisely, that the synergistic effort is a computationally tractable decision. Particularly in large networks, which are the focus of my study, socializing is not equivalent to elaborating an exhaustive list of intended relationships, as in much of the literature on socio-economic networks. In spite of its parsimony that ensures tractability, the model retains enough richness to replicate a relatively broad range of empirical regularities displayed by social and economic networks, and is directly estimable to recover its structural parameters. For example, in homogeneous populations, the model predicts the emergence of a giant component in the emergent graph, a feature that is documented from a number of empirical studies.

I also investigate the use of network-based policies for improving the equilibrium performance of networks [4]. The mechanisms investigated belong to the class of indirect mechanisms, which I consider because of the limited scope for intervention by a social planner in the motivating applications. One of these policies, the key player policy, is a network disruption strategy based on the exclusion of a single agent. In the game with exogenous spillovers, I show that this policy is a provably optimal mechanism to mitigate the discrepancy between best- and worst-case equilibrium performance when confronted by graphs that admit multiple equilibria.

3. A Game-Theoretic Analysis of Race Strategy in Formula One (F1)

The difference between success and failure in F1 can measure in the hundredths of a second. A poor decision on the part of either driver or pit crew is punished immediately and irrevocably; a good decision can lead to victory. Race strategy, which includes a team's choice of when to call its drivers into the pits and how much fuel to add, impacts heavily on many aspects of a team's performance: it strongly determines the car's optimal aerodynamic and mechanical set-up and so its handling and drivability at different stages of the race, the amount of fuel onboard in qualifying and so the driver's starting grid position, his pace relative to his direct competitors throughout the race, even his expected, best and worst finishing position possibilities. On the back of a new qualifying format introduced for 2006, strategy has become even more critical to the outcome of an F1 race and has quickly become one of the most technologically sophisticated aspects of the sport.

For the last two years, I have worked as Chief Race Strategist for the ING Renault F1 Team Ltd. In my position, I have led the design and development of Renault F1 Team's strategy software infrastructure. To that end, I have developed game-theoretic and financial options calculations to assess the optimality and corresponding risks of race strategy decisions. In particular, I have formulated a game-theoretic description of an F1 race based on a variation of ranking games closely related to the class of multi-player competitive games known as *unilaterally competitive games* [6]. An important feature of our formulation is that the *price of cautiousness*, defined as the worst-possible loss an agent may incur by playing his maximin (security level) strategy instead of the worst (quasi-strict) Nash equilibrium, is bounded, which is not the case in general ranking games. In fact, in practice we find that under certain conditions, maximin will yield strategies that are very close to optimal, thereby suggesting a polynomial-time routine to compute one's strategy in the face of the intractability of computing a Nash equilibrium strategy.

As part of this work, I have also used empirical data to parameterize the relative risk aversion of competitors based on historical data, a technique that can be applied to many settings where sufficient data is available, notably to portfolio selection problems. The application of game-theoretic techniques to predictive modelling of complex endogenous phenomena is relatively rare, but proved remarkably effective in the case of optimizing race strategy. The success of my approach subsequently motivated The Boeing Company to invite me to work on a portfolio selection problem for them (discussed immediately below).

4. Portfolio Selection in High Technology Industry

In this work [7], performed for The Boeing Company's Phantom Works division while working at the ING Renault F1 Team Ltd., I studied the use of a game-theoretic approach to solving an R&D portfolio selection management problem. Assuming a fixed budget, the management problem is twofold: first, decide which projects to start, continue or terminate; second, decide how much to invest in each project holding. A game-theoretic approach is motivated by the typically unsatisfactory performance of real-options approaches in settings with very few stakeholders, long development cycles, and very uncertain returns. The latter conditions are emblematic of the aerospace industry and other high-technology industries, and exhibit a number of (approximately) zero sum conditions. I developed a game-theoretic description of the problem by formulating a multi-player competitive game. More precisely, I consider a variation (properly a subset) of the well-studied class of *almost strictly competitive games*.

Under the market conditions loosely described above, the optimization routine predicated on this game-theoretic model outperforms Boeing's own in-house algorithms both in terms of expected return and volatility in the revenue stream.

Future Work: Structural Factors in Organizational Performance and the Effective Use of Collaborative Technologies

A number of empirical studies have established that innovation and the propagation of best practice within the firm travels along organizational structural lines. The conclusions of this work generally suggest that relatively flat or well-connected organizational arrangements are more effective at the dissemination of information. Yet such demands place a strain on the organization's information-processing requirements. Indeed, variables in the design of organizations, such as structural self-containment and vertical information systems, should provide sufficient capacity to meet the organizational work unit's information needs. I am interested in studying how organizational structure affects knowledge diffusion and organizational learning, and how both knowledge propagation processes and organizational structure can be optimized against the competing requirements of the firm's business.

These problems naturally admit a network-based formulation, which is significant for providing a single coherent paradigm in which to study a broad array of organizational phenomena as well as their embedded optimization problems. Furthermore, networks, dynamic processes on networks, and network design are very well studied in many fields, notably economics and computer science; this suggests a large literature to borrow from in the derivation of models and optimization policies.

One of the most compelling facets of this research agenda involves the effective deployment of collaborative technology to help with information aggregation and exchange. Insofar as these technologies manipulate the flow of information within the firm, I view these technologies as potentially powerful instruments for fostering flexible, responsive, and efficient organizational structures. Collaborative technologies can help with information exchange and broadcast as well as with mitigating processing requirements by aggregation. My research will address effective information technology strategies and how these may vary depending on the organizational structure and information-processing requirements of the firm.

I am also interested in exploring the potential of information technology and electronic communications data, such as email, to resolve some of the measurement difficulties associated with studying human interactions and organizational dynamics.

References

[1] J. Corbo, D. C. Parkes. "The Price of Selfish Behavior in Bilateral Network Formation," In Proceedings of the 24th ACM Symposium on Principles of Distributed Computing (PODC), Las Vegas, Nevada, USA, pp. 99-107, July 2005.

[2] J. Corbo, T. Petermann. "Selfish Routing and Peering in the Internet," In Eric Jen, editor, Proceedings of the Santa Fe Institute School of Complexity Series, Santa Fe, New Mexico, USA, vol. 14, p. 141-147, August 2004.

[3] J. Corbo, S. Jain, M. Mitzenmacher, D. C. Parkes. "An Economically-Principled Model of the Internet Topology" Harvard EECS Technical Report Series, TR-17-07, December 2007.

Earlier versions of this work appeared as:

J. Corbo, S. Jain, M. Mitzenmacher, D. C. Parkes. "An Economically Principled Generative Model of AS Graph Connectivity," In Proceedings of the International Joint Workshop on The Economics of Networked Systems and Incentive-Based Computing, June 2007.

[4] J. Corbo, A. Calvó-Armengol, D. C. Parkes. "Network Effects in Local Contribution Economies: Identification and Regulation," UAB Working Paper Series 413, Universitat Autonoma de Barcelona, October 2007.

Earlier versions of this work appeared as:

J. Corbo, A. Calvó-Armengol, D. C. Parkes. "A Study of the Nash Equilibrium in contribution games for peer-to-peer networks," In ACM SIGOPS Operating Systems Review (40) 3, July 2006, pp.61-66.

J. Corbo, A. Calvó-Armengol, D. C. Parkes. "The Importance of Network Topology in Local Contribution Games," In X. Deng and F. Chung Graham, editors, Proceedings of the International Workshop on Internet and Network Economics (WINE), Lecture Notes in Computer Science (LNCS). Springer-Verlag, December 2007.

[5] J. Corbo, A. Calvó-Armengol. "Investment Decisions in the Creation of Synergies: Lessons from Socio-Economic Networks" UAB Working Paper Series 382, Universitat Autonoma de Barcelona, October 2007.

[6] J. Corbo, D. C. Parkes. "A Game-Theoretic Analysis of Race Strategy in F1: Lessons For the Real World" Renault Technical Report Series 104921a, December 2007.

[7] J. Corbo. "R&D Portfolio Selection as the Solution to a Strictly Competitive Game" Renault Technical Report Series 104937, December 2007.