

RESEARCH STATEMENT

Victor Amelkin

The **central theme** of my research is **network science**, with an emphasis on **social and economic networks**, viewed from the perspectives of **analysis, modeling, and control**. My work spans applications and methods from Computer Science (algorithm design and analysis, data mining, applied machine learning), Systems and Industrial Engineering (dynamical systems, applied linear algebra), Operations Research, Management Science, and Economics (optimization, game theory, optimal decision making, probability), as well as Computational Social Science (socio-psychological models). In this statement, I describe my contributions to network science and outline my vision of the evolution of the cross-disciplinary studies of social and economic networks.

1 DATA-DRIVEN MODEL-BASED SOCIAL NETWORK ANALYSIS

PAST WORK [1, 2]: Two general types of analyses useful with social network data are **anomaly detection** (e.g., detecting when the evolution of user opinions drastically deviates from the expected word-of-mouth dynamics, potentially indicating the start of a viral marketing campaign); and **future state prediction** (e.g., prediction of how user opinions will change in the future). Straightforward domain-oblivious approaches towards these analyses—such as a domain-independent feature extraction followed by anomaly detection or extrapolation in the obtained vector space—are simple yet not promising, since, to tell whether an observed change in the state of a network process is anomalous, we need a domain-specific definition of what is “expected”, that is, **the data-driven analyses should rely upon a model**.

The key question, answering which would enable above mentioned analyses is, **given two states of a network process, what is the likelihood that the process has transitioned between these states under a given state evolution model?** For example, assuming that the spread of opinions through a social network follows a fitted instance of the Independent Cascade model, we are after the likelihood of the transition between two observed user opinion distributions, where the pathways for the opinion evolution are determined by the model and the network. The exact computation of this likelihood—or the distance between two network process states—is unfeasible. An estimate, however, can be efficiently computed, as shown in my past work.

I have designed a **method to estimate the likelihood of a network process’ switching between two states** in the context of opinion dynamics in large social networks. This method—a distance measure for network process states termed Social Network Distance (SND)—computes the likelihood of the most likely transition between two observed large user opinion distributions. SND reduces the state transition likelihood to the solution of a transportation problem or, alternatively, Earth Mover’s Distance. While the direct solution of the obtained transportation problem would take (super-)cubic time, SND exploits the special structure of the transportation problem and effectively uses a combination of radix-Fibonacci heap-based Dijkstra algorithm and bi-push min-cost network flow algorithm for unbalanced bipartite graphs, resulting in SND’s computability in time linear in the network’s size. In experiments with Twitter data, SND **has shown to be effective at detecting polarizing events**—such as the introduction of “Obamacare” in the US in 2010; it has also shown to be **effective at predicting future opinions of a small subset of network users**.

FUTURE WORK: There is a cornucopia of research opportunities in model-driven analysis of network data. One pertinent general problem in economic network analysis is **connecting theoretical economic models with real-world data**. For example, given the knowledge of what supply chain networks arise in theory (see Sec. 2), we should also ask to what extent the structure of real-world supply chains follows the structure prescribed by the models. Besides buttressing the models, it would tell us to what extent the assumptions upon which the models are built are realistic. Additionally, while most supply chain models are static, it is of even more importance to **account for the dynamics of supply chain network formation**. Validating such theoretical models with data has an obvious value, yet, comes with a great challenge—many real-world supply chain networks evolve slowly, and analyzing dynamics of such networks would require—besides new analysis methods accounting for temporal data sparsity—new datasets obtained through collaboration with industry.

Another open problem in the analysis of network data is **efficient forecasting of the future state of an observed network process**. Consider a situation when, having observed how the political opinions of a social network’s users have evolved in the past, we want to predict how these opinions will change at the time of the upcoming elections. While I have addressed this application in the past [1, 2] through estimating the likelihood of a specific change

in a network process' state, the bottleneck of the designed method is the need to assess a large number of future network state candidates. To tackle it, we need to rely upon the domain-specific structure of the state space of the network process, and find a way to effectively prune the state space, for example, by exploiting semi-metricity of the distance measure used in the analysis of the network state series.

RELATED PAPERS:

- [1] V. Amelkin, P. Bogdanov, and A. K. Singh. A distance measure for the analysis of polar opinion dynamics in social networks. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 13(4), 2019. [pdf], [doi], [slides].
- [2] V. Amelkin, P. Bogdanov, and A. K. Singh. A distance measure for the analysis of polar opinion dynamics in social networks. In *Proc. International Conference on Data Engineering (ICDE)*, pages 159–162. IEEE, 2017. [pdf], [doi], [poster].

2 ECONOMIC NETWORK MODELS AND RESILIENCE

PAST WORK [3, 4]: Supply chains are the backbone of the modern economy, yet, they suffer from uncertainty and, in particular, production disruptions, which can be very costly. The latter stems in part from the degree of uncertainty¹. However, and perhaps more importantly, as uncertainty and disruptions can cascade through a supply chain network, complex **supply chain networks have an inherent ability to amplify inefficiency**. Thus, a fundamental question raised in supply chain management is how to structure a supply chain network to guarantee its high performance. The question posed this way implies a centrally planned supply chain. An efficient implementation of the latter, however, requires a high degree of coordination and awareness among the supply chain members, which may be difficult to achieve in practice, as indicated, for example, by the presence of bullwhip effect in real-world supply chains. Thus, I am interested in **how efficient supply chain networks resilient to disruption can be strategically formed by rational self-interested agents in the absence of coordination**. To that end, I have contributed two supply chain network formation models, in both of which a supply chain can be viewed, for simplicity, as a bipartite network of (downstream) retailers and (upstream) suppliers.

In the first model [3], suppliers, having uncertain supplies, strategically announce wholesale prices, and retailers, competing à la Cournot, respond by strategically forming links with suppliers. I show that, in this environment, **retailers insufficiently diversify their supplier bases at pure strategy Nash equilibrium**, leading to inefficiency. I also show how different improvements of supplies affect supplier and consumer welfare.

In the second model [4], both suppliers and retailers are subject to production failures and competition, and market prices are formed under the market clearance assumption. Competing suppliers are subject to yield uncertainty and congestion—which can also be interpreted as a soft constraint on supply capacity—while competing and also unreliable retailers must decide which suppliers to link to based on both price and reliability. I have characterized the structure of the formed pure strategy Nash equilibrium networks and shown that, **in the presence of yield uncertainty only, the resulting supply chain networks are sparse, with retailers' concentrating their links on a single supplier**, counter to the idea that retailers should mitigate yield uncertainty through supply base diversification. This suggests that **competition among quantity-controlled markets will amplify output uncertainty**². **When congestion is added, the resulting equilibrium networks resemble bipartite expander graphs that have been known to serve as resilient supply chain networks** from the central network planning perspective. I also show that a supplier's investments in improved yield can make it worse off, as price-minimizing retailers can undersaturate the market with product.

FUTURE WORK: In two above described supply chain models, retailers strategically link to suppliers, yet set the quantities matching their downstream demands. The next natural question to target—assuming that retailers' strategies now include both links and quantities—is **what should be the optimal balance between boosting order quantities vs. signing redundant contracts when faced with production uncertainty?** The corresponding supply chain model will be technically complex, being similar to a k -partite network version of the newsvendor model, yet, will provide a valuable managerial insight with an immediate impact upon supply chain engineering.

Another related question currently being addressed in my ongoing research is **how the structure of a network of firms in the presence of economic shocks affects the firms' value distribution at equilibrium, as well as contagion and**

¹For example, in digital hardware industry, production yield can swing to as low a value as 50%.

²Market quantity-based control's driving output uncertainty up complements the insight from Weitzman's "Prices vs. Quantities" (1974) about quantity control's being more advantageous when production uncertainty is large.

systemic risk in such networks, answering which will provide an insight for policy makers into how and whether to perform firm restructuring or buyouts in the face of potential cascading bankruptcies, thereby, **improving economic systems’ resilience to individual firms’ failures**. More specifically, I focus on characterizing the dynamics of the financial well-being distribution in “typical” instances of random networks of firms owning each other’s equity shares. A firm’s value depends, besides the firm’s own cash reserves, on the dividends from its investments in other firms. A firm whose value falls short experiences a shock, affecting its value; these shocks can propagate through and be amplified by the network. The goals are to characterize the set of equilibrium points³ as a function of the model’s parameters and the network’s structure; as well as characterize the dynamics of firm values towards these equilibria. My preliminary results show how the set of equilibrium points changes with the firms’ cash reserves in a class of regular networks, whose behavior is tightly connected to that of the typical instances of random networks; and I am currently finalizing the characterization of firm value dynamics towards equilibria. **The main insight that can already be drawn from these results is that economic networks most often function in a safe regime, when contagion is unlikely, and there is no systemic risk.** The specific conditions under which systemic risk in financial networks emerges are upcoming.

RELATED PAPERS:

- [3] V. Amelkin and R. Vohra. Yield uncertainty and strategic formation of supply chain networks. *In submission to OPERATIONS RESEARCH (July, 2019)*. [pdf], [arxiv].
- [4] V. Amelkin and R. Vohra. Strategic formation and reliability of supply chain networks. *In submission to MANAGEMENT SCIENCE (September, 2019)*. [pdf], [arxiv].

3 SOCIAL NETWORK MODELS

I aim to design non-linear models for social processes relying on first principles and amenable to theoretical analysis.

PAST WORK [5]: I have proposed a class of non-linear opinion formation models, having incorporated **dynamic opinion-dependent susceptibility to persuasion** into the opinion formation process⁴. These models are particularly suitable for the case of polar opinions—such as pro-iOS vs. pro-Android—when people’s shifting towards extreme opinions makes them less or more malleable. For example, a person shifting towards the Republican ideology is harder to persuade to invert her or his political stance. Alternatively, in a society with strong social norms, people holding conservative opinions, following the norms, are harder to convince, while extreme opinions are volatile. The main theoretical question here is **how the opinion distribution in the society will evolve in a long term?** I have answered this question for the above defined class of non-linear models, having brought up the tools from non-smooth analysis and non-linear control theory. One finding is an analytical expression for the asymptotic opinion distribution over the users of an online social network—dependent upon the structure of the network as well as the placement of “stubborn” agents in it, yet independent of the initial opinions of susceptible users. Another qualitative finding is that, as long as a person is at least to some extent susceptible to persuasion, that specific extent does not affect the long-term opinion adoption behavior of that person.

FUTURE WORK: The above described models assume that interpersonal relationships between people do not change with time (even though, people can become more or less susceptible to persuasion over time). In real-world online social networks, however, links and their quality can rapidly change based on the users’ published content. Hence, the question is **what can be said about the dynamics of user opinions in an online social network if the individual links are allowed to change?** While that general a model would not be tractable, one potentially tractable specialization is a model where social ties evolve in time based on people’s positions in the society. For example, a rising politician’s opinion can gain weight even outside of his or her network neighborhood solely due to that person’s growing reputation. I aim to design a non-linear opinion formation model, where all social ties in the network continuously evolve based upon the agents’ centrality; and where the agents weight their social ties based on locally available information. Such a model would be a good fit for real-world large-scale social networks, which are inherently incomplete, and where members make decisions based on a small observed part of the network. The theoretical analysis of this model will combine the well-established dynamical system theory together with

³The model has a large number of equilibrium points due to its non-linearity.

⁴The origins of the idea that susceptibility to persuasion is a function of the held opinion can be traced back at least to Abelson’s 1964 work “Mathematical models of the distribution of attitudes under controversy”, and has been periodically [revisited](#) since then. However, a theoretically analyzed network model had been absent prior to [5].

my recent theoretical results connecting the change of the structure of a social network to the resulting network’s users’ centrality change (see Sec. 4).

Another question to address in the future is **how to mix information diffusion and rational decision-making in the same model?** Consider an organization, where [decisions can be made based on the practices that historically showed to be successful](#), making the decision adoption process similar to the process of opinion acquisition in a social network. Yet, these decisions—having an immediate impact upon decision-makers’ careers—are clearly made under rational assessment of their potential economic consequences. Understanding the balance between diffusion and rational decision-making in both theory and data is not only technically challenging pursuit, but also an avenue towards a better organization management.

RELATED PAPERS:

- [5] V. Amelkin, F. Bullo, and A. K. Singh. Polar opinion dynamics in social networks. *IEEE Transactions on Automatic Control (TAC)*, 62, 2017. [\[pdf\]](#), [\[doi\]](#), [\[poster\]](#), [\[slides\]](#).

4 NETWORK CONTROL

A deep embedding of network processes in our lives makes it important to be able to control them. How to prevent malicious interventions in the process of information adoption in online social networks? How to assess the limits of control of online social networks over the public opinion and how to ensure that such control cannot harm the society? How to ensure that an economic system is functioning efficiently? Below, I discuss how I address these questions in my past and planned future work.

PAST WORK [6]: As online social networks have pervaded our lives, it is lucrative for marketing and political technologists to tap into these networks, aiming to shift the public opinion distribution according to certain business objectives. If we assume that an external adversary selectively influences a subset of nodes in the network—with the goal for this influence to subsequently efficiently spread through the network—we may ask a question of **how an online social network can defend itself against malicious external attacks upon the opinion distribution**. I recently proposed [6] a method to address that problem. More specifically, I provided theory and efficient algorithms that an online social network can use to recover the (eigenvector) centrality-weighted average user opinion from external influence maximization-like attacks by using the friend recommendation infrastructure: having recognized a malicious change of the average opinion, the method finds—in pseudo-linear time—which links to strategically recommend to network users (in addition to the rest of the links recommended for non-strategic reasons) to return the affected average opinion back to its state before the attack. The technical core of the method is a perturbation analysis that relies on the theory of Markov chains and shows how eigenvector centrality and, hence, the problem’s objective function, change in response to a network’s augmentation with new links.

FUTURE WORK: One conclusion of the above described study [6] is that a mechanism as seemingly innocuous as friend recommendation in online social networks may have a dramatic impact upon the formation of public opinion. It raises a whole set of questions: **How can we make sure that the link recommendation is not (even unintentionally) misused** by online social networks, and the process of information spread through the network is “fair”? Can a small number of sensor nodes placed in the network with the purpose of monitoring link recommendation succeed at detecting such misuse? More generally, **what are the limits of control that online social networks have over the user opinions?** For example, if we assume a limited ability of online social networks to *strategically* compose “newsfeeds”, to what extent the network can influence the users’ opinion distribution?

Complementary to the questions of controlling social networks, there is a multitude of questions related to the control—or optimal (mechanism) design—of economic networks. **How can we make sure that the supply chain networks formed in an uncoordinated fashion are both efficient and resilient to disruptions? How to ensure that, in an economic network of firms, linked through equity share cross-holding, a small number of firm bankruptcies will not trigger the collapse of the entire economy**(e.g., which firms should be restructured or marked as “risky” by the regulator)?

RELATED PAPERS:

- [6] V. Amelkin and A. K. Singh. Fighting opinion control in social networks via link recommendation. In *Proc. of ACM SIGKDD Conference of Knowledge Discovery and Data Mining (KDD)*, pages 677–685, Anchorage, AK, US, Aug. 2019. [\[pdf\]](#), [\[doi\]](#), [\[supplement\]](#), [\[poster\]](#), [\[slide\]](#), [\[video\]](#).