Optimization and resilience of complex supply-demand networks

2015 New J. Phys. 17 063029

(http://iopscience.iop.org/1367-2630/17/6/063029)

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
26 January 2015

REVISED
27 March 2015

ACCEPTED FOR PUBLICATION
1 May 2015

PUBLISHED
23 June 2015

PAPER

Optimization and resilience of complex supply-demand networks

Si-Ping Zhang[1], **Zi-Gang Huang**[1,2], **Jia-Qi Dong**[1], **Daniel Eisenberg**[3], **Thomas P Seager**[3] **and Ying-Cheng Lai**[2]

[1] Institute of Computational Physics and Complex Systems, Key Laboratory for Magnetism and Magnetic Materials of the Ministry of Education, Lanzhou University, Lanzhou 730000, China
[2] School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287, USA
[3] School of Sustainable Engineering and Built Environment, Arizona State University, Tempe, AZ 85287, USA

E-mail: huangzg@lzu.edu.cn **and** Ying-Cheng.Lai@asu.edu

Keywords: supply-demand networks, cascading failure, optimization, resilience

Abstract

Supply-demand processes take place on a large variety of real-world networked systems ranging from power grids and the internet to social networking and urban systems. In a modern infrastructure, supply-demand systems are constantly expanding, leading to constant increase in load requirement for resources and consequently, to problems such as low efficiency, resource scarcity, and partial system failures. Under certain conditions global catastrophe on the scale of the whole system can occur through the dynamical process of cascading failures. We investigate optimization and resilience of *time-varying* supply-demand systems by constructing network models of such systems, where resources are transported from the supplier sites to users through various links. Here by optimization we mean minimization of the maximum load on links, and system resilience can be characterized using the cascading failure size of users who fail to connect with suppliers. We consider two representative classes of supply schemes: load driven supply and fix fraction supply. Our findings are: (1) optimized systems are more robust since relatively smaller cascading failures occur when triggered by external perturbation to the links; (2) a large fraction of links can be free of load if resources are directed to transport through the shortest paths; (3) redundant links in the performance of the system can help to reroute the traffic but may undesirably transmit and enlarge the failure size of the system; (4) the patterns of cascading failures depend strongly upon the capacity of links; (5) the specific location of the trigger determines the specific route of cascading failure, but has little effect on the final cascading size; (6) system expansion typically reduces the efficiency; and (7) when the locations of the suppliers are optimized over a long expanding period, fewer suppliers are required. These results hold for heterogeneous networks in general, providing insights into designing optimal and resilient complex supply-demand systems that expand constantly in time.

1. Introduction

Supply-demand processes associated with various types of resources ranging from mass and energy to information are key to modern social, technological, and eco-systems. The network of services in a modern infrastructure such as hospitals, schools, firehouses, post offices, stores, power and water stations, etc is one example. Data networks in the world-wide-web, the underlying physical networks (i.e., the Internet), and online social media are other examples. In an ecosystem, the energy transportation processes among different species in a food chain can also be regarded as a supply-demand process. Mathematically, the dynamical properties of a supply-demand process can be studied in the context of complex networks [1, 2]. Such a network is typically time varying because of the constant addition of new suppliers into the system in response to rising demand. To optimize the new suppliers in terms of their number and locations to ensure high efficiency is of great interest. For example, for governmental social welfare agencies and commercial service industries, it is desirable to be able to provide better services to more people with fewer facilities. This optimization problem is mathematically

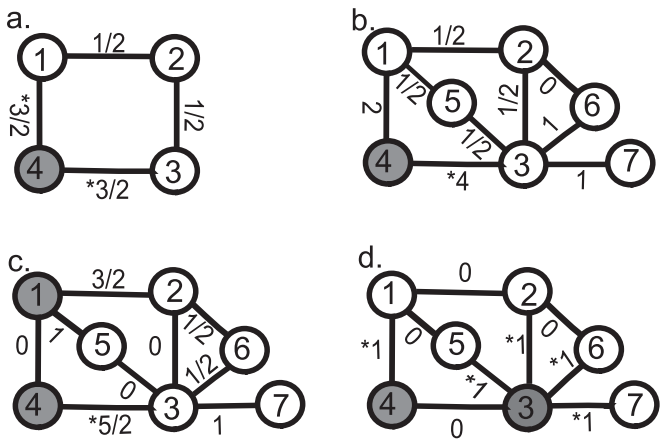

Figure 1. *Illustration of an expanding supply-demand network.* Filled gray and open circles denote the sites occupied by suppliers and demanders, respectively. Each demander orders one unit of resource from the nearest suppliers, with load divided uniformly among all the shortest paths. (a) Demanders 1, 2, and 3 receive resources from the supplier 4 through the respective shortest paths. The maximum edge load $L_{\max} = 3/2$ is labeled by '$\star$'. (b) For the expanded system of size seven, $L_{\max}$ is increased to 4. (c) Addition and optimization of the location of one new supplier to reduce $L_{\max}$. (d) Optimal configuration of suppliers.

challenging, attracting continuous interest of researchers from various disciplines such as business, economics, systems engineering, computer science, geography, and even biology [3–11].

In this paper, we investigate optimization and resilience of complex demand-supply networks from the dynamical point of view, motivated by the fact that, in general, rapid expansion of any networked system will inevitably affect the various dynamical processes that it supports. For a supply-demand network, expansion can lead to increasing load requirement for resources, causing problems such as low efficiency, resource scarcity, and small and large scale failures. Of particular interest is the dynamical process of cascading failures, which has been studied extensively in the past but mostly for static networks [12–26]. There were also previous works on dynamical processes on time-varying networks [27], in specific contexts such as genomics [28], oscillator synchronization [29–32], opinion dynamics and evolutionary games [33, 34]. These works, however, mainly addressed the dynamics of the co-evolving systems of stable size. From the standpoint of time varying networks, the distinct feature of a supply-demand network lies in the rapid expansion of its size. To our knowledge, the effects of such expansion on network optimization and resilience have not been studied. Specifically, by optimization we mean minimization of the maximum load $L_{\max}$ on links, and by resilience we mean the system's ability to sustain cascading failures. When such failures occur, some demanders will be separated from the suppliers. The number of the separated demanders corresponds to the size of the cascading failure, which can be used as a quantitative measure to characterize the resilience of the system. An optimized supply-demand network is more robust to perturbation such as disabling or removal of links .

Due to the rapid expanding nature of supply-demand networks, analytic treatment of cascading dynamics is extremely difficult. We thus rely on systematic numerical computations. Our main results are the following. We find that the specific route to cascading failures depends sensitively on the location of the perturbed link and its capacity (cf, figure 3). An intuitive approach to mitigating cascading failures is to have 'redundant' links in the network, links that are free of load. However, we find that these links play a 'double-sword' role: they can help reroute the traffic but can also increase the final failure size of the system (cf, figure 4). The links that handle neither too large nor too small load have a higher probability to trigger large scale cascading failures upon perturbation (cf, figure 5). By considering various types of expansion and optimization schemes, we also find that expansion typically reduces efficiency because it makes the present optimal locations of suppliers immediately less optimal (cf, figures 7–9). To maintain efficient function of the system, the locations of the suppliers need to be adjusted frequently over a larger region of candidate sites in response to expansion.

In section 2, we define supply-demand networks and introduce two types of supply schemes for systems under expansion: load driven supply (LDS) and fixed fraction supply (FFS). In section 3, we study the interplay between optimization and resilience in terms of cascading failures triggered by removal of a single link. In section 4 we provide an understanding, through extensive numerics, of how the expansion affects optimization. In section 5, we present conclusions and discussions.

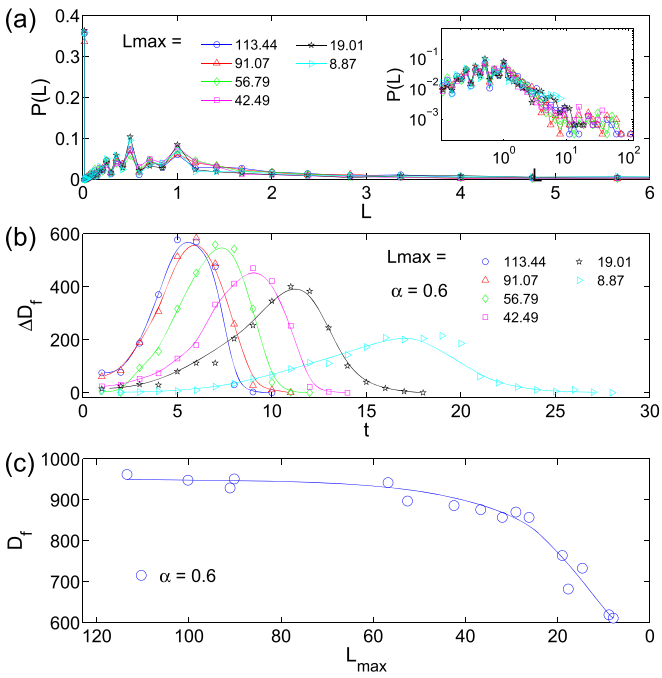

Figure 2. *Cascading failures of a supply-demand network under different levels of optimization.* (a) Edge-load distribution of the system optimized to a given value of L_{max} ranging from 113.44 to 8.87. The inset shows the corresponding plot on a double-logarithmic scale. (b) Number of failed demanders at each time step triggered by an attack targeted at the maximum-load edge at $t = 0$. (c) Total number of failed demanders D_f versus L_{max}. The solid lines are for eye guide. The underlying network is scale-free with size $N = 1000$, $S = 10$, average degree $\langle k \rangle = 6$, and total number of edges $E = 3000$. The edge capacity parameter is $\alpha = 0.6$.

2. Models of supply-demand networks

A supply-demand network consists of two components: suppliers that provide certain types of resources or services, and demanders that exploit the resources or use the services. Resources can be, e.g., data packets in Internet, electric power, water supply in an urban system, public transportation devices, etc. To conveniently quantify the resources, we conceive them as being composed of packets [2] that flow from suppliers' sites (sources) to demanders' sites (destinations). The suppliers and demanders are connected together through a networked structure — a kind of complex transportation network.

We assume that the amount of resource supported by each supplier is unlimited, and each demander orders one unit of resource from the nearest supplier(s) through the shortest path(s) in the underlying network. The unit resource to each demander is equally divided among suppliers with identical shortest paths, as illustrated in figure 1. If there are x shortest paths, regardless of the number of suppliers, the share or weight of each path is $1/x$. The load L on a given edge is the sum over the shares of all the paths through it [2], as shown in figure 1. The load is thus a variant of the link betweenness [35] with respect to sources and destinations. For the realistic situation where the traffic capacity C on every edge is limited, the maximum edge load L_{max} is an important parameter determining the performance of the supply-demand system. The optimal locations of suppliers subject to minimization of L_{max} can be found through methods such as simulated annealing [36, 37] and genetic algorithms [38, 39].

In an expanding system of population growth, more resources are required from time to time, introducing more load to the underlying supply-demand network. As illustrated in figure 1(a), the network grows from a given optimized initial state with one supplier (filled circle) and three demanders (open circles). The number of sites is thus $N = S + D = 4$, with S and D denoting the numbers suppliers and demanders, respectively. The maximum edge load is $L_{max} = 3/2$ (marked by $\star$). When three new nodes are introduced into the system, as shown in figure 1(b), the maximum load becomes $L_{max} = 4$, which does not necessarily occur on the original maximum-load edge. As a new supplier is added to the system to relieve edge overloading, its location plays an important role in minimizing L_{max}. Figure 1(c) and (d) illustrate the two outcomes for the two possible locations of the new supplier, where the location in figure 1(d) is the optimal one.

We consider two types of expansion mechanisms: LDS and FFS. The system with an increasing number of demanders may cause certain edges to become overloaded. Through LDS, once L_{max} exceeds a pre-assigned upper bound of edge capacity C, new suppliers are added and optimized in system one by one until L_{max} becomes smaller than C. Addition and optimization of suppliers take place on the same time scale as the

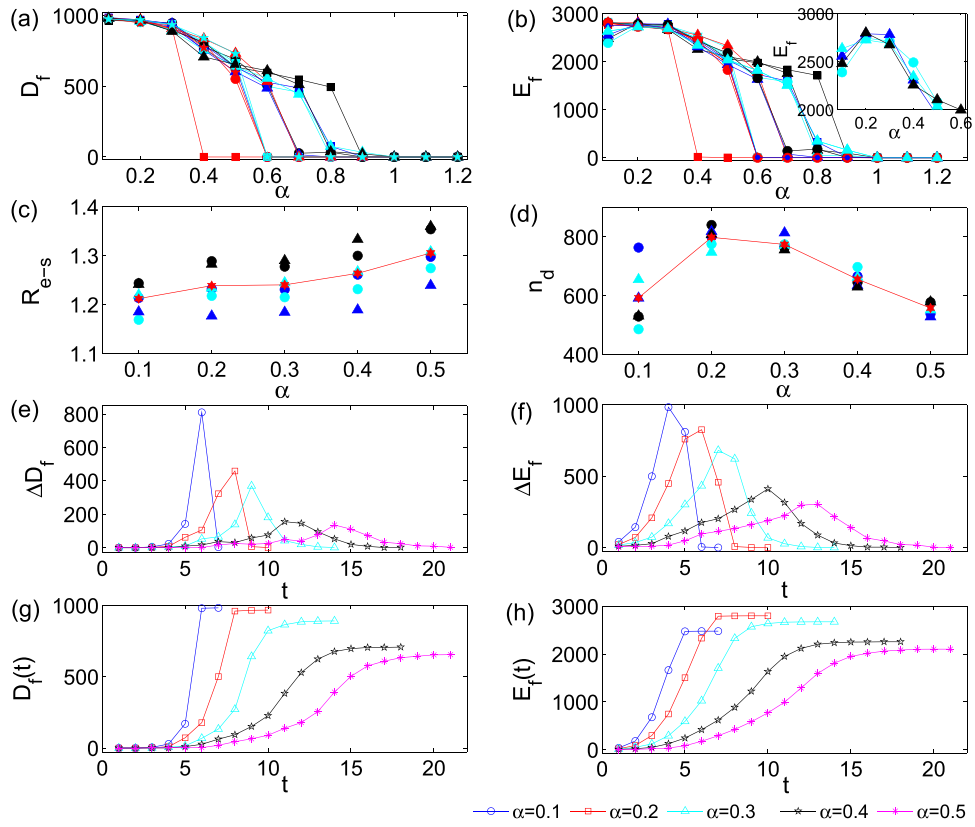

Figure 3. *Effects of edge capacity on cascading failures.* (a), (b) Numbers of failed demanders and disabled edges, denoted as D_f and E_f, respectively, versus the tolerance parameter α. The results from the ten different network configurations are marked by different symbols. (c) Average distance from disabled edges to the closest suppliers, denoted by R_{e-s}, versus α. This results indicates that, for systems with small values of α, the edges near the suppliers are more likely to be involved in the cascading process. (d) Number of clusters of failed demanders, denoted as n_d, versus α, where the solid symbols indicate the results from ten different network configurations, and the symbols connected by lines correspond to the average values. (e), (f) Temporal behaviors of ΔD_f and ΔE_f, and (g), (h) accumulated amounts of $D_f(t)$ and $E_f(t)$, respectively, with α ranges from 0.1 to 0.5 (denoted by different open symbols). The supply-demand network has $N = 1000$ nodes and $E = 3000$ edges.

expansion of the network. For FFS, a fixed fraction of nodes are arranged to be suppliers as the system expands, i.e., the number of suppliers increases proportionally with the system size, and the locations of the new suppliers are optimized over the whole system, or in the newly established region composed of the latest set of nodes added to the system. In this case, there is separation in time scales in that the network can expand much faster than suppliers are added into the system and optimized, where the expansion rate ΔN is an externally adjustable parameter. We employ the simulated annealing algorithm [36, 37] to optimize the locations of the new suppliers. The representative growing scale-free network model [40] is adopted to describe the underlying expanding supply-demand network, where adding nodes in the course of network growth corresponds to introducing more demanders and thus more load into the system. The network growth rule is set according to the two types of expansion mechanisms, LDS and FFS. In LDS, nodes (with m initial links) are added into the system one by one. In FFS, a group of nodes are added before each optimization process. The so generated scale-free network has power-law degree distribution with the scaling exponent $\gamma \simeq 3$ and average degree $\langle k \rangle = 2m$. In addition, after optimizing the locations of the suppliers under different scenarios (see details in section 4), suppliers are found to have a preference to large degree sites.

3. Optimization and resilience

In general, random errors or an intentional attack can trigger cascading failures. To understand how such failures can occur in a supply-demand network provides a way to assess the resilience of the system.

For a static supply-demand network with a given configuration of suppliers, the load on each edge is known *a priori*. A reasonable assumption [12] is that the capacity C_i of edge i is proportional to its load L_i:

$$C_i = (1 + \alpha) L_i, \tag{1}$$

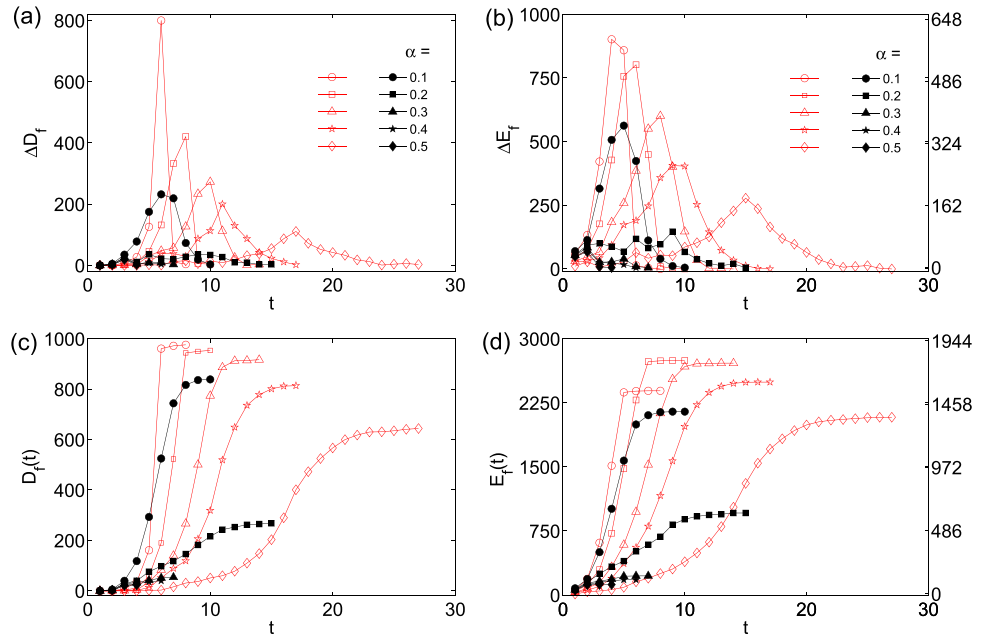

Figure 4. *Effect of redundant edges on cascading process.* For a supply-demand network and its variant in which all redundant edges are removed, (a) number of failed demanders and (b) number of failed edges versus time in a cascading event. Results for the original network are plotted using red open symbols while those for the variant network are represented by black filled symbols. (c), (d) The corresponding accumulated numbers of $D_{\rm f}\,(t)$ and $E_{\rm f}\,(t)$, respectively. In (b) and (d), the results from the original network (with $E\,=\,3000$ edges) are plotted with y-axis labeling on the left, and the results from the variant network without redundant edges (with $E\,=\,1944$ edges) correspond to y-axis labeling on the right.

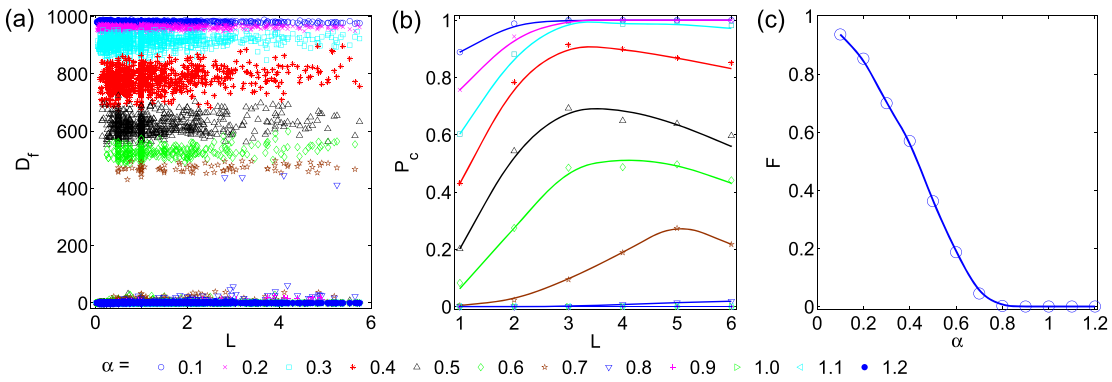

Figure 5. *Cascading size and probability triggered by single attack on edges.* For a supply-demand network of size $N = 1000$, $E = 3000$ and $S = 10$, (a) cascading size $D_{\rm f}$ triggered by a single attack on edges of load L for different values of α and (b) the corresponding probability P_c of cascading process. (c) Fraction of loaded edges that trigger the cascading process. The simulation results (denoted by symbols) in (b) and (c) are averaged over ten network realizations, and the solid curves are for eye guidance.

where the parameter $\alpha > 0$ is an adjustable tolerance parameter. When one edge fails to work (due either to random failure or to an intentional attack), the set of paths passing through this edge will no longer be available, leading to a global redistribution of load over the whole system. Any edge with new load $L_i > C_i$ will fail to deliver the resources to the demanders, and this causes the load to redistribute again, and so on. The cascade of overload failures can cut off a large number of paths from suppliers to demanders, leading to catastrophic failures of the demanders. A feature that distinguishes this type of cascading failures from previously studied ones [12–26] is that here, the failures are result of edge overload instead of node overload.

To characterize the extent of edge-overload induced cascading failures in a supply-demand network, we use the quantity $D_{\rm f}$, the number of demanders that are not connected to any supplier and thus fail to function, due to the network's inability to deliver the required resources to them. For convenience, we call them *failed demanders*.

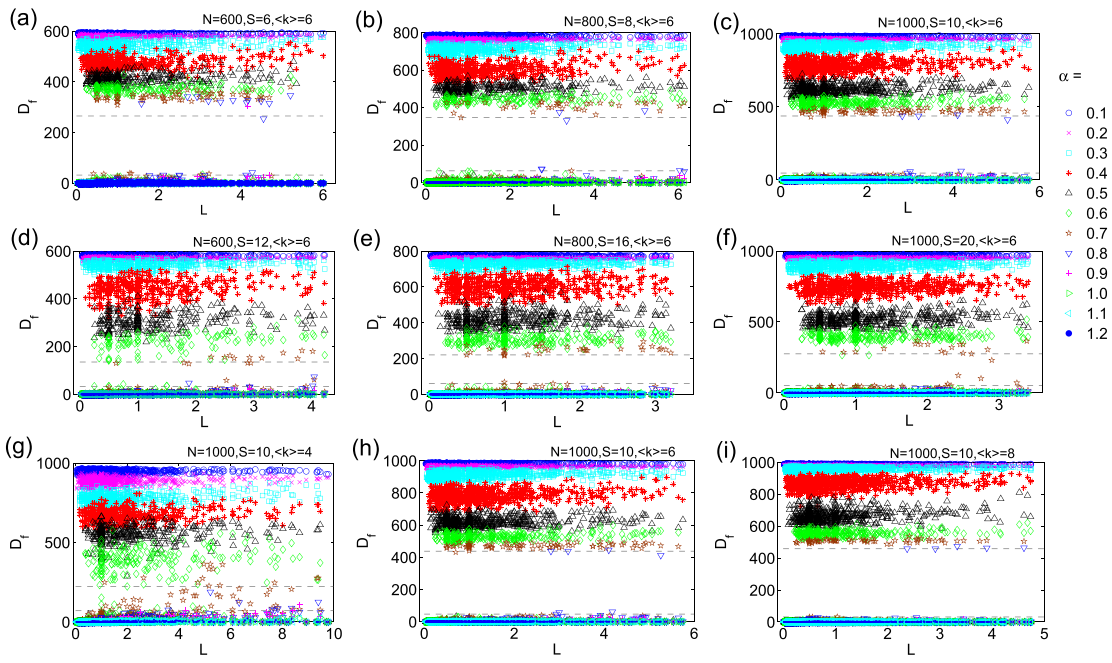

Figure 6. *Origin of gap in D_f.* Size D_f of the cascading process triggered by a single attack on edge of load L for different values of α (denoted by different symbols) in the supply-demand networks with various combinations of the network parameters N, S and $\langle k \rangle$. The upper (a)–(c) and middle (d)–(f) rows of panels have $S = 0.01N$ and $S = 0.02N$, respectively. The bottom row of panels compares the results from the network with increasing average degree $\langle k \rangle$.

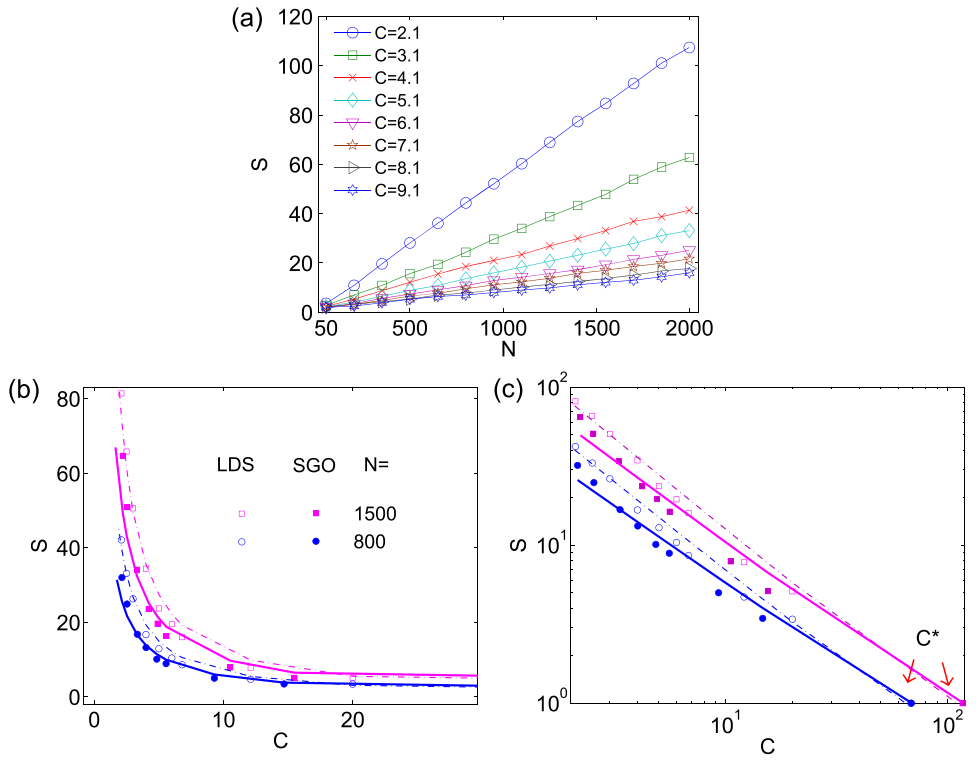

Figure 7. *Required suppliers in expanding systems with load driven supply (LDS).* (a) Number of suppliers S as a function of the system size N. (b) Linear plots and (c) log–log plots of S versus C for $N = 800$ and 1500 under LDS (open symbols fitted by dashed lines) and SGO (filled symbols fitted by solid lines). The power-law decay of S towards the critical value C^* with exponents $\beta = 1.3$ and 1.0 for LDS and SGO are plotted, respectively. The underlying scale free network has the average degree of $\langle k \rangle = 6$. The results are averaged over ten random network realizations.

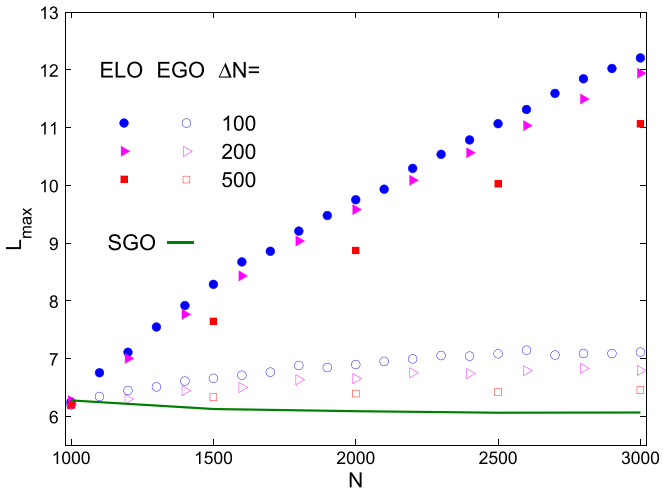

Figure 8. *Maximum edge load under fixed fraction supply (FFS).* For network expansion under FFS, maximum edge load L_{max} versus the system size N, for expansion rate $\Delta N = 100, 200$, and 500. The optimization schemes considered are ELO (filled symbols) and EGO (open symbols). For comparison, the results from SGO (solid curve) for the corresponding static systems are also shown. Results are averaged over 40 statistical realizations, and the fraction of suppliers is $a = 0.1$.

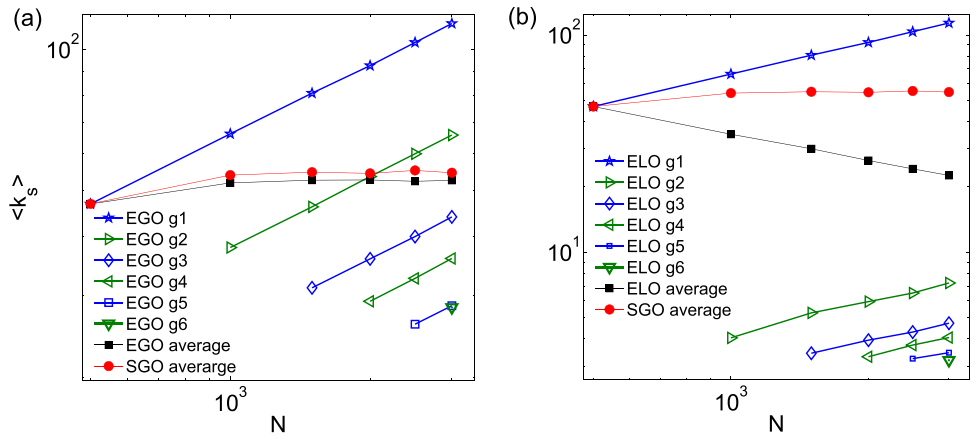

Figure 9. *Average degree of suppliers in expanding systems.* The average degree $\langle k_s \rangle$ of all the suppliers (black filled squares) and those at each time step (open symbols labeled by g_1–g_6) in systems under optimization schemes (a) EGO and (b) ELO, respectively. The expansion rate of the systems are $\Delta N = 500$ and $\Delta S = 5$. The SGO results for the static system with the corresponding values of N and S (red filled circles) are also included for comparison.

3.1. Optimization and resilience

In a supply-demand network, optimization of suppliers' locations leads to minimally possible value of L_{max}. What is then the interplay between optimization and resilience? Figure 2(a) shows the load distributions of a system with different values of L_{max}, from a random initial configuration of suppliers with $L_{max} = 113.44$ (open circles) to optimal configuration with $L_{max} = 8.87$. We see that, in the optimization process of reducing L_{max}, the qualitative features of the load distribution remain mostly unchanged. In particular, the region of the distribution remains broad. Furthermore, for random or optimal configuration of suppliers, over 35% of the edges have load $L = 0$. Intuitively, these 'redundant' edges provide 'room' for the system to recover when the number of edges that carry load is reduced due to failures. However, a more careful examination (see figure 4) shows that, while the redundant edges may help to reroute the traffic, an undesirable consequence is that they also promote the propagation of cascading failures and lead to larger cascading size.

Single attack upon nonzero-load edge may trigger a cascade of failures. For simplicity, we first consider attacks targeted at the maximum-load edge. As an example, we show in figures 2(b) and (c) the incremental number (denoted by ΔD_f) of failed demanders versus time and the corresponding asymptotic numbers, respectively, for networks with different values of L_{max}. The edge capacities C_i in each network are set according to the load definition in equation (1). We see that, the better the system is optimized (i.e., with smaller value of L_{max}), the cascading process is relatively more benign in the sense that the failure spreading is slower. In contrast,

if a system has a large value of L_{max} as for the case where the locations of suppliers are arranged randomly, the cascading process is more devastating in the sense that many more demanders fail to reach suppliers. We thus see that, although the original objective of optimizing the locations of the suppliers does not seem to be directly related to system resilience, a better optimized network is apparently more resilient against intentional attacks. The results shown in figure 2 keep unchanged qualitatively when the parameters of the system such as N, S, and E are changed. In addition, for a given system, once a cascading process is triggered by an attack on a single edge, the failure size D_f is independent of the specific location of the attack, as shown in figures 5(a) and 6. To gain more insights into the interplay between optimization and resilience, in the following we study the *optimal* system with varying edge capacities, i.e., one with minimal value of L_{max}, from the perspective of perturbation on edges with different load.

3.2. Cascading failures in systems of varying edge capacities

We study the role of edge capacity C_i in the cascading process, which can be varied systematically through the tolerance parameter α (equation (1)). Figures 3(a) and (b) show the asymptotic number of the failed demanders (D_f, the cascading size), and the number of disabled edges E_f at the end of the cascading process, respectively, versus α for ten different network realizations with identical values of the parameters N, S, E and $\langle k \rangle$, but generated from different random seeds. We see that, as the edge capacity C_i is increased, D_f decreases monotonically, but E_f exhibits a nonmonotonic behavior around $\alpha = 0.2$ for some network realizations as shown in the inset of figure 3(b). In addition, the details of the cascading process is quite sensitive to the specific topology of the network, resulting in different critical values of α above which the network is free of cascading dynamics.

The nonmonotonic behavior of E_f in figure 3(b) signifies a counterintuitive phenomenon: increasing the edge capacity can reduce the number of failed demanders but can simultaneously cause more edges to fail. A detailed check of the underlying cascading process reveals that varying edge capacity can affect the route (or trajectory) of the cascading process in the network. Figures 3(c) and (d) show the average distance from failed edges to their nearest suppliers, denoted by R_{e-s}, and the number of clusters of failed demanders, denoted by n_d, respectively. For $\alpha = 0.1$, the edges near the suppliers (smaller distance R_{e-s}) fail rapidly and the failed demanders separated from the suppliers form a few large clusters and many small clusters. This is indication that large amount of edges among demanders are not involved in the cascading process. For relatively larger α values (e.g., 0.2), more edges fail (corresponding to larger E_f values in (b)) but the distance R_{e-s} to suppliers becomes larger, as shown in figure 3(c), implying failure of edges among demanders that leads to the emergence of smaller clusters of failed demanders, as indicated in figure 3(d) through the larger values of n_d. Overall, in contrast to the monotonically decreasing behavior of D_f, the nonmonotonic behavior in E_f with a peak at about $\alpha = 0.2$ indicates a strong variance in the cascading trajectory through the network. In particular, small edge capacities induce local edge failures close to suppliers and result in large cascading size, while larger edge capacity leads to more edge failures but relatively smaller cascading size.

Examples of the temporal behaviors of D_f and E_f are shown in figures 3(e)–(h) for $\alpha = 0.1$–0.5. We see that cascading dynamics in systems with smaller values of α are more severe with larger final failure size and shorter duration, while for large values of α cases, the process spreads more slowly. We also see two factors that contribute to the cascading failures: (1) failures of edges and subsequent load redistribution that can trigger overload of the remaining edges in a cascading manner, and (2) reduction of total traffic flow in the system due to disconnections of demanders from the suppliers. The final extent of the cascading process is result of the balance of these two factors.

3.3. The role of redundant edges in cascading dynamics

From figure 2, we see that a considerable fraction of edges are in fact free of load for various degree of optimization. Are the redundant edges useful to mitigate overloading and cascading failures? To address this question, we calculate the numbers of failed demanders and failed edges for the original network and for its variant in which all the redundant edges are removed. The results are shown in figure 4. Surprisingly, we see that removal of all the redundant edges can always inhibit the cascading process. This counterintuitive phenomenon can be explained, as follows. The redundant edges serve to provide more rerouting paths from the suppliers to the demanders in load redistribution when some nonzero-load edges are disabled. As shown in figure 4, the original system with redundant edges (red open symbols) has smaller values of D_f and ΔD_f for several initial time steps during the cascading process as compared to the variant system without redundant edges (black filled symbols). However, the 'saved' demanders that are connected to the suppliers via new paths through redundant edges will bring more loads to the whole system, leading to more dramatic cascading failures.

For the network without redundant edges, there are fewer rerouting paths available. As a result, even though some demanders would fail initially, the path structure between the suppliers and demands are relatively more

stable, making the extent of load redistribution less severe and effectively inhibiting the cascading dynamics. In addition to this load redistribution issue, we find that for the network with redundant edges, isolation of the supplier, namely failures of *all* the edges towards demanders for a given supplier, occurs with a higher probability.

3.4. Effect of edge load on cascading process

The results obtained so far are for cascading failures triggered by a single attack on the edge with the maximum load $L_{\max}$. A question is whether an attack on an arbitrary edge of load $L < L_{\max}$ can induce a cascading process. Figure 5(a) shows the cascading size D_{f} versus the load of the attacked edge for systems with different values of α. We find that D_{f} is independent of the load of the attacked edge as the value of D_{f} is distributed randomly in a small interval. This means that, the location of the initial edge failure has little effect on the final cascading size, once the process has taken place. As α is increased, D_{f} decreases.

Interestingly, in figure 5(a), a gap of D_{f} in the range $[20,\ 400]$ can be observed. Extensive simulations are carried out on networks with different values of parameters N, E, and S to understand the gap. Figure 6 plots a typical set of results. We find that the systems with larger values of N, smaller values of S, or larger values of $\langle k\rangle$ have larger gaps in D_{f}. More specifically, we observe the following: (1) the width of the gap is proportional to the size N (see figures 6(a)–(f)), (2) larger number of suppliers can reduce the width of the gap (comparing figures 6(a), (d), figures 6(b), (e), and figures 6(c), (f)), and (3) an increase in the average degree $\langle k\rangle$, i.e., larger number of edges will enlarge the width of the gap (figures 6(g)–(i)). These results imply that the emergence of the D_{f}-gap can be attributed to the *tree structure* of suppliers which result in *strong correlations* among the nodes in the cascading process, where the supply tree of a given supplier is composed of all the paths along which the supplier provides resources to demanders. Once a cascading process is triggered, the strong correlation among nodes (through the supply trees) will induce a relatively large failure size D_{f}, rather than a continuous increase in D_{f} from zero. A gap in D_{f} thus emerges between the cases with and without cascading failures. Furthermore, for a system with fewer suppliers (small S), the supply tree has longer paths and larger size on average, which induces stronger correlation among the nodes. Compared to the opposite case of larger value of S (panels (d)–(f) in figure 6], in the small S systems (panels (a)–(c) in figure 6), an initial single attack can trigger a cascading process of larger size, generating a larger gap in D_{f}. In addition, the existence of redundant links enlarges the cascading size. An increase in the average degree $\langle k\rangle$ (panels (g)–(i) in figure 6), which leads to an increasing number of redundant links, results in a larger value of D_{f} and a larger gap width in D_{f}.

Figure 5(b) shows the probability P_c for the occurrence of cascading failures versus load L of the initially failed edge. Equivalently, P_c is the fraction of edges with load L on which a single attack triggers cascading. There is a non-monotonic relation between P_c and L, indicating that an attack on some edge with median load is more likely to trigger a cascading process. Additionally, as shown in figure 5(b), for the case of small α values, e.g., 0.1 or 0.2, the system is fragile in the sense that an attack on any edge with $L > 4$ will trigger cascading failures with probability one. For larger α values, e.g., $\alpha = 0.3$, the non-monotonic behavior of P_c becomes apparent.

Figure 5(c) shows the total fraction F of edges that can trigger a cascading process versus α, which corresponds to the probability cascading process due to random edge failure. We see that, as α is increased through a critical value $\alpha^{\star} \approx 0.8$, the system is immune to cascading failures under any single-edge attack. This can also be seen in figures 5(a) and (b) where, for $\alpha > 0.8$, no cascading occurs and both the failure size D_{f} and the cascading probability P_c approach zero.

4. Optimization of growing supply-demand networks

We consider the standard growing, scale-free network model [40] for two supply scenarios: (1) LDS and (2) FFS. For LDS, arrangement and optimization of suppliers take place on the same time scale as that of expansion of the network. For FFS, the rate of network expansion is larger than that of optimization.

4.1. Scenario of LDS

When a system expands, loads on edges increase with more demanders, and new suppliers are required due to the limited edge capacity C. The new supplier can be anywhere in the network except for those locations already occupied by previous suppliers. The goal is to select optimal locations for the new suppliers which minimize $L_{\max}$. If $L_{\max}$ is larger than C after addition of one supplier, another supplier can be added into the system at some optimal location. This process continues until $L_{\max} < C$. Since the amount of resource supported by each supplier is unlimited, the system with fewer suppliers (smaller value of S) would satisfy all of the demanders through more long range paths, provided that no edge is overloaded. When too many paths are needed, some edges will inevitably be overburdened, requiring more suppliers at appropriate locations.

Figure 7(a) shows a typical relationship between the number S of suppliers and system size N for LDS, where the edge capacity limit C ranges from 2.1 to 9.1. Because of the linear increase of S with N, the average output of each supplier $O \equiv (N - S)/S$ is approximately a constant for any given capacity C. However, the output depends on C due to the dependence of S on C. For comparison, we also analyze the strategy of static global optimization (SGO) where, for a given static network of size N and identical edge capacity C, the locations of all S suppliers are optimized synchronously over the whole system to obtain the lowest value of $L_{\max}$. In SGO, if optimization of the S suppliers is unable to meet the condition $L_{\max} \leqslant C$, one more supplier will be added and the locations of *all* the $S + 1$ suppliers will be recalculated. This process is repeated iteratively until the system satisfies the constraint $L_{\max} \leqslant C$. Notably, different from LDS, in determining the supplier locations under the SGO strategy, information about the evolutionary history of the network is not needed, i.e., the 'elder' suppliers are not fixed.

Figures 7(b) and (c) respectively show the linear and log–log plots of the value of S versus C for two instants of time when the system size is $N = 800$ and 1500 (circles and squares, respectively) both for LDS (open symbols) and SGO (filled symbols). We observe the following power-law decay of S with C:

$$S \sim NC^{-\beta}, \tag{2}$$

where the power-law decay exponent is $\beta = 1.3$ for LDS and 1.0 for SGO. The dashed curves and solid curves respectively are the least square fittings to the results from LDS and SGO. Take LDS as an example, the systems in the lower-left region below the dashed curve in figures 7(b) and (c), i.e., those with inadequate suppliers or too small values of C, will have a large number of overloaded edges, while those corresponding to upper regions above the curves have redundant capacities. If C is adjustable, one can see from figure 7(b) that the initial increase in C (e.g., from 1 to about 6) dramatically reduces the number of additional suppliers, making enhancing C a highly efficient strategy for avoiding overloading. However, increasing C in the larger capacity region is not effective at reducing the number of new suppliers. In addition, the larger S of LDS compared to SGO implies that, for a static system of a given size, the SGO strategy requires fewer suppliers as compared to an expanding system evolved to the same size, in which the addition and optimization of suppliers are driven by overload events. This can be attributed to the *memory* effect in the expanding system, e.g., immobility of the existent suppliers.

To gain further insights, we consider the extreme case of one supplier, i.e., $S = 1$. For such a system, edge capacity C approaching N is sufficient for the system to avoid overload. The critical value C^*, below which overload on edge occurs, is indicated in figure 7(c). We find that the C^* values for the systems with LDS and SGO coincide with each other, which can be attributed to the fact that the two different optimization schemes have no effect on the one-supplier system. As C is decreased further, the SGO scheme requires fewer suppliers and consequently performs better. The power-law exponents from the LDS and SGO schemes respectively are $\beta = 1.3$ and 1.0, implying their different responses to decreasing C. The simple relation $S \cdot C \sim N$, which holds for the SGO scheme, is due to its sufficient and global utilization of edges. The LDS scheme, however, generates small inhomogeneous distribution of loads and thus requires more suppliers to avoid overloading. In addition, as C is decreased to the extreme case of $C < 1$, the number of suppliers S diverges with the system size for both schemes.

These results suggest that, to avoid overloading in an expanding supply-demand system, an effective scheme needs to simultaneously take into account two factors: (1) enhancement of edge capacity limit C, and (2) addition of suppliers.

4.2. Fixed fraction supply

In supply-demand networks under LDS, the number of demanders expands one at a time, i.e., the expansion rate is $\Delta N = 1$. In this case, the edge load is sensitive to each unit increment of N. In a realistic situation, the expansion rate ΔN for a system subject to optimization can be large. For example, a group of suppliers can be added into the system simultaneously. It is thus of interest to generalize the LDS scheme. To capture the essential features of this variant, we study the simple scheme denoted by FFS where a fixed fraction of suppliers is added to the system constantly. That is, we assume $S = aN$ or, equivalently, $\Delta S = a\Delta N$. The advantage of this setting is that it is not necessary to consider the relatively complicated situation of overloading under limited edge capacity. In this approach, $L_{\max}$ is effectively a measure of the system performance, and we focus on how $L_{\max}$ is affected by the value of the expansion rate ΔN. A useful indicator is the available optimization region for suppliers. The solution will be somewhat trivial if the locations of all suppliers can be optimized over the whole system without memory at any time—SGO scheme. However, this is over simplified because, in a real situation, the cost to add a new supplier (e.g., a hospital, a fire house or a school) in an already established region (e.g., an old urban district) can often be much higher than that to have the supplier in a newly developed region. Motivated by this consideration, we propose two realistic optimization schemes in which the new ΔS suppliers

at each time step are located in regions of (a) all sites except those already occupied by the elder suppliers (evolving global optimization) (EGO) and (b) the ΔN newly added sites (evolving local optimization) (ELO).

We carry out a comparative analysis of the results from EGO, ELO, and SGO schemes under system expansion. Figure 8 shows $L_{\max}$ for expansion rate $\Delta N = 100, 200$, or 500, where the system evolves from a small initial size N_0 to a larger size $N = 3000$. The number of suppliers is fixed at $S = 0.1N$ ($a = 0.1$). We see that the ELO scheme (filled symbols) generally leads to higher values of $L_{\max}$, implying lower efficiency as compared with the EGO cases (open symbols). In fact, expansion of the system generally changes the global supply-demand configuration. It cannot function optimally for the present system by simply combining the elder suppliers distribution which was optimized to fit within the original system and the new suppliers distribution which was optimized separately in the new region. However, under the EGO scheme, the locations of the new suppliers are optimized in the whole system, leading to much smaller values of $L_{\max}$ so as to have a better fit with the new configuration. That is, the larger optimization region for new suppliers associated with EGO can yield higher efficiency for the supply-demand process. We also see that, for both ELO and EGO, the systems with larger expansion rate perform better than those with smaller rate, which can also be attributed to the larger optimization region for new suppliers. The size of the optimization region for ΔS new suppliers is $N(t) - S(t) + \Delta N$ for the EGO case, and ΔN for the ELO case, both increasing with the expansion rate ΔN.

In comparison to the two evolving optimization schemes, EGO and ELO where the locations of the elder suppliers are constrained due to the prior system evolution, the SGO scheme requires the smallest number of suppliers, as shown in figure 8 (the solid curve). The locations of suppliers in both EGO and ELO can satisfy the demands of the system but only temporarily and partially. A disadvantage of SGO in spite of its higher efficiency, lies in cost because the elder sites occupied previously by demanders or suppliers need to be reestablished.

The scale free topology we assume for the supply-demand system has a heterogeneous degree distribution. Based on extensive simulations, we find that the optimal locations for the suppliers under SGO in static systems tend to favor the hub nodes. Even for ELO and EGO, new suppliers have a preference to large degree sites. Figures 9(a) and (b) show the average degrees of suppliers from EGO and ELO, respectively, for $\Delta N = 500$ and $\Delta S = 5$. As the system expands continuously, with each new generation having 495 demanders and 5 suppliers, the average degree $\langle k_s \rangle$ of suppliers in each generation (labeled as $g1$–$g6$ with open symbols, respectively) exhibits a power-law scaling with N as $\langle k_s \rangle \sim N^{1/2}$, which can be attributed to the degree preferential attachment process [40]. In particular, in the continuum limit the degree of the ith site added to the system at t_i increases as

$$\frac{\partial k_i}{\partial t} = m\frac{k_i}{\sum_j k_j} \approx \frac{k_i(t_i)}{2t}, \quad k_i(t_i) = m, \tag{3}$$

for which the solution is $k_i(t) = m(t/t_i)^{\beta}$ with $\beta = 1/2$ and t corresponding to the number of sites N. This leads to the observed scaling relation $\langle k_s \rangle \sim N^{1/2}$. However, the average degree over all the existing suppliers for EGO, ELO (black filled squares), and SGO (red filled circles) exhibits a somewhat different behavior. Especially, suppliers from EGO have the same rising trend but a smaller average degree with respect to SGO. The average degree associated with ELO still exhibits a power-law decay behavior, since the new suppliers are constrained within the newly generated small-degree sites.

5. Conclusion

Rapid expansion of infrastructure is ubiquitous in the modern time, in which various supply-demand processes take place. Does expansion make the system more fragile or the opposite or, more generally, *what is the interplay between expansion and resilience*? In this paper, we systematically investigate the expansion, optimization, and resilience of supply-demand networks. Firstly, we study the effects of optimization on the locations of suppliers, and those of enhancement of edge capacity on the resilience of the system via characterization of cascading failures of demanders triggered by perturbation to links. We find that, in general, the optimized systems (with smaller values of the maximum edge load) are more robust because the size of cascading failures is typically smaller. For edges with median load, there is a higher probability that a single attack can trigger cascading failures. Once a cascading process is initiated, its size does not depend on the specific location of the original link that triggers the process. The pattern of cascading failures also depends strongly upon the capacity of links, where a smaller capacity can lead to more rapid and massive cascading failures of demanders and the disabled edges are closer to the locations of suppliers on average. We also find that the 'redundant' edges with zero load play a paradoxical role, i.e., while they can help reroute the traffic so as to ensure that demanders are connected to suppliers, they can undesirably increase the failure size of the whole system. Taking into account various types of expanding and supply schemes, we study the effect of size expansion on the system efficiency. Under the LDS

scheme where suppliers are added one by one into the system in responses to overloading, the required number of suppliers S scales with the capacity limit C as a power law. For the FFS scheme, both local and global optimization strategies require more suppliers in comparison with the result of global optimization in static systems of the same size. In general, system expansion makes the present optimal location of suppliers quickly non-optimal, reducing the system efficiency. If the locations of the suppliers are optimized over a larger region of available sites, fewer suppliers are required. Extensive simulations show that these results hold for heterogeneous networks in general.

The supply-demand systems with heterogeneous structures numerically investigated in this paper can be a prototype model for real world infrastructure systems under constant expansion, such as supply chains, logistic networks [41], flight networks, and the Internet. Our results provide initial insights into the resilience of such systems, for which further efforts are justified due to the importance of the problem. In particular, in the real world there are supply-demand networks that do not possess the scale-free topology, such as urban traffic systems and power grids. The issues associated with weighted nodes and directed-weighted edges taking into account the nonhomogeneous capacities and specific function of suppliers and edges are also important, as well as multiple layer or interdependent structures with more complicated coupling among different supply-demand processes.

Acknowledgments

This work was supported by NSF under Grant No. 1441352. SPZ and ZGH were supported by NSF of China under Grants No. 11135001 and No. 11275003. ZGH thanks Prof Liang Huang and Xin-Jian Xu for helpful discussions.

References

[1] Rubido N, Grebogi C and Baptista M S 2014 *Phys. Rev.* E **89** 012801
[2] Chen Y H, Wang B H, Zhao L C, Zhou C and Zhou T 2010 *Phys. Rev.* E **81** 066105
[3] Yook S H, Jeong H and Barabsi A L 2002 *Proc. Natl Acad. Sci. USA* **99** 13382–6
[4] Lloyd P and Dicken P 1977 *Location in Space: A Theoretical Approach to Economic Geography* (London: Joanna Cotler)
[5] Stenphan G 1997 *Science* **196** 523–4
[6] Vining D, Yang C and Yeh S 1979 *Science* **205** 219
[7] Stephan G E 1988 *J. Reg. Sci.* **28** 29–40
[8] Hamilton M, Milne B, Walker R and Brown J 2007 *Proc. Natl Acad. Sci. USA* **104** 4765–9
[9] Jetz W, Carbone C, Fulford J and Brown J 2004 *Science* **306** 266–8
[10] Haskell J, Ritchie M and Olff H 2002 *Nature* **418** 527–30
[11] Um J, Son S, Lee S, Jeong H and Kim B 2009 *Proc. Natl Acad. Sci. USA* **106** 14236–40
[12] Motter A and Lai Y C 2002 *Phys. Rev.* E **66** 065102
[13] Motter A E 2004 *Phys. Rev. Lett.* **93** 098701
[14] Gallos L K, Cohen R, Argyrakis P, Bunde A and Havlin S 2005 *Phys. Rev. Lett.* **94** 188701
[15] Schäfer M, Scholz J and Greiner M 2006 *Phys. Rev. Lett.* **96** 108701
[16] Simonsen I, Buzna L, Peters K, Bornholdt S and Helbing D 2008 *Phys. Rev. Lett.* **100** 218701
[17] Wang W X and Chen G R 2008 *Phys. Rev.* E **77** 026101
[18] Wang W X and Lai Y C 2009 *Phys. Rev.* E **80** 036109
[19] Lehmann J and Bernasconi J 2010 *Phys. Rev.* E **81** 031129
[20] Mirzasoleiman B, Babaei M, Jalili M and Safari M 2011 *Phys. Rev.* E **84** 046114
[21] Huang X Q, Gao J X, Buldyrev S V, Havlin S and Stanley H E 2011 *Phys. Rev.* E **83** 065101
[22] Hackett A, Melnik S and Gleeson J P 2011 *Phys. Rev.* E **83** 056107
[23] Yehezkel A and Cohen R 2012 *Phys. Rev.* E **86** 066114
[24] Liu R R, Wang W X, Lai Y C and Wang B H 2012 *Phys. Rev.* E **85** 026110
[25] Dong G, Gao J, Du R J, Tian L X, Stanley H E and Havlin S 2013 *Phys. Rev.* E **87** 052804
[26] Mizutaka S and Yakubo K 2013 *Phys. Rev.* E **88** 012803
[27] Gross T and Blasius B 2008 *J. R. Soc. Interface* **5** 259–71
[28] Bornholdt S and Sneppen K 1998 *Phys. Rev. Lett.* **81** 236
[29] Ito J and Kaneko K 2001 *Phys. Rev. Lett.* **88** 028701
[30] Bornholdt S and Rohlf T 2000 *Phys. Rev. Lett.* **84** 6114
[31] Stilwell D J, Bollt E M and Roberson D G 2006 *SIAM J. Appl. Dyn. Syst.* **5** 140
[32] Porfiri M, Stilwell D J, Bollt E M and Skufca J D 2006 *Physica* D **224** 102
[33] Holme P and Ghoshal G 2006 *Phys. Rev. Lett.* **96** 098701
[34] Huang Z G, Wu Z X, Xu X J, Guan J Y and Wang Y H 2007 *Eur. Phys. J.* B **58** 493–8
[35] Fekete A, Vattay G and Kocarev L 2006 *Phys. Rev.* E **73** 046102
[36] Kirkpatrick S, Gelatt C D and Vecchi M P 1983 *Science* **220** 671–80
[37] Aarts E and Korst J 1989 *Simulated Annealing and Boltzmann Machines* (New York: Wiley)
[38] John H 1992 *Adaptation in Natural and Artificial Systems* (Cambridge, MA, USA: MIT Press)
[39] Goldberg D E 1989 *Genetic Algorithms in Search, Optimization and Machine Learning* 1st eds (Boston, MA: Addison-Wesley Longman)
[40] Barabási A L and Albert R 1999 *Science* **286** 509–12
[41] Sun H and Wu J 2005 *Mod. Phys. Lett.* B **19** 841–8