Network connectivity under node failure

Billand, P., Bravard, C., Iyengar, S. S., Kumar, R., & Sarangi, S. (2016). Network connectivity under node failure. DOI: 10.1016/j.econlet.2016.10.028

Published in:
Economics Letters

Document Version:
Peer reviewed version

Network Connectivity under Node Failure

Pascal Billand* Christophe Bravard† Sitharama S. Iyengar‡
Rajnish Kumar§¶ Sudipta Sarangi‖

October, 2016

Abstract

We examine a non-cooperative model of network formation where players may stop functioning. We identify conditions under which Nash and efficient networks will remain connected after the loss of k nodes by introducing the notion of k-Node Super Connectivity.

Keywords: Connectivity; node failure; Nash networks; efficient networks.
JEL Codes: C7; D8; R4

1 Introduction

In this paper, we focus on a model of network formation where players (also called nodes) in the network can fail with a certain exogenous probability. Examples of networks being affected by node failure abound in the real world. Consider for instance, the social and economic networks in a region hit by a natural calamity like a hurricane, or a business network where some firms exit the industry. This can also occur in a network of servers or a sensor network either due to mechanical failure or a malicious attack. Connectivity in the network is important in all these examples, suggesting that strategic agents will have an incentive to create multiple paths between themselves.

The literature on strategic reliability in economics has mostly focused on the possibility of links failures. The model incorporating reliability in networks was introduced in a paper by Bala and Goyal (2000) where all links are allowed to fail with a given exogenous probability. The authors then proceed to provide a partial characterization of Nash and efficient networks in this context. Subsequently, Haller and Sarangi (2005) and Billand, Bravard and Sarangi (2011) allow for heterogeneity in the link failure probabilities and values that can be obtained from other players respectively. While full characterization of the equilibrium networks is again shown to be elusive, both papers provide an "anything goes" result which shows that with only a bit of heterogeneity (two different

*Université de Lyon, St Etienne, France. (Email: pascal.billand@univ-st-etienne.fr)
†GAEL, Univ. Grenoble-Alpes, France. (Email: christophe.bravard@upmf-grenoble.fr)
‡Florida International University, FL, USA. (Email: iyengar@cis.fiu.edu)
§Queen's University Belfast | Management School, United Kingdom
¶Corresponding author: Rajnish Kumar, Queen's University Belfast | Management School, 185 Stranmillis Road, Belfast, BT9 5EE, United Kingdom (Email: rajnish.kumar@qub.ac.uk, Phone: +44 (0)28 9097 5424, Fax: +44 (0)28 9097 4201)
‖Virginia Tech, VA, USA; DIW Berlin, Germany (Email: ssarangi@vt.edu)

parameter values) any essential network can be supported as a Nash equilibrium. In a slightly different framework, De Jaegher and Hoyer (2016) study a game between a network designer and a network disruptor and find the equilibrium network architectures for different levels of link costs. These types of game have also been studied by Dziubiński and Goyal (2013), Goyal and Vigier (2014) and Haller (2016). Note that in these models the attacker typically removes links or nodes. Ours is the first paper to study node failure when agents form links in a decentralized way – a phenomenon quite different from link failure. Under link failure, as well as under node failure, agents can create alternate paths by forming costly links with other agents. However, the logic for creating alternate paths is quite different. Consider Figure 1.

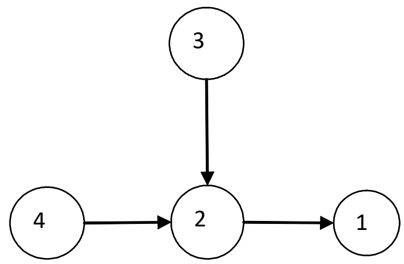

Figure 1: Node failure vs link failure: An illustration

In Figure 1, player 1 has formed a link with player 2 and player 2 has formed links with players 3 and 4. In this figure, under link failure, the formation of an additional link by player 1 with player 3 can allow player 1 to access the resources of players 2, 3 and 4, in situation where the link between 1 and 2 fails. By contrast, under node failure, this additional link can never allow player 1 to get access to the resources of players 2 and 4 when player 2 fails; it can only allow access to the resources of player 3. Moreover, under node failure, in Figure 1, player 1's payoff will not change if players 3 and 4 add a link between themselves. However, this link will improve her expected payoff under link failure. Finally, under node failure, player 1 also needs to take into account her own survival probability while computing her payoffs before adding a costly link. Thus, ensuring connectivity in the two different models may require different strategies.

Our focus in this note is on connectivity in the network. To study this we introduce the notion of k-Node Superconnectivity which checks whether a network is still connected after the deletion of any k nodes. Using this definition, we then identify sufficient conditions for both Nash and efficient networks.

2 Preliminaries

Graph-theoretic concepts. A (simple directed) **network** $\boldsymbol{g}$ is a pair of sets (N, E) where N is a set of nodes and $E \subset N \times N$ is a set of links with $(i, i) \notin E$ for all $i \in N$. We denote by $(i, j) \in E$ the link from i to j. Let G be the set of all (simple directed) networks whose set of vertices is N. A **chain** in $\boldsymbol{g}$ between node j and node $i \neq j$, is an alternating sequence of distinct nodes $i_0, i_1, .., i_m$ such that $i_0 = i, i_m = j$, and an alternating sequence of distinct links such that for $k = 0, ..., m-1$, $(i_k, i_{k+1}) \in E$ or $(i_{k+1}, i_k) \in E$. A network $\boldsymbol{g}$ is **connected** if there is a chain in $\boldsymbol{g}$ between all nodes $i, j \in N$. A subnetwork $\boldsymbol{g}^{N'} = (N', E')$ induced by $N' \subseteq N$ consists of a set of nodes N'

and a set of links $E' \subset N' \times N'$ such that $(i,j) \in E'$ if and only if $(i,j) \in E$ for every pair $(i,j) \in N' \times N'$. Let $S(\boldsymbol{g})$ be the set of all subnetworks of $\boldsymbol{g}$ induced by all subsets $N' \subseteq N$. Note that for each subset $N' \subseteq N$ there is a unique subnetwork of $\boldsymbol{g}$ which belongs to $S(\boldsymbol{g})$. A **component** $\boldsymbol{g}^{N'}$ of $\boldsymbol{g}$ is a connected (induced) subnetwork of $\boldsymbol{g}$ such that for all $N'' \subseteq N$ with $N'' \supset N'$, $\boldsymbol{g}^{N''}$ is not connected. Finally, a network $\boldsymbol{g} \in G$ is **essential** if $(i,j) \in E$ implies $(j,i) \notin E$.

We now present two definitions that will play an important role in our analysis. A set of nodes $N' \subseteq N$ in a connected network $\boldsymbol{g}$ is **critical** if $\boldsymbol{g}^{N \setminus N'}$ is not connected. A network is **k-Node Super Connected** (k-NSC) if no set of k nodes or less is critical. To avoid triviality, we set $k < n-2$.

Players and strategies. The set of players is identified with the set of nodes $N = \{1, \ldots, n\}$, $n \geq 3$. For each player $i \in N$, a pure strategy is a vector $\boldsymbol{g}_i = (g_{i,1}, ..., g_{i,i-1}, 0, g_{i,i+1}, ..., g_{i,n}) \in \{0,1\}^n$. Here $g_{i,j} = 1$ means that player i forms a link with player j, whereas $g_{i,j} = 0$ means that i does not form this link. Let $\boldsymbol{g}_{-i} = (\boldsymbol{g}_1, \ldots, \boldsymbol{g}_{i-1}, \boldsymbol{g}_{i+1}, \ldots, \boldsymbol{g}_n)$ be the profile of strategies of all players except i. We focus only on pure strategies. The set of all pure strategies of player i is denoted by $\mathcal{G}_i$, with $\mathcal{G}_i = \{0,1\}^{N \setminus \{i\}}$. The joint strategy space is denoted by $\mathcal{G} = \mathcal{G}_1 \times \ldots \times \mathcal{G}_n$. Note that there is a one-to-one correspondence between $\mathcal{G}$ and G the set of simple directed networks with vertex set N. Hence with a slight abuse of notation, we identify the strategy profile $(\boldsymbol{g}_1, \ldots, \boldsymbol{g}_n) \in \mathcal{G}$ with the network $\boldsymbol{g} = (N, E)$ where $g_{i,j} = 1$ if and only if $(i,j) \in E$.

Payoff. Player i incurs a cost $c > 0$ for each link she forms. We consider the **two-way flow of information** model, where both the agents involved in a link can access the resources (or information) of the other agent regardless of which agent initiates the link. Moreover, player i obtains resources from player j if there exists a chain between i and j. We denote by $N_i(\boldsymbol{g}) = \{j \in N : j \neq i, \text{ there exists a chain in } \boldsymbol{g} \text{ between } i \text{ and } j\}$ the set of players whom i can access or "observe" in network $\boldsymbol{g}$.
In our context, players may stay put (i.e. node failure occurs) or appear (i.e. node failure does not occur). It follows that the network formed by the players can be different from the actual network observed. Hence we introduce the notion of realization to capture the effects of this assumption. Formally, a realization of $\boldsymbol{g}$, $\boldsymbol{g}^{N'} \in S(\boldsymbol{g})$, is a subnetwork of $\boldsymbol{g}$ where all players in N' are functioning and all players in $N \setminus N'$ are not functioning. Following the strategic reliability literature, assume the probability of node failure to be identical and independent, where the survival probability of every node is given by $p \in (0,1)$. Given $\boldsymbol{g}$, the probability of subnetwork $\boldsymbol{g}^{N'}$ being realized is:

$$\lambda(\boldsymbol{g}^{N'}) = p^{|N'|}(1-p)^{n-|N'|}.$$

Note that for $\boldsymbol{g}, \boldsymbol{h} \in G$ we have $\lambda(\boldsymbol{g}^{N'}) = \lambda(\boldsymbol{h}^{N'})$ for all $N' \in 2^N$. This property is important for establishing Proposition 2.
We now define the function $B_i(\boldsymbol{g})$ as the expected benefit of player i in network $\boldsymbol{g}$. Summing over all possible realizations of the network we get:

$$B_i(\boldsymbol{g}) = V \sum_{N' \in 2^N} \lambda(\boldsymbol{g}^{N'}) |N_i(\boldsymbol{g}')|, \tag{1}$$

where V is the value of information that i gets from an agent with whom he is connected to directly or indirectly. Wlog we set $V = 1$. Using equation (1) we define i's expected payoff, that takes into account both costs and benefits as:

$$u_i(\boldsymbol{g}) = B_i(\boldsymbol{g}) - c\sum_{j\in N} g_{i,j} \tag{2}$$

Nash networks. With a slight abuse of notation, we identify the pair $(\boldsymbol{g}_i, \boldsymbol{g}_{-i})$ with the network $\boldsymbol{g}$. A strategy $\boldsymbol{g}_i$ is a **best response** of player i to $\boldsymbol{g}_{-i}$ if

$$u_i(\boldsymbol{g}_i, \boldsymbol{g}_{-i}) \geq u_i(\boldsymbol{g}'_i, \boldsymbol{g}_{-i}), \text{ for all } \boldsymbol{g}'_i \in \mathcal{G}_i.$$

Let $BR_i(\boldsymbol{g}_{-i})$ denote the set of player $i's$ best responses to $\boldsymbol{g}_{-i}$. A network $\boldsymbol{g}$ is a **Nash network** if $\boldsymbol{g}_i \in BR_i(\boldsymbol{g}_{-i})$ for each $i \in N$. Note that if $\boldsymbol{g}$ is a Nash network, then it must be essential. This follows from the fact that each link is costly while information flow is two-way and independent of which player invests in forming the link.

Efficient Networks. A network $\boldsymbol{g}$ is efficient if

$$\boldsymbol{g} \in \arg\max_{\boldsymbol{g}'\in G} \sum_{i\in N} B_i(\boldsymbol{g}') - c\sum_{i\in N}\sum_{j\in N} g_{i,j}.$$

Network $\boldsymbol{g}$ may be considered as the social planner's objective.

3 Results

We start by exploring the impact of node failure on connectivity in the network. In all the following results we will assume that the payoffs are given by equation 2. Our first result shows that non-empty Nash networks are connected.

Proposition 1 *A Nash network is either empty, or connected.*

Proof. Consider a Nash network $\boldsymbol{g}$ that is neither empty nor connected. Then there exist three players, say i, j and j' such that i and j lie in a component $\boldsymbol{g}^{N_1}$ with $g_{i,j} = 1$, and j' lies in a different component $\boldsymbol{g}^{N_2}$. Let $\boldsymbol{g}^{-i}$ be identical to the network $\boldsymbol{g}$ except that any links between i and other players are deleted and let $\boldsymbol{g}^{-ij}$ be identical to the network $\boldsymbol{g}$ except that the link between i and j is deleted. Moreover, let $\boldsymbol{h}^{-i}$ be identical to the network $\boldsymbol{g}^{-i}$ except that a link from i to j is added. Let b_1 be the marginal expected benefit of i associated with the link she has formed with j in $\boldsymbol{g}^{-ij}$. Since $\boldsymbol{g}$ is a Nash network $b_1 \geq c$.
The marginal expected benefit obtained by i due to her link with j in $\boldsymbol{g}^{-i}$ is $b_2 = p^2(1 + \sum_{\ell\in N_j(\boldsymbol{g}^{-i})} Q^{\ell}_{\boldsymbol{g}^{-i}}(p))$, where $Q^{\ell}_{\boldsymbol{g}^{-i}}(p)$ is a polynomial function showing the probability that player j will obtain the resources of other players ℓ (assuming that j is alive) in the network $\boldsymbol{g}^{-i}$.
We now compare b_2 and b_1. Let ρ_ℓ be the probability that there is a chain between i and ℓ in $\boldsymbol{g}^{-ij}$, and let q_ℓ be the probability that there is a chain between i and ℓ in $\boldsymbol{h}^{-i}$. The probability that there is a chain between i and ℓ in $\boldsymbol{g}$ is $\rho'_\ell \leq \rho_\ell + q_\ell$. Since the probability that there is a chain between i and ℓ in $\boldsymbol{g}^{-i}$ is zero, the increase in the probability that there is a chain between i and ℓ is higher when a link from i to j is added in the network $\boldsymbol{g}^{-i}$ than when a link from i to j is added in the network $\boldsymbol{g}^{-ij}$. It follows that $b_2 \geq b_1$.

We now compare the marginal expected benefits that player j' obtains if she forms a link with j in $\boldsymbol{g}$ and the marginal expected benefits that payer i obtains when she forms a link with j in $\boldsymbol{g^{-i}}$. The marginal expected benefit to j' of forming a link with j in $\boldsymbol{g}$ is $b_3 = b_2 + p^3 > b_2 \geq b_1 \geq c$ since $p > 0$. Therefore j' improves her payoff if she forms a link with player j. Hence $\boldsymbol{g}$ is not Nash, a contradiction. ■

Note that an equilibrium network can be empty. This will be the case for situations where building links is very costly and is true whether or not we account for node failure. Indeed, if $p < c$, then the expected benefits associated with a link in the empty network is lower than the cost of this link. Moreover, it is worth noting that a Nash network is connected when $p > c$. The key thing to be noted here is that this connectivity result in equilibrium is obtained under decentralized decision making when heterogeneity on probabilities is introduced.[1] Note that it is easy to obtain sufficient conditions under which certain common architectures will be Nash networks in our model.

First, let us define some specific networks. A network $\boldsymbol{g}$ is called a **star** if there is a vertex i_s , such that for all $j \neq i_s, \max\{g_{i_s,j}, g_{j,i_s}\} = 1$ and for all $k \notin \{i_s, j\}, g_{k,j} = 0$. Moreover a star, where $g_{j,i_s} = 0$ for all $j \neq i_s$ is a Center Sponsored Star (CSS), a star where $g_{i_s,j} = 0$ for all $j \neq i_s$, is a Periphery Sponsored Star (PSS), other stars are called mixed stars. A network $\boldsymbol{g}$ is **complete** if $\max\{g_{i,j}, g_{j,i}\} = 1$ for any $i, j \in N$, $j \neq i$.

If $p^2 > c$ and $p^2(1-p) < c$, then all types of stars are Nash networks. If $p^2 + p^3(n-2) > c$ and $p^2(1-p) < c$, then a PSS is a Nash Network. If $p^2(1-p)^{n-2} > c$, then a Nash network is a complete network. The complete characterization of the set of Nash networks is difficult under node failure. Having established the basis for connectivity, we now dig deeper into the level of connectivity. As one would suspect, due to possibility of node failure, the agents will form backup paths whenever it is worth doing so. In other words, equilibrium networks will have more links than a minimally connected graph. Hence using the notion of node superconnectivity introduced earlier, our next proposition provides sufficient conditions for Nash and efficient networks.

Proposition 2 *(i) If $p^2(1-p)^k > c$, then a Nash network is k-NSC. (ii) If $p^2(1-p)^k > c/2$, then an efficient network is k-NSC.*

Proof. First, we prove part (i). Suppose that $p^2(1-p)^k > c$. To introduce a contradiction, let $\boldsymbol{g}$ be a Nash network that is not k-NSC. Since $\boldsymbol{g}$ is not k-NSC, there exists a set of players $K \subset N$, with $|K| = k$, which is critical. So there are two distinct players $i, j \in N \setminus K$ that are unconnected in any subnetwork of $\boldsymbol{g}$ where all players in K fail and i and j are functioning. Note that i and j have no link in $\boldsymbol{g}$.

Let $\mathcal{N}_1 \subset 2^N$ be the set of subsets of players that contain both players i and j, and do not contain any player in K. The probability that a realization, where all players in K fail and both players i and j are active, is given by $\sum\limits_{N' \in \mathcal{N}_1} \lambda(\boldsymbol{g}^{N'}) = p^2(1-p)^k$.

We denote by $\boldsymbol{h}$ the network identical to $\boldsymbol{g}$ except that a link from i to j is added in $\boldsymbol{g}$. The marginal benefit obtained by i when she forms a link with j in $\boldsymbol{g}$ is $\Delta_i(\boldsymbol{h}, \boldsymbol{g}) = B_i(\boldsymbol{h}) - B_i(\boldsymbol{g})$, with $B_i(\boldsymbol{h}) = \sum_{N'^N \setminus \mathcal{N}_1} \lambda(\boldsymbol{h}^{N'})|N_i(\boldsymbol{h}^{N'})| + \sum_{N' \in \mathcal{N}_1} \lambda(\boldsymbol{h}^{N'})|N_i(\boldsymbol{h}^{N'})|$ and $B_i(\boldsymbol{g}) = \sum_{N'^N \setminus \mathcal{N}_1} \lambda(\boldsymbol{g}^{N'})|N_i(\boldsymbol{g}^{N'})|+$

[1] A simple continuity argument on cost and benefits can be used for this. See for instance Haller and Sarangi (2005).

$\sum_{N'\in\mathcal{N}_1} \lambda(\boldsymbol{g}^{N'})|N_i(\boldsymbol{g}^{N'})|$. We have:

$$\begin{aligned}\Delta_i(\boldsymbol{h},\boldsymbol{g}) &= \textstyle\sum_{N'^N\setminus\mathcal{N}_1} \lambda(\boldsymbol{h}^{N'})|N_i(\boldsymbol{h}^{N'})| + \sum_{N'\in\mathcal{N}_1} \lambda(\boldsymbol{h}^{N'})|N_i(\boldsymbol{h}^{N'})| \\ &\quad - \textstyle\sum_{N'^N\setminus\mathcal{N}_1} \lambda(\boldsymbol{g}^{N'})|N_i(\boldsymbol{g}^{N'})| - \sum_{N'\in\mathcal{N}_1} \lambda(\boldsymbol{g}^{N'})|N_i(\boldsymbol{g}^{N'})| \\ &= \textstyle\sum_{N'^N\setminus\mathcal{N}_1} \lambda(\boldsymbol{h}^{N'})(|N_i(\boldsymbol{h}^{N'})| - |N_i(\boldsymbol{g}^{N'})|) \\ &\quad + \textstyle\sum_{N'\in\mathcal{N}_1} \lambda(\boldsymbol{h}^{N'})(|N_i(\boldsymbol{h}^{N'})| - |N_i(\boldsymbol{g}^{N'})|).\end{aligned}$$

The second equality follows from the fact that $\lambda(\boldsymbol{g}^{N'}) = \lambda(\boldsymbol{h}^{N'})$. Note that, $|N_i(\boldsymbol{h}^{N'})| - |N_i(\boldsymbol{g}^{N'})| \geq 1$ for all $N' \in \mathcal{N}_1$ and $\sum_{N'\in\mathcal{N}_1} \lambda(\boldsymbol{h}^{N'}) = p^2(1-p)^k$. Hence the second term is bounded below by $p^2(1-p)^k$. The first term is non negative. Consequently, the marginal benefit obtained by i when she forms a link with j is $\Delta_i(\boldsymbol{h},\boldsymbol{g}) \geq p^2(1-p)^k > c$. It follows that player i has an incentive to form the link with j and $\boldsymbol{g}$ cannot be a Nash network, a contradiction.

Second, we prove part (ii). Suppose $p^2(1-p)^k > c/2$, and let $\boldsymbol{g}$ be an efficient network in which a set K of players is critical, with $|K| = k$. So there are two distinct players $i, j \in N \setminus K$ that are unconnected in any subnetwork of $\boldsymbol{g}$ where all players in K fail and i and j are active. Again, let $\boldsymbol{h}$ be the network identical to $\boldsymbol{g}$ except that a link from i to j is added. Using the same argument as in the proof of part (i) above, we obtain that $B_i(\boldsymbol{h}) - B_i(\boldsymbol{g}) \geq p^2(1-p)^k$ and $B_j(\boldsymbol{h}) - B_j(\boldsymbol{g}) \geq p^2(1-p)^k$. Moreover, any additional link formed by i cannot decrease the expected benefits of players in $N\setminus\{i,j\}$. Therefore, $\sum_{\ell\in N} u_\ell(\boldsymbol{h}) - \sum_{\ell\in N} u_\ell(\boldsymbol{g}) = \sum_{\ell\in N} B_\ell(\boldsymbol{h}) - \sum_{\ell\in N} B_\ell(\boldsymbol{g}) - c \geq 2p^2(1-p)^k - c > 0$. Thus $\boldsymbol{g}$ is not efficient, a contradiction. ■

Proposition 2 lays out the relationship between costs, benefits and the level of connectivity. Proposition 2.(i) identifies that trade-offs involved in making these decisions from a decentralized perspective while Proposition 2.(ii) identifies the conditions that a network designer will have to satisfy. From Proposition 2.(i) it is easy to see that when $p^2(1-p) > c$, then a Nash network is node superconnected. The same condition holds for efficient networks by adjusting the cost parameter.

4 Conclusion

Most of the analysis of models of network formation occurs in deterministic settings. Our paper is the first strategic model of network formation that allows for node failure and identifies conditions for network connectivity. Given the fact that node failure is a common cause for concern, this is important for studying several new types of problems. Clearly, following earlier work on link failure, one can always introduce heterogeneity in values and probabilities in the model. The more interesting question would be to simultaneously analyze the possibility of node and link failure in the network. This line of research has important implications for the defense of critical networks.

References

[1] Bala, V. and S. Goyal (2000) "A strategic analysis of network reliability," *Review of Economic Design 5*, 205-228.

[2] P. Billand, C. Bravard and S. Sarangi, (2011) "Heterogeneity and Link Imperfections in Nash Networks." *International Game Theory Review*, 13, 181-194.

[3] M. Dziubiński and S. Goyal. Network design and defense. *Games and Economic Behavior*, 79:30-43, 2013.

[4] S. Goyal and A. Vigier. Attack, defense, and contagion in networks. *The Review of Economic Studies*, 81(4):1518-1542, 2014.

[5] De Jaegher, K.J.M. and B. Hoyer (2016). Strategic network disruption and defense, forthcoming *Journal of Public Economic Theory* .

[6] H. Haller. Network vulnerability: a Designer-Disruptor game (2016), *Working paper*. Available at: https://ideas.repec.org/p/vpi/wpaper/e07-50.html

[7] Haller, H. and S. Sarangi (2005) "Nash Networks with Heterogeneous Links," *Mathematical Social Sciences, 50*, 181- 201.