Accelerating networks

David M D Smith\textsuperscript{1}, Jukka-Pekka Onnela\textsuperscript{1,2} and Neil F Johnson\textsuperscript{1}

\textsuperscript{1} Clarendon Laboratory, Physics Department, Oxford University, Oxford, UK
\textsuperscript{2} Laboratory of Computational Engineering, Helsinki University of Technology, Finland

Abstract. Evolving out-of-equilibrium networks have been under intense scrutiny recently. In many real-world settings the number of links added per new node is not constant but depends on the time at which the node is introduced in the system. This simple idea gives rise to the concept of accelerating networks, for which we review an existing definition and—after finding it somewhat constrictive—offer a new definition. The new definition provided here views network acceleration as a time dependent property of a given system as opposed to being a property of the specific algorithm applied to grow the network. The definition also covers both unweighted and weighted networks. As time-stamped network data becomes increasingly available, the proposed measures may be easily applied to such empirical datasets. As a simple case study we apply the concepts to study the evolution of three different instances of Wikipedia, namely, those in English, German, and Japanese, and find that the networks undergo different acceleration regimes in their evolution.
## Contents

1. Introduction .................................................. 2
2. The framework of Gagen and Mattick (GM) .............. 3
3. Rethinking the accelerating network ....................... 5
4. Defining accelerating networks ........................... 7
   4.1. Note on hyper-acceleration ............................ 9
   4.2. Accelerating weighted networks ...................... 10
   4.3. Stochastic accelerating networks .................... 10
5. Case study: Wikipedia ...................................... 11
6. Conclusion ................................................... 12
Acknowledgments .............................................. 13
References ...................................................... 13

### 1. Introduction

In many real-world evolving networks the rate at which links are added is different from the rate at which nodes are added, and this seems to be the case in particular in ‘functionally organized systems whose operation is reliant on the integrated activity of any or all of its component nodes’ [1]. This results in accelerating or decelerating networks, of which the Internet is one example [2, 3]. Other examples of such systems include the arrangement of components on an integrated chip, configuration of computers in clusters, structure of integrated production systems, configuration of human social and communication systems, and the organization of regulatory proteins controlling gene expression in bacteria [1].

Consider an arbitrary evolving network that may explore any allowed trajectory in its phase space of node number \( N(t) \) (number of nodes at time \( t \)) and link number \( M(t) \) (number of links at time \( t \)). A constructed example of such a network is shown in figure 1. In an undirected network of size \( N(t) \), the total number of possible links is \( \frac{N(t)(N(t) - 1)}{2} \) of which \( M(t) \) exist at some time \( t \). The ratio of the two is described by

\[
q(t) = \frac{2M(t)}{N(t)(N(t) - 1)}.
\]

In many conventional non-equilibrium evolving (growing) networks this quantity is expected to decrease in time [4]. As time-stamped network data becomes increasingly available, we expect to find that the dynamical evolution of such quantities in most networks will exhibit non-trivial behaviour [5, 6].

The notion of accelerating networks has been considered before [7, 8], and here we consider the concept as discussed by Gagen and Mattick (GM) in [7]. Although interesting, the definitions in [7] seem constrictive in that they are applicable only to a specific algorithm, and the concept of acceleration is static in nature, i.e. a given network is considered to be either accelerating, non-accelerating, or decelerating throughout its evolution. Contrary to this, the notion of acceleration in physical systems refers to the nature of the evolution of a system at a specific moment in time, suggesting that acceleration in the context of complex networks should be reconsidered. In addition, it is imperative that network acceleration be measurable for empirical networks.
Figure 1. The trajectory of an evolving network might exist anywhere in the allowable region of the phase space corresponding to its number of nodes $N(t)$ and number of links $M(t)$. This might be a network evolving according to some algorithm or an empirically observed network. It is possible that at any stage in its evolution the network might have many components and the criterion $M(t) < (N(t) - 1)$ imposes this condition.

We review the GM framework in section 2 and critically examine the proposed concept of network acceleration in section 3. This is followed by a new definition of network acceleration in section 4. This might be used to describe a network at any stage of its evolution (whether decaying, growing, or neither), and applies to both directed and undirected networks. We then extend the analysis to incorporate weighted links, an increasingly scrutinized aspect of evolving networks [9, 10]. We demonstrate the use of the new definition in a simple empirical case study in section 5, and conclude our discussion in section 6.

2. The framework of Gagen and Mattick (GM)

In the GM framework as in many other conventional out-of-equilibrium evolving network models, the system evolves by introducing exactly one new node at each time step [7]. The key feature of the model is that the new node connects to the existing network through a time-dependent number of links, whereas in most evolving network models [4] the new node attaches to the existing network with a fixed, time-independent number of links $m$, as demonstrated in figure 2.

Here we shall consider the scenario whereby the new node attaches with undirected binary (unweighted) links. Specifically, we assume that each new node attaches with on average $m(t)$ links. Whether or not the number of links added per time step is a stochastic process, the actual evolution of the system will be if a stochastic attachment algorithm (random, preferential, or otherwise) is employed. Clearly the maximum number of links with which the new node can connect to the existing network, is equal to the number of nodes within the existing network.
Figure 2. Evolution of a network. In most out-of-equilibrium network models a new node attaches to the existing network with $m$ links at each time step via some arbitrary attachment mechanism. These links might be directed or undirected, weighted or unweighted.

Figure 3. The evolution of the expected number of links added per new node, $m(t) = p(N(t))^\alpha$, within the GM framework of accelerating networks. Three examples comprising non-acceleration ($\alpha = 0$), acceleration ($\alpha = 0.8$) and hyper-acceleration ($\alpha = 1.5$) are displayed [7].

GM actually specify the functional form describing the expected rate of link addition $m(t)$ as

$$m(t) = p(N(t))^\alpha. \quad (2)$$

Here $\alpha$ is described as an acceleration parameter and $p$ as a probability constant with the constraint of $0 \leq p \leq 1$. GM characterize network behaviour in terms of $\alpha$ as either decelerating ($\alpha < 0$), non-accelerating ($\alpha = 0$), accelerating ($0 < \alpha < 1$), or hyper-accelerating ($\alpha \geq 1$). Some examples are shown in figure 3. We shall revisit such definitions later.

We note that for a non-accelerating network ($\alpha = 0$) within this description, the expected number of new links added is simply $p \in [0, 1]$ such that, on an average, at most one link may be added per time step. Note also that the hyper-accelerating network has a finite time scale for which it can evolve in that the maximum number of links that can possibly be added to the system is $N(t)$. This sets an upper bound on $m(t)$ after which the network cannot hyper-accelerate, i.e.
in order to continue hyper-acceleration, one would have to introduce more than \( N(t) \) links per time step, which is impossible without allowing for multiple links between any pair of nodes.

We also note that so far the actual attachment mechanism, random or otherwise, has not been discussed, meaning that the accelerating nature of the network is simply related to the rate of link addition. This is an important point to consider because although the expected number of links added for the new node is \( m(t) \), the variance in this will differ between microscopic attachment mechanisms. Indeed, the variance could be zero for integer \( m(t) \), corresponding to exactly \( m(t) \) links being added every time step, or the number of links to add could be drawn from a probability distribution with mean \( m(t) \). Even though the number of links to be added might be deterministic, one must still specify the algorithm to determine to which nodes within the existing network these links are attached.

3. Rethinking the accelerating network

The notion of network acceleration is applicable to situations in which the rates of node and link addition are not stochastic. In the rest of this paper, we will concentrate on the scenario in which the number of nodes and links which are added per time step is not an inherently stochastic process.

Consider some arbitrary evolving network whose evolution we can observe. This might be a realization of an algorithm or an empirically observed evolving network. At time \( t = 0 \) we have an initial configuration of \( N(0) \) nodes and \( M(0) \) links. At each time step, a quantity of new nodes \( n(t) \) are added and they are connected to the pre-existing network with some number of binary links \( m(t) \). Here, these are both integer quantities. In this scenario no new links are formed between existing nodes within the network, although this feature could be incorporated. For now, we shall assume that \( n(t) = 1 \) such that exactly one new node is added per time step\(^3\). As such, at time \( t = 1 \) the number of nodes is \( N(1) = N(0) + 1 \). The maximum number of links that could have been introduced on attaching the \( N(1) \) th node is clearly \( N(0) \) as the new node can link to each previous node with at most one link. Similarly, for a node added at time \( t \), the total number of nodes is \( N(t) = N(0) + t \) and the maximum number of links that could have been used to connect it is \( N(t-1) \). For this process, by which one new node is added per time step, we know that \( N(t-1) = N(t) - 1 \). This is the upper limit for the time-dependent number of added links, \( m(t) \), at each time step. This region is illustrated in figure 4.

Clearly, the functional form of \( m(t) \) could be any function that exists within the allowable region of \( m(t) \leq (N(t) - 1) \) for the addition of one node per time step. For the addition of \( n(t) \) nodes per time step, the constraint becomes

\[
\frac{m(t)}{n(t)} \leq N(t - 1). \tag{3}
\]

One can envisage any number of such functions for the time-dependent number of links added. We note that once the function reaches the constraint, the network can no longer evolve according to such a process.

\(^3\) Often, there is little merit in adding more than one node per time step as this might be equivalent to adding one node per time step over a longer period. However, situations might arise, virtual or empirical, in which several nodes are added per time step.
Figure 4. The allowable number of links (shaded area) added with each new node is capped by the total number of nodes in the existing network (line). Here, as one new node is added per time step, the growth of the network in terms of number of nodes is linear.

Figure 5. Left panel: A fairly standard example of a non-accelerating network in which more than one link is used to link the new node to the existing network. Right panel: Although this evolving network asymptotes to non-accelerating behaviour, clearly, initially, it is accelerating.

A simple example might be such that a single new node is added with \( m(t) = N(0) \) links per time step with \( N(0) > 1 \) as illustrated in figure 5. This is a non-accelerating network that could not be described by the GM framework that only allows at most one new link per new node for non-accelerating networks, corresponding to \( p = 1 \) and \( \alpha = 0 \) in their model. Another example function is one that is initially increasing but asymptotes to a constant value as illustrated in figure 5. This might be some empirically observed network growth or the behaviour of some
Figure 6. A schematic illustration of a network evolving over fourteen time steps in the link-node phase space (solid line). The mean degree of the network is given by $\langle k \rangle = 2M(t)/N(t)$, and the dashed line represents half the mean degree of the system at time step 8 (and time step 4). At time step 8 the network accelerates but the mean degree actually decreases.

growth algorithm. Should this be described as an accelerating, decelerating or non-accelerating network? Clearly, different regimes of this particular network’s evolution might satisfy differing descriptions. As such, we must re-define the accelerating network framework to encompass this feature.

4. Defining accelerating networks

One might expect that one could identify the regimes of acceleration, non-acceleration and deceleration with relative ease by writing these phenomena in terms of the mean degree $\langle k \rangle = 2M(t)/N(t)$ of the network. Specifically, one might expect that if the addition of a new node via some number of links results in an increase in the mean degree of the entire network, the network would be described as accelerating. Likewise, if the mean degree remains constant, the network is not accelerating and, if the mean degree decreases with the addition of new nodes, the network could be considered to be decelerating.

Although these ideas are intuitively appealing, one can envisage a scenario whereby a network might be accelerating without increasing its mean degree. This can occur if a network has undergone rapid deceleration such that the rate of link addition is very low even though it might have been high previously. At some point the ratio of number of links added to the number of nodes added over a time step, $[M(t) - M(t - 1)]/[N(t) - N(t - 1)]$, might be less than that for the existing network as a whole, $M(t)/N(t)$, thereby decreasing the average degree $\langle k \rangle$ while still constituting network acceleration. This is evident in figure 6, where the network accelerates at time step $t = 8$, even though the mean degree of the system decreases as the ratio $M(t)/N(t)$
decreases. In order for the mean degree to increase the trajectory would have to exceed the dashed line whose gradient represents half the mean degree of the system at time $t = 8$.

In order to identify the different regimes of network acceleration, we must relate the rate of increase in the number of links with the rate of addition of new nodes, denoting the rate of link addition and node addition by the approximate derivatives

$$m(t) = \frac{dM(t)}{dt} \approx M(t) - M(t - 1),$$

$$n(t) = \frac{dN(t)}{dt} \approx N(t) - N(t - 1).$$

We can then define the different regimes of network acceleration. The important ratio is the ratio of the rate of link addition to the rate of node addition, $m(t)/n(t)$, the evolution of which prescribes a network measure. We define network acceleration $a(t)$ as

$$a(t) \equiv \frac{d}{dt} \left( \frac{m(t)}{n(t)} \right).$$

We approximate the discrete values with continuous derivatives and define the following three regimes:

$$\begin{cases} 
    a(t) < 0 & \text{decelerating,} \\
    a(t) = 0 & \text{non-accelerating,} \\
    a(t) > 0 & \text{accelerating.}
\end{cases}$$

As such, a single evolving network might navigate all regimes during its evolution. Note that the definition of $a(t)$ allows more than one node to be added per time step. It is interesting to note that within this definition of network acceleration, a decaying network (losing nodes) could accelerate. Also, we note that the definition holds for directed graphs. The above definition alludes to the notion of network velocity $v(t)$, which we define as

$$v(t) \equiv \frac{dM(t)}{dN(t)} = \frac{m(t)}{n(t)}.$$

This velocity is simply the gradient of the network trajectory in the link-node phase space as in figure 6.

These definitions are motivated by non-equilibrium (growing) networks where $n(t) \neq 0$ by definition [4]. The definitions are intrinsically related to the nature of that growth although in light of their macroscopic make up, they can be applied when nodes and/or links are being added, removed or neither4. The possibility that links might be added when nodes are not (i.e. $n(t) = 0$) results in the velocity of equation 6 being undefined. This would correspond to a vertical trajectory in the link-node phase space of figure 6 in the positive direction for link addition and negative direction for link removal. Within the context of the definitions outlined here, positive network acceleration describes the process by which the trajectory becomes more vertical in the positive direction. As such, we postulate that, for a process in which no nodes are added or removed, the process of repeated link addition (or removal) constitutes non-accelerating behaviour.

4 One could simply describe network velocity and acceleration as vectors, specifically the time derivatives of the network’s position vector $\mathbf{r}(t) = (N(t), M(t))$ within the link-node phase space. However, meaningful interpretation of these quantities is less clear.
Figure 7. Superimposing the appropriate contour lines, we observe hyper-acceleration between times $A$ and $B$, acceleration between $B$ and $C$ and non-acceleration between $C$ and $D$.

Figure 8. Different regimes of network acceleration.

4.1. Note on hyper-acceleration

We note the existence of a turning point in the accelerating $a(t) > 0$ regime of network evolution. The acceleration regime $a(t) > n(t)$ cannot be sustained indefinitely as the number of added links per new node would have to exceed the number of existing nodes $N(t)$ which is not possible. As such, this behaviour is deemed hyper-acceleration if

$$a(t) > n(t).$$

If we reconsider the function $m(t)$ being an initially increasing function of time, asymptoting to a constant value, and then superimpose the appropriate contour lines, i.e. $y = \text{const}$ and $y = nx + \text{const}$, where coefficient $n$ corresponds to the constant rate of node addition $n(t) = n$, one can clearly identify the acceleration regimes for this particular evolving network. This evolution is illustrated in figure 7. We observe hyper-acceleration between times $A$ and $B$, acceleration between $B$ and $C$ and non-acceleration between $C$ and $D$. These acceleration regimes are shown schematically in figure 8.
4.2. Accelerating weighted networks

Having defined the accelerating unweighted evolving network, we now extend the concept to encompass weighted networks, an aspect of evolving networks attracting increasing attention [9, 10]. This is a relatively simple process. The key components in the definition of network acceleration for unweighted graphs were the rates of node and link addition, both of which are macroscopic properties of the system. Similarly, we can observe the macroscopic weight of the system, denoted \( L(t) \), which reflects the total weight of all the links within the network expressed as

\[
L(t) = \sum_{i=1}^{M(t)} w_i.
\]  

In a similar manner to the unweighted scenario (see equation 5 for comparison), we make use of the approximate derivatives and write

\[
l(t) = \frac{dL(t)}{dt} \approx L(t) - L(t - 1),
\]

\[
n(t) = \frac{dN(t)}{dt} \approx N(t) - N(t - 1).
\]  

For useful comparison between networks, it is important to normalize the weights such that for any link \( w \in [0, 1] \)—otherwise \( l(t) \) would vary enormously according to the type of network under consideration\(^5\). The total weight of the evolving network at time \( t \) is then constrained by the total number of possible links, given by \( L_{\text{max}}(t) = N(t)(N(t) - 1)/2 \).

The evolution of these rates for weighted graphs are then used to define velocity and acceleration as

\[
\tilde{v}(t) \equiv \frac{l(t)}{n(t)},
\]

\[
\tilde{a}(t) \equiv \frac{d}{dt} \left( \frac{l(t)}{n(t)} \right).
\]  

Note that the above weighted definitions of equation 11 will reduce to the unweighted definitions of equation 5 if the weights are made binary, i.e. \( \tilde{v}(t) \rightarrow v(t) \) and \( \tilde{a}(t) \rightarrow a(t) \). The above definition is then possibly the most general definition in that it can be applied to both weighted and unweighted, as well as directed and undirected, evolving out-of-equilibrium networks.

4.3. Stochastic accelerating networks

The definitions (weighted and unweighted) outlined earlier have been introduced for the scenario whereby the rates of node and link addition are not stochastic. However, they can be applied easily to stochastic scenarios. In this case, the corresponding measures would be the expected

\(^5\) The method of normalizing the weights is likely to be situation specific. A simple division by the largest weight in the network may suffice or some cumulative binning process might be appropriate. In either case, it is necessary to take care with respect to the statistical significance of outliers.
velocity $\langle v(t) \rangle$ and expected acceleration $\langle a(t) \rangle$ where the expectation $\langle \cdots \rangle$ is computed over an ensemble of network realizations. In certain cases, it might be possible to achieve this by simply replacing $n(t)$, $m(t)$ and $l(t)$ by their expectation values although in general this will not be suitable. For this to be appropriate we require rather specific constraints on the evolution of the system, namely that the rate of node addition $n(t)$ is deterministic and the rate of link (weight) addition is path independent. That is, $m(t)$ is not dependent on the number of links added at the last time step.

In general, to evaluate the required quantities properly, we must consider all possible contributing trajectories of the system’s evolution and their corresponding probabilities. We must incorporate the possibility that the rates of node and link addition at a given time step might not be independent of each other and that their outcomes might also influence the evolution of the network at the next time step. For the unweighted case, this would give

$$\langle v(t) \rangle = \sum_{a,b} P[(m(t) = a) \cap (n(t) = c)] \frac{a}{c},$$

$$\langle a(t) \rangle = \sum_{a,b,c,d} \left( P[(m(t+1) = b) \cap (n(t+1) = d) \cap (m(t) = a) \cap (n(t) = c)] \left( \frac{b}{d} - \frac{a}{c} \right) \right).$$

Evaluation of these quantities would be specific to the algorithm employed to generate the network.

5. Case study: Wikipedia

As a simple example application of the concepts introduced above, we look at the evolution of Wikipedia in three different languages, namely, in English, German and Japanese. Each of these is a distinct evolving network such that the nodes correspond to the articles within each language and the links correspond to the hyperlinks between those articles. The data is available in the public domain. The evolution through the macroscopic link-node phase space is shown in the upper panel of figure 9. All three networks appear to converge on the same non-accelerating behaviour.

While the public domain data is somewhat imprecise, the plot of the evolution of the velocity $v(t)$ of the system does indicate that all networks show an initial accelerating trend before non-accelerating behaviour is reached as shown in the lower plot of figure 9. There are no negative velocities as the total number of links and the total number of nodes are increasing in time for all three networks. It is interesting to note that acceleration of the growing Japanese site far exceeded that of the English and German sites. This simple example demonstrates that it is reasonable to consider network velocity and acceleration as time-dependent properties of a network’s evolution as opposed to considering them as static properties as suggested in [7]. It is also very straightforward to measure the introduced characteristics, velocity and acceleration, for empirical networks, which further supports their role in network characterization.

---

6 See http://stats.wikimedia.org/EN/Sitemap.htm for Wikipedia site statistics.
Figure 9. The evolution of Wikipedia for three different languages. In these networks, the nodes correspond to articles in English, German and Japanese. The links in each network correspond to hyperlinks connecting articles to others of the same language. The evolution through the macroscopic link-node phase space of total node and link numbers is shown in the top plot and the network velocity in the lower.

6. Conclusion

We have revisited the framework of network acceleration suggested by Gagen and Mattick [7]. We have explored the limits of the proposed definition of network acceleration and, based on our findings, have provided an alternative definition for accelerating networks. Perhaps most important is the conceptual difference between the two definitions: the concept of network acceleration as introduced in this paper refers to the properties of the network at a particular moment in time as opposed to an algorithm governing the evolution of the network as suggested in [7]. In addition to introducing the related concept of network velocity, we have augmented the definition of network acceleration to cover weighted networks as well. As such, the definition put forward in this paper holds for both weighted and unweighted, as well as directed and undirected networks. We have demonstrated the utility of these concepts by their simple application to study the evolution of Wikipedia in three different languages. Although the data obtained from
the public domain is not very precise, the results clearly support the conclusion that networks undergo different regimes of acceleration throughout their evolution. Since measurement of the proposed characteristics for empirical networks is very simple, we hope that the measures will find their use in the study of network evolution, in particular as time-stamped network data becomes increasingly available in the future.

Acknowledgments

DMDS acknowledges support from the European Union under the MMCOMNET program and J-PO by a Wolfson College Junior Research Fellowship (Oxford, UK).

References

[1] Mattick J S and Gagen M J 2005 Accelerating networks Science 307 856
[2] Dorogovtsev S N and Mendes J F F 2001 Effect of the accelerating growth of communications networks on their structure Phys. Rev. E 63 25101
[3] Dorogovtsev S N and Mendes J F F 2001 Scaling properties of scale-free evolving networks: Continuous approach Phys. Rev. E 63 056125
[4] Dorogovtsev S N and Mendes J F F 2003 Evolution of Networks: From Biological Nets to the Internet and WWW (Oxford: Oxford University Press)
[5] Vázquez A, Oliveira J G and Barabási A-L 2005 Inhomogeneous evolution of subgraphs and cycles in complex networks Phys. Rev. E 71 025103
[6] Palla G, Barabási A-L and Vicsek T 2007 Quantifying social group evolution Nature 446 664–7
[7] Gagen G M and Mattick J S 2005 Accelerating, hyperaccelerating and decelerating probabilistic networks Phys. Rev. E 72 16123
[8] Sen P 2004 Accelerated growth in outgoing links in evolving networks: deterministic versus stochastic picture Phys. Rev. E 69 46107
[9] Barrat A, Barthélemy M and Vespignani A 2004 The architecture of complex weighted networks Proc. Natl Acad. Sci. USA 101 3747
[10] Costa L F, Rodrigues F A, Travieso G and Villas Boas P R 2006 Characterization of complex networks: a survey of measurements Preprint cond-mat/0505185v5