

Some Models of Reproducing Graphs: II Age Capped Vertices

Richard Southwell, Chris Cannings

School of Mathematics and Statistics, University of Sheffield, Sheffield, United Kingdom

E-mail: bugzsouthwell@yahoo.com

Received June 26, 2010; revised August 10, 2010; accepted August 12, 2010

Abstract

In the prequel to this paper we introduced eight reproducing graph models. The simple idea behind these models is that graphs grow because the vertices within reproduce. In this paper we make our models more realistic by adding the idea that vertices have a finite life span. The resulting models capture aspects of systems like social networks and biological networks where reproducing entities die after some amount of time. In the 1940's Leslie introduced a population model where the reproduction and survival rates of individuals depends upon their ages. Our models may be viewed as extensions of Leslie's model-adding the idea of network joining the reproducing individuals. By exploiting connections with Leslie's model we are to describe how many aspects of graphs evolve under our systems. Many features such as degree distributions, number of edges and distance structure are described by the golden ratio or its higher order generalisations.

Keywords: Reproduction, Graph, Population, Leslie, Golden Ratio

1. Introduction

Networks are everywhere, wherever a system can be thought of as a collection of discrete elements, linked up in some way, networks occur. With the acceleration of information technology more and more attention is being paid to the structure of these networks, and this has led to the proposal of many models [1-3].

In many situations networks grow-expanding in size as material is produced from the inside, not added from outside. To study network growth we introduced a class of pure reproduction models [4,5], where networks grow because the vertices within reproduce. These models can be applied to many situations where entities are introduced which derive their connections from pre existing elements. Most obviously they could be used to model social networks, collaboration networks, networks within growing organisms, the internet and protein-protein interaction networks. One of our systems (model 3) has also been introduced independently [6], proposed as a model for the growth of online social networks.

In our pure reproduction models networks grow endlessly in a deterministic fashion. This allows a rigorous analysis, but costs a degree of realism. Nature includes birth and death and entities may be destroyed for

reasons of conflict, crowding or old age. In this paper we consider age; and extend our models by including vertex mortality.

2. The Models

In [5] we defined a set $\{F_m : m \in \{0,1,\dots,7\}\}$ of eight different functions F_m which map graphs to graphs. $F_m(G)$ is the graph obtained by simultaneously giving each of G 's vertices an offspring vertex and then adding edges according to some rule. The connections given to offspring depend upon the binary representation $\alpha\beta\gamma$ of m (i.e. $m = 4\alpha + 2\beta + \gamma$) as follows:

$\gamma = 1 \Leftrightarrow$ offspring are connected to their parent's neighbours,

$\beta = 1 \Leftrightarrow$ offspring are connected to their parents,

$\alpha = 1 \Leftrightarrow$ offspring are connected to their parent's neighbour's offspring.

In our age capped reproduction models we think of the vertices as having ages. Graphs grow under these models exactly as before, except that vertices grow and then die when their age exceeds some pre-specified integer Q . Our new update operator $T_{m,Q}$ is defined so that $T_{m,Q}(G)$ is the graph obtained by taking the graph G and performing the following process;

- 1) Increase the age of each vertex by one.
- 2) Give every vertex an age zero offspring, born with connectivity dependant upon m , as above (*i.e.* the new graph is $F_m(G)$).
- 3) Remove every vertex with age greater than the age cap Q .

We are interested in the sequence $\{G_t\}$ of graphs which evolve from an initial structure $G_0 = (V_0, E_0)$ in such a way that $G_{t+1} = T_{m,Q}(G_t), \forall t \geq 0$. We always suppose that initial vertices have age $\leq Q$.

3. The Number of Vertices

The number of vertices $|G_t|$ in G_t is deeply connected with the golden ratio and its generalisations. The number n_i^t of age i vertices in G_t can be conveniently described in terms of Leslie matrices.

In Leslie's population model [7,8] individuals of age i have a survival rate s_i and fertility rate f_i . The expected number of individuals of a given age, at a given time, is kept track of via repeated multiplication of the state vector with the 'Leslie matrix'

$$\begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_Q \\ s_0 & 0 & 0 & \dots & 0 \\ 0 & s_1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & s_Q & 0 \end{pmatrix}$$

In our case $n^{t+1} = L.n^t$ where $n^t = (n_0^t, n_1^t, \dots, n_Q^t)^T$ and L is the Leslie matrix with $s_i = f_i = 1, \forall i$.

L is a primitive matrix with characteristic polynomial

$$\sum_{i=0}^Q x^i = x^{Q+1}$$

and principle eigenvalue λ_Q (also known as the $Q+1$ step Fibonacci constant). The golden ratio is $\lambda_1 = \frac{1+5^{1/2}}{2}$, $\lambda_2 = 1.8393, \lambda_3 = 1.9276, \lambda_4 = 1.96559$ and $\lambda_Q \rightarrow 2$ as $Q \rightarrow \infty$. When t is large $n^{t+1} = \lambda_Q n^t$ where

$n^t = c(1, (\lambda_Q)^{-1}, (\lambda_Q)^{-2}, \dots, (\lambda_Q)^{-Q})^T$ is the stable age

distribution and c is a constant which depends upon the initial state of the system.

Let $d_i = (\delta_{i,0}, \delta_{i,1}, \dots, \delta_{i,Q})^T$ where $\delta_{i,j}$ is the Kronecker delta. The n step Fibonacci numbers $f_i^{[n]}$ are natural generalisations of the famous Fibonacci numbers [9] which can be generated by repeatedly multiplying d_0 by L . When G_0 is age zero (*i.e.* all its vertices have zero age) $n_i^t = |G_0| \cdot (L^t d_0)_i$, where $(L^t d_0)_i = f_{t+1-i}^{[Q+1]}$. In such a case G_t will have $|G_0| \cdot f_{t+2}^{[Q+1]}$ vertices.

4. Binary Strings

As we update our graphs, their vertex sets will grow, and a good way to keep track of these vertex sets is to use binary strings. Suppose v is a vertex of G_0 . When we update G we write $(v,0)$ and $(v,1)$ to denote v 's offspring, and v itself (respectively), in the graph G_1 . This means, for example, that $((v,0),0)$ is the grand child of $((v,1),1)$ in G_2 . We use short hand by omitting the parenthesis, so for example we write $((v,0),0)$ as $v00$. An example of the evolution of model 2 is shown in **Figure 1**.

When our age cap $Q = \infty$ an initial graph $G_0 = (V_0, E_0)$ will evolve in exactly the same way as in pure reproduction *i.e.* $G_t = F_m^t(G_0)$; this will have vertex set $V_0 \times \{0,1\}^t$ and edge set as specified in [5]. When Q is finite the situation is more complex, but our binary string notation allows us to keep track of the ages of vertices in a convenient way.

Let ab denote the concatenation of binary strings a and b and let a^t denote the string obtained by concatenating a with itself t times. Suppose $v(a01^n)$ is a vertex of G_t for some $a \in \{0,1\}^{t-1-n}$. Now $v(a0)$ is a new born offspring in G_{t-n} and every subsequent 1 in $v(a01^n)$'s name corresponds to an update within which

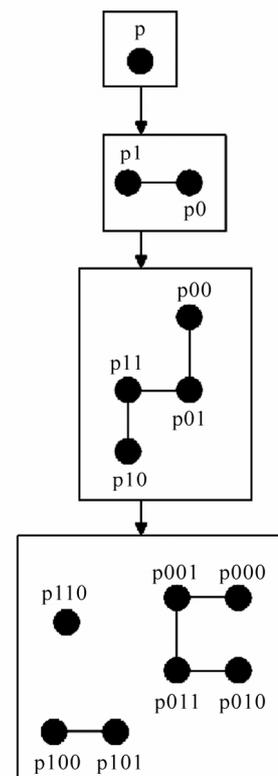


Figure 1. A depiction of the evolution of model 2 when $Q = 2$, starting with an isolated vertex named p .

this vertex survives as a parent and gets older by one. It follows that $v(a01^n)$ will be an age n vertex in G_t .

Theorem 1

If every vertex of the initial graph $G_0 = (V_0, E_0)$ is age zero then G_t will be the subgraph of $F_t(G_0)$ induced upon the vertex set $V_0 \times W_Q^t$, where $W_Q^t \subseteq \{0,1\}^t$ denotes the set of all t length binary strings which do not contain a run of $Q+1$ consecutive 1's.

Proof

Suppose G_t is the subgraph of $F_t(G_0)$ induced upon $V_0 \times W_Q^t$, as is clearly the case when $t = 0$. An age n vertex va in G_t , with $v \in V_0$, will produce an offspring $v(a0)$ in G_{t+1} . va will also survive to become $v(a1)$ iff $n < Q$. Such an a must be of the form 01^n or 1^n . It follows that G_{t+1} will be the subgraph of $F_{t+1}(G)$ induced upon $V_0 \times X$ where X is the set of all ax with $a \in W_Q^t$ and $x \in \{0,1\}$ such that $a_{t-n+1}a_{t-n+2} \dots a_t x \neq 1^{Q+1}$. Clearly X is W_Q^{t+1} so we can use induction with t to prove the result \square .

If the initial graph holds vertices of non-zero age; G_t can be obtained by taking the structure described in theorem 1 and removing every vertex of the form $v(1^n a)$, where n plus the age of v (in G_0) is greater than Q .

5. How Edges Connect Vertices of Different Ages

To keep track of the number of edges of G_t it helps to consider how vertices of different ages link to one another. Let S denote the Leslie matrix with all survival rates set at one and all fertility rates set at zero. Let F denote the Leslie matrix with all survival rates set at zero and all fertility rates set at one (note $L = S + F$). Let us define the age sampling vector $X = (X_0, X_1, \dots, X_Q)^T$ of a vertex to be such that X_i is the number of neighbours it has of age i .

Applying the $T_{m,Q}$ update will cause an age $n \leq Q$ vertex to have an offspring with age sampling vector

$$o_{m,n}(X) = (\gamma S + \alpha F).X + \beta(1 - \delta_{n,Q}).d_{n+1} \tag{1}$$

and also, provided $n < Q$, this vertex will also survive the $T_{m,Q}$ update to become a parent with age sampling vector

$$p_m(X) = (S + \gamma F).X + \beta.d_0. \tag{2}$$

Equations (1) and (2) describe how the age sampling vector of a vertex determines the age sampling vector of itself and its offspring on the next time step. Repeatedly using these equations allows us to understand how the history of a vertex relates to its connectivity. The sequence of zeros and ones in a tell us the sequence of birth and survival stages which lead to the creation of a vertex

va in G_t . Given this information one can compute the age sampling vector of va by performing the corresponding sequence of $o_{m,n}$ and p_m operations-starting with the age sampling vector of the initial vertex, v , in G_0 .

Let $e_{i,j}^t$ denote the number of edges of G_t that connect vertices of age i to vertices of age j . We consider how $e_{i,j}^t$ evolves in order to describe the growth rate of the number of edges in the different models. The vector $e_i^t = (e_{i,0}^t, e_{i,1}^t, \dots, e_{i,Q}^t)$ is equal to the sum of the age sampling vectors of all of G_t 's age i vertices and hence satisfies the equations

$$e_0^{t+1} = \sum_{j=0}^Q (\gamma S + \alpha F)e_j^t + \beta(1 - \delta_{j,Q})n_j^t.d_{j+1}, \tag{3}$$

$$i > 0 \Rightarrow e_i^{t+1} = (S + \gamma F).e_{i-1}^t + \beta.n_{i-1}^t.d_0. \tag{4}$$

Since the graphs we are concerned with are undirected we have $e_{i,j}^t = e_{j,i}^t, \forall i, j$.

The average asymptotic rates of increase of the minimal and maximal degrees for the different models are given in **Table 1**. We use the term *average* because, under some models, these extremal degrees increase at varying rates dependant upon the time modulo $Q+1$. These rates were found by determining which binary string describes a vertex with maximal (or minimal) degree and using Equations (1) and (2). For example, suppose the initial graph G_0 is age zero, and holds a vertex v with maximal degree, $deg(v)$, also suppose $t = n.(Q+1) + c$, for $c \leq Q$. When $m=1$ the vertex va , with $a = (1^{Q0})^n 1^c$, will have maximal degree in G_t . This vertex will have age sampling vector $L^c(S.L^Q)^n(deg(v).d_0)$. The degree of the vertex with this form will increase by $(\lambda_{Q-1})^Q$ every subsequent $Q+1$ time steps, and so it follows that the average asymptotic rates of increase of the maximal degree when $m=1$ is $(\lambda_{Q-1})^{Q/(Q+1)}$.

Table 1. A table showing the average asymptotic growth rates of the minimal and maximal degrees under the different models m . The notation $LIN(x)$ indicates that the extremal degrees increase linearly (as opposed to exponentially) with time with gradient x .

m	growth rate of the minimal degree	growth rate of the maximal degree
0	0	0
1	0	$(\lambda_{Q-1})^{Q/(Q+1)}$
2	0	1
3	1	$(\lambda_{Q-1})^{Q/(Q+1)}$
4	1	1
5	λ_Q	λ_Q
6	$LIN(Q)$	$LIN(Q+1)$
7	λ_Q	λ_Q

6. Connectivity, Degrees and Distances in Specific Models

In this section we will focus on reproduction mechanisms with $m \in \{1, 2, 3, 5, 6, 7\}$, one after another, and discuss the development of: connected components, number of edges, degree distributions, average path length and diameter. We do not discuss the dynamics when $m = 0$ or $m = 4$ because they are relatively uninteresting.

Before we discuss the specifics it is worth pointing out an effect that occurs under many models. We say that a graph is *age mixed* when each of its edges connect a pair of vertices with different ages.

If $\alpha = 0$ and $t > Q$ then G_t will be age mixed. The reason is that when $\alpha = 0$ offspring are not born connected to one another. So when $t > Q$ all of the initial vertices will be dead, and G_t will never again produce linked vertices with the same age.

Saying that G_t is age mixed has many implications, for example it means that G_t has chromatic number $\leq Q$ because its vertices may be coloured according to their ages.

6.1. Aspects of Model 1

Suppose $m = 1$ and we begin with a connected graph. G_t will typically consist of a growing connected component and lots of isolated vertices.

In the special case when $t > Q = 1$, updates do not cause the connected part of G_t 's structure to change. The reason is that G_t is age mixed and every new born vertex either has a dead parent, which it replaces, or no surviving neighbours.

Suppose $G_0 = (V_0, E_0)$ is an age zero graph with $u \in V_0$. Any vertex ua of G_t will be isolated iff a holds a run of $Q + 1$ consecutive zeros.

To see this note that theorem 1, together with results from [5], imply that vb will be a neighbour of ua iff $\{u, v\} \in E_0$, $b \in W_Q^t$ and $a_i = 0 \Rightarrow b_i = 1, \forall i$. Now if a holds a run of $Q + 1$ consecutive zeros then this means b holds a run of $Q + 1$ consecutive ones, which means $b \notin W_Q^t$, so no neighbour vb can actually exist. On the other hand if a does not hold a run of $Q + 1$ consecutive zeros, and $\{u, v\} \in E_0$ then $\{ua, vb\} \in E_t$, where $b_i = 1 - a_i, \forall i$.

Let Y_i^t denote the set of t length binary strings of the form ax^i where a is any $t - i$ length string which does not hold a run of $Q + 1$ consecutive 0's or a run of $Q + 1$ consecutive 1's, and $x \in \{0, 1\}: x \neq a_{t-i}$. By our argument above the number of non-isolated vertices in G_t will be $|G_0| \sum_{i=1}^Q |Y_i^t|$. For example consider a string $100110 \in Y_1^6$ (when $Q = 2$), this string can be

thought of as being responsible for generating the strings $1001101 \in Y_1^7$ and $1001100 \in Y_2^7$ at the next time step. Following similar reasoning one can see that, for generic Q , we have the difference equation;

$$\begin{pmatrix} |Y_1^{t+1}| \\ |Y_2^{t+1}| \\ |Y_3^{t+1}| \\ \vdots \\ |Y_Q^{t+1}| \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} |Y_1^t| \\ |Y_2^t| \\ |Y_3^t| \\ \vdots \\ |Y_Q^t| \end{pmatrix},$$

which involves the $Q \times Q$ Leslie matrix. It follows that the number of non-isolated vertices in G_t increases asymptotically at a rate of λ_{Q-1} (whilst the total number of vertices grows at a rate of λ_Q), meaning eventually almost every vertex of G_t will be isolated.

Although vertices do get destroyed during the the $T_{1,Q}$ they are always replaced by offspring which link to their old neighbours. The only way that a component of G_0 could be disconnected under the update (in a non-trivial way) is if G_0 has a cutset of edges that connect pairs of age Q vertices. This can only happen during the first Q updates.

Regarding the edges, Equations (3) and (4) lose their dependence upon n_i^t and $t > Q \Rightarrow e_{i,i}^t = 0, \forall i$. Given these considerations we can reduce (3) and (4) to the following system of linear difference equations:

$$\forall i, j \in \{1, 2, \dots, Q\} \text{ such that } i < j,$$

$$e_{0,i}^{t+1} = \sum_{n=0}^{i-2} e_{n,i-1}^t + \sum_{n=i}^Q e_{i-1,n}^t \tag{5}$$

$$e_{i,j}^{t+1} = e_{i-1,j-1}^t. \tag{6}$$

We can cast this system as a matrix difference equation which describes the evolution of

$$\begin{pmatrix} e_{0,1}^t \\ e_{0,2}^t \\ \vdots \\ e_{0,Q}^t \\ e_{1,2}^t \\ e_{1,3}^t \\ \vdots \\ e_{1,Q}^t \\ e_{2,3}^t \\ \vdots \\ e_{2,Q}^t \\ \vdots \\ e_{Q-1,Q}^t \end{pmatrix} \tag{7}$$

The matrix which describes how (7) changes is primitive with principle eigenvalue μ_Q so the number of edges in G_t increases at a rate of μ_Q asymptotically. $\mu_1 = 1$, $\mu_2 = \lambda_2 = 1.8393$, $\mu_3 = 2.3336$ and $\mu_4 = 2.6077$. In general μ_Q increases with Q and $\mu_\infty = 3$.

Let $D_t(n)$ denote the number of vertices of degree n in G_t . Computer simulations suggest that when $1 \ll t < n$ we have $D_t(\lfloor \lambda_{Q-1} n \rfloor) = (\lambda_{Q-1})^{-1} \cdot D_t(n)$ so it appears at the high end, as if the distribution obeys a geometric law. Whilst it seems there is some pattern in the degree distribution at the high degrees, the behaviour of the distribution of the lower degree vertices is more mysterious. For example it appears that when $Q > 1$ there will be less degree 1 vertices than degree 2 vertices when t is large.

Global notions of distance (such as diameter) do not really make sense when $m = 1$ because the structure is disconnected, with many isolated vertices.

6.2. Aspects of Model 2

Introducing an age cap into the $m = 2$ model leads to fascinating self replicative behaviour. Whatever graph we begin with we end up with a set of special tree graphs that grow up and break into more tree graphs. Let S_Q^t denote the graph obtained by starting with an age zero isolated vertex and evolving updating it with $T_{2,Q}$, t times. This graph will have vertex set W_Q^t and a pair of vertices a, b will be adjacent iff $\exists k \in \{1, 2, \dots, t\}$ such that $\forall i \in \{1, 2, \dots, t\}$ we have $i < k \Rightarrow a_i = b_i$, $i = k \Rightarrow a_i \neq b_i$ and $i > k \Rightarrow a_i = b_i = 1$.

To understand the self replicative behaviour in S_Q^t it helps to understand the self similarity of S_∞^t , the oldest and most central vertex of which is 1^t . Consider any neighbour $1^{t-n-1}01^n$ of 1^t . Since structures grow out of every vertex in the same way; the subgraph, X_n^t , induced upon the vertices $\{1^{t-n-1}0x : x \in \{0, 1\}^n\}$ which grew out of $1^{t-n-1}01^n$, over n time steps, will be age-isomorphic to the graph S_∞^n which grew out of the initial vertex, over n time steps (by *age-isomorphic* we mean there is a one to one mapping, from one vertex set to the other, which preserves the adjacency, non-adjacency and ages of the vertices).

More generally $S_Q^t = S_\infty^t$ when $t \leq Q$, and S_Q^{Q+1} is the graph obtained by taking S_∞^{Q+1} and removing the oldest vertex, 1^{Q+1} . Since S_Q^{Q+1} is a tree, the removal of 1^{Q+1} causes the graph to break into numerous components, namely $X_0^{Q+1}, X_1^{Q+1}, \dots, X_Q^{Q+1}$. Since X_n^{Q+1} is age-isomorphic to S_∞^n , it follows that S_Q^{Q+1} consists of $Q+1$ different connected components, one age-isomorphic to S_∞^n for each $n \leq Q$. Each of these connected components will evolve in the same manner-growing until the age of its central vertex exceeds Q , at which point it will

fragment into yet more of these special trees.

Any initial graph will evolve to become a set of these trees after $Q+1$ time steps. The reason is that when $t = Q+2$ all of the initial vertices will have died. This means the oldest surviving ancestor of any vertex in G_{Q+2} will be a vertex which was born when $t > 0$. If a pair of vertices lie within the same connected component in G_{Q+2} then they will have the same oldest surviving ancestor, and it follows that every connected component of G_{Q+2} is a tree structure which grew out of a vertex that was not initially present, and is hence age isomorphic to S_∞^n for $n \leq Q$.

Let C_n^t denote the number of connected components of G_t which are age-isomorphic to S_∞^n for $n \leq Q$. The vector $C^t = (C_0^t, C_1^t, \dots, C_Q^t)^T$ satisfies the matrix equation $C^{t+1} = \Gamma \cdot C^t$, where

$$\Gamma = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & 0 & 1 \\ 0 & 1 & \dots & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 1 \end{pmatrix}$$

is equivalent to the transpose of the Leslie matrix L . It follows that when t is large C_n^t will increase at a rate of λ_Q and the probability that a random connected component is age-isomorphic to S_∞^i will be

$$P_i = \frac{\sum_{x=0}^i \lambda_Q^{-x}}{\sum_{y=0}^Q \sum_{x=0}^y \lambda_Q^{-x}} = \frac{1 - \lambda_Q^{-(i+1)}}{Q+1 + \left(\frac{1 - \lambda_Q^{-(Q+1)}}{1 - \lambda_Q} \right)} \quad (8)$$

The number of edges is described by the equations:

$$e_0^{t+1} = \sum_{j=0}^{Q-1} n_j^t \cdot d_{j+1}, \quad (9)$$

$$i > 0 \Rightarrow e_i^{t+1} = S \cdot e_{i-1}^t + n_{i-1}^t \cdot d_0. \quad (10)$$

When $t > Q$ we will have $e_{i,i}^t = 0$, $\forall i$. This implies that $\forall i, j \in \{0, 1, \dots, Q\} : i < j$ we have $e_{i,j}^t = n_{j-i-1}^{t-i-1}$ and so the number of edges in G_t increases at a rate of λ_Q asymptotically.

We can gain the asymptotic form of the degree distribution of G_t . First note that the graph S_∞^i has 2^i vertices. The number of degree k vertices in S_∞^i will be

$$\Delta_k^i = \left(\lfloor 2^{i-k} \rfloor + \delta_{k,i} \right) (1 - \delta_{k,0}) + \delta_{i,0} \cdot \delta_{k,0}. \quad (11)$$

Now the probability $\frac{D_t(k)}{|G_t|}$ that a randomly selected

vertex of G_t will be of degree k will be equal to [the probability that a randomly selected vertex of G_t be-

longs to a connected component isomorphic to S_∞^i] times [the probability that a randomly selected vertex of S_∞^i will have degree k], summed over all $i \in \{0, 1, \dots, Q\}$.

For large t the probability that a randomly selected vertex of G_t belongs to a connected component isomorphic to S_∞^i is

$$\frac{p_i \cdot 2^i}{\sum_{r=0}^Q p_r \cdot 2^r} = \frac{2^i (1 - \lambda_Q^{-(i+1)})}{\Omega_Q} \tag{12}$$

where

$$\Omega_Q = 2^{Q+1} - 1 + \lambda_Q^{-1} \left(\frac{1 - \left(\frac{2}{\lambda_Q}\right)^{Q+1}}{\left(\frac{2}{\lambda_Q}\right) - 1} \right). \tag{13}$$

The probability that a randomly selected vertex of S_∞^i will be of degree k will be

$$\frac{(\lfloor 2^{i-k} \rfloor + \delta_{k,i})(1 - \delta_{k,0}) + \delta_{i,0} \cdot \delta_{k,0}}{2^i}. \tag{14}$$

Hence as $t \rightarrow \infty$ we have $\frac{D_t(k)}{|G_t|}$ will be equal to

$$\sum_{i=0}^Q \frac{2^i (1 - \lambda_Q^{-(i+1)}) (\lfloor 2^{i-k} \rfloor + \delta_{k,i})(1 - \delta_{k,0}) + \delta_{i,0} \cdot \delta_{k,0}}{\Omega_Q \cdot 2^i}. \tag{15}$$

Suppose t is large.

$$\frac{D_t(0)}{|G_t|} = \frac{1 - \lambda_Q^{-1}}{\Omega_Q}, \tag{16}$$

$$\frac{D_t(Q)}{|G_t|} = \frac{2(1 - \lambda_Q^{-(Q+1)})}{\Omega_Q}, \tag{17}$$

If $1 \leq k \leq Q-1$ then $\frac{D_t(k)}{|G_t|}$ will be

$$\frac{2}{\Omega_Q} \left(2^{Q-k} - \lambda_Q^{-(k+1)} - \lambda_Q^k \cdot \left(\frac{\left(\frac{2}{\lambda_Q}\right)^{Q-k} - 1}{\left(\frac{2}{\lambda_Q}\right) - 1} \right) \right), \tag{18}$$

and if $k \notin \{0, 1, 2, \dots, Q\}$ then $\frac{D_t(k)}{|G_t|} = 0$.

Once again we do not discuss distances because global notions of distance do not really make sense upon graphs which constantly disconnect.

6.3. Aspects of Model 3

Growth model 3 produces complicated structures; we

can say a little about their connectivity using reasoning like that used when $m=1$. Since newborn vertices are never linked, $t > Q$ implies that G_t will not hold any linked vertices with the same age. If G_0 is connected then G_t will usually be connected. When G_0 has a cutset of edges connecting pairs of vertices with the same age then $T_{3,Q}(G_0)$ will be disconnected. This is the only way that structures can become disconnected, and it can only happen during the first Q updates.

With respect to edge numbers, there are many similarities in the way that G_t evolves when $m=1$ and $m=3$. The only difference is that when $m=3$ offspring are connected to their parents, and this means that the equations which describe the evolution of $e'_{i,j}$ gain a dependance upon the number of vertices. When $t > Q$ we will have $e'_{i,i} = 0, \forall i$, and we will hence have that:

$$\forall i, j \in \{1, 2, \dots, Q\} \text{ such that } i < j,$$

$$e'_{0,i}{}^{t+1} = \sum_{n=0}^{i-2} e'_{n,i-1}{}^t + \sum_{n=i}^Q e'_{i-1,n}{}^t + n'_{i-1}{}^t \tag{19}$$

$$e'_{i,j}{}^{t+1} = e'_{i-1,j-1}{}^t \tag{20}$$

In the $m=1$ case the number of edges increase asymptotically at a rate of μ_Q . The $m=3$ case is similar except that the number of edges is bolstered by the number of vertices n'_i , which increases at a lesser rate of λ_Q . For large t the effect of these additional edges is hence negligible and the number of edges again increases at a rate of μ_Q .

Like the $m=1$ case computer simulations again suggest that when $1 \ll t < n$ we have

$$D_t(\lfloor \lambda_{Q-1} n \rfloor) = (\lambda_{Q-1})^{-1} \cdot D_t(n), \tag{21}$$

so the distribution again obeys a geometric law at the high end.

When $Q=1$ we can describe the evolution of the degree distribution exactly for any initial graph G_0 that is age mixed with no isolated vertices. Applying $T_{3,1}$ to G_0 is equivalent to changing the age of each vertex (from 0 to 1, and from 1 to 0) and then, for each age 1 vertex v , adding an age zero vertex that is only adjacent to v .

Let $N'_x(d)$ denote the number of vertices of age x and degree d in G_t . $\forall t \geq 0$ we have

$$N_1^{t+1}(1) = 0, \tag{22}$$

$$N_0^{t+1}(1) = N_1^t(1) + \sum_{i=1}^\infty N_0^t(i) = N_1^t(1) + n_0^t, \tag{23}$$

$$d > 1 \Rightarrow N_0^{t+1}(d) = N_1^t(d), \tag{24}$$

$$N_1^{t+1}(d) = N_0^t(d-1). \tag{25}$$

Solving these equations implies $\forall t, d > 0 \quad \forall x \in \{0, 1\}$

that when $d > \frac{t+x+1}{2}$ we have

$$N'_x(d) = (1 - \delta_{1,x} \delta_{1,d}) N^0_{t+x \bmod 2} (d - \lfloor (t+x)/2 \rfloor), \quad (26)$$

when $d = \frac{t+x+1}{2}$ we have

$$N'_x(d) = n_0^0 + N_1^0(1), \quad (27)$$

and when $d < \frac{t+x+1}{2}$ we have

$$N'_x(d) = n_0^{t+x+1-2d}. \quad (28)$$

When we introduce mortality our graphs seem to get longer. Diameter and average path length become greater. This is a result of the death of old vertices (which tend to be more central), this decreases the ease with which one can travel between the extremities.

Let L_t denote the sum of $d(u,v)$ for each ordered pair of vertices (u,v) in G_t . The average length l_t is equal to $L_t / |G_t|^2$. Let $U^t_{i,j}$ denote the sum of the distances $d(u,v)$ between each ordered pair of vertices (u,v) from G_t such that either u is of age i , v is of age j or u is of age j , v is of age i . When $Q=1$ the average length is given by

$$l_t = \frac{U^t_{0,0} + U^t_{0,1} + U^t_{1,1}}{(n_0^t + n_1^t)^2}. \quad (29)$$

When $m=3$ it seems as if both the average length and diameter of G_t increase linearly with t whenever Q is finite. In the special case where $Q=1$ we can gain an exact description.

Suppose that $G=(V,E)$ is a connected age mixed graph; if u and v are age zero vertices of G then, after applying the $T_{3,1}$ update, $d(u1,v1) = d(u,v)$, $d(u0,v1) = d(u,v) + 1$ and (provided $u \neq v$), $d(u0,v0) = d(u,v) + 2$. If u is age zero and v is age one in $T_{3,1}(G)$ then after the update we have $d(u1,v0) = d(u,v)$ and $d(u0,v0) = d(u,v) + 1$. If u and v are both of age one then after the update $d(u0,v0) = d(u,v)$. The diameter of G_t will increase by two every two time steps and moreover the system obeys the equations

$$U^{t+1}_{0,0} = U^t_{0,0} + U^t_{0,1} + U^t_{1,1} + 2n_0^t(n_0^t + n_1^t - 1), \quad (30)$$

$$U^{t+1}_{0,1} = U^t_{0,1} + 2(U^t_{0,0} + n_0^t n_0^t) \quad (31)$$

$$U^{t+1}_{1,1} = U^t_{0,0}. \quad (32)$$

These equations imply that L_{t+1} is equal to

$$2.L_t + 2.L_{t-1} - L_{t-2} + 4(n_0^t)^2 + 2.n_1^t(3.n_0^t - 2) + 2.n_1^{t-1}(1 - n_0^{t-1}), \quad (33)$$

which means that when t is large, the average length

increases linearly with

$$l_{t+1} = l_t + \frac{8 + 14.\lambda_1}{10 + 15.\lambda_1}. \quad (34)$$

For $Q > 1$ we have that $T'_{3,1}(G)$ is a partial subgraph of $T'_{3,Q}(G)$ which is a partial subgraph of $T'_{3,\infty}(G)$. This implies that the curve which describes l_t , for generic Q is bounded below by a constant (because of the $Q = \infty$ case, see [5]) and bounded above by a straight line.

6.4. Aspects of Model 5

In this case, when G_0 is age zero, G_t may be obtained by replacing each vertex v of G_0 with a cluster C_v of $f^{[Q+1]}_{t+2}$ isolated vertices, and then connecting each vertex of C_u to each vertex of C_v whenever u and v where adjacent in G_0 . It follows that

$$D_t(n) = f^{[Q+1]}_{t+2} . D_0 \left(\frac{n}{f^{[Q+1]}_{t+2}} \right). \quad (35)$$

Equations (3) and (4) which describes the development of the edges may be cast as the matrix equation

$$\begin{pmatrix} e_0^{t+1} \\ e_1^{t+1} \\ e_2^{t+1} \\ \vdots \\ e_Q^{t+1} \end{pmatrix} = \begin{pmatrix} L & L & L & \cdots & L \\ L & 0 & 0 & \cdots & 0 \\ 0 & L & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & L & 0 \end{pmatrix} \begin{pmatrix} e_0^t \\ e_1^t \\ e_2^t \\ \vdots \\ e_Q^t \end{pmatrix}.$$

The matrix involved is clearly the Kronecker product of L with itself, it is hence primitive with principle eigenvalue λ_Q^2 . It follows that the asymptotic growth rate of the number of edges will be λ_Q^2 .

Suppose our initial graph is connected, non-trivial and age zero. G_t can be obtained by replacing each vertex by a cluster of $f^{[Q+1]}_{t+2}$ vertices. This means every ordered pair (u,v) such that $d(u,v) = k$, in the initial graph gives rise to $f^{[Q+1]}_{t+2} . f^{[Q+1]}_{t+2}$ ordered pairs, spaced by distance k , in G_t . In addition to this, every cluster adds $2.f^{[Q+1]}_{t+2}(f^{[Q+1]}_{t+2} - 1)$ to the total distance, by the fact that every pair of distinct vertices within a given cluster will be spaced by distance 2. It follows that the total distance of G_t will be

$$L_t = f^{[Q+1]}_{t+2} . f^{[Q+1]}_{t+2} L_0 + 2.f^{[Q+1]}_{t+2} (f^{[Q+1]}_{t+2} - 1) |G_0|. \quad (36)$$

This means (irrespective of Q) that when t is large the average length approaches the constant $l_t = l_0 + \frac{2}{|G_0|}$.

The diameter of G_t will be the maximum of the diameter

of G_0 and 2.

6.5. Aspects of Model 6

In this case G_t will be a connected graph that can be obtained by taking the t dimensional hypercube graph and removing some vertices. The next graph in a sequence can be obtained by fusing together previous structures. For example when $t > Q = 1$, G_{t+1} can be obtained by taking the disjoint union of G_t and G_{t-1} , choosing an isomorphism f from G_{t-1} to the subgraph of G_t induced upon its age zero vertices (such an isomorphism always exists) and adding an edge from each v vertex of G_{t-1} to $f(v)$. The age one vertices of G_{t+1} will be those which came from G_{t-1} , vertices which came from G_t will be age zero.

The number of edges $\|G_t\|$, in G_t , satisfies $e_{0,0}^{t+1} = \|G_t\|$, $e_{i,j}^{t+1} = e_{i-1,j-1}^t$ and $e_{0,i}^{t+1} = n_{i-1}^t$, $\forall i, j > 0$. This implies

$$\|G_{t+1}\| = \sum_{i=0}^Q \sum_{j=i}^Q e_{i,j}^{t+1}, \tag{37}$$

we can split the sums to get

$$\|G_{t+1}\| = \sum_{i=0}^Q e_{i,i}^{t+1} + \sum_{i=0}^{Q-1} \sum_{j=i+1}^Q e_{i,j}^{t+1} \tag{38}$$

making substitutions we find

$$\|G_{t+1}\| = \sum_{i=0}^Q \|G_{t-i}\| + \sum_{i=0}^{Q-1} \sum_{j=0}^{Q-i-1} n_j^{t-i}. \tag{39}$$

When t is large the minimal degree of G_t becomes large-implying that the average degree also becomes large. This implies

$$\sum_{i=0}^Q \|G_{t-i}\| \sum_{i=0}^{Q-1} \|G_{t-i}\| > \sum_{i=0}^{Q-1} \sum_{j=0}^{Q-i-1} n_j^{t-i} \tag{40}$$

and so the asymptotic growth rate of the number of edges will be λ_Q .

Determination of the degree distribution when $m = 6$ appears to be a difficult problem. Although some progress can be made when $Q = 1$ the resulting formulae are long and complicated.

With respect to distances it appears that the diameter and average length of G_t increase linearly when t is large. We can show explicitly that this is the case when $Q = 1$.

We say a graph is zero spanning if there is a shortest path between each pair of age zero vertices that only passes through age zero vertices. Updating any connected graph with $T_{6,1}$ will always yield a zero spanning graph. Supposing that G_t is a zero spanning graph, if u and v are age zero vertices of G_t then after updating with $T_{6,1}$ we will have $d(u1, v1) = d(u0, v0) = d(u, v)$ and

$d(u0, v1) = d(u, v) + 1$. If u is age zero and v is age one in G then after the update we will have

$d(u1, v0) = d(u, v) + 1$ and $d(u0, v0) = d(u, v)$. If u and v are both age one vertices of G then after updating we will have $d(u0, v0) = d(u, v)$. This implies that the system obeys the equations:

$$U_{0,0}^{t+1} = U_{0,0}^t + U_{0,1}^t + U_{1,1}^t \tag{41}$$

$$U_{0,1}^{t+1} = U_{0,1}^t + 2(U_{0,0}^t + n_0^t(n_0^t + n_1^t)) \tag{42}$$

$$U_{1,1}^{t+1} = U_{0,0}^t \tag{43}$$

These equations imply that as $t \rightarrow \infty$ the average length increases linearly with

$$l_{t+1} = l_t + \frac{2}{5}$$

The reasoning behind this is very similar to that when $Q = 1$ and $m = 3$. The diameter of G_t will increase by one every time step once the graph becomes zero spanning.

6.6. Aspects of Model 7

When our initial graph G_0 is age zero G_t may be obtained by replacing each vertex v of G_0 with a complete graph K_v on $f_{t+2}^{[Q+1]}$ vertices, and then connecting each vertex of K_u to each vertex of K_v whenever u and v where adjacent in G_0 . It follows that

$$D_t(n) = f_{t+2}^{[Q+1]} \cdot D_0 \left(\frac{n+1 - f_{t+2}^{[Q+1]}}{f_{t+2}^{[Q+1]}} \right). \tag{44}$$

With respect to the edges this case is similar to the $m = 5$ case, except that there is an extra dependence upon n_i^t caused by the presence of edges linking offspring to their parents.

$$\begin{pmatrix} e_0^{t+1} \\ e_1^{t+1} \\ e_2^{t+1} \\ \vdots \\ e_Q^{t+1} \\ n^{t+1} \end{pmatrix} = \begin{pmatrix} L & L & L & \cdots & L & S \\ L & 0 & 0 & \cdots & 0 & B^0 \\ 0 & L & 0 & \cdots & 0 & B^1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & L & 0 & B^{Q-1} \\ 0 & 0 & \cdots & 0 & 0 & L \end{pmatrix} \begin{pmatrix} e_0^t \\ e_1^t \\ e_2^t \\ \vdots \\ e_Q^t \\ n^t \end{pmatrix}$$

where B^n is the $(Q+1) \times (Q+1)$ matrix such that $\forall i, j \in \{0, 1, \dots, Q+1\}$ we have $B_{i,j}^n = 0$, except that $B_{0,n}^n = 1$.

In the $m = 5$ case the number of edges increase asymptotically at a rate of λ_Q^2 . This $m = 7$ case is similar except that the number of edges is bolstered by the number of vertices n_i^t , which increases at a lesser rate of λ_Q . For large t the effect of these additional edges is hence negligible and the number of edges again increases at a rate of λ_Q^2 .

Suppose our initial graph is connected, non-trivial and age zero. G_t can be obtained by replacing each vertex by a complete graph on $f_{t+2}^{[Q+1]}$ vertices. This means every ordered pair (u, v) such that $d(u, v) = k$, in the initial graph gives rise to $f_{t+2}^{[Q+1]} \cdot f_{t+2}^{[Q+1]}$ ordered pairs, spaced by distance k , in G_t . In addition to this, every cluster adds $f_{t+2}^{[Q+1]} \cdot (f_{t+2}^{[Q+1]} - 1)$ to the total distance, by the fact that every pair of distinct vertices within a given cluster will be spaced by distance 1. It follows that the total distance of G_t will be

$$L_t = f_{t+2}^{[Q+1]} \cdot f_{t+2}^{[Q+1]} L_0 + f_{t+2}^{[Q+1]} (f_{t+2}^{[Q+1]} - 1) |G_0|. \quad (45)$$

Interestingly this means that when t is large l_t loses its dependence upon Q and approaches the constant $l_t = l_0 + \frac{1}{|G_0|}$. The diameter of G_t will be equal to the diameter of G_0 .

7. Discussion

We have discussed many properties of age capped models, however many open problems remain. These include describing degree distribution when $m \in \{1, 3, 6\}$ and demonstrating the linearity of average length when $m \in \{3, 6\}$ (for generic Q).

There are many directions in which our models may be expanded. As highlighted by theorem 1, our models may be regarded as an extension of pure reproduction models by adding restrictions upon the language of binary strings which the vertices can possess. Many other restrictions could be considered, e.g. forbidding the subword $1^{Q+1}0$ (which would correspond to saying vertices of age $> Q$ become infertile).

Our models can be viewed as an extension of Leslie's population model, introducing the idea of a network which connects the reproducing individuals. We will further develop this connection by considering the evolution of generic Leslie matrices (so that individuals of a given age can have differing numbers of offspring and chances of survival). Taking this approach and considering connectivity as stochastic (so that α, β and γ are probabilities, rather than binary integers) should yield models which directly simulate the development of animal social networks and other phenomena.

This paper demonstrates how our original reproducing graph models can be generalised in different directions whilst remaining analytically tractable. Perhaps the main reason these models are amenable to analysis is that the growth of one part of a graph is not influenced by the structure of another. This spatial independence allows

one to understand the evolution of generic structures by studying the evolution of simple ones.

There are many extensions of these models that it would be interesting to consider. In the future papers we will discuss the fascinating dynamics which can ensue when game theory is incorporated into these models. In this case we lose the spatial independence and dynamics of immense complexity become possible. It is also possible to extend many of the results here to cases where individuals produce several offspring-connected up in different ways. This kind of generalisation allows one model how the social networks of specific types of organisms grow in a more direct way.

8. Acknowledgements

The authors gratefully acknowledge support from the EPSRC (Grant EP/D003105/1) and helpful input from Dr Jonathan Jordan, and others in the Amorph Research Project (www.amorph.group.shef.ac.uk).

9. References

- [1] P. Erdős and A. Rényi, "On Random Graphs. I," *Publications Mathematicae*, Vol. 6, 1959, pp. 290-297.
- [2] G. U. Yule, "A Mathematical Theory of Evolution, Based on the Conclusions of Dr. J. C. Willis, F. R. S.," *Philosophical Transactions of the Royal Society of London, B*, Vol. 213, 1925, pp. 21-87.
- [3] D. J. Watts and S. H. Strogatz, "Collective Dynamics of 'Small-World' Networks," *Nature*, Vol. 393, No. 6684, 1998, pp. 440-442.
- [4] R. Southwell and C. Cannings, "Games on Graphs that Grow Deterministically," *Proceedings of International Conference on Game Theory for Networks GameNets '09*, Istanbul, Turkey, 2009, pp. 347-356.
- [5] R. Southwell and C. Cannings, "Some Models of Reproducing Graphs: I Pure Reproduction," *Journal of Applied Mathematics*, Vol. 1, No. 3, 2010, pp. 137-145.
- [6] A. Bonato, N. Hadi, P. Horn, P. Praalat and C. Wand, "Models of On-Line Social Networks," To appear in *Internet Mathematics*, 2010.
- [7] P. H. Leslie, "The Use of Matrices in Certain Population Mathematics," *Biometrika*, Vol. 30, 1945, pp. 183-212.
- [8] P. H. Leslie, "Some Further Notes on the Use of Matrices in Population Mathematics," *Biometrika*, Vol. 35, No. 3-4, 1948, pp. 213-245.
- [9] T. D. Noe, "Primes in Fibonacci n-step and Lucas n-step Sequences," *Journal of Integer Sequences*, Vol. 8, 2005, pp. 1-12.