# Introduction

The idea of mimicking the propagation of biological epidemics to achieve diffusion of useful information was first proposed in the late 1980s, the decade that also saw the appearance of computer viruses. Back then, these viruses propagated by copies on floppy disks and caused much less harm than their contemporary versions. But it was already noticed that they evolved and survived much as biological viruses do, a fact that prompted the idea of putting these features to good (rather than evil) use. The first application to be considered was synchronisation of distributed databases.

Interest in this paradigm received new impetus with the advent of peer-to-peer systems, online social systems and wireless mobile ad hoc networks in the early 2000s. All these scenarios feature a complex network with potentially evolving connections. In such large-scale dynamic environments, epidemic diffusion of information is especially appealing: it is decentralised, and it relies on randomised decisions which can prove as efficient as carefully made decisions. Detailed accounts of epidemic algorithms can be found in papers by Birman *et al.* [12] and Eugster *et al.* [33]. Their applications are manifold. They can be used to perform distributed computation of global statistics in a spatially extended environment (e.g. mean temperature seen by a collection of sensors), to perform real-time delivery of video data streams (e.g. to users receiving live TV via peer-to-peer systems over the internet) and to propagate updates of dynamic content (e.g. to mobile phone users whose phone operating system requires patching against vulnerabilities).

This book is meant as an introduction for applied mathematicians and computer scientists to the study of epidemic propagation over complex networks. The book's first purpose is to provide the reader with an accessible introduction to the elementary models of epidemic propagation, and develop an understanding of the basic *phase transition* phenomena (also called *threshold* phenomena) that are typical of epidemic behaviour. To this end it introduces the

relevant generic models, analytical tools, and mathematical results. This book also aims to explain the role played by network topology in the propagation of information.

This book can be used both to accompany courses for graduate students in computer science and applied probability and to provide an overview of the probabilistic techniques widely used to study random processes on complex networks. It evolved from a graduate course given by the authors at Cambridge University from 2005 to 2007. The mathematical prerequisites are reasonable maturity in probability theory at undergraduate level. The main tools introduced in the text consist of large deviations inequalities, coupling of stochastic processes and Poisson approximation.

Readers who want to delve deeper will find thorough treatments of random graphs in Bollobás [13] and Janson *et al.* [42]. The recent book by Durrett [28] also expands on some of the material here, in several directions, in particular on the behaviour of random walks on graphs. Andersson and Britton [3] deals with epidemic modelling in the biological context. Specific suggestions for further reading are also provided at the end of each chapter.

## A tour of the book

The text is divided into two distinct but related parts: *shapeless networks* and *structured networks*. The first part (Chapters 1–5) presents techniques for analysing "homogeneous mixing" epidemic processes. In this setting the large population in which the disease spreads has no particular (or totally random, symmetric) topological structure; that is, each individual is the "neighbour" of every other individual. The infection can therefore spread from any infected individual to any healthy one. In particular, this part of the book introduces branching processes, Erdős–Rényi random graphs and so-called Reed–Frost epidemics. Issues such as ultimate outreach and time to global infection are considered.

The second part (Chapters 6–9) covers recent results on the spread of epidemics in structured networks wherein individuals interact with a limited set of neighbours, and where the corresponding topology can exhibit rich structure. It introduces models of such topologically structured networks, including power-law random graphs and navigable small-world graphs. It gives explanatory models for the emergence of such structures, and addresses the impact of structure on the behaviour of epidemics. It also touches upon the algorithmic issue of maximising outbreak as a function of the initial infectives. In what follows we describe the content of the book in more detail and give some

perspective on the relevance of each topic to the analysis of epidemic processes on networks.

Chapter 1 reviews elementary results on the Galton–Watson branching process. Branching processes arise naturally in the study of epidemics, as they provide an accurate description of epidemic behaviour in large, shapeless populations, at least in the early stages of dissemination. They are arguably the simplest class of model that exhibits a *phase transition*: the phenomenon by which an infinitesimal variation of microscopic characteristics (here, the offspring distribution) can lead to a macroscopic change in system behaviour (here, infinite survival of the epidemic).

The results of Chapter 1 are then exploited to establish a similar phase transition in a classical epidemic model, namely the Reed–Frost model. In particular, Chapters 2–4 exploit a parallel between the Reed–Frost epidemic model and Erdős–Rényi random graphs to gain insight into the behaviour of the former. Chapter 2 describes a phase transition appearing in the fraction of ultimately infected individuals. Chapter 3 identifies under which parameter ranges an epidemic starting from one infected individual eventually spreads to the whole population. In Chapter 4 we derive an upper bound on the time needed for the epidemic to reach the whole population. Chapter 5 describes a setting in which the microscopic behaviour of the epidemic (how it spreads randomly in the population) can be approximated by a set of differential equations describing the macroscopic or mean-field dynamics of the system. It also introduces some classical models of epidemic spread.

The chapters of the second part of the book cover more advanced topics. These use some techniques introduced in the first part, in particular large deviations inequalities, but are otherwise largely self-contained. Chapter 6 proposes two models and corresponding analyses of the small-world phenomenon, introducing the notion of navigability of a graph. Chapter 7 focuses on another phenomenon observed in many real-world networks, namely the *power-law* distribution of the degree sequence (or number of neighbours) of the graph. The most salient feature of a sequence with a power-law distribution is that it typically contains samples with very high values. In contrast, classical network models such as the Erdős–Rényi graph have degree sequences highly concentrated around their mean. Chapter 7 describes processes for generating such power laws together with their analysis. Chapter 8 covers recent results on the threshold behaviour of classical models of epidemics on general networks, identifying thresholds with graph properties of the underlying network topology. Finally, Chapter 9 approaches the algorithmic optimisation problem of maximising the spread of an epidemic on a general network by targeting nodes that are likely to yield a large cascade of infections.

**Acknowledgements**

We first would like to thank Frank Kelly for encouraging us to develop our initial lecture notes into this book. We would also like to thank Diana Gillooly of Cambridge University Press for her comments and advice in improving the manuscript.

# PART I

# SHAPELESS NETWORKS

# 1

# Galton–Watson branching processes

## 1.1 Introduction

The branching process model was introduced by Sir Francis Galton in 1873 to represent the genealogical descendance of individuals. More generally it provides a versatile model for the growth of a population of reproducing individuals in the absence of external limiting factors. It is an adequate starting point when studying epidemics since, as we shall see in Chapter 2, it describes accurately the early stages of an epidemic outbreak. In addition, our treatment of so-called *dual* branching processes paves the way for the analysis of the *supercritical* phase in Chapter 2. Finally, the present chapter gives an opportunity to introduce *large deviations* inequalities (and notably the celebrated Chernoff bound), which is instrumental throughout the book.

A Galton–Watson branching process can be represented by a tree in which each node represents an individual, and is linked to its parent as well as its children. The "root" of the tree corresponds to the "ancestor" of the whole population. An example of such a tree is depicted in Figure 1.1.

In the following we consider three distinct ways of exploring the so-called Galton–Watson tree, each well suited to establishing specific properties.

In the *depth-first* view, we start by exploring one child in the first generation, then explore using the same method recursively the subtree of its descendants, before we move to the next child of the first generation. This view is used to establish fixed-point equations satisfied by quantities of interest, such as the extinction probability or the probability distribution of the total population.

The *breadth-first* view stems from the standard breadth-first search of trees and graphs. It consists of exploring first the children of the root, then the children of these children, and so on. In other words, the population is explored generation by generation. This method of exploration allows us to characterise exactly the extinction probability. In particular we will show that the

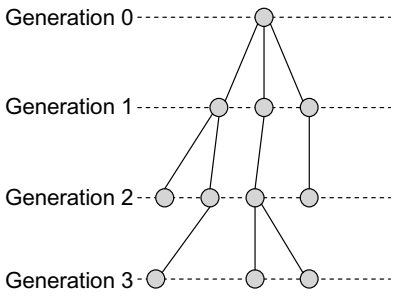


Figure 1.1 Galton–Watson branching process

population can either grow indefinitely (the supercritical case) or go extinct (the subcritical case) depending on whether the mean number of children per individual is above or below 1. This kind of behaviour, namely a qualitative change in global behaviour driven by a tiny change in a system parameter, is known as a phase transition and is widely studied in statistical physics. The emergence of unlimited population growth as the mean number of children goes above 1 is arguably the simplest example of phase transition. Several other examples will be discussed and analysed throughout the book.

Finally, the *one-by-one* view consists of exploring individual nodes' direct children starting from the ancestor. It is used to introduce the notion of dual branching processes. This duality property corresponds to a one-to-one mapping between supercritical and subcritical processes. We will also use one-by-one exploration in our study of Erdős–Rényi graphs in Chapter 2. The one-by-one view yields a description of the total population of branching processes in terms of random walks, which in turn provides simple characterisations of the total population size.

Let us now define formally the object of interest in this chapter. The Galton–Watson branching process is characterised by the probability distribution of the number of children (also called offspring) of each individual. It is a distribution on $\mathbb{N}$, denoted $\{p_k\}_{k\in\mathbb{N}}$. Starting from one individual (ancestor) at generation 0 and denoting by $X_n$ the number of individuals at generation $n$, one then has

$$X_{n+1} = \sum_{i=1}^{X_n} \xi_{n,i}\,, \tag{1.1}$$

where $\xi_{n,i}$ is the number of children of the $i$th individual of the $n$th generation.

By assumption, the $\{\xi_{n,i}\}_{i,n\in\mathbb{N}}$ are independent and identically distributed (i.i.d.), distributed according to $\{p_k\}_{k\in\mathbb{N}}$; that is, for all $k \geq 0$, $\mathbf{P}(\xi = k) = p_k$.

Now draw the tree, as in Figure 1.1, where there is one edge from each individual to each of its children. This is the so-called Galton–Watson tree. In graph-theoretic terms it is an oriented tree spanning the descendants of the ancestor and rooted at the ancestor. We shall denote this tree by $\mathcal{T}$.

## 1.2 Depth-first exploration

Considering depth-first exploration of the tree $\mathcal{T}$, we observe that for any child of the root, the subtree rooted at this child has the same statistical properties as $\mathcal{T}$. Thus conditional on $X_1$ (the number of children of the ancestor), $\mathcal{T}$ is obtained by connecting the root to the individual roots of $X_1$ rooted trees, $\mathcal{T}_1, \ldots, \mathcal{T}_{X_1}$, that are mutually independent, and distributed as $\mathcal{T}$.

Denote by $p_{\text{ext}}$ the extinction probability, i.e. the probability that after some finite $n$, $X_n = 0$. Denote by $|\mathcal{T}|$ the number of nodes of the tree $\mathcal{T}$. We thus have

$$\begin{aligned} p_{\text{ext}} &= \mathbf{P}(|\mathcal{T}| < \infty) \\ &= \sum_{k=0}^{\infty} p_k \, \mathbf{P}(|\mathcal{T}_1| < \infty, \ldots, |\mathcal{T}_k| < \infty) \\ &= \sum_{k=0}^{\infty} p_k p_{\text{ext}}^k \, . \end{aligned}$$

Thus, denoting by

$$\phi_\xi(s) := \sum_{k\in\mathbb{N}} p_k s^k, \quad s \in \mathbb{R}$$

the generating function of the offspring distribution, the extinction probability $p_{\text{ext}}$ is a solution of the equation

$$x = \phi_\xi(x) \, . \tag{1.2}$$

A similar argument yields the following result.

**Theorem 1.1** (Total population) *Consider $\mathcal{X}$, the total population of the branching process, given by $\mathcal{X} = \sum_{k=0}^{\infty} X_k$. Let $\phi_{\mathcal{X}}$ denote its generating function. Then*

$$\phi_{\mathcal{X}}(s) = s\phi_\xi(\phi_{\mathcal{X}}(s)) \, . \tag{1.3}$$

*Proof* Conditional on the event $X_1 = k$, i.e. that there are $k$ individuals in the first generation, let $\mathcal{X}_j$, $j = 1, \ldots, k$ denote the total number of descendents of the $j$th child in the first generation (also counting the $j$th child itself). The variables $(\mathcal{X}_j)_{j=1,\ldots,k}$ are i.i.d. with law equal to that of $\mathcal{X}$. Moreover,

$$\mathcal{X} = 1 + \sum_{j=1}^{k} \mathcal{X}_j \,.$$

Hence

$$\begin{aligned}
\phi_{\mathcal{X}}(s) &:= \mathbf{E}(s^{\mathcal{X}}) \\
&= \sum_{k\geq 0} p_k \, \mathbf{E}(s^{\mathcal{X}} \,\big|\, X_1 = k) \\
&= \sum_{k\geq 0} p_k \, \mathbf{E}(s^{1+\sum_{j=1}^{k} \mathcal{X}_j} \,\big|\, X_1 = k) \\
&= \sum_{k\geq 0} p_k \, s \, \mathbf{E}(s^{\mathcal{X}})^k = s\phi_\xi(\phi_{\mathcal{X}}(s)).
\end{aligned}$$

□

In Section 1.4 we provide a characterisation of the total population $\mathcal{X}$ in terms of the sequence of numbers of children, $\{\xi_{n,i}\}_{i,n\in\mathbb{N}}$, and use it to obtain explicit bounds on the probability that the total population exceeds any particular size.

## 1.3 Breadth-first exploration

The breadth-first view consists of drawing the tree generation after generation. This viewpoint, which is reflected by the defining equation (1.1), allows us to characterise exactly the extinction probability $p_{\text{ext}}$ previously introduced. We will study the existence of solutions of the equation $s = \phi_\xi(s)$ on $[0, 1]$ and characterise them in terms of the offspring distribution $\{p_k\}_{k\geq 0}$.

**Theorem 1.2** (Survival vs. extinction) *The extinction probability $p_{\text{ext}}$ is the smallest solution of equation (1.2) in $[0, 1]$. Denoting by $\mu := \mathbf{E}(\xi)$ the average number of children per individual, one further has the following regimes:*

*(i) Subcritical regime: If $\mu < 1$, then $p_{\text{ext}} = 1$.*
*(ii) Critical regime: If $\mu = 1$ and $p_1 < 1$, then $p_{\text{ext}} = 1$.*
*(iii) Supercritical regime: If $\mu > 1$, then $p_{\text{ext}} < 1$.*

*Proof* Let $p_{\text{ext}}^{(n)} = \mathbf{P}(X_n = 0)$ be the probability that extinction has occurred at or before the $n$th generation. The sequence of events $(\{X_n = 0\})_{n\geq 0}$ is increasing, i.e. $\{X_n = 0\} \subseteq \{X_{n+1} = 0\}$, and converges as $n$ goes to infinity to the extinction event. Thus the probability $p_{\text{ext}}^{(n)} := \mathbf{P}(X_n = 0)$ must converge to $p_{\text{ext}}$ as $n \to \infty$. Let $\phi_n(s) := \mathbf{E}(s^{X_n})$. Then $\mathbf{P}(X_n = 0) = \phi_n(0)$. Using the i.i.d. property of the sequence of offspring of distinct individuals, we have

$$\begin{aligned}\phi_n(s) = \mathbf{E}(s^{X_n}) &= \sum_{k=0}^{\infty} \mathbf{E}(s^{X_n} \mid X_1 = k) p_k \\ &= \sum_{k=0}^{\infty} \left(\mathbf{E}(s^{X_{n-1}})\right)^k p_k \,.\end{aligned}$$

Taking $s = 0$ yields $p_{\text{ext}}^{(n)} = \phi_\xi(p_{\text{ext}}^{(n-1)})$.

Let us prove that $p_{\text{ext}}$ is the smallest solution of equation (1.2) in $[0, 1]$. Let $\psi \in [0, 1]$ be a solution of equation (1.2). First note that $p_{\text{ext}}^{(0)} = 0 \leq \psi$. Then, by induction and using the monotonicity of $\phi_\xi$ on $[0, 1]$ (which holds in view of the expression $\phi'_\xi(s) = \sum_{k\geq 1} k p_k s^{k-1}$), we have that $p_{\text{ext}}^{(n)} \leq \psi$ for $n \geq 0$ and, by taking the limit when $n$ goes to infinity, $p_{\text{ext}} \leq \psi$.

Let us now use this characterisation to establish properties of $p_{\text{ext}}$. Note first that the function $\phi_\xi$ is non-decreasing and convex ($\phi''_\xi(s) = \sum_{k\geq 2} k(k-1)s^{k-2}p_k$ is non-negative on $[0, 1]$), and such that its derivative $\phi'_\xi(1)$ equals the average number of children, $\mu$. It is moreover strictly convex whenever $p_0 + p_1 < 1$.

In the case where $p_0 = 0$, necessarily $p_{\text{ext}} = 0$ is the smallest solution. This can also be seen directly: each individual has at least one child, so the process survives forever.

Assume now that $p_0 > 0$. Since $\{p_k\}_{k\geq 0}$ is a probability distribution, necessarily $p_1 < 1$.

Consider then the following cases.

(i) $\mu < 1$: In this case $\phi'_\xi(1) = \mu < 1$ so that, for small enough $\epsilon > 0$, $\phi_\xi(1 - \epsilon) \sim 1 - \mu\epsilon > 1 - \epsilon$. Thus, by convexity of $\phi_\xi$, the only solution of equation (1.2) is $x = 1$ so that $p_{\text{ext}} = 1$.

(ii) $\mu = 1$: Here the previous assumption that $p_0 > 0$ is equivalent to $p_1 < 1$. In this situation, $p_0 + p_1 < 1$ so that $\phi_\xi$ is strictly convex. It is therefore strictly above its tangent at $x = 1$, whose equation is $y = 1 + \mu(x - 1)$ or equivalently $y = x$. Thus the only solution to equation (1.2) is again $x = 1$.

(iii) $\mu > 1$: In this case, for small enough $\epsilon > 0$, $\phi_\xi(1 - \epsilon) \sim 1 - \mu\epsilon < 1 - \epsilon$. Thus by continuity of the function $x \to \phi_\xi(x) - x$, which takes a positive value $p_0$ at $x = 0$ and a negative value at $x = 1 - \epsilon$, there must exist a