Scale-free networks may prove to be a very useful type of overlay networks for cloud computing.

The degree distribution of scale-free networks follows a power law. We only consider the discrete

case when the probability density function is p(k) = a f (k) with f (k) = k−γ and the constant a is

a = 1/ζ(γ, kmin). Thus,

p(k) = 1

ζ(γ, kmin)

k−γ . (7.4)

In this expression, kmin is the smallest degree of any vertex, and for the applications we discuss in this

chapter kmin = 1; ζ is the Hurvitz zeta function8

ζ(γ, kmin) =

∞

n=0

1

(kmin + n)γ

=

∞

n=0

1

(1 + n)γ . (7.5)

Many physical and social systems are interconnected by a scale-free network. Indeed, empirical data

available for power grids, theWeb, the citation of scientific papers, or social networks confirm this trend:

The power grid of the Western United States has some 5,000 vertices representing power-generating

stations, and in this case γ ≈ 4. For the World Wide Web the probability that m pages point to one

page is p(k) ≈ k−2.1 [40]. Recent studies indicate that γ ≈ 3 for the citation of scientific papers. The

collaborative graph of movie actors in which links are present if two actors were ever cast in the same

movie follows the power law with γ ≈ 2.3. The larger the network, the closer a power law with γ ≈ 3

approximates the distribution [39].

A scale-free network is nonhomogeneous; the majority of the vertices have a low degree, and only

a few vertices are connected to a large number of edges (see Figure 7.16). On the other hand, an

exponential network is homogeneous since most of the vertices have the same degree. The average

distance d between the N vertices, also referred to as the diameter of the scale-free network, scales as

ln N; in fact, it has been shown that when kmin > 2, a lower bound on the diameter of a network with

2 < γ < 3 is ln N [88]. A number of studies have shown that scale-free networks have remarkable

properties such as robustness against random failures [40], favorable scaling [10,11], resilience to

congestion [140], tolerance to attacks [352], small diameter [88], and small average path length [39].

The moments of a power-law distribution play an important role in the behavior of a network. It has

been shown that the giant connected component (GCC) of networks with a finite average vertex degree

and divergent variance can only be destroyed if all vertices are removed; thus, such networks are highly

resilient against faulty constituents [249]. These properties make scale-free networks very attractive for interconnection networks in many applications, including social systems [256], peer-to-peer systems

[8,318], sensor networks [231], and cloud computing.

As an example, consider the case γ = 2.5 and the minimum vertex degree, xmin = 1. We first

determine the value of the zeta function ζ(γ, xmin) and approximate ζ(2.5, 1) = 1.341; thus, the

distribution function is p(k) = k−2.5/1.341 = 0.745   (1/k2.5), where k is the degree of each vertex.

The probability of vertices with degree k > 10 is Prob(k > 10) = 1 − Prob(k   10) = 0.015. This

means that at most 1.5% of the total number of vertices will have more than 10 edges connected to them.

We also see that 92.5% of the vertices have degree 1, 2, or 3. Table 7.2 shows the number of vertices of

degrees 1 to 10 for a very large network, N = 108.

Another important property is that the majority of the vertices of a scale-free network are directly

connected to the vertices with the highest degree. For example, in a network with N = 130 vertices and

m = 215 edges, 60% of the nodes are directly connected to the five vertices with the highest degree,

whereas in a random network fewer than half, 27%, of the nodes have this property [11].

Thus, the nodes of a scale-free network with a degree larger than a given threshold (e.g., k = 4 in

our example) could assume the role of control nodes, and the remaining 92.5% of the nodes could be

servers; this partition is autonomic. Moreover, most of the server nodes are at distance 1, 2, or 3 from

a control node that could gather more accurate state information from these nodes and with minimal

communication overhead.

We conclude that a scale-free network is an ideal interconnect for a cloud. It is not practical to construct

a scale-free physical interconnect for a cloud, but we can generate instead a virtual interconnect with the

desired topology. We pay a small penalty in terms of latency and possibly bandwidth when the nodes

communicate through the virtual interconnect, but this penalty is likely to become smaller and smaller

as new networking technologies for cloud computing emerge.

An Algorithm for the Construction of Graphs with Power-Law Degree Distribution. Consider an

Erd s-R nyi (ER) graph GER with N vertices.Vertex i has a unique label from a compact set i ∈ {1, N}.

We want to rewire this graph and produce a new graph GSF in which the degrees of the vertices follow

a power-law distribution. The procedure we discuss consists of the following steps [212]:

1. We assign to each node i a probability:

pi = i−α

 Nj

=1 j−α

= i−α

ζN (α)

with 0 < α < 1 and ζN (α) =

 N

i=1

i−α. (7.6)

2. We select a pair of vertices i and j and create an edge between them with probability

pi j = pi p j = (i j)

−α

ζ2N

(α)

(7.7)

and repeat this process n times.

Then the probability that a given pair of vertices i and j is not connected by an edge hi j is

pNC

i j

= (1 − pi j )n ≈ e−2npi j (7.8)

and the probability that they are connected is

pC

i j

=

1 − pNC

i j

= 1 − e−2npi j . (7.9)

Call ki the degree of vertex i ; then the moment-generating function of ki is

gi (t) =

j  =i

pNC

i j

+ tpC

i j

. (7.10)

The average degree of vertex i is

  ki = t

d

dt

gi (t)|t=1 =

j  =i

pC

i j . (7.11)

Thus,

  ki =

j  =i

(1 − e−2npi j ) =

j  =i

1 − e

−2n (i j)

−α

ζ2N

(α)

≈

j  =i

2n

(i j)

−α

ζ2N

(α)

= 2n

ζ2N

(α)

j  =i

(i j)

−α. (7.12)

This expression can be transformed as

  ki = 2n

ζ2N

(α)

j  =i

(i j)

−α =

2ni−α

j  =i j−α

ζ2N

(α)

=

2ni−α

ζN (α) − i−α

ζ2N

(α)

. (7.13)

The moment-generating function of ki can be written as

gi (t) =

j  =i

pNC

i j

+ tpC

i j

= e(1−t)   ki =

j  =i

e−(1−t)pC

i j

≈

j  =i

[1 − (1 − t)pC

i j

= e(1−t)

j  =i pC

i j = e(1−t)   ki . (7.14)

We conclude that the probability that ki = k is given by

pd,i (k) = 1

k!

dk

dtk gi (t)|t=0 ≈

  ki

k! e−  ki . (7.15)

When N → ∞then ζN (α) =  N

i=1 i−α converges to the Riemann zeta function ζ(α) for α > 1

and diverges as N1−α

1−α if 0 < α < 1. For 0 < α < 1 Eq. (7.6) becomes

pi = i−α

ζN (α)

= 1 − α

N1−α

i−α. (7.16)

When N →∞, 0 < α < 1, and the average degree of the vertices is 2m, then the degree of vertex i is

k = pi   mN = 2mN

1 − α

N1−α

i−α = 2m(1 − α)

i

N

 −α

. (7.17)

Indeed, the total number of edges in the graph is mN and the graph has a power-law distribution. Then

i = N

k

2m(1 − α)

 − 1

α

. (7.18)

From this expression we see that there is a one-to-many correspondence between the unique label of

the node i in the GER graph and the degree k of the vertices in the GSF graph. This reflects the fact

that multiple vertices may have the same degree k. The number of vertices of degree k is

n(k) = N

k

2m(1 − α)

 − 1

α

− N

k − 1

2m(1 − α)

 − 1

α

= N

k − 1

2m(1 − α)

 − 1

α

1 + 1

k

 − 1

α

− 1

. (7.19)

We denote γ = 1 + 1

α and observe that

1 + 1

k

 − 1

α

= 1 +

−1

α

1

k

 − 1

α

+ 1

2

−1

α

−1

α

− 1

1

k

 − 1

α

−1

+      . (7.20)

We see that

n(k) = N

(k − 1)(γ − 1)

2m(γ − 2)

 −γ+1

(1 − γ )

1

k

 −γ+1

− γ (1 − γ )

2

1

k

 −γ

+

. (7.21)

We conclude that to reach the value predicted by the theoretical model for the number of vertices of

degree k, the number of iterations is a function of N, of the average degree 2m, and of γ , the degree of

the power law. Next we discuss an algorithm for constructing scale-free networks using biased random

walks.

Biased Random Walks. A strategy used successfully to locate systems satisfying a set of conditions

in applications such as peer-to-peer systems is based on biased random walks [33]. Random walks are

reported to bemore efficient in searching for nodes with desirable properties than other methods, such as

flooding [135].

Unfortunately, the application of random walks in a large network with an irregular topology is

unfeasible because a central authority could not maintain accurate information about a dynamic set of

members. A solution is to exploit the fact that sampling with a given probability distribution can be

simulated by a discrete-time Markov chain. Indeed, consider an irreducible Markov chain with states

(i , j ) ∈ {0, 1, . . . , S}, and let P = [pi j ] denote its probability transition matrix, where

pi j = Prob[X(t + 1) = j |X(t) = i ], (7.22)

with X(t) the state at time t. Let π = (π0, π1, . . . , πS) be a probability distribution with nonzero

probability for every state, πi > 0, 0   i   S. The transition matrix P is chosen so that π is its unique

stationary distribution; thus, the reversibility condition π = π P holds.When g(   ) is a function defined

on the states of the Markov chain and we want to estimate

E =

 S

i=0

g(i)πi , (7.23)

we can simulate the Markov chain at times t = 1, 2, . . . , N, and the quantity

 E =

 N

i=1

f (X(t))

N

(7.24)

is a good estimate of E – more precisely,  E  → E when N  → ∞. Hastings [160] generalizes the

sampling method of Metropolis [243] to construct the transition matrix given the distribution π. He

starts by imposing the reversibility condition πi pi j = πj p ji. If Q = [qi j ] is the transition matrix of an

arbitrary Markov chain on the states {0, 1, . . . , S}, it is assumed that

pi j = qi jαi j if i  = j and pii = 1 −

j  =i

pi j . (7.25)

Two versions of sampling are discussed in [160]: that of Metropolis and one proposed by Baker[36].

The quantities αi j are, respectively,

αM

i j

=

1 if πj

πi   1

πj

πi

if πj

πi

< 1

and αB

i j

= πj

πi + πj

. (7.26)

For example, consider a Poisson distribution πi = λi e−λ/i !.We choose qi j = 1/2 if j = i −1, i  = 0

or j = i + 1, i  = 0 and q00 = q01 = 1/2. Then, using Baker’s approach, we have

pi j =

λ/(λ + i + 1) if j = i + 1, i  = 0

i/(i + λ) if j = i − 1, i  = 0

(7.27)

and p00 = 1/2, and p01 = λe−λ/(1 + λe−λ).

The algorithm to construct scale-free overlay topologies with an adjustable exponent presented in

[319] adopts the equilibrium model discussed in [140]. The algorithm is based on random walks in a

connected overlay network G(V, E), viewed as a Markov chain with state space V and a stationary

distribution with a random walk bias configured according to a Metropolis–Hastings chain [160]. In

this case

pi = i−α, with 1   i   N, α∈ [0, 1) (7.28)

and add an edge between two vertices a and b with probability

pa

  N

i=1

pi   pb

  N

i=1

pi (7.29)

if none exists; they repeat the process until mN edges are created and the mean degree is 2m. Then the

degree distribution is

p(k) ∼ k−γ , with γ = (1 + α)/α. (7.30)

The elements of the transition matrix P = [pi j ] are

pi j =

⎧⎪⎪⎪⎨

⎪⎪⎪⎩

1

ki

min

1j

  1

γ−1 ki

k j

, 1

(i , j ) ∈ E

1 − 1

ki

(l,i )∈E cil i = j

0 (i , j) /∈ E

(7.31)

with ki the degree of vertex i . An upper bound for the number of random walk steps can be determined

from a lower bound for the second smallest eigenvalue of the transition matrix, a nontrivial problem.

A distributed rewiring scheme for constructing a scale-free overlay topology with an adjustable

exponent is presented in [319]. An alternative method of creating the scale-free overlay network could

be based on the gossip-based peer sampling discussed in [181]. The distributed algorithm for constructing

a scale-free network in [319] is based on the method of constructing a random graph with a power-law

distribution sketched in [140,212].

Estimation of the Degree of a Power-Law Network. The question we address next is how to estimate

the degree distribution of any scheme for the construction of a power-law network [43]. The estimation

of the degree distribution from empirical data is analyzed in [86]. According to this study, a good

approximation for γ for a discrete power-law distribution for a network with P vertices and kmin = 1 is

γˆ ≈ 1 + P

   P

i=1

ln

ki

kmin − 1/2

 −1

= 1 + P

 P

i=1 2ki

. (7.32)

Several measures exist for the similarity and dissimilarity of two probability density functions of

discrete random variables, including the trace distance, fidelity, mutual information, and relative entropy

[92,198]. The trace distance (also called Kolmogorov or L1 distance) of two probability density functions,

pX (x) and pY (y), and their fidelity are defined as

D(pX (x), pY (x)) = 1

2

x

|pX (x) − pY (x)| (7.33)

and

F(pX (x), pY (x)) =

x

pX (x)pY (x). (7.34)

The trace distance is a metric: It is easy to prove nonnegativity, symmetry, the identity of indiscernibles,

and the triangle inequality. On the other hand, the fidelity is not a metric, since it fails to

satisfy the identity of indiscernibles,

F(pX (x), pX (x)) =

x

pX (x)pX (x) = 1  = 0, (7.35)

respectively. Determining either the L1 distance between the distribution calculated based on Eq. (7.4) or

the one produced by the algorithm discussed earlier requires information about the degree of all vertices.

From Table 7.2 we see that the degree-one vertices represent a very large fraction of the vertices of a

power-law network and may provide a reasonable approximation of the actual degree distribution.