Entropy of Deterministic Networks and Network Ensembles

Justin P. Coon

in collaboration with
A. Cika, C. P. Dettmann, O. Georgiou, D. Simmons, P. J. Smith

Short Course on Complex Networks and Point Processes with Applications

Oriel College
UNIVERSITY OF OXFORD

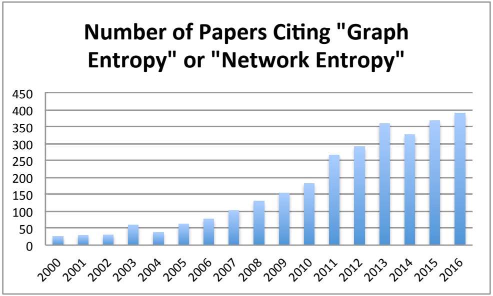
Overview

- History and motivation
- Complexity measures for deterministic networks
- Introduction to entropy
- Entropy measures for deterministic networks
- Entropy of nonspatial network ensembles
- Entropy of spatial network ensembles
- Applications

History and Motivation

A Brief History

The task of measuring graph structure has been a worthwhile objective for a number of years in many disciplines, including **chemistry**, **sociology**, **computer science**, and **ecology**. Efforts to characterise complexity in networks have gathered pace since the dawn of the Internet.



A Brief History

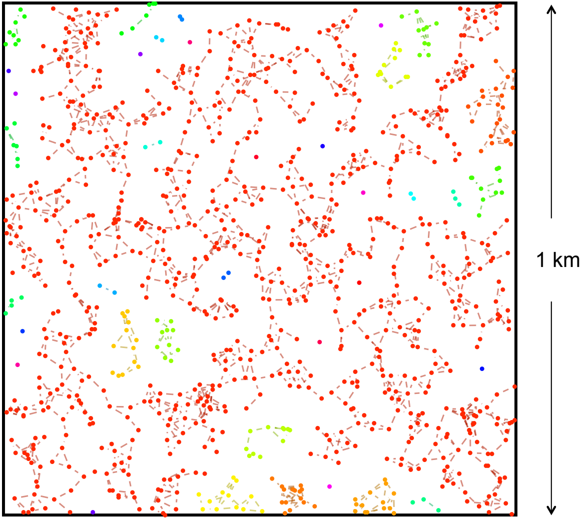
Mathematical studies of graph structure and complexity date back to Rashevsky (1955) and Mowshowitz (1968).

This formalism was taken up in earnest within the **biological** and **chemical** sciences to classify **chemical structures** by Bonchev et al (1977, 1983).

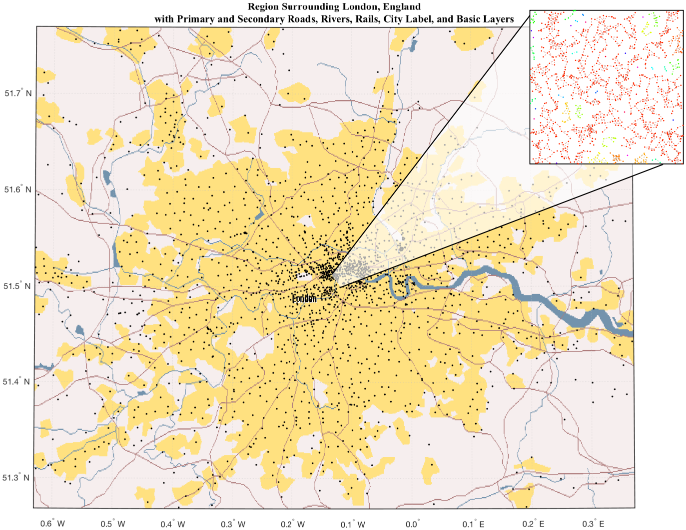
At the turn of the millennium, the **network physics** community began to explore network entropy. Notable work was published by Park and Newman (2004), Anand and Bianconi (2009).

Around the same time, graph entropy was employed to study **social networks** (Butts, 2001) and **ecological networks** (Solé et al, 2001; Ulanowicz, 2004).

Why study network entropy?

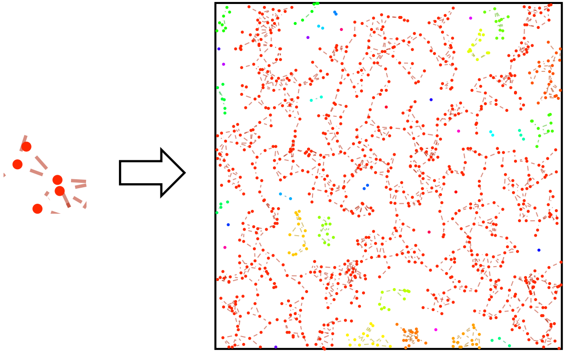


Why study network entropy?



Why study network entropy?

Local behaviour $\mapsto$ Global complexity



Complexity Measures for Deterministic Networks

Measuring Complexity through Structure

Substructure Count

Idea: Complex networks contain a rich substructure.

Example: Complexity is quantified as the subgraph count

$$\text{complexity}(G) := \sum_{k=0}^{|\mathcal{E}|} \text{subgraphCount}_k(G)$$

Notation and terminology: $G = (\mathcal{V}, \mathcal{E})$ and we use "graph" and "network" interchangeably.

Measuring Complexity through Generation

Generative

Idea: A large number of operations must be performed on a set of protographs to construct a complex network.

Example: Given a set of protographs isomorphic to stars, complexity is quantified as the number of unions and intersections of these elemental graph structures required to generate the edge set of a given network.

Measuring Complexity through Encoding

Encoding

Idea: A large number of yes/no questions (bits) are required to describe a complex network.

Example: A four-node network would require *at most* six bits to describe the topology. A 400 node network would require *at most* 79,800 bits to describe its topology. (Different encoding schemes lead to different results.)

Do better encoding schemes exist?

A Journey into Theoretical Computer Science

Encoding measures are related to *Kolmolgorov complexity* and *minimum description length*.

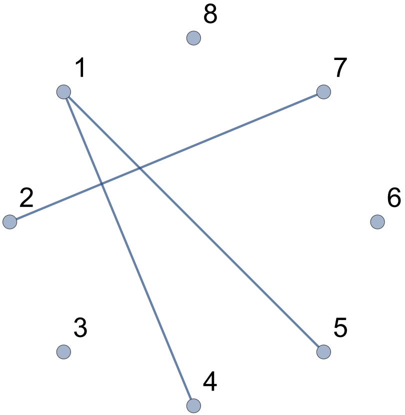
Definition (Kolmolgorov Complexity)

The Kolmolgorov complexity of an object is defined as the smallest possible description of that object using a fixed, universal description language.

Definition (Minimum Description Length)

The principle that the best encoding of a dataset is the one that compresses it the most.

Encoding Example



This graph has 28 possible edges, and hence we can encode the graph using a 28-bit string. Alternatively, we could list the edges as vertex pairs (in binary) and terminate on one end with the sequence 111111 as follows

111111 001100 001101 010111

Touching on Communication/Information Theory

Theorem (Compression via Source Coding)

Consider a discrete memoryless source with symbols denoted by the random variable X. Suppose groups of J symbols are encoded into sequences of N bits. Let P_e be the probability that a block of symbols is decoded in error. Then P_e can be made arbitrarily small if

$$R = \frac{N}{J} \geq \text{entropy}(X) + \epsilon$$

for some $\epsilon > 0$ and J sufficiently large. Conversely, if

$$R \leq \text{entropy}(X) - \epsilon$$

then P_e becomes arbitrarily close to 1 as J grows large.

Making the Link: From Complexity to Entropy

$$\text{Complexity} \mapsto \text{MDL} \mapsto \text{Entropy}$$

Entropy has a rich mathematical history in physics and information theory. Hence, this formalism provides us with a much more complete set of tools for analysing network structure and complexity.

Introduction to Entropy

Shannon Entropy

Shannon Entropy

Shannon entropy is defined with respect to a probability distribution $\{p_0, p_1, \ldots\}$ as

$$H(p_0, p_1, \ldots) = -c \sum_i p_i \log p_i$$

where c is a constant that determines the base of the log.

When $\{p_0, p_1, \ldots\}$ describes the distribution of a discrete random variable X, we often write

$$H(X) = \mathbb{E}[-\log P(X)]$$

If X is a continuous random variable with density $p(x)$, the *differential entropy* is defined as

$$H(X) = \mathbb{E}[-\log p(X)] = -\int_{\text{supp(X)}} p(x) \log p(x) \, \mathrm{d}x$$

Properties of Shannon Entropy

- **Positive and Finite**

$$0 \leq H(X) \leq \log |\mathrm{supp}(X)|$$

- **Concave**

$$H(\lambda p_0 + (1-\lambda)p_1) \geq \lambda H(p_0) + (1-\lambda)H(p_1), \quad \lambda \in [0,1]$$

- **Joint Entropy**

$$H(X,Y) = \mathbb{E}[-\log P(X,Y)]$$

- **Conditional Entropy**

$$H(X|Y) = \mathbb{E}[-\log P(X|Y)]$$

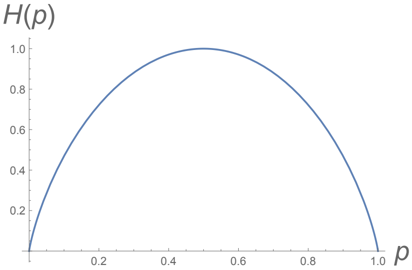
- **Independence**

$$H(X|Y) = H(X) \Rightarrow H(X,Y) = H(X) + H(Y|X) = H(X) + H(Y)$$

Example: Bernoulli Experiment

Consider a coin toss where the probability of a heads is p and the probability of a tails is $1 - p$. The entropy is

$$H(p) = -p \log p - (1 - p) \log(1 - p)$$



Entropy Variants

Other definitions of entropy exist, which are suitable for use in the context of ensembles, but comparatively little work related to the entropy of graph ensembles using these measures has been reported in the literature.

Rényi Entropy

Rényi entropy of order α is defined for the probability distribution $\{p_0, p_1, \ldots\}$ as

$$H_\alpha(p_0, p_1, \ldots) := \frac{1}{1-\alpha} \log \left(\sum_i p_i^\alpha \right).$$

It generalises Shannon entropy ($H = \lim_{\alpha \to 1} H_\alpha$), max-entropy ($H_0$), collision entropy ($H_2$), and min-entropy ($\lim_{\alpha \to \infty} H_\alpha$). It is Schur concave.

Entropy Variants

Von Neumann Entropy

The von Neumann entropy of a quantum mechanical system described by the density ρ is defined as

$$H(\rho) := -\mathrm{tr}(\rho \ln \rho)$$

For the eigendecomposition $\rho = \sum_i \omega_i |i\rangle \langle i|$, the von Neumann entropy takes the Shannon form

$$H(\rho) = -\sum_i \omega_i \ln \omega_i$$

Entropy Measures for Deterministic Networks

General Approach

The key to using entropy as a measure of structural information or complexity in the context of networks is to **define a probability distribution on an appropriate set of graph invariants** associated with the network.

Take a **graph invariant** $\mathcal{Z}$ (i.e., a property of the graph that is invariant under isomorphisms) and **define an equivalence relation** that induces a set of **equivalence classes** $\{\mathcal{Z}_i\}$. In general, we can use the Shannon entropy formalism to define the entropy of the graph with respect to that equivalence relation as

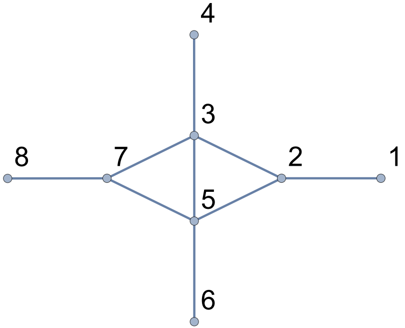
$$H(G) := -\sum_i \frac{|\mathcal{Z}_i|}{|\mathcal{Z}|} \log \frac{|\mathcal{Z}_i|}{|\mathcal{Z}|}$$

Topological Information Content

Consider the **automorphism group** of a graph, i.e., the set of graphs derived from vertex permutations whereby edges in the original graph are contained in the set of edges in the permuted graph. The equivalence classes of the graph under automorphisms are called **vertex orbits**. Let $\mathcal{V}_i$ denote the ith orbit, and K is the number of different orbits. Rashevsky (followed by Mowshowitz) formalised the notation of **topological information content** as

$$H_a(G) := -\sum_{i=1}^{K} \frac{|\mathcal{V}_i|}{|\mathcal{V}|} \log \frac{|\mathcal{V}_i|}{|\mathcal{V}|}$$

Automorphism Group Example



Automorphisms

(1)(2)(3)(4)(5)(6)(7)(8)
(1 8)(2 7)(3)(4)(5)(6)
(1)(2)(3 5)(4 6)(7)(8)
(1 8)(2 7)(3 5)(4 6)

Vertex orbits

(1 8)(2 7)(3 5)(4 6)

Chromatic Information Content

The **chromatic number** of a graph, denoted by $\chi(G)$, is the smallest number of colours required to colour the vertices such that no two adjacent vertices share the same colour. Let $\hat{\mathcal{V}} = \{\mathcal{V}_i | 1 \leq i \leq \chi(G)\}$ denote the a chromatic decomposition of a graph. Then $\hat{\mathcal{V}}$ forms a set of equivalence classes, and we can define the **chromatic information content** of a graph as (Mowshowitz, 1968)

$$H_c(G) := \min_{\hat{\mathcal{V}}} \left\{ - \sum_{i=1}^{\chi(G)} \frac{|\mathcal{V}_i|}{|\mathcal{V}|} \log \frac{|\mathcal{V}_i|}{|\mathcal{V}|} \right\}$$

Radial Centric Information Content

The **eccentricity** σ_v of a vertex v is the maximum distance (path length) from that vertex to any other vertex in the network. Let $\mathcal{V}_\sigma$ denote the set of vertices with eccentricity σ. Denote the diameter of the graph as D. Then $\{\mathcal{V}_\sigma\}$ form a set of equivalence classes, and we can define the **radial centric information content** of a graph as (Bonchev, 1983)

$$H_r(G) := -\sum_{\sigma=1}^{D} \frac{|\mathcal{V}_\sigma|}{|\mathcal{V}|} \log \frac{|\mathcal{V}_\sigma|}{|\mathcal{V}|}$$

Vertex Degree Information Content

Let $\mathcal{V}_\delta$ denote the set of vertices with degree δ. Denote the maximum degree of the vertices by $\bar{\delta}$. Then $\{\mathcal{V}_\delta\}$ form a set of equivalence classes, and we can define the **vertex degree information content** of a graph as (Bonchev, 1983)

$$H_v(G) := -\sum_{\delta=0}^{\bar{\delta}} \frac{|\mathcal{V}_\delta|}{|\mathcal{V}|} \log \frac{|\mathcal{V}_\delta|}{|\mathcal{V}|}$$

Parametric Graph Entropy

The measures discussed previously are **intrinsic** to the network in question. **Extrinsic** measures can also be defined by placing a value of some description to features of the graph. The most general and accessible approach that has been proposed is to use an **information function** $f : \mathcal{V} \mapsto \mathbb{R}_+$ that acts on the individual vertices. Maintaining the need for a probabilistic interpretation yields (Dehmer, 2011)

$$H_p(G) := -\sum_{i=1}^{|\mathcal{V}|} \frac{f(v_i)}{\sum_{j=1}^{|\mathcal{V}|} f(v_j)} \log \frac{f(v_i)}{\sum_{j=1}^{|\mathcal{V}|} f(v_j)}$$

Different Interpretations for Different Measures

Example (Vertex Degree Information Content vs. Parametric Degree Entropy)

Let $f(v_i) := \delta(v_i)$. For the **cycle graph** C_N with N vertices, the **parametric degree entropy is** $\log N$. Yet, there is a single nonempty degree equivalence class corresponding to $\delta = 2$, and thus the **vertex degree information content is zero**.

The **vertex degree information content** is the **entropy of the degree distribution**. It is upper bounded by $\log(1 + \bar{\delta})$.

The **parametric degree entropy**, which can be written as

$$H_{p,\delta}(G) = -\sum_{i=1}^{|\mathcal{V}|} \frac{\delta(v_i)}{2|\mathcal{E}|} \log \frac{\delta(v_i)}{2|\mathcal{E}|}$$

quantifies the **uniformity of the vertex degrees** across the set $\mathcal{V}$.

Entropy of Nonspatial Network Ensembles

Refining the Definition for Ensembles

When working with *ensembles* of networks, a probability distribution P can often be naturally defined on the set of graphs. The graph G then becomes a random variable, and the entropy of the ensemble $\mathcal{G}$ is defined as

$$H(\mathcal{G}) := \mathbb{E}[-\log P(G)]$$

in its most general form.

Erdős-Rényi (ER) Ensembles

Introduced in 1959, this model concerns an ensemble of graphs $\mathcal{G}_{N,E}$, where each graph is formed of N vertices and E edges. The random graph (i.e., random variable) $G_{N,E}$ is drawn uniformly from this set.

There are $\binom{N(N-1)/2}{E}$ graphs in the ensemble. Hence, the entropy is

$$H(\mathcal{G}_{N,E}) = \log \binom{N(N-1)/2}{E}$$

Gilbert Ensembles

Also introduced in 1959, this model concerns an ensemble of graphs $\mathcal{G}_{N,p}$, where each graph is formed of N vertices and where each edge exists with probability p independent of all other edges. This is often referred to as the ER ensemble since Erdős and Rényi considered the model as well.

The entropy of the random graph $G_{N,p}$ is equivalent to the joint entropy of the edges, i.e.,

$$H(\mathcal{G}_{N,p}) = H(X_{1,2}, X_{1,3}, \ldots, X_{N-1,N})$$

where $P(X_{i,j} = 1) = p$. Independence implies

$$H(\mathcal{G}_{N,p}) = \sum_{i<j} H(X_{i,j}) = \frac{N(N-1)}{2} H(p)$$

where $H(p) = -p \log p - (1-p) \log(1-p)$.

Exponential Random Graphs (ERG) Ensembles

The ERG model, originally studied in the context of social network analysis (Holland and Leinhardt, 1981; Frank and Strauss, 1986), assumes a random graph G has a distribution

$$P(G) = \frac{\exp(\sum_i \theta_i z_i(G))}{\kappa(\{\theta_i\})}$$

where $\{\theta_i\}$ are model parameters and $\{z_i(G)\}$ are statistical observables. The parameter $\kappa(\{\theta_i\})$ is a normalisation constant that ensures $\sum_{G \in \mathcal{G}} P(G) = 1$.

From an equilibrium statistical physics perspective, this distribution can be derived more constructively...

Exponential Random Graphs (ERG) Ensembles

To derive the ERG distribution using the principles of equilibrium statistical mechanics, we seek the function P that maximises the Gibbs entropy functional

$$S[P] = -\sum_{G \in \mathcal{G}} P(G) \ln P(G)$$

subject to the constraints

$$\langle z_i \rangle = \sum_{G \in \mathcal{G}} P(G) z_i(G) \qquad \text{and} \qquad 1 = \sum_{G \in \mathcal{G}} P(G).$$

Forming the Lagrangian with multipliers $\{\lambda_i\}$, yields the required condition

$$\frac{\partial}{\partial P(G)} \left\{ S[P] + \lambda_0 \left(1 - \sum_{X \in \mathcal{G}} P(X) \right) + \sum_i \lambda_i \left(\langle z_i \rangle - \sum_{X \in \mathcal{G}} P(X) z_i(X) \right) \right\} = 0$$

Exponential Random Graphs (ERG) Ensembles

Functional differentiation results in

$$\ln P(G) + 1 + \lambda_0 + \sum_i \lambda_i z_i(G) = 0$$

which leads to

$$P(G) = \frac{\exp\left(-\sum_i \lambda_i z_i(G)\right)}{Z(\{\lambda_i\})}$$

with $Z(\{\lambda_i\}) = \mathrm{e}^{1+\lambda_0}$.

To recover the relationship to the ERG model and equilibrium statistical mechanics, note that $\lambda_i = -\theta_i$, $Z = \kappa$ is the *partition function*, and $\sum_i \lambda_i z_i(G)$ is the *Hamiltonian*. The entropy of the ERG ensemble is

$$H(\mathcal{G}) = \left(1 + \lambda_0 + \sum_i \lambda_i \langle z_i \rangle\right) \log \mathrm{e}$$

Entropy of Spatial Network Ensembles

Basic Definitions

Definition (Point Process)

A **point process** is a mathematical model for a set of random distributed points in some space.

Definition (Binomial Point Process (BPP))

Let λ denote a spatial density defined on a set $\mathcal{B} \subseteq \mathcal{S}$. A point process with N points independently and identically distributed in $\mathcal{B}$ according to λ is called a **binomial point process**.

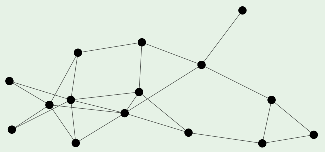
Definition (Pair Distance Distribution)

The **pair distance distribution** gives a statistical description of the distance between two arbitrary points in some space.

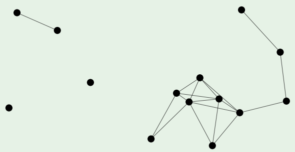
Spatial Network Model

- Embed N nodes randomly (BPP) in $\mathcal{K}_d \subset \mathbb{R}^d$.
- Connect nodes i and j, which are separated by a distance $r_{i,j}$, with probability $p(r_{i,j})$.
- Doing this many times yields the a graph ensemble $\mathcal{G}$.

Example (Erdős-Rényi)



Example (Spatial)



Upper Bound on Entropy

Theorem

Given a point process situated in $\mathcal{K}_d \subset \mathbb{R}^d$ admitting the pair distance density $f(r)$, and assuming a pair connection function $p(r)$, the entropy of the resulting graph ensemble $\mathcal{G}$ satisfies

$$H(\mathcal{G}) \leq \frac{N(N-1)}{2} H(\overline{p})$$

where

$$\overline{p} = \int_0^D p(r) f(r) \, \mathrm{d}r$$

and D is the diameter of $\mathcal{K}_d$.

Proof.

The proof uses the correspondence between a graph and its edge set. The bound follows from an argument based on the concavity of entropy. □

Lower Bound on Entropy

Theorem

Given a point process situated in $\mathcal{K}_d \subset \mathbb{R}^d$ admitting the pair distance density $f(r)$, and assuming a pair connection function $p(r)$, the entropy of the resulting graph ensemble $\mathcal{G}$ satisfies

$$H(\mathcal{G}) \geq \frac{N(N-1)}{2} \int_0^D H(p(r)) f(r) \, \mathrm{d}r =: H(\mathcal{G}|\mathcal{P})$$

The lower bound $H(\mathcal{G}|\mathcal{P})$ is called the **conditional entropy** *of the ensemble $\mathcal{G}$ given the distribution of the point process $\mathcal{P}$.*

Proof.

The proof follows from the classical definition of conditional entropy (in the Shannon sense) and the fact that conditioning cannot increase uncertainty. □

Consequences of the Bounds

The upper and lower bounds yield the following immediate results:

Property 1 The entropy of an ensemble arising from a soft connection function scales like N^2.

Property 2 Classical random geometric graphs (with a hard connection function) have zero conditional entropy.

Property 3 The reduction in uncertainty of the ensemble (topology) given the statistical properties of the point process quantifies the **mutual information** between $\mathcal{G}$ and $\mathcal{P}$

$$I(\mathcal{P};\mathcal{G}) := H(\mathcal{G}) - H(\mathcal{G}|\mathcal{P})$$

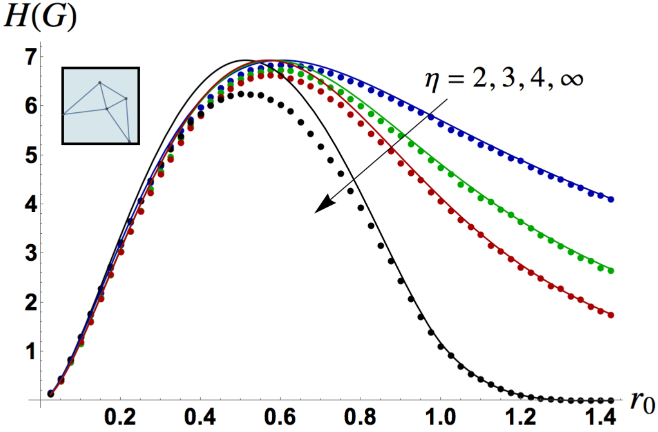
Controlling Complexity as Networks Scale

If a system has a *typical connection range* r_0, it is possible to derive scaling laws that provide insight about how to *control* the complexity of a network ensemble as the number of vertices grows large.

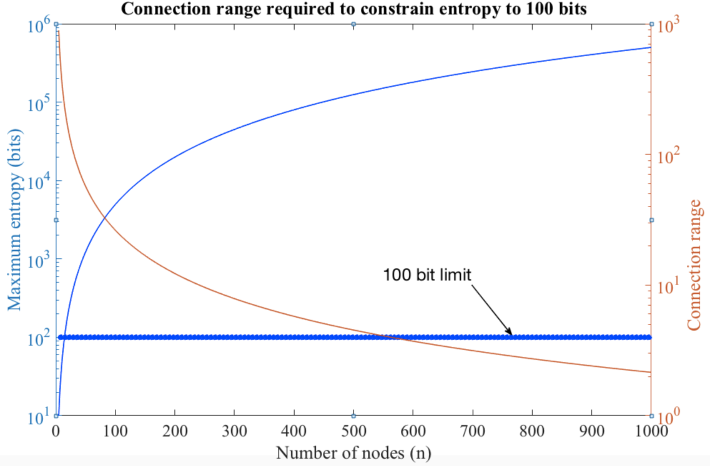
Test function:

$$p(r) = \exp\left(-(r/r_0)^{\eta}\right)$$

Controlling Complexity as Networks Scale



Controlling Complexity as Networks Scale



Statistical Characterisation of Sets

Definition (Set Covariance)

In d dimensions, the **set covariance** of a convex set $\mathcal{K}_d$ is given by

$$c_{\mathcal{K}_d}(\mathbf{r}) = \int_{\mathbb{R}^d} \mathbf{1}_{\mathcal{K}_d}(\mathbf{x}) \mathbf{1}_{\mathcal{K}_d}(\mathbf{x} - \mathbf{r}) \, \mathrm{d}\mathbf{x}, \qquad \mathbf{r} \in \mathbb{R}^d$$

where $\mathbf{1}_{\mathcal{K}_d}(\mathbf{x})$ is the indicator function for $\mathbf{x} \in \mathcal{K}_d$.

Definition (Isotropised Set Covariance)

The **isotropised set covariance** is given by

$$\overline{c}_{\mathcal{K}_d}(r) = \int_{\mathcal{S}_{d-1}} c_{\mathcal{K}_d}(r\mathbf{u}) U(\,\mathrm{d}\mathbf{u}), \qquad r \geq 0$$

where $\mathbf{u}$ is a vector denoting a point on the unit sphere $\mathcal{S}^{d-1}$.

Another Look at Pair Distance

Theorem (Pair Distance Density Function)

The **pair distance probability density function** *is given by*

$$f(r) = \frac{2\pi^{d/2} r^{d-1} \overline{c}_{\mathcal{K}_d}(r)}{\Gamma(d/2)\,\mathrm{vol}(\mathcal{K}_d)^2}.$$

Theorem (Small Argument)

For small distances, the pair distance density can be approximated to first order as

$$f(r) \simeq \mathrm{vol}(\mathcal{K}_d) - r \lim_{r\to 0} \int_{\mathcal{S}_{d-1}} \mathrm{vol}((\mathcal{K}_d \cap (\mathcal{K}_d + r\mathbf{u}))_{\mathbf{u}^\perp}) U(\mathrm{d}\mathbf{u})$$

Scaling Law in 2D

Theorem

As $N \to \infty$, the entropy of a graph in $\mathcal{K}_2$ can be bounded away from zero only if

$$r_0^2 \log\left(\frac{1}{r_0}\right) = \Omega\left(\frac{1}{N^2}\right).$$

The entropy bound will tend to a limit $\ell_h > 0$ as $N \to \infty$ if

$$\begin{aligned} r_0(N) &= \exp\left(\frac{1}{2} W_m\left(-\frac{2\ell_h}{u_2 N(N-1)}\right)\right) \\ &= \left(\frac{\ell_h}{u_2 N^2 \log N}\right)^{\frac{1}{2}} \left(1 + O\left(\frac{1}{\log N}\right)\right) \end{aligned}$$

where $W_m(x)$ is the lower branch ($-1/e \leq x < 0$ and $W_m \leq -1$) of the solution to $x = W \exp W$.

Scaling Law in 3D

Theorem

As $N \to \infty$, the entropy of a graph in $\mathcal{K}_3$ can be bounded away from zero only if

$$r_0^3 \log\left(\frac{1}{r_0}\right) = \Omega\left(\frac{1}{N^2}\right).$$

The entropy bound will tend to a limit $\ell_h > 0$ as $N \to \infty$ if

$$\begin{aligned} r_0(N) &= \exp\left(\frac{1}{3} W_m\left(-\frac{2\ell_h}{u_3 N(N-1)}\right)\right) \\ &= \left(\frac{\ell_h}{u_3 N^2 \log N}\right)^{\frac{1}{3}} \left(1 + O\left(\frac{1}{\log N}\right)\right) \end{aligned}$$

where $u_3 = 4\pi\Gamma(3/\eta)/(\eta V(\mathcal{K}_3))$ is the fractional volume of a soft unit ball.

Maximum Entropy Pair Connection Function (Conditional Entropy)

As with nonspatial networks, it is natural to study the maximum entropy properties of the pair connection function for spatial ensembles. One must first define a set of meaningful constraints

$$\int_0^D \theta_\ell(r)p(r)\,\mathrm{d}r, \qquad \ell = 1, \ldots, L$$

Solving the associated constrained variational problems yields the entropy maximising function

$$p(r) = \frac{1}{\mathrm{e}^{\psi(\mathrm{r})} + 1}$$

where

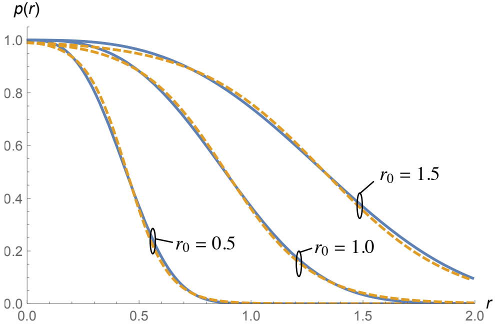
$$\psi(r) = \frac{2}{N(N-1)f(r)} \sum_{\ell=1}^{L} \lambda_\ell \theta_\ell(r)$$

Applications

Applications: Measuring Complexity in Practice

Example (Wireless Networks)

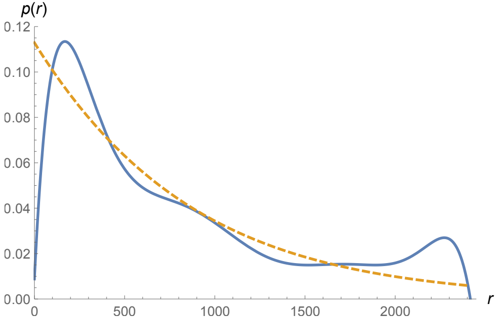
The pair connection function in wireless networks is typically well defined. Hence, we can quantify how complex engineered networks are.



Applications: Measuring Complexity in Practice

Example (Airline Networks)

Flight path data is available for primary airports. Here, we show that the (conditional) entropy of the primary conterminous US airline network is nearly maximal, despite the difference in pair connection functions.



Applications: Communication Network Protocols

Example (Routing Protocol Design)

The parametric graph entropy model has been adopted to design routing protocols in communication networks. The idea is to design a time-dependent stability metric for each link (S_l^t). A probability distribution for each link in the routing table of the ith node can then be defined

$$p_l = \frac{S_l^t}{\sum_{l \in \mathcal{R}_i^t} S_l^t}$$

and the entropy at time t and the ith node calculated to be

$$H_i(t) = -\frac{1}{\log |\mathcal{R}_i^t|} \sum_l p_l \log p_l.$$

Routing will be executed such that stability (measured through node entropy) is maximised.

Applications: Graph Compression

Example (Compressing Graphical Structures)

Choi and Szpankowski (2012) studied the problem of structural entropy (defined with respect to isomorphic graphs) in Gilbert graph ensembles and used this to derive structural compression algorithms. Applications include the compression of biological or medical datasets and topographical maps. It was shown that for $p \gg \ln N/N$ and $1-p \gg \ln N/N$, the structural entropy of the ensemble is

$$H_S(\mathcal{G}_{N,p}) \sim \frac{N(N-1)}{2} H(p) - \log N!.$$

The compression algorithm is iterative: the number of neighbours of a vertex are stored and the remaining vertices are partitioned into disjoint sets depending on their relationship to this vertex; a neighbour of the original vertex is then selected and the process repeats, where partitioning is executed by taking into account all parent vertices.

References

- An, B., & Papavassiliou, S. (2002). An entropy-based model for supporting and evaluating route stability in mobile ad hoc wireless networks. IEEE Communications Letters, 6(8), 328-330.
- Anand, K., & Bianconi, G. (2009). Entropy measures for networks: Toward an information theory of complex topologies. Physical Review E, 80(4), 045102.
- Anand, K., Bianconi, G., & Severini, S. (2011). Shannon and von Neumann entropy of random networks with heterogeneous expected degree. Physical Review E, 83(3), 036109.
- Boushaba, M., Hafid, A., & Gendreau, M. (2017). Node stability-based routing in wireless mesh networks. Journal of Network and Computer Applications, 93, 1-12.
- Bonchev, D. (1983). Information theoretic indices for characterization of chemical structures (No. 5). Research Studies Press.
- Choi, Y., & Szpankowski, W. (2012). Compression of graphical structures: Fundamental limits, algorithms, and experiments. IEEE Transactions on Information Theory, 58(2), 620-638.
- Cika, A., Coon, J. P., & Kim, S. (2017). Effects of directivity on wireless network complexity. In Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2017 15th International Symposium on (pp. 1-7). IEEE.
- Coon, J. P. (2016, December). Topological Uncertainty in Wireless Networks. In Global Communications Conference (GLOBECOM), 2016 IEEE (pp. 1-6). IEEE.
- Coon, J., & Smith, P. J. (2017). Topological entropy in wireless networks subject to composite fading.
- Coon, J. P., Dettmann, C. P., & Georgiou, O. (2017). Entropy of Spatial Network Ensembles. arXiv preprint arXiv:1707.01901.
- de Beaudrap, N., Giovannetti, V., Severini, S., & Wilson, R. (2016). Interpreting the von Neumann entropy of graph Laplacians, and coentropic graphs. A Panorama of Mathematics: Pure and Applied, 658, 227.

References

- Dehmer, M. (2008). Information processing in complex networks: Graph entropy and information functionals. Applied Mathematics and Computation, 201(1), 82-94.
- Dehmer, M., & Mowshowitz, A. (2011). A history of graph entropy measures. Information Sciences, 181(1), 57-78.
- Everett, M. G. (1985). Role similarity and complexity in social networks. Social Networks, 7(4), 353-359.
- Frank, O., & Strauss, D. (1986). Markov graphs. Journal of the American Statistical Association, 81(395), 832-842.
- Guo, J. L., Wu, W., & Xu, S. B. (2011). Study on Route Stability Based on the Metrics of Local Topology Transformation Entropy in Mobile Ad Hoc Networks. In Advanced Engineering Forum (Vol. 1, pp. 288-292). Trans Tech Publications.
- Halu, A., Mukherjee, S., & Bianconi, G. (2014). Emergence of overlap in ensembles of spatial multiplexes and statistical mechanics of spatial interacting network ensembles. Physical Review E, 89(1), 012806.
- Holland, P. W., & Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. Journal of the American Statistical Association, 76(373), 33-50.
- Lu, J. L., Valois, F., Dohler, M., & Barthel, D. (2008). Quantifying organization by means of entropy. IEEE communications letters, 12(3).
- Mowshowitz, A., & Dehmer, M. (2012). Entropy and the complexity of graphs revisited. Entropy, 14(3), 559-570.
- Mowshowitz, A. (1968). Entropy and the complexity of graphs: I. an index of the relative complexity of a graph. The bulletin of mathematical biophysics, 30(1), 175-204.
- Mowshowitz, A. (1968). Entropy and the complexity of graphs: II. The information content of digraphs and infinite graphs. Bulletin of Mathematical Biology, 30(2), 225-240.

References

- Mowshowitz, A. (1968). Entropy and the complexity of graphs: III. Graphs with prescribed information content. Bulletin of Mathematical Biology, 30(3), 387-414.
- Mowshowitz, A. (1968). Entropy and the complexity of graphs: IV. Entropy measures and graphical structure. Bulletin of Mathematical Biology, 30(4), 533-546.
- Park, J., & Newman, M. E. (2004). Statistical mechanics of networks. Physical Review E, 70(6), 066117.
- Simmons, D. E., Coon, J. P., & Datta, A. (2017). Symmetric Laplacians, Quantum Density Matrices and their von Neumann Entropy. arXiv preprint arXiv:1703.01142.
- Simonyi, G. (1995). Graph entropy: A survey. Combinatorial Optimization, 20, 399-441.
- Timo, R., Blackmore, K., & Hanlen, L. (2005). On entropy measures for dynamic network topologies: Limits to MANET. In Communications Theory Workshop, 2005. Proceedings. 6th Australian (pp. 95-101). IEEE.
- Zayani, M. H. (2012). Link prediction in dynamic and human-centered mobile wireless networks (Doctoral dissertation, Institut National des Tlcommunications).