Entropy in Complex Networks (An Application of Shannon Entropy)

Marko S. Suchy

4/9/23

Shannon's "Information" Entropy

For an Ensemble $X(R, p_i)$

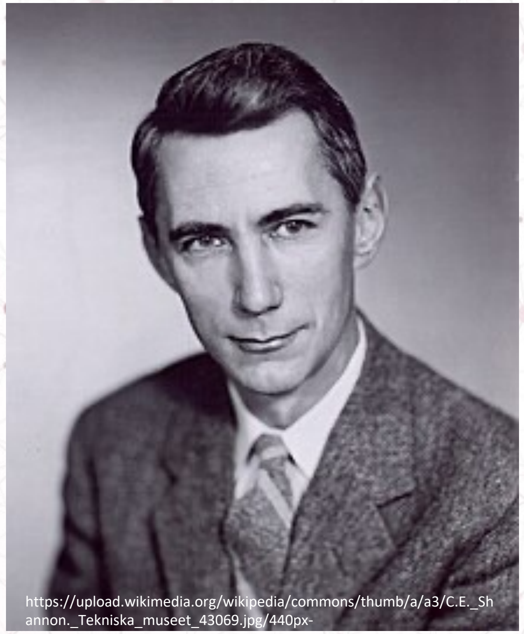
- Where R is the set of possible outcomes
- And p_i is the probability of a particular microstate
- X is a discrete random variable

- Originally came from information theory, 1948
- Underpins the limits of data compression
- Quantifies the Uncertainty in a set of outcome
- Very similar to Gibbs expression of entropy:

$$S = -k_b \sum_i p_i \ln p_i$$

The Shannon entropy of X…

$$H(X) = -\sum_{i=1}^{n} p_i \log_2 p_i$$



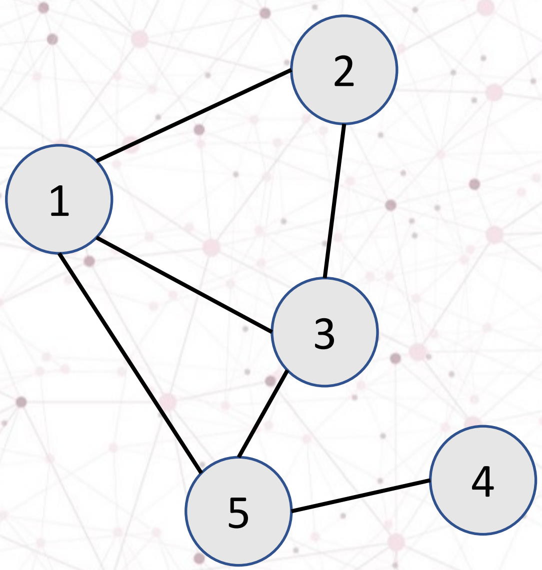
https://upload.wikimedia.org/wikipedia/commons/thumb/a/a3/C.E._Shannon._Tekniska_museet_43069.jpg/440px-

Preliminary Network Information

A Graph G with Adjacency Matrix A

With…

- A finite set of nodes, V
- $V = \{v_0, v_1, \dots, v_N\}$
- A set of edges, E



$$A_{i,j} = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

‘a pattern of interconnections among a set of things’ [1]

- Social network analysis
- Communications routing
- Protein interaction
- Transportation networks

…

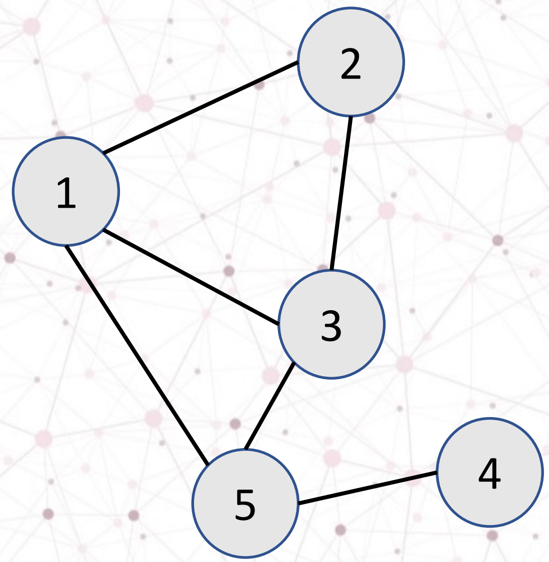
Graphs can also be…

- Weighted
- Directed

Shannon's Entropy in the Context of Networks

$$H(X) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

p_i, the probability distribution, can be based on many attributes:



- Degree (number of edges attached)
- Betweenness (the fraction of times in which a node v falls on the geodesic path between any two other nodes i and j.)
- Paths (number paths a node falls on)
- More!

Shannon's Entropy in the Context of Networks

$$H(X) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

Two examples of Shannon's entropy formulations, using degree:

Node Level Entropy:

$$H(i) = -\sum_{j=1}^{N} \frac{a_{ij}}{k_i} \log(\frac{a_{ij}}{k_i})$$ [2]

Where:

i = a node
$a_{ij} \in \{1, 0\}$ (element in A)
k_i = degree of node i

Network Level Entropy:

$$H(G) = -\sum_{i=1}^{N} \frac{k_i}{2N} \log_2(\frac{k_i}{2N})$$ [3]

Where:

G = a graph
N = number of nodes in system
k_i = degree of node i

A measure of uncertainty!

Shannon's Entropy in the Context of Networks

$$H(X) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

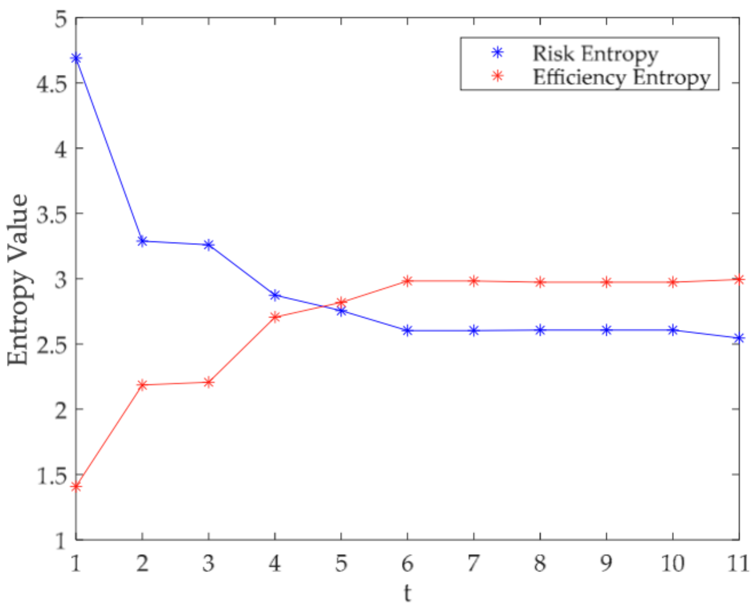
Wang, Zeng, and Tu… [4]

$$\mathrm{H}_R = \sum_{j=1}^{s} \frac{n_j}{n} \mathrm{H}_{\mathrm{RM}_j}$$

where $\mathrm{H}_{\mathrm{RM}_j} = -\sum_{i=1}^{n_j} p_i \ln p_i$; and $p_i = \frac{k_i^w}{\sum_{i=1}^{n_j} k_i^w}$

, k_i^w is the weighted degree value of node v_i.

- "The systematic risk of [an IT project] portfolio depends on the project elements and their relationships "
- Model IT projects as biological agents who exist on a biological system (the project portfolio)
- Use network entropy to find optimal balance in uncertainty between efficiency and risk.



Open Research Questions for Entropy in Networks?

"it can be fairly argued that the field is in its infancy." [5]

> Many more probability distributions to explore (eigenvector centrality, k-shells, and/or clustering coefficient, among others)

> directed and/or weighted graphs

References

[1] D. Easley and J. Kleinberg. Networks, Crowds, and Markets: Reasoning about a Highly Connected World. Cambridge University Press, 2010.

[2] Ortiz-Arroyo, D.; Hussain, D.A. An information theory approach to identify sets of key players. In Proceedings of the European Conference on Intelligence and Security Informatics, Esbjerg, Denmark, 10.3–5 December 2008, Springer: Berlin/Heidelberg, Germany, 2008; pp. 15–26.

[3] Wiedermann, M.; Donges, J.F.; Kurths, J.; Donner, R.V. Mapping and discrimination of networks in the complexity-entropy plane. *Phys. Rev. E* **2017**, *96*, 042304.

[4] Wang, Q., Zeng, G., & Tu, X. (2017). Information technology project portfolio implementation process optimization based on complex network theory and entropy. Entropy, 19(6), 287.

[5] Omar, Y. M., & Plapper, P. (2020). A survey of information entropy metrics for complex networks. Entropy, 22(12), 1417.

Questions?