Complexity of Families of Multigraphs

Ove Frank[*]   Termeh Shafie[†]

Abstract

 This article describes families of finite multigraphs with labeled or unlabeled edges and vertices. It shows how size and complexity vary for different types of equivalence classes of graphs defined by ignoring only edge labels or ignoring both edge and vertex labels. Complexity is quantified by the distribution of edge multiplicities, and different complexity measures are discussed. Basic occupancy models for multigraphs are used to illustrate different graph distributions on isomorphism and complexity. The loss of information caused by ignoring edge and vertex labels is quantified by entropy and joint information that provide tools for studying properties of and relations between different graph families.

Key Words: labeled graph, edge multiplicity, complexity measure, entropy, joint information, isomorphism

1. Introduction

Typical applications of graphs consider sequences of edges associated with vertex pairs. For instance, records of telephone calls, internet connections, money transactions or business contacts during a period of time and their distributions on pairs of individuals, addresses, bank accounts or companies are four such applications. Multigraphs appear natural in many such contexts. A random multigraph is a family of multigraphs with a probability distribution, and appropriately chosen it can be a model for the application. Information theoretic tools can be used to describe, evaluate and compare different models, and they are particularly useful to analyze variability and dependence structures in multivariate data of network type. A survey of such information theoretic tools based on entropy measures is given by Frank(2011a) . Statistical analysis of network data is treated in a book by Kolaczyk (2009) and in survey articles by Frank (2005, 2011b). Many other issues concerning network analysis are also found in the encyclopedia edited by Carrington, Scott and Wasserman (2005), Meyers (2009), and Scott and Carrington (2011).

This article focuses on basic occupancy models adapted to fit multigraphs. The complexity of a multigraph is defined as its multiplicity distribution, that is the frequencies of vertex pairs with different numbers of multiple edges. The relationships between labeled and unlabeled graphs, isomorphism and complexity are specified in the next section. The numbers of graphs of various types are given and illustrated in Sections 3 and 4. Section 5 describes different complexity measures. Uniform graph models are analyzed and illustrated in Sections 6 to 8. Some other models are presented in Section 9 together with some comments on extensions and references.

2. Basic Concepts and Notation

A finite graph g with n labeled vertices and m labeled edges associates with each edge an ordered or unordered vertex pair. Let $V = \{1, \ldots, n\}$ and $E = \{1, \ldots, m\}$ be the sets of vertices and edges labeled by integers, and denote by R the set of available sites for the

---

[*]Department of Statistics, Stockholm University, SE-106 91 Stockholm, Sweden
[†]Department of Statistics, Stockholm University, SE-106 91 Stockholm, Sweden

edges. For directed graphs $R = V^2$ or $R = \{(i,j) \in V^2 : i \neq j\}$ depending on whether or not loops are allowed. For undirected graphs we use the site space $R = \{(i,j) \in V^2 : i \leq j\}$ or $R = \{(i,j) \in V^2 : i < j\}$ and consider (i,j) with $i \leq j$ as a canonical representation for the unordered vertex pair. Let r be the number of sites so that $r = n^2$, $n(n-1)$, $\binom{n+1}{2}$ or $\binom{n}{2}$ for the cases mentioned. The graph $g : E \to R$ is an injective map that is represented by an ordered sequence

$$\mathbf{g} = (g_1, \ldots, g_m) \in R^m$$

of m sites for the edges, or, equivalently, by an ordered partition

$$\mathbf{M} = (M_{ij} : (i,j) \in R)$$

of r disjoint subsets of edges for the sites. Here

$$M_{ij} = \{k \in E : g_k = (i,j)\} \quad \text{for} \quad (i,j) \in R \ .$$

Edges at the same site are called multiple edges, and the number of multiple edges at site (i,j) is the multiplicity denoted by m_{ij} for $(i,j) \in R$. We use the notation

$$\mathbf{g} \leftrightarrow \mathbf{M}$$

for the bijection between the two representations of the graph g.

If edges are not distinguished, their labels can be ignored, and order in $\mathbf{g}$ is irrelevant. A representation for the graph with labeled vertices but unlabeled edges is denoted by $\mathbf{g}^*$ and defined by listing the sites in $\mathbf{g}$ in some canonical order such as

$$(1,1) < (1,2) < \ldots < (1,n) < (2,1) < (2,2) < \ldots \ .$$

A convenient shorthand notation is

$$\mathbf{g}^* = ((i,j)^{m_{ij}} : (i,j) \in R) \ .$$

There is a bijection between the unordered site sequence for the edges and the multiplicity sequence for the edges:

$$\mathbf{g}^* \leftrightarrow \mathbf{m} = (m_{ij} : (i,j) \in R) \ .$$

If both vertex labels and edge labels are ignored, the isomorphic unlabeled graph is represented by $\mathbf{G}$. The unordered version of $\mathbf{M}$ is an unordered partition $\mathbf{M}^*$ of the edge set into r subsets. The unordered version of the multiplicity sequence $\mathbf{m}$ is an unordered partition $\mathbf{m}^*$ of m into r non-negative integers. There is a bijection between this partition and the sequence of frequencies of sites with multiplicities $0, 1, \ldots, m$ given by $\mathbf{r} = (r_0, \ldots, r_m)$ where

$$r_k = \sum_{(i,j) \in R} I(m_{ij} = k) \quad \text{for} \quad k = 0, 1, \ldots, m \ .$$

Thus,

$$\mathbf{m}^* \leftrightarrow \mathbf{r} \ ,$$

and the sequence $\mathbf{r}$ is called the complexity of the graph g. Figure 1 shows a schematic view of bijections and other functional relationships between the various concepts introduced here. The functional relationships comprise canonizing ordering (denoted by $*$), specifying multiplicities $\mathbf{m} = \mathbf{m}(\mathbf{g})$, specifying isomorphism $\mathbf{G} = \mathbf{G}(\mathbf{m})$ which is a function of $\mathbf{m}$, and specifying complexity $\mathbf{r} = \mathbf{r}(\mathbf{G})$ which is a function of $\mathbf{G}$. With an abuse of notation we also write $\mathbf{G} = \mathbf{G}(\mathbf{m}) = \mathbf{G}(\mathbf{g})$ and $\mathbf{r} = \mathbf{r}(\mathbf{G}) = \mathbf{r}(\mathbf{m}) = \mathbf{r}(\mathbf{g})$.
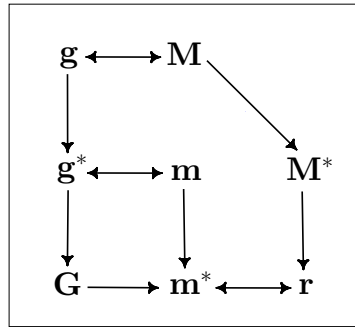
Figure 1: Relationships between graphs, multiplicity and complexity.

3. A Numerical Example

The different concepts introduced are illustrated and visualized by considering a simple example before we turn to general formulae for numbers of graphs and equivalence classes of different kinds.

Consider undirected graphs with $n = 4$ labeled vertices and $m = 3$ labeled edges with loops not allowed so that $r = 6$. Here $m < r$ and all partitions of 3 into positive integers can be used to find the possible multiplicity sequences. Thus the multiplicity sequences divide into three equivalence classes corresponding to permutations of $(3, 0, 0, 0, 0, 0)$, of $(2, 1, 0, 0, 0, 0)$, and of $(1, 1, 1, 0, 0, 0)$. Shorthand notation is $\sim 30^5$, $\sim 210^4$, and $\sim 1^3 0^3$. The classes have complexity sequences $(5, 0, 0, 1)$, $(4, 1, 1, 0)$, and $(3, 3, 0, 0)$. The classes consist of 1,2, and 3 non-isomorphic graphs, and the isomorphisms are shown in Figure 2.
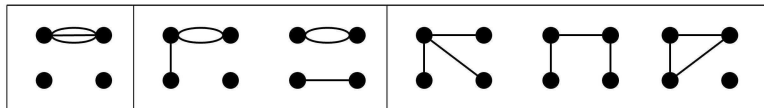

Figure 2: Unlabeled graphs according to complexity.

Vertex labels can be assigned to the non-isomorphic graphs in 6, 24, 6, 4, 12, and 4 ways, and edge labels can be assigned to each vertex labeled graph with the same complexity in 1, 3, and 6 ways in the order shown in Figure 2. Table 1 lists the number of unlabeled graphs $\#(\mathbf{G}|\mathbf{r})$, the number of vertex labeled graphs $\#(\mathbf{m}|\mathbf{r})$, and the number of fully labeled graphs $\#(\mathbf{g}|\mathbf{r})$ for each complexity sequence $\mathbf{r}$. Table 2 gives the numbers $\#(\mathbf{m}|\mathbf{G})$ and $\#(\mathbf{g}|\mathbf{G})$ of vertex labeled and fully labeled graphs for each isomorphism class.

Table 1: Distributions on complexity for graphs with 4 vertices, 3 edges and no loops.

Complexity	(5,0,0,1)	(4,1,1,0)	(3,3,0,0)	Total
Unlabeled graphs	1	2	3	6
Vertex labeled graphs	6	30	20	56
Fully labeled graphs	6	90	120	216

Table 2: Distributions on isomorphism for graphs with 4 vertices, 3 edges and no loops.

Isomorphism							Total
Vertex labeled graphs	6	24	6	4	12	4	56
Fully labeled graphs	6	72	18	24	72	24	216

4. Numbers of Graphs

The number of multigraphs with n labeled vertices, m labeled edges and r available sites of vertex pairs for the edges is given by the number of sequences $\mathbf{g}$, which is denoted $\#(\mathbf{g}) = r^m$. When edge labels are ignored, the number of graphs is given by the number of multiplicity sequences $\mathbf{m}$, which is the number of ordered partitions of m into r non-negative integers:

$$\#(\mathbf{m}) = \binom{m+r-1}{m}.$$

Each graph with only vertex labels can be edge labeled in the number of ways that $\mathbf{g}^*$ can be permuted, which is equal to

$$\#(\mathbf{g}|\mathbf{m}) = \binom{m}{\mathbf{m}} = \frac{m!}{\prod_{(i,j)\in R} m_{ij}!}.$$

The total $r^m = \sum_{\mathbf{m}} \binom{m}{\mathbf{m}}$ is a sum over $\binom{m+r-1}{m}$ terms.

Different fully labeled graphs are isomorphic if they are equal when vertex labels as well as edge labels are ignored. The number of isomorphic fully labeled graphs is given by $\binom{m}{\mathbf{m}}$ multiplied by the number of isomorphic vertex labeled graphs with no edge labels. Formally,

$$\#(\mathbf{g}|\mathbf{G}) = \binom{m}{\mathbf{m}} \#(\mathbf{m}|\mathbf{G})$$

since $\binom{m}{\mathbf{m}}$ is invariant for graphs isomorphic to $\mathbf{G}$.

Multiplicity sequences have the same complexity if they are permutations of the same $\mathbf{m}^*$. There are $\binom{r}{\mathbf{r}}$ such permutations where $\mathbf{r} \leftrightarrow \mathbf{m}^*$. Thus,

$$\#(\mathbf{m}|\mathbf{r}) = \binom{r}{\mathbf{r}} = \frac{r!}{\prod_{k=0}^{m} r_k!}$$

is the number of graphs with vertex labels but no edge labels having complexity $\mathbf{r}$. The number of fully labeled graphs with complexity $\mathbf{r}$ is obtained by multiplying $\#(\mathbf{m}|\mathbf{r})$ with $\#(\mathbf{g}|\mathbf{m})$:

$$\#(\mathbf{g}|\mathbf{r}) = \binom{m}{\mathbf{m}}\binom{r}{\mathbf{r}} = \frac{m!\, r!}{\prod_{k=0}^{m} k!^{r_k}\, r_k!}.$$

The number of different complexity sequences $\mathbf{r}$ is the same as the number of unordered partitions of m into r non-negative integers. This number is the sum of the numbers of unordered partitions of m into k positive integers for $k = 1, 2, \ldots, \min(r, m)$. If a_{mk} denotes the number of partitions of m into k positive integers and $a_m = a_{m1} + \ldots + a_{mm}$ is the number of partitions of m, it is possible to show that $a_{mk} = a_{m-k}$ for $k \geq m/2$.

Tables of a_m and a_{mk} for $k < m/2$ and $m = 1, 2, \ldots$ can be used to find

$$
\#(\mathbf{r}) = \begin{cases}
\sum_{k=1}^{r} a_{mk} & \text{for } r \leq \frac{m}{2} \\
\sum_{k<m/2}(a_k + a_{mk}) & \text{for } \frac{m}{2} < r < m \\
a_m & \text{for } r \geq m.
\end{cases}
$$

Such tables are available, for instance, in Comtet (1974).

5. Complexity of Graphs

The complexity sequence $\mathbf{r}$ contains the distribution of multiplicities among the sites. Summary measures of this distribution might be of interest as measures of complexity focusing on special properties of the graph. For instance, the proportion of multiple sites

$$
\frac{(r - r_0 - r_1)}{r}
$$

or the average multiplicity among multiple sites

$$
\frac{(m - r_1)}{r - r_0 - r_1}
$$

are simple measures of complexity focusing on any kind of deviation from graphs without multiple edges. If loops are forbidden, this amounts to deviation from graph simplicity.

A measure that linearly combines the frequencies of different multiplicities is given by

$$
\sum_{k=2}^{m} \binom{k}{2} r_k \, ,
$$

which counts the number of pairs of edges associated with the same site. If loops are forbidden, this measure is positive if and only if the graph is not simple. Another linear measure with this property is

$$
\sum_{k=0}^{m} r_k \log k! \, ,
$$

which is the logarithm of the common number of permutations that leave the edge sequence $\mathbf{g}$ invariant for graphs of complexity $\mathbf{r}$.

A special class of complexity measures focuses on how many graphs of different kinds that have the same complexity $\mathbf{r}$. Since these numbers might be very large, it is convenient to consider logarithmic measures. For vertex labeled and fully labeled graphs with the same complexity, the measures are

$$
\log \#(\mathbf{m}|\mathbf{r}) = \log r! - \log \mathbf{r}! = \log r! - \sum_{k=0}^{m} \log r_k!
$$

and

$$
\log \#(\mathbf{g}|\mathbf{r}) = \log m! + \log r! - \sum_{k=0}^{m} (r_k \log k! + \log r_k!) \, .
$$

These measures are similar to measures based on entropy, which characterize flatness of the relative frequency distributions $\mathbf{m}/m$ and $\mathbf{r}/r$. One might think of the entropy as a

measure of the range or dimension of a latent flat distribution (see Section 7). The entropy of the relative edge frequencies at different sites is given by

$$h(\mathbf{m}/m) = \sum_{(i,j) \in R} \varphi(m_{ij}/m)$$

where

$$\varphi(p) = \begin{cases} -p \log p & \text{for } p > 0 \\ 0 & \text{for } p = 0. \end{cases}$$

The entropy of the relative site frequencies for different multiplicities is given by

$$h(\mathbf{r}/r) = \sum_{k=0}^{m} \varphi(r_k/r) .$$

It follows that the entropy of $\mathbf{m}/m$ is equal to

$$h(\mathbf{m}/m) = \log m - \frac{1}{m} \sum_{k=2}^{m} r_k k \log k .$$

This entropy is non-negative, zero only for all edges at the same site, and it attains a maximum value of $\log r$ only if m is a multiple of r and the multiplicity distribution is uniform with the same multiplicity m/r at all sites. For $m < r$, the maximum is $\log m$ and is attained for all edges at different sites. For other cases with $m > r$ the maximal value is somewhat below $\log r$ and attained for an almost uniform distribution.

It also follows that the entropy of $\mathbf{r}/r$ is equal to

$$h(\mathbf{r}/r) = \log r - \frac{1}{r} \sum_{k=0}^{m} r_k \log r_k .$$

This entropy is non-negative, zero only for all multiplicities equal, and it attains a maximum value of $\log(m + 1)$ only in the degenerate case $m = 1, r = 2$. The maximal values for other cases are lower but not easily found.

For large values of m, Stirling's formula can be used to show that

$$h(\mathbf{m}/m) = \frac{1}{m} \log \binom{m}{\mathbf{m}} + O\left(\frac{\log m}{m}\right) \approx \frac{1}{m} \log \#(\mathbf{g}|\mathbf{m})$$

so that the entropy of $\mathbf{m}/m$ is approximately equal to the average number of bits (provided logarithms to base 2 are used) per edge needed to generate all $\mathbf{g}$ corresponding to $\mathbf{m}$.

6. Uniform Graph Models

The classical occupancy models with equal or unequal objects distributed among equal or unequal sites can be modified to fit graph data with its special combinatorial structure for the sites. We focus here on uniform distributions for different families of graphs. Families of graphs are conveniently specified as random graphs. The uniform distributions might be null models used to test or explore empirical graph families. The range of applications for such null models is conveniently extended to families of subgraphs induced by vertices of special kinds.

Assume that $\boldsymbol{\xi}$ is the edge sequence of a random graph that is uniform with probabilities

$$P(\boldsymbol{\xi} = \mathbf{g}) = \frac{1}{r^m} \quad \text{for} \quad \mathbf{g} \in R^m.$$

In this case the probability distributions of the different functions $\mathbf{m}(\boldsymbol{\xi})$, $\mathbf{G}(\boldsymbol{\xi})$, and $\mathbf{r}(\boldsymbol{\xi})$ are simply given as the relative frequencies of outcomes of $\boldsymbol{\xi}$ that are consistent with the outcomes of the functions. Thus

$$P(\mathbf{m}(\boldsymbol{\xi}) = \mathbf{m}) = \frac{\binom{m}{\mathbf{m}}}{r^m} \,,$$

$$P(\mathbf{G}(\boldsymbol{\xi}) = \mathbf{G}) = \frac{\#(\mathbf{g}|\mathbf{G})}{r^m} \,,$$

$$P(\mathbf{r}(\boldsymbol{\xi}) = \mathbf{r}) = \frac{m! \, r!}{r^m \prod_{k=0}^{m} k!^{r_k} \, r_k!} \,.$$

The entropy of a random variable is the same as the entropy of its probability distribution, so

$$H(\boldsymbol{\xi}) = \sum_{\mathbf{g}} \varphi(P(\boldsymbol{\xi} = \mathbf{g})) = m \log r \,.$$

Using calculation rules for entropy (given for instance in Frank, 2011a) it follows that

$$H(\mathbf{m}(\boldsymbol{\xi})) = H(\boldsymbol{\xi}) - E\left[\log\binom{m}{\mathbf{m}(\boldsymbol{\xi})}\right] \,,$$

$$H(\mathbf{G}(\boldsymbol{\xi})) = H(\boldsymbol{\xi}) - E\left[\#(\mathbf{g}|\mathbf{G}(\boldsymbol{\xi}))\right] \,,$$

$$H(\mathbf{r}(\boldsymbol{\xi})) = H(\boldsymbol{\xi}) - E\left[\log\binom{m}{\mathbf{m}(\boldsymbol{\xi})}\right] - E\left[\log\binom{r}{\mathbf{r}(\boldsymbol{\xi})}\right] \,.$$

Using that $m_{ij}(\boldsymbol{\xi})$ is binomially distributed with parameters m and $1/r$, the entropy of the multiplicity sequence can be expressed as

$$H(\mathbf{m}(\boldsymbol{\xi})) = m \log r - \log m! + r \sum_{k=2}^{m} \binom{m}{k} \left(\frac{1}{r}\right)^k \left(1 - \frac{1}{r}\right)^{m-k} \log k! \,.$$

The entropies of $\mathbf{G}(\boldsymbol{\xi})$ and $\mathbf{r}(\boldsymbol{\xi})$ can be numerically evaluated but seem to have no explicit formulae.

Consider now an alternative model with edge sequence $\boldsymbol{\eta}$ assuming that $\mathbf{m}(\boldsymbol{\eta})$ is uniform and that $\boldsymbol{\eta}$ conditional on $\mathbf{m}(\boldsymbol{\eta})$ is uniform. In this case

$$P(\mathbf{m}(\boldsymbol{\eta}) = \mathbf{m}) = \frac{1}{\binom{m+r-1}{m}} \,,$$

$$P(\mathbf{r}(\boldsymbol{\eta}) = \mathbf{r}) = \frac{\binom{r}{\mathbf{r}}}{\binom{m+r-1}{m}} \,,$$

$$P(\boldsymbol{\eta} = \mathbf{g}) = \frac{1}{\binom{m}{\mathbf{m}(\mathbf{g})}\binom{m+r-1}{m}} \,,$$

and

$$H(\mathbf{m}(\boldsymbol{\eta})) = \log\binom{m+r-1}{m} \,,$$

$$H(\mathbf{r}(\boldsymbol{\eta})) = \frac{1}{\binom{m+r-1}{m}} \sum_{\mathbf{r}} \binom{r}{\mathbf{r}} \log\binom{r}{\mathbf{r}} - \log\binom{m+r-1}{m} \,,$$

$$H(\boldsymbol{\eta}) = \frac{1}{\binom{m+r-1}{m}} \sum_{\mathbf{m}} \log\binom{m}{\mathbf{m}} + \log\binom{m+r-1}{m} \,.$$

The two uniform models considered are in physics referred to as the Maxwell-Boltzmann model with uniform distribution of unequal particles in unequal cells, and the Bose-Einstein model with uniform distribution of equal particles in unequal cells.

For the fully labeled graphs the entropy of $\boldsymbol{\xi}$ is maximal, and the entropy of $\boldsymbol{\eta}$ deviates by

$$D_1 = H(\boldsymbol{\xi}) - H(\boldsymbol{\eta})$$

from it. For the vertex labeled graphs the entropy of $\mathbf{m}(\boldsymbol{\eta})$ is maximal and the entropy of $\mathbf{m}(\boldsymbol{\xi})$ deviates by

$$D_2 = H(\mathbf{m}(\boldsymbol{\eta})) - H(\mathbf{m}(\boldsymbol{\xi}))$$

from it. Therefore the reductions in entropy caused by skipping edge labels is larger for $\boldsymbol{\xi}$ than for $\boldsymbol{\eta}$,

$$H(\boldsymbol{\xi}) - H(\mathbf{m}(\boldsymbol{\xi})) \geq H(\boldsymbol{\eta}) - H(\mathbf{m}(\boldsymbol{\eta})) \,,$$

and the difference between the reductions is equal to the sum $D_1 + D_2$ of the two deviations from maximal entropy. This can also be expressed as the following ordering of the entropies

$$H(\mathbf{m}(\boldsymbol{\xi})) \leq H(\mathbf{m}(\boldsymbol{\eta})) \leq H(\boldsymbol{\eta}) \leq H(\boldsymbol{\xi}) \,.$$

Some of the simplified complexity measures mentioned in Section 5 rely on the frequencies of sites with no or single occupancy only. The distributions of $r_0(\boldsymbol{\xi})$ and $r_1(\boldsymbol{\xi})$ are obtained as marginal distributions of $\mathbf{r}(\boldsymbol{\xi})$. For $r_0(\boldsymbol{\xi})$ the marginal probabilities are given by

$$P(r_0(\boldsymbol{\xi}) = r_0) = \frac{r!}{r_0!\, r^m} \sum S_m(r_1, \ldots, r_m)$$

$$= \frac{r!\, S(m, r - r_0)}{r_0!\, r^m} \qquad \text{for } r_0 = 0, 1, \ldots, r - 1 \,.$$

Here the sum extends over $(r_1, \ldots, r_m)$ satisfying $\sum_{k=1}^{m} r_k = r - r_0$ and $\sum_{k=1}^{m} k\, r_k = m$. The term

$$S_m(r_1, \ldots, r_m) = \frac{m!}{\prod_{k=1}^{m} k!^{r_k}\, r_k!}$$

counts the number of partitions of the edge set into r_1 singletons, r_2 parts of size 2, etc. The sum is equal to $S(m, r - r_0)$, which is a Stirling number of the second kind for the number of partitions of an m-set into $r - r_0$ non-empty disjoint subsets.

For the bivariate distribution of $(r_0(\boldsymbol{\xi}), r_1(\boldsymbol{\xi}))$ the probabilities for $r_0 < r$ and $r_1 \leq \min(m, r - r_0)$ are obtained as

$$P(r_0(\boldsymbol{\xi}) = r_0, r_1(\boldsymbol{\xi}) = r_1) = \frac{r!}{r_0!\, r_1!\, r^m} \sum S_m(0, r_2 \ldots, r_m)$$

where the sum extends over $(r_2, \ldots, r_m)$ satisfying $\sum_{k=2}^{m} r_k = r - r_0 - r_1$ and $\sum_{k=2}^{m} k\, r_k = m - r_1$. Thus, to evaluate the sum we need to specify the partitions of $m - r_1$ into $r' = r - r_0 - r_1$ integers larger than 2 and find the terms separately. The number of terms is the same as the number of partitions of $m' = m - r_1 - r' = m - r + r_0$ into r' positive integers, that is the number $a_{m'r'}$ given at the end of Section 4.

Upper and lower bounds to the bivariate probability can be found much easier, and they are based on that

$$S_m(0, r_2 \ldots, r_m) = \frac{m!\, S_{m'}(r_2 \ldots, r_m)}{m'!\, \prod_{k=2}^{m} k^{r_k}}$$

where

$$m' = \sum_{k=1}^{m-1} k\, r_{k+1} = \sum_{k=2}^{m} (k - 1) r_k = m - r + r_0 \,.$$

Moreover, the geometric mean of the multiplicities is bounded between 2 and the arithmetic mean, which implies that there are bounds α and β so that

$$\alpha = 2^{r'} \leq \prod_{k=2}^{m} k^{r_k} \leq \left[(m'+r')/r'\right]^{r'} = \beta$$

for $r' > 0$ and $m' > 0$. Therefore,

$$\frac{m!\, S_{m'}(r_2, \ldots, r_m)}{m'!\, \beta} \leq S_m(0, r_2, \ldots, r_m) \leq \frac{m!\, S_{m'}(r_2, \ldots, r_m)}{m'!\, \alpha}$$

and, consequently,

$$\frac{r!\, m!\, S(m', r')}{r_0!\, r_1!\, m'!\, r^m\, \beta} \leq P(r_0(\boldsymbol{\xi}) = r_0, r_1(\boldsymbol{\xi}) = r_1) \leq \frac{r!\, m!\, S(m', r')}{r_0!\, r_1!\, m'!\, r^m\, \alpha} \, ,$$

where the lower bound to the probability is often quite accurate.

The probability that there are no multiple edges is given by

$$P(r_1(\boldsymbol{\xi}) = m) = P(r_0(\boldsymbol{\xi}) = r - m, r_1(\boldsymbol{\xi}) = m) = \frac{m! \binom{r}{m}}{r^m} \quad \text{for } r \geq m \, .$$

The distributions of $r_0(\boldsymbol{\eta})$ and $r_1(\boldsymbol{\eta})$ for the model with uniform vertex labeled graphs are given by

$$P(r_0(\boldsymbol{\eta}) = r_0) = \frac{\binom{r}{r_0}\binom{m-1}{r-r_0-1}}{\binom{m+r-1}{m}} \quad \text{for } r_0 = 0, 1, \ldots, r - 1$$

and

$$P(r_0(\boldsymbol{\eta}) = r_0, r_1(\boldsymbol{\eta}) = r_1) = \frac{\binom{r}{r_0}\binom{r-r_0}{r_1}\binom{m-r+r_0-1}{r-r_0-r_1-1}}{\binom{m+r-1}{m}}$$

for $r_0 < r$ and $r_1 \leq \min(m, r - r_0)$. The first case is proved by noticing that when the r_0 empty sites have been chosen, the remaining $r' = r - r_0$ sites should have at least one edge per site, and the remaining $m' = m - r'$ edges can be distributed in any of $\binom{m'+r'-1}{m'}$ ways. Similarly, in the second case, when the r_0 empty sites and the r_1 single occupancy sites have been chosen, the remaining $r' = r - r_0 - r_1$ sites should have at least two edges per site, and the remaining $m' = m - r_1 - 2r'$ edges can be distributed in any of $\binom{m'+r'-1}{m'}$ ways.

In this model the probability that there are no multiple edges is given by

$$P(r_1(\boldsymbol{\eta}) = m) = P(r_0(\boldsymbol{\eta}) = r - m, r_1(\boldsymbol{\eta}) = m) = \frac{\binom{r}{m}}{\binom{m+r-1}{m}} \quad \text{for } r \geq m \, .$$

Now $\binom{m+r-1}{m} = r(r+1)\cdots(r+m-1)/m! \geq r^m/m!$, so obviously the graph property of having no multiple edges has a smaller probability under the $\boldsymbol{\eta}$-model than under the $\boldsymbol{\xi}$-model.

The entropy of $(r_0(\boldsymbol{\eta}), r_1(\boldsymbol{\eta}))$ is smaller than the entropy of the complete complexity sequence $\mathbf{r}(\boldsymbol{\eta})$. The difference reveals how much information is lost by using the simpler complexity measure. A simple illustration showing that simple complexity summaries can be quite satisfactory is given in Table 3. We see that the outcomes of (r_0, r_1) match $\mathbf{r}$ quite well, so that there is not much uncertainty about $\mathbf{r}$ when (r_0, r_1) is known. In fact, for the $\boldsymbol{\xi}$-model the entropies are $H(\mathbf{r}(\boldsymbol{\xi})) = 2.82$ and $H(r_0(\boldsymbol{\xi}), r_1(\boldsymbol{\xi})) = 2.78$ so only

about one percent of the information about complexity is lost by using the simpler complexity measure. The univariate entropies $H(r_0(\boldsymbol{\xi})) = 1.95$ and $H(r_1(\boldsymbol{\xi})) = 2.50$ are also retaining almost the same information as the bivariate entropy. For the $\boldsymbol{\eta}$-model we find $H(\mathbf{r}(\boldsymbol{\eta})) = 3.63$, $H(r_0(\boldsymbol{\eta}), r_1(\boldsymbol{\eta})) = 3.38$, $H(r_0(\boldsymbol{\eta})) = 2.11$, and $H(r_1(\boldsymbol{\eta})) = 2.50$ which implies that about 7% of the information about complexity is lost by the simpler measure.

Table 3: Number of outcomes of complexity $\mathbf{r} = (r_0, r_1, \ldots, r_m)$ for given numbers r_0, r_1 of empty and single occupancy sites in graphs with 5 vertices, 8 edges and no loops.

r_0	r_1 0	1	2	3	4	5	6	7	8
2									1
3							1		
4				1	1				
5			1	1	1				
6	1	1	2	1					
7	2	2	1						
8	3	1							
9	1								

7. Entropy and Joint Information

Entropies are convenient measures of variation for general random variables but are also useful to determine dependence and other relationships between several random variables. This possibility can be intuitively understood by considering entropy as a measure of information, and interpreting it as the number of informative binary dimensions in a bijective representation of the outcomes. The technical interpretation of entropy as information refers to a property of latent codes. It is known that repeated independent outcomes of a random variable with N different possible outcomes and entropy H can be assigned binary sequences of different lengths according to a prefix code that requires in the long run no more than H binary digits (bits) per outcome. This corresponds to 2^H latent code sequences with uniform probabilities instead of N outcomes with arbitrary probabilities. The length of the latent codes, the entropy H, is called the information in the outcomes, and the extra length that a binary code would require for the outcomes, $\log N - H$, is called the redundancy in the outcomes. When two random variables ξ and η have common bits in their latent codes, they are related, and this relationship is measured by the joint information or joint entropy

$$J(\xi, \eta) = H(\xi) + H(\eta) - H(\xi, \eta) .$$

Joint information is zero if and only if the variables are independent and do not reveal any information about each other. Joint information is maximal when one of the variables is completely determined by the other. Joint information is an alternative to correlation and other measures that require numerical variables and specify linear or special non-linear regression relationships. Arbitrary functional relationships as well as various conditional dependence structures can be specified by different combinations of entropy measures. See Frank (2011a) for further details about such possibilities.

The total information in (ξ, η) minus the information in η is the expected remaining information in ξ when η is provided,

$$H(\xi, \eta) - H(\eta) = E[H(\xi|\eta)] ,$$

and the joint information is equal to the original information minus the remaining information in any of the variables according to

$$J(\xi, \eta) = H(\xi) - E[H(\xi|\eta)] = H(\eta) - E[H(\eta|\xi)] \ .$$

If η is determined by ξ, the difference $H(\xi) - H(\eta)$ is equal to the remaining information in ξ when η is provided, or, in other words, the information in ξ that is lost if nothing more than η is released.

Consider the edge sequence $\boldsymbol{\xi}$ of a random multigraph. The entropy of $\mathbf{G}(\boldsymbol{\xi})$ measures variation from uniformity or flatness in the probability distribution over the unlabeled graphs. The joint information of the multiplicity sequence $\mathbf{m}(\boldsymbol{\xi})$ and the complexity sequence $\mathbf{r}(\boldsymbol{\xi})$ is trivially equal to the entropy of $\mathbf{r}(\boldsymbol{\xi})$ because complexity is determined by multiplicity. Less transparent relationships between network properties might be between number of loops and number of multiple sites or any other network characteristics of special interest for the applications. Joint entropies reveal such relationships. Sometimes it is possible to give explicit expressions for the measures. A few examples are given in Section 9.

8. Some Further Illustrations

Numerical algorithms have been developed to handle distributions of edges among vertex pairs for arbitrary values of m and n. Here these algorithms are used for the case with $n = 6$ and $m = 4$ in order to visualize how families of multigraphs are composed of isomorphisms of varying complexity. We also illustrate the distributions on isomorphism and complexity of fully labeled and vertex labeled graphs. We evaluate the possibilities to gain information about them by using partial information.

Table 4 shows complexity distributions for uniform distributions over the fully labeled graphs, over the vertex labeled graphs, and over the unlabeled graphs. Random variables generating these graph families are the earlier defined edge sequences $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ together with an edge sequence $\boldsymbol{\zeta}$ with uniform distribution of $\mathbf{G}(\boldsymbol{\zeta})$ over the isomorphisms and with $\boldsymbol{\zeta}$ uniform conditional on $\mathbf{G}(\boldsymbol{\zeta})$.

Table 4: Distributions on complexity for graphs with 6 vertices, 4 edges and no loops.

Complexity	(14,0,0,0,1)	(13,1,0,1,0)	(13,0,2,0,0)	(12,2,1,0,0)	(11,4,0,0,0)	Total
Unlabeled graphs	1	2	2	7	9	21
Vertex labeled graphs	15	210	105	1365	1365	3060
Fully labeled graphs	15	840	630	16380	32760	50625

The complexity distributions have entropies $H(\mathbf{r}(\boldsymbol{\xi})) = 1.11$, $H(\mathbf{r}(\boldsymbol{\eta})) = 1.51$ and $H(\mathbf{r}(\boldsymbol{\zeta})) = 1.91$. Maximum entropy is here equal to $\log 5 = 2.32$. Thus, the complexity distributions have redundancies of 52%, 35%, and 18%, and all distributions exhibit a clear concentration towards simplicity with no or few multiple edges.

From the distributions on isomorphisms in Figure 3 it follows that $H(\mathbf{G}(\boldsymbol{\xi})) = 3.87$, $H(\mathbf{G}(\boldsymbol{\eta})) = 3.95$, and $H(\mathbf{G}(\boldsymbol{\zeta})) = 4.39$. The unlabeled graphs have maximal entropy $\log 21 = 4.39$ obtained for the $\boldsymbol{\zeta}$-model. The vertex labeled graphs have maximal entropy for the $\boldsymbol{\eta}$-model, and 34% of that entropy is retained by $\mathbf{G}(\boldsymbol{\eta})$. The fully labeled graphs have maximal entropy for the $\boldsymbol{\xi}$-model, and 25% of that entropy is retained by $\mathbf{G}(\boldsymbol{\xi})$. A more complete and systematic view of how the information content in different kinds of data varies for the three models is given in Table 5. The models are constructed to have no redundancy for one of the data levels. All other redundancies are between 4% and

12%, except for the complexity level which has higher redundancies. If maximal entropy is rounded upwards, a rough common feature of the models is apparent. Of the 16 binary dimensions required for fully labeled graphs, about 3 are informative about complexity, another 2 are informative about how the sites need to be ordered to achieve graph structure, another 7 are informative about vertex labeling, and another 4 about edge labeling.
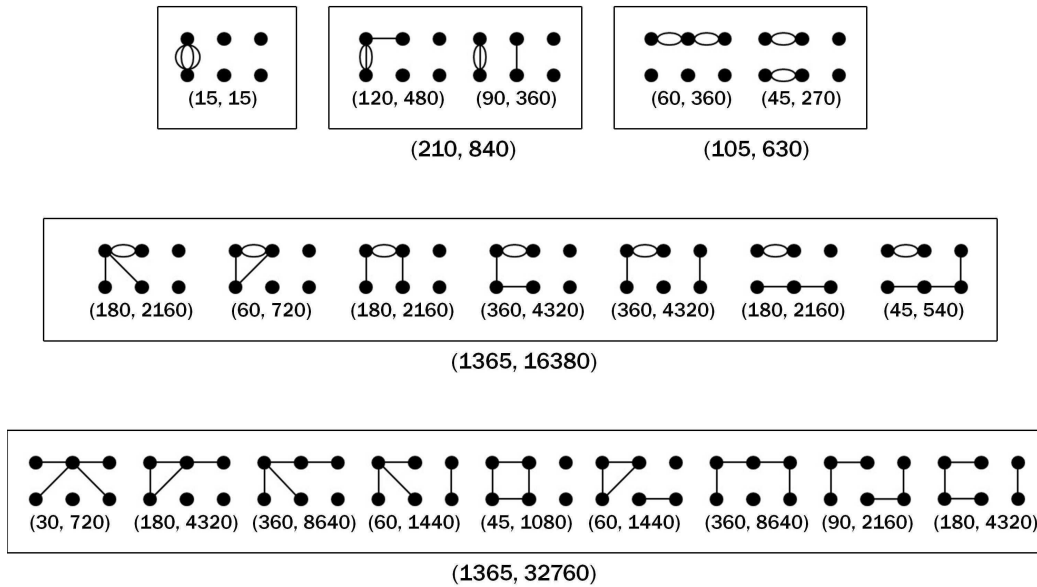

Figure 3: Number of labeled and fully labeled graphs for different isomophisms and complexities with 6 vertices, 4 edges and no loops.

Table 5: Entropy and maximal entropy of graph data under three uniform random models for graphs with 6 vertices, 4 edges and no loops.

	Entropy of data according to			
Data	ξ-model	η-model	ζ-model	Maximal entropy
Fully labeled graph	15.63	15.45	14.67	15.63
Vertex labeled graph	11.44	11.58	11.07	11.58
Unlabeled graph	3.87	3.95	4.39	4.39
Graph Complexity	1.11	1.51	1.91	2.32

9. Other graph models

A natural generalization of the uniform model for the sequence $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m)$ of sites for the edges is to assume that edges are independently assigned to sites according to a common arbitrary probability distribution

$$\mathbf{p} = (p_{ij} : (i, j) \in R)$$

over the possible sites of vertex pairs. Thus,

$$P(\boldsymbol{\xi} = \mathbf{g}) = \mathbf{p}^{\mathbf{m}(\mathbf{g})} = \prod_{(i,j) \in R} p_{ij}^{m_{ij}(\mathbf{g})} \qquad \text{for} \quad \mathbf{g} \in R^m .$$

The multiplicity sequence $\mathbf{m}(\boldsymbol{\xi})$ is multinomially distributed with parameters m and $\mathbf{p}$ so that

$$P(\mathbf{m}(\boldsymbol{\xi}) = \mathbf{m}) = \binom{m}{\mathbf{m}} \mathbf{p}^{\mathbf{m}}$$

for the $\binom{m+r-1}{m}$ different ordered partitions $\mathbf{m}$ of m into r non-negative integers. The complexity sequence $\mathbf{r}(\boldsymbol{\xi})$ has probabilities given by

$$P(\mathbf{r}(\boldsymbol{\xi}) = \mathbf{r}) = \sum_{\mathbf{m}|\mathbf{r}} \binom{m}{\mathbf{m}} \mathbf{p}^{\mathbf{m}} = \frac{m!}{\prod_{k=0}^{m} k!^{r_k}} \sum_{\mathbf{m}|\mathbf{r}} \mathbf{p}^{\mathbf{m}}$$

so, unless $\mathbf{p}$ is uniform, the sum needs a specification of all multiplicity sequences that have complexity $\mathbf{r}$. It is straightforward to find the entropies

$$H(\boldsymbol{\xi}) = -E\left[\log \mathbf{p}^{\mathbf{m}(\boldsymbol{\xi})}\right] = -m\mathbf{p}\log\mathbf{p} = m \sum_{(i,j)\in R} \varphi(p_{ij}) = m\, h(\mathbf{p})$$

and

$$H(\mathbf{m}(\boldsymbol{\xi})) = -E\left[\log \binom{m}{\mathbf{m}(\boldsymbol{\xi})} \mathbf{p}^{\mathbf{m}(\boldsymbol{\xi})}\right] = m\, h(\mathbf{p}) - E\left[\log \binom{m}{\mathbf{m}(\boldsymbol{\xi})}\right]$$

$$= m\, h(\mathbf{p}) - \log m! + \sum_{(i,j)\in R} E\left[\log m_{ij}(\boldsymbol{\xi})!\right]$$

$$= m\, h(\mathbf{p}) - \log m! + \sum_{(i,j)\in R} \sum_{k=0}^{m} \binom{m}{k} p_{ij}^k (1-p_{ij})^{m-k} \log k! \, .$$

For $m > r$ there has to be some multiplicity larger than 1, but for $m \leq r$ it might be of interest to find the probability of no multiple edges. If loops are forbidden, this is the same as the probability of graph simplicity. If loops are allowed, the number of loops

$$m_1 = \sum_{i=1}^{n} m_{ii}$$

and the number of sites r_0 and r_1 with no and single occupancy are statistics that suffice to specify graph simplicity. For the $\boldsymbol{\xi}$-model with common component distribution $\mathbf{p}$, the number of loops $m_1(\boldsymbol{\xi})$ is binomially distributed with parameters m and $p_1 = \sum_{i=1}^{n} p_{ii}$. The number $r_k(\boldsymbol{\xi})$ of sites with occupancy k has expected value

$$E\left[r_k(\boldsymbol{\xi})\right] = \sum_{(i,j)\in R} \binom{m}{k} p_{ij}^k (1-p_{ij})^{m-k} \qquad \text{for} \quad k = 0, 1, \ldots, m \, .$$

This implies the following expected values for the simple complexity measures given by the number of multiple occupancy sites and the number of multiple edges:

$$E\left[r - r_0(\boldsymbol{\xi}) - r_1(\boldsymbol{\xi})\right] = r - \sum_{(i,j)\in R} (1-p_{ij})^m - m \sum_{(i,j)\in R} p_{ij}(1-p_{ij})^{m-1}$$

$$E\left[m - r_1(\boldsymbol{\xi})\right] = m\left(1 - \sum_{(i,j)\in R} p_{ij}(1-p_{ij})^{m-1}\right) \, .$$

The probability that there are no multiple edges is given by the sum of all ordered different products of m of the r probabilities in $\mathbf{p}$, that is by

$$P(r_1(\boldsymbol{\xi}) = m) = m! \sum \prod_{(i,j)\in R} p_{ij}^{m_{ij}}$$

where the sum extends over all permutations of $\mathbf{m}^* = (0^{r-m}1^m)$.

For many applications, an important generalization of independent assignments of edges to vertex pairs is obtained by introducing stochastic processes that generate edge sequences. A simple setup is to define r independent Poisson point processes that generate edges at the different sites with intensities λ_{ij} for $(i,j) \in R$. The sequence $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m)$ has components ξ_k that record the sites in the order the edges occur during a fixed period of time. Such an approach is to be discussed elsewhere.

An investigation of entropy measures for occupancy models similar to those considered here is described in an article by Frank and Nowicki (1989). They introduce a graph on objects corresponding to our edges with their edges specifying whether or not the objects occupy the same site. Thus, this graph has complete connected components and is closely related to the concepts discussed here. They also develop asymptotic results for various entropies. Of special interest is the asymptotic entropy for the multinomial distribution, which implies that the multiplicities of the fully labeled graphs have an entropy $H(\mathbf{m}(\boldsymbol{\xi}))$ that for large m and r with r^2/m tending to zero is given by

$$H(\mathbf{m}(\boldsymbol{\xi})) = \frac{1}{2} \log \left[(2\pi em)^{r-1} \prod_{(i,j)\in R} p_{ij} \right] + O\left(\frac{r^2}{m}\right) .$$

Complexity is a general property considered in many different contexts and used with or without a specific definition. Complexity in graphs has been given different definitions in the literature focusing on other graph properties than edge multiplicity. For instance, Karreman (1955) and Mowshowitz (1968) are references that deal with completely different complexity properties of graphs used as models for molecules with chemical bonds between atoms. A common feature of many complexity concepts is that they seem to be well described and analyzed by information measures based on entropy.

REFERENCES

Carrington, P., Scott, J. and Wasserman, S. (eds.) (2005), *Models and Methods in Social Network Analysis*, New York: Cambridge University. Press.

Comtet, L. (1974), *Advanced Combinatorics: The Art of Finite and Infinite Expansions*, Dordrecht: Reidel Publishing Company.

Frank, O. (2005), "Network Sampling and Model Fitting," in *Models and Methods in Social Network Analysis*, eds. P. Carrington, J. Scott and S. Wasserman, New York: Cambridge University Press, pp. 31–56.

Frank, O.(2011a), "Statistical Information Tools for Multivariate Discrete Data," in *Modern Mathematical Tools and Techniques in Capturing Complexity*, eds. L. Pardo, N. Balakrishnan and M. Ángeles Gil, Berlin: Springer Verlag, pp. 177–190.

Frank, O. (2011b), "Survey Sampling in Networks," in *Handbook of Social Network Analysis*, eds J. Scott and P. Carrington, London: Sage Publications.

Frank, O. and Nowicki, K. (1989), "On Entropies of Occupancy Distributions," in *Combinatorics and Graph Theory*, eds. Z. Skupien, M. Borowiecki, Warsaw: Banach Center Publications, Volume 25, PWN-Polish Scientific Publishers, pp. 71–86.

Karreman, G. (1955), "Topological Information Content and Chemical Reactions," *Bulletin of Mathematical Biophysics*, 17, pp. 279–285.

Kolaczyk, E. (2009), *Statistical Analysis of Network Data*, New York: Springer Verlag.

Meyers, R. (ed.) (2009), *Encyclopedia of Complexity and Systems Science*, New York: Springer Verlag.

Mowshowitz, A. (1968), "Entropy and the Complexity of Graphs: I. An Index of the Relative Complexity of a Graph," *Bulletin of Mathematical Biophysics*, 30, pp. 175–204.

Scott, J. and Carrington, P. (eds.) (2011), *Handbook of Social Network Analysis*, London: Sage Publications.