Symmetry in Complex Networks

Ben D. MacArthur
Bone and Joint Research Group, University of Southampton,
Southampton General Hospital, Southampton, SO16 6YD, UK

Rubén J. Sánchez-García*
Mathematisches Institut, Heinrich-Heine Universität
Düsseldorf, Universitätsstr 1, 40225, Düsseldorf, Germany

James W. Anderson
School of Mathematics, University of Southampton,
Southampton, SO17 1BJ, UK

March 28, 2008

Abstract

We consider the size and structure of the automorphism groups of a variety of empirical 'real-world' networks and find that, in contrast to classical random graph models, many real-world networks are richly symmetric. We construct a practical network automorphism group decomposition, relate automorphism group structure to network topology and discuss generic forms of symmetry and their origin in real-world networks. We also comment on how symmetry can affect network redundancy and robustness.

keywords: Complex Network, Symmetry, Automorphism Group, Redundancy

*Ben MacArthur and Rubén Sánchez-García contributed equally to this work.

1 Introduction

The use of complex networks to model the underlying topology of 'real-world' complex systems – from social interaction networks such as scientific collaboration networks[32, 33] to biological regulatory networks[25] and technological networks such as the internet[39] – has attracted much current research interest[2, 34, 43]. Previous studies have highlighted the fact that seemingly disparate networks often have certain features in common including (amongst others): the 'small-world' property[45]; the power-law distribution of vertex degrees[4]; and network construction from motifs[31].

Identification of universal structural properties such as these allows generic network properties to be decoupled from system-specific features. In this present work we consider the symmetry structure of a variety of real-world networks and find that a certain degree of symmetry is also ubiquitous in complex systems. Although the symmetry structure of some types of well-ordered networks has received some attention[17, 28], a systematic study of the symmetry structure of real-world complex networks – which typically contain ordered and disordered elements – has not yet been undertaken.

This paper therefore investigates the origin and form of real-world network symmetry and its effect on network function. We consider network symmetry via the automorphism group of the underlying graph. Firstly, we identify 'essential' network symmetries and use these symmetries to derive a natural direct product decomposition of the automorphism group into irreducible factors. This decomposition is *per se* a very efficient way to handle large automorphism groups of real-world networks. We then associate with each factor in this decomposition a *symmetric subgraph* – the subgraph on which the factor subgroup acts non-trivially – and investigate the generic structure of symmetric subgraphs. Finally, by considering orbits of the automorphism group we investigate the relationship between network symmetry and redundancy.

2 Network Automorphism Groups

Mathematically, a *network* is a graph, $\mathscr{G} = (V(\mathscr{G}), E(\mathscr{G}))$, with vertex set, $V(\mathscr{G})$ (of size $N_{\mathscr{G}}$), and edge set, $E(\mathscr{G})$ (of size $M_{\mathscr{G}}$) where vertices are said to be *adjacent* if there is an edge between them. An *automorphism* is a permutation of the vertices of the network which preserves adjacency. The set of automorphisms under composition forms a group $\mathrm{Aut}(\mathscr{G})$ of size $a_{\mathscr{G}}$ which compactly describes network symmetry[11]. Throughout this discussion we shall let $\mathscr{G}$ refer to a generic network, and G to a generic group. If the network is a multi-digraph, we remove weights and directions and consider the automorphism group of the *underlying* graph.

Here, the `nauty` program[30] – which includes one of the most efficient graph isomorphism algorithms available[18] – is used to calculate the size and generating sets of the various automorphism groups.

Table 1 gives the order of the automorphism groups of some representative real-world complex networks, which includes a broad selection of biological, technological and social networks. In all cases these complex real-world networks have a nontrivial symmetry structure. Since almost all large graphs (including, for example, the classical Erdös-Rényi random graphs) are asymmetric[12] the ubiquity of symmetry in real-world systems is somewhat unexpected.

Network	$N_{\mathcal{G}}$	$M_{\mathcal{G}}$	$a_{\mathcal{G}}$	$r_{\mathcal{G}}$	% BSS
Biological Networks					
Human B Cell Genetic (BCell)[5]	5,930	64,645	5.9374×10^{13}	0.96	97.4
C. elegans Genetic (Cele)[46]	2,060	18,000	6.9985×10^{161}	0.85	98.7
BioGRID Human (BGHum)[40]	7,013	20,587	1.2607×10^{485}	0.87	99.5
BioGRID *S. cerevisiae* (BGScerev)[40]	5,295	50,723	6.8622×10^{64}	0.98	100
BioGRID *Drosophila* (BGDros)[40]	7,371	25,043	3.0687×10^{493}	0.87	99.2
BioGRID *Mus musculus* (BGMus)[40]	209	393	5.3481×10^{125}	0.32	100
Yeast Protein Interactions (Yeast)[24]	1,458	1,948	1.0667×10^{254}	0.70	95.1
c. elegans metabolic (CeleMet)[16]	453	2,040	1.9327×10^{10}	0.92	100
Technological Networks					
Internet (AS Level) (IntAS)[21]	22,332	45,392	$1.2822 \times 10^{11,298}$	0.51	98.4
US Power Grid (USPow)[45]	4,941	6,594	5.1851×10^{152}	0.90	88.1
US Airports (USAir)[36]	332	2,126	2.5916×10^{24}	0.83	93.3
www California search subnet (Calif)[27]	5,925	15,770	$1.2414 \times 10^{1,298}$	0.68	98.7
www EPA.gov subnet (EPA)[26]	4,253	8,897	$1.2772 \times 10^{2,321}$	0.52	98.0
www Political Blogs (PolBlog)[1]	1,222	16,714	2.3995×10^{35}	0.95	100
Social Networks					
Email[22]	1,133	5,452	1.5288×10^{9}	0.98	100
PGP users network (PGP)[10]	10,680	24,316	$4.4963 \times 10^{1,251}$	0.74	94.4
Media ownership (Media)[35]	4,475	4,652	$3.3638 \times 10^{4,818}$	0.20	90.1
Geometry Co-authorship (Geom)[8]	3,621	9,461	1.8994×10^{320}	0.77	96.7
Erdös Collaboration (Erdös)[7]	6,927	11,850	$3.4610 \times 10^{4,222}$	0.34	99.6
PhD network (PhD)[15, 38]	1,025	1,043	2.9810×10^{292}	0.50	86.7

Table 1: **The size of the automorphism group of some representative real-world networks**. The size of the automorphism group of the largest connected component is given (to 5 significant figures). Additionally, the percentage of geometric factors corresponding to basic symmetric subgraphs (BSSs) (see section 3) is given, as is the degree of structural redundancy present in the network, as quantified by $r_{\mathcal{G}}$ (see section 4). Note that almost all real-world symmetry is due to the presence of basic symmetric subgraphs and many networks contain a substantial degree of structural redundancy. Connected components were extracted using `Pajek`[6].

Many networks – for example the internet and the world wide web – are 'growing'[4] (that is, new vertices are added to the network over time). Generically, any growth process in which allows for new vertices to be added to the network *one at a time* naturally leads to a network with locally tree-like regions. Such locally tree-like areas are common in real-world networks and their presence is important because, while the majority of large graphs are asymmetric, it is common for large random trees to exhibit a high degree of symmetry[23], deriving from the presence of identical branches about the same fork. Thus we expect a certain degree of tree-like symmetry to be present in many real-world networks. In the following sections we determine the extent to which real-world symmetry is locally tree-like. We begin by considering the structure of network automorphism groups.

Consider the permutations of a set of n points $X = \{x_1, \ldots, x_n\}$. The *support* of a permutation p is the set of points which p moves, $\mathrm{supp}(p) = \{x_i \,|\, p(x_i) \neq x_i\}$. Two permutations p and q are *disjoint* if their supports are non-intersecting. If p and q are disjoint then they commute (with respect to the composition of permutations). Similarly, two sets of permutations P and Q are *support-disjoint* if every pair of permutations $p \in P$ and $q \in Q$ have disjoint supports.

Let $\mathscr{G}$ be a network with automorphism group $\mathrm{Aut}(\mathscr{G})$. Let $1 \notin S$ be a set of generators of $\mathrm{Aut}(\mathscr{G})$. Suppose that we partition S into n support-disjoint subsets $S = S_1 \cup \ldots \cup S_n$ such that each S_i cannot itself be decomposed into smaller support-disjoint subsets. Call H_i the subgroup generated by S_i. Since S is a generating set and elements from different factors H_i, H_j commute, this procedure gives a direct product decomposition:

$$\mathrm{Aut}(\mathscr{G}) = H_1 \times H_2 \times \ldots \times H_n. \tag{1}$$

Note that, in general, the choice of generators of a group is not unique and different choices of generating sets may give different decompositions. Thus, for this decomposition to be well-defined, we need to show that it is unique and the factors in Eq. (1) are 'irreducible'; that is, they cannot themselves be written as $K \times L$ with K and L support-disjoint subgroups.

A group G is *support-indecomposable* if it cannot be written as $K \times L$ with $K \neq 1$ and $L \neq 1$ support-disjoint subgroups. Similarly, a set S is support-indecomposable if it cannot be written as $S_1 \cup S_2$ with $S_1, S_2 \neq \emptyset$ both support-disjoint subsets.

Proposition 2.1. *The subgroups in Eq.(1) are independent of the choice of generators (that is, unique) and support-indecomposable (that is irreducible) when the generating set S satisfies the following two conditions*

- $(*)$ *S does not contain elements in the form $s = gh$ with $g, h \neq 1$ and g, h support-disjoint;*
- $(**)$ *if a subset $S' \subset S$ generates a subgroup $H \leq G$ such that $H = H_1 \times H_2$ with H_1 and H_2 support-disjoint then there exits a partition $S' = S_1 \cup S_2$ such that S_i generates H_i.*

We say that a set of generators satisfying these two conditions is *essential.* Note that these conditions are ensured if, for example, the `nauty` algorithm is used to calculate the generators of $\mathrm{Aut}(\mathscr{G})$ (see parts (1) and (2) of Theorem 2.34 in [30]). The proof of proposition 2.1 can be considered in two parts: irreducibility and uniqueness.

Proposition 2.2. *(irreducibility) Let S be a finite set of permutations and H the group generated by S. If H is support-indecomposable as a group, then so is S as a set. The converse is also true when S satisfies $(*)$.*

Proof. The first claim is clear. For the converse, suppose that $S = \{s_1, \ldots, s_n\}$ is support-indecomposable as a set but $H = K \times L$ (K, L support-disjoint). Then $s_1 = kl$ for $k \in K$, $l \in L$. By condition $(*)$, $k = 1$ or $l = 1$, that is, $s_1 \in K$ or $s_1 \in L$, and similarly for $s_2, \ldots, s_n$. Thus $S = (S \cap K) \cup (S \cap L)$. Since S is support-indecomposable as a set, one of $S \cap K$ or $S \cap L$ is empty, that is, $S \subseteq K$ or $S \subseteq L$. Hence $H = K$ and $L = 1$ or $H = L$ and $K = 1$. $\square$

Proposition 2.3. *(uniqueness) Suppose that X and Y are two sets of generators of a permutation group G, with associated direct product decompositions*

$$\begin{aligned} G &= H_1 \times \ldots \times H_n, \\ G &= K_1 \times \ldots \times K_m\,. \end{aligned}$$

If both X and Y are essential, then $n = m$ and there is a permutation σ of the factors such that $H_i \cong K_{\sigma(i)}$ for $i = 1, \ldots, n$.

Proof. (*sketch*) Firstly, generalize condition ($**$) to a finite number of subgroups $H_1, \ldots, H_n$, by induction on n. Then apply this to the first set of generators X with respect to the second decomposition. We then have a partition $X = X_1 \cup \ldots \cup X_m$ such that X_i generates K_i ($1 \leq i \leq m$). Suppose that H_1 is generated by a set $\{x_1, \ldots, x_t\} \subseteq X$. Since H_1 and $X_1, \ldots, X_m$ are support-indecomposable, we must have $\{x_1, \ldots, x_t\} \subseteq X_{i_1}$ for some i_1. That is, $H_1 \subset K_{i_1}$. Since K_{i_1} is support-indecomposable this implies $H_1 = K_{i_1}$. The same argument applies for $H_2, \ldots, H_n$. $\square$

Thus, the decomposition given in Eq. (1) is well-defined if (for example) the `nauty` algorithm is used. We shall refer to this decomposition as the *geometric* decomposition, and note that it is a simple variation of the Krull-Schmidt factorization into the direct product of indecomposable subgroups[37]. A GAP[19] procedure which calculates the geometric decomposition for an arbitrary permutation group is available from the authors on request.

In general the geometric decomposition is coarser than the Krull-Schmidt decomposition since non-disjoint permutations may still commute. The Krull-Schmidt decomposition may easily be obtained from the geometric decomposition using a computational group theory package such as GAP. The main advantage to using the geometric decomposition is that it provides a computationally efficient way to calculate the structure of large real-world networks and relates more intuitively to graph topology than the Krull-Schmidt factorization. For all the real-world networks we considered the automorphism group was factorized efficiently using this method, while a direct 'brute-force' factorization was not computationally feasible.

The geometric decompositions of some representative real-world networks are given in Table 2. In all cases the geometric factors are either symmetric groups or wreath products of symmetric groups (wreath products are a mild generalization of direct products, see [37] for a definition and examples).

Remark: it is a result of Pólya that automorphism groups of trees belong to the class of permutation groups which contains the symmetric groups and is closed under taking direct and wreath products[9]. Thus, the automorphism groups of many real-world networks belong to the same class of groups as the automorphism groups of trees. Note however, this does not necessarily mean that real-world symmetry is tree-like (for example, the complete graphs also belong to this class). In the following section we relate automorphism group structure to network topology in order to determine the extent to which real-world symmetry is, in fact, tree-like.

3 Automorphism Group Structure and Symmetric Subgraphs

The *induced subgraph* on a set of vertices $S \in \mathscr{G}$ is the graph obtained by taking S and any edges whose end points are both in S. We define a *symmetric subgraph* as the induced subgraph on the support of a geometric factor H (that is, on the points with non-trivial action by H). It is natural to ask whether there are any properties of symmetric subgraphs which are generic.

From Table 2 it is clear that most of the geometric factors found in real-world networks are isomorphic to

Network	Aut($\mathscr{G}$)
Biological Networks	
BCell	$C_2^{36} \times S_3^2 \times S_4$
Cele	$C_2^{95} \times S_3^{27} \times S_4^9 \times S_5^5 \times S_6^5 \times S_7^2 \times S_8 \times S_9^2 \times S_{10} \times S_{11} \times S_{33} \times (C_2 \wr C_2)$
BGHum	$C_2^{286} \times S_3^{80} \times S_4^{30} \times S_5^{14} \times S_6^{10} \times S_7 \times S_8^2 \times S_9^5 \times S_{10} \times S_{11} \times S_{12}^3 \times S_{15} \times S_{16}^2 \times S_{17} \times S_{23} \times S_{26} \times S_{44}$
BGScerev	$C_2^{42} \times S_3^8 \times S_4^5 \times S_5^2 \times S_6^2 \times S_7 \times S_{14} \times S_{17}$
BGDros	$C_2^{289} \times S_3^{86} \times S_4^{35} \times S_5^{19} \times S_6^{11} \times S_7^{10} \times S_8^5 \times S_9^3 \times S_{10}^3 \times S_{11}^3 \times S_{12}^3 \times S_{14} \times S_{16} \times S_{20} \times S_{30}$
BGMus	$C_2^7 \times S_3^4 \times S_4 \times S_5 \times S_6^3 \times S_8 \times S_{10} \times S_{11} \times S_{12} \times S_{26} \times S_{44}$
Yeast	$C_2^{90} \times S_3^{26} \times S_4^{16} \times S_5^8 \times S_6^6 \times S_7^5 \times S_8^2 \times S_9^2 \times S_{10}^3 \times S_{11}^2 \times S_{12} \times S_{13} \times S_{46} \times (C_2 \wr C_2)$
CeleMet	$C_2^{23} \times S_3^2 \times (C_2 \wr C_2)^2$
Technological Networks	
IntAS	$C_2^{955} \times S_3^{352} \times S_4^{197} \times S_5^{120} \times S_6^{83} \times S_7^{56} \times S_8^{55} \times S_9^{47} \times S_{10}^{32} \times S_{11}^{24} \times S_{12}^{14} \times S_{13}^{13} \times S_{14}^{13} \times S_{15}^8 \times S_{16}^9 \times S_{17}^7 \times S_{18}^3 \times S_{19}^{12} \times S_{20}^7 \times S_{21}^{10} \times S_{22} \times S_{23} \times S_{24}^6 \times S_{25} \times S_{26}^3 \times S_{27}^3 \times S_{28} \times S_{29}^2 \times S_{30}^2 \times S_{31}^2 \times S_{32}^2 \times S_{33} \times S_{34}^4 \times S_{35} \times S_{36} \times S_{37} \times S_{38}^2 \times S_{41} \times S_{42} \times S_{43} \times S_{46} \times S_{48}^2 \times S_{50} \times S_{51} \times S_{52}^2 \times S_{54} \times S_{56} \times S_{58} \times S_{59} \times S_{60} \times S_{62} \times S_{64} \times S_{70} \times S_{71} \times S_{76} \times S_{79} \times S_{82} \times S_{95} \times S_{112} \times S_{137} \times S_{138} \times S_{147} \times S_{167} \times S_{170} \times S_{194} \times S_{202} \times S_{216} \times S_{276} \times S_{318} \times S_{356} \times (C_2 \wr C_2)^2$
USPow	$C_2^{228} \times S_3^{44} \times S_4^{14} \times S_5^4 \times S_6^2 \times S_7 \times S_9 \times (C_2 \wr C_2)^8$
USAir	$C_2^{21} \times S_3^5 \times S_4^4 \times S_{12}$
Calif	$C_2^{276} \times S_3^{74} \times S_4^{40} \times S_5^{30} \times S_6^{21} \times S_7^{18} \times S_8^{19} \times S_9^{10} \times S_{10}^{10} \times S_{11}^6 \times S_{12}^4 \times S_{13}^4 \times S_{14}^5 \times S_{15}^3 \times S_{16}^4 \times S_{17}^2 \times S_{18}^3 \times S_{19}^2 \times S_{20}^2 \times S_{21} \times S_{23} \times S_{26} \times S_{31} \times S_{43} \times S_{46}$
EPA	$C_2^{74} \times S_3^{20} \times S_4^{20} \times S_5^{10} \times S_6^7 \times S_7^3 \times S_8 \times S_9^6 \times S_{10}^6 \times S_{11}^6 \times S_{12} \times S_{13}^4 \times S_{14} \times S_{15}^2 \times S_{16}^4 \times S_{17}^2 \times S_{18}^2 \times S_{19} \times S_{20} \times S_{21} \times S_{25}^2 \times S_{26} \times S_{27}^3 \times S_{30} \times S_{36} \times S_{37} \times S_{42} \times S_{44} \times S_{45} \times S_{48} \times S_{49} \times S_{51} \times S_{56} \times S_{57}^2 \times S_{61} \times S_{62} \times S_{76} \times S_{103}^2 \times S_{106} \times S_{115} \times (C_2 \wr C_2)^4$
PolBlog	$C_2^{13} \times S_3^2 \times S_4^2 \times S_5^2 \times S_8 \times S_{20}$
Social Networks	
Email	$C_2^{15} \times S_3^6$
PGP	$C2^{905} \times S_3^{247} \times S_4^{98} \times S_5^{46} \times S_6^{28} \times S_7^{13} \times S_8^9 \times S_9^5 \times S_{10}^3 \times S_{11}^2 \times S_{12}^2 \times S_{13} \times S_{14}^2 \times S_{15}^2 \times S_{16}^2 \times S_{18}^2 \times S_{27} \times S_{28} \times S_{31} \times S_{32} \times S_{40} \times (C_2 \wr C_2)^{16} \times (C_2 \wr S_5) \times (S_3 \wr C_2) \times (C_2^4 \wr C_2)$
Media	$C_2^{81} \times S_3^{41} \times S_4^{25} \times S_5^8 \times S_6^{13} \times S_7^9 \times S_8^8 \times S_9^3 \times S_{10}^5 \times S_{11}^2 \times S_{12}^2 \times S_{13}^2 \times S_{14}^4 \times S_{15}^6 \times S_{16}^4 \times S_{17} \times S_{18}^3 \times S_{19}^2 \times S_{20} \times S_{21}^5 \times S_{22}^3 \times S_{23} \times S_{24} \times S_{26} \times S_{28} \times S_{29} \times S_{30} \times S_{31}^2 \times S_{35} \times S_{39}^2 \times S_{42} \times S_{47} \times S_{52} \times S_{54} \times S_{55} \times S_{56} \times S_{63} \times S_{69} \times S_{72} \times S_{75} \times S_{84} \times S_{86} \times S_{91} \times S_{117} \times S_{122} \times S_{132} \times S_{145} \times S_{152} \times S_{545} \times (S_3 \wr C_2) \times (S_4 \wr C_2) \times (S_6 \wr C_2)$
Geom	$C_2^{370} \times S_3^{91} \times S_4^{25} \times S_5^{12} \times S_6^5 \times S_7 \times S_8^3 \times S_9 \times S_{10}^3 \times S_{13} \times (C_2 \wr C_2)^{11} \times (S_3 \wr C_2)$
Erdös	$C_2^{135} \times S_3^{67} \times S_4^{32} \times S_5^{48} \times S_6^{29} \times S_7^{26} \times S_8^{19} \times S_9^{20} \times S_{10}^{16} \times S_{11}^{19} \times S_{12}^{16} \times S_{13}^8 \times S_{14}^7 \times S_{15}^{11} \times S_{16}^8 \times S_{17}^9 \times S_{18}^5 \times S_{19}^8 \times S_{20}^4 \times S_{21}^4 \times S_{22}^6 \times S_{23}^2 \times S_{24}^3 \times S_{25}^4 \times S_{26}^2 \times S_{27}^4 \times S_{28}^5 \times S_{29}^2 \times S_{30}^3 \times S_{31}^5 \times S_{32} \times S_{33} \times S_{35}^2 \times S_{36}^2 \times S_{37} \times S_{41} \times S_{42}^3 \times S_{47} \times S_{51} \times S_{52} \times S_{53} \times S_{54} \times S_{57} \times S_{58} \times S_{86} \times S_{142}$
PhD	$C_2^{43} \times S_3^{27} \times S_4^{16} \times S_5^{11} \times S_6^{10} \times S_7^4 \times S_8^5 \times S_9^6 \times S_{10} \times S_{11}^3 \times S_{12}^2 \times S_{13}^2 \times S_{35} \times (C_2 \wr C_2)^3 \times (S_5 \wr C_2)$

Table 2: **The geometric decomposition of the automorphism group of some representative real-world networks**. In all cases, the automorphism group can be decomposed into direct and wreath products of symmetric groups.

S_n, the symmetric group on n letters (for some n). Furthermore, almost all of these symmetric factors act transitively on their supports. We shall refer to transitive symmetric factors as *basic factors* and associated symmetric subgraphs as *basic symmetric subgraphs* (BSSs). We shall refer to all other factors as *complex factors* and their associated symmetric subgraphs as *complex symmetric subgraphs*. Table 1 shows that, in all the wide range of representative cases we considered, almost all factors are basic and therefore that almost all symmetry is due to the presence of basic symmetric subgraphs.

Since a graph $\mathscr{G}$ on n vertices with $\mathrm{Aut}(\mathscr{G}) \cong S_n$ is either empty or complete[29] it is immediate that BSSs are also either empty or complete. Furthermore, transitivity ensures that for a given BSS $\mathscr{B}$ and a given vertex $v \in \mathscr{G} - \mathscr{B}$, *all* vertices in $\mathscr{B}$ are adjacent to v or *none* are. This means that most real-world symmetry is due to the presence of symmetric *cliques* (complete subgraphs invariant under $\mathrm{Aut}(\mathscr{G})$) and symmetric *bicliques* (complete bipartite subgraphs invariant under $\mathrm{Aut}(\mathscr{G})$).

In practice, for all the real-world networks we considered, bicliques other than *stars* (a k-star is a subgraph consisting of a vertex of degree $> k$ adjacent to k vertices of degree 1), although occasionally present, were rare (see Fig. 2 for some examples). In fact, we found that stars were the predominant symmetry structure present in all the networks we considered, although symmetric cliques were also significantly present in a number of networks. For example, the *c. elegans* genetic regulatory network[46] – which was constructed by inferring connections from multiple datasets across multiple organisms and is thus arguably one of the most well-characterized biological networks available – contains multiple symmetric cliques, including one on 33 vertices corresponding to the largest geometric subgroup in the decomposition of its automorphism group. This example (and those in Fig. 2) illustrate the fact that although much real-world symmetry is tree-like (and thus can be related to generic growth processes) a certain degree is not. In particular, a significant proportion of real-world symmetry originates in symmetric cliques. Since cliques and bicliques are topologically very similar (they are both complete multipartite graphs), the presence of symmetric cliques in complex networks may derive from similar growth processes to those that produce stars in combination with local clustering.

Fig. 1 gives a typical arrangement of symmetric subgraphs (basic and complex) found in many real world networks, illustrating the relationship between these symmetric subgraphs and the structure of the network automorphism group. Since complex symmetric subgraphs can potentially take *any* form it is not possible to say anything more general about their structure. However, since they are rare they may be considered on a case-by-case basis. Fig. 2 shows the complex symmetric subgraphs present in the US power grid, illustrating that in some real-world networks a certain degree of complex symmetry is present.

Note: in this study we focus on automorphism groups of undirected networks. In the case that a network is directed, its automorphism group is necessarily a subgroup of that of the underlying undirected network. Although this subgroup may be trivial, we expect that generally some of the underlying symmetry structure is carried to the directed network. For example directed-stars are commonly present in generic regulatory networks. Here the protein product of a central hub gene regulates a group of downstream target genes of in-degree 1; however the protein products of the target genes do not regulate the hub's gene expression (as is the case for an undirected star). Since directed-stars have the same symmetry group of their undirected counterparts, we expect that a certain degree of 'star-like' symmetry is also present in such directed networks.

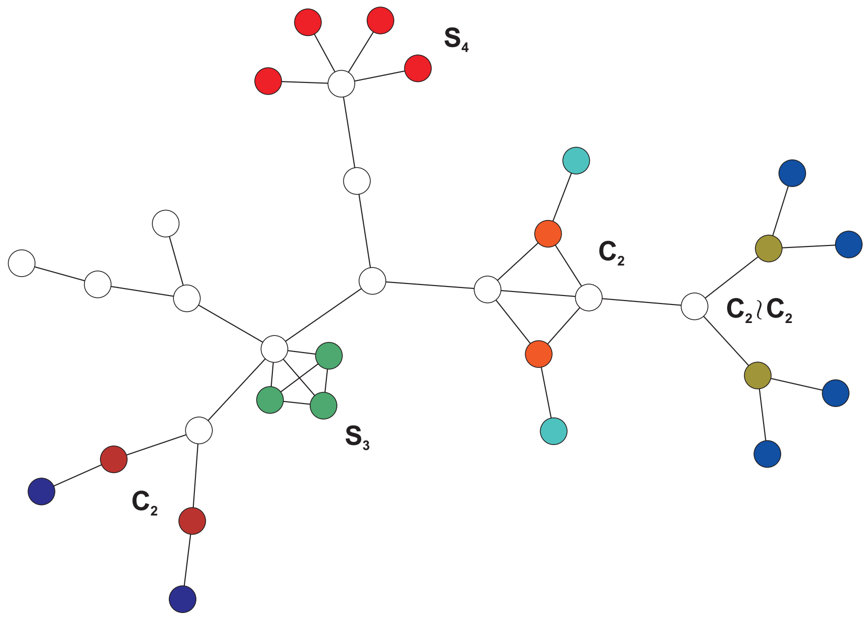

Figure 1: **A typical arrangement of symmetric subgraphs**. The geometric decomposition of the automorphism group of this graph is $\mathrm{Aut}(\mathscr{G}) = C_2^2 \times S_3 \times S_4 \times (C_2 \wr C_2)$. This example illustrates how different symmetric subgraphs contribute to the automorphism group, as well as showing common 'non-treelike' real-world symmetry. In particular note the 4-star (red) and the 3-clique (green) which correspond to the factors S_4 and S_3 respectively in the geometric decomposition of $\mathrm{Aut}(\mathscr{G})$. We found that wreath product factors generally associate with extended branches (see the far right of this figure), although this is not always the case (see the starred subgraph in Fig. 2 for example). Vertices are colored by orbit, fixed points are in white.

4 Symmetry and Redundancy

Tolerance to attack is of crucial importance to the effective functioning of many networks. Consequently, some considerable attention has been paid in the literature to understanding network *robustness*: the ability of a network to tolerate vertex deletions and still function effectively[2, 3]. Redundancy naturally reinforces against attack by providing functional 'backups' should network elements fail[44]. Thus network robustness is related to network redundancy. Since network automorphisms permute vertices without altering vertex adjacency, symmetric networks necessarily contain a certain degree of structural redundancy, and automorphism group structure may be used to precisely quantify this redundancy.

The *orbit* of a vertex $v \in V(\mathscr{G})$ is the set $\Delta(v) = \{\pi v \ \in V(\mathscr{G}) : \pi \in \mathrm{Aut}(\mathscr{G})\}$[13]. Automorphism group orbits naturally partition network vertices into disjoint structural equivalence classes. Since two vertices in the same orbit may be permuted without altering network adjacency they are structurally equivalent in the strongest possible sense: they play exactly the same structural role in the network. Thus, nontrivial orbits are associated with structural redundancy, while trivial orbits are associated with structurally unique elements. A network's orbit structure may therefore be used to assess the degree to which it is constructed from structurally unique elements, and the degree to which it is constructed from repetitions of structurally equivalent elements. In particular, we may quantify network redundancy by calculating the ratio:

$$r_{\mathscr{G}} = \frac{N_{\mathscr{Q}} - 1}{N_{\mathscr{G}}},$$

where $N_{\mathscr{Q}}$ is the number of network orbits and $N_{\mathscr{G}}$ is the number of vertices in the network.

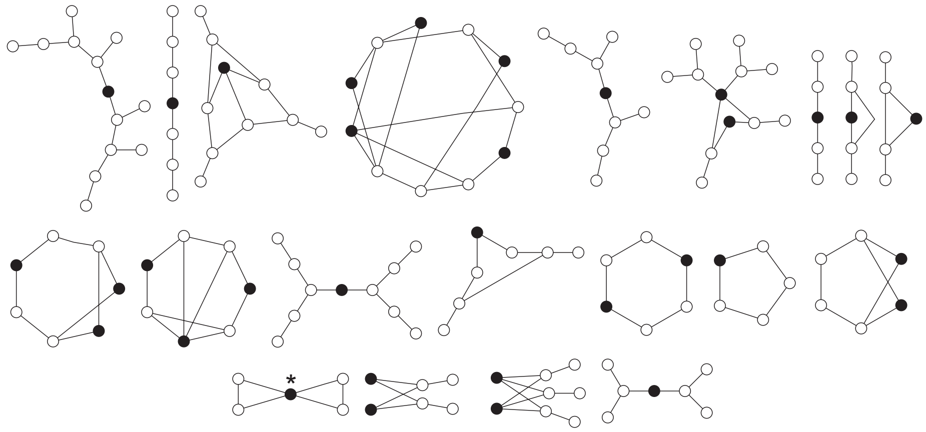

Figure 2: **Complex symmetric subgraphs in the US power grid.** Vertices in white correspond to those in the symmetric subgraphs. Vertices in black are those adjacent to those in the symmetric subgraph, and are shown to clarify subgraph structure. The starred subgraph has automorphism group $C_2 \wr C_2$, illustrating that wreath products do not associate exclusively with extended branches such as in the example in Fig. 1

Networks which have a transitive automorphism group (and therefore possess only one orbit with all vertices playing the same structural role) have $r_{\mathscr{G}} = 0$; while asymmetric networks (which have a trivial automorphism group and in which all vertices play a unique structural role) have $r_{\mathscr{G}} = 1 - 1/N_{\mathscr{G}}$. Thus, $0 \leq r_{\mathscr{G}} < 1$: the smaller the value of $r_{\mathscr{G}}$ the more the network is constructed from repetition of structurally identical elements; while the larger the value of $r_{\mathscr{G}}$ the more the network is constructed from structurally unique elements. Table 1 gives $r_{\mathscr{G}}$ for a representative selection real-world networks. It is clear that while some real-world networks are primarily constructed from unique structural elements, others contain a significant amount of structural redundancy.

5 Conclusions

We have considered the automorphism groups of a wide variety of real-world networks and have found that a certain degree of symmetry is ubiquitous. We have constructed a practical network automorphism group decomposition (the geometric decomposition), and found that automorphism groups of real-world networks can typically be decomposed into direct and wreath products of symmetric groups. We have shown that each geometric factor can be associated with a symmetric subgraph, and demonstrated that most factors can be related to either a symmetric clique or symmetric biclique. Thus, we find that these two types of subgraph generically account for almost all real-world network symmetry. We have also considered the relationship between symmetry and redundancy and found that many real-world networks carry a significant amount of structural redundancy. Thus, we conclude that symmetry is present in many real-world empirical networks, it arises from natural growth processes, commonly has a simple generic form and can affect network properties such as robustness.

To conclude, we note that while symmetry is ubiquitous in many real-world complex systems, many networks also contain elements which are almost, but not exactly, symmetric. For example, in a biological context, growth with partial duplication of structural motifs[14] naturally gives rise to elements which are almost symmetric. Such 'near' symmetry is not captured in the network automorphism group, although it can have a profound effect on network structure and behavior. In order to investigate near symmetry, some authors have considered weaker descriptions of structural equivalence than that provided by the automorphism group. For example, in their consideration of networks of differential equations, Golubitsky, Stewart and co-workers weaken group axioms and consider network symmetry groupoids[20, 41, 42]. They show how symmetry groupoids can significantly effect the dynamics of coupled cell networks, giving rise to patterns of synchrony for instance[42].

The automorphism group approach we have taken in this paper has the advantage that it enables application of the powerful theory of groups to network analysis. We anticipate that further investigation of automorphism groups, groupoids, and alternative measures of structural equivalence in networks will enhance our understanding of the structure and function of complex systems.

6 Acknowledgments

This work was funded by the EPSRC and by a London Mathematical Society scheme 6 grant.

References

[1] L. A. Adamic and N. Glance, *The political blogosphere and the 2004 US election: divided they blog*, Proceedings of the 3rd international workshop on Link discovery (2005), 36–43.

[2] R. Albert and A.-L. Barabási, *Statistical mechanics of complex networks*, Rev. Modern Phys. **74** (2002), no. 1, 47–97.

[3] R. Albert, H. Jeong, and A.-L. Barabási, *Error and attack tolerance of complex networks*, Nature **406** (2000), no. 6794, 378–382.

[4] A.-L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science **286** (1999), no. 5439, 509–512.

[5] K. Basso, A. A. Margolin, G. Stolovitzky, U. Klein, R. Dalla-Favera, and A. Califano, *Reverse engineering of regulatory networks in human b cells*, Nat. Genet. **37** (2005), no. 4, 382–390.

[6] V. Batagelj and A. Mrvar, *Pajek - program for large network analysis*, Connections **21** (1998), no. 2, 47–57.

[7] ———, *Some analyses of Erdös collaboration graph*, Social Networks **22** (2000), no. 2, 173–186.

[8] V. Batagelj and M. Zaveršnik, *The collaboration network in computational geometry*, `http://vlado.fmf.uni-lj.si/pub/networks/data/collab/geom.htm`.

[9] N. Biggs, *Algebraic graph theory*, Cambridge University Press, London, 1974.

[10] M. Boguñá, R. Pastor-Satorras, A. Díaz-Guilera, and A. Arenas, *Models of social networks based on social distance attachment*, Phys. Rev. E **70** (2004), no. 5, 056122.

[11] B. Bollobás, *Modern graph theory*, Graduate Texts in Mathematics, vol. 184, Springer-Verlag, New York, 1998.

[12] ———, *Random graphs*, second ed., Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001.

[13] P. J. Cameron, *Permutation groups*, London Mathematical Society Student Texts, vol. 45, Cambridge University Press, Cambridge, 1999.

[14] F. Chung, L. Lu, T. G. Dewey, and D. J. Galas, *Duplication Models for Biological Networks*, J. Comput. Biol. **10** (2003), no. 5, 677–687.

[15] W. de Nooy, A. Mrvar, and V. Batagelj, *Exploratory Social Network Analysis with Pajek*, Cambridge University Press, 2005.

[16] J. Duch and A. Arenas, *Community detection in complex networks using extremal optimization*, Phys. Rev. E **72** (2005), 027104.

[17] Y. Q. Feng, *Automorphism groups of Cayley graphs on symmetric groups with generating transposition sets*, J. Combinator. Theor. B **96** (2006), no. 1, 67–72.

[18] P. Foggia, C. Sansone, and M. Vento, *A performance comparison of five algorithms for graph isomorphism*, Proc. of the 3rd IAPR TC-15 Workshop on Graph-based Representations in Pattern Recognition (2001), 188–199.

[19] The GAP Group, *Gap – groups, algorithms, and programming, version 4.4*, 2006, `http://www.gap-system.org`.

[20] M. Golubitsky and I. Stewart, *Nonlinear dynamics of networks: the groupiod formalism*, Bul. Amer. Math. Soc. **43** (2006), no. 3, 305–364.

[21] The CAIDA Group, *The CAIDA AS relationships dataset*, 19/06/2006, `http://www.caida.org/data/active/as-relationships`.

[22] R. Guimerà, L. Danon, A. Díaz-Guilera, F. Giralt, and A. Arenas, *Self-similar community structure in a network of human interactions*, Phys. Rev. E **68** (2003), no. 6, 065103.

[23] F. Harary and E. M. Palmer, *The probability that a point of a tree is fixed*, Math. Proc. Cambridge Philos. Soc. **85** (1979), no. 3, 407–415.

[24] H. Jeong, S. P. Mason, A.-L. Barabási, and Z. N. Oltvai, *Lethality and centrality in protein networks*, Nature **411** (2001), 41–42.

[25] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabási, *The large-scale organization of metabolic networks*, Nature **407** (2000), 651–654.

[26] J. M. Kleinberg, *Pages linking to epa.gov*, `http://www.cs.cornell.edu/courses/cs685/2002fa`.

[27] ———, *Pages matching the query 'california'*, `http://www.cs.cornell.edu/courses/cs685/2002fa`.

[28] S. Lakshmivarahan, J. S. Jwo, and S. K. Dhall, *Symmetry in interconnection networks based on Cayley graphs of permutation groups: a survey.*, Parallel Computing **19** (1993), no. 4, 361–407.

[29] J. Lauri and R. Scapellato, *Topics in graph automorphisms and reconstruction*, London Mathematical Society Student Texts, vol. 54, Cambridge University Press, Cambridge, 2003.

[30] B. D. McKay, *Practical graph isomorphism*, Congr. Numer. **30** (1981), 45–87.

[31] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, *Network motifs: Simple building blocks of complex networks*, Science **298** (2002), no. 5594, 824–827.

[32] M. E. J. Newman, *Scientific collaboration networks I: Network construction and fundamental results*, Phys. Rev. E **64** (2001), no. 1, 016131.

[33] ______, *Scientific collaboration networks. II: Shortest paths, weighted networks, and centrality*, Phys. Rev. E **64** (2001), no. 1, 016132.

[34] ______, *The structure and function of complex networks*, SIAM Rev. **45** (2003), no. 2, 167–256.

[35] K. Norlen, G. Lucas, M. Gebbie, and J. Chuang, *EVA: Extraction, Visualization and Analysis of the Telecommunications and Media Ownership Network*, Proceedings of International Telecommunications Society 14th Biennial Conference (ITS2002) (2002).

[36] Bureau of Transportation Statistics, *North american transportation atlas data*, 1997, `http://www.bts.gov`.

[37] J. J. Rotman, *An introduction to the theory of groups*, fourth ed., Graduate Texts in Mathematics, vol. 148, Springer-Verlag, New York, 1994.

[38] SIGACT, *The theoretical computer science genealogy*, `http://sigact.acm.org/genealogy`.

[39] G. Siganos, M. Faloutsos, P. Faloutsos, and C. Faloutsos, *Power laws and the as-level internet topology*, IEEE/ACM Transactions on Networking **11** (2003), no. 4, 514–524.

[40] C. Stark, B.-J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz, and M. Tyers, *BioGRID: a general repository for interaction datasets*, Nucleic Acids Res. **34** (2006), D535–D539.

[41] I. Stewart, *Networking opportunity*, Nature **427** (2004), 601–604.

[42] I. Stewart, M. Martin Golubitsky, and M. Pivato, *Symmetry Groupoids and Patterns of Synchrony in Coupled Cell Networks.*, SIAM J. Applied Dynamical Systems **2** (2003), no. 4, 609–646.

[43] S. H. Strogatz, *Exploring complex networks*, Nature **410** (2001), 268–276.

[44] G. Tononi, O. Sporns, and G. M. Edelman, *Measures of degeneracy and redundancy in biological networks*, Proc. Natl. Acad. Sci. USA **96** (1999), no. 6, 3257–3262.

[45] D. J. Watts and S. H. Strogatz, *Collective dynamics of 'small-world' networks*, Nature **393** (1998), no. 6684, 440–442.

[46] W. Zhong and P. W. Sternberg, *Genome-wide prediction of c. elegans genetic interactions*, Science **311** (2006), no. 5766, 1481–1484.