Overrepresented and Underrepresented Patterns in System Architectures Across Diverse Engineering Systems

A. S. Shaja

Doctoral Student, Department of Aerospace Engineering, IIT Bombay, Mumbai, India - 400076,

Email: shaja@aero.iitb.ac.in

K. Sudhakar

Professor, Department of Aerospace Engineering, IIT Bombay, Mumbai, India - 400076,

Email: sudhakar@aero.iitb.ac.in

Abstract

Recent years have witnessed new research interest in the study of complex systems architectures, in domains like biological systems, social networks etc. Seminal works covering each of these systems have appeared in high impact journals like Nature, Science, etc. Unifying principles have emerged and helped in gaining new understanding in a domain by extending the understanding gained in other domains. These developments in complex systems open up possibilities in the research into architectures of complex engineering systems. Complex engineering systems are synthesized from large number of components giving it a physical architecture. We abstract the physical architecture of different engineering systems as a network/graph, where the nodes/vertices correspond to components and edges correspond to interconnections between them. Complex systems research in biology defines motifs as recurring sub-graphs from which the network is built. They also argue motifs as simple building blocks of complex networks, offering a way to understand the basic functionality of a system. In this paper, we explore 32 arbitrarily chosen engineering systems architectures for motifs. We discover motifs within each system and also interesting motif templates across systems.

Keywords

Systems architecture, motifs.

1. Introduction

Recent years have witnessed new research interest in the study of complex systems architectures, in domains like biological systems, social networks etc. [Duncan J Watts , Newman MEJ]. Unifying principles have emerged and helped in gaining new understanding in a domain by extending the understanding gained in other domains [Boccaletti S et all]. Researchers in other areas have commented on the hesitation of researchers in complex engineering systems to look at their problems in the light of emerging ideas in complex systems in general. "Engineering should be at the centre of these developments, and contribute to the development of new theory and tools" [J.M. Ottino]; "Engineers seem a little bit indifferent as if engineering is at the edge of the science of complexity" [Zhi-Qiang Jiang at all].

Architecture is the fundamental structure of components of a system - the roles they play, and how they are related to each other and to their environment [ANSI IEEE Standard 1471]. The dictionary definition of complexity refers to – consisting of interconnected/interwoven components. Complexity of a system scales with the number of components, number of interactions, complexities of the components & complexities of interactions [Edward Crawley et all]. Complex engineering systems are synthesized from large number of components coupled to each other giving it a physical architecture; they are evolved through a design process that is best represented by large number of connected tasks giving it a technical architecture; and they are evolved by collaborating groups of people giving it an organizational architecture. These architectural views pose interesting possibilities in respect of searching for new understanding in complex engineering systems. Product architectures are considered complex systems [Tyson R Browning]. Architecture of a system can be abstracted as a network/graph, where the nodes/vertices correspond to components in the system and edges correspond to interconnection between them.

Complex systems research in biology defines motifs as recurring sub-graphs from which the network is built. In biology, the analysis of network motifs has led to interesting insights in the areas of protein-protein interaction prediction [Albert L and Albert R] and analysis of temporal gene expression patterns [M Ronen et all , S.S. Shen-Orr et all]. Research in biology also argues motifs as simple building blocks of complex networks whose selection may possibly be one way to understand the basic functionality of a system.

2. Motifs

Motifs are considered to be functional building blocks of a network. “Motifs are recurring sub-graphs of interactions from which the networks are built” [Milo R et all]. These are patterns of interconnections occurring in real networks in numbers that are considered significant. . Motifs can be of any size from n=2 to N-1, where N is the total number of nodes in the network. Let us consider a directed network with N nodes and look for motifs of size n=3. There are ${}^{N}C_3$ different combinations of triplets of nodes in an N-noded network. Some triplets out of ${}^{N}C_3$ need not form a connected graph, and are not sub-graphs (an example is when out of 3 nodes 2 nodes are connected to each other and the third does not have an edge with the first two). A connected triplet is a 3-noded sub-graph. For a 3-noded sub-graph there are 13 patterns possible as shown in Fig 3.1.

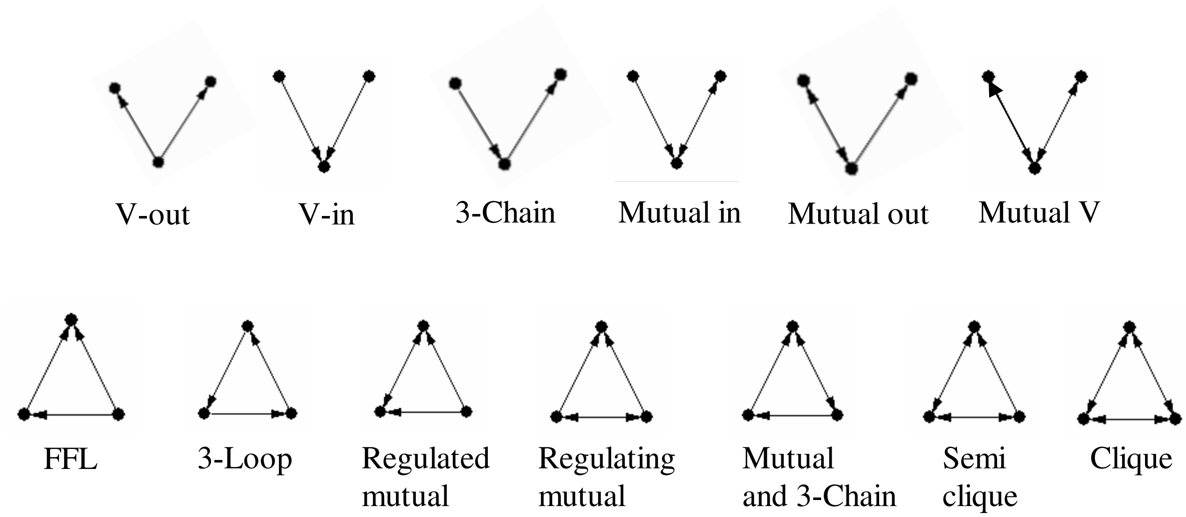

Fig 3.1

Each of the ${}^{N}C_3$ triplets, if it is a sub-graph, will assume one of the 13 patterns. One can count the occurrence of each pattern for all ${}^{N}C_3$ triplets and define a vector, $\mathbf{P}_{real}$, of size 13. In a network the count for a particular pattern may be high, which by itself is not considered important. It is possible that such high count for that pattern is unavoidable for a network synthesized using the N nodes that preserve the degree distribution of the real network. To investigate this, randomized networks are created [Milo R et all] using same N nodes, ie. number of nodes and their degree distribution is preserved. Each randomized network defines a pattern count vector, $\mathbf{P}_{rand\text{-}i}$. Large number of randomized networks (i=1 to m) will define a vector of mean, $\boldsymbol{\mu}_{rand}$ and a vector of standard deviation, $\boldsymbol{\sigma}_{rand}$, of 13 patterns. For the real network we can check the significance of j^{th} pattern by, $S_j = (P_{real\text{-}j}$ -

$\mu_{rand-j})/\sigma_{rand-j}$ for j=1 to 13. For a normally distributed random number, value of S_j greater than 3 or less than 3 implies a rare occurrence (3σ limit). Any pattern with its $S_j > 2$ is considered a motif [Milo R et all], and is an over-represented pattern. Any pattern with its $S_j < -2$ is an anti-motif, and is an under-represented pattern.

2.1 Motif Significance Profile

S is a vector of size 13 that defines significance of 13 patterns in the real network. Milo R et all argue that **S** is influenced by the size of the network and propose normalization of **S** to make it largely independent of network size. Thus, significance profile vector, **Z** is defined as $Z_j = S_j / |\mathbf{S}|$. This makes comparison of networks of varying sizes possible.

2.2 Correlation of Motif Significance Profiles

[Milo R et all] have reported similarities in significant profiles of systems. They propose the standard correlation coefficients (Pearson correlation coefficient) between **Z** vectors of two systems as a measure of similarity between their significance profiles. The correlation coefficient can vary from -1 to +1. A value of +1 implies that the 13 patterns are present to the same extent in both systems, ie if a particular pattern is over-represented (under-represented) in one system it will be over-represented (under-represented) in the other system to the same extent. A value of -1 means that if a pattern is over-represented (under-represented) in one system the same will under-represented (over-represented) in the other system.

3. Engineering Systems

In this paper we consider 32 arbitrarily chosen engineering systems and study their architectures. Systems considered range from aircraft engine [Manuel E Sosa et all], softwares [Software graph data for specified software systems], electronic circuits [ISCAS High level models, ISCAS'89 benchmark data], robot [Amro M. Farid and Duncan C. McFarlane], refrigerator [Homas U. Pimmler and Steven D] etc. Table 3.1 briefly identifies each of the 32 systems. The 32 systems are of vastly different sizes (ranging from minimum 16 components to maximum 23843 components). We create 1000 random networks for each considered system using same N nodes, ie. number of nodes and their degree distribution is preserved. Adequacy of 1000 samples for estimating μ and σ of patterns is confirmed. For each real network we compute the significance of each of the 13 patterns of 3-noded sub-graphs, $S_j = (P_{real-j} - \mu_{rand-j})/\sigma_{rand-j}$; j=1 to 13. For example, the aircraft engine has **S** = [-5.02, -1.86, -10.04, -3.90, 0.72, 4.04, -11.49, -20.71, -0.23, 1.94, 5.34, 8.41, 16.86]. **S** vectors are in fact computed for 3-noded, 4-noded and 5-noded sub-graphs and results are available at our website [Shaja AS, Sudhakar K, CASMots:]. (It may be noted that the size of **S** vector for 4-noded is 199). Further study in this paper is restricted to 3-noded sub-graphs only. The significance profiles for all 32 systems are now computed as, $Zj = Sj / |S|$. For example, **Z** for the aircraft engine is [-0.099, -0.036, -0.198, -0.077, 0.014, 0.079, -0.227, -0.408, -0.004, 0.038, 0.105, 0.166, 0.332].

Similarities in significance profiles across all 32 systems are now investigated by computing correlation coefficient between each pair. This information is presented in Fig 3.2 as a square matrix of size 32. Diagonal elements of the matrix represent similarity of significance of profile of a system with itself and are always +1. Off diagonal elements can take values in the range -1 to +1. In Figure 3.2 the full range of values (-1 to +1) is grouped into 3 regions and indicated by 3 different colors for visual impact.

- Correlation coefficient of -0.65 to +0.65, Weak or no correlation, light green color
- Correlation coefficient of +0.65 to +1.00, Positively correlated, red color
- Correlation coefficient of -0.65 to -1.00, Negatively correlated, blue color

The matrix in Fig 3.2 appears checkered and it is not easy to discern similarities between groups of systems that are hidden within. We perform clustering using a standard clustering algorithm (partitioning around medoids algorithm [L Kaufman, P.J Rousseeuw, 1990, Finding Groups in Data: An introduction to cluster analysis]), where edge weights are clustering coefficients. The row-columns of the correlation matrix are now re-ordered based on that standard clustering algorithm and shown in Figure 3.3. Grouping of systems based on similarity of significant profiles is clearly visible now. There are 4 distinct groups as revealed by 4 red colored blocks along the diagonal. These 4 blocks contain systems whose significance profiles are all positively correlated with respect to each other. It is extremely interesting to note that the above grouping coincides with groupings of systems based on whether they are software, electrical or mechanical. Each of these 4 groups is discussed in the next section.

System no	System Name	System no	System Name	System no	System Name
S1	Apword	S12	Traffic control system (s382)	S23	Robot
S2	Linux	S13	ALU (c880)	S24	Vtk
S3	Aircraft Engine	S14	Digital Fractional Multiplier (s838)	S25	PLD (s832)
S4	ALU (c7552)	S15	ECAT (c1908)	S26	Digital Fractional Multiplier (s208)
S5	Forword logic chips (s13207)	S16	Traffic control system (s400)	S27	Forword logic chips (s15850)
S6	ALU (74181)	S17	PLD (s832)	S28	Mysql
S7	ECAT (c1355)	S18	Xmms	S29	Forword logic chips (s38417)
S8	Forword logic chips (s9234)	S19	Forword logic chips (s38584)	S30	PLD (s713)
S9	Refrigerator	S20	Digital Fractional Multiplier (s420)	S31	ALU (c3540)
S10	Traffic control system (s420)	S21	PLD (s641)	S32	ALU (c2670)
S11	ECAT (c499)	S22	Traffic control system (s562)		

Table 3.1

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20	S21	S22	S23	S24	S25	s26	S27	S28	S29	S30	S31	S32
S1	1.0	0.9	0.5	-0.8	0.8	-0.9	0.8	0.8	0.1	0.8	-0.8	-0.8	-0.8	0.3	-0.8	-0.8	-0.8	0.8	0.8	0.3	-0.8	-0.6	0.2	0.9	-0.8	0.2	0.8	0.9	0.8	0.8	-0.8	0.8
S2	0.9	1.0	0.2	-0.8	1.0	-1.0	0.9	0.9	0.0	0.9	-1.0	-1.0	-1.0	0.3	-0.9	-1.0	-1.0	1.0	1.0	0.2	-0.9	-0.8	0.0	1.0	-1.0	0.1	0.9	0.9	1.0	0.9	-0.9	0.9
S3	0.5	0.2	1.0	-0.1	0.2	-0.1	0.2	0.2	0.8	0.1	-0.1	-0.1	-0.1	0.0	-0.1	-0.1	-0.1	0.1	0.1	0.0	-0.1	-0.1	0.9	0.4	-0.1	0.0	0.2	0.3	0.2	0.1	-0.1	0.1
S4	-0.8	-0.8	-0.1	1.0	-0.7	0.9	-0.7	-0.7	0.0	-0.7	0.9	0.9	0.9	-0.8	0.9	0.9	0.9	-0.8	-0.8	-0.7	0.9	0.3	0.0	-0.8	0.9	-0.6	-0.7	-0.7	-0.7	-0.7	1.0	-0.6
S5	0.8	1.0	0.2	-0.7	1.0	-1.0	1.0	1.0	0.0	1.0	-1.0	-1.0	-0.9	0.1	-0.9	-1.0	-1.0	1.0	1.0	0.1	-0.9	-0.9	0.0	0.9	-1.0	0.0	1.0	0.9	1.0	1.0	-0.9	1.0
S6	-0.9	-1.0	-0.1	0.9	-1.0	1.0	-1.0	-1.0	0.0	-0.9	1.0	1.0	1.0	-0.4	1.0	1.0	1.0	-1.0	-1.0	-0.3	1.0	0.7	0.0	-0.9	1.0	-0.2	-1.0	-0.9	-1.0	-0.9	1.0	-0.9
S7	0.8	0.9	0.2	-0.7	1.0	-1.0	1.0	1.0	0.0	1.0	-1.0	-1.0	-0.9	0.1	-0.9	-0.9	-1.0	1.0	1.0	0.1	-0.9	-0.9	0.0	0.9	-1.0	0.0	1.0	0.9	1.0	1.0	-0.9	1.0
S8	0.8	0.9	0.2	-0.7	1.0	-1.0	1.0	1.0	0.0	1.0	-0.9	-1.0	-0.9	0.1	-0.9	-0.9	-1.0	1.0	1.0	0.1	-0.9	-0.9	0.0	0.9	-1.0	-0.1	1.0	0.9	1.0	1.0	-0.9	1.0
S9	0.1	0.0	0.8	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
S10	0.8	0.9	0.1	-0.7	1.0	-0.9	1.0	1.0	0.0	1.0	-0.9	-0.9	-0.9	0.0	-0.9	-0.9	-0.9	1.0	1.0	0.0	-0.9	-0.9	0.0	0.9	-0.9	-0.1	1.0	0.8	1.0	1.0	-0.8	1.0
S11	-0.8	-1.0	-0.1	0.9	-1.0	1.0	-1.0	-0.9	0.0	-0.9	1.0	1.0	1.0	-0.4	1.0	1.0	1.0	-1.0	-1.0	-0.4	1.0	0.7	0.0	-0.9	1.0	-0.3	-0.9	-0.9	-1.0	-0.9	1.0	-0.9
S12	-0.8	-1.0	-0.1	0.9	-1.0	1.0	-1.0	-1.0	0.0	-0.9	1.0	1.0	1.0	-0.4	1.0	1.0	1.0	-1.0	-1.0	-0.4	1.0	0.7	0.0	-0.9	1.0	-0.2	-1.0	-0.9	-1.0	-0.9	1.0	-0.9
S13	-0.8	-1.0	-0.1	0.9	-0.9	1.0	-0.9	-0.9	0.0	-0.9	1.0	1.0	1.0	-0.5	1.0	1.0	1.0	-1.0	-1.0	-0.4	1.0	0.7	0.0	-0.9	1.0	-0.3	-0.9	-0.9	-0.9	-0.9	1.0	-0.9
S14	0.3	0.3	0.0	-0.8	0.1	-0.4	0.1	0.1	0.0	0.0	-0.4	-0.4	-0.5	1.0	-0.5	-0.4	-0.3	0.3	0.2	1.0	-0.5	0.4	0.0	0.2	-0.4	1.0	0.1	0.2	0.1	0.0	-0.6	0.0
S15	-0.8	-0.9	-0.1	0.9	-0.9	1.0	-0.9	-0.9	0.0	-0.9	1.0	1.0	1.0	-0.5	1.0	1.0	1.0	-1.0	-1.0	-0.4	1.0	0.6	0.0	-0.9	1.0	-0.3	-0.9	-0.9	-0.9	-0.9	1.0	-0.9
S16	-0.8	-1.0	-0.1	0.9	-1.0	1.0	-0.9	-0.9	0.0	-0.9	1.0	1.0	1.0	-0.4	1.0	1.0	1.0	-1.0	-1.0	-0.4	1.0	0.7	0.0	-0.9	1.0	-0.3	-0.9	-0.9	-1.0	-0.9	1.0	-0.9
S17	-0.8	-1.0	-0.1	0.9	-1.0	1.0	-1.0	-1.0	0.0	-0.9	1.0	1.0	1.0	-0.3	1.0	1.0	1.0	-1.0	-1.0	-0.3	1.0	0.7	0.0	-0.9	1.0	-0.2	-1.0	-0.9	-1.0	-0.9	1.0	-0.9
S18	0.8	1.0	0.1	-0.8	1.0	-1.0	1.0	1.0	0.0	1.0	-1.0	-1.0	-1.0	0.3	-1.0	-1.0	-1.0	1.0	1.0	0.2	-0.9	-0.8	0.0	1.0	-1.0	0.1	1.0	0.9	1.0	1.0	-0.9	0.9
S19	0.8	1.0	0.1	-0.8	1.0	-1.0	1.0	1.0	0.0	1.0	-1.0	-1.0	-1.0	0.2	-1.0	-1.0	-1.0	1.0	1.0	0.1	-0.9	-0.8	0.0	1.0	-1.0	0.0	1.0	0.9	1.0	1.0	-0.9	1.0
S20	0.3	0.2	0.0	-0.7	0.1	-0.3	0.1	0.1	0.0	0.0	-0.4	-0.4	-0.4	1.0	-0.4	-0.4	-0.3	0.2	0.1	1.0	-0.5	0.4	0.0	0.2	-0.3	1.0	0.1	0.2	0.1	0.0	-0.5	-0.1
S21	-0.8	-0.9	-0.1	0.9	-0.9	1.0	-0.9	-0.9	0.0	-0.9	1.0	1.0	1.0	-0.5	1.0	1.0	1.0	-0.9	-0.9	-0.5	1.0	0.6	0.0	-0.9	1.0	-0.4	-0.9	-0.8	-0.9	-0.9	1.0	-0.8
S22	-0.6	-0.8	-0.1	0.3	-0.9	0.7	-0.9	-0.9	0.0	-0.9	0.7	0.7	0.7	0.4	0.6	0.7	0.7	-0.8	-0.8	0.4	0.6	1.0	0.0	-0.8	0.7	0.5	-0.9	-0.7	-0.9	-0.9	0.5	-0.9
S23	0.2	0.0	0.9	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
S24	0.9	1.0	0.4	-0.8	0.9	-0.9	0.9	0.9	0.1	0.9	-0.9	-0.9	-0.9	0.2	-0.9	-0.9	-0.9	1.0	1.0	0.2	-0.9	-0.8	0.1	1.0	-0.9	0.1	0.9	0.9	0.9	0.9	-0.9	0.9
S25	-0.8	-1.0	-0.1	0.9	-1.0	1.0	-1.0	-1.0	0.0	-0.9	1.0	1.0	1.0	-0.4	1.0	1.0	1.0	-1.0	-1.0	-0.3	1.0	0.7	0.0	-0.9	1.0	-0.2	-1.0	-0.9	-1.0	-0.9	1.0	-0.9
s26	0.2	0.1	0.0	-0.6	0.0	-0.2	0.0	-0.1	0.0	-0.1	-0.3	-0.2	-0.3	1.0	-0.3	-0.3	-0.2	0.1	0.0	1.0	-0.4	0.5	0.0	0.1	-0.2	1.0	0.0	0.1	0.0	-0.1	-0.4	-0.2
S27	0.8	0.9	0.2	-0.7	1.0	-1.0	1.0	1.0	0.0	1.0	-0.9	-1.0	-0.9	0.1	-0.9	-0.9	-1.0	1.0	1.0	0.1	-0.9	-0.9	0.0	0.9	-1.0	0.0	1.0	0.9	1.0	1.0	-0.9	1.0
S28	0.9	0.9	0.3	-0.7	0.9	-0.9	0.9	0.9	0.0	0.8	-0.9	-0.9	-0.9	0.2	-0.9	-0.9	-0.9	0.9	0.9	0.2	-0.8	-0.7	0.0	0.9	-0.9	0.1	0.9	1.0	0.9	0.8	-0.8	0.8
S29	0.8	1.0	0.2	-0.7	1.0	-1.0	1.0	1.0	0.0	1.0	-1.0	-1.0	-0.9	0.1	-0.9	-1.0	-1.0	1.0	1.0	0.1	-0.9	-0.9	0.0	0.9	-1.0	0.0	1.0	0.9	1.0	1.0	-0.9	1.0
S30	0.8	0.9	0.1	-0.7	1.0	-0.9	1.0	1.0	0.0	1.0	-0.9	-0.9	-0.9	0.0	-0.9	-0.9	-0.9	1.0	1.0	0.0	-0.9	-0.9	0.0	0.9	-0.9	-0.1	1.0	0.8	1.0	1.0	-0.8	1.0
S31	-0.8	-0.9	-0.1	1.0	-0.9	1.0	-0.9	-0.9	0.0	-0.8	1.0	1.0	1.0	-0.6	1.0	1.0	1.0	-0.9	-0.9	-0.5	1.0	0.5	0.0	-0.9	1.0	-0.4	-0.9	-0.8	-0.9	-0.8	1.0	-0.8
S32	0.8	0.9	0.1	-0.6	1.0	-0.9	1.0	1.0	0.0	1.0	-0.9	-0.9	-0.9	0.0	-0.9	-0.9	-0.9	0.9	1.0	-0.1	-0.8	-0.9	0.0	0.9	-0.9	-0.2	1.0	0.8	1.0	1.0	-0.8	1.0

Legend

	-1 to -0.66
	-0.65 to 0.65
	0.66 to 1

Fig 3.2

		Group 1			Group 2														Group 3												Group 4		
		S26	S0	S14	S1	S2	S28	S24	S18	S10	S30	S32	S7	S8	S5	S27	S29	S19	S16	S25	S12	S6	S17	S11	S13	S4	S21	S15	S31	S22	S23	S3	S9
Group 1	S26	1.0	1.0	1.0	0.2	0.1	0.1	0.1	0.1	-0.1	-0.1	-0.2	0.0	-0.1	0.0	0.0	0.0	0.0	-0.3	-0.2	-0.2	-0.2	-0.2	-0.3	-0.3	-0.6	-0.4	-0.3	-0.4	0.5	0.0	0.0	0.0
	S0	1.0	1.0	1.0	0.3	0.2	0.2	0.2	0.2	0.0	0.0	-0.1	0.1	0.1	0.1	0.1	0.1	0.1	-0.4	-0.3	-0.4	-0.3	-0.3	-0.4	-0.4	-0.7	-0.5	-0.4	-0.5	0.4	0.0	0.0	0.0
	S14	1.0	1.0	1.0	0.3	0.3	0.2	0.2	0.3	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.2	-0.4	-0.4	-0.4	-0.4	-0.3	-0.4	-0.5	-0.8	-0.5	-0.5	-0.6	0.4	0.0	0.0	0.0
Group 2	S1	0.2	0.3	0.3	1.0	0.9	0.9	0.9	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	-0.8	-0.8	-0.8	-0.9	-0.8	-0.8	-0.8	-0.8	-0.8	-0.8	-0.8	-0.6	0.2	0.5	0.1
	S2	0.1	0.2	0.3	0.9	1.0	0.9	1.0	1.0	0.9	0.9	0.9	0.9	0.9	1.0	0.9	1.0	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-0.8	-0.9	-0.9	-0.9	-0.8	0.0	0.2	0.0
	S28	0.1	0.2	0.2	0.9	0.9	1.0	0.9	0.9	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.7	-0.8	-0.9	-0.8	-0.7	0.0	0.3	0.0
	S24	0.1	0.2	0.2	0.9	1.0	0.9	1.0	1.0	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	1.0	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.8	-0.9	-0.9	-0.9	-0.8	0.1	0.4	0.1
	S18	0.1	0.2	0.3	0.8	1.0	0.9	1.0	1.0	1.0	1.0	0.9	1.0	1.0	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-0.8	-0.9	-1.0	-0.9	-0.8	0.0	0.1	0.0
	S10	-0.1	0.0	0.0	0.8	0.9	0.8	0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.7	-0.9	-0.9	-0.8	-0.9	0.0	0.1	0.0
	S30	-0.1	0.0	0.0	0.8	0.9	0.8	0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.7	-0.9	-0.9	-0.8	-0.9	0.0	0.1	0.0
	S32	-0.2	-0.1	0.0	0.8	0.9	0.8	0.9	0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.6	-0.8	-0.9	-0.8	-0.9	0.0	0.1	0.0
	S7	0.0	0.1	0.1	0.8	0.9	0.9	0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-0.9	-1.0	-1.0	-1.0	-1.0	-1.0	-0.9	-0.7	-0.9	-0.9	-0.9	-0.9	0.0	0.2	0.0
	S8	-0.1	0.1	0.1	0.8	0.9	0.9	0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-0.9	-1.0	-1.0	-1.0	-1.0	-0.9	-0.9	-0.7	-0.9	-0.9	-0.9	-0.9	0.0	0.2	0.0
	S5	0.0	0.1	0.1	0.8	1.0	0.9	0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-0.9	-0.7	-0.9	-0.9	-0.9	-0.9	0.0	0.2	0.0
	S27	0.0	0.1	0.1	0.8	0.9	0.9	0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-0.9	-1.0	-1.0	-1.0	-1.0	-0.9	-0.9	-0.7	-0.9	-0.9	-0.9	-0.9	0.0	0.2	0.0
	S29	0.0	0.1	0.1	0.8	1.0	0.9	0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-0.9	-0.7	-0.9	-0.9	-0.9	-0.9	0.0	0.2	0.0
	S19	0.0	0.1	0.2	0.8	1.0	0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-0.8	-0.9	-1.0	-0.9	-0.8	0.0	0.1	0.0
Group 3	S16	-0.3	-0.4	-0.4	-0.8	-1.0	-0.9	-0.9	-1.0	-0.9	-0.9	-0.9	-0.9	-0.9	-1.0	-0.9	-1.0	-1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9	1.0	1.0	1.0	0.7	0.0	-0.1	0.0
	S25	-0.2	-0.3	-0.4	-0.8	-1.0	-0.9	-0.9	-1.0	-0.9	-0.9	-0.9	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9	1.0	1.0	1.0	0.7	0.0	-0.1	0.0
	S12	-0.2	-0.4	-0.4	-0.8	-1.0	-0.9	-0.9	-1.0	-0.9	-0.9	-0.9	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9	1.0	1.0	1.0	0.7	0.0	-0.1	0.0
	S6	-0.2	-0.3	-0.4	-0.9	-1.0	-0.9	-0.9	-1.0	-0.9	-0.9	-0.9	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9	1.0	1.0	1.0	0.7	0.0	-0.1	0.0
	S17	-0.2	-0.3	-0.3	-0.8	-1.0	-0.9	-0.9	-1.0	-0.9	-0.9	-0.9	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9	1.0	1.0	1.0	0.7	0.0	-0.1	0.0
	S11	-0.3	-0.4	-0.4	-0.8	-1.0	-0.9	-0.9	-1.0	-0.9	-0.9	-0.9	-1.0	-0.9	-1.0	-0.9	-1.0	-1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9	1.0	1.0	1.0	0.7	0.0	-0.1	0.0
	S13	-0.3	-0.4	-0.5	-0.8	-1.0	-0.9	-0.9	-1.0	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9	1.0	1.0	1.0	0.7	0.0	-0.1	0.0
	S4	-0.6	-0.7	-0.8	-0.8	-0.8	-0.7	-0.8	-0.8	-0.7	-0.7	-0.6	-0.7	-0.7	-0.7	-0.7	-0.7	-0.8	0.9	0.9	0.9	0.9	0.9	0.9	0.9	1.0	0.9	0.9	1.0	0.3	0.0	-0.1	0.0
	S21	-0.4	-0.5	-0.5	-0.8	-0.9	-0.8	-0.9	-0.9	-0.9	-0.9	-0.8	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9	1.0	1.0	1.0	0.6	0.0	-0.1	0.0
	S15	-0.3	-0.4	-0.5	-0.8	-0.9	-0.9	-0.9	-1.0	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9	1.0	1.0	1.0	0.6	0.0	-0.1	0.0
	S31	-0.4	-0.5	-0.6	-0.8	-0.9	-0.8	-0.9	-0.9	-0.8	-0.8	-0.8	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.5	0.0	-0.1	0.0
	S22	0.5	0.4	0.4	-0.6	-0.8	-0.7	-0.8	-0.8	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.8	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.3	0.6	0.6	0.5	1.0	0.0	-0.1	0.0
Group 4	S23	0.0	0.0	0.0	0.2	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.9	1.0
	S3	0.0	0.0	0.0	0.5	0.2	0.3	0.4	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2	0.2	0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	0.9	1.0	0.8
	S9	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.8	1.0

Legend	
	-1 to -0.66
	-0.65 to 0.65
	0.66 to 1

4. Discussion

Group I :

All systems from Group I in Fig 3.3 are positively correlated to each other with correlation coefficients that average at 0.94. They all display one strong motif, also referred as 3-loop (id position 8 in fig:3.1) and one anti-motif, also referred as "V-in" (id position:2 V-in in fig:3.1). Interestingly systems in this group turned out to be Electronic Circuits* of the type Digital Fractional Multiplier

* All Electronics circuits are picked up randomly. Nodes represent component gates and edges represent the interconnection between gates.

namely s208, s420 & s838 (systems S26, S20 & S14 respectively using terminology used in figure 3.2 & 3.3).

Group II :

All systems from Group II in Fig 3.3 are positively correlated to each other with correlation coefficients that average at 0.99. They display one strong motif, also referred as "feed forward loop" (id position:7 in fig:3.1) and one anti-motif, also referred as "V-in" (id position:2 in fig:3.1). Interestingly systems in this group turned out to be all 5 Softwares[†] namely Apword (S1), Linux (S2), MySQL (S28), Vtx (S24), XMMS (S18) and few Electronic Circuit[*] namely Traffic System s444 (S10), PLD s713 (S30), ALU c2670 (S32), ECAT c1355 (S7), Forward Logic Chains s9234, s13207, s15850, s38417, s38584(S8, S5, S27, S29, S19). This feed forward loop has been shown to perform signal processing functions like acceleration of transcription response, pulse generation etc for transcription regulation family in biology [Milo R et all].

Group III :

All systems from Group III in Fig 3.3 are positively correlated to each other with correlation coefficients that average at 0.92. They display one strong motif, also referred as "V-in" (id position:2 in fig:3.1) and one anti-motif, also referred as "feed forward loop" (id position:7 in fig:3.1). Interestingly systems in this group turned out to be Electronic Circuits[*] namely Traffic System s400, s382, s526 (S16, S12, S22), PLD s820, s832, s641 (S25, S17, S21), ALU 74181, c880, c7552, c3540 (S6, S13, S4, S31) and ECAT c499, c1908 (S11, S15).

Group IV :

All systems from Group IV in Fig 3.3 are positively correlated to each other with correlation coefficients that average at 0.92. They display one strong motif, also referred as "clique" (id position:13 in fig:3.1) and one anti-motif, also referred as "Mutual V" (id position:6 in fig:3.1). Interestingly systems in this group turned out to be all Mechanical[‡] systems namely Robot (S23), Aircraft Engine (S3) and Refrigerator (S9).

---

† These Softwares chosen randomly from open source for our study. Nodes represent a software class and edges represents reference between classes.

‡ All Mechanical systems picked up randomly for our study. Nodes represent physical components and edges represent exchange of energy, material or signal between components.

The correlation across groups is also interesting. Group I is unlike other 3 groups. Group IV is similarly unlike other 3 groups. Group II and Group III are highly negatively correlated with respect to each other, while being unlike Group I or Group II. There are a few systems that fall marginally outside of this observation.

A pictorial comparison of significance profiles of pairs of systems that are positively correlated, negatively correlated, not correlated is shown from Fig 3.4 to 3.6. (The 13 centered symbols from left to right represent Z_j for j=1 to 13. The lines joining the 13 symbols have no meaning and are present only to create a visual impact). When systems are positively correlated all the 13 patterns are over (under) represented to the same extent. When systems are negatively correlated all the 13 patterns are over (under) represented in an inverse manner. When systems are not correlated the 13 patterns do not show any such relation.

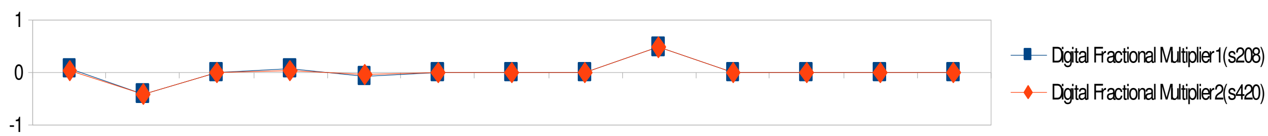

Fig 3.4 Significance Profile of 2 systems taken from Group I – Positive Correlation between them

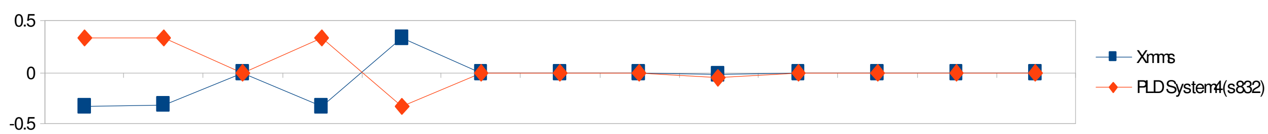

Fig 3.5 Significance Profile of one system from Group II & III – Negative Correlation between them

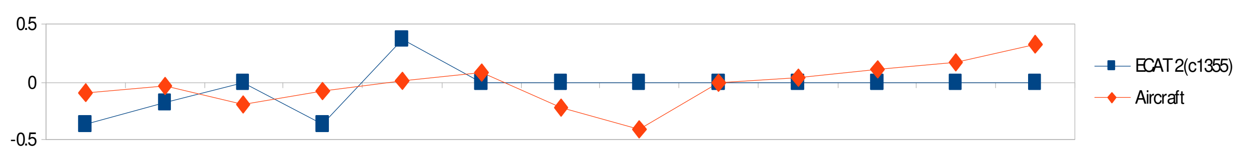

Fig 3.6 Significance Profile of one system from Group III & IV – No correlation between them

5. Conclusion & Directions

Ideas related to complex system architectures may give insight into previously complex and poorly understood phenomena in engineering domain. Albert Barabasi argues that, "The science of networks is experiencing a boom. But despite the necessary multidisciplinary approach to tackle the theory of complexity, scientists remain largely compartmentalized in their separate disciplines" [Albert László Barabási]. The application of this complex system architectures theory is still in infancy and has very recently entered into study of engineering systems or their design. This paper has calculated motifs and significance profile for system architectures based on components across 32 diverse engineering systems. Interesting motifs are seen in all systems. Motif significance profiles across systems has indicated interesting grouping of systems that coincides with their grouping as software, electrical and mechanical. This study has thrown some insights about motif being a possible building block to understand complex engineering systems.

6. Acknowledgements

We thank Aeronautics Research and Development Board (ARDB), India for this project grant. We are also thankful to Mr Mahesh for his support in computing.

7. References

Albert László Barabási. 2005, Taming Complexity. *Nature Physics* (1): 162, 68 – 70.

Albert L and Albert R. 2004, Conserved Network Motifs Allow Protein-Protein Interaction Prediction, *Bioinformatics* (20): 18, pp 3346-3352.

Amro M. Farid and Duncan C. McFarlane, An approach to the application of the design structure matrix for assessing reconfigurability of distributed manufacturing systems, *Proceedings of the IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications (DIS'06)* .

Amro M. Farid and Duncan C. McFarlane. 2006*,* An Approach to the Application of the Design Structure Matrix for Assessing,Reconfigurability of Distributed Manufacturing Systems, *Proceedings of the IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications*.

ANSI IEEE Standard 1471. http://www.iso-architecture.org/ieee-1471/

Boccaletti S, Latora V, Moreno Y, Chavez M. and Hwang D. 2006, Complex networks: Structure and dynamics, *Physics Reports,* 424 (4): 175-308.

Dan Braha and Yaneer Bar-Yam. 2007, Statistical Mechanics of Complex Product Development, *Management Science*, 53 (7): 1127-1145.

Duncan J Watts. 2004, Six Degrees: The Science of a Connected Age. Norton and Company, NewYork.

Edward Crawley, Olivier de Weck, Steven Eppinger, Christopher Magee, Joel Moses, Warren Seering, Joel Schindall, David Wallace and Daniel Whitney. 2004, The Influence of Architecture in Engineering Systems. *MIT Engineering System Division Monograph*.

Homas U. Pimmler and Steven D. September 1994 Eppinger Integration analysis of product decompositions, *ASME Design Theory and Methodology Conference Minneapolis.*

ISCAS High level models: http://www.eecs.umich.edu/~jhayes/iscas.restore/

ISCAS'89 benchmark data : http://www.pld.ttu.ee/~maksim/benchmarks/iscas89/bench/

J.M. Ottino. 2004, Engineering Complex Systems, *Nature*, **427**, 399 .

Manuel E Sosa, Steven D Eppinger and Craig M. Rowles. 2003, Identifying Modular and Integrative Systems and Their Impact on Design Team Interactions, *Journal of Mechanical Design*, 125 (2): 240-252.

Mathworld website, Pearson Clustering Coefficient http://mathworld.wolfram.com/CorrelationCoefficient.html

Milo R, Shen-Orr, Itzkovitz S, Kashtan N, Chklovskii D and Alon. 2002 , Network Motifs: Simple Building Blocks of Complex Networks. *Science*, : 298, 824-827.

M Ronen, R Rosenberg, B.I. Shraiman and U. Alon. 2002, Assigning Numbers to the Arrow: Parameterizing a Gene Regulation Network by Using Accurate Expression Kinetics, *Proceedings of the National Academy of Sciences* (99) : 16, 10555-10560.

Newman MEJ. 2003, The structure and function of complex networks, *SIAM Review*, 45 (2): 167-256.

L Kaufman, P.J Rousseeuw, 1990, Finding Groups in Data: An Introduction to cluster Analysis, New york, *John wiley & Sons Inc.*

Preferential Attachment Wiki: http://en.wikipedia.org/wiki/Preferential_attachment

Ron Milo, Shalev Itzkovitz, Nadav Kashtan, Reuven Levitt, Shai Shen-Orr, Inbal Ayzenshtat, Michal Sheffer, Uri Alon. March 2004 Superfamilies of Evolved and Designed Networks, *Science* (303): 5663, pp. 1538 – 1542.

Shaja AS, Sudhakar K, CASMots: http://www.casde.iitb.ac.in/complexsystems/motifs/

Snijders T and Borgatti S. 1999, Non-Parametric standard errors and tests for network statistics, *Connections*, 22 (2):161–170.

S.S. Shen-Orr, R. Milo, S. Mangan and U. Alon. 2002, “Network Motifs in the Transcriptional Regulation Network of Escherichia Coli” *Nature Genetics* (31): 1, pp 64-68.

Software graph data for specified software systems: http://www.tc.cornell.edu/~myers/Data/SoftwareGraphs/index.htm

Tyson R Browning. 2001, Applying the Design Structure Matrix to System Decomposition and Integration Problems: A Review and New Directions, IEEE *Transactions on Engineering Management*, 48 (3) : 292-306.

Thomas U. Pimmler and Steven D. Eppinger,Integration analysis of product decompositions, *ASME Design Theory and Methodology Conference*

Vito Latora and Massimo Marchiori. 2001, Efficient Behavior of Small-World Networks. *Physical Review Letters*: 87, 198701.

Watts DJ. 1999, Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology*, 10 (2) : 493-527.

Yael Artzy-Randrup, Sarel J. Fleishman, Nir Ben-Tal, and Lewi Stone. August 2004, Comment on Network Motifs: Simple Building Blocks of Complex Networks and Superfamilies of Evolved and Designed Networks, *Science*,1107c.

Zhi-Qiang Jiang, Wei-Xing Zhou, Bing Xu and Wei-Kang Yuan. 2007, Process flow diagram of an Ammonia Plant as a Complex Network, *AIChE Journal*, 53 (2): 423-428.

8. Biography

Shaja. A. S is currently a doctoral student in Aerospace Engineering Department, Indian Institute of Technology Bombay. He previously did his M.Tech in Indian Institute of Technology Bombay. His recent research interests include Multidisciplinary Design Optimization, Systems Engineering, Networks etc.

Sudhakar. K is a Professor in Aerospace Engineering Department, Indian Institute of Technology Bombay, for the last 25 years. He was previously associated with Aerospace Industry as a Design Expert for 12 years. He completed his B.Tech, M.Tech and Doctorate from Indian Institute of Technology. Some of his recent research interests include optimization in design, Flight Mechanics, Systems Engineering, Mini Aerial Vehicle, Modeling and simulation etc.