6 PRODUCT DEVELOPMENT PROCESS ARCHITECTURE OF DIVERSE ENGINEERING SYSTEMS — AN EMPIRICAL STUDY

A. S. Shaja[a] and K. Sudhakar[b]

Department of Aerospace Engineering, IIT Bombay, Bombay, India. E-mail: [a]shaja@aero.iitb.ac.in, [b]sudhakar@aero.iitb.ac.in

Engineering design process is used by engineers to develop products and comprises of enormous number of interdependent tasks. Literature shows that understanding the interaction of tasks is important to improve the design process which in turn impacts the success of the product. In this paper we consider task interdependency dimension (Product Development Process Architecture) of diverse engineering systems. The last few years have seen new research interest in the study of complex system architectures, like social networks, biological systems, etc. Seminal works covering each of these systems have appeared in high impact journals like Nature, Science, etc. Unifying principles have helped in gaining new understanding or extending the understanding gained in one domain to another. This paper is inspired by these developments. We consider a variety of product development systems ranging from vehicle design, construction, software, aircraft engine etc. We show that the considered systems are small-worlds. We determine motifs for each of the considered system and investigate whether these systems share any motifs across each other. We calculate some centrality metrics for the considered systems.

Keywords: Product Development Process Architecture, Motifs, Small World, Metrics.

1. INTRODUCTION

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.[1] 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.[2] The last few years have seen a new research interest in the study of complex system architectures, across domains like social networks, biological systems, etc.[3,4] Unifying principles have helped in gaining new understanding or extending the understanding gained in one domain to another[5] But "Engineers seem a little bit indifferent as if engineering is at the edge of the science of complexity" is a growing feeling.[6]

Engineering design process is used by engineers to develop products. The development of any complex engineering product involves enormous number of interdependent tasks. Literature[7,8] shows that understanding the interaction between tasks is important to improve the design process which in turn impacts the success of the product. This paper views the task dependency dimension (Product Development Process Architecture) of diverse engineering systems. We consider 5 product development system designs like an aircraft engine design,[7] a vehicle design, hospital facility design, an operating software design and a pharmaceutical facility design.[8] We abstract the product development process architecture of the considered systems as a complex network/graph, where the node/vertex corresponds to tasks in the system and edges correspond to dependency between tasks. We show that all the considered product networks are small world networks. In this paper, we discover motifs of sizes 3 to 5 and tabulate the similar motifs observed across the considered systems. Some centrality measures with respect to considered systems from complex system architecture are also analyzed.

2. MOTIFS

Motifs have been suggested to be the functional building blocks of network complexity. "Motifs are recurring sub graphs of interactions from which the networks are built".[11] These are patterns of interconnections (over represented sub graphs) occurring in complex networks at numbers significantly higher than those in randomized networks. Milo *et al.*[11] have proposed an algorithm to detect network motifs. As per their method each network is scanned for all the possible m-node sub graphs and the number of times each sub graph occurred is noted (Count real). To concentrate on those sub graphs that are important, a comparison of real network with suitably generated random network is made. Each randomized network is generated subject to two conditions: (i) Degree distribution of the real network is conserved in the randomized network (ie number of nodes having a particular count of edges connected to it will be same in real and randomized network, in a statistical sense) (ii) Each randomized network which is used to calculate the significance of m-node sub graph has the same number of appearances of (m-1) node sub graphs as in real network.

We created 1000 random networks for each considered system with the above criteria. We calculated the number of times each sub graph occurs and estimated its mean, Countrand and its standard deviation, σ. As a measure of statistical significance, we calculate Z-score as (Count real — Count rand)/σ also[a]. We discovered all the motifs of size 3 to 5 for each considered system, out of which, only those having high Z-score are tabulated in Table 1. It is interesting to observe common motifs shared across the considered diverge systems. We have archived the detailed results for all motifs with size from 3 to 5 in our website.[12]

3. SMALL WORLD EFFECT

This concept emerged from the "small-world" experiments of Milgram – where letters passed from one person to other was able to reach designated target in small number of hops.[3] Small world effect has been studied and verified in different types of networks.[9] The mathematical characterization of small-world behavior is based on of 2 quantities, viz the average path length L and the clustering coefficient C. Geodesic path is the average of shortest path associated with each pair of vertices of the network.[4] The average geodesic path L for a graph G is defined as

$$L = \frac{1}{1/2n(n-1)} \sum_{1 \le i,j \le n} d_{ij}$$

where d_{ij} shortest path from vertex i to vertex j and n is the number of vertices.

In many networks it has been observed that if vertex A1 is connected to vertex A2 and vertex A2 to vertex A3, then there is more probability that vertex A1 will also be connected to vertex A3. In other words, Clustering Coefficient C (also referred in literature as transitivity) means the presence of sets of three vertices each of which is connected to each of the others.[4]

$$C = \frac{3 * number\ of\ triangles}{number\ of\ connected\ triples\ of\ vertices}$$

where 'connected triple' means a single vertex with edges connecting to a pair of others. In other words C indicates how much the adjacent vertices of the adjacent vertices of x are adjacent vertices of x.[3]

Watts and Strogatz proposed a model for small world. The model starts with a ring of n vertices, each connected to its k nearest neighbors. In this model, a vertex and the edge that connects it to its nearest neighbors (in a clockwise sense) is chosen. With probability p, this edge is reconnected to a vertex chosen uniformly at random over the entire ring, with duplicate edges forbidden. This process is repeated by moving clockwise around the ring, considering each vertex in turn until one round

[a] Only motifs with significant Z score (Z score $>$ 2) are considered in our experiment

is completed. Next, the edges that connect vertices to their second-nearest neighbors clockwise are considered. As before, each of these edges are randomly rewired with probability p. This process is continued, circulating round the ring and proceeding to distant vertex in the ring after each round, until each edge in the original lattice has been considered once. As p increases, the graph becomes increasingly disordered until for p = 1, all edges are rewired randomly. Small world networks can be regarded as special networks which are superposition of regular and random networks.[10] The graph as plotted using above procedure reveals small world regions — the clustering coefficient is high yet with geodesic path is small (intermediate values of p). We constructed several random graphs following the above procedure for each considered system from p=0 to p=1. A graph for one of the considered systems is shown in Figure 1. L & C for all the considered systems are tabulated in Table 2. As per, Ref. 9 a network is small world when Lregular > Lreal > Lrandom and Cregular > Creal >> Crandom.

This condition is satisfied by all the systems considered for study. The small geodesic path and large clustering coefficient observed for real networks considered here might imply that a change created by a task may appear to be encapsulated within its cluster, but in reality might navigate quickly across the entire network.

4. CENTRALITY MEASURES

A metric/measure is an indicator of a system characteristic. Centrality measures helps to determine the relative importance of a node/edge within the graph. Degree[b] centrality is defined as the number of links incident upon a node. The betweenness centrality of a node/edge is defined by the number of geodesic paths (shortest paths) passing through the respective node/edge.[13] We calculated the centrality measures for all considered systems (Centreal).

We then created 1000 corresponding random graphs (preserving the number of edges and nodes of corresponding real network) and calculated the mean centrality measure Centrand and the standard deviation, σ. As a measure of statistical significance, Y-score = (Centreal — Centrand)/σ is also calculated. We tabulate the top three measures for each system in Table 3. Detailed numerical values

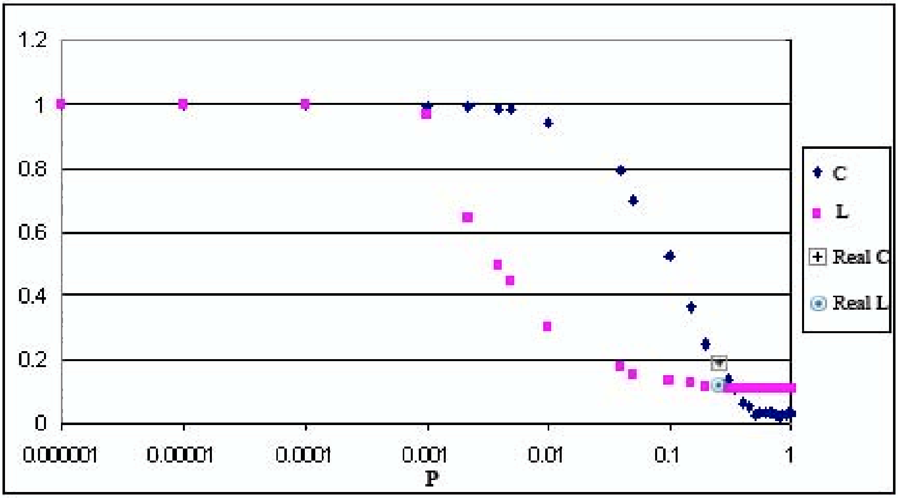


Figure 1. Avg Shortest path (L) and Clustering coeff (C) for family of Randomly rewired graphs[c].

[b]Indegree is count of the number of ties directed to the node, and outdegree is the number of ties that the node directs to others.

[c]This graph is for one of the considered system — software design. C and L are normalized with C &L at p=0. Logarithmic horizontal scale is used to resolve rapid drop in L.

Table 1. Degree and Between Centrality measures for considered systems.

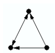
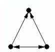
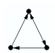
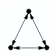
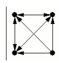
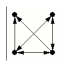
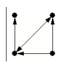
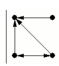
System		Aircraft engine design			Vehicle design			Hospital facility design			Operating software design			Pharmaceutical facility design		
		N=54, E=353			N=120, E=416			N=820, E=7751			N=386, E=1221			N=579, E=3568		
		$Count_{real}$	$Count_{rand}\pm\sigma$	Z score	$Count_{real}$	$Count_{rand}\pm\sigma$	Zscore	$Count_{real}$	$Count_{rand}\pm\sigma$	Zscore	$Count_{real}$	$Count_{rand}\pm\sigma$	Zscore	$Count_{real}$	$Count_{rand}\pm\sigma$	Zscore
Motif 3		2.70%	1.2161%±0.0016263	9.093	0.23%	0.069202%±0.00047978	3.36	0.08%	0.058746%±0.000068222	2.806	0.77%	0.24863%±0.0004847	10.85	0.16%	0.07699%±0.000063229	12.847
		2.18%	1.1448%±0.0014285	7.278	0.30%	0.057055%±0.00041464	5.77	0.22%	0.19422%±0.0001117	2.247	0.86%	0.25289%±0.00051681	11.76	0.27%	0.035591%±0.000056665	41.57
		2.13%	1.2712%±0.0016513	5.186	NonSignificant	NonSignificant	NonSignificant	0.06%	0.04866%±0.000058362	2.065	0.41%	0.27104%±0.00051036	2.785	NonSignificant	NonSignificant	NonSignificant
		7.49%	0.32499%±0.0011481	62.4	NonSignificant	NonSignificant	NonSignificant	NonSignificant	NonSignificant	NonSignificant	0.51%	0.0066831%±0.000083306	60.16	0.004%	0.0000000055555%±0.00000017568	261.01
Motif 4		0.14%	0.00070198%±0.000014875	93.01	NonSignificant	NonSignificant	NonSignificant	0.001737%	0.000000000000362%±0.0000000011	1517.1	0.01%	0.000000031826±0.00000037925	149.07	0.000159%	0.000000000000766%±0.000000002425	659.14
		0.47%	0.00057474%±0.000012755	64.622	NonSignificant	NonSignificant	NonSignificant	0.000000226%	0.0000000000384%±0.0000001013	2.198	0.06%	0.00021323%±0.0000046651	118.59	0.0002%	0.0000000000023%±0.0000000072	351.52
		0.41%	0.17531%±0.0003759	6.2116	0.09%	0.032738%±0.00025811	2.4092	0.06%	0.024559%±0.000037489	8.669	0.46%	0.15951%±0.0004062	7.302	NonSignificant	NonSignificant	NonSignificant
		0.31%	0.17859%±0.00024259	5.2485	0.16%	0.031079%±0.0002964	4.289	0.15%	0.11889%±0.000074258	4.657	0.10%	0.057378%±0.00013688	2.8334	0.008%	0.0034569%±0.0000058402	8.382

Table 1. Degree and Between Centrality measures for considered systems. *(continued)*

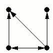
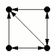
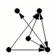
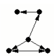
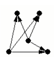
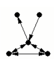
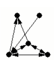
		0.59%	0.34786%±0.0005238	4.614	0.14%	0.031484%±0.00025477	4.422	0.33%	0.27218%±0.00018637	2.882	0.88%	0.32594%±0.00080134	6.925	0.28%	0.048347%±0.000089951	25.73
		NonSignificant	NonSignificant	NonSignificant	0.02%	0.00075833%±0.00001971	10.317	0.02%	0.0038254%±0.0000071161	22.067	0.02%	0.002744%±0.000016818	7.54	0.005%	0.00000000359%±0.00000022992	215.7
Motif 5		0.02%	0.0103%±0.000040285	2.045	0.009%	0.0002%±0.000005186	16.04	0.01%	0.00249%±0.000010541	8.166	0.009%	0.00202%±0.000014366	4.64	0.03%	0.000307%±0.0000044376	57.76
		0.0047%	0.000798%±0.0000071936	5.465	0.002%	0.000232%±0.0000076554	2.01	0.007%	0.000828%±0.0000037343	18.31	0.005%	0.000703%±0.0000061953	6.11	0.002%	0.000234%±0.0000026505	5.292
		0.05%	0.01039%±0.000029908	14.41	0.04%	0.0121%±0.00011086	2.42	0.005%	0.002998%±0.0000063594	3.265	0.05%	0.02%±0.0000537	6.17	0.01%	0.00608%±0.0000269	2.96
		NonSignificant	NonSignificant	NonSignificant	0. 04%	0.00387%±0.0000227	15.767	0.2%	0.098923%±0.00022	80.94	0.004%	0.001106%±0.00000646	4.11	0.2%	0.078%±0.0003005	66.42
		0.2%	0.00723%±0.0000481	41.97	NonSignificant	NonSignificant	NonSignificant	0.000032%	0.000000962%±0.0000000962	3.217	0.1%	0.001471%±0.00002434	43.382	0.000115%	0.000000702%±0.000000045873	25.1
		0.04%	0.000929%±0.000010389	40.37	NonSignificant	NonSignificant	NonSignificant	0.0000174%	0.000000089%±0.0000000089096	19.44	0.02%	0.0000682%±0.0000019511	75.409	0.000115%	0.0000000678%±0.000000006781	170.79

Table 2. C and L for regular, real and random networks across 5 systems.

No	System	Nodes	Edges	$C_{regular}$	$L_{regular}$	C_{real}	L_{real}	C_{random}	L_{random}
1	Aircraft engine design	54	353	0.6923	2.41509	0.513342	1.99441	0.228732	1.80293
2	Vehicle design	120	416	0.64285	7.94117	0.12913	2.8781	0.06756	2.5724
3	Hospital facility design	820	7751	0.71052	20.9755	0.10636	3.11754	0.02399	2.5843
4	Operating software design	386	1221	0.6	32.7519	0.13903	3.70019	0.01147	3.5341
5	Pharmaceutical facility design	579	3568	0.69230	21.14533	0.11799	2.7276	0.02217	2.70054

of centrality measures are not tabulated in this paper due to page constraints. All detailed numerical values of all centrality measures of all nodes are archived in Ref. 12.

5. CONCLUSION AND DIRECTIONS

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".[14] The application of this complex system architectures theory is still in infancy stage and has very recently entered into study of engineering systems or their design. This study has revealed that 5 systems considered have revealed 'small world' behaviour. Motifs of size 3, 4 and 5 have been observed. Some of these motifs are even common to almost all the 5 systems. For the considered systems, we cannot precisely say whether the above parameters are the only important quantities to measure. Trying to answer the questions on significance of observed motifs in above systems, finding new metrics, motifs, models with respect to engineering systems etc. form future work. Some processes taking place in system, like phase transition, spreading etc, will be interesting to experiment. Ideas related to complex system architectures may give insight into previously complex and poorly understood phenomena in engineering domain. The generic interaction between considered system and systems from other domains like biology etc will be an interesting area to explore.

Table 3. Degree and Between Centrality measures for considered systems.

S. No	Centrality Measures		Aircraft engine design			Vehicle design			Hospital facility design			Operating software design			Pharmaceutical facility design		
			$Cent_{real}$	$Cent_{rand} \pm \sigma$	Y_{score}	$Cent_{real}$	$Cent_{rand} \pm \sigma$	Y_{score}	$Cent_{real}$	$Cent_{rand} \pm \sigma$	Y_{score}	$Cent_{real}$	$Cent_{rand} \pm \sigma$	Y_{score}	$Cent_{real}$	$Cent_{rand} \pm \sigma$	Y_{score}
1	Betweenness	Node	252.78	107.9±18.326	7.902736	4623.82	1375.376±376.37	8.631	112852.17	3448.17±1113.1	98.28	20827.2	8055.3±658.4	19.3	58702.01	5508.2±525.8	101.16
			237.65	93.27±7.19	20.07065	3995.5	1016.01±166.07	17.94	86669.78	3123.51±1014.5	82.35	12200.86	6983.5±797.4	6.54	37149.005	5206.3±661.25	48.3
			227.99	87.85±6.56	21.35421	2802	970.07±169.07	10.84	83251.08	3035.1±974.6	82.30	10378.76	6409.4±614.2	6.46	35897.57	4874.6±680.93	45.559
		Edge	55.815	22.811±1.675	19.698	3960.12	533.55±77.65	44.127	76515.49	243.937±8.82	8638.49	3438.25	2666.6±156.2	4.93	31201.91	1228.3±48.223	621.55
			39.11	22.12±1.89	8.988	2224.98	496.6±75.42	22.91	28046.27	235.75±10.51	2644.54	2944.38	2566.5±176.7	2.13	26420.95	1180.02±71.04	355.26
			37.66	20.76±2.07	8.15	2136.96	477.54±81.672	20.31	18021.004	231.88±10.21	1741.7	2897.36	2476.1±183.8	2.29	23061.41	1135.9±57.63	380.4
3	Degree	In	23	17±1.22	4.898	14	10±1	4	122	36±1.92	44.7	31	11±0.89	22.36	37	15±0.83	26.29
			21	16±1	5	12	8±0.54	7.302	103	35±1.14	59.63	29	9±0.44	44.72	24	14±0.7	14.14
			21	16±0.5	9.1287	12	8±0.707	5.65	102	34±1.516	44.83	22	8±0.54	25.5	24	14±0.44	22.36
		Out	27	17±1.92	5.1987	24	10±0.836	16.7	261	37±1.140	196.46	52	10±1.303	32.21	343	15±0.836	392.03
			26	16±1.14	8.7705	15	9±0.83	7.17	158	33±0.447	279.50	47	9±0.89	42.48	136	14±0.44	272.8
			24	16±1.3	5.9628	12	8±0	8.94	147	33±0.7	161.22	32	8±0.44	53.6	136	13±0.54	224.56
		Total	48	30±1.3	13.805	26	15±1.64	6.69	308	64±1.94	125.16	83	16±1.51	44.17	343	24±0.707	451.1
			47	28±1.2	15.5134	24	14±1.22	8.16	198	63±1.78	75.46	61	14±1.09	42.9	151	24±1.14	111.3
			38	28±1.3	7.669	21	13±0.70	11.31	183	61±1.3	93.56	52	14±0.54	69.3	145	23±1	122

ACKNOWLEDGEMENTS

We thank Center for Aerospace Systems Design and Engineering (CASDE), Aerospace Engineering Department, IIT Bombay for the research environment. The analysis for small world and centrality measures makes some use of an open source library of igraph (http://igraph.sf.net/). The analysis of motifs makes use of open source code named FanMod (http://bioinformatics.oxfordjournals.org/cgi/content/full/22/9/1152). We are also thankful to Mr Balan and Mr Mahesh for their support in computing.

REFERENCES

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

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

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

[4] Newman MEJ. “The structure and function of complex networks”, SIAM Review, 2003, **45** (2), 167–256.

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

[6] 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.

[7] 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.

[8] Dan Braha and Yaneer Bar-Yam (2007) “Statistical Mechanics of Complex Product Development”, Management Science, **53** (7), 1127–1145.

[9] Watts D. J. (1999). Networks, dynamics, and the small-world phenomenon”, American Journal of Sociology, **10** (2), 493–527.

[10] Vito Latora and Massimo Marchiori (2001). Efficient Behavior of Small-World Networks”, Physical Review Letters, 87, 198701.

[11] 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.

[12] Shaja AS, Sudhakar K, Empirical results for 5 product development systems. http://www.casde.iitb.ac.in/complexsystems/proddev/results/

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

[14] Albert László Barabási. Taming Complexity. 2005, Nature Physics 1, 162, 68–70.