



Bridge and brick network motifs: Identifying significant building blocks from complex biological systems

Chung-Yuan Huang^{a,b,*}, Chia-Ying Cheng^{c,1}, Chuen-Tsai Sun^{c,1}

^a Department of Computer Science and Information Engineering, Chang Gung University, 259 Wen Hwa 1st Road, Taoyuan 333, Taiwan

^b Research Center for Emerging Viral Infections, Chang Gung University, 259 Wen Hwa 1st Road, Taoyuan 333, Taiwan

^c Department of Computer Science, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 300, Taiwan

Received 20 November 2006; received in revised form 22 July 2007; accepted 24 July 2007

KEYWORDS

Small-world properties;
Local clustering;
Strong/weak-tie link;
Network-oriented approach;
Complex biological systems;
Network motif

Summary

Objective: A major focus in computational system biology research is defining organizing principles that govern complex biological network formation and evolution. The task is considered a major challenge because network behavior and function prediction requires the identification of functionally and statistically important motifs. Here we propose an algorithm for performing two tasks simultaneously: (a) detecting global statistical features and local connection structures in biological networks, and (b) locating functionally and statistically significant network motifs.

Methods and material: Two gene regulation networks were tested: the bacteria *Escherichia coli* and the yeast eukaryote *Saccharomyces cerevisiae*. To understand their structural organizing principles and evolutionary mechanisms, we defined bridge motifs as composed of weak links only or of at least one weak link and multiple strong links, and defined brick motifs as composed of strong links only.

Results: After examining functional and topological differences between bridge and brick motifs for predicting biological network behaviors and functions, we found that most genetic network motifs belong to the bridge category. This strongly suggests that the weak-tie links that provide unique paths for signal control significantly impact the signal processing function of transcription networks.

* Corresponding author at: Department of Computer Science and Information Engineering, Chang Gung University, 259 Wen Hwa 1st Road, Taoyuan 333, Taiwan. Tel.: +886 3 2118800x3474; fax: +886 3 2118700.

E-mail address: gscott@mail.cgu.edu.tw (C.-Y. Huang).

¹ Tel.: +886 3 5712121x56612; fax: +886 3 5711490.

Conclusions: Bridge and brick motif content analysis can provide researchers with global and local views of individual real networks and help them locate functionally and topologically overlapping or isolated motifs for purposes of investigating biological system functions, behaviors, and similarities.

© 2007 Elsevier B.V. All rights reserved.

1. Introduction

Network-oriented approaches are receiving significant attention in computational biology. Increasingly powerful computers and Internet accessibility make it easy for biologists to share their latest research results via genome and proteome databases such as the Human Genome Project. In these and other large databases, system biologists integrate network theory with computer simulations in order to solve complex problems. As the borders between scientific disciplines become more fluid, system biologists regularly apply research approaches developed by computer scientists and statistical physicists to analyze local and global features, system functions, and dynamic behaviors in complex biological networks.

Commonalties have emerged from studies of complex networks in fields ranging from biology to social and computer sciences. Three global features in complex networks have been identified and investigated: highly clustered connections [1,3,6], small-world properties [2–6], and the scale-free phenomenon [1,9,40–42]. Approaches based on quantitative and qualitative analyses of the topological properties of complex networks are being utilized to study how the global features of network topological structures affect the dynamic behavior of networks [1,13,18,22,36,39,40]. This is currently considered one of the field's most important and challenging research topics [1–6,25,26,36,39,40].

Some local structural motifs (building blocks) reveal unique and statistically significant patterns when compared with random [7,8,10–15], biological [7,11,15,16,18], and food web [7,17] motifs; all are thought to contain important information. However, simple motifs of complex networks that are statistically significant but functionally unimportant are clearly inadequate for investigating network functions and dynamic behaviors [7,8,12,15,19–22]. We therefore propose an algorithm to perform two tasks: simultaneously detecting global features and local structures in complex networks, and identifying functionally and statistically significant network building blocks from complex networks.

When considering the global features and local structural motifs of biological networks, it is worth noting that link properties (weights) exert strong impacts on network functions and dynamic behaviors [1,17,23,24,44]. Examples include the role of weak

links in the six degrees of separation (i.e., small-world) effect of interpersonal networks [23,24] and the strength of predator–prey interactions that determine the stability of ecological communities [17]. Network researchers have reported that a weighted value representing interaction strength can be assigned to each link (edge) in a real network [1,25,27]. We therefore took into consideration network motif link strength in terms of two categories: bridge motifs and brick motifs (Fig. 1). Bridge motifs consist of weak links only or a minimum of one weak link; brick motifs consist of strong links only (Table 1). Bridge motifs connect clusters and reduce the average degree of separation, and brick motifs exhibit the local clustering phenomenon in biological networks.

2. Background

Although small-world properties and the scale-free phenomenon are common statistical features of complex networks, they can have very different local structures, thus underscoring the importance of understanding the local structures of real networks and the design principles and evolutionary mechanisms that generate them. Taking gene regulation networks as an example, recent studies have shown that gene biochemical functions operate in a highly modular manner [34–36,39], with independent tasks being performed by modules consisting of

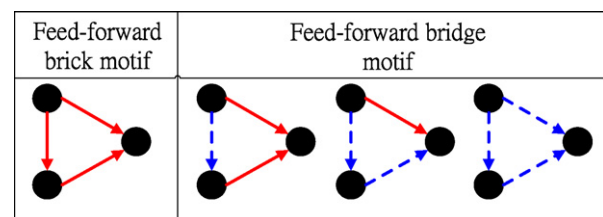















Figure 1 Network motifs example. Network motifs can be separated into two categories: bridge and brick. Using the three-point feed-forward motif as an example, it can be divided into two categories: a three-point feed-forward brick motif (left box) composed of three strong (red) links, and a three-point feed-forward bridge motif (right box) composed of at least one weak (blue) link and a maximum of two strong (red) links. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Table 1 Thirteen directed three-node subgraphs without weighted values (i.e., simple network motifs)

ID	Motif
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	

single groups of physically or functionally connected nodes [34,39,40,43]. Many real networks contain large numbers of modules—for instance, groups of coworkers in a corporation or World Wide Web hyperlinks on a specific topic.

In many complex and large-scale systems, reusable components and highly modular structures represent important design principles [1,34,40]. Examples of modularized functions and structures in ecosystems include food webs, gene transcriptions, protein–protein interactions, and real networks distributed across various bases, genes, amino acids, proteins, cells, organs, and food chains. These modules direct biochemical reactions, growth and development, and evolutionary processes [7,18,28,43].

Understanding complex network modularity requires appropriate measurement tools and the identification of modular relationships within a

given real network [25,26,40]. Module identification is considered difficult due to contradictions between the modularity concept and the scale-free phenomenon of small-world networks. However, most nodes in scale-free networks are connected to central or core nodes—a structure that affects the appearance of isolated modules.

Results from recent studies suggest that motifs can serve as basic modular elements in complex networks [7,12,37,40]. The clustering degrees of most real networks are much higher than those of randomized networks with the same number of nodes and edges, indicating that real networks can contain multiple subgraphs composed of strongly connected nodes. However, multiple subgraphs in real networks may not hold equal importance, with some subgraphs appearing much more frequently than others in complex networks. These statistically significant subgraphs are called motifs [7].

Based on the original motif concept, Milo et al. proposed a method for comparing local structures among different complex networks according to their significance profiles (SPs) [28]. The method can be used to classify real networks in which similar SP vector values belong to the same superfamily. There is evidence indicating that similar types of real networks not only have similar network motifs, but also have similar relative significance within real networks [7,28]. At the same time, a network superfamily may contain networks consisting of vastly different scales and functions.

3. Bridge and brick network motif-detecting algorithms

As shown in Eq. (1), a link-weighted value dependent on the number of all possible paths between two linked nodes equals the summation of the reciprocal values of all possible path lengths except for the link itself (Fig. 2)

$$\text{weight}(a, b) = \sum_i \frac{1}{\text{length}(\text{path}_i(a, b))} \quad (1)$$

where $\text{path}_i(a, b)$ indicates the i th path from node a to node b ; $\text{path}_i(a, b) \neq \text{edge}(a, b)$; and $\text{length}(\text{path}_i(a, b)) \leq \text{average network diameter}$. The length of one path represents its total number of nodes. Average network diameter and $\text{ShortestPath}(a, b)$ can be expressed as the following equations:

average network diameter

$$= \frac{\sum_{a,b \in N \wedge a \neq b} \text{ShortestPath}(a, b)}{(|N| \times (|N| - 1))/2} \quad (2)$$

$$\text{ShortestPath}(a, b) = \text{Min}(\text{length}(\text{path}_i(a, b))) \quad (3)$$

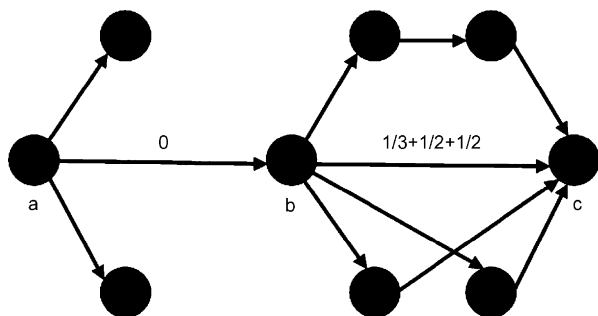


Figure 2 Link-weighted value calculating example. The link-weighted value $\text{weight}(a, b)$ of edge (a, b) is 0 while $\text{weight}(b, c)$ is $4/3$.

This definition implies clustering, with any increase in the number of possible paths resulting in an increase in the clustering degree between two linked nodes. Furthermore, the concepts and algorithms discussed in this paper are generalizable to non-directed networks.

To ensure that our method can be applied to any complex network, the link-weighted values calculated by our network motif detection method are derived from the number of all possible paths between two linked nodes within all network topological and local connection structures (no preset link quantity). This definition is similar to that of betweenness—that is, effects that result from the removal of a link from a network. Accordingly, we believe our proposed link-weighted value calculation method represents the importance of each link in a real network.

We also considered the interactive strength of each link in the quantitative real network, and attempted to validate our proposed weighted links by comparing them with quantitative links. However, interactive quantitative links are defined by functions that are category-specific (e.g., proteins, genes, species). It is difficult to specify the overall impact of these links on protein–protein interaction networks and food webs. For example, the number of links between tigers and wild oxen does not reflect the significance of their connection within an overall food web. Furthermore, each type of complex network has its own measure for interactive strength.

We used a switching algorithm (i.e., $A \rightarrow B$, $C \rightarrow D$ becomes $A \rightarrow D$, $C \rightarrow B$ if $A \rightarrow D$ and $C \rightarrow B$ do not exist) to create random networks according to any given degree sequence [7,28]. Results from previous studies indicate that these random networks have the same number of nodes and edges, as well as node in-degrees (incoming edges) and out-degrees (outgoing edges) that are identical to those of real networks. Furthermore, randomized

networks preserve the same number of appearances of all $(n - 1)$ node subgraphs as in the real (original) network [7]. The threshold that determines the strength of an edge (link) is the mean weighted value of all edges in a random network ensemble. Accordingly, we generated 1000 random networks to serve as a control. An edge was labeled “weak” when its weighted value in these or real networks was smaller than the threshold subtract a double standard deviation ($P = 0.01$); all other edges were labeled “strong”. Researchers can define criteria for strong and weak links according to their own needs. Finally, we located all possible motifs and compared their distributions in real and random networks.

We expanded Milo et al.’s methodology [7] for identifying bridge and brick motifs in complex networks to include the following steps:

1. Calculate the weighted value of each link in a network of interest and an ensemble of random networks to calculate the significance of n -node subgraphs. The purpose is to maintain the same number of appearances for all $(n - 1)$ node subgraphs as in the original network.
2. Label all weighted links in the network of interest and random network ensemble as strong or weak according to a benchmark of two standard deviations from the mean weighted value of all links in the ensemble. Links with weighted values below the benchmark are labeled as weak.
3. Identify all n -node bridge/brick subgraph types in the network of interest and random network ensemble.
4. Mark all n -node bridge/brick subgraph types by calculating their numbers in the network of interest and random network ensemble. Each n -node bridge/brick subgraph type is selected as a representative motif only if its frequency in the network of interest far exceeds its frequency in the ensemble.

These steps can assist researchers in their efforts to understand the functions and roles of identified motifs in a real network and to analyze the dynamic behaviors of complex networks. Regarding method robustness, our proposed approach emphasizes the global and local topological properties of each real network rather than the specific functions of different network types.

Motif frequency can be used to measure levels of similarity between two networks of interest. In addition, it is possible to calculate the Z_{score} for all bridge/brick motifs and significance profiles (SPs) in a network by expanding Milo et al.’s [7,28] methods. In Eq. (4), $Z_{\text{score}}(\text{Bridge}_i)$ represents

the statistical significance of the i th kind of bridge motif in a network:

$$Z_{\text{score}}(\text{Bridge}_i) = \frac{N_{\text{real}}(\text{Bridge}_i) - \langle N_{\text{random}}(\text{Bridge}_i) \rangle}{\text{STD}(N_{\text{random}}(\text{Bridge}_i))} \quad (4)$$

where $N_{\text{real}}(\text{Bridge}_i)$ represents the time of appearance of the i th type of bridge motif in a network and $\langle N_{\text{random}}(\text{Bridge}_i) \rangle$ and $\text{STD}(N_{\text{random}}(\text{Bridge}_i))$, respectively represent the mean and standard deviation of the time of appearance of the i th type of bridge motif in a randomized network ensemble. In Eq. (5), $\text{SP}(\text{Bridge}_i)$ is the vector of $Z_{\text{score}}(\text{Bridge}_i)$ normalized to a length of 1. The normalization emphasizes the relative significance of the i th type of bridge motif rather than the absolute significance. As shown in Eqs. (6) and (7), $Z_{\text{score}}(\text{Brick}_i)$ and $\text{SP}(\text{Brick}_i)$ can be derived in the same manner

$$\text{SP}(\text{Bridge}_i) = \frac{Z_{\text{score}}(\text{Bridge}_i)}{(\sum Z_{\text{score}}(\text{Bridge}_i)^2)^{1/2}} \quad (5)$$

$$Z_{\text{score}}(\text{Brick}_i) = \frac{N_{\text{real}}(\text{Brick}_i) - \langle N_{\text{random}}(\text{Brick}_i) \rangle}{\text{STD}(N_{\text{random}}(\text{Brick}_i))} \quad (6)$$

$$\text{SP}(\text{Brick}_i) = \frac{Z_{\text{score}}(\text{Brick}_i)}{(\sum Z_{\text{score}}(\text{Brick}_i)^2)^{1/2}} \quad (7)$$

4. Experiments

We performed validation experiments to confirm the definitions of weighted links and network motifs. Due to the links' non-directional characteristic, only two kinds of motifs were identified for the three-node scenarios: ID = 8 and ID = 13 (Table 1). We looked at four well-known types of theoretical complex networks with specific topological properties to validate our algorithm: regular, scale-free, random, and Watts and Strogatz's small-world (Table 2) [3]. Due to their small-world properties, we found more bridge than brick motifs in scale-free and random networks. Regular networks with a Moore neighborhood structure only contain brick motifs due to the structure's high clustering property (minus any shortcuts). Watts and Strogatz's small-world networks are formed by rewiring 1% of the links of regular networks containing only a few bridge motifs; when more than 5% of the links are rewired, bridge motifs outnumber brick motifs (Fig. 3). We therefore suggest that bridge motifs indicate the presence of small-world properties and brick motifs the presence of local clustering properties as follows:

1. *Regular*. We applied the Moore neighborhood concept to a two-dimensional lattice, with each node linked to its eight adjacent cells. For this type of network we found brick motifs only. To maintain the same in- and out-degree distributions in random and regular networks, individual nodes in random networks can link with any other cell except its eight adjacent cells. As clustering in a random network decreases, the threshold of the weighted value of its links also decreases. Therefore, all links in regular networks turn out to be strong (exclusively brick motifs).
2. *Scale-free*. Here the degree of distribution (i.e., the number of edges per node) obeys a long-tailed power-law distribution in which the majority of nodes have only a few links, but a small number of nodes have many links. We found that scale-free networks are composed of many bridge motifs and very few brick motifs consisting of nodes with high degrees of separation.
3. *Random*. As predicted, we failed to find a dominant motif during our comparison of 1000 random networks. Accordingly, random networks served as a successful null hypothesis for our algorithm.
4. *Small-world*. We rewired links in two-dimensional regular networks with Moore neighborhood structure using rewiring percentages of 0.01, 0.05, 0.1 and 0.5 of all links. In the 0.01 trial we found that some of the brick motifs became bridge motifs. As the rewiring percentage increased, the number of bridge motifs increased and number of brick motifs decreased. At a rewiring ratio of 1, small-world networks change into random networks. Brick motifs appear to play an important role in reducing the degree of separation and increasing the degree of clustering in scale-free networks.

Several engineering (electronic circuit) and social networks (Table 3) were used to demonstrate that our proposed motif detection method is both general-purpose and robust. We also compared the method to Milo et al.'s original method for complex network analysis. In electronic circuits consisting of digital fractional multipliers (data from an ISCA89 benchmark) [28], nodes represent logic gates. Flip-flops and edges represent directed electronic transmission paths. Experimental results indicate that the s208, s420, and s838 electronic circuit networks contain significant numbers of bridge motifs. Here the low degree of clustering is considered trivial because designers often try to simplify connection structures and numbers of electronic components [4]. The feedback bridge motif we identified (consisting of weak-tie links only) fulfills this requirement as described by Kundu et al. [38] (ID = 9)

Table 2 Bridge and brick subgraph frequencies in four complex network categories (for validation purposes)

Category	Nodes	Edges	Subgraph type	ID	N_{real}	$N_{random} \pm STD$	Z_{score}
Regular	900	7200	Bridge	8	0	24983.2 ± -39.0	-640.61
			Brick	8	14,400	40.8 ± 17.4	824.81
			Bridge	13	0	0.0 ± 0.0	0.00
			Brick	13	3,600	58.6 ± 8.2	430.13
Scale-free	900	1800	Bridge	8	4,355	4099.7 ± 53.7	4.75
			Brick	8	45	258.9 ± 47.1	-4.54
			Bridge	13	2	7.0 ± 2.6	-1.95
			Brick	13	0	8.8 ± 3.5	-2.54
Random	900	1800	Bridge	8	1,229	1226.1 ± 27.9	0.11
			Brick	8	537	536.9 ± 27.8	0.01
			Bridge	13	0	0.4 ± 0.7	-0.64
			Brick	13	0	0.6 ± 0.7	-0.81
WS small-world #1 (rewiring % = 0.01)	900	7200	Bridge	8	2,399	25029.7 ± 38.0	-595.92
			Brick	8	12,573	58.3 ± 18.6	674.60
			Bridge	13	320	7.6 ± 2.8	113.47
			Brick	13	3,111	51.5 ± 7.1	430.65
WS small-world #2 (rewiring % = 0.05)	900	7200	Bridge	8	9,434	24713.2 ± 73.5	-207.80
			Brick	8	8,100	656.3 ± 64.5	115.49
			Bridge	13	991	25.1 ± 5.3	182.65
			Brick	13	1,681	35.0 ± 6.1	268.59
WS small-world #3 (rewiring % = 0.10)	900	7200	Bridge	8	13,386	24047.0 ± 111.4	-95.69
			Brick	8	6,089	1519.0 ± 99.9	45.73
			Bridge	13	1,063	30.9 ± 4.9	209.23
			Brick	13	1,029	30.7 ± 5.6	179.17
WS small-world #4 (rewiring % = 0.50)	900	7200	Bridge	8	22,649	22935.4 ± 148.5	-1.93
			Brick	8	3,973	4244.9 ± 153.7	-1.77
			Bridge	13	213	56.2 ± 8.9	17.64
			Brick	13	47	17.7 ± 4.1	7.11

(Fig. 4, Table 3). As its name implies, the feedback bridge motif indicates the existence of a feedback structure without redundancy in the three named electronic circuits—again as proven by Kundu et al. [38]. That research team reported that redundant circuits seldom appear in simple electronic circuits such as s208, s420, and s838. However, they also

note that redundant wires and components are frequently added to more complex electronic circuits (e.g., s15850, s35932, s38417 and s38584) to prevent accidental system failures. The over-simplification of electronic circuits can result in large numbers of errors [4] or complete system breakdowns when one component fails. Accordingly, it is

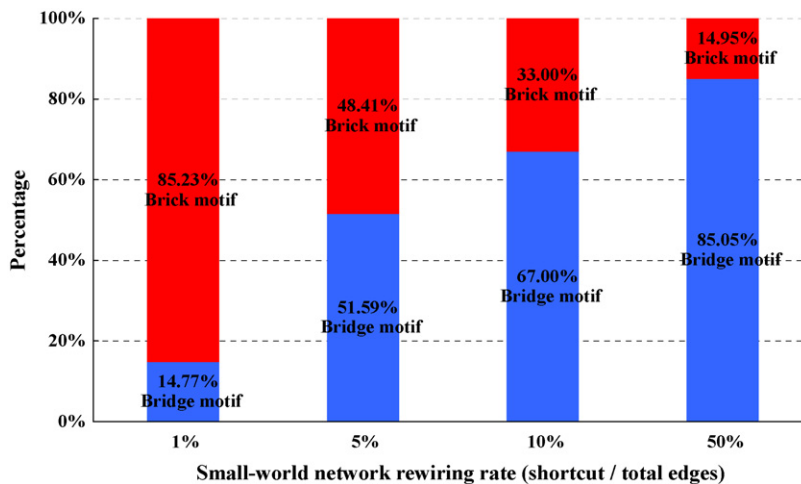


Figure 3 Percentages of bridge and brick motifs in small-world networks according to different rewiring ratios.

Table 3 Brick and bridge motifs in 14 real world networks, including edge and node definitions, network sizes, and references

Category	Common feature	Directed network	Nodes	Links	Motif type	ID	N_{real}	$N_{random} \pm STD$	Z_{score}
Gene regulation (transcription) [12,33]	Directed graph in which nodes represent genes and edges are directed from one gene to another, regulated by the transcription factor	<i>Escherichia coli</i>	424	519	Bridge	5	42	7.5 ± 3.1	11.14
		Yeast	688	1079	Bridge	5	67	13.8 ± 3.8	14.04
Social [31,32]	Directed graph in which nodes represent people and edges indicate friendships between two persons	Leader	32	96	Brick	7	38	22.1 ± 9.5	1.67
					Brick	11	5	1.5 ± 1.3	2.59
		Prisoner	67	182	Bridge	6	11	2.0 ± 1.4	6.42
					Brick	12	5	0.5 ± 0.7	6.26
Food webs [29]	Seven different ecosystems.	LittleRock	92	984	Bridge	11	93	41.3 ± 6.2	8.33
		Ythan	83	391	Bridge	2	1182	850.1 ± 86.0	3.86
	Directed graph in which nodes represent groups of species and edges connect predator and prey nodes	St. Martin	42	205	Bridge	5	244	180.4 ± 20.0	3.18
		Chesapeake	31	67	Bridge	5	21	11.2 ± 4.0	2.42
	Coachella	29	243	Bridge	2	275	192.5 ± 14.8	5.57	
				Bridge	4	252	110.3 ± 15.1	9.38	
				Bridge	6	110	68.1 ± 5.3	7.84	
				Bridge	13	10	6.2 ± 1.4	2.83	
				Bridge	2	181	140.1 ± 11.3	3.63	
				Bridge	4	234	115.2 ± 33.4	3.56	
	Skipwith	25	189	Bridge	2	181	140.1 ± 11.3	3.63	
			Bridge	4	234	115.2 ± 33.4	3.56		
	B. Brook	25	104	Bridge	1	266	123.5 ± 31.2	4.57	
			Bridge	2	181	103.1 ± 23.1	3.37		
Electrical circuits [4]	ISCAS89 benchmark set of sequential logic electronic circuits. Directed graph in which nodes represent logic gates and flip-flops	s208	122	189	Bridge	9	10	0.90 ± 1.0	9.23
		s420	252	399	Bridge	9	20	0.9 ± 0.9	20.13
		s838	512	819	Bridge	9	40	0.9 ± 1.3	30.2

necessary to add an appropriate level of redundancy as a means of bypassing failed components or substituting for the original path [4,38]. Strong-tie links represent alternative paths and weak-tie links represent simplified electronic circuits. Combined, simplification and duplication help prevent unexpected system breakdowns.

In the two social networks we analyzed, nodes represent individuals in a group and edges represent positive sentiments directed from one group member to another based on responses to questionnaire items. We found similar characteristics between two networks, one consisting of prison inmates ($N = 67$ nodes and $E = 110$ edges) and the other

college students in a leadership course ($N = 32$ and $E = 96$). The inmates responded to the question “Who are your closest friends in your cellblock?” The students were asked to name three classmates they would invite to serve on a committee (correlation coefficient $c = 0.92-0.96$ [31–32]). According to Milo et al.’s [7,28] methods, both social networks belong to the same superfamily. Strong similarities between the two networks were also identified according to the triad significance profile (TSP) of bridge motifs ($c = 0.92$) (Fig. 5, Table 3), but not according to the TSP of brick motifs ($c = 0.6$) (Fig. 6, Table 3). We also found a significantly higher number of bridge motifs (i.e., more “nodding acquaint-

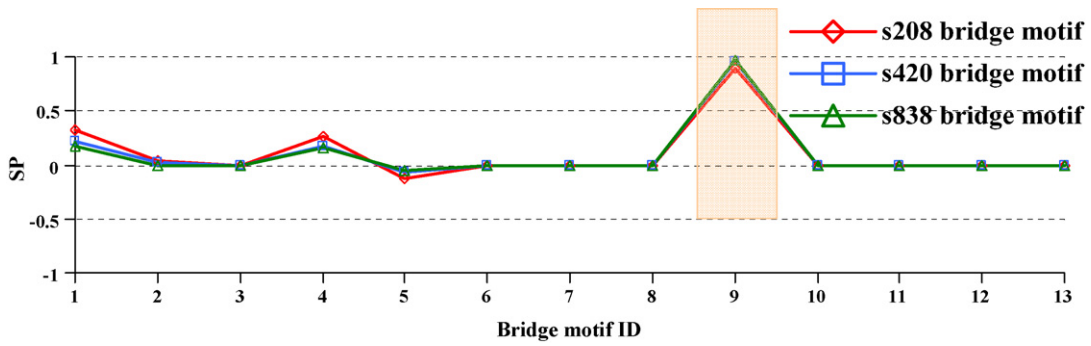


Figure 4 Bridge motif ratio profiles for three electrical circuits (s208, s420, and s838).

tances”) in the prisoner network. The significantly larger number of brick motifs in the leadership class network indicates that small, strong groups are easily formed. The bridge and brick motifs can be used to further analyze network topological structures, functions, and differences.

We applied our proposed method to several biochemistry (transcriptional gene regulation) and ecology (food web) networks to identify bridge and brick network motifs. Networks and sources are listed in Table 3. All data and programs (including source code) are available online at <http://ftp://www.csie.cgu.edu.tw/bbm/>.

In gene regulation networks for one bacteria (*Escherichia coli*) and one eukaryote (the yeast *Saccharomyces cerevisiae*) [28], each node represents a gene or operon that encodes a transcription

factor (TF); edges denote the TFs themselves. Many TFs are encoded within operons, therefore directed links represent direct transcriptional modulation from a TF to an operon or from a TF-contained operon to another operon [28]. In both networks we found more bridge than brick subgraphs (they are not called motifs until they reach statistical significance). Furthermore, the two transcription networks had the same feed-forward bridge motif (ID = 5), indicating that the transcription networks have, at minimum, non-replaceable interactions without intermediate interactions with other genes (Figs. 1 and 7; Table 3). We suggest that the weak-tie link that provides a unique path for controlling the signal exerts a significant impact on the signal processing function of transcription networks [13,28]. We also tried to learn more about the relationship

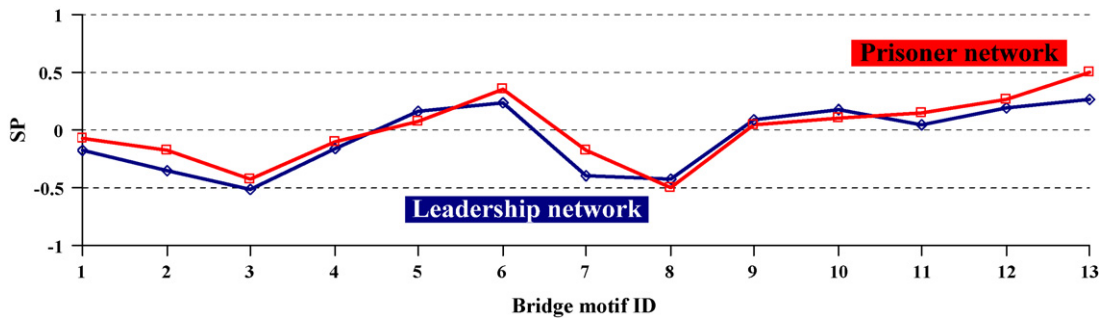


Figure 5 Bridge motif ratio profiles for two social networks.

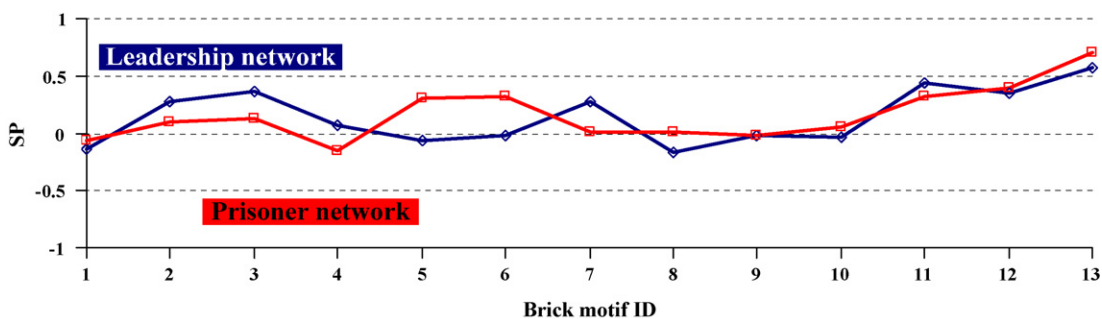


Figure 6 Brick motif ratio profiles for two social networks.

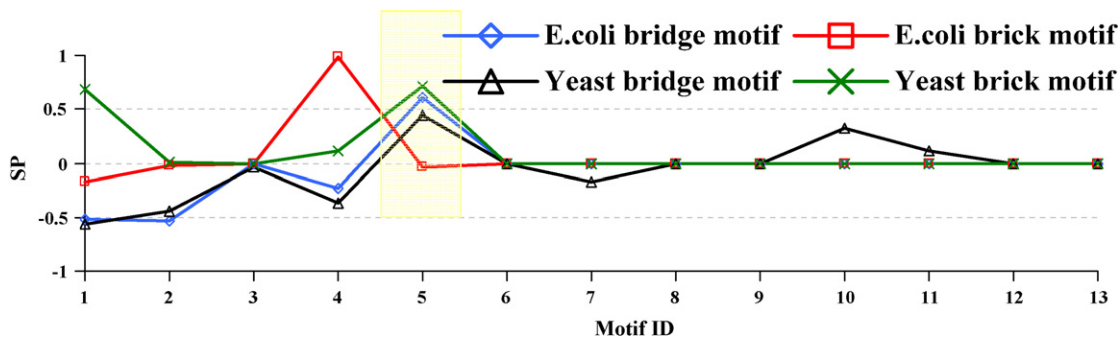


Figure 7 Brick-bridge motif ratio profiles for two regulation networks (one bacteria and one eukaryote).

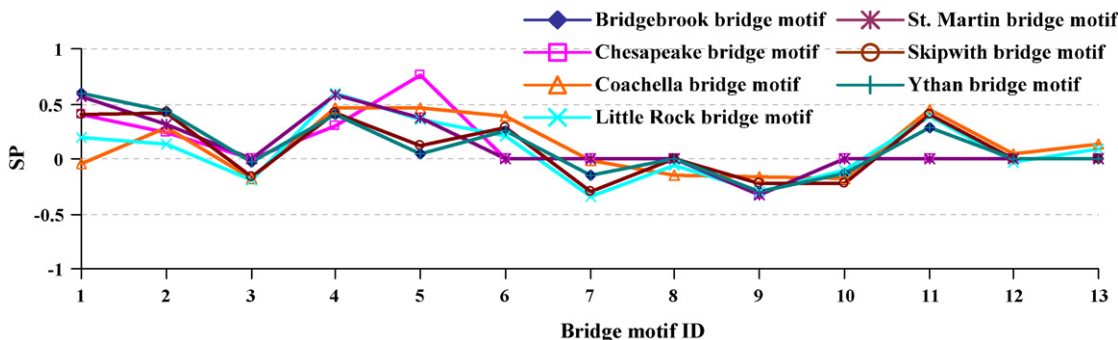


Figure 8 Bridge motif ratio profiles for seven food webs.

between coherent (incoherent) FFLs and brick (bridge) FFLs and identified *E. coli*'s 34 coherent and 8 incoherent FFLs (Table 3) [39]. Therefore, differences in coherent (incoherent) FFL frequencies cannot be explained simply in terms of the relative abundances of bridge and brick motifs in a network.

In the seven food webs we analyzed [29], nodes represent groups of species and edges connecting predator and prey nodes. Recent studies have shown that strong interactions (similar to our definition of weak-tie links) between two consecutive levels of a trophic chain have a significant effect on food web stability and dynamics [17,30]. A strong interaction indicates a strong predator preference for one prey species and a low potential for intermediate species—a phenomenon that supports our claim that weak-tie links exert certain impacts on food webs. Also in the seven food webs, the numbers of bridge motifs were significantly higher than the numbers of brick motifs, especially feedback (ID = 5) and three-point chains (ID = 2) (Fig. 8, Table 3). This confirms Jordi's [8] claim that these two motifs exert significant impacts on ecosystem food webs. We believe the reason why ecosystems' containing these two types of bridge motifs easily become unbalanced is because they have many weak links—in other words, it is difficult to find substitute nodes or links for the purpose of preserving ecosystem stability.

5. Conclusion

According to our definitions of weighted links and network motifs and the results of our validation experiments using theoretical complex networks, we conclude that the presence of bridge and brick motifs in a network is closely associated with network topological (especially local connection) structures, but not with network size (i.e., number of nodes). In summary, we tested three experimental predictions to verify the importance and function of bridge and brick network motifs: (a) regular networks with Moore neighborhood structures only contain brick motifs due to the structure's strong clustering property; (b) as rewiring percentages increase in regular networks, the number of bridge motifs increases and the number of brick motifs decreases, with the rewiring process contributing to the formation of networks that exhibit small-world and clustering properties; and (c) the combination of more bridge motifs and fewer brick motifs means that a network is less prone to cluster formation.

Our proposed method combines two measures, each with its own merits (i.e., determining the topological properties of links in real networks and identifying statistically significant motifs in real networks). The combined measures can be used to explore the functions and roles of real network

motifs. To locate statistically significant network motifs, Milo et al. [1] proposed comparing the real network in question with suitably randomized networks, then selecting patterns (subgraphs) that appear at significantly higher frequencies in the real network. Compared to Milo et al.'s approach, our method simultaneously detects global features and local structures in complex networks and locates functionally and statistically significant network motifs. We believe that the combination of these two methods can (a) assist in locating motifs; (b) help researchers find clusters between bridge motifs and within the brick motifs of complex networks for the purpose of identifying real network functions, behaviors, and similarities; and (c) provide global and local views of the real network in question. Most network motif functions can be identified via network topological structures. Combining a motif structure with its function can help identify complex network properties. Motifs with special topological structures reveal the global features of real networks and significant local structural patterns. This information can help researchers working with design principles and network evolution.

References

- [1] Newman MEJ. The structure and function of complex networks. *SIAM Rev* 2003;4:167–256.
- [2] Newman MEJ, Watts DJ. Scaling and percolation in the small-world network model. *Phys Rev E* 1999;60:7332–42.
- [3] Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. *Nature* 1998;393:440–2.
- [4] Cancho RFI, Janssen C, Sole RV. Topology of technology graphs: small world patterns in electronic circuits. *Phys Rev E* 2001;6404:046119.
- [5] Barrat A, Weigt M. On the properties of small-world network models. *Eur Phys J: B* 2000;13:547–60.
- [6] Strogatz SH. Exploring complex networks. *Nature* 2001;410:268–76.
- [7] Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. Network motifs: simple building blocks of complex networks. *Science* 2002;298:824–7.
- [8] Artzy-Randrup Y, Fleishman SJ, Ben-Tal N, Stone L. Comment on "network motifs: simple building blocks of complex networks" and "superfamilies of evolved and designed networks". *Science* 2004;305:1107.
- [9] Goh KI, Oh E, Jeong H, Kahng B, Kim D. Classification of scale-free networks. *Proc Natl Acad Sci USA* 2002;99:12583–8.
- [10] Vazquez A, Dobrin R, Sergi D, Eckmann JP, Oltvai ZN, Barabasi AL. The topological relationship between the large-scale attributes and local interaction patterns of complex networks. *Proc Natl Acad Sci USA* 2004;101:17940–5.
- [11] Moon HS, Bhak J, Lee KH, Lee D. Architecture of basic building blocks in protein and domain structural interaction networks. *Bioinformatics* 2005;21:1479–86.
- [12] Shen-Orr SS, Milo R, Mangan S, Alon U. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat Genet* 2002;31:64–8.
- [13] Mangan S, Alon U. Structure and function of the feed-forward loop network motif. *Proc Natl Acad Sci USA* 2003;100:11980–5.
- [14] Berg J, Lassig M. Local graph alignment and motif search in biological networks. *Proc Natl Acad Sci USA* 2004;101:14689–94.
- [15] Middendorf M, Ziv E, Wiggins CH. Inferring network mechanisms: the *Drosophila melanogaster* protein interaction network. *Proc Natl Acad Sci USA* 2005;102:3192–7.
- [16] Alon U. Biological networks: the tinkerer as an engineer. *Science* 2003;301:1866–7.
- [17] Bascompte J, Melian CJ, Sala E. Interaction strength combinations and the overfishing of a marine food web. *Proc Natl Acad Sci USA* 2005;102:5443–7.
- [18] Rice JJ, Kershenbaum A, Stolovitzky G. Lasting impressions: motifs in protein–protein maps may provide footprints of evolutionary events. *Proc Natl Acad Sci USA* 2005;102:3173–4.
- [19] Hethcote HW. The mathematics of infectious diseases. *SIAM Rev* 2000;42:599–653.
- [20] Eckmann JP, Moses E. Curvature of co-links uncovers hidden thematic layers in the World Wide Web. *Proc Natl Acad Sci USA* 2002;99:5825–9.
- [21] Manrubia SC, Poyatos JF. Motif selection in a model of evolving replicators: the role of surfaces and limited transport in network topology. *Europhys Lett* 2003;64:557–63.
- [22] Mangan S, Zaslaver A, Alon U. The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. *J Mol Biol* 2003;334:197–204.
- [23] Granovetter MS. The strength of weak ties: a network theory revisited. *Sociological theory. Sociol Theory* 1983;1:201–33.
- [24] Davidsen J, Ebel H, Bornholdt S. Emergence of a small world from local interactions: modeling acquaintance networks. *Phys Rev Lett* 2002;88:128701.
- [25] Girvan M, Newman MEJ. Community structure in social and biological networks. *Proc Natl Acad Sci USA* 2002;99:7821–6.
- [26] Newman MEJ, Girvan M. Finding and evaluating community structure in networks. *Phys Rev E* 2004;69:026113.
- [27] Zhu H, Huang ZX. Navigation in a small world with local information. *Phys Rev E* 2004;70:036117.
- [28] Milo R, Itzkovitz S, Kashtan N, Levitt R, Shen-Orr S, Ayzenshtat I, et al. Superfamilies of evolved and designed networks. *Science* 2004;303:1538–42.
- [29] Williams RJ, Martinez ND. Simple rules yield complex food webs. *Nature* 2000;404:180–3.
- [30] Shurin JB, Borer ET, Seabloom EW, Anderson K, Blanchette CA, Broitman B, et al. A cross-ecosystem comparison of the strength of trophic cascades. *Ecol Lett* 2002;5:785–91.
- [31] MacRae Jr D. Direct factor analysis of sociometric data. *Sociometry* 1960;23:360–71.
- [32] Zeleny LD. Adaptation of research findings in social leadership to college classroom procedures. *Sociometry* 1950;13:314–28.
- [33] Costanzo MC, Crawford ME, Hirschman JE, Kranz JE, Olsen P, Robertson LS, et al. WormPD: model organism volumes of the BioKnowledge library, an integrated resource for protein information. *Nucleic Acids Res* 2001;29:75–9.
- [34] Hartwell LH, Hopfield JJ, Leibler S, Murray AW. From molecular to modular cell biology. *Nature* 1999;402:C47–52.
- [35] Han JDJ, Bertin N, Hao T, Goldberg DS, Berriz GF, Zhang LV, et al. Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* 2004;430:88–93.
- [36] Stelling J, Sauer U, Szallasi Z, Doyle III JZ, Doyle J. Robustness of cellular functions. *Cell* 2004;118:675–85.

- [37] Yeager-Lotem E, Sattath S, Kashtan N, Itzkovitz S, Milo R, Pinter RY, et al. Network motifs in integrated cellular networks of transcription-regulation and protein-protein interaction. *Proc Natl Acad Sci USA* 2004;101:5934–9.
- [38] Kundu S, Huisman LM, Nair I, Iyengar V, Reddy LN. A small test generator for large designs. In: *Proceedings of International Test Conference*, vol. 30; 1992. p. 30–40.
- [39] Dobrin R, Beg QK, Barabasi AL, Oltvai ZN. Aggregation of topological motifs in the *Escherichia coli* transcriptional regulatory network. *BMC Bioinform* 2004;5:10.
- [40] Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet* 2004;5:101–13.
- [41] Albert R. Scale-free networks in cell biology. *J Cell Sci* 2005;118:4947–57.
- [42] Barabasi AL, Albert R. Emergence of scaling in random networks. *Science* 1999;286:509–12.
- [43] Babu MM, Luscombe NM, Aravind L, Gerstein M, Teichmann SA. Structure and evolution of transcriptional regulatory networks. *Curr Opin Struct Biol* 2004;14:283–91.
- [44] Teichmann SA, Babu MM. Gene regulatory network growth by duplication. *Nat Genet* 2004;36:492–6.