

Bridge and brick motifs in complex networks

Chung-Yuan Huang^{a,*}, Chuen-Tsai Sun^b, Chia-Ying Cheng^b, Ji-Lung Hsieh^b

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

^b*Department of Computer Science, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 300, Taiwan, Republic of China*

Received 2 September 2006; received in revised form 31 October 2006

Available online 1 December 2006

Abstract

Acknowledging the expanding role of complex networks in numerous scientific contexts, we examine significant functional and topological differences between bridge and brick motifs for predicting network behaviors and functions. After observing similarities between social networks and their genetic, ecological, and engineering counterparts, we identify a larger number of brick motifs in social networks and bridge motifs in the other three types. We conclude that bridge and brick motif content analysis can assist researchers in understanding the small-world and clustering properties of network structures when investigating network functions and behaviors.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Small-world properties; Clustering; Network motif; Complex network; Strong/weak-tie link

1. Introduction

Three global features in complex networks have been identified and investigated: highly clustered connections in which nodes have many mutual neighbors [1,3,6]; the small-world properties of short paths between any two nodes [2–6]; and long-tailed (power-law) degree distributions in which many nodes have only one or very few connections and a few nodes have a much larger than mean number of connections [1,7–9]. Some local structural motifs reveal their own statistically significant patterns when compared with motifs in random [7,8,10–15], biological [7,11,15,16], food web [7,17], and electronic circuit networks [7,16,18]—all of which are thought to contain important information. It is necessary to identify functionally and statistically important motifs for exploring and predicting network behaviors and functions [1–3,6–8,15,19,20]. Instead of analyzing functions according to their real network features and effects [10], researchers have shown a tendency to analyze them in terms of their theoretical pattern structures [7,12,15,21,22]. Here we propose an algorithm to perform two tasks: to simultaneously detect global features and local structures in complex networks, and locate functionally and statistically significant network motifs.

*Corresponding author. Tel.: +886 3 2118800x3474; fax: +886 3 2118700.

E-mail addresses: gscott@mail.cgu.edu.tw (C.-Y. Huang), ctsun@cs.nctu.edu.tw (C.-T. Sun), gis92566@cs.nctu.edu.tw (C.-Y. Cheng), gis91572@cs.nctu.edu.tw (J.-L. Hsieh).

When considering the global features and local structures of complex networks, it is worth noting that link properties (or weights) strongly impact network functions and behaviors [1,17,23,24]. 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 believe network motifs can be separated into two categories: bridge motifs (composed of weak links only or at least one weak link and multiple strong links) and brick motifs (composed of strong links only) (Table 1, Fig. 1). Whereas bridge motifs connect clusters and reduce the average degree of separation (network diameter), brick motifs reveal the local clustering phenomenon in complex networks [1].

2. Backgrounds: simple network motifs and superfamilies

Although small-world properties and long-tailed degree distributions are common statistical features of complex networks, at times they have very different local structures. It is therefore important to understand 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 [28–30], with independent tasks performed by modules consisting of single groups of physically or functionally connected nodes [28]. Many real networks contain large numbers of modules—for instance, groups of coworkers in a corporation or World Wide Web hyperlinks on similar topics.

Understanding complex network modularity requires appropriate measurement tools and the identification of modular relationships within a given real network. However, module identification is considered difficult due to contradictions between the power-law degree distributions of small-world networks and modularity concepts. According to the definition of a module, their presence infers isolated clusters of nodes within a network [28]. However, most nodes in scale-free networks are connected to central or core nodes, which affects the appearance of isolated modules.

Recent studies suggest that motifs can serve as basic modular elements in complex networks [7,12,31]. 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. From a micro-viewpoint, each subgraph denotes a special mutual connection pattern in a given network. However, multiple subgraphs in real networks may not hold equal

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

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

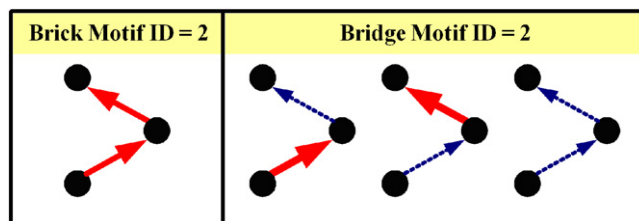


Fig. 1. The three-point chain motif can be divided into two categories: a three-point chain brick motif composed of two strong links (left box) and three-point chain bridge motifs composed of at least one weak link and a maximum of one strong link (middle box).

importance—for instance, some subgraphs appear much more frequently than others in complex networks. These statistically significant subgraphs are called *motifs* [7].

Randomized networks must be generated for purposes of comparison in order to determine if a subgraph i in a real network is a motif [7,32]. Using a directed network as an example, an algorithm for generating randomized networks takes real networks and repeatedly switches selected connection pairs until the entire network is sufficiently randomized (e.g., the connections $X_1 \rightarrow Y_1$ and $X_2 \rightarrow Y_2$ are replaced by $X_1 \rightarrow Y_2$ and $X_2 \rightarrow Y_1$). No further changes are made if at least one of the resulting connections already exists in the network. Each node in the randomized network has the same out-degree and in-degree values as those of its corresponding node in the real network.

Based on the motif detection algorithm proposed by Milo et al. [7], three conditions must be met in order to identify a subgraph i as a motif:

1. It is highly unlikely that the average number of appearances of subgraph i in the corresponding random network ensemble exceeds that of the real network. The possibility should be below a threshold of $P = 0.01$.
2. The number of appearances of subgraph i in the real network $N_{real}(Subgraph_i)$ should be greater than or equal to a lower bound $U = 4$.
3. The $N_{real}(Subgraph_i)$ number of appearances of subgraph i in the real network should exceed the average $\langle N_{random}(Subgraph_i) \rangle$ number of appearances of subgraph i in the random network ensemble. This requirement is expressed as $N_{real}(Subgraph_i) > 1.1 \times \langle N_{random}(Subgraph_i) \rangle$.

As shown in Fig. 2, a feed-forward motif (ID = 5) in a given real network consisting of 16 nodes and 19 edges can be identified via comparisons with four corresponding randomized networks. It should be noted that the number of different motifs is exponential to the number of nodes in a motif. Using directed subgraphs as an example, there are 13 possible three-node subgraphs (Table 1) and 199 possible four-node subgraphs. When dealing with large subgraphs, researchers must take care to avoid problems associated with combination explosions. One way to avoid such problems is to extract clusters of highly connected nodes from the real network and combine each cluster with its corresponding functionality.

Since small-world properties and long-tailed degree distributions are common global features of most real networks, it may be possible that different networks share similar local structures. Network motifs can help people understand the design principles and evolutionary mechanisms of complex networks from a local structure point of view. Based on the original simple motif concept, Milo et al. proposed a method for comparing local structures among different complex networks according to their significance profiles (SPs)

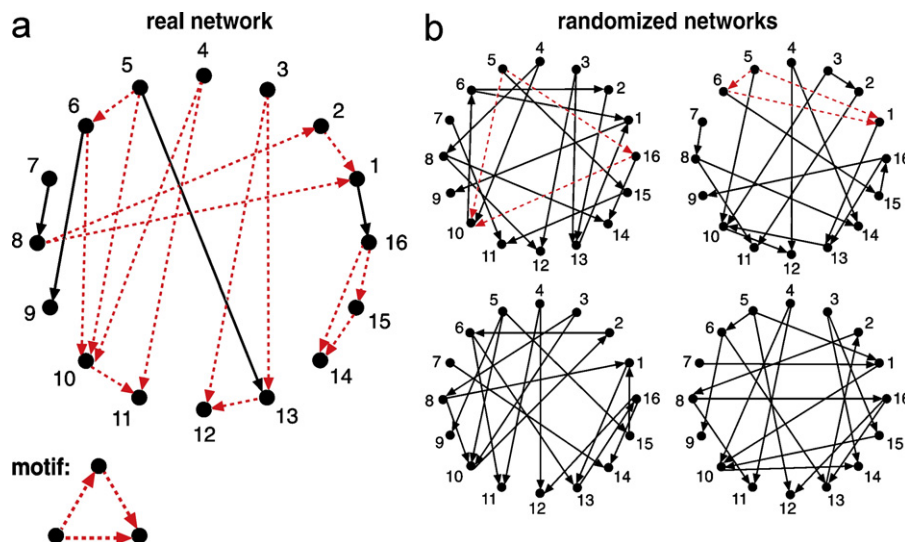


Fig. 2. Network motif detection example from [7].

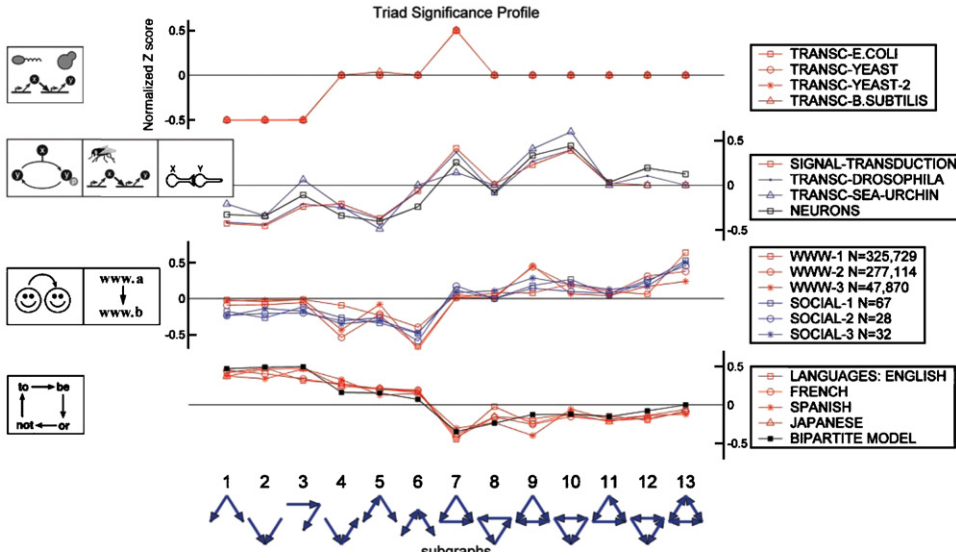


Fig. 3. Triad significance profile (TSP) for networks from various disciplines [32].

[32]. Eq. (1) is used to compute the statistical importance of a subgraph i in a given real network

$$Z_{Score}(Subgraph_i) = \frac{N_{real}(Subgraph_i) - \langle N_{random}(Subgraph_i) \rangle}{STD(N_{random}(Subgraph_i))}, \quad (1)$$

where $N_{real}(Subgraph_i)$ and $N_{random}(Subgraph_i)$ represent the number of appearances of subgraph i in the real and randomized networks, respectively, and $\langle N_{random}(Subgraph_i) \rangle$ and $STD(N_{random}(Subgraph_i))$ represent the mean and standard deviation of $N_{random}(Subgraph_i)$. As shown in Eq. (2), $SP(Subgraph_i)$ can be calculated by normalizing $Z_{Score}(Subgraph_i)$:

$$SP(Subgraph_i) = \frac{Z_{Score}(Subgraph_i)}{(\sum Z_{Score}(Subgraph_i)^2)^{1/2}}. \quad (2)$$

The purpose of normalization is to emphasize the relative significance of each subgraph. Since motifs in large networks usually have higher Z_{Score} values than those in small networks, normalization is an important operation for networks with different scales. The method can be used to classify real networks in which similar SP vector values belong to the same superfamily (Fig. 3). 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 [7,32]. At the same time, a network superfamily may contain networks with vastly different scales and functions.

3. Bridge and brick motifs detecting methods

As shown in Eq. (3), 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.

$$weight(a, b) = \sum_i \frac{1}{length(path_i(a, b))}, \quad (3)$$

where $path_i(a, b) \neq edge(a, b)$ and $length(path_i(a, b)) \leq average\ network\ diameter$. The length of one path represents its total number of nodes. The *average network diameter* and *ShortestPath(a, b)* can be expressed as

$$average\ network\ diameter = \frac{\sum_{a,b \in N \wedge a \neq b} ShortestPath(a, b)}{|N| \times (|N| - 1) / 2}, \quad (4)$$

$$ShortestPath(a, b) = Min(length(path_i(a, b))). \quad (5)$$

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. This definition is similar to that of *betweenness*—that is, effects resulting from the removal of a link from a network. Furthermore, the concepts and algorithms discussed in this paper are generalizable to non-directed networks.

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 [7,32]. Results from previous studies indicate that these random networks have the same number of nodes and edges and identical node in-degrees (incoming edges) and out-degrees (outgoing edges) as real networks. Furthermore, the randomized networks preserve the same number of appearances of all $(n - 1)$ node subgraphs as in the real network [7]. The threshold that determines the strength of an edge (link) is the mean weighted value of all edges in 1000 random networks. Accordingly, we generated 1000 random networks to serve as a control. When the weighted value of an edge in these or real networks was smaller than the threshold plus a double standard deviation, the edge was labeled “weak”; 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 random network 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 the frequency in the ensemble.

Motif frequency can be used to measure the similarity of two networks of interest. In addition, it is possible to calculate the Z_{score} for all bridge/brick motifs and SPs in a network by expanding Milo et al.’s [7,32] methods. In Eq. (6), $Z_{score}(Bridge_i)$ represents the statistical significance of the i th kind of bridge motif in a network:

$$Z_{Score}(Bridge_i) = \frac{N_{real}(Bridge_i) - \langle N_{random}(Bridge_i) \rangle}{STD(N_{random}(Bridge_i))}, \quad (6)$$

where $N_{real}(Bridge_i)$ represents the time of appearance of the i th type of bridge motif in a network and $\langle N_{random}(Bridge_i) \rangle$ and $STD(N_{random}(Bridge_i))$, 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. (7), the $SP(Bridge_i)$ is the vector of $Z_{score}(Bridge_i)$ 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. (8) and (9), $Z_{score}(Brick_i)$ and $SP(Brick_i)$ can be derived in the same manner.

$$SP(Bridge_i) = \frac{Z_{Score}(Bridge_i)}{(\sum Z_{Score}(Bridge_i)^2)^{1/2}}, \quad (7)$$

$$Z_{Score}(Brick_i) = \frac{N_{real}(Brick_i) - \langle N_{random}(Brick_i) \rangle}{STD(N_{random}(Brick_i))}, \quad (8)$$

$$SP(Brick_i) = \frac{Z_{Score}(Brick_i)}{(\sum Z_{Score}(Brick_i)^2)^{1/2}}. \quad (9)$$

Since it is more difficult to replace a weak link than a strong one, removing all bridge motif types will result in a significant increase in the average degree of separation. In other words, bridge motifs contribute to reducing a network's average degree of separation and revealing its small-world properties. Bridge motifs also connect high-density clusters or low-density clusters that serve as a network's foundation. Since brick motifs consist of strong links only, removing them has little effect on the average degree of separation.

4. Results

We applied our proposed method to several biochemistry (transcriptional gene regulation), ecology (food webs), engineering (electronic circuits) and social networks to identify network bridges and brick motifs (Table 2). All networks and sources are listed in Table 3. All data and programs (including source code) are available at ftp://www.csie.cgu.edu.tw/~gis89802/bridge_and_brick_motifs/data_and_programs.zip.

In gene regulation networks for one bacteria (*Escherichia coli*) and one eukaryote (the yeast *Saccharomyces cerevisiae*) [32], each node represents a gene and each edge connects a gene that encodes a transcription factor protein to a gene that is transcriptionally regulated by that factor. Our algorithm-generated results indicate that the two transcription networks have the same feed-forward bridge motif (ID = 5) (Fig. 4 and Table 3), the appearance of which indicates that the transcription networks have at minimum non-replaceable interactions without intermediate interactions with other genes. 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 [13,32].

In electronic circuits consisting of digital fractional multipliers (data from an ISCA89 benchmark) [32], nodes represent logic gates and flip-flops and edges represent directed electronic transmission paths. Experimental results indicate that 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

Table 2
Descriptions of 14 complex networks, including edge and node definitions, network sizes, and references

Network type	Common feature	Directed network	Nodes	Edges	Description
Gene regulation (transcription)	Directed graph in which nodes represent genes and edges are directed from one gene to another, regulated by the transcription factor.	<i>E. coli</i>	424	519	<i>Escherichia coli</i> [12]
		Yeast	685	1052	<i>Saccharomyces cerevisiae</i> [33]
Social	Directed graph in which nodes represent people and edges indicate friendships between two persons.	Leader	32	96	College students [34]
		Prisoner	67	182	Prisoners [35]
Food webs	Directed graph in which nodes represent groups of species and edges connect predator and prey nodes.	Little Rock	92	984	Seven different ecosystems [36]
		Ythan	83	391	
		St. Martin	42	205	
		Chesapeake	31	67	
		Coachella	29	243	
		Skipwith	25	189	
Electrical networks	Directed graph in which nodes represent logic gates and flip-flops.	s208	122	189	ISCAS89 benchmark set of sequential logic electronic circuits [4]
		s420	252	399	
		s838	512	819	

Table 3
Brick and bridge motifs in fourteen real networks

Category	Network	Nodes	Links	Motif type	ID	N_{Real}	$N_{Random} \pm STD$	Z_{Score}
Biochemistry	<i>E. 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 network	Leader	32	96	Brick	11	5	1.5 ± 1.3	2.59
				Brick	7	38	22.1 ± 9.5	1.67
	Prisoner	67	182	Bridge	6	11	2.0 ± 1.4	6.42
				Brick	12	7	1.5 ± 1.2	4.53
Food webs	Little Rock	92	984	Bridge	11	93	41.3 ± 6.2	8.33
	Ythan	83	391	Bridge	2	1182	850.1 ± 86.0	3.86
	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 ± 14.1	5.57
				Bridge	4	252	110.3 ± 15.1	9.38
				Bridge	6	110	68.1 ± 5.3	7.84
	Skipwith	25	189	Brick	13	10	6.2 ± 1.4	2.83
				Bridge	2	181	140.1 ± 11.2	3.63
				Bridge	4	234	115.2 ± 33.4	3.56
Bridge				2	181	103.1 ± 23.2	3.37	
B. Brook	25	104	Bridge	6	266	123.5 ± 31.2	4.57	
			Bridge	9	10	1.0 ± 1.0	8.73	
			Bridge	9	20	0.9 ± 0.9	20.13	
Electrical circuits	s208	122	189	Bridge	9	10	1.0 ± 1.0	8.73
	s420	252	399	Bridge	9	20	0.9 ± 0.9	20.13
	s838	512	819	Bridge	9	40	0.9 ± 1.3	30.2

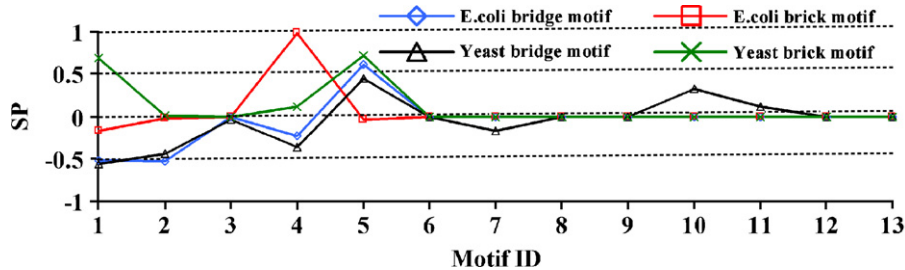


Fig. 4. Brick-bridge motif ratio profiles for two regulation networks (one bacteria and one eukaryote).

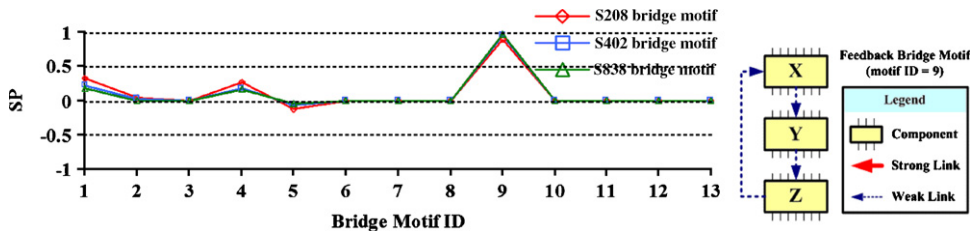


Fig. 5. Bridge motif ratio profiles for three engineering networks (s208, s420 and s838).

the connection structure and number 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. [37] (ID = 9) (Fig. 5, Table 3). However, the over-simplification of electronic circuits can result in large numbers of errors [4] and complete system breakdowns when one component fails. Accordingly, it is necessary to add an appropriate level of redundancy as a means of bypassing failed components or substituting for the original path [4,37].

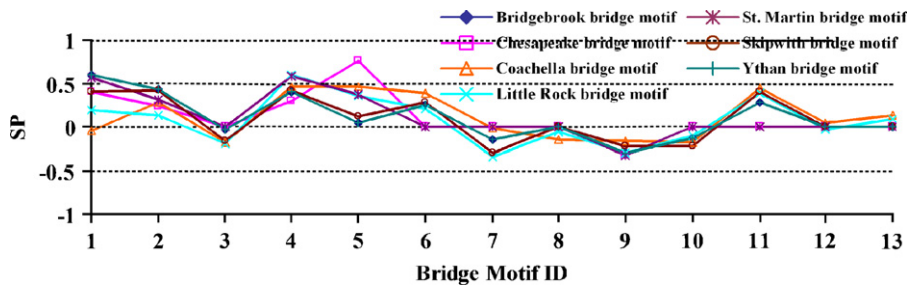


Fig. 6. Bridge motif ratio profiles for seven food webs.

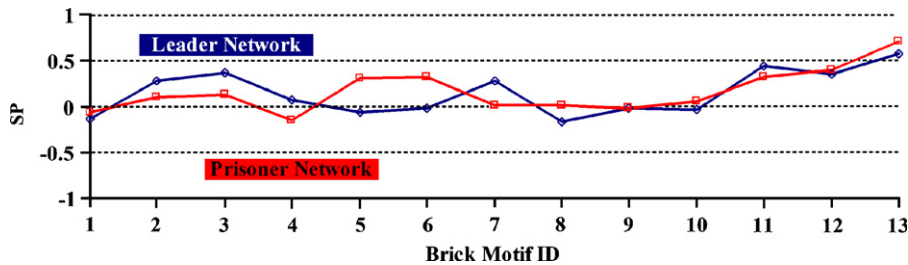


Fig. 7. Brick motif ratio profiles for two social networks.

The strong-tie links represent alternative paths and the weak-tie links represent simplified electronic circuits. Combined, the simplification and duplication helps prevent unexpected system breakdowns.

In the seven food webs we analyzed [36], nodes represent groups of species and edges connect predator and prey nodes. Recent studies have shown that strong interactions between two consecutive levels of a trophic chain have a significant effect on food web stability and dynamics [17,38]. 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. 6 and Table 3). This result 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 kinds of bridge motifs easily lose their balance 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.

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 the 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 on the tier?” The students identified three classmates they would want to have serve on a committee (correlation coefficient $c = 0.92$ to 0.96 [34,35]). According to Milo et al.'s [7,32] 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 ($c = 0.92$), but not according to the TSP of brick motifs ($c = 0.6$) (Fig. 7 and Table 3). We also found a significantly higher number of bridge motifs (i.e., more “nodding acquaintances”) in the prisoner network. The significantly larger number of brick motifs in the leadership class network indicates that small and strong groups are easily formed. The bridge and brick motifs can be used to further analyze network topological structures, functions, and differences.

5. Validation

Due to the links' non-directional characteristic, only two kinds of motifs (ID = 8, 13) were identified for the three-node scenarios. We looked at four types of theoretical complex networks with specific topological

properties to validate our network motif detection algorithm: regular, scale-free, random, and Watts and Strogatz's small-world (Table 4) [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. 8). We therefore conclude 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 on 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).

Table 4
Bridge and brick motif frequencies in four complex network categories (for validation purposes)

Category	Nodes	Edges	Motif type	ID	N_{Real}	$N_{Random} \pm STD$	Z_{Score}
Regular	900	7200	Bridge	8	0	24983.2 ± 39.0	-640.61
			Brick	8	14400	40.8 ± 17.4	824.81
			Bridge	13	0	0.0 ± 0.0	0.00
			Brick	13	3600	58.6 ± 8.2	430.13
Scale-free	900	1800	Bridge	8	4355	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	1229	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	2399	25029.7 ± 38.0	-595.92
			Brick	8	12573	58.3 ± 18.6	674.60
			Bridge	13	320	7.6 ± 2.8	113.47
			Brick	13	3111	51.5 ± 7.1	430.65
WS Small-world #2 (rewiring % = 0.05)	900	7200	Bridge	8	9434	24713.2 ± 73.5	-207.80
			Brick	8	8100	656.3 ± 64.5	115.49
			Bridge	13	991	25.1 ± 5.3	182.65
			Brick	13	1681	35.0 ± 6.1	268.59
WS Small-world #3 (rewiring % = 0.1)	900	7200	Bridge	8	13386	24047.0 ± 111.4	-95.69
			Brick	8	6089	1519.0 ± 99.9	45.73
			Bridge	13	1063	30.9 ± 4.9	209.23
			Brick	13	1029	30.7 ± 5.6	179.17
WS Small-world #4 (rewiring % = 0.5)	900	7200	Bridge	8	22649	22935.4 ± 148.5	-1.93
			Brick	8	3973	4244.9 ± 153.7	-1.77
			Bridge	13	213	56.2 ± 8.9	17.64
			Brick	13	47	17.7 ± 4.1	7.11

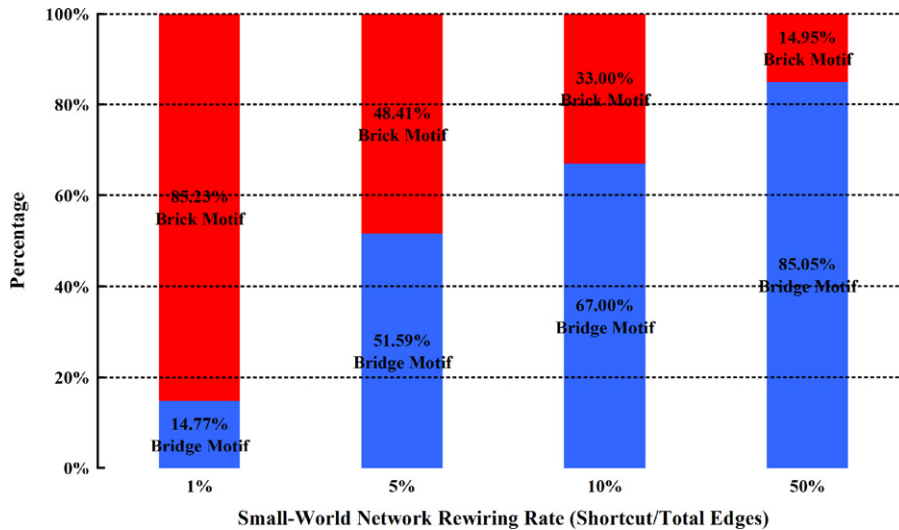


Fig. 8. Percentages of bridge and brick motifs in small-world networks according to different rewiring ratios.

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. Brick motifs appear to play an important role in reducing the degree of separation and increasing the degree of clustering in scale-free networks.
3. *Random*: As we 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.

6. Conclusion

We believe our detection algorithm can 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. 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 in their work with design principles and network evolution.

Acknowledgement

This work was supported in part by National Science Council, Taiwan, Republic of China under grant NSC 94-2520-S-009-002 and NSC 95-2221-E-182-076.

References

- [1] M.E.J. Newman, SIAM Rev. 45 (2003) 167.
- [2] M.E.J. Newman, D.J. Watts, Phys. Rev. E 60 (1999) 7332.
- [3] D.J. Watts, S.H. Strogatz, Nature 393 (1998) 440.

- [4] R.F. Cancho, C. Janssen, R.V. Sole, *Phys. Rev. E* 6404 (2001) 046119.
- [5] A. Barrat, M. Weigt, *Eur. Phys. J. B* 13 (2000) 547.
- [6] S.H. Strogatz, *Nature* 410 (2001) 268.
- [7] R. Milo, et al., *Science* 298 (2002) 824.
- [8] Y. Artzy-Randrup, S. Fleishman, N. Ben-Tal, L. Stone, *Science* 305 (2004) 1107.
- [9] K. Goh, E. Oh, H. Jeong, B. Kahng, D. Kim, *Proc. Natl. Acad. Sci. U.S.A.* 99 (2002) 12583.
- [10] A. Vazquez, et al., *Proc. Natl. Acad. Sci. U.S.A.* 101 (2004) 17940.
- [11] H. Moon, J. Bhak, K. Lee, D. Lee, *Bioinformatics* 21 (2005) 1479.
- [12] S. Shen-Orr, R. Milo, S. Mangan, U. Alon, *Bioinformatics* 31 (2002) 64.
- [13] S. Mangan, U. Alon, *Proc. Natl. Acad. Sci. U.S.A.* 100 (2003) 11980.
- [14] J. Berg, M. Lassig, *Proc. Natl. Acad. Sci. U.S.A.* 101 (2004) 14689.
- [15] M. Middendorf, E. Ziv, C. Wiggins, *Proc. Natl. Acad. Sci. U.S.A.* 102 (2005) 3192.
- [16] U. Alon, *Science* 301 (2003) 1866.
- [17] J. Bascompte, C. Melian, E. Sala, *Proc. Natl. Acad. Sci. U.S.A.* 102 (2005) 5443.
- [18] J. Rice, A. Kershenbaum, G. Stolovitzky, *Proc. Natl. Acad. Sci. U.S.A.* 102 (2005) 3173.
- [19] H. Hethcote, *SIAM Rev.* 42 (2000) 599.
- [20] J. Eckmann, E. Moses, *Natl. Acad. Sci. U.S.A.* 99 (2002) 5825.
- [21] S. Manrubia, J. Poyatos, *Europhys. Lett.* 64 (2003) 557.
- [22] S. Mangan, A. Zaslaver, U. Alon, *J. Mol. Biol.* 334 (2003) 197.
- [23] M.S. Granovetter, *Sociological Theory* 1 (1983) 201.
- [24] J. Davidsen, H. Ebel, S. Bornholdt, *Phys. Rev. Lett.* 88 (2002) 128701.
- [25] M. Girvan, M. Newman, *Natl. Acad. Sci. U.S.A.* 99 (2002) 7821.
- [26] M. Newman, M. Girvan, *Phys. Rev. E* 69 (2004).
- [27] H. Zhu, Z. Huang, *Phys. Rev. E* 70 (2004).
- [28] L.H. Hartwell, et al., *Nature* 402 (1999) C47.
- [29] J.D. Han, et al., *Nature* 430 (2004) 88.
- [30] J. Stelling, et al., *Cell* 118 (2004) 675.
- [31] E.Y. Lotem, et al., *Proc. Natl. Acad. Sci.* 101 (2004) 5934.
- [32] R. Milo, et al., *Science* 303 (2004) 1538.
- [33] M. Costanzo, et al., *Nucl. Acids Res.* 29 (2001) 75.
- [34] L.D. Zeleny, *Sociometry* 13 (1950) 314.
- [35] D. Macrae, *Sociometry* 23 (1960) 360.
- [36] R.J. Williams, N. Martinez, *Nature* 404 (2000) 180.
- [37] S. Kundu, et al, in: *Proceedings of the IEEE International Test Conference*, vol. 30, 1992.
- [38] J. Shurin, et al., *Ecol. Lett.* 5 (2002) 785.