Developing a Framework for Exploring Clustering Coefficient to Evaluate System Coupling

John W. Dahlgren, MITRE Corp

dahlgren@mitre.org

28 May 2007

Approved for Public Release Distribution Unlimited #07-0202

Purpose

- This briefing discusses the concept of the Clustering Coefficient as it applies to people and organizations
- The briefing will then discuss an analogy between the Clustering Coefficient and what the author calls the Coupling Coefficient that is meant to be applied to technical systems

MITRE

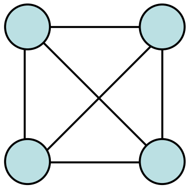
Clustering Coefficient

- Clustering is a measure, or at least a heuristic, to define the level of connectivity between a group of people
- Clustering Coefficient (CC) is:

  CC = number of close links/number of possible close links

  Number of possible close links = N(N-1)/2

  Assumes Full Duplex connectivity



- Four people, all tightly linked, have a CC = 1.0
  - All of the 6 possible connections are tight, CC = 6/6
- If only 4 tight connections existed, then CC = .667
- Both *Linked* and *The Agile Organization* discuss the need to have week connections to outside organizations
  - *Linked* discussed that jobs are suggested by friends, but by friends of friends

MITRE

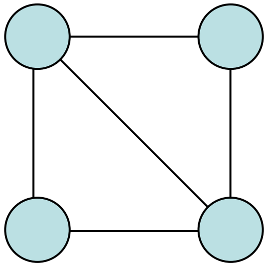
Optimal Clustering Coefficient

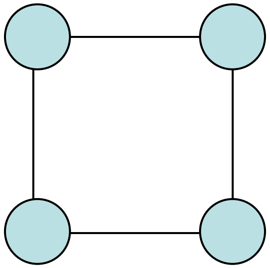
- Author is still searching for an answer to how to optimize network performance to have a Small World Network that takes into account:
  - How the formula for CC can be amended to take into account the optimal range of tight connections
  - How the formula for CC can be amended to take into account the benefits of medium and weak connections
  - How the answers to the CC and the above items relate to the survivability of a Small World Network
  - How a formula can evaluate the gradual degradation of performance as various links are removed in a Small World Network

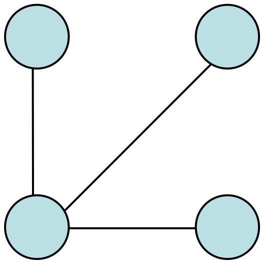
MITRE

Range of CC Values

Close Connections	Clustering Coefficient	Maximum # of Hops	Average # of Hops	Resiliency of Network
1	.167	Undefined	Undefined	0
2	.333	Undefined	Undefined	0
3	.5	2	1.5	Minimal to 0
4	.667	2	1.33	Good
5	.833	2	1.17	Best
6	1.0	1	1	Good to Avg







MITRE

Outside Connections

- Previous slide discussed resiliency of the internal network nodes to each other
- Fails to consider need and value of connections to outside world
- If each node can handle up to 3 tight connections (see diagram with all 4 nodes tightly connected), then…
- Looking at the 5 connection diagram shows 2 nodes with the capacity for outside connections
- The 4 connection diagram shows all 4 nodes have capacity for 1 additional connection to the outside world
- Appears to be a tradeoff between additional tight connections with other close nodes and additional loose connections to the outside world
- There is probably a point of Diminishing Marginal Returns and Negative Total Returns for more tight connections

Diminishing Marginal Return: $\Delta V_n < \Delta V_{n-1}$

Negative Returns: Total Value $(TV)_n < TV_{n-1}$

MITRE

Real Options and Network Connectivity

- Each strong connection in a small network should be evaluated against adding multiple weak outside connections
- Each weak outside connection represents a Real Option the network can exercise when needed
  - Each tight internal connection should be evaluated according to Diminishing Marginal Returns and Negative Total Returns
  - Economic theory often assumes an average value and cost for each tight connection and each weak connection
- Adding the value of weak outside connections to the previous discussion on resiliency, the 4-connection network is probably the most resilient
- Information value from each connection is also an important consideration

	Tight Connection	Loose Connection
High Quality Information	+	++
Low Quality Information	--	-

Relating this to the Clustering Coefficient

Optimal Network Design ⊓ (union of)

.6 ≥ CC ≥ .83

Design2 > # of Weak Links than Design1

CC = number of close links/ number of possible close links

Number of possible close links = N(N-1)/2

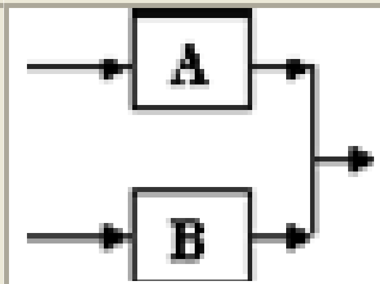
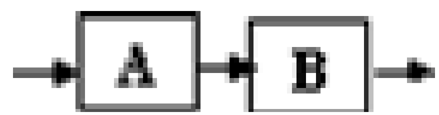
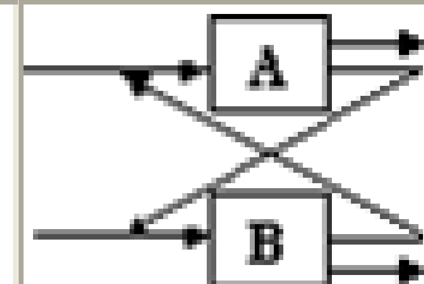
MITRE

Loose Coupling

- Difficult to define, although term is often used
- Some forums have implied that any tight coupling is bad
- Research by Konstantinos Kalligeros (MIT) has shown that some tight coupling, intentionally done to provide platforming opportunities, can be good. Platforms can provide:
  - Faster/cheaper deployment of new variants or system increments
  - Modularity of design and use, leading to operational flexibility
  - Interface standardization, thus providing interchangeability
- Kalligeros readily admits incorrect platforming can have some negative consequences, such as:
  - Strategic commitment and possible sub-optimality of long-term design
  - Locking-in with expertise and a given supply chain
  - Dominant standards that subsequently limit innovation
- Design Structure Matrices can help identify degrees of coupling

Design Structure Matrix (DSM)

- **DSMs identify dependencies and aid in understanding complexity**
- **Start in a column and ask the question, “How does a change in the component from this column impact the component in the row?”**

<table>
<tr><th colspan="4">Three Configurations that Characterize a System</th></tr>
<tr><td>Relationship</td><td>Parallel</td><td>Sequential</td><td>Coupled</td></tr>
<tr><td>Graph Representation</td><td>
</td><td>
</td><td>
</td></tr>
<tr><td>DSM Representation</td>
<td><table><tr><td></td><td>A</td><td>B</td></tr><tr><td>A</td><td></td><td></td></tr><tr><td>B</td><td></td><td></td></tr></table></td>
<td><table><tr><td></td><td>A</td><td>B</td></tr><tr><td>A</td><td></td><td></td></tr><tr><td>B</td><td>X</td><td></td></tr></table></td>
<td><table><tr><td></td><td>A</td><td>B</td></tr><tr><td>A</td><td></td><td>X</td></tr><tr><td>B</td><td>X</td><td></td></tr></table></td></tr>
<tr><td colspan="4">DSM Component Relationships (DSM Web Site 2006)</td></tr>
</table>

Degrees of Coupling

- Hypothetically, a system of 30 components could have 435 tight couples
- Does the previous heuristic for human networks apply to component systems? This would lead to 261 – 361 tight couples.
  – This level of complexity will still inhibit system evolution
- Goal may be to have each component coupled in some way to only one other component and not to every other component

MITRE

Generalized Satellite Terminal – Satellite – Terminal DSM

- H=High, M=Medium, L=Low degree of coupling
- Tight couplings between subsystems can lead to customized design rules that make the total system very tightly coupled and difficult to design, maintain and spiral develop
- Tight couples between components can be similar to tight couples in social networks, with a tradeoff to loose couples

		1	2	3	4	5	6	7	8	9
Baseband -- User 1	1		L	H						H
Input Port -- User 1	2			H						
Terminal -- User 1	3	H	H		M	H	L			
Uplink Channel -- User 1	4			H		H	M			
Satellite Communications Payload	5			M	L		L			
Downlink Channel -- User 2	6			M	M	H				
Terminal -- User 2	7			H	L	H	M		H	H
Output Port -- User 2	8							H		
Baseband -- User 2	9	H		M				H	L	

MITRE

Standards and Interfaces

- Coupling to standard interfaces can lessen complexity and overall system coupling
- The OSI stack can provide a common set of standards for many systems to be coupled to and thus lessen coupling between the subsystems
  - Coupled subsystems may need to be evaluated at each level of the OSI stack
- Often systems become coupled to Commercial Off The Shelf products
  - As COTS products gain a significant market share they can become the dominant standard and the system can become tightly coupled to a given product line
  - COTS suppliers often choose to upgrade their products every 5 years, forcing users to upgrade
  - While a COTS product might initially represent an *Option* to be exercised if the owner chooses, dominant COTS products make themselves an *Obligation*

MITRE

Coupling Coefficient

- Coupling Coefficient formula for hardware and software should relate to the tightness of coupling for each system to other systems, to standards, and to upgrade cycles
- For computer networks, the coupling to standards would be evaluated at each layer of the OSI stack if each layer is required to exchange information
- A system can be tightly coupled to another system if the coupling is accomplished using a non-standard linkage
- Systems are loosely coupled to each other if they are linked using a standard that is readily available to and used by the consumer base for this product
- Therefore, systems do not necessarily need to use standards linked directly to COTS products

$$\text{Coupling Coefficient} = \sum_{i=1}^{7} OSI_i \left[\sum_{\substack{i=1 \\ j=1}}^{n} S_i S_j \right] \quad \text{number of possible tight couplings}$$

$$\text{COTS Portion of Coupling Coefficient} = \sum_{\substack{i=1 \\ j=1}}^{n} S_{(COTS)i} S_{(COTS)j} \,/\text{Years to upgrade}$$

Conclusions

- It appears that a Coupling Coefficient for components can be derived from lessons learned regarding the Clustering Coefficient for social networks
- The use of applying tight versus loose couplings to a social network can aid in determining the resiliency of the network
- The Coupling Coefficient relates to the number of tight couplings between components, the use of standards, and the use of (dominant) COTS products
- The use of standards represents an option on being able to upgrade subsystems as technology evolves without having to upgrade every other subsystem it interfaces with
- The use of COTS products represents an option that can become an obligation
- Some tight couplings will still be needed, if only to standards

MITRE

Bibliography

- **Atkinson, Simon Reay, and James Moffatt. 2005. *The agile organization: From informal networks to complex effects and agility.* Washington DC: The Command and Control Research Program (CCRP).**
- **Barabási, Albert-László. 2002. *Linked: How everything is connected to everything else and what it means for business, science, and everyday life.* New York: Plume.**
- **Dahlgren, J.W. Real options and value driven design in spiral development. Paper presented at 2006 Command and Control Research and Technology Symposium (CCRTS), Coronado, Ca. 2006**
- **Design Structure Matrix Web Site. www.dsmweb.org, 2006.**
- **Kalligeros, K., Richard de Neufville, Olivier de Weck, David Geltner and Patrick Jalliet. Platforming and real options in large scale system design. PhD dissertation presentation. Cambridge, MA 2006**