Carnegie Mellon

CASOS 2006 Summer Institute

Kathleen M. Carley
Center for Computational Analysis of Social and Organizational Systems

Carnegie Mellon University

www.casos.cs.cmu.edu

kathleen.carley@cmu.edu
412 268 6016

Center for Computational Analysis of Social and Organizational Systems

http://www.casos.cs.cmu.edu/

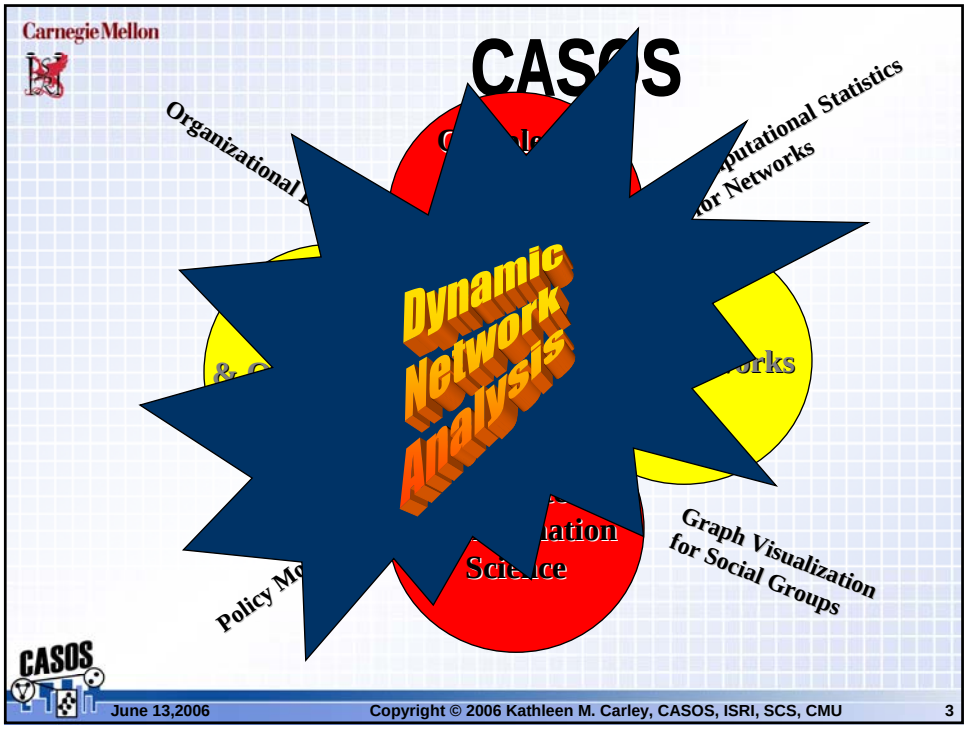
Carnegie Mellon

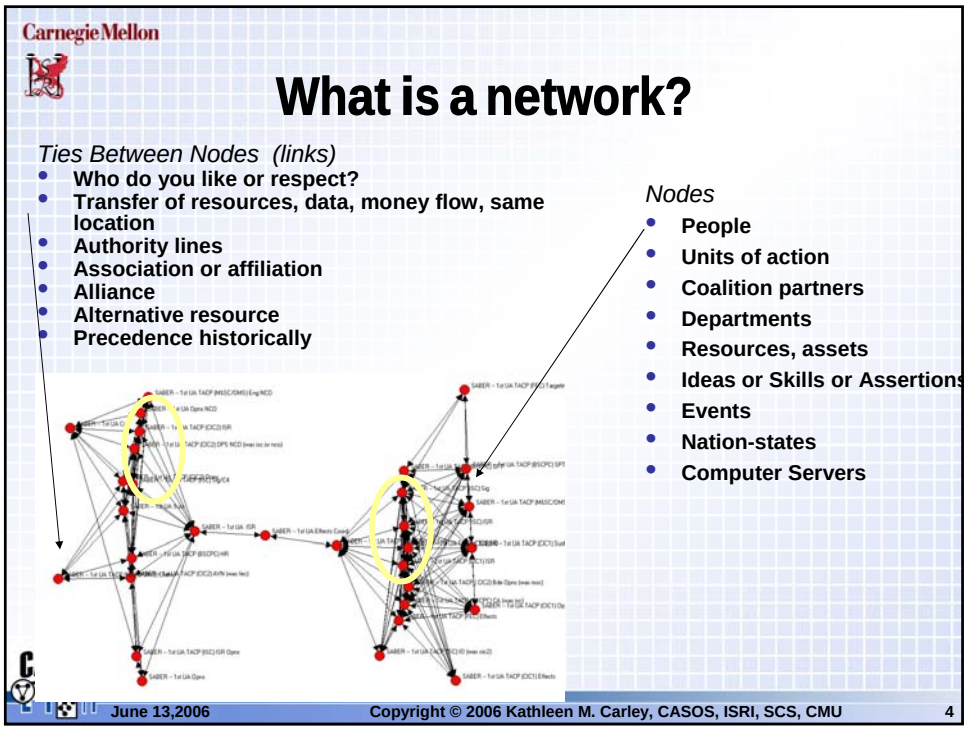
Purpose of Institute

Learn:

- What is computational social & organizational science
- How to link network analysis and simulation
- Move beyond traditional social network analysis to dynamic network analysis
- What is the state of the art
- Process of
  - Designing models
  - Analyzing models
  - Validating models
- Become a knowledgeable consumer of computational models
- Introduction to a set of models and tools
- Learn appropriate and inappropriate critiques

NOT about programming


Carnegie Mellon
CASOS
Organizational
Computational Statistics
Networks
Dynamic
Network
Analysis
Graph Visualization
for Social Groups
Policy
Science
CASOS
June 13,2006
Copyright © 2006 Kathleen M. Carley, CASOS, ISRI, SCS, CMU
3


Carnegie Mellon
What is a network?
Ties Between Nodes (links)
Who do you like or respect?
Transfer of resources, data, money flow, same location
Authority lines
Association or affiliation
Alliance
Alternative resource
Precedence historically
Nodes
People
Units of action
Coalition partners
Departments
Resources, assets
Ideas or Skills or Assertions
Events
Nation-states
Computer Servers
June 13,2006
Copyright © 2006 Kathleen M. Carley, CASOS, ISRI, SCS, CMU
4

Carnegie Mellon

Why is Network Analysis Important

- Network analysis enables measurement of socio-cultural environment
- Relations constrain and enable behavior
  - Network structure impacts performance, information diffusion, technology adoption, disease spread …
  - Individual's position in network influences chance of promotion, ability gather and relay information, ability to influence and be influenced
  - "Health" of the organization can be assessed by examining the networks
  - When information is uncertain people rely on their networks to make decisions
  - Organization's position in inter-organizational network controls access to resources and information
  - Manager's perception of inter-organizational network influences alliance decisions such as mergers and joint ventures

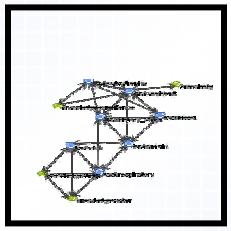
CASOS

Carnegie Mellon

Network Assessment

Who is Key

What are the groups



What kind of network is this

How can some one or some group be influenced

CASOS

Carnegie Mellon

So – why is this hard?

- The Network
  - Vast quantities of data
  - Multi-mode – people, events, etc.
  - Multi-plex – many connections e.g. financial and authority
- The Information
  - Intentional misinformation – e.g., aliases
  - Inaccurate information – e.g., typos
  - Out-of-date information
  - Incomplete information
- Dynamic
  - Learning
  - Recruitment
  - Attrition
  - …

Carnegie Mellon

What is Dynamic Network Analysis?

- **Combines**
  - **Social network analysis**
  - **Link analysis**
  - **Multi-agent modeling**
- **Applied to networks that are (meta-networks)**
  - **Large**
  - **Multi-mode**
  - **Multi-link**
  - **Dynamic**
  - **Uncertain**
- **Using**
  - **real world empirical data**
  - **Social, behavioral, organizational research findings**
- **Resulting in multi-agent network modeling (MAS_DNA)**

- **Sub Areas**
  - **Metric development**
  - **Network assessment**
  - **Network forecasting**

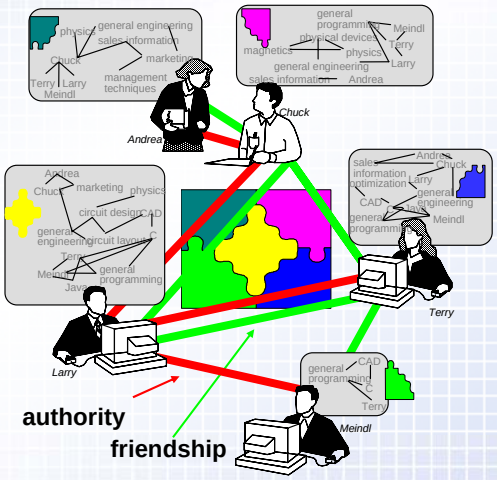
Carnegie Mellon

Networks Interlink

- **Social networks**
  - **Who to whom**
- **Information networks**
  - **What to what**
- **Knowledge networks**
  - **Who to what**

These can be inter-linked at either the individual, group, or corporate level.

These can be inter-linked in terms of words, specific pieces of information, or general bodies of knowledge.

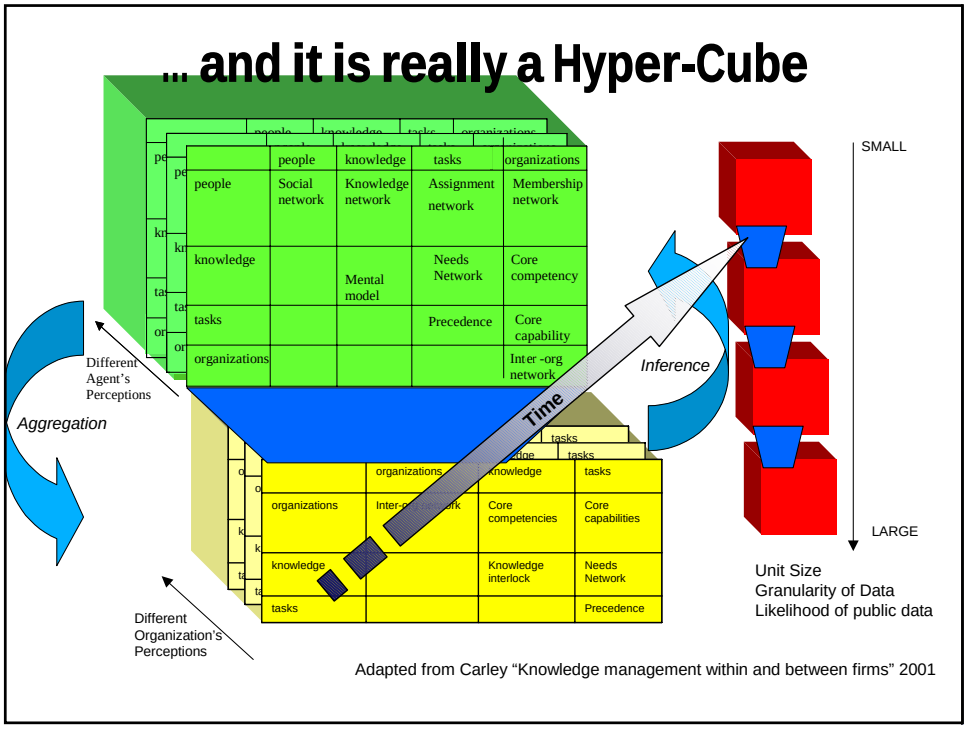


Carnegie Mellon

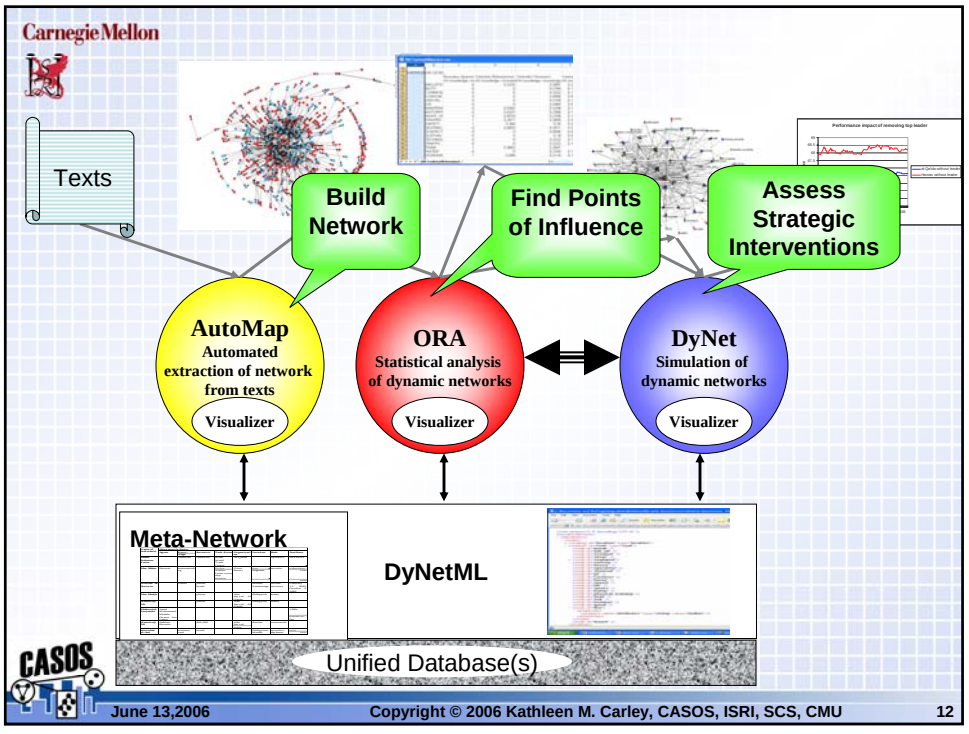
Meta-Matrix:

connections among multiple entities at varying strength

	People / Agents	Knowledge / Resources	Tasks / Events	Group/ Organizations
People / Agents	Social Network	Knowledge Network	Assignment Network	Membership Network
Knowledge / Resources		Information Network / Substitutes	Needs Network	Core Capabilities
Tasks / Events			Precedence Ordering	Institutional Relation
Group/ Organizations				Inter-organizational Network



Adapted from Carley "Knowledge management within and between firms" 2001



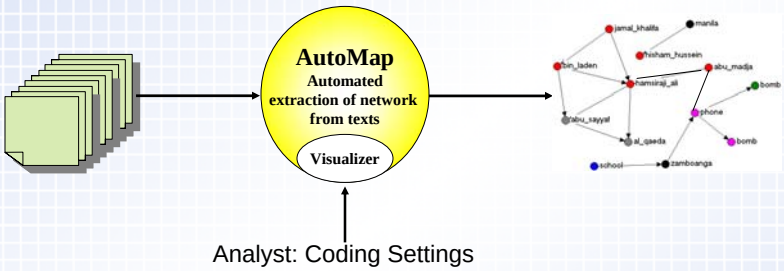
Carnegie Mellon

Steps in a Structural Analysis

1. Collect network data.
   - Use automated text analysis.
   - Parse from excel files.
2. Generate report
3. Enter data into network statistical tool.
4. Visualize.
5. Analyze
   - If multiple networks create combined measures.
   - If needed look at some measures more indepth.
   - Possibly drop isolates and pendants
   - Check interpretations.
6. Generate report.
7. Immediate impact analysis.
8. Generate report.
9. Near term impact analysis.
   - Use simulation.
10. Generate report.

Carnegie Mellon

From texts to networks: process and methodology



Analyst: Coding Settings

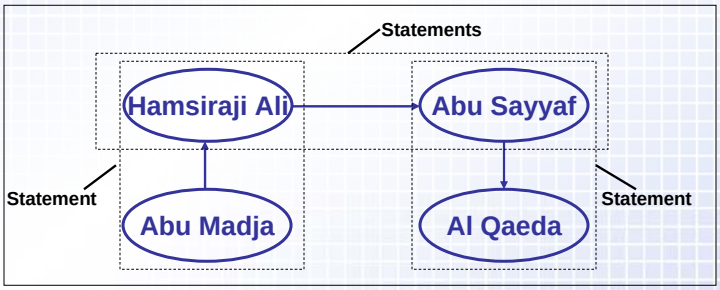
That mobile phone also registered calls to Abu Madja and Hamsiraji Ali, leaders of Abu Sayyaf, Al Qaeda's Philippine branch.

- Map Analysis (AutoMap, Diesner & Carley 2004)
- Analyze meaning of texts by identifying relationships between concepts and categories (Carley 1997)

Carnegie Mellon

Distance based Map Analysis in AutoMap

- That mobile phone also registered calls to Abu Madja and Hamsiraji Ali, leaders of Abu Sayyaf, Al Qaeda's Philippine branch



CASOS

15

Carnegie Mellon

Example Text Files

Hisham Al Hussein

… the Philippine government booted the second secretary at Iraq's Manila embassy, Hisham Al Hussein, on February 13, 2003, after discovering that the same mobile phone that reached his number on October 3, 2002, six days later rang another cell phone strapped to a bomb at the San Roque Elementary School in Zamboanga.

Abu Madja and Hamsiraji Ali

That mobile phone also registered calls to Abu Madja and Hamsiraji Ali, leaders of Abu Sayyaf, Al Qaeda's Philippine branch.

Abdurajak Janjalani

It was launched in the late 1980s by the late Abdurajak Janjalani, with the help of Jamal Mohammad Khalifa, Osama bin Laden's brother-in-law.

Hamsiraji Ali

… Hamsiraji Ali, an Abu Sayyaf commander on the southern island of Basilan, bragged that his group received almost $20,000 annually from Iraqis close to Saddam Hussein.

CASOS

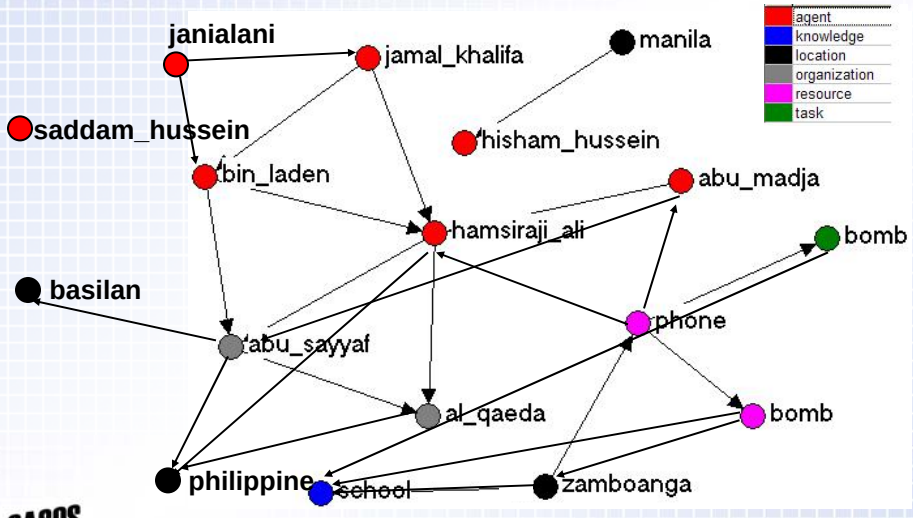
Carnegie Mellon

Place Concepts in Meta-Matrix

Agents	Knowledge	Resources	Tasks	Locations	Organizations
abu_madja	school	bomb	bomb	basilian	al_qaeda
bin_laden		phone		philippine	bu_sayyaf
hamsiraji_ali				manila	
hisham_hussein				zaboanga	
janialani					
jaml_khalifa					
saddam_hussein					

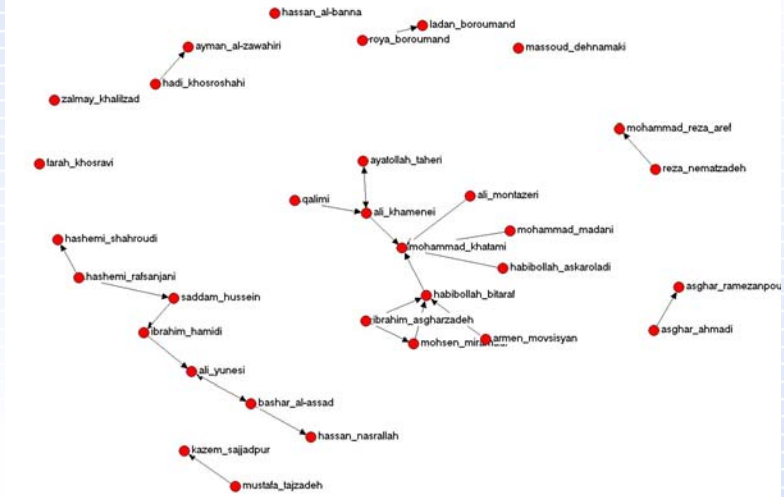
Carnegie Mellon

Resultant Meta-Network



Carnegie Mellon

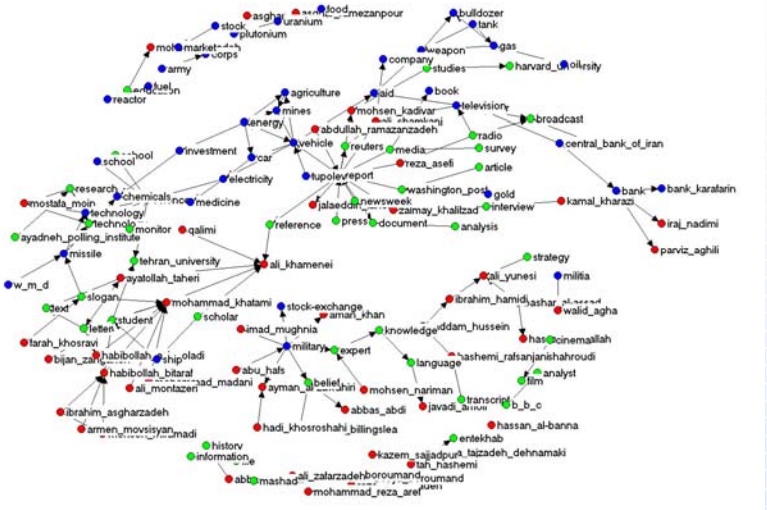
Mid East 2002



CASOS

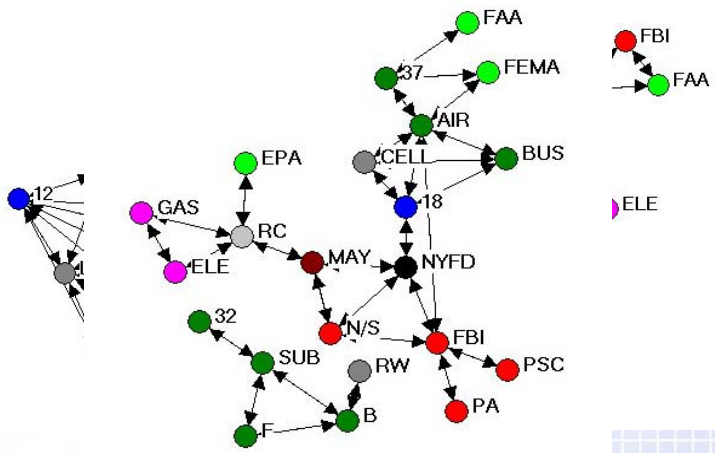
Carnegie Mellon

... add knowledge and resources – and connections are found



Carnegie Mellon

NYC response network

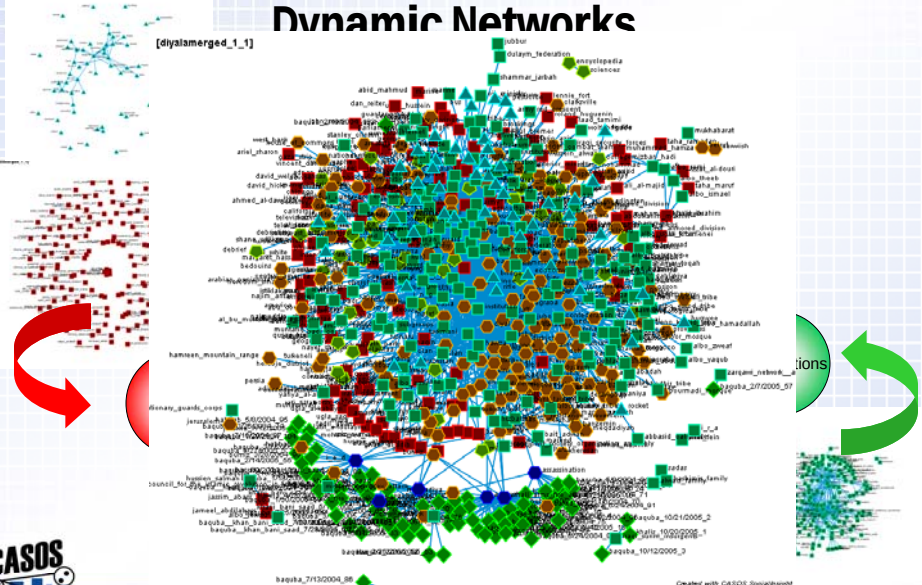


9-19-2001

Carnegie Mellon

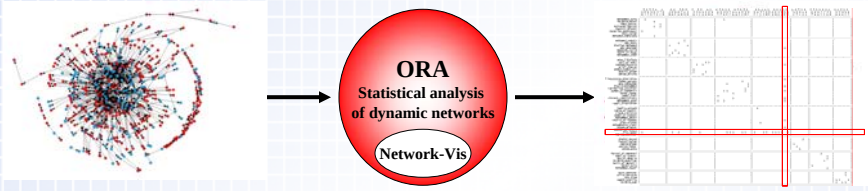
FOUO

Dynamic Network Analysis: Interacting Dynamic Networks



Carnegie Mellon

ORA: From networks to patterns

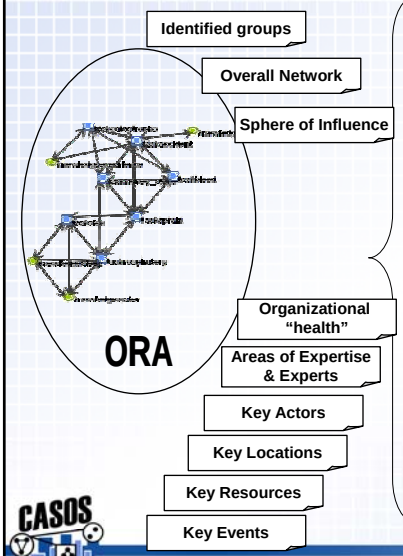


ORA: a DNA statistical analysis tool for locating patterns and identifying vulnerabilities

- Intelligence Report
- Management Report
- Sub-Group Report
- Context Report
- Sphere of Influence
- Mental Model Report
- Immediate Impact Report
- Near Term Impact Report
- All Measures

Carnegie Mellon

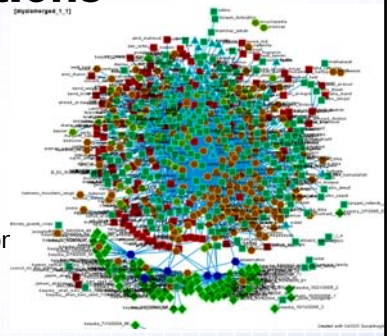
Potential Patterns that Can be Discovered



- **Identification of key actors**
  - **Knowledge Hubs**
  - **High Strength Relationships**
  - **Boundary Spanners**
  - **Power brokers**
  - **Emergent leaders**
- **Paths between actors**
- **Identification of working-groups or teams**
- **Identification of actor's sphere of influence**
- **Identification of key locations, events, resources**
- **Assessment of team "health" - workload, cognitive demand, shared SA, performance …**
- **Impact assessment**
  - **Immediate**
  - **Near term**
- **Text**
  - **Key semantics**
  - **Relational concepts**
  - **Hidden expertise**

Carnegie Mellon

Analyst Questions



- Intelligence report
  - Who to target (vulnerabilities)
    - What groups or individuals stand out
  - How to influence
    - Are there important connections
- Management report
  - What is the "health" of the organization
  - Where might there be missing data
- Comparison reports
  - How different are two groups – or two sources – or the same group at two different times
- Context report
  - What kind of network is this?
- Sphere of influence
  - What do we know about X?
- Immediate and near term impact reports
  - What is the impact of an intervention

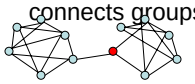
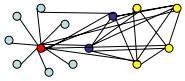
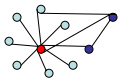
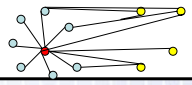
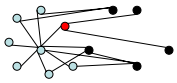
Carnegie Mellon

Simple SNA Measures

Measure	Definition	Meaning	Usage
Centrality	Number nodes connected to	In the know	Identifying sources for intel; Reducing information flow
Betweenness	Likelihood of paths through	Connects groups	Typically has political influence, but may be too constrained to act
Closeness	Nearness to all other nodes	Rapid access to all information	Identifying sources to acquire/transmit information
Betweenness - Closeness	High in betweenness but not closeness	Connects disconnected groups	Reduction in activity

Carnegie Mellon

Individuals Who Stand Out

Degree Centrality in the know 	High Betweenness and not Degree connects groups 	Cognitive Demand emergent leader 	Task exclusivity critical ability 
Eigenvector central core 	Betweenness many paths 	Resource exclusivity Mobilize resources 	Knowledge exclusivity Mobilize info 

Carnegie Mellon

Network Measures Are General Purpose

Measures can be calculated on any matrix
if cell values are correct (integer, binary, symmetric)
BUT ...
interpretation is different
and depends on type of tie and node

Three examples
1) Information networks
2) Inter-organizational networks
3) Task precedence networks

Carnegie Mellon

Who to target (vulnerabilities)

- Centralities
  - Communication
    - Degree – most connected
    - Betweenness – most paths
- Exclusivities
  - Expertise
    - Knowledge – special expertise
    - Task – special experience
- Demands/Loads
  - Roles
    - Cognitive demand – emergent leader
    - Workload

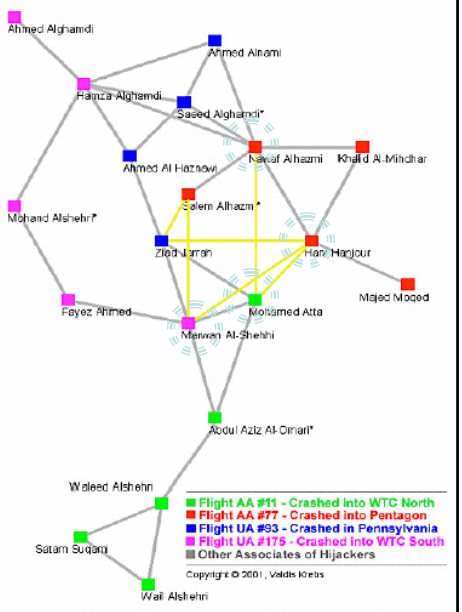
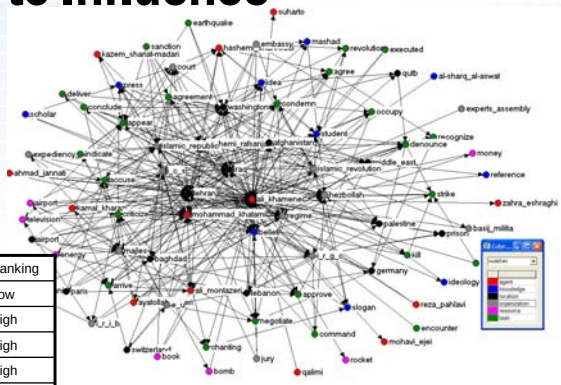


Figure 3 Trusted Prior Contacts + Meeting Ties [shortcuts]

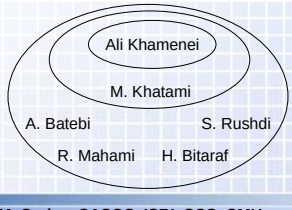
Carnegie Mellon

How to Influence

- For the target
  - Who are they connected to
  - What groups are they in
  - What do they know
  - What resources do they control
  - What activities are they involved in

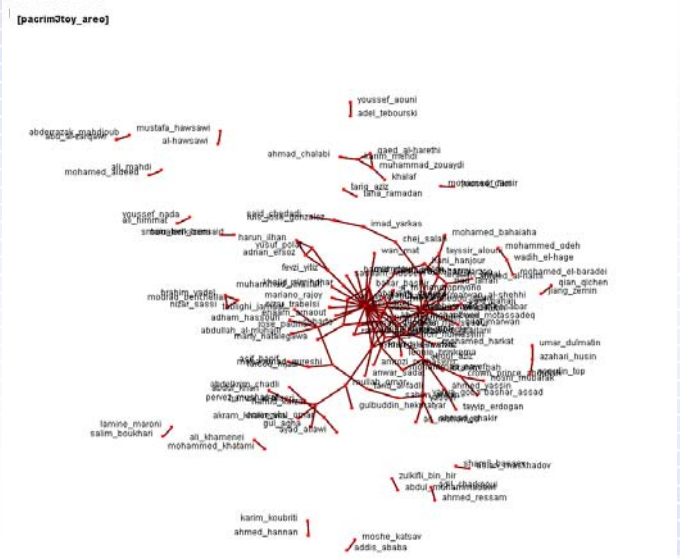


Attribute	Score	Ranking
Interacts with	1	Low
Knowledge areas	10	High
Resource areas	7	High
Organizations associated with	17	High
Density of ego net	.097	High
Task exclusivity	.0226	High
Resource exclusivity	.010	High
Knowledge exclusivity	.0254	High
Degree Centrality	.017	Norm
Betweenness	.0024	High
Cognitive Demand	.033	High
Eigenvector Centrality	.001	Low



Carnegie Mellon

Network Visualization

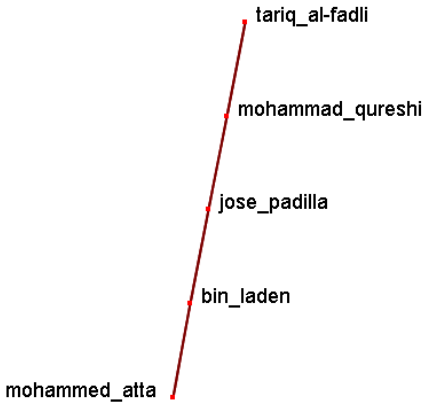


1. See
2. Drop isolate
3. Label
4. …

CASOS

Carnegie Mellon

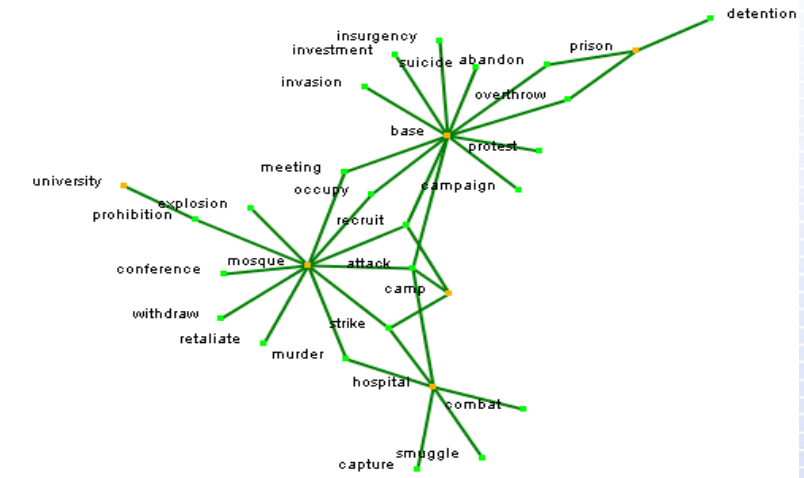
Path Finder



CASOS

Carnegie Mellon

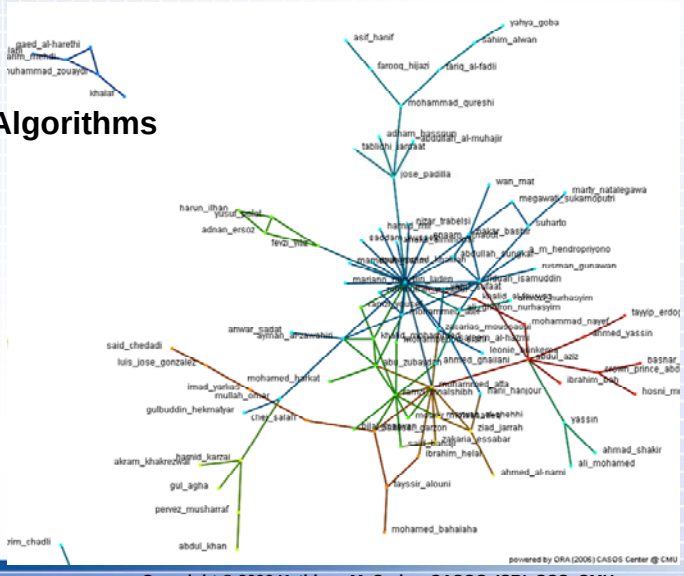
Difference in Events by Locations



Carnegie Mellon

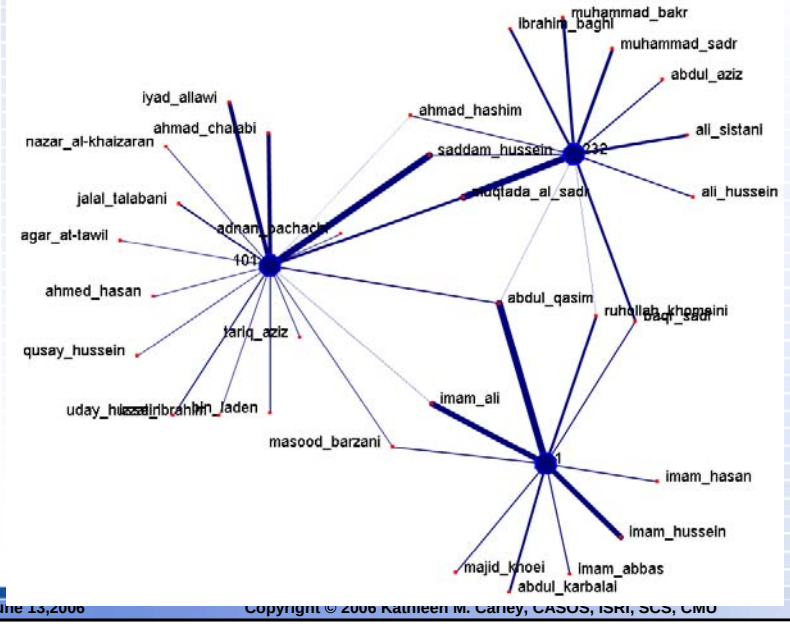
Locate Groups

3 Grouping Algorithms
Newman
Concor
FOG



Carnegie Mellon

FOG: Fuzzy Groups on Central Core



CASOS

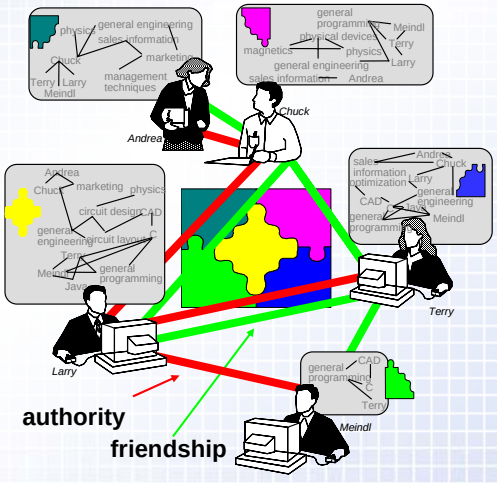
Carnegie Mellon

Networks Interlink

- Social networks
  - Who to whom
- Information networks
  - What to what
- Knowledge networks
  - Who to what

These can be inter-linked at either the individual, group, or corporate level.

These can be inter-linked in terms of words, specific pieces of information, or general bodies of knowledge.



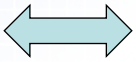
CASOS

Carnegie Mellon

Comparison of Actual and Expected Communication Networks

Actual Communication Network

```
  ACLTM
A 01000
C 10110
L 01010
T 01101
M 00010
```



Expected Communication Network

```
  ACLTM
A 01110
C 10110
L 01010
T 01100
M 00110
```

Difference

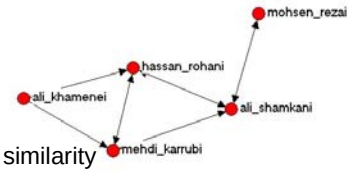
```
  ACLTM
A ==---=
C =====
L =====
T ====+
M ==-==
```

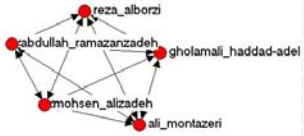
Potential Concerns are those interacting less/more than they should

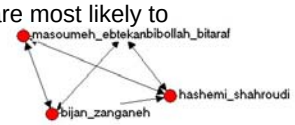
Carnegie Mellon

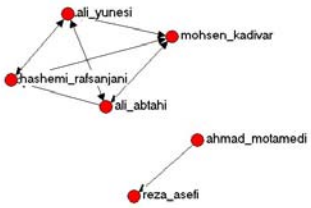
Estimating possible missing data or errors Who Should be Interacting?

- Static –
  - Relative similarity
  - Relative expertise
- Dynamic –
  - Simulate network evolution – which ties are most likely to occur



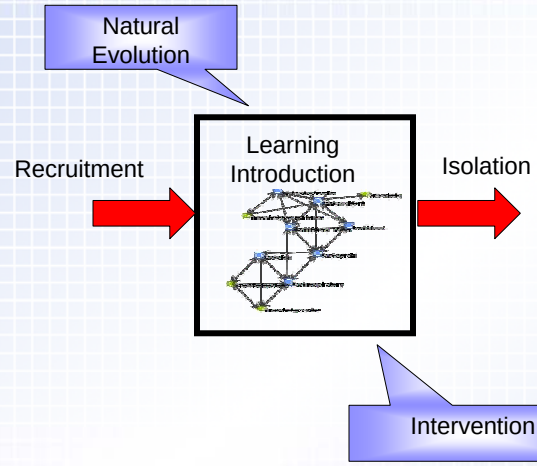






Carnegie Mellon

Change in Networks

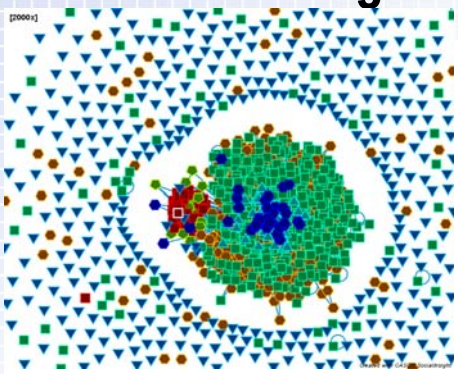


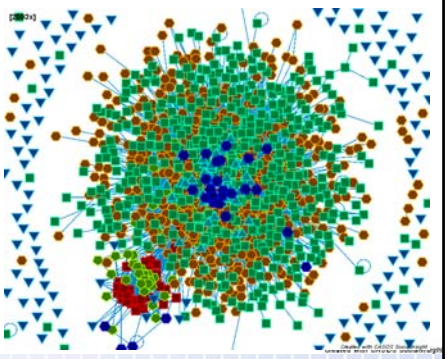
Three Ways to Study Dynamics

1. Comparison over time
   - Look at real data
2. Immediate Impact
   - Comparative statics
3. Near Term Impact
   - Utilizes simulation - DyNet

Carnegie Mellon

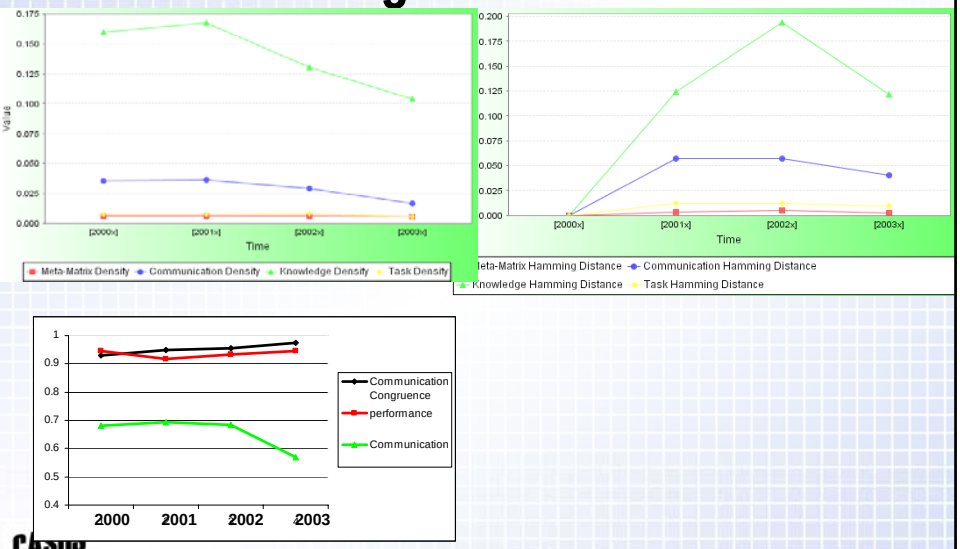
Change 2000 - 2003





Carnegie Mellon

Change 2000-2003



Carnegie Mellon

Immediate Impact - Prediction

- *What if ? Remove top 5 emergent leaders*
- Change in performance
  - Anticipated drop – 4% percentage difference
- Change in information diffusion
  - Anticipated increase – 67% percentage difference
- New emergent leaders
  1. 0.0174 said_mortazavi
  2. 0.0137 kamal_kharazi
  3. 0.0127 reza_asefi
  4. 0.0120 morteza_sarmadi
  5. 0.0100 hashemi_shahroudi
- Value of “lowest” old emergent leader was .0246

Carnegie Mellon

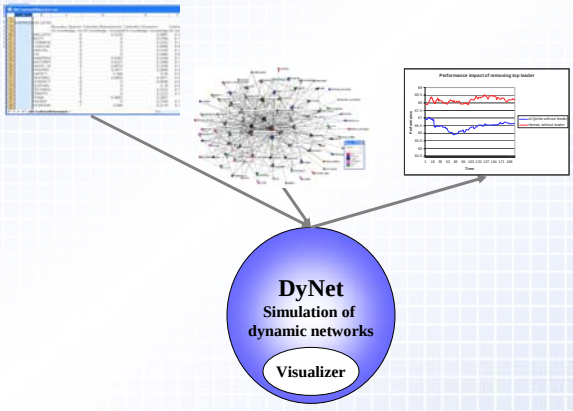
Who are likely to be new Emergent Leaders?

	Who is Isolated	Baasyir	Bin Laden	High Cognitive Demand	Weak Boundary Spanner
	1	Gokhan	Gokhan	Kandari	Kandari
	2	MaFadli	Al Ha Ghamdi	Aufi	Nawar
	3	Tabarak	Benyaich	Benali	Aufi
	4	Al Ha Ghamdi	Maqbul	HaGhamdi	Jabarah1
	5	Aufi	MaFadli	MOShehri	Ameroude

Carnegie Mellon

Moving Beyond ORA to DyNet Simulation

DyNet – from Patterns and Identified COA to Near Term Impact



Carnegie Mellon

Change in Networks

Natural Evolution

Recruitment

Learning Introduction

Isolation

Intervention

Carnegie Mellon

What is Computational Modeling?

- Analysis of complex socio-technical systems using simulation
- Types
  - Multi-agent
  - Expert System
  - General System
  - Markov
  - …
- Why
  - Complexity
  - Dynamics
  - Time
  - Ethics

Carnegie Mellon

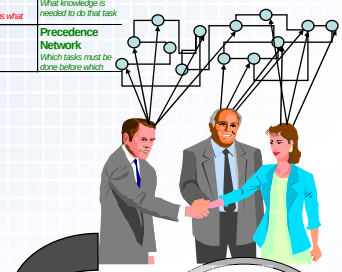
Use Simulation to Look at Change

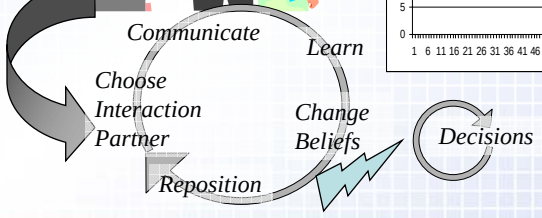
- Removal of nodes
  - Only immediate impact is seen using SNA
  - Need simulation for near or long term change
- Addition of nodes
  - Can't be done with SNA
- In general change means understanding process
- Common change processes
  - Learning
    - Relative Similarity, Relative Expertise
  - Entrance
    - Birth
  - Removal
    - Death

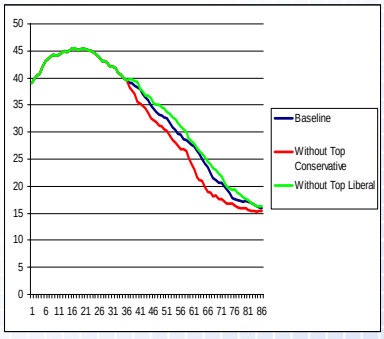
Carnegie Mellon

Change in Reformism – Near Term Prediction

	People	Knowledge	Tasks
People Relation	**Social Network** *Who knows who*	**Knowledge Network** *Who knows what*	**Assignment Network** *Who does what*
Knowledge Relation		**Information Network** *What informs what*	**Needs Network** *What knowledge is needed to do that task*
Tasks Relation			**Precedence Network** *Which tasks must be done before which*

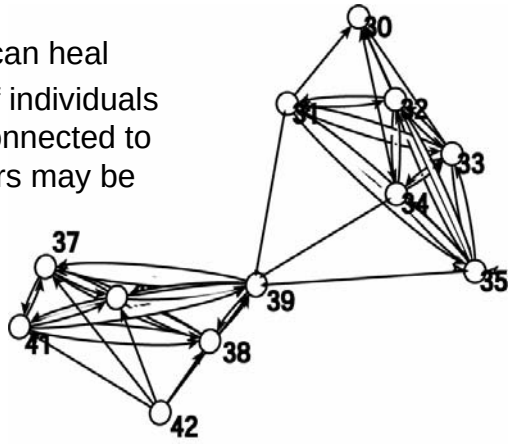






What is the near term effect (can it heal)

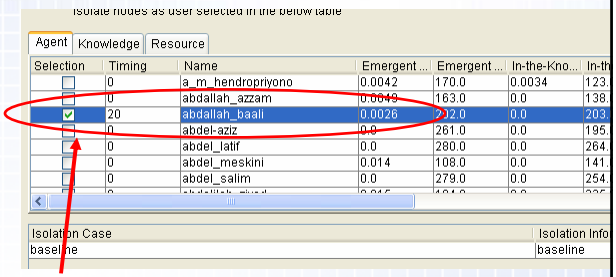
- Networks can heal
- Isolation of individuals who are connected to many others may be ineffective



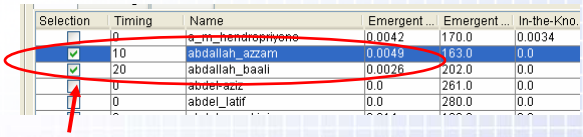
6. New Cell Leader Has Emerged

Near Term Analysis main window - Advanced Mode

- With third 'Add simulations', you can setup a customized isolation case.
  1) Find an agent to isolate from 'Name' column
  2) Check the 'Selection' box that is at the left side of the name
  3) Double click the 'Timing' box and type in the isolation timing
  4) repeat this as many as you can
  5) Click the third 'Add simulations' button to save the isolation case you just setup



'abdallah_baali' will be isolated at time 20

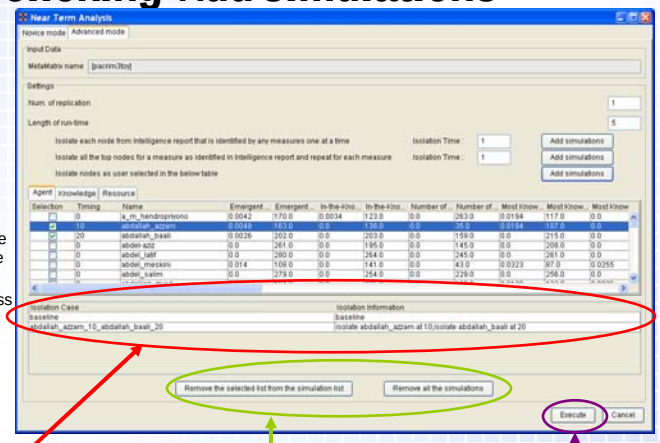


'abdallah_baali' will be isolated at time 20, and 'abdallah_azzam' will be isolated at time 10, too.

Carnegie Mellon

Near Term Analysis main window - After clicking 'Add simulations'

- Whenever you click 'Add simulations' button, you will see the added isolation cases at the bottom of the window.
- If you want to remove some isolation cases, please click or select them and press 'Remove the selected list from the simulation list' button
- If you want to remove all the isolation cases and start from the beginning, press 'Remove all the simulations' button
- If you finished for the setup, press 'Execute' button



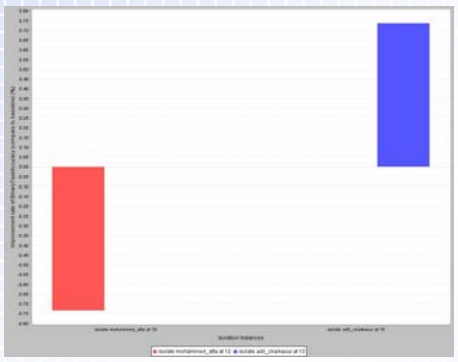
You can see the added isolation cases

You can remove isolation cases that are selected in the above list.

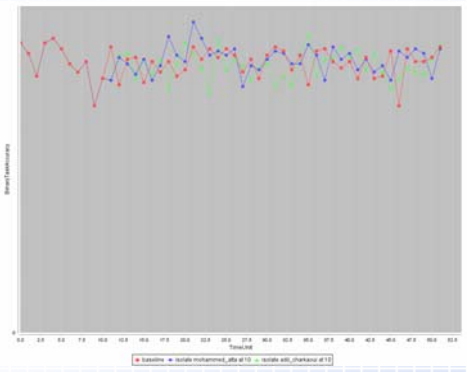
Click this button to execute

Carnegie Mellon

Impact of isolation



Ending Performance



Performance Over Time

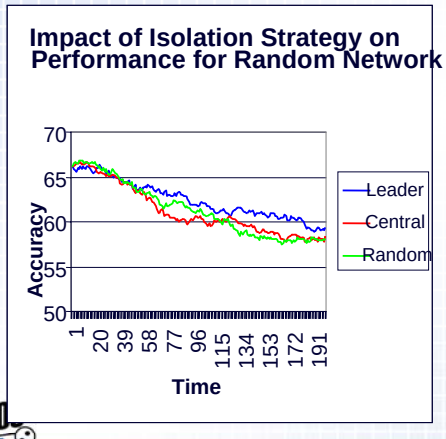
DyNet Analysis Output

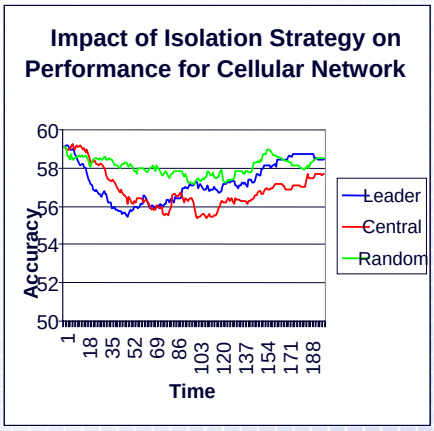
Bin Laden Isolation→ New Network Relationships

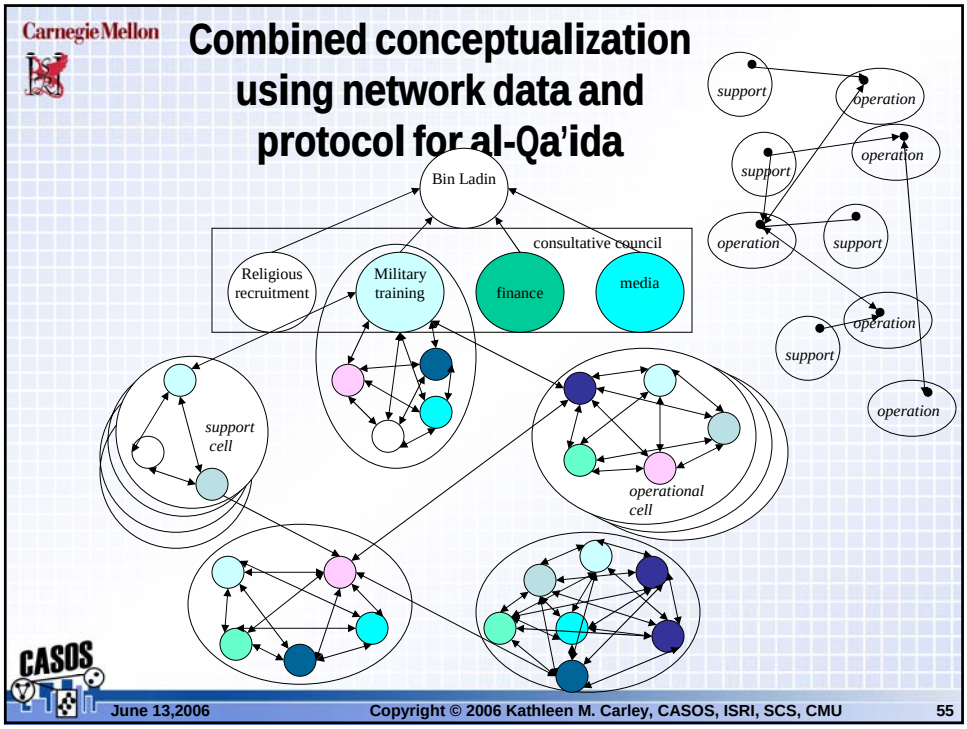
New Relationships based on bin Laden Isolation

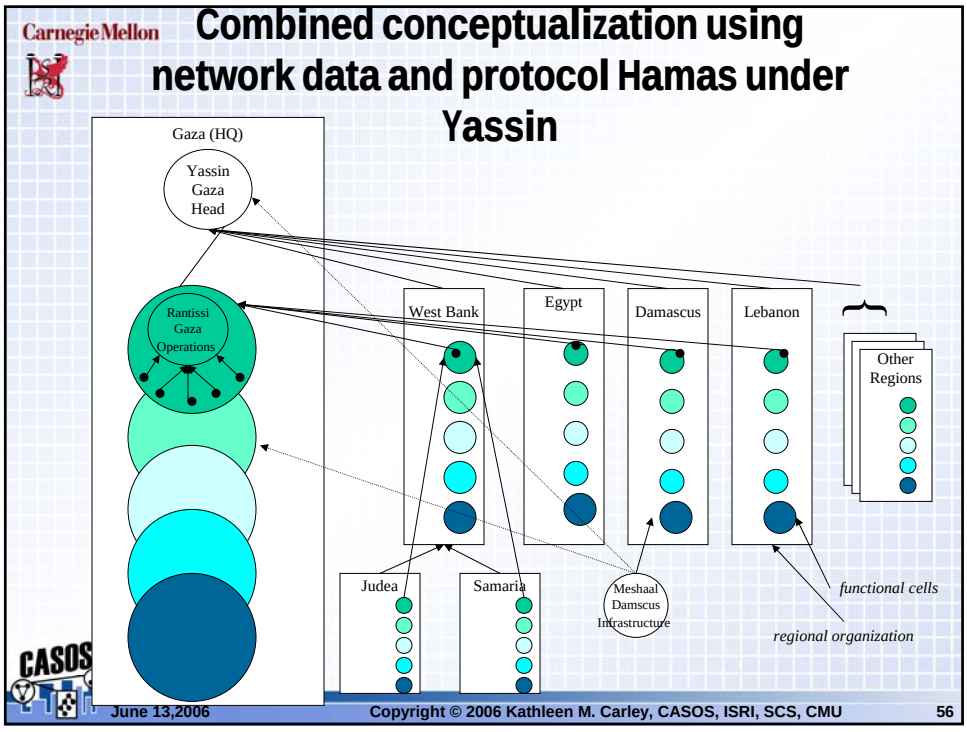
From	To	Value	NormValue
Salah	Sulaeman	0.07378	1
Nashiri	Sliti	0.016457	0.223052
Ghoni	Haroun	0.012465	0.168949
Zubaydah	Faiz	0.012083	0.163765
Sabour	Khalfaoui	0.011504	0.155918
Ghoni	Ikhlef	0.011457	0.155283
Nashiri	Lillie	0.01078	0.146111
Ghoni	Atmani	0.01067	0.144618
Doha	Ghoni	0.009806	0.132915
Mahdjoub	Hannachi	0.009584	0.129906
Roche	Muqrin	0.009504	0.128819
Ghoni	Hannachi	0.009374	0.12705
Doha	Faris	0.009175	0.124359

Impact is exacerbated when forecasting



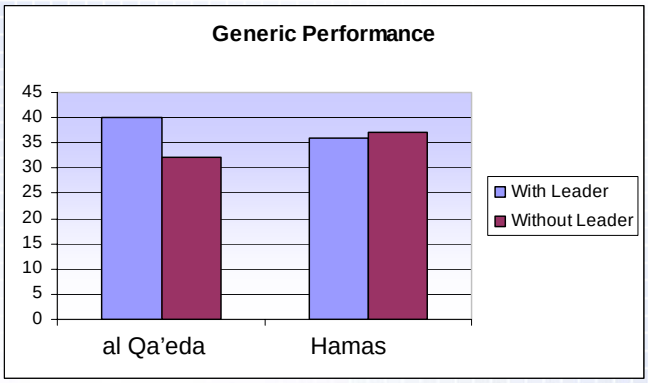



Carnegie Mellon
Combined conceptualization using network data and protocol for al-Qa'ida
Bin Ladin
consultative council
Religious recruitment
Military training
finance
media
support cell
operational cell
support
operation
support
operation
operation
support
support
operation
operation
CASOS
June 13,2006
Copyright © 2006 Kathleen M. Carley, CASOS, ISRI, SCS, CMU
55


Carnegie Mellon
Combined conceptualization using network data and protocol Hamas under Yassin
Gaza (HQ)
Yassin Gaza Head
Rantissi Gaza Operations
West Bank
Egypt
Damascus
Lebanon
Other Regions
Judea
Samaria
Meshaal Damscus Infrastructure
functional cells
regional organization
CASOS
June 13,2006
Copyright © 2006 Kathleen M. Carley, CASOS, ISRI, SCS, CMU
56

Carnegie Mellon

What If Analysis



Carnegie Mellon

Portions of ORA analysis: Characteristics of Key Actors

Characteristics	Meaning	al-Qaeda	Hamas
Complexity	**Very low then probably major amounts of missing data, possibly cells are self directed. Very high then system is tightly coupled and possibly prone to group think.**	**slightly more complex .096 Overall – very low density**	**slightly less complex .053 Overall – very low density**
Highest in degree centrality	**Individual most likely to diffuse new information, isolation of this person will be slightly crippling for a short time.**	**Bin Ladin .028**	**Yassin .011** 
Highest in cognitive demand	**Individual most likely to be an emergent leader, isolation of this person will be moderately crippling for a medium time**	**Bin Ladin .015** 	**Rantissi .087** 

Carnegie Mellon

Where are Dynamic Network Analysis Models Used

- Designing adaptive teams
- Performance evaluations
- Evaluating organizational structures and evaluating changes such as downsizing
  - E.g., hospitals, health departments NY Dutchess County
- Estimating effectiveness and adaptability of new structures
  - E.g., SSG – Concargru, Army Unit of Action, CPOF (IRAQ)
- Estimating size, shape and vulnerabilities in organizational designs and covert networks
  - E.g., NASA, Counter-terrorism, counter-narcotics, terrorist, tax-avoiders
- Examining impact of IT on effectiveness
  - E.g., NASA, Knowledge Wall in JTF
- Impact analysis of actions in asymmetric warfare situations
- Impact on cities of weaponized biological or chemical attacks
- Identifying key actors and emergent groups
  - E.g., Counter terrorism, Health Units, Merchant Marine
- Prevention and intervention
  - E.g., IRS tax avoidance interventions

CASOS

Carnegie Mellon

Where are Dynamic Network Analysis Models Used

- Designing adaptive teams for Command and Control
- Evaluating organizational structures
  - E.g., hospitals
- Estimating impact of changes such as downsizing
  - E.g., nursing staffing in critical care units
- Estimating effectiveness of new structures
  - E.g., SSG – Concargru, Army Unit of Action
- Evaluating risks in organizational designs
  - E.g., NASA, Counter-terrorism
- Examining impact of IT on effectiveness
  - E.g., NASA, Knowledge Wall in JTF
- Impact analysis of actions in asymmetric warfare situations
- Impact on cities of weaponized biological attacks
- Estimating size/shape of covert networks
- Destabilization of covert networks

CASOS

Carnegie Mellon

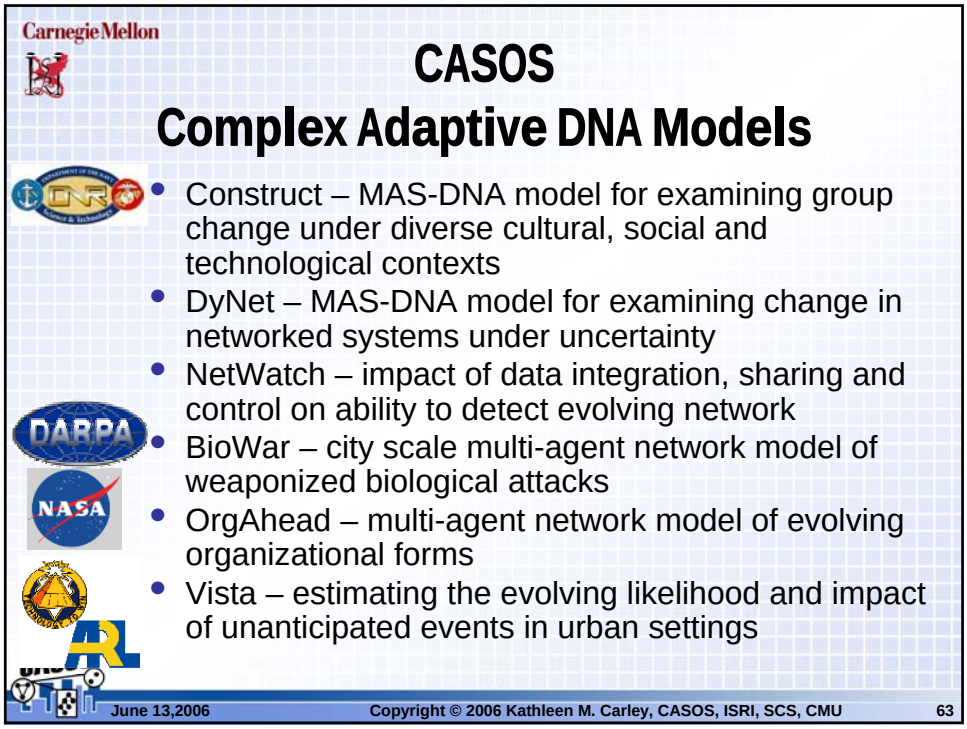
Tools (CMU & *Others*)

- **Automap – automated text analysis tool – converting text to networks**
- **ORA – DNA statistical tool kit**
- **Construct – impact of IT on groups and organizations**
- ***UCINET – Social network statistical tool kit***
- ***KeyPlayer – Identification of elite in social network***
- **OrgAhead – multi-agent network model of evolving organizational forms**
- **DyNet – evaluation of network dynamics & destabilization policies**
- **BioWar – city scale multi-agent network model of weaponized biological attacks**

Carnegie Mellon

CASOS Tools

- ORA – statistical toolkit for meta-matrix, identifies vulnerabilities, key actors (including emergent leaders), and network characteristics of groups, teams, organizations, C2 – used with army battle labs, risk estimation NASA
- DyNetML – XML based interchange language for relational data
- AutoMap – Semi-automated text analysis
- Social Insight – network visualization


Carnegie Mellon
CASOS
Complex Adaptive DNA Models
Construct – MAS-DNA model for examining group change under diverse cultural, social and technological contexts
DyNet – MAS-DNA model for examining change in networked systems under uncertainty
NetWatch – impact of data integration, sharing and control on ability to detect evolving network
BioWar – city scale multi-agent network model of weaponized biological attacks
OrgAhead – multi-agent network model of evolving organizational forms
Vista – estimating the evolving likelihood and impact of unanticipated events in urban settings
DARPA
NASA
ARL
June 13,2006
Copyright © 2006 Kathleen M. Carley, CASOS, ISRI, SCS, CMU
63