On Agent-Based Models of Mode Social Networks

Yasmin H. Said
Edward J. Wegman
Walid K. Sharabati

George Mason University
and Purdue University

Swarmfest 2009
June 28, 2009

1. Introduction
   - Network Terminology
2. Network Theory
   - Two-Mode and One-Mode Networks
   - Three-Mode Networks
   - $N-$Mode Networks
3. Conclusion
   - Closing Remarks

What is a (Social) Network?

- Network: Collection of entities (actors) together with a set of relations defined on them.

What is a (Social) Network?

- Network: Collection of entities (actors) together with a set of relations defined on them.
- Graphs are mathematical structures to visualize networks. Actors are represented by vertices and interactions between actors are represented by edges (ties).
  - A graph $G(V, E)$ is a set of vertices V and edges E.
  - $V = \{v_1, v_2, \cdots, v_{|V|}\}$, $E = \{e_1, e_2, \cdots, e_{|E|}\}$.

What is a (Social) Network?

- Network: Collection of entities (actors) together with a set of relations defined on them.
- Graphs are mathematical structures to visualize networks. Actors are represented by vertices and interactions between actors are represented by edges (ties).
  - A graph $G(V, E)$ is a set of vertices V and edges E.
  - $V = \{v_1, v_2, \cdots, v_{|V|}\}$, $E = \{e_1, e_2, \cdots, e_{|E|}\}$.
- Adjacency matrix is a mathematical tool to quantify a network.

What is a (Social) Network?

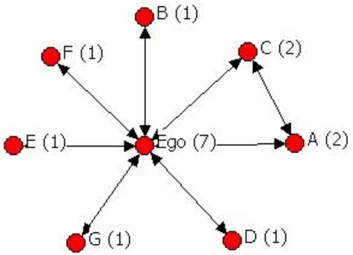
- Network: Collection of entities (actors) together with a set of relations defined on them.
- Graphs are mathematical structures to visualize networks. Actors are represented by vertices and interactions between actors are represented by edges (ties).
    - A graph $G(V, E)$ is a set of vertices V and edges E.
    - $V = \{v_1, v_2, \cdots, v_{|V|}\}$, $E = \{e_1, e_2, \cdots, e_{|E|}\}$.
- Adjacency matrix is a mathematical tool to quantify a network.
- Network Theory: Science of analyzing relations and linkages among actors within a network or among networks.

What is a (Social) Network?

- Network: Collection of entities (actors) together with a set of relations defined on them.
- Graphs are mathematical structures to visualize networks. Actors are represented by vertices and interactions between actors are represented by edges (ties).
    - A graph $G(V, E)$ is a set of vertices V and edges E.
    - $V = \{v_1, v_2, \cdots, v_{|V|}\}$, $E = \{e_1, e_2, \cdots, e_{|E|}\}$.
- Adjacency matrix is a mathematical tool to quantify a network.
- Network Theory: Science of analyzing relations and linkages among actors within a network or among networks.
- Network Types:
    - One-mode vs Multi-mode Networks: One class of entities vs several classes of entities.
    - Stationary Networks vs Evolutionary Networks. Static vs Dynamic. Time-Series.
    - Binary (Dichotomous) vs Weighted (Valued).
    - Social vs other types.

Graph vs Matrix Representations

One-to-one correspondence between graphs and adjacency matrices.

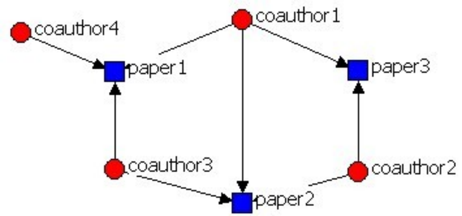


	Ego	A	B	C	D	E	F	G
Ego	0	1	1	1	1	1	1	1
A	1	0	0	1	0	0	0	0
B	1	0	0	0	0	0	0	0
C	1	1	0	0	0	0	0	0
D	1	0	0	0	0	0	0	0
E	1	0	0	0	0	0	0	0
F	1	0	0	0	0	0	0	0
G	1	0	0	0	0	0	0	0

One-Mode and Two-Mode Networks

- One-mode: One set of vertices.
- Two-mode (or also known as bipartite): Two sets of vertices.

Example (Co-authorship Networks)



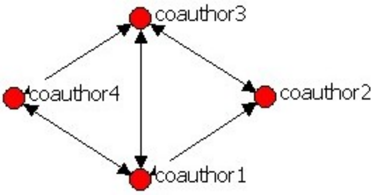


Figure: Left: Two-mode author-by-paper social network. Right: One-mode author-by-author social network.

One-Mode and Two-Mode Networks

Example (Continued)

The author-by-paper adjacency matrix *AP* for the previous example is

	paper1	paper2	paper3
coauthor1	1	0	1
coauthor2	0	1	1
coauthor3	1	1	0
coauthor4	1	0	0

$$\Rightarrow AP = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}_{4\times 3}$$

The author-by-author adjacency matrix *AA* for the previous example is

	author1	author2	author3	author4
author1	0	1	1	1
author2	1	0	1	0
author3	1	1	0	1
author4	1	0	1	0

$$\Rightarrow AA = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

Mathematical Model

Let AP be the adjacency matrix of size $m \times n$ representing the graph of the network, with $m =$ number of coauthors, and $n =$ number of papers. Then,

$$AA_{m\times m} = AP_{m\times n} \cdot AP^T_{n\times m} = \text{ coauthorship adjacency matrix, and}$$

$$PP_{n\times n} = AP^T_{n\times m} \cdot AP_{m\times n} = \text{ paper-by-paper adjacency matrix.}$$

where,

$$aa_{ii} = \sum_{j=1}^{n} ap_{ij} = \text{number of papers author } i \text{ published,}$$

$$pp_{jj} = \sum_{i=1}^{m} a_{ij} = \text{number of coauthors coauthored paper } j\text{, and}$$

$$aa_{ij} = \text{edge weight (tie-strength) between coauthors } i \text{ and } j.$$

If $D_{m\times m} = AA^2_{m\times m}$ then $d_{ii} =$ vertex degree of coauthor i.

Formal Model

Definition

Consider the bipartite graph $G\left(V^a, V^b, E\right)$ with sets of vertices $V^a = \left\{v_1^a, v_2^a, \cdots, v_i^a, \cdots v_{|V^a|}^a\right\}$ of type A; $|V^a| = n$, $V^b = \left\{v_1^b, v_2^b, \cdots, v_j^b, \cdots v_{|V^b|}^b\right\}$ of type B; $|V^b| = m$, and a set of edges $E = \left\{e_1, e_2, \cdots, e_{|E|}\right\}$ connecting vertices of types A and B respectively.

The adjacency matrix $AB_{n\times m}$, also known as Edmonds matrix, is defined by

$$ {}^{w}AB_{n\times m} = \begin{cases} ab_{ij}, & \left(v_i^a, v_j^b\right) \in E \\ 0, & \left(v_i^a, v_j^b\right) \notin E. \end{cases} $$

$1 \le i \le n, \quad 1 \le j \le m$, and the indeterminate $ab_{ij} \in \mathbb{R}$.
If AB is binary then $ab_{ij} = 1$.

Notation: ${}^{w}AB$, ${}^{b}AB$.

Methodology

- We can mathematically reduce a two-mode network to a one-mode network.

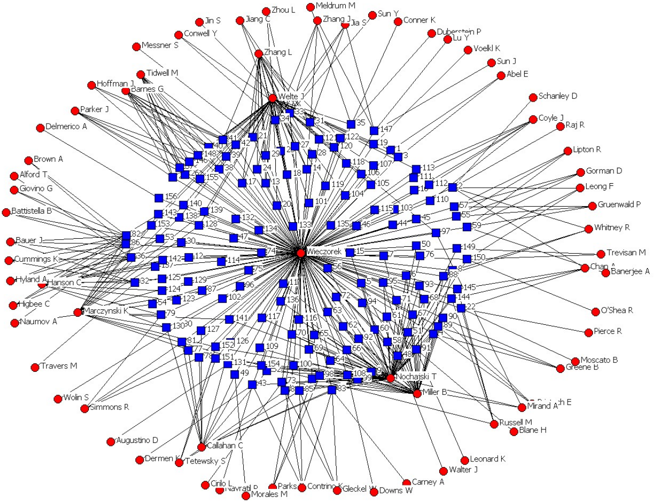
Relational Networks

Let ${}^{b}AB$ be a two-mode **binary** relational matrix of size $n \times m$ with respect to types A and B respectively, then

$${}^{w}AA_{n\times n} = {}^{b}AB_{n\times m} \cdot {}^{b}AB^{T}_{n\times m}$$

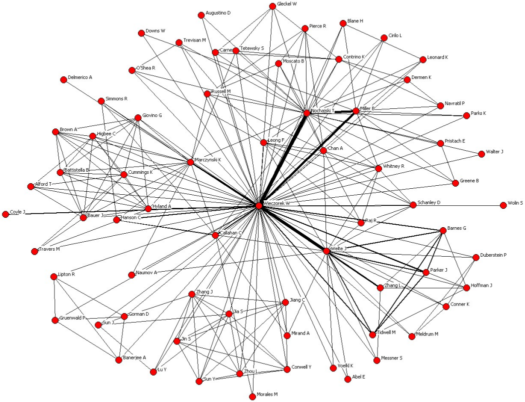
is the one-mode weighted matrix of vertices of type A related through vertices of type B.

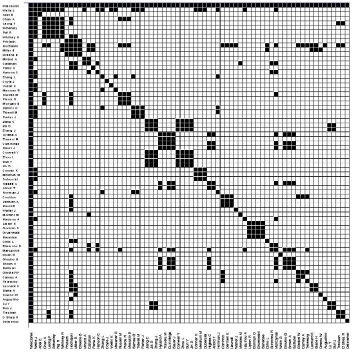
Two-mode Coauthorship Social Network



One-mode Coauthorship Social Network

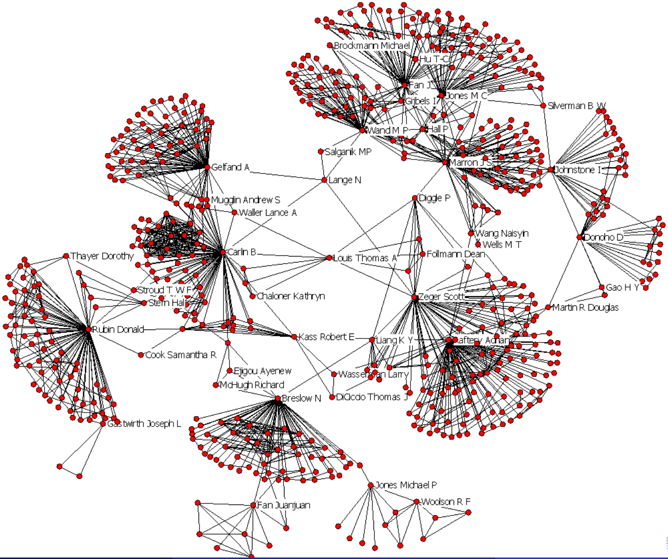
$$^{w}AA =^{b} AP \times^{b} AP^{T}$$



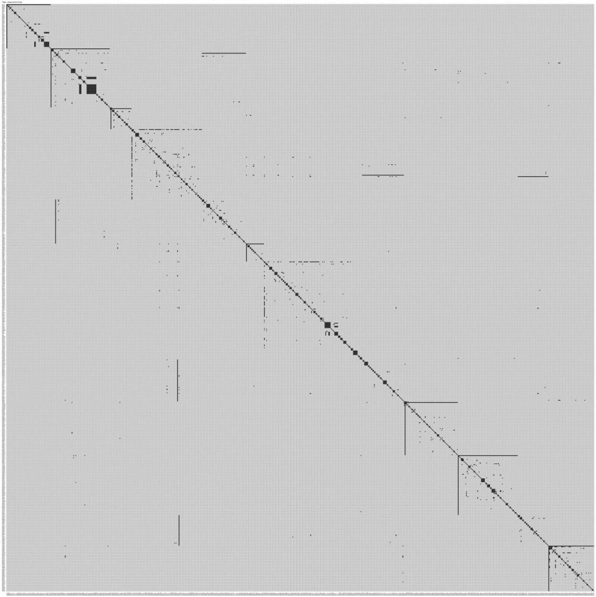


Another Coauthorship Social Networks

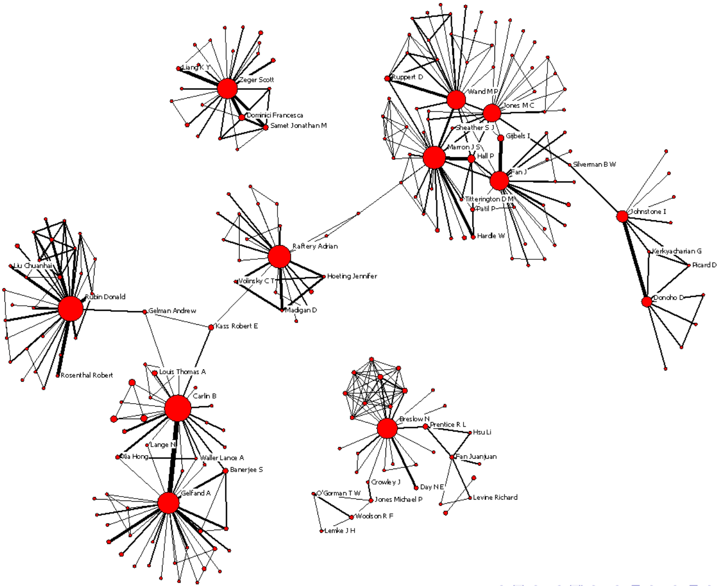
Source: CIS database. 1767 published papers and 874 unique authors.



Coauthorship Social Networks



Coauthorship Social Networks



Modeling Tripartite Networks

Definition

Let $V^a = \left\{v_1^a, v_2^a, \cdots, v_i^a, \cdots, v_{|V^a|}^a\right\}$ be the set of vertices of type a, $V^b = \left\{v_1^b, v_2^b, \cdots, v_j^b, \cdots, v_{|V^b|}^b\right\}$ be the set of vertices of type b, and $V^c = \left\{v_1^c, v_2^c, \cdots, v_k^c, \cdots, v_{|V^c|}^c\right\}$ be the set of vertices of type c.
Furthermore, let $E = \left\{e_1, e_2, \cdots, e_{|E|}\right\}$ be the set of edges connecting types a, b, c vertices; e_i is a hyperedge. Assume $|V^a| = n$, $|V^b| = m$, and $|V^c| = p$.
The binary adjacency matrix ${}^bABC_{n\times m\times p}$ corresponding to the finite graph $G\left(V^a, V^b, V^c, E\right)$ is defined by

$$abc_{ijk} = \begin{cases} 1, & \left(v_i^a, v_j^b, v_k^c\right) \in E \\ 0, & \text{otherwise.} \end{cases}$$

$1 \leq i \leq n, \quad 1 \leq j \leq m, \quad 1 \leq k \leq p.$

Manipulating Three-Mode Networks

- A cuboid is a three-dimensional object.
- More network features are revealed.
- There are several matrix arithmetic operations to perform on a cuboid matrix.
- Some result in a rectangular matrix; others result in a cuboid matrix.

Example

- A network of Author-by-Paper-by-Institution.
- A network of Alcohol outlets-by-Zip code-by-Income.

2D Projection

Consider the tripartite dichotomous matrix ${}^{b}ABC_{n\times m\times p}$.

${}^{b}ABC_{n\times m\times p}$ is a tensor of rank 3; I will use the term cuboid instead.

The marginal bipartite weighted matrix ${}^{w}AB_{n\times m}$ is

$$ {}^{w}ab_{ij} = \sum_{k=1}^{p} abc_{ijk}, \quad 1 \leq k \leq p. $$

- This is equivalent to projecting the 3D cuboid matrix onto the 2D planar matrix.
- The result is the marginal bipartite distribution for vertices of types A and B.

3D Transpose

- Unlike a 2D rectangular matrix (has only two faces), a 3D cuboid has six faces.
- Six different ways to view the block in terms of size; namely, $n \times m \times p$, $n \times p \times m$, $m \times n \times p$, $m \times p \times n$, $p \times n \times m$, and $p \times m \times n$.
- 3D transpose can be performed in six different ways.

Example (3D Transpose)

$$ABC^{T_{cba}}_{n\times m\times p} = CBA_{p\times m\times n}, \text{ with } cba_{ijk} = abc_{kji},$$

and

$$ABC^{T_{bac}}_{n\times m\times p} = BAC_{m\times n\times p}, \text{ with } bac_{ijk} = abc_{jik}.$$

3-D Two-Mode Network

Traditional 3D matrix product giving another 3D matrix.

$$^{w}AAC_{n\times n\times p} = {}^{b}ABC_{n\times m\times p} \cdot {}^{b}ABC^{T_{bac}}_{n\times m\times p},$$

$AB_{n\times m} \cdot BA_{m\times n}$ must be well-defined,

$$^{w}aac_{k} = {}^{b}ABC_{k} \cdot {}^{b}BAC_{k}, \quad 1 \leq k \leq p.$$

$^{w}AAC_{n\times n\times p}$ is a two-mode graph (network) represented by a 3D matrix.

3-D One-Mode Network

Given ${}^{w}AAC_{n\times n\times p}$.

$$ {}^{w}AAA_{n\times n\times n} = {}^{w}AAC_{n\times n\times p} \cdot {}^{w}AAC^{T_{caa}}_{n\times n\times p}, $$

$AC_{n\times p} \cdot CA_{p\times n}$ must be well-defined,

$$ {}^{w}aaa_{i} = {}^{w}AAC_{i} \cdot {}^{b}CAA_{i} \ . $$

- ${}^{w}AAA_{n\times n\times n}$ is the one-mode 3D matrix of triadic vertices (triplets) of type A related through vertices of types B and C.
- ${}^{w}AAA_{n\times n\times n}$ contains hyperedges connecting three vertices.

The Hyper Product $A \circ B$

Hyper Cuboid

$$^{w}ABBA_{n\times m\times m\times n} = {}^{b}ABC_{n\times m\times p} \circ {}^{b}ABC^{T_{cba}}_{n\times m\times p},$$

$BC_{m\times p} \cdot CB_{p\times m}$ must be defined.

$$^{w}abba_{kl} = ^{b} ABC_{kl} \cdot ^{b} CBA_{kl}, \quad 1 \leq k, l \leq p.$$

- $^{w}ABBA_{n\times m\times m\times n}$ is the hyper-cuboid two-mode adjacency matrix of pair of vertices (v_i, v_j) of types A and B related through the set of vertices v_k of type C.
- The 4D hyper-matrix can be represented using 3D matrix by stacking n cuboid matrices each of size $m \times m \times n$, resulting in a 3D matrix of size $m \times m \times n^2$.

From Three-Mode to One-Mode in One Step

2-D One-Mode Network

$$^{w}CC_{p\times p} = {}^{b}ABC_{n\times m\times p} \odot {}^{b}ABC^{T_{bac}}_{n\times m\times p},$$

$ABC_{n\times m} \cdot BAC_{m\times n}$ must be well-defined,

$$^{w}cc_{ij} = \sum_{q=1}^{n}\sum_{r=1}^{n} {}^{b}ABC_{(n\times m)_i} \cdot {}^{b}BAC_{(m\times n)_j}$$

$$= \sum_{q=1}^{n}\sum_{r=1}^{n} {}^{w}AA_{(qr)_{ij}}, \quad 1 \leq i,j \leq p.$$

- $^{w}CC_{p\times p}$ is the one-mode 2D matrix of pairwise vertices of type C related through vertices of types A and B.
- $^{w}CC_{p\times p}$ represents diadic relations.

Generalized $N-$Mode Model

- $N-$hyper cuboid matrix (tensor of rank N).
- Multidimensional Networks.
- Use hyper graphs (cliques).
- Lower mode relations may be retrieved in the same fashion we convert three-mode to two-mode and two-mode to one-mode networks.

Generalized $N-$Mode Model

Definition

Let $V^1 = \left\{v_1^1, v_2^1, \cdots, v_{i_1}^1, \cdots, v_{|V^1|}^1\right\}$, $V^2 = \left\{v_1^2, v_2^2, \cdots, v_{i_2}^2, \cdots, v_{|V^2|}^2\right\}$, $V^3 = \left\{v_1^3, v_2^3, \cdots, v_{i_3}^3, \cdots, v_{|V^3|}^3\right\}, \cdots,$ $V^N = \left\{v_1^N, v_2^N, \cdots, v_{i_N}^N, \cdots, v_{|V^N|}^N\right\}$ be the sets of vertices of type $1, 2, 3, \cdots, N$ respectively. Furthermore, let $E = \left\{e_1, e_2, \cdots, e_{|E|}\right\}$ be the set of edges connecting types $1, 2, 3, \cdots, N$ vertices. Once again, e_i is a hyperedge. Assume $|V^i| = n_i, \forall\ 1 \le i \le N$.

The binary adjacency matrix $A_{n_1 \times n_2 \times \cdots \times n_N}$ for the finite graph $G(V^1, V^2, \cdots, V^N, E)$ is defined by

$$a_{i_1 i_2 \cdots i_N} = \begin{cases} 1, & \left(v_{i_1}^1, v_{i_2}^2, \cdots, v_{i_N}^N\right) \in E \\ 0, & \text{otherwise.} \end{cases}$$

$1 \le i_j \le n_j$, for $1 \le j \le N$.

Reducing the $N-$mode network to a two-mode network

Consider the hyper-matrix $A_{n_1 \times n_2 \times \cdots \times n_N}$.

Then, ${}^{w}A_{n_i \times n_j}$, $1 \leq i, j \leq N$, $i \neq j$, is the two-mode matrix where,

$$a_{ij} = \sum_{k_1=1}^{n_{k_1}} \cdots \sum_{k_{N-2}=1}^{n_{k_{N-2}}} a_{i_1 i_2 \cdots i_N}, \quad 1 \leq k_l \leq n_l.$$

This is equivalent to projecting from $N-$dimensional hyper-cuboid matrix onto the 2D planar matrix giving the marginal $N-$mode distribution for types i and j.

$N-$D Transpose

Suppose $A_{n_1 \times n_2 \times \cdots \times n_N}$ is a binary $N-$partite matrix. Then,

$$A^{T_{m_1 \times m_2 \times \cdots \times m_N}}_{n_1 \times n_2 \times \cdots \times n_N} = A_{m_1 \times m_2 \times \cdots \times m_N}, \quad a_{i_1 i_2 \cdots i_N} = a_{j_1 j_2 \cdots j_N},$$

where $j_1 j_2 \cdots j_N$ is a permutation of $i_1 i_2 \cdots i_N$.

Reducing the $N-$mode network to a one-mode network

Let ${}^{b}A_{n_1 \times n_2 \times \cdots \times n_N}$ be a binary $N-$mode matrix. Assume $m_1 = n_i$, the $(N-1)-$mode matrix for a mode i is

$$\begin{aligned} {}^{w}A_{m_1 \times m_1 \times m_3 \times \cdots \times m_N} &= {}^{b}A_{m_1 \times m_2 \times m_3 \times \cdots \times m_N} \cdot {}^{b}A^{T_{m_2 m_1 m_3 \cdots m_N}}_{m_1 \times m_2 \times m_3 \times \cdots \times m_N} \\ &= {}^{b}A_{m_1 \times m_2 \times m_3 \times \cdots \times m_N} \cdot {}^{b}A_{m_2 \times m_1 \times m_3 \times \cdots \times m_N}, \end{aligned}$$

The product of the sub-matrices $A_{m_1 \times m_2} \cdot A_{m_2 \times m_1}$ must be well-defined.

Thus, m_2 mode is being eliminated.

The one-mode matrix ${}^{w}A_{m_1 \times m_1 \times \cdots \times m_1}$ for a mode i is obtained in $N-1$ matrix multiplications and transpose.

${}^{w}A_{n_i \times n_i \times \cdots \times n_i} = {}^{w}A_{m_1 \times m_1 \times \cdots \times m_1}$ is the one-mode $N-$dimensional matrix of $N-$cliques of type i related through all other types.

Hyperedge connecting $N-$vertices.

Problem!

The traditional method of multiplying the matrix by its transpose fails to give meaningful one-mode edge weights if the two-mode matrix is not binary; i.e. edges have real values.

Applications

- Online Friendship Social Networks.
  - Facebook, Livejournal, MySpace.
- Coauthorship Social Networks.
  - "Who-Wrote-With-Whom".
- Covert/Espionage (Alliances) Social Networks.
  - "Who-Talked-With-Whom"
- Disease Propagation Social Networks.
- Computer Networks.
- Term-Document Networks.
- Protein Networks.

Current Research

- Developing mathematics underpinning networks
- Infinite Networks and Multidimensional Networks.
- Mechanism to store networks using matrices.
- Estimating diadic relations: Edges and Vertices.
- Estimating triadic relations: Hyperedges.
- Expressing weighted matrices in terms of binary matrices.
- Multi-Mode Networks.
- Modeling the generalized $N-$mode networks on real data.
- Estimate missing edges and vertices.
- Studying relationships between cliques and hypergraphs.
- Evolutionary networks
- Preferential attachment.
- Stochastic models.
- Text mining.

Acknowledgements

This is joint work with Dr. Walid K. Sharabati and Dr. Yasmin H. Said.
This work is based on Dr. Sharabati's Ph.D. dissertation.

Contact Information

Edward J. Wegman
Center for Computational Data Sciences
George Mason University
Fairfax, VA 22030

Email: ewegman@gmu.edu

Shameless Propaganda

Look for the Third International Conference on Social Computing, Behavioral Modeling and Prediction
SBP(2010)

To be held in Washington, DC
March-April 2010

Contacts:
Edward J. Wegman, ewegman@gmail.com
Ysmin H. Said, ysaid99@hotmail.com