Revisited Experimental Comparison of Node-Link and Matrix Representations

Mershack Okoe, Radu Jianu, Stephen Kobourov

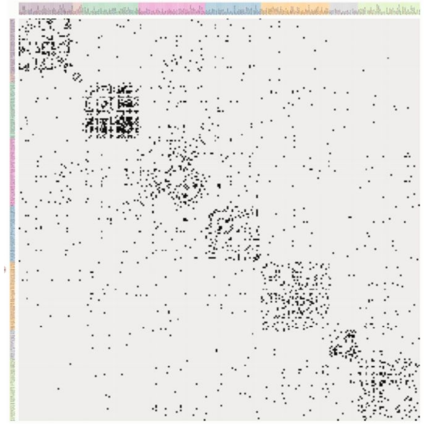
FIU
FLORIDA INTERNATIONAL UNIVERSITY

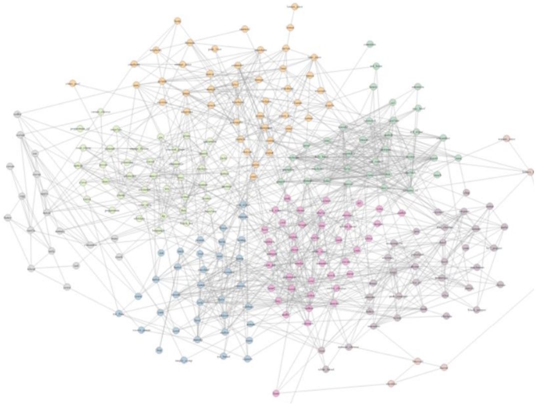
CITY UNIVERSITY LONDON

THE UNIVERSITY OF ARIZONA

Background

- Many studies compare Node-link diagrams (NL) and Adjacency matrices (AM)
    - **Ghoniem et al. (2004)**
    - **Keller et al. (2006)**
    - **Abuthawabeh et al. (2013)**
    - **Alper et al. (2013)**
    - **Christensen et al. (2014)**
- Various aspects of the problem, studied by different groups in various settings:
    - Varying the size of graphs
    - Varying the tasks
    - Varying NL algorithms
    - Varying AM algorithms





Why do we need yet another study of NL and AM?

- Cover a **broad spectrum of tasks**.
- Use new tasks such as **cluster-based** and **memorability** tasks.
- Use a **small-world, clustered, sparse graph,** more similar to real-life networks.
- Measure **beyond time and error** (e.g., memorability).
- Use **basic interactions** (harder to do well, but more realistic than past studies)
- Use **more participants** than typical of such studies

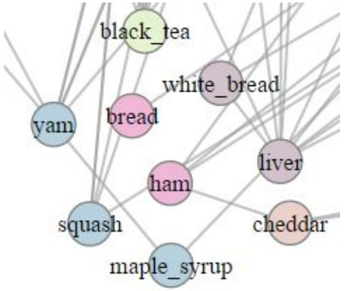
Motivation

- Networks are used to solve increasingly complex problems, and there is an expanding range of tasks that are relevant in real applications.
- Earlier studies show the effectiveness of NL and AM representations depends on the properties of the datasets and the tasks.
- We hypothesize that there might be differences depending on the structure of the network (e.g., clustered, small-world) and for new tasks such as group and memorability tasks).

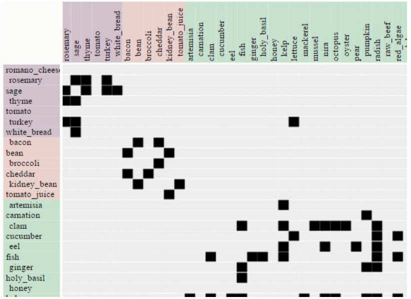
*Compare the effectiveness of **NL** and **AM** on a **broader spectrum** of tasks, using a **large dataset** representative of a **real-life network**, leveraging **crowdsourcing,** going **beyond time and error**.*

Study Design: Data

- A single network with 258 nodes (cooking ingredients) and 1090 edges (ingredients frequently used together in recipes).

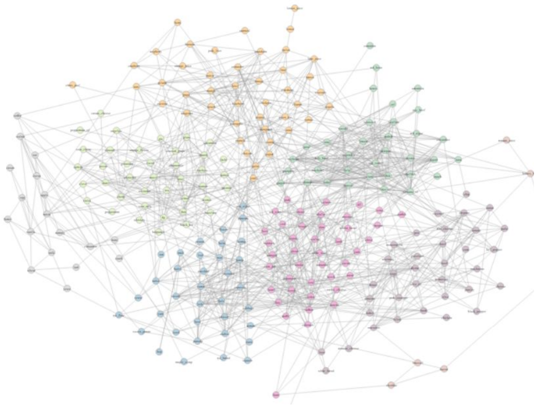


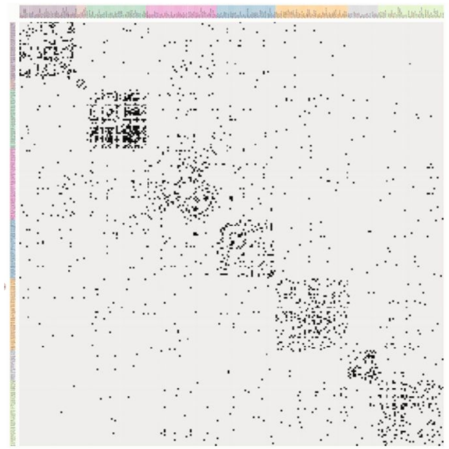
- ***Motivation***
    - Larger graph than those evaluated by prior studies.
    - Representative of many networks found in real life (small-world, sparse).
    - Involves labeled nodes (cooking ingredients): a realistic and relatable example for participants. Instead of using node numbers.



Study Design: Visual Encoding

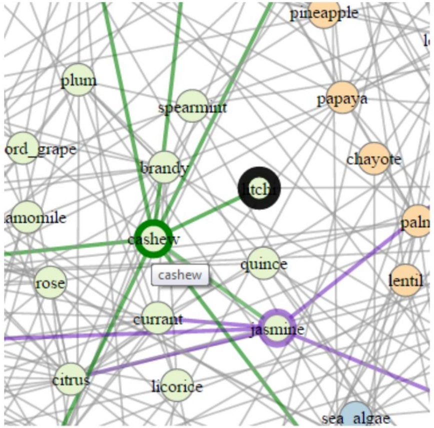
- We evaluated two visual encodings
  - **Node-link diagrams (NL)** drawn using the `neato` algorithm.
  - **Adjacency matrix (AM)** sorted to reveal clusters using the Barycenter algorithm.
- We clustered the network using **modularity clustering** from **GMap** and encoded this information using color.

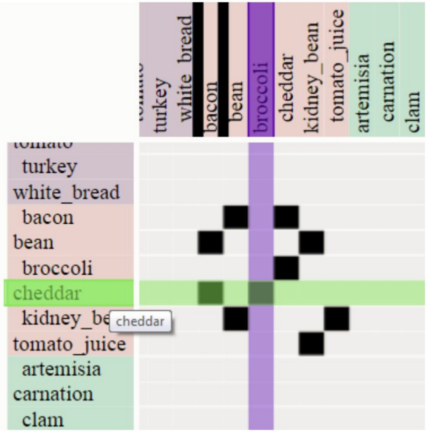




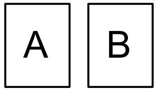
Study Design: Interactions

- Both visual encodings support **panning and zooming**, clicking, hovering, selecting answers.
- Multiple nodes can be selected by clicking on them, and can be deselected.
- Nodes can be moved around in NL.





Study Design: Design



- **Between subjects experiment with**
  - Independent variable: network type (NL and AM)
  - Dependent variables: task accuracy and completion time

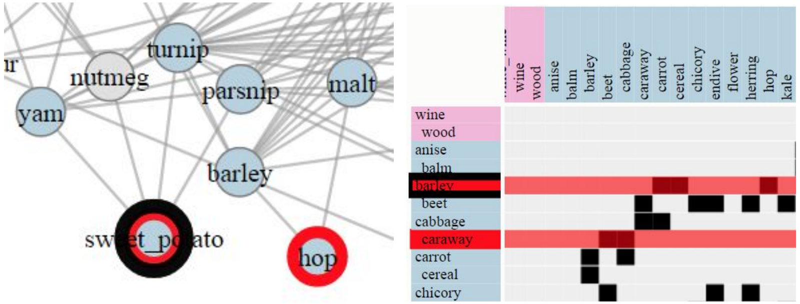
- **Procedure**:
  - Used Amazon Mechanical Turk to crowdsource our study to a broad population.
  - Ran conditions in parallel and directed incoming participants to conditions in a round-robin assignment.
  - Used a color blindness test to filter participants, provided an introduction with sample questions and answers, and instructions on how to interact with the visualization.
  - Provided a training session which involved solving two instances of each type of task.
  - Followed by the main study.

Tasks

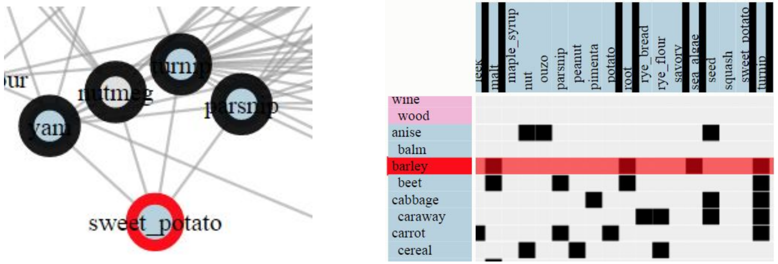
- 14 tasks, divided into 5 experimental groups, covering 3 dimensions

Group	Target	Lee et al. Taxonomy	Amar et al. Taxonomy
1	Node, Edge, **Clusters**, **Cliques**	Topology (adjacency, accessibility), Overview(connectivity)	Retrieve value, Sort, Filter, Cluster
2	Edge, Path	Topology (shared neighbor), Overview (connectivity)	Retrieve value, filter, Derive value, sort
3	**Clusters**, Node	Overview (connectivity), Attribute-based	Derive value, Filter, Sort, Correlate
4	Path, Edge, **Memorability**	Topology (adjacency, connectivity)	Retrieve value, Derive value, Filter
5	Edge, **Memorability**	Topology (shared neighbor, accessibility)	Retrieve value, Derive value, Filter

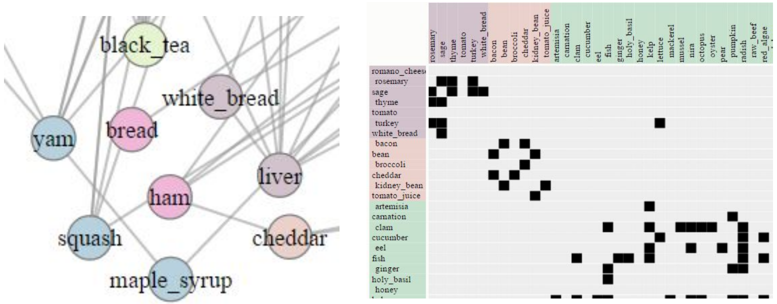
T1: Given two highlighted nodes, select the one with the larger degree (#Instances: 10, Time: 15s).



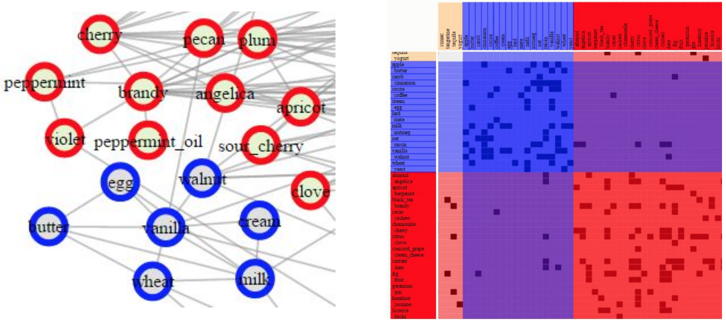
T2: Given a highlighted node, select all its neighbors (#Instances: 10, Time: 25s)



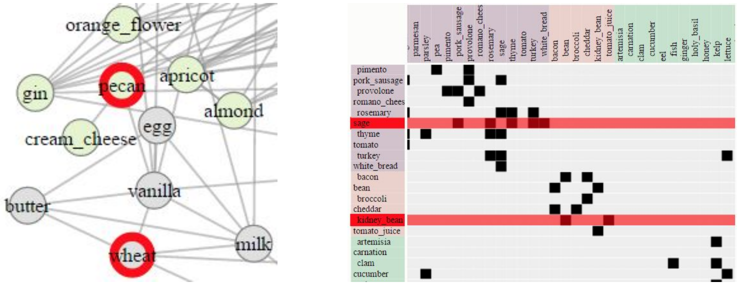
T6: How many clusters are there in the visualization? (#Instances: 1, Time: 10s).



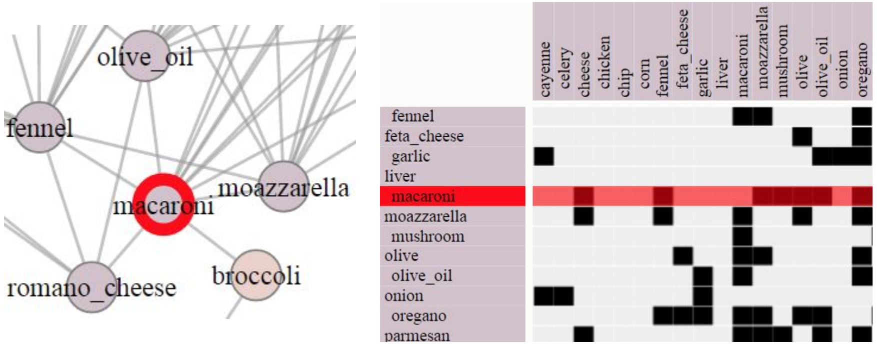
T7: Given two groups of highlighted nodes, estimate which group is larger(#Instances: 10, Time: 10s).



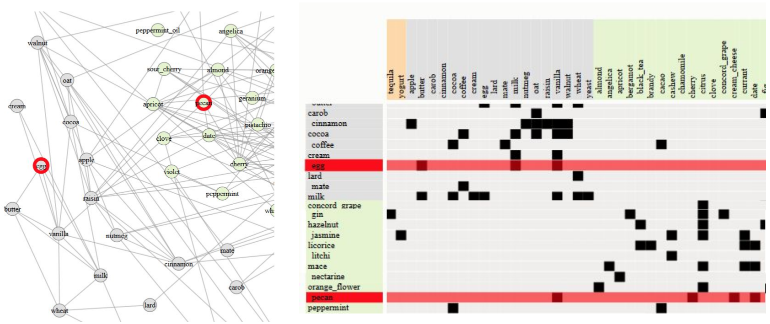
T8: Given two highlighted nodes decide whether they belong to the same cluster (#Instances: 10, Time 10s).



T9: Given one highlighted node and one named node, are they connected? (#Instances: 5, Time: 20).

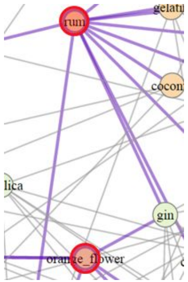


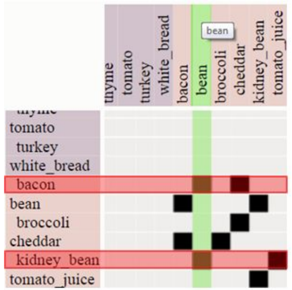
T10: Given two highlighted nodes, how long is the shortest path between them? (#Instances: 5, Time: 60s).



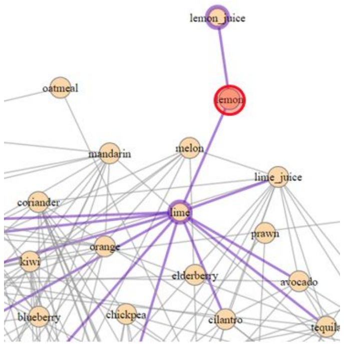
T11: (Memorability) - After spending several minutes on T10, can participants remember the answers they gave to T9, without access to the visualization? (#Instances: 5, Time: unlimited)

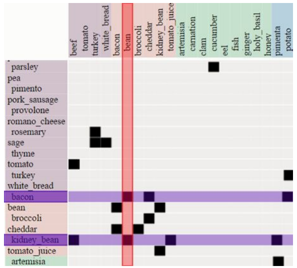
T12: Given two highlighted nodes and three named ones, which of the named nodes is connected to both highlighted nodes? (#Instances: 5, Time: 60s).





T13: Given a selected node, how many nodes are within two edges reach? (#Instances: 5, Time: 60s).





T14: (Memorability) - After spending several minutes on tasks 13, can participants remember which nodes were highlighted as part of task 12, if showed the visualization with the answers they gave to task 13 highlighted? (#Instances: 5, Time: Unlimited).

Number of participants

- We collected responses from a total of **557** individual participants.
- We removed **28** responses (participants who spent at most an **avg. of 2 seconds** on tasks and had **accuracy** in the **bottom 10 percentile**).
- Duration: 10-15mins on average.

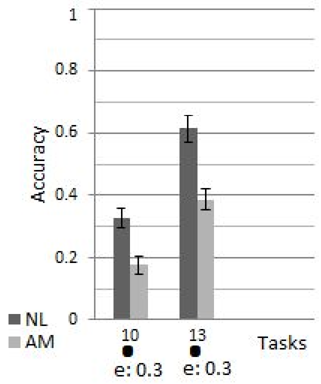
Groups	Condition	User Size	Valid Data
1	NL	65	63
	AM	62	58
2	NL	58	54
	AM	53	50
3	NL	55	53
	AM	55	52
4	NL	52	50
	AM	53	50
5	NL	54	52
	AM	50	47

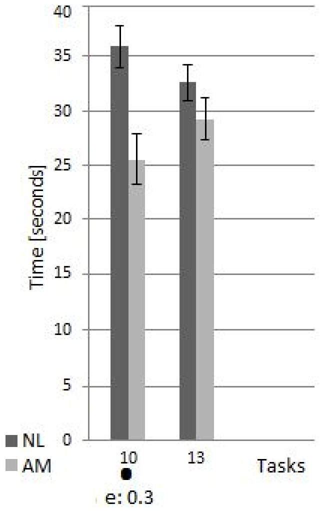
Results: Confirming Previous Claims

NL Wins!

- Ghoniem et al. found AM performs poorly on long path tasks. **T10** and **T13** confirms that.

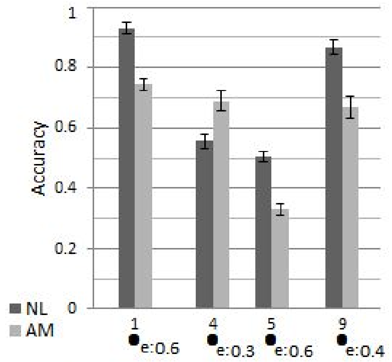
- **Interestingly**, average time of AM is significantly lower than NL in **T10**
    - AM users give up on solving the tasks early on.

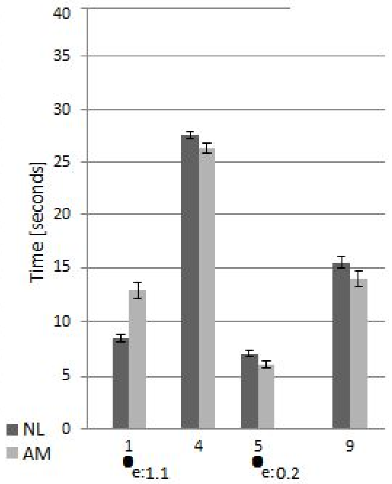




Results: Differing from Previous Claims

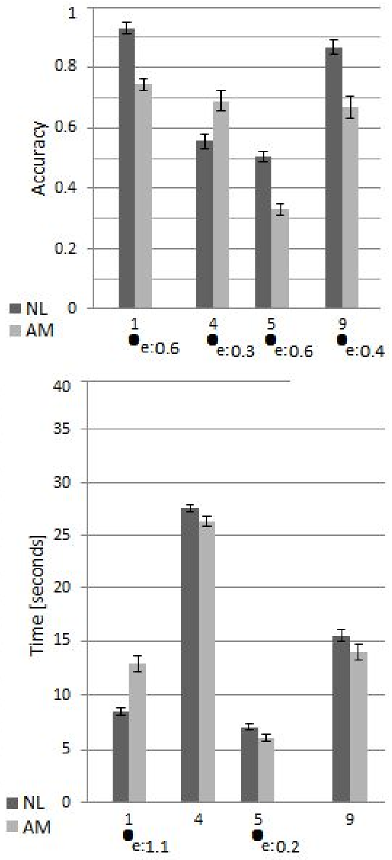
- T1: **NL Wins!**
  - NL required less zooming for nodes to become legible and selected accurately.
  - Matrices favor dense networks and not sparse ones
- T4: **AM Wins!**
  - AM eliminate occlusion and ambiguity problems
  - Occlusion is common in NL.
- T5: **NL Wins!**
  - NL places nodes so that their network distance matches their embedded distance.
  - Matrices are constrained by a single dimension





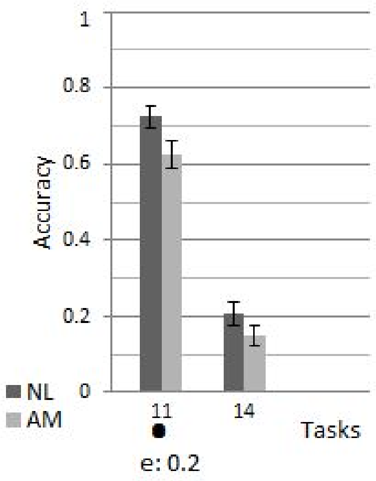
Results: Differing

- T9: **NL Wins!**
  - NL represents nodes and connections together
  - Finding endpoints of nodes in AM involves horizontal and vertical traces.

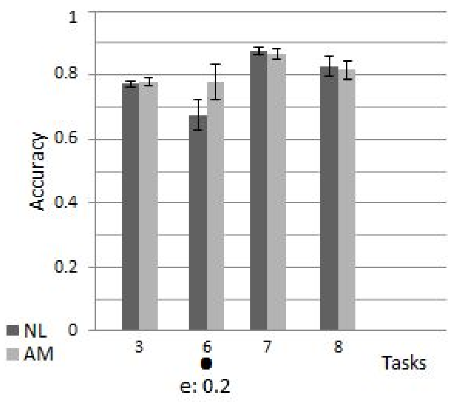


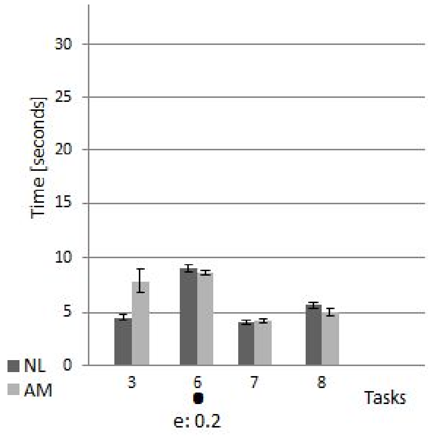
Results: New

- Memorability tasks: **NL Wins!**



- Group tasks: **NL and AM Tied**





Summary of Results

- NL outperforms AM for most types of connectivity tasks.
- NL and AM give similar results for group tasks
  - Except one in which AM outperforms NL.
  - AM is better for estimating the number of clusters rather than their interconnectivity.
- NL outperforms AM results on memorability tasks.
- NL can be more compact than AM (especially for sparse graphs).
- NL draws a node’s glyph and connections together.
- AM eliminate some occlusions and ambiguity problems.

Discussion: Limitations

there are many limitations to this work...

- We used one type of network and a single instance of graph structure.
  - We used this approach due to the overhead associated with preparing multiple appropriate real-world networks and multiple dataset specific tasks.
- Density of our network was lower than that of Ghoniem et al. and Keller et al.
  - But networks of similar density are quite common, Melancon (2006).
- Visualizations were interactive and it is difficult to ensure that all interactions are fair to both visualizations.
  - For example moving a node can be done in NL and not in AM. We opted for ecological validity.
- Study participants were crowdsourced. Inherent crowdsourced limitations include difficulty in controlling what the participants do.
  - However, crowdsourcing has been used for studies extensively in Vis and have been used to replicate several lab studies.

Conclusions

- Interesting results confirming old observations
- Interesting results contradicting old observations
- New results on cluster-based tasks and memorability tasks
- Potential for more and better NL and AM comparisons
- Most importantly for the GD community, NL wins or ties AM for most tasks

Thank You!

Questions?