

IBM i2 Analyst's Notebook Social Network Analysis

Discover new insights and actionable intelligence from social network data



Contents

- 2 Introduction
- 2 Who should read this white paper
- 2 The background of Social Network Analysis
- 3 Why use Social Network Analysis?
- 4 Social Network Analysis—an analytical tool, not an answer
- 4 A description of the Social Network Analysis measures implemented in Analyst's Notebook
- 11 Where Social Network Analysis functions are located
- 14 SNA glossary
- 15 Technical description

IBM® i2® Analyst's Notebook® delivers powerful assisted analysis and visualization capabilities designed to help increase analyst productivity and reduce the time required to deliver high value intelligence within quickly growing data sets.

IBM has implemented Social Network Analysis (SNA) capabilities within Analyst's Notebook. These can be used with existing Analyst's Notebook functions to help examine and analyze group structures and communication flows within networks. This enables users to better understand relationships between entities in Analyst's Notebook charts.

This white paper describes: a brief background on the origins of SNA, how it can be used to conduct intelligence analysis, a description of the SNA measures implemented in Analyst's Notebook and where the key features of this function are located.

Providing quantitative techniques that help users gain better insight and understanding of complex networks of entities.

Who should read this white paper

This white paper is intended for users who are familiar with Analyst's Notebook and wish to learn about the SNA functions delivered by Analyst's Notebook. Managers can also benefit from learning about these new features so that they can assign tasks to their analysts more effectively.

This document relates to Analyst's Notebook version 9.0.0.

Combines organizational theories with mathematical models to map complex networks by measuring and weighting the interactions within them.

The background of Social Network Analysis

Social Network Analysis (SNA) emerged from Social Sciences as a useful method to enable academics to study how and why social groups operate, interact and behave in particular ways. This quantitative technique enables mapping and measuring complex networks of entities such as people and organizations, by measuring the interactions between them. It is suggested that SNA may help people to understand and predict future network behaviors, such as their likely courses of actions and intentions in certain situations.

SNA combines organizational theories with mathematical models to help people to better understand the dynamics of groups and organizations in which they are interested. The structure of a network can determine:

- The performance of the network as a whole and its ability to achieve its key goals
 - Network characteristics that are not immediately obvious, such as the existence of smaller sub-networks operating within the network
 - The relationships between prominent people of interest whose position may provide the greatest influence over the rest of the network
 - How directly and quickly information flows between people in different parts of the network
- Looking beyond the network structure into its dynamics— To identify characteristics of networks that are not immediately apparent and to also analyze how those networks change over time
 - Finally, our users recognize that in social networks not all connections are equal. They need to be able to use methods such as weighting relationships between entities to take account of how such links affect a network

SNA techniques may help overcome such issues by providing users with the means to optimize their understanding of the data that they collect. This may then enable them to better evaluate future courses of action against target networks such as how to disrupt and destabilize networks more efficiently.

Why use Social Network Analysis?

During conversations with customers, we noted a number of real-world issues:

- Data deluge—Users are dealing with ever-growing data sets, which means that they need capabilities that can help to filter network information faster and more efficiently
- Understanding target networks—Due to the dynamism of target networks, users need to quickly identify potential key individuals/groups for better prioritization of often limited resources

Social Network Analysis— An analytical tool, not an answer

SNA is an analytical tool in the intelligence analysis toolbox, not a “silver bullet”. It merely provides users with a starting point for areas that warrant further analysis. The use of SNA can be extremely effective when used as an aid to human analytical judgment. However users need to be mindful that their data could be incomplete and may not take into account the wider context of those networks.

A description of the Social Network Analysis measures implemented in Analyst's Notebook

Centrality

Centrality is a key concept in SNA. A highly centralized network is dominated by one person who controls information flow and may become a single point of communication failure. A less centralized network has no single point of failure, so people can still pass on information even if some communication channels are blocked.

Analyst's Notebook allows users to calculate betweenness, closeness, degree and eigenvector (including hub and authority) centrality measures to provide different perspectives on the social relationships within the network. It is also possible to further refine centrality measures by taking into account the direction of links and the weightings applied to them.

Betweenness

Betweenness centrality measures the number of paths that pass through each entity. This may identify entities with the ability to control information flow between different parts of the network. These are called gatekeeper entities. Gatekeepers may have many paths running through them, allowing them to channel information to most of the others in the network. Alternatively, they may have few paths running through them but still play a powerful communication role if they exist between different network clusters.

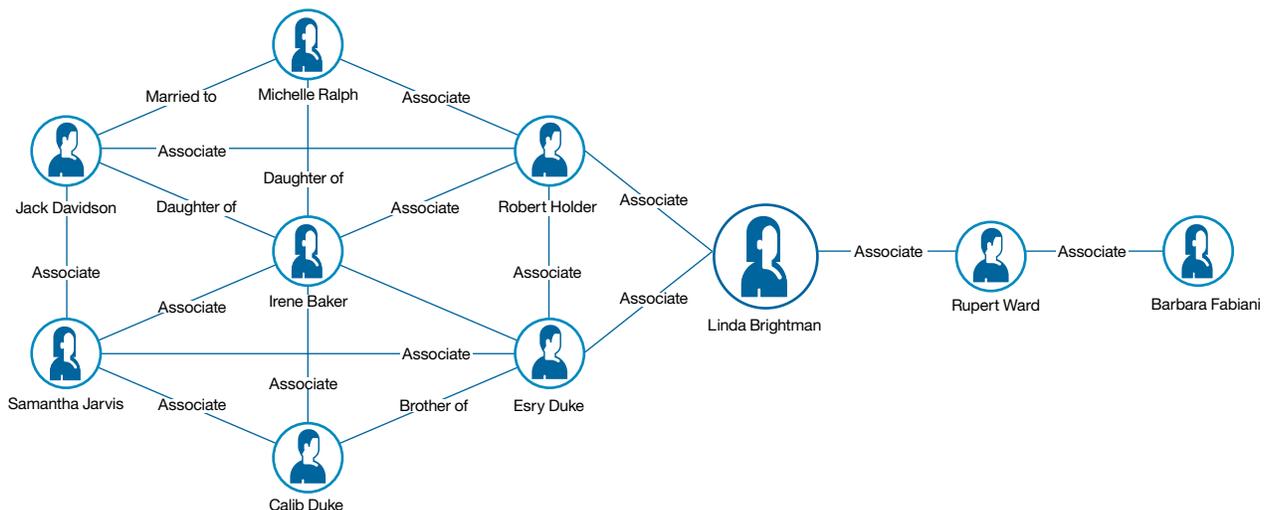


Figure 1: Betweenness—Linda Brig is the person with the highest betweenness score as she is the link between two distinct parts of the network.

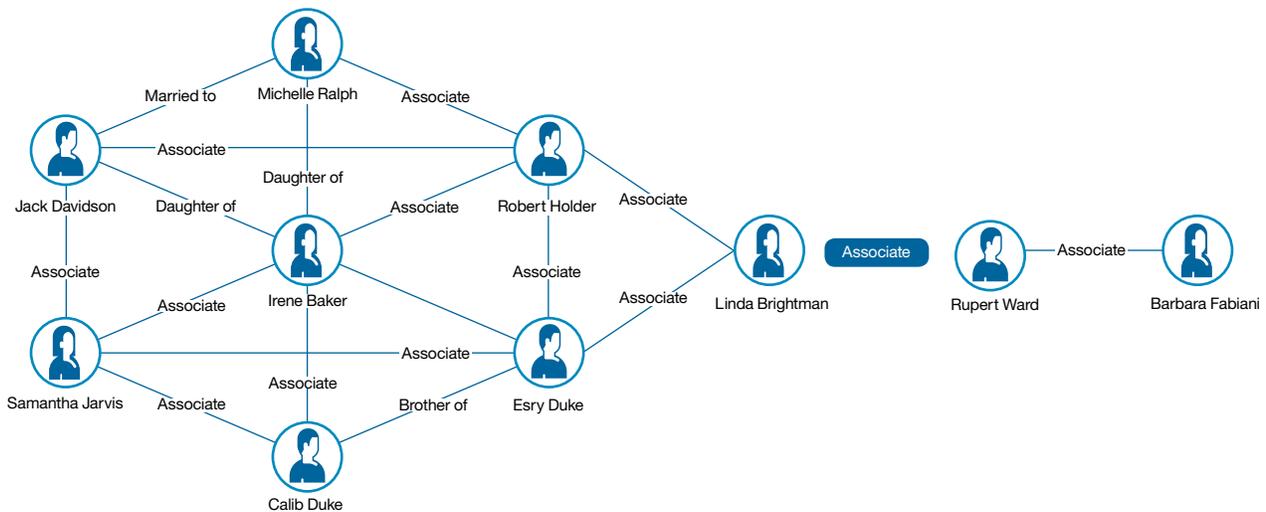


Figure 2: Link Betweenness—the key connection of influence is highlighted in red. If this link were broken, a key channel for information exchange may no longer operate and the network could be separated into two distinct parts.

Link Betweenness

Link betweenness centrality measures the number of paths that pass through each link. This can help identify key connections of influence within the network. A link through which many paths pass may be a significant route for information exchange between entities.

Closeness

Closeness centrality measures the proximity of an entity to other entities in the social network. An entity with a high measure of closeness centrality has the shortest paths to the other entities, allowing them to pass on and receive

communications more quickly than others in the organization. Information travels much further to and from an entity on the edge of a network that is attached to few other entities, so they will have a lower measure of closeness centrality. Closeness centrality measures both direct and indirect closeness:

- Direct closeness occurs when two entities are connected by a link
- Indirect closeness exists when information can only pass from one entity to another via a path that runs through one or more entities

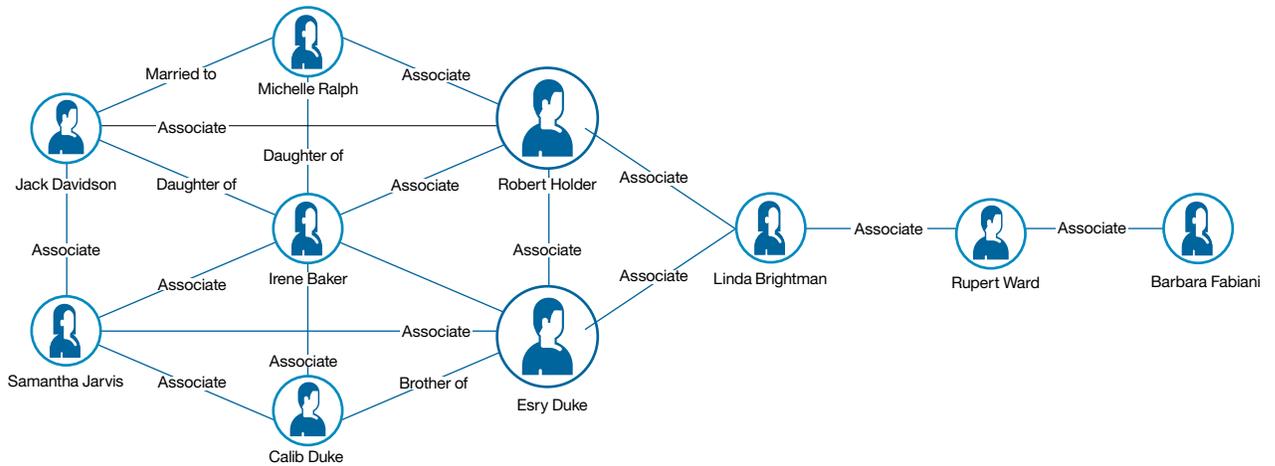


Figure 3: Closeness—Esry Duke and Robert Holder have the highest closeness score; they have the best access to the majority of other members of the network.

Degree

Degree centrality measures how well connected an entity is by counting the number of direct links each entity has to others in the network. This can reveal how much activity is going on and who are its most active members.

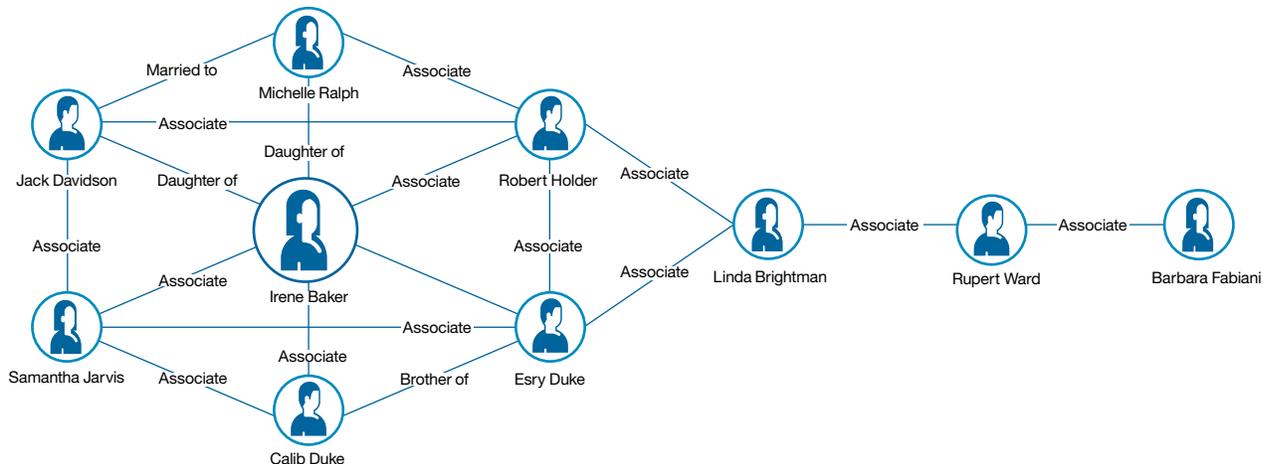


Figure 4: Degree—Irene Baker is the most central person in the network. She has the highest number of connections to other people in the network.

Eigenvector

Eigenvector centrality measures how well connected an entity is and how much direct influence it may have over the most active entities in the network. It does this by taking into consideration the centrality scores of the entities to which it is connected. For example, a person with high eigenvector centrality is likely to be at the center of a cluster of key entities that also have high centrality. That person can communicate directly with those key entities compared with a person with a low eigenvector score on the periphery of the network.

Hubs and authorities are the terms used to describe the two eigenvector centrality scores calculated in networks containing directed links. Hubs refer to the scores for outbound links, and authorities refer to the scores for inbound links. There is a reciprocal relationship between the two. A high-scoring hub has many outbound links to high-scoring authorities, and a high-scoring authority has many inbound links from high-scoring hubs.

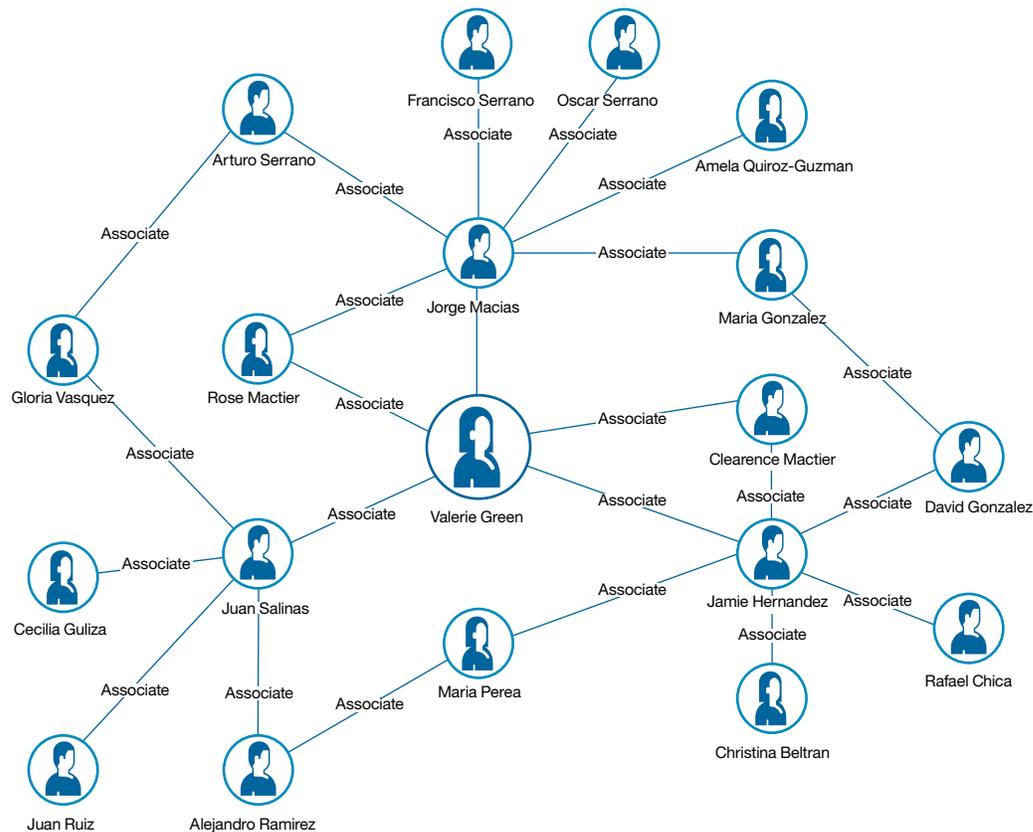


Figure 5: Eigenvector—Valerie Green has the highest measure of eigenvector centrality because she is connected to entities that are the most active in the network. Her position at the heart of the central cluster in the network means that she has more direct links to key entities than any other entity. She may exercise influence over them more quickly than anyone else.

Link Direction

The use of link direction on a chart is often useful in assessing how information and commodities flow through a network. A link with arrows added to it represents the directed flow of information between entities; either in a single direction or in both directions. This may have an important bearing on how quickly information is passed from one part of the network to

another. For example, a person may receive information from many others in the network but only send information to a select few. The centrality measures for an entity through which information is channeled in both directions will be higher than the measures for an entity through which information is channeled one way. Directed links can be included in the calculation of centrality measures against network charts.

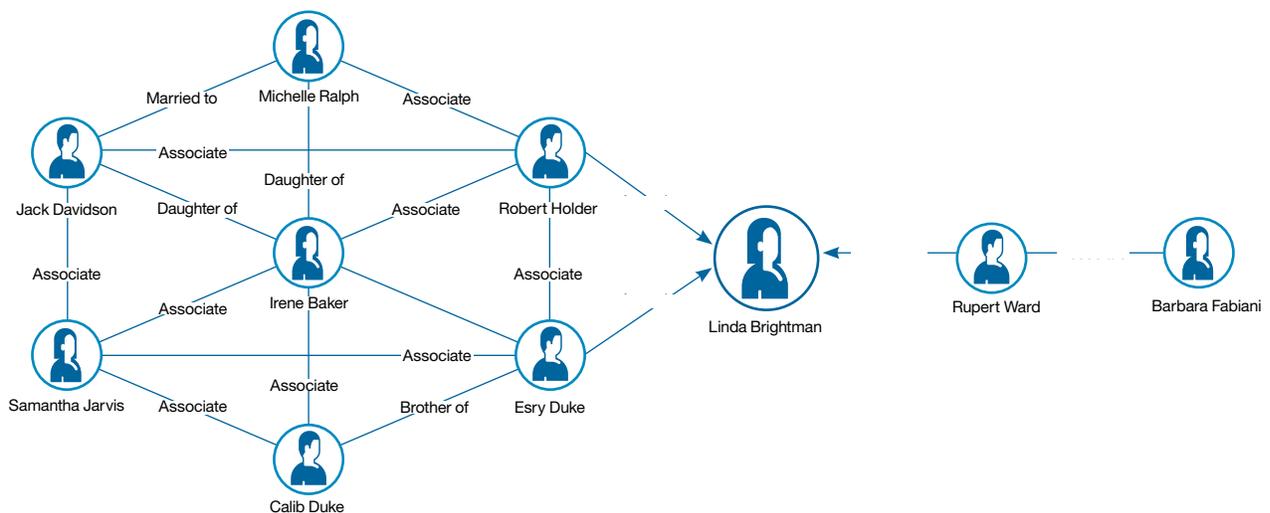


Figure 6: Link Direction—Linda BRIGHTMAN appears to be capable of receiving information from others, but is not passing it on to other parts of the network.

Link Weightings

SNA can also be enhanced by the use of weightings to indicate the strength of differing relationships (links), each of which has an effect on a target network. This helps deliver a more real-world indication of the dynamics and structure of a given target network. As mentioned earlier, not all relationships in a network are equal. For example, qualitatively, the link

between two people connected through a family relationship may be stronger than a link between two business associates. These links can be weighted so that they represent real-world strengths when carrying out SNA. Weighting key paths in the network may also infer that the entities using them to channel information have key roles to play. For this reason, centrality measure results are affected by link weightings.

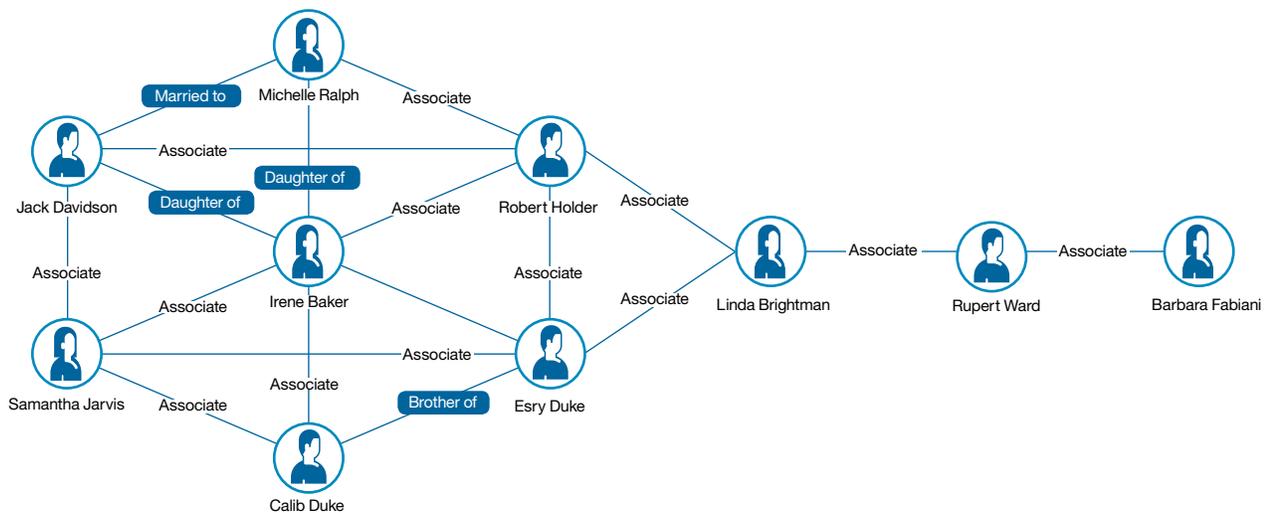


Figure 7: Link Weightings—Differing relationships can have weightings applied to them to allow SNA results to reflect the different strengths of real world relationship types.

Conditional Formatting

Once SNA calculations have been run against a chart, users may wish to emphasize certain aspects of that information within it. This is especially helpful when the chart is visually complex. The new conditional formatting capability in Analyst's Notebook can be used, for example, to illustrate key entities in a social network. Formatting rules can be created and then run against the data contained within the chart to highlight key entities and links by enlarging their size, thickening and coloring their links so they clearly stand out in complex charts.

The following example shows how the use of conditional formatting can help more clearly visualize who and where potential key individuals are placed in a network and how a network is structured. The larger the circle (Icon Frame) on an entity, the higher its' betweenness score is.

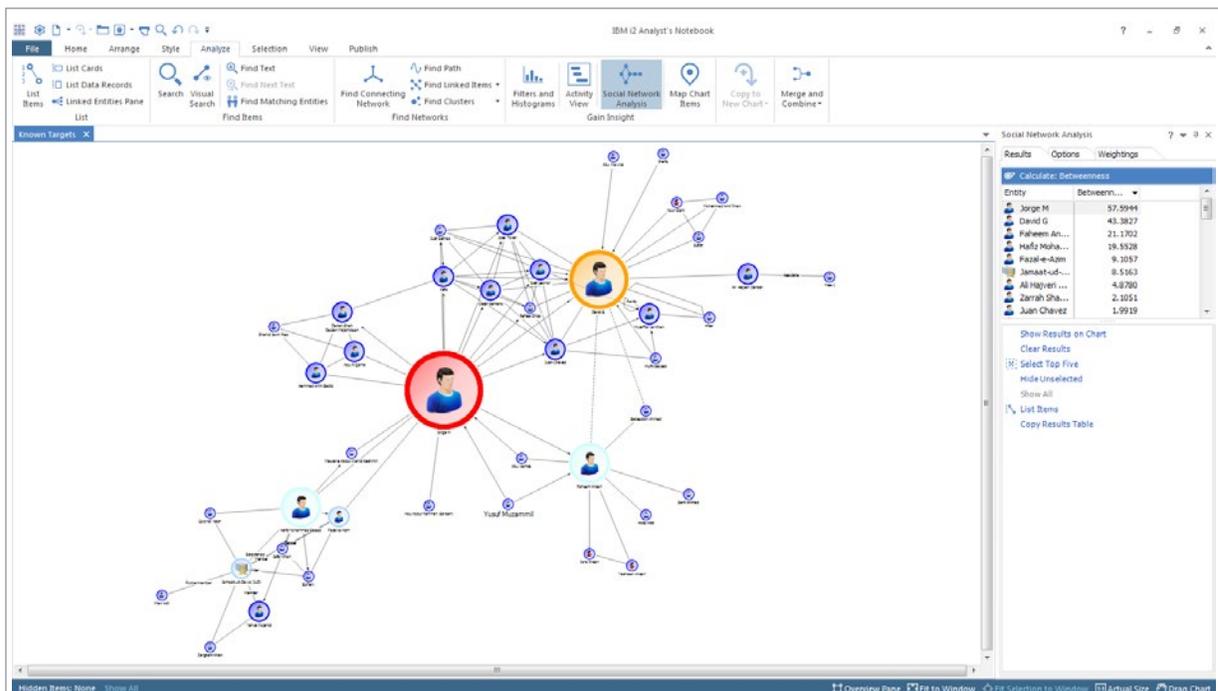


Figure 8: Conditional Formatting helps to clearly emphasize key entities in a network based, in this example, on the value of SNA results.

Where Social Network Analysis functions are located in Analyst's Notebook and how they can be used

To carry out SNA, you need to select the Social Network Analysis tab in the Task Pane.

The Options Page

To set the centrality measures you want to run, select the Options tab to display the Options Page. From here, you can select single or multiple centrality measures and choose to take into account directed links.

The Enhanced Analysis option allows users to control over how Analyst's Notebook deals with charts with several unconnected networks, or directed links that block paths between certain entities.

Normalize Results to adjust the calculations and display them as percentages in the Results table. If you turn off this option, no adjustment is made to the calculations and results are displayed as raw data. Normalizing the results makes it easier to compare the centrality results of items in different networks and charts.

Actions on Completion

Different actions can be run on completion of SNA calculations to aid in visualizing the results.

- They can also show results on the chart by choosing Show Results on Chart.
- Conditional Formatting can also be applied on completion, to emphasize information such as enlarging and coloring icons by betweenness. To apply this select the Apply Conditional Formatting option, then select a Conditional Formatting specification from the drop-down list.

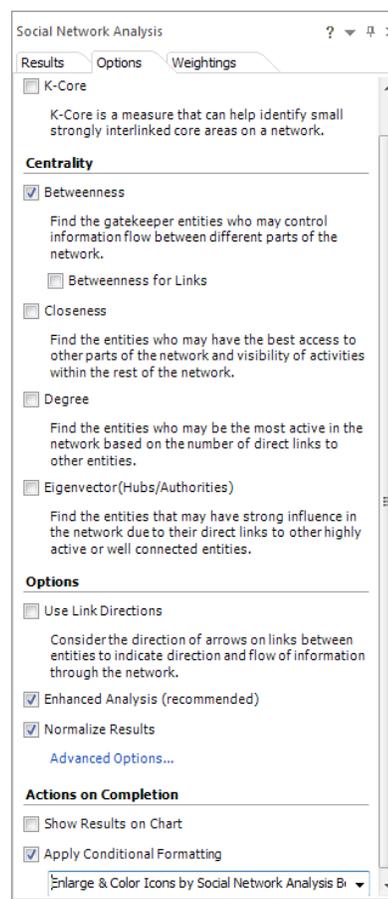


Figure 9: Users have control over what centrality measures to run.

The Weightings Page

SNA can be enhanced by the use of weightings to indicate the strength of differing relationships such as links in social networks. In Analyst's Notebook, weightings can be manually set on a chart-by-chart basis, or can be implemented by using pre-configured weightings file. Weightings can be used by selecting the Use Link Weightings option on the Options page, and then:

- Selecting entities and using the Manually Set Value option to enter a specific weighting value.
- Selecting an existing numeric Custom Attribute and using that attribute value as the weighting value.
- Selecting details From a Weightings File. Click the Create a New File option to create a new weightings file, or Select a File to import an existing one.
- Users can opt to display the weightings that they have used on the chart by selecting Show Weightings on chart.

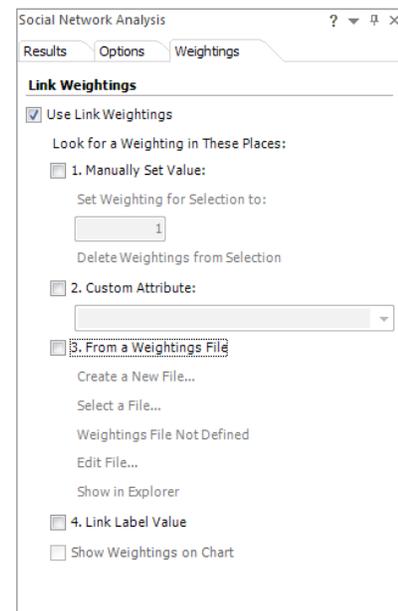
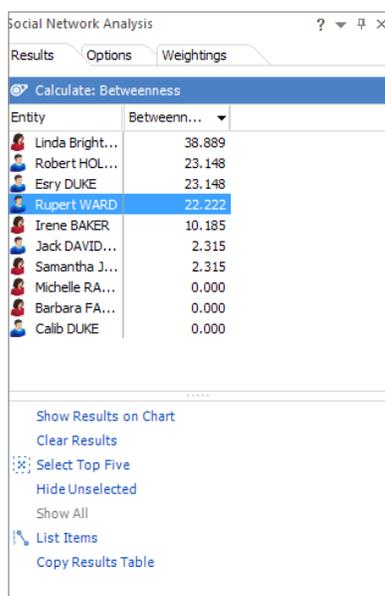


Figure 10: Weightings can be applied using a variety of methods.

Calculating Results

Start generating SNA results by clicking the Calculate button. Once results have been generated, they are automatically displayed in the Results table. Results for entities and links are displayed in different tables, with the entity Results table displayed by default. In the list view, the SNA results are then displayed. You can re-sort the data by clicking one of the column headers.



The screenshot shows the 'Social Network Analysis' application window with the 'Results' tab selected. A 'Calculate: Betweenness' dialog is open, displaying a table of results. Below the table is an actions panel with several options.

Entity	Betweenness...
Linda Bright...	38.889
Robert HOL...	23.148
Esry DUKE	23.148
Rupert WARD	22.222
Irene BAKER	10.185
Jack DAVID...	2.315
Samantha J...	2.315
Michelle RA...	0.000
Barbara FA...	0.000
Calib DUKE	0.000

Actions Panel:

- Show Results on Chart
- Clear Results
- Select Top Five
- Hide Unselected
- Show All
- List Items
- Copy Results Table

Figure 11: The results pane lists the SNA results.

The Results Page

To calculate and analyze your SNA results, select the Results tab to display the Results page.

Calculating Results

Start generating SNA results by clicking the Calculate button. Once results have been generated, they are automatically displayed in the Results table. Results for entities and links are displayed in different tables, with the entity Results table displayed by default. In the list view, the SNA results are then displayed. You can re-sort the data by clicking one of the column headers.

Actions Panel

After generating results, users have several options from which to choose from regarding what they wish to do with their results:

- They can show the results on their chart by selecting Show Results on Chart.
- It is also possible to select the top five results in the list so that they are highlighted on the chart. This is done by choosing Select Top Five, and opting to Hide Unselected items from this point.
- Finally, users can also copy the Results Table for inclusion into other applications such as Microsoft Word or Excel by selecting Copy Results Table.

Social Network Analysis glossary:

Entity

The entity icon followed by the label.

Betweenness

A value representing how many of the shortest paths pass through each entity. For example, an entity with the lowest betweenness value has the lowest number of shortest paths running through it compared with the other entities in the network. If Use Link Direction was selected in the Options page, the arrow direction determines which paths can be measured between entities.

There are no separate columns for In and Out betweenness results because the number of directed inbound paths would be identical to the number of directed outbound paths if the arrows were reversed. This is because the same connecting paths are used for the betweenness calculation regardless of direction.

Closeness

A value representing how close each entity is to the others in the network. For example, an entity with the highest closeness value has the shortest paths to the majority of other entities in the network. If Use Link Direction was selected in the Options page, the closeness result is displayed in two separate columns:

- Closeness (In): how close an entity is, based on the number of inbound paths
- Closeness (Out): how close an entity is, based on the number of outbound paths

Degree

The number of links each entity has to others in the network. If Use Link Direction was selected in the Options page, the degree result is displayed in two separate columns:

- Degree (In): the total number of inbound links
- Degree (Out): the total number of outbound links

An entity's degree result is affected if the links connecting it have been weighted. For example, a link with a weighting of four is classed as four times more important than a link without a weighting. It counts as four links, whereas a link without an assigned weighting counts as one by default. The degree for an entity connected by two links, where one link has no weighting and the other has a weighting of four, would be five.

Eigenvector

A value representing how well connected or influential an entity is, based on its direct links to the other active entities in the network. An entity with the highest eigenvector score is directly connected to many other entities with high centrality scores. If Use Link Direction was selected in the Options page, two eigenvector results are calculated and displayed in two separate columns:

- Authority (In): how well connected an entity is, based on its inbound links
- Hub (Out): how well connected an entity is, based on its outbound links

An entity's eigenvector centrality result is affected if the links connecting it to other entities have been weighted. For example, a link weighting of three means that the link is treated as three times more important than a link with no weighting.

Link

The link icon followed by a summary of the link in the format: Entity A label <arrow> link label <arrow> Entity B label.

Link Betweenness

A value representing how many of the shortest paths pass through each link. For example, a link with the highest betweenness value has the highest number of paths running through it compared with the other links in the network.

Weighting

The weighting value assigned to each link.

Selecting a result in the table also selects and zooms to the corresponding entity or link on the chart. Results can be selected and columns resized and sorted in the usual way. Results can also be displayed on chart items.

Technical description

Product architecture

Analyst's Notebook is a standalone desktop product designed to provide users with a powerful visualization and analytical tool.

Analyst's Notebook has an Application Programming Interface (API) to enable programmatic control of the application via the IBM i2 Analyst's Notebook SDK.

What documentation is provided?

The documentation supplied with Analyst's Notebook is as follows:

- Analyst's Notebook Release Notes
- Analyst's Notebook Online Help

The documentation (other than the on-line help) is provided in electronic form only. To display the documentation, a PDF viewer must be present on the installation system.

Is Analyst's Notebook available in languages other than English?

Analyst's Notebook is developed in US English. It is supported on Western Europe, US, Central European, Baltic, Cyrillic, Turkic, Arabic, Japanese, Korean, Simplified Chinese and Traditional Chinese regional versions of the supported operating systems.

For more details on the availability of national language versions, please contact your regional sales representative.

For more information

To learn more about IBM i2 Analyst's Notebook, contact your IBM sales representative, or visit: ibm.com/software/products/en/analysts-notebook



© Copyright IBM Corporation 2017

IBM Corporation
IBM Security
Route 100
Somers, NY 10589

Produced in the United States of America
March 2017

IBM, the IBM logo and ibm.com are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at www.ibm.com/legal/copytrade.shtml.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED “AS IS” WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

The client is responsible for ensuring compliance with laws and regulations applicable to it. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the client is in compliance with any law or regulation.



Please Recycle