

Matrix to Coordinates (M2C) 4.0 Documentation

Elbirt Technologies
© Copyright 2010

Table of Contents

TABLE OF CONTENTS	2
SOFTWARE LICENSE AGREEMENT	3
I PRODUCT SPECIFICATION	6
PRODUCT OVERVIEW	6
ABOUT THE AUTHOR / BUG REPORTING	6
MINIMUM HARDWARE/SOFTWARE REQUIREMENTS	6
SOFTWARE INSTALLATION	6
FUTURE ENHANCEMENTS	6
MEMORY EXPANSION AND CONSUMPTION	7
II HOW IT WORKS	9
STRUCTURAL EQUIVALENCE DISTANCE CONVERSION	9
RIEMANN VS. EUCLIDEAN GEOMETRY	9
COORDINATES (MULTI-DIMENSIONAL SCALING)	9
COMPARING SPACES & ROTATION	10
DETERMINING COORDINATE ERROR	12
III INPUT FILES	13
DATA HEADER	13
MULTIPLES	13
PAIRS	14
MATRIX	14
IV APPLICATION INTERFACE	15
PROJECT MENU	15
INPUT FILES MENU	16
CONFIGURATIONS MENU	17
PROGRESS BAR	18
V OUTPUT FILES	19
COORDINATE DATA FILE	19
CENTRALITIES FILE	20
COORDINATE ERROR FILE	21
SE (STRUCTURAL EQUIVALENCE) FILE	22
ET-V© (VISUALIZER) FILE	23
TIME SERIES DATA FILE	24
CENTRALITIES TIME SERIES DATA FILE	25
VI TIME SERIES ROTATION ALGORITHM - EVIDENCE OF VALIDITY	26
VII BIBLIOGRAPHY	29
APPENDIX A – MEMORY RECOMMENDATIONS	30
APPENDIX B – ERROR DEBUGGING	31
HOW TO VIEW THE ERROR OUTPUT TO THE SYSTEM COMMAND LINE	31
CATASTROPHIC FAILURE - APPLICATION AUTO TERMINATION	31
DATA OUTPUT SPEEDS	32

Software License Agreement

PLEASE READ THIS SOFTWARE LICENSE AGREEMENT CAREFULLY BEFORE INSTALLING OR USING THE SOFTWARE.

BY OPENING THE PACKAGE, DOWNLOADING THE PRODUCT, OR USING THE EQUIPMENT THAT CONTAINS THIS PRODUCT, YOU ARE CONSENTING TO BE BOUND BY THIS AGREEMENT. IF YOU DO NOT AGREE TO ALL OF THE TERMS OF THIS AGREEMENT, RETURN THE PRODUCT TO THE PLACE OF PURCHASE FOR A FULL REFUND, OR DO NOT DOWNLOAD THE PRODUCT.

Single User License Grant: Benjamin Elbirt ("Owner") grants to Customer ("Customer") a nonexclusive and nontransferable license to use **Elbirt Technologies Software Suite 1.0** ("Software") in object code form solely on a single central processing unit owned or leased by Customer.

The **Elbirt Technologies Software Suite 1.0** ("Software") consists of multiple applications. This license grants to Customer ("Customer") use of the applications **Visualizer (ET-V) 1.0**, **Matrix to Coordinates (ET-M2C) 4.0**, and **Unicode Semantic Network Analysis Application (USNAA) 1.0**.

The Software is provided for 32 and 64 bit Windows™ Operating Systems.

Multiple-Users License Grant: Benjamin Elbirt ("Owner") grants to Customer ("Customer") a nonexclusive and nontransferable license to use **Elbirt Technologies Software Suite 1.0** ("Software") in object code form: (i) installed in a single location on a hard disk or other storage device of up to the number of computers owned or leased by Customer for which Customer has paid a license fee ("Permitted Number of Computers"); or (ii) provided the Software is configured for network use, installed on a single file server for use on a single local area network for either (but not both) of the following purposes: (a) permanent installation onto a hard disk or other storage device of up to the Permitted Number of Computers; or (b) use of the Software over such network, provided the number of computers connected to the server does not exceed the Permitted Number of Computers. Customer may only use the programs contained in the Software (i) for which Customer has paid a license fee (or in the case of an evaluation copy, those programs Customer is authorized to evaluate. Customer grants to Owner or its independent accountants the right to examine its books, records and accounts during Customer's normal business hours to verify compliance with the above provisions. In the event such audit discloses that the Permitted Number of Computers is exceeded, Customer shall promptly pay to Owner the appropriate licensee fee for the additional computers or users. At Owner option, Owner may terminate this license for failure to pay the required license fee.

Customer may make one (1) archival copy of the Software provided Customer affixes to such copy all copyright, confidentiality, and proprietary notices that appear on the original.

EXCEPT AS EXPRESSLY AUTHORIZED ABOVE, CUSTOMER SHALL NOT: COPY, IN WHOLE OR IN PART, SOFTWARE OR DOCUMENTATION; MODIFY THE SOFTWARE; REVERSE COMPILE OR REVERSE ASSEMBLE ALL OR ANY PORTION OF THE SOFTWARE; OR RENT, LEASE, DISTRIBUTE, SELL, OR CREATE DERIVATIVE WORKS OF THE SOFTWARE.

Customer agrees that aspects of the licensed materials, including the specific design and structure of individual programs, constitute trade secrets and/or copyrighted material of Owner. Customer agrees not to disclose, provide, or otherwise make available such trade secrets or copyrighted material in any form to any third party without the prior written consent of Owner. Customer agrees to implement reasonable security measures to protect such trade secrets and copyrighted material. Title to Software and documentation shall remain solely with Owner.

LIMITED WARRANTY. Owner warrants that for a period of ninety (90) days from the date of shipment from Owner: (i) the media on which the Software is furnished will be free of defects in materials and workmanship under normal use; and (ii) the Software substantially conforms to its published specifications. Except for the foregoing, the Software is provided AS IS. This limited warranty extends only to Customer as the original licensee. Customer's exclusive remedy and the entire liability of Owner under this limited warranty will be, at Owner option, repair, replacement, or refund of the Software if reported (or, upon request, returned) to the party supplying the Software to Customer. In no event does Owner warrant that the Software is error free or that Customer will be able to operate the Software without problems or interruptions.

This warranty does not apply if the software (a) has been altered, except by Owner, (b) has not been installed, operated, repaired, or maintained in accordance with instructions supplied by Owner, (c) has been subjected to abnormal physical or electrical stress, misuse, negligence, or accident, or (d) is used in ultra-hazardous activities.

DISCLAIMER. EXCEPT AS SPECIFIED IN THIS WARRANTY, ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS, AND WARRANTIES INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, NONINFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE, ARE HEREBY EXCLUDED TO THE EXTENT ALLOWED BY APPLICABLE LAW.

IN NO EVENT WILL Owner BE LIABLE FOR ANY LOST REVENUE, PROFIT, OR DATA, OR FOR SPECIAL, INDIRECT, CONSEQUENTIAL, INCIDENTAL, OR PUNITIVE DAMAGES HOWEVER CAUSED AND REGARDLESS OF THE THEORY OF LIABILITY ARISING OUT OF

THE USE OF OR INABILITY TO USE THE SOFTWARE EVEN IF Owner HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. In no event shall Owner liability to Customer, whether in contract, tort (including negligence), or otherwise, exceed the price paid by Customer. The foregoing limitations shall apply even if the above-stated warranty fails of its essential purpose. SOME STATES DO NOT ALLOW LIMITATION OR EXCLUSION OF LIABILITY FOR CONSEQUENTIAL OR INCIDENTAL DAMAGES.

The above warranty DOES NOT apply to any beta software, any software made available for testing or demonstration purposes, any temporary software modules or any software for which Owner does not receive a license fee. All such software products are provided AS IS without any warranty whatsoever.

This License is effective until terminated. Customer may terminate this License at any time by destroying all copies of Software including any documentation. This License will terminate immediately without notice from Owner if Customer fails to comply with any provision of this License. Upon termination, Customer must destroy all copies of Software.

Software, including technical data, is subject to U.S. export control laws, including the U.S. Export Administration Act and its associated regulations, and may be subject to export or import regulations in other countries. Customer agrees to comply strictly with all such regulations and acknowledges that it has the responsibility to obtain licenses to export, re-export, or import Software.

This License shall be governed by and construed in accordance with the laws of the State of New York, United States of America, as if performed wholly within the state and without giving effect to the principles of conflict of law. If any portion hereof is found to be void or unenforceable, the remaining provisions of this License shall remain in full force and effect. This License constitutes the entire License between the parties with respect to the use of the Software.

I Product Specification

Product Overview

M2C 4.0© is a Java application that converts matrix and pair type data inputs into coordinate systems using Structural Equivalence distance conversion, Multidimensional Scaling (*MDSJ*) (Algorithmics Group, 2009; Brandes & Pich, 2007; Woelfel & Fink, 1980) coordinate generation and a custom Orthogonal Procrustes (Schonemann, 1966) rotation algorithm for longitudinal alignment.

Outputs are in *Microsoft Excel XLSX™* files and include a coordinate file, time series file, **Elbirt Technologies Visualizer (ET-V)©** data file, structural equivalence pair data file and coordinate to structural-equivalence error values.

About the Author / Bug Reporting

M2C 4.0© is a product of Benjamin Elbirt / Elbirt Technologies. Use the following e-mail address to report bugs, request changes or ask questions. Be sure to check [Appendix B – Error Debugging](#) for more information on common errors and how to collect error information.

Website : <http://www.elbirttechnologies.com>

E-mail: elbirt@elbirttechnologies.com

Minimum Hardware/Software Requirements

Microsoft Windows 32 (XP or later) or 64 bit™
Latest version of the Java Runtime Environment
1.0 GB Memory

Software Installation

M2C 4.0© requires the installation of the latest Java environment which is available at <http://www.java.com/>.

The software is installed as part of the **Elbirt Technologies Software Suite©** and can be found at <http://www.elbirttechnologies.com>.

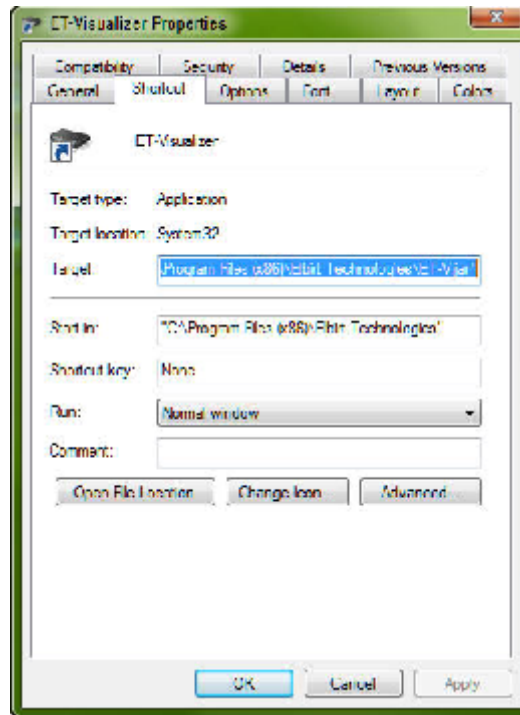
Future Enhancements

1. Additional Layout Coordinate Methods.
2. Additional Centrality and Network Analysis measures.
3. Additional Time-Series Analysis statistics.
4. Enhanced ET-V output files.

Memory Expansion and Consumption

M2C 4.0© uses the memory specified in the default java runtime environment if no options are specified. The installation provides Windows short-cut link files with the proper java execution variables for expanding the system memory. These files can be found in the installation directory (\$INSTALL_PATH) and on your Desktop if the desktop installation option was selected.

- **32 bit Windows** – C:\Program Files\Elbirt Technologies\
- **64 bit Windows** – C:\Program Files (x86)\Elbirt Technologies\



The short-cut properties can be accessed by right clicking on the short-cut and selecting "Properties" from the menu. The properties window will open and provide multiple tabs of information. Select the "Shortcut" tab; which contains the following *Target*:

C:\Windows\System32\java.exe -Djava.library.path="\$INSTALL_PATH\lib" -Xms1000M -Xmx1500M -jar "\$INSTALL_PATH\M2C.jar"

The **–Xms1000M** specifies a minimum virtual memory setting of 1,000 MB or 1GB. The **–Xmx1500M** specifies a maximum virtual memory setting of 1,500 MB or 1.5GB. Please see [Appendix A – Memory Recommendations](#) for more information on what these values should be based on your system memory size.

You should EDIT the short-cut “Target” value to contain the proper Xms and Xmx values for your system. The **Xmx** value will need to be larger if you have out-of-memory errors appear in the application execution window.

II How it Works

M2C 4.0© is a Java application that converts matrix and pair type data inputs into coordinate systems using Structural Equivalence distance conversion, Multidimensional Scaling (MDS) coordinate generation (Brandes & Pich, 2007; Woelfel & Fink, 1980) and a custom Orthogonal Procrustes (Schonemann, 1966) rotation algorithm for longitudinal alignment.

Structural Equivalence Distance Conversion

M2C 4.0© converts the input matrix data into a distance matrix using Structural Equivalence (SE). The resulting distance matrix is necessary to satisfy MDS generation requirements – symmetrical and all values must be greater than or equal to zero.

This method compares the object rows/columns to each other to determine how different they are. The difference is calculated using the Pythagorean Theorem such that the distance between two nodes is the square root of the sum of all differences squared. This process will work on both binary and non-binary input matrix values preserving the non-binary relationships.

The final matrix is a symmetrical matrix of distances representing the distance between objects based on the differences between their rows and columns in the original matrix. The smaller the value the closer two objects are to each other structurally; larger values are less structurally similar and thus more distant.

$$diff(i, j) = diff(j, i) = \sqrt{\sum_{k=1}^n \left(((i, k) - (j, k))^2 + ((k, i) - (k, j))^2 \right)}$$

Riemann vs. Euclidean Geometry

MDS algorithms exist for creating Euclidean geometries; however not all networks (especially cognitive networks) neatly fit within a standard Euclidean geometry (Woelfel & Fink, 1980). Riemann geometry is the accepted alternative for representing non-Euclidean geometries (Woelfel & Fink, 1980; Brandes & Pich, 2007).

Coordinates (Multi-Dimensional Scaling)

MDS (Multi-dimensional scaling) is the methodology used to convert distance matrixes to coordinate systems and has been used by many software packages across multiple disciplines. This software uses the Algorithmics Group (2009) *MDSJ* java code exclusively based on the following assessments.

A comparison of the Woelfel & Fink method with *MDSJ* yielded nearly identical results with the differences being less than 0.0001. This difference can be attributed to the use of Fortran (Woelfel & Fink) with four decimal precision vs. Java (*MDSJ*) that has up to 11 decimal point precision.

A major difference between the two methods is the ability of *MDSJ* to produce only the desired dimensionality where as the Woelfel & Fink method requires the coordinate system generated to be the same size as the matrix inputs. This significantly reduces processing time for solutions when a subset of the total dimensionality is needed.

Another important difference between the two software packages is the unlimited capabilities of *MDSJ* (limited by memory/cpu available on the computer) vs. the Woelfel & Fink software that is limited to 999 concepts. *MDSJ* allows for the expansion of the software capabilities without code modifications.

Finally, *MDSJ* are integrated into other java code where as the Woelfel & Fink software is an entire package that cannot be customized or integrated with other code. Although the Woelfel & Fink software (Galileo/CatPac) provides additional functionality/computations that the *MDSJ* does not; these calculations can be added to future versions of *MDSJ* where as the Woelfel & Fink software has not seen a significant update since the early 1990's.

Comparing Spaces & Rotation

M2C 4.0© provides two uses of the Orthogonal Procrustes (Schonemann, 1966) rotation algorithm; time series rotation and optimal choice rotation. Further, the option to scale during rotation is provided. Finally, the **M2C 4.0©** algorithm allows for the rotation of coordinate systems that contain unique objects so long as at least two (2) objects are shared between each system.

Why Rotate?

MDSJ and MDS algorithms in general find a best fit solution of coordinates for each unique data matrix provided. As a result the coordinate system generated does not consider the previously generated data sets and can be oriented differently. Such orientation differences can include scale differences, dimensional rotations and axis inversion.

Another way to consider the problem is that any coordinate system space can be thought of as a solid, clear plastic ball (the space), with specs of black on the surface and throughout the inside of the ball (objects). You can spin the ball, flip it over, etc. and although you are looking at the same ball your orientation provides a different point of view.

The variations in orientation can suggest a larger or smaller amount of change from one coordinate system to another when considering multiple data matrixes as either a time series or a set of data for comparison. Imagine two identical balls, however one is flipped over and rotated 30 degrees along one axis. Although the balls are exactly the same, the orientation suggests that they are not and the amount of difference will be a function of the orientation of the second ball.

Orthogonal Procrustes rotation finds the best rotation of a given coordinate system that map onto another coordinate system being rotated to. The final map is a set of coordinates that have the least difference to the original system being rotated to. Thus, the rotation will re-orient the second ball to the first and will do so without warping or distorting the objects.

This is very important when comparing two coordinate systems or looking at a time series of coordinates. Orthogonal Procrustes rotation removes the orientation problem and its affect on the resulting coordinate systems.

Scaling Rotations

Many rotation methods scale the coordinate system to be rotated as part of the rotation. Thus, more compact coordinate systems will be expanded to fit larger systems rotated to and larger systems will be compressed to smaller systems rotated to.

In most time series data cases the scaling will reduce the amount of change between points in time even though the change is a function of the input data. Turning the Scale Rotation option OFF will prevent the rotation from applying scale and maintains the compression / expansion of the data.

Unique Object Rotation

Most rotation algorithms require both data sets to contain the same objects. In some time series cases objects appear and disappear from the data over time. The **M2C 4.0©** rotation method accommodates for missing objects by rotating the coordinates based on shared objects only. The rotations affect the entire coordinate system to be rotated; however they are determined by the shared coordinates among the two systems.

Time Series Rotation

Time series rotation considers the coordinate system rotation serially. Coordinate system 2 will be rotated to 1, 3 to 2, 4 to 3 and so on.

Optimal Choice Rotation

Optimal choice rotation determines a list and order of rotation based on shared objects. Rotation of a given coordinate system will be done based on the

coordinate system it shares the most objects to rather than the previous coordinate system in the data set (which may contain fewer shared objects from other systems).

Determining Coordinate Error

M2C 4.0© provides the option for generating distance errors. These errors compare the distances between objects as a function of the coordinate system generated to the distances determined with structural equivalence.

The distance error of two objects is the absolute difference of the coordinate distance to the structural equivalence distance. Thus, if the coordinates put two objects 4.0 distance from each other and the structural equivalence found a distance of 3.8 there will be a 0.2 distance error.

The error rate is important because the best solution is always the highest coordinate dimension solution; i.e. where all coordinate dimensions are generated. However, this can be computationally expensive both for memory and time when dealing with many objects (25+).

The more objects in the system the more dimensionality is needed to compute a full solution. Using less coordinate dimensions results in higher error because more of the missing distances are found in the dimensions not used; which grows as the object volume grows.

Additionally, only a few dimensions are needed for visualization purposes; up to 3 for coordinates and possibly more for colors and other visualization attributes. The **Elbirt Technologies Visualizer (ET-V)©** can use 3 for coordinates, 3 for colors, 1 for size, 1 for spin, and up to 5 for intonation (music) for a total of 13¹.

The error outputs provide a symmetrical distance matrix of error distances that can be used in other applications. One example is to take the mean error of a given object as related to all other objects as the size of the object in the **ET-V©** with a Sphere shape. The object size will then represent the possible locations for the object given the average error related to all other objects; larger objects have more error.

¹ Colors are normalized to values between 0.0 and 1.0, size must be a positive non-zero value, spin must be a positive value greater than or equal to zero and the musical values are all integers normalized to a value range based on the attribute (Note, Velocity, Channel, Instrument, Beats).

III Input Files

M2C 4.0 uses Text delimited and *Microsoft Excel™* files as input. Text delimiters can be Tab or Comma only. Both XLS and XLSX versions of *Excel™* are valid. Each spreadsheet is considered a unique data set and can be in one of two formats; pairs or matrix.

Data Header

The first line of each *Excel™* spreadsheet should contain in the first cell (A1) the word **Matrix** or **Pairs** (case insensitive). This is used by the application to determine the input format expectations and validation methods.

Text files should contain the **Matrix/Pairs** value on the first line without any other textual content including delimiters.

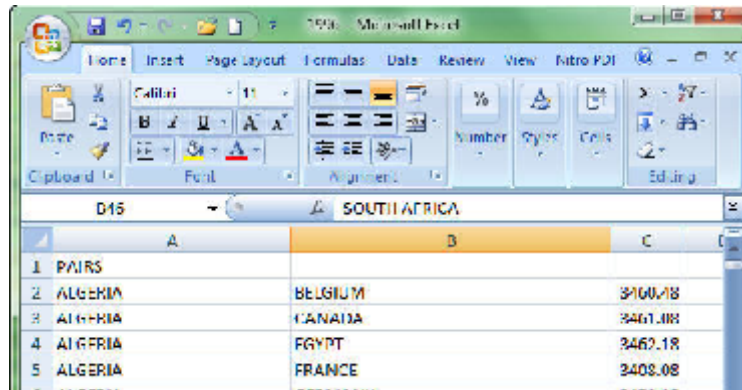
Multiples

Each file provided will be processed based on the [configurations](#) specified. Excel files containing multiple spreadsheets will be processed as though each spreadsheet is a separate file of data allowing the mix of *pairs* and *matrix* input types.

Text files must all use the same delimiter specified in [configurations](#), however they can mix between *pairs* and *matrix* input types.

The processing will occur in the order shown in the Input File List. Files with multiple spreadsheets will be processed completely prior to moving on to the next input file. Rotations will occur in the order the files appear with the order of the spreadsheets taken into account when processing files with multiple data sheets.

Pairs

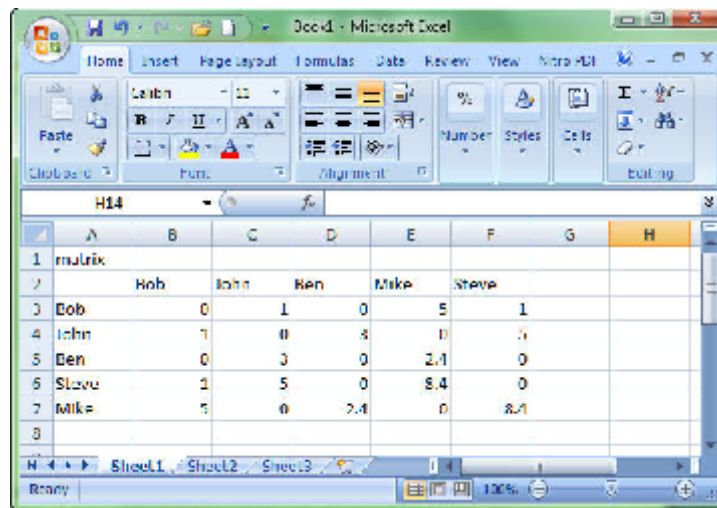


	A	B	C
1	PAIRS		
2	ALGERIA	BELGIUM	3460.18
3	ALGERIA	CANADA	3461.08
4	ALGERIA	EGYPT	3462.18
5	ALGERIA	FRANCE	3408.08

Pair input expects the first column to contain the matrix row label, the second column to contain the matrix column label and the third column to contain the positive matrix value. Zero values are ignored and a missing value is treated as binary (1.0).

Text files should contain the same format using the desired delimiter between each column of values.

Matrix

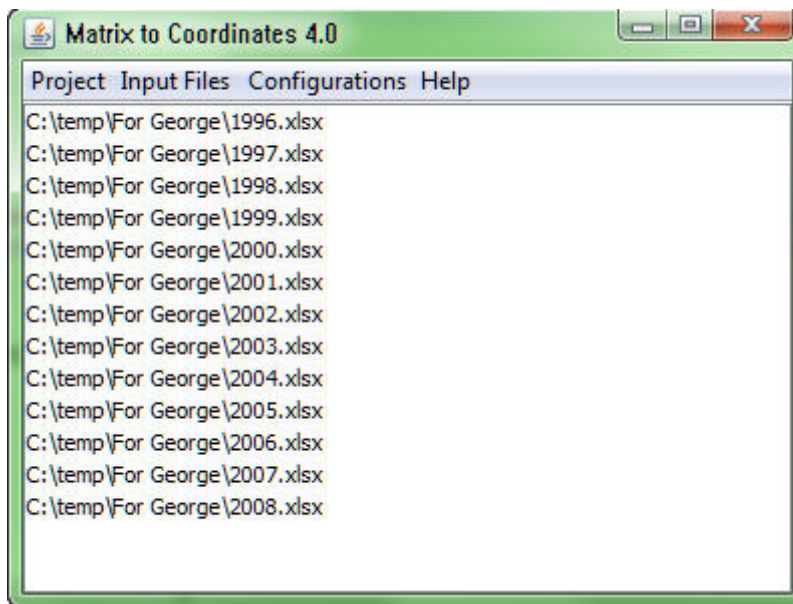


	A	B	C	D	E	F	G	H
1	matrix							
2		Bob	John	Ben	Mike	Steve		
3	Bob	0	1	0	5	1		
4	John	1	0	3	0	0		
5	Ben	0	3	0	2.4	0		
6	Steve	1	5	0	3.4	0		
7	Mike	5	0	2.4	0	3.4		

Matrix input expects the next row to contain the column labels with the first column having no value. Each row that follows should contain the row label in the first column and the matrix value in the corresponding row/column.

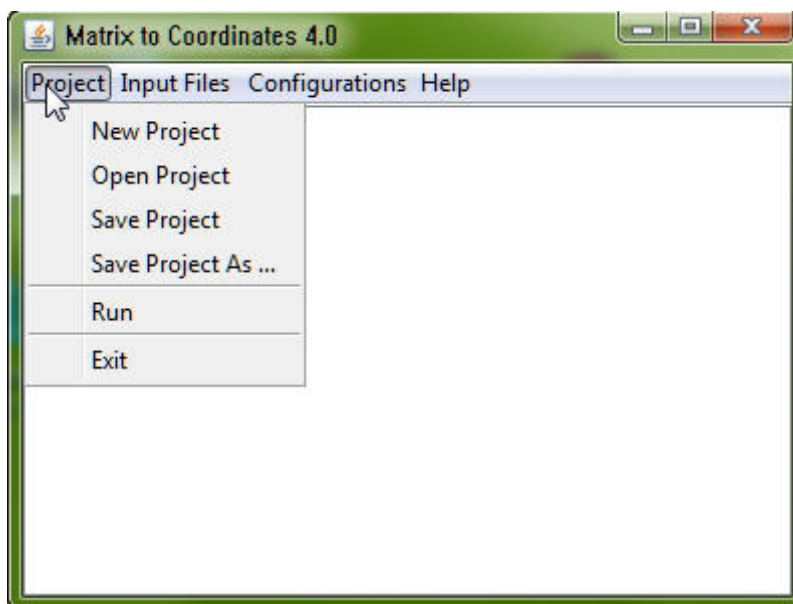
Text files should contain the same format using the desired delimiter between each column of values. Each row should be on a unique line in the file separated by the EOL (end of line) delimiter.

IV Application Interface



The main **M2C 4.0©** interface provides four menus and the input file list window as shown above. Multiple files can be selected using the mouse and control/shift keys on the keyboard. The control key allows for multiple individual selections and the shift key allows for interval selection of multiple files.

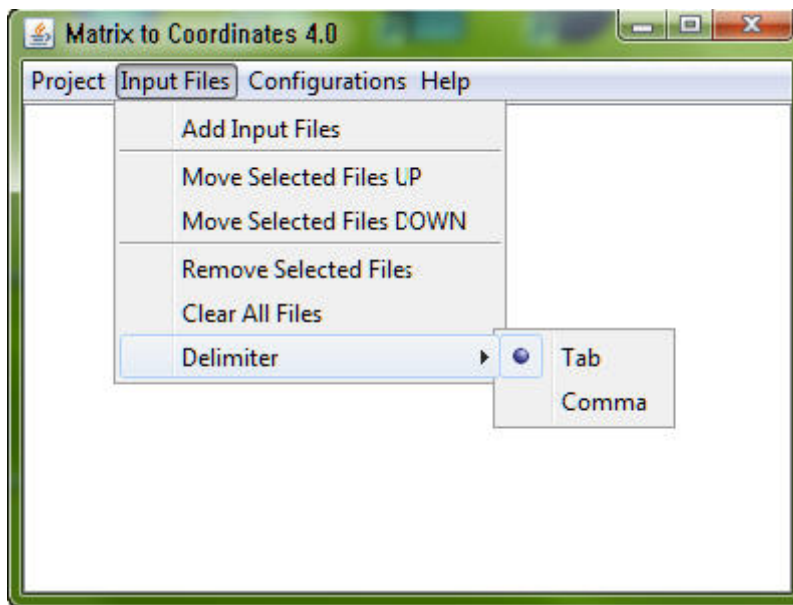
Project Menu



The *Project Menu* provides access to options for saving, opening and clearing the existing project. The *Run* option will execute the Coordinate Generation process on the files provided and the contents there in. *Exit* will terminate the application.

A dialogue window will appear asking for the directory to put the output files after selecting the *Run* option. A directory named *M2COutput* will be created within the directory specified to hold all the files created by the application. A numerical index will be added if this directory already exists; i.e. *M2COutput1*.

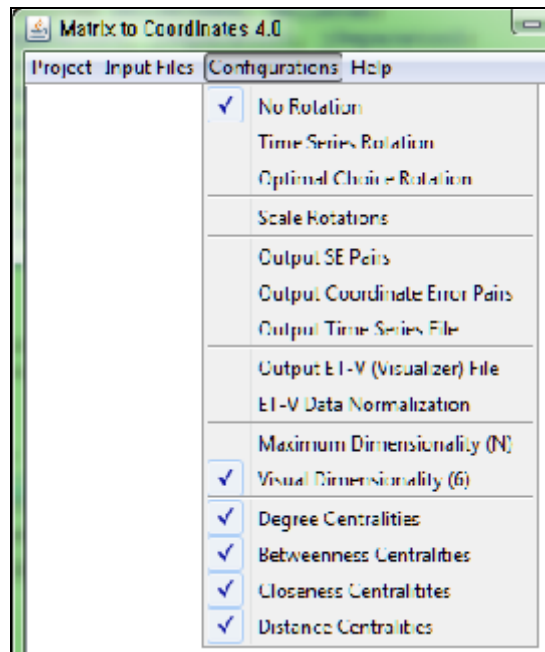
Input Files Menu



This menu provides the ability to *Add Input Files*, *Remove Selected Files*, *Clear All Files* and the ability to re-order the files via *Move Selected Files UP* and *Move Selected Files DOWN*.

The *Delimiter* option provides a sub menu where the delimiter type, *Tab* or *Comma* can be selected. Only one delimiter type can be chosen from the list.

Configurations Menu



The *Configurations Menu* provides controls for the Centrality Calculation, MDS generation, Procrustes Rotation and Output functionality. Rotation options include *No Rotation*, *Time Series Rotation*, *Optimal Choice Rotation* and *Scale Rotations*. Dimensionality can be either *Maximum Dimensionality* or *Visual Dimensionality* (6 dimensions).

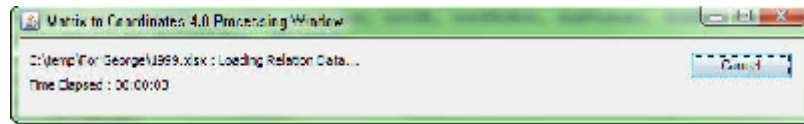
Optional outputs include *SE Pairs* (structural equivalences), *Coordinate Error Pairs*, **ET-V©** File, and a *Time Series File*. The *Coordinate Data File* is the default output file for all runs. Please see the section on [M2C 4.0 Output Files](#) for more information regarding the contents of each of these files.

An additional option, *ET-V Data Normalization*, allows for the normalization of coordinate data when creating the ET-V file. Normalizations are automatically to a world size of +/- 500.0; colors are normalized to 1.0. This is highly recommended so the final outputs are better suited for **ET-V©**. Do not use this option to obtain the raw values.

All normalizations are done across the entire data set provided to avoid improper scaling for time series and comparison data sets.

Finally, options are provided to use JUNG (Madadhain, et. al., 2005) to calculate Distance, Closeness and Betweenness centrality measures. Degree is calculated internally by the application.

Progress Bar



The progress bar will show while the application is running the MDS, Rotation and Output processes with a message of the current activity and the Time Elapsed since the process started. A *Cancel* button is provided to terminate the process in the current state.

Please be aware that cancelling during file outputs will result in a delay prior to the cancellation of the process and may result in output file generation prior to cancellation of process.

V Output Files

Outputs are in *Microsoft Excel XLSX™* files and include a coordinate file, centralities file, time series file, centrality time series file, **Elbirt Technologies Visualizer (ET-V)©** data file, structural equivalence pair data file and coordinate to structural-equivalence error values.

A dialogue window will appear asking for the directory to put the output files after selecting the *Run* option. A directory named *M2COutput* will be created within the directory specified to hold all the files created by the application. A numerical index will be added if this directory already exists; i.e. *M2COutput1*.

Each file will contain a spreadsheet with the relevant information for each data set found among all the files. The spreadsheets are named based on the naming used in the input files or the input file name itself when using text files.

Coordinate Data File

1	Label	Crd 1	Crd 2	Crd 3	Crd 4	Crd 5
2	ALBANIA	15181.73	-10637	-9351.84	-4638.75	-43
3	ALGERIA	12481.59	-12481.6	1137.33	-8203.42	717
4	ANDORRA	13921.03	-12594.1	-943.417	-6612.29	-43
5	ANGOLA	13861.57	-11619	1327.694	-5533.87	-26
6	ANTIGUA	15281.84	-10256.4	-9119.58	-4641.9	-43
7	ARGENTIN	9967.716	-5788.66	5663.443	5409.308	-49
8	ARMENIA	14292.7	-14292.7	-12461.9	-5893.96	-95

File Name : \$OUT_DIRECTORY\M2COutput\Coordinates.xlsx

The coordinate data file provides each coordinate set (magenta) on a different spreadsheet for each input data set (yellow highlight). The first row contains the

Centralities File

K29																	
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Label	Binary Cx	Binary Cy	Strength	Strength	Normalized	Normalized	Normalized	Normalized	Binary Sx	Weighted	Binary Ay	Weighted	Binary Bx	Weighted	Binary Cx	Weighted
2	ADAME DA	1	0	1	0	2.1E 00	0	1.0E 00	0	1	1	2	2	0	0	1	1
3	AETHER M	1	0	1	0	2.1E 00	0	1.0E 00	0	1	1	2	2	0	0	1	1
4	ALBERT FF	2	0	3	0	4.2E 00	0	3.0E 00	0	0.071429	0.002032	0.5	0.068421	0	0	0.428571	0.117649
5	ALBRECHT	1	1	1	1	2.1E 00	2.1E 00	1.0E 00	1.0E 00	NUM!	NUM!	NUM!	NUM!	0	0	NUM!	NUM!
6	ALEXANDI	2	1	2	1	4.2E 00	2.1E 00	2.1E 00	1.0E 00	1	1	2	2	0	0	1	1
7	ANGELUS	1	0	1	0	2.1E 00	0	1.0E 00	0	1	1	2	2	0	0	1	1
8	ANIKWA B	1	0	1	0	2.1E 00	0	1.0E 00	0	1	1	2	2	0	0	1	1
9	ANIKWA B	1	0	1	0	2.1E 00	0	1.0E 00	0	1	1	2	2	0	0	1	1
10	ANIKWA B	3	0	3	0	6.3E 00	0	3.0E 00	0	0.000000	0.000000	0.000000	0.000000	0	0	0.000000	0.000000
11	ANIKWA B	5	0	5	0	10.5E 00	0	5.0E 00	0	0.00	0.00	1.00	1.00	1	1	1	1
12	ANIKWA B	5	0	5	0	10.5E 00	0	5.0E 00	0	0.000000	0.000000	0.000000	0.000000	0	0	0.000000	0.000000
13	ANIKWA B	1	0	1	0	2.1E 00	0	1.0E 00	0	1	1	2	2	0	0	1	1
14	ANIKWA B	1	0	1	0	2.1E 00	0	1.0E 00	0	1	1	2	2	0	0	1	1
15	ANIKWA B	1	1	1	1	2.1E 00	2.1E 00	1.0E 00	1.0E 00	ANUM!	ANUM!	ANUM!	ANUM!	0	0	ANUM!	ANUM!
16	BLIPLINS	1	1	1	1	2.1E 00	2.1E 00	1.0E 00	1.0E 00	ANUM!	ANUM!	ANUM!	ANUM!	0	0	ANUM!	ANUM!
17	BEN ARIA	1	1	1	1	2.1E 00	2.1E 00	1.0E 00	1.0E 00	NUM!	NUM!	NUM!	NUM!	0	0	NUM!	NUM!
18	BIAN YAN	1	1	1	1	2.1E 00	2.1E 00	1.0E 00	1.0E 00	NUM!	NUM!	NUM!	NUM!	0	0	NUM!	NUM!
19	BIEN WAL	1	1	1	1	2.1E 00	2.1E 00	1.0E 00	1.0E 00	NUM!	NUM!	NUM!	NUM!	0	0	NUM!	NUM!
20	BIENST	0	1	0	1	2.1E 00	0	1.0E 00	0	0.000000	0.000000	0.000000	0.000000	0	0	0.0	0.0
21	BITENC KA	4	0	4	0	8.42E 00	0	4.2E 00	0	0.002778	0.002778	0.416667	0.416667	0	0	0.388889	0.388889
22	BLUMBER	1	1	1	1	2.1E 00	2.1E 00	1.0E 00	1.0E 00	NUM!	NUM!	NUM!	NUM!	0	0	NUM!	NUM!
23	BOZU DAK	1	0	1	0	2.1E 00	0	1.0E 00	0	1	1	2	2	0	0	1	1
24	BOZU DAK	1	0	1	0	2.1E 00	0	1.0E 00	0	1	1	2	2	0	0	1	1
25	BOZU DAK	1	0	1	0	2.1E 00	0	1.0E 00	0	0.0	0.0	0.0	0.0	0	0	0.0	0.0
26	BOZU DAK	1	1	1	1	2.1E 00	2.1E 00	1.0E 00	1.0E 00	ANUM!	ANUM!	ANUM!	ANUM!	0	0	ANUM!	ANUM!
27	BRANDIS	1	1	1	1	2.1E 00	2.1E 00	1.0E 00	1.0E 00	ANUM!	ANUM!	ANUM!	ANUM!	0	0	ANUM!	ANUM!
28	BRANDIS	2	0	2	0	4.2E 00	0	2.0E 00	0	0.0	0.0	1.0	1.0	0	0	1	1
29	BURNS KA	0	1	0	1	0	2.1E 00	0	1.0E 00	1	1	2	2	0	0	1	1
30	CANTER D	0	1	0	1	0	2.1E 00	0	1.0E 00	1	1	2	2	0	0	1	1
31	CAPOBAR	1	0	1	0	2.1E 00	0	1.0E 00	0	0.008824	0.000000	0.417000	0.388889	0	0	0.302841	0.388889
32	CARNABU	1	1	1	1	2.1E 00	2.1E 00	1.0E 00	1.0E 00	NUM!	NUM!	NUM!	NUM!	0	0	NUM!	NUM!
33	CASCIARD	1	1	1	1	2.1E 00	2.1E 00	1.0E 00	1.0E 00	NUM!	NUM!	NUM!	NUM!	0	0	NUM!	NUM!
34	CASTRO J	2	0	2	0	4.2E 00	0	2.0E 00	0	0.000000	0.000000	0.388889	0.388889	0	0	0.000000	0.000000
35	CHANY	4	0	4	0	8.42E 00	0	4.2E 00	0	0.000000	0.000000	1	1	0	0	0.388889	0.388889
36	CHEN RUI	5	0	5	0	10.5E 00	0	5.0E 00	0	0.00	0.00	1.00	1.00	0	0	1	1
37	CHEN RUI	2	1	2	1	4.2E 00	2.1E 00	2.1E 00	1.0E 00	0.0	0.0	1	1	0	0	0.0	0.0
38	CHEN RUI	2	0	2	0	4.2E 00	0	2.0E 00	0	0.0	0.0	1.0	1.0	0	0	1	1
39	CHEN RUI	1	1	1	1	2.1E 00	2.1E 00	1.0E 00	1.0E 00	ANUM!	ANUM!	ANUM!	ANUM!	0	0	ANUM!	ANUM!
40	CHEN RUI	1	1	1	1	2.1E 00	2.1E 00	1.0E 00	1.0E 00	ANUM!	ANUM!	ANUM!	ANUM!	0	0	ANUM!	ANUM!

A unique centralities file is generated for each data set containing the centrality scores selected from the [Configurations Menu](#). The first row (green) contains the header labels. The data (magenta) is stored beneath and input files with multiple spreadsheets will contain the centrality scores for each spreadsheet in a separate spreadsheet (yellow) within the output file.

Coordinate Error File

	NAME	DATE	TIME	STATUS	REMARKS
1	ANITA	2023-01-01	08:00	Present	
2	ANITA	2023-01-02	08:00	Present	
3	ANITA	2023-01-03	08:00	Present	
4	ANITA	2023-01-04	08:00	Present	
5	ANITA	2023-01-05	08:00	Present	
6	ANITA	2023-01-06	08:00	Present	
7	ANITA	2023-01-07	08:00	Present	
8	ANITA	2023-01-08	08:00	Present	
9	ANITA	2023-01-09	08:00	Present	
10	ANITA	2023-01-10	08:00	Present	
11	ANITA	2023-01-11	08:00	Present	
12	ANITA	2023-01-12	08:00	Present	
13	ANITA	2023-01-13	08:00	Present	
14	ANITA	2023-01-14	08:00	Present	
15	ANITA	2023-01-15	08:00	Present	
16	ANITA	2023-01-16	08:00	Present	
17	ANITA	2023-01-17	08:00	Present	
18	ANITA	2023-01-18	08:00	Present	
19	ANITA	2023-01-19	08:00	Present	
20	ANITA	2023-01-20	08:00	Present	
21	ANITA	2023-01-21	08:00	Present	
22	ANITA	2023-01-22	08:00	Present	
23	ANITA	2023-01-23	08:00	Present	
24	ANITA	2023-01-24	08:00	Present	
25	ANITA	2023-01-25	08:00	Present	
26	ANITA	2023-01-26	08:00	Present	
27	ANITA	2023-01-27	08:00	Present	
28	ANITA	2023-01-28	08:00	Present	
29	ANITA	2023-01-29	08:00	Present	
30	ANITA	2023-01-30	08:00	Present	

File Name : \$OUT_DIRECTORY\M2COutput\\$DATA_SET_Errors.xlsx

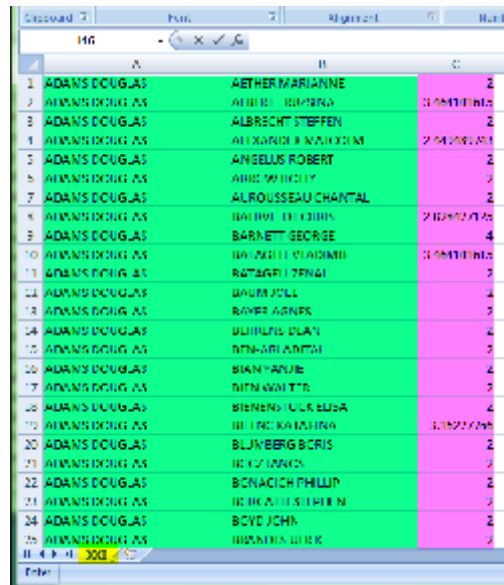
A unique error file is generated for each data set containing the error matrix with a single spreadsheet (yellow highlight). The data is presented in pairs (green and magenta highlight); only one set of each pair will be provided given the symmetrical nature of this data. Values (magenta highlight) are the distance error for the two objects when comparing the coordinates generated to the Structural Equivalence (SE) data calculated. The rows and columns (green highlight) are labeled accordingly.

The number of rows generated will be based on the number of objects (N):

$$Rows = N(N - 1)$$

1,000 Objects = 990,000 rows; 10,000 objects = 99,990,000 rows; etc.

SE (Structural Equivalence) File



	A	B	C
1	ADAMS DOUGLAS	AETHER MARIANNE	2
2	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
3	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
4	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
5	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
6	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
7	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
8	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
9	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
10	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
11	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
12	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
13	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
14	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
15	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
16	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
17	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
18	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
19	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
20	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
21	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
22	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
23	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
24	ADAMS DOUGLAS	ALBRECHT STEFFEN	2
25	ADAMS DOUGLAS	ALBRECHT STEFFEN	2

File Name : \$OUT_DIRECTORY\M2COutput\\$DATA_SET_SE.xlsx

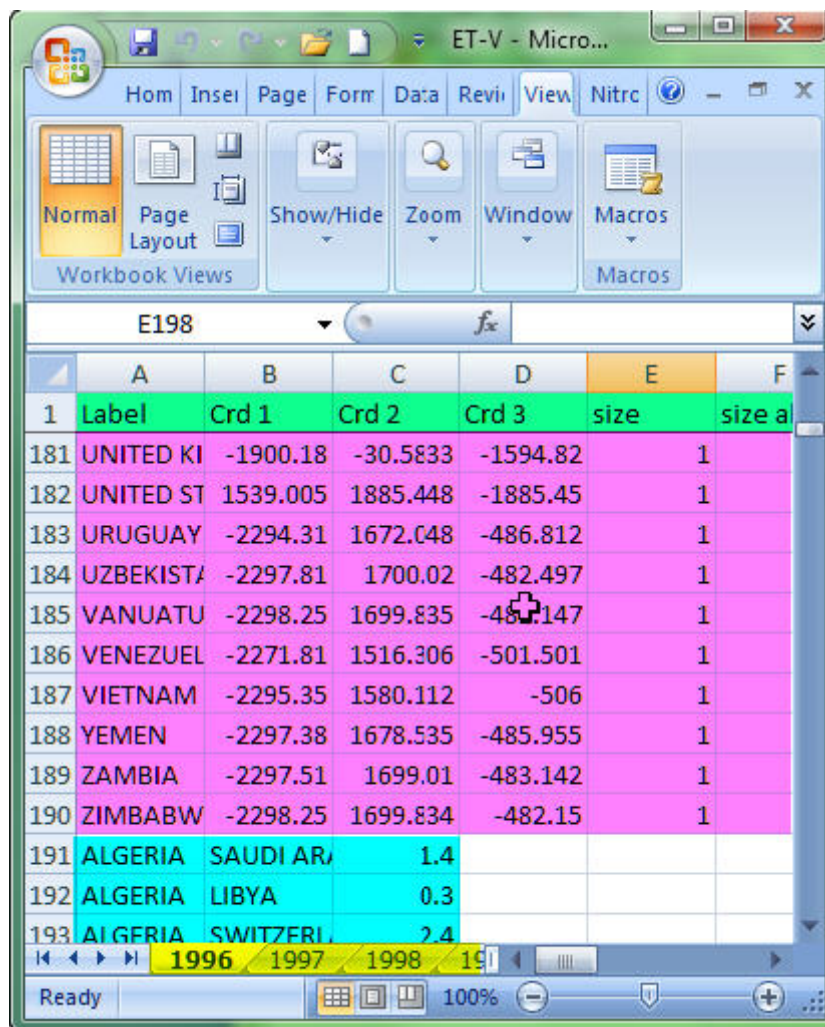
A unique error file is generated for each data set containing the SE matrix with a single spreadsheet (yellow highlight). The data is presented in pairs (green and magenta highlight); only one set of each pair will be provided given the symmetrical nature of this data. Values (magenta highlight) are the SE distances between the object labels provided. The rows and columns (green highlight) are labeled accordingly.

The number of rows generated will be based on the number of objects (N):

$$Rows = N(N - 1)$$

1,000 Objects = 990,000 rows; 10,000 objects = 99,990,000 rows; etc.

ET-V© (Visualizer) File



	A	B	C	D	E	F
1	Label	Crd 1	Crd 2	Crd 3	size	size a
181	UNITED KI	-1900.18	-30.5833	-1594.82	1	
182	UNITED ST	1539.005	1885.448	-1885.45	1	
183	URUGUAY	-2294.31	1672.048	-486.812	1	
184	UZBEKIST/	-2297.81	1700.02	-482.497	1	
185	VANUATU	-2298.25	1699.835	-482.147	1	
186	VENEZUEL	-2271.81	1516.306	-501.501	1	
187	VIETNAM	-2295.35	1580.112	-506	1	
188	YEMEN	-2297.38	1678.535	-485.955	1	
189	ZAMBIA	-2297.51	1699.01	-483.142	1	
190	ZIMBABW	-2298.25	1699.834	-482.15	1	
191	ALGERIA	SAUDI AR	1.4			
192	ALGERIA	LIBYA	0.3			
193	ALGERIA	SWITZERL	2.4			
	1996	1997	1998	1999		

File Name : \$OUT_DIRECTORY\M2COutput\ET-V.xlsx

The **Elbirt Technologies Visualizer (ET-V)©** file provides each visualization series on a different spreadsheet for each input data set (yellow highlight). The first row contains the column labels, in the required order, used by **ET-V©** (green highlight) and the object attribute data is provided on the rows that follow (magenta highlight). Coordinate dimensions 1 to 3 are used for the X, Y and Z coordinates. Dimensions 4 through 6 are used for colors if available.

Finally, the pair data for the given data set is provided (blue highlight) for every non zero pair in the dataset. These are used by **ET-V©** to create the edge/relation visualization.

Time Series Data File

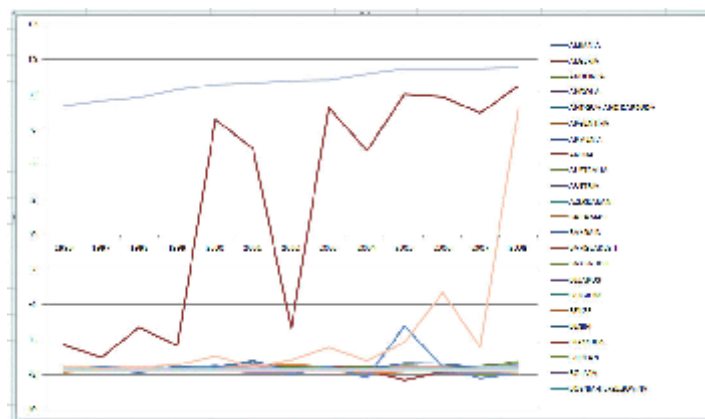
	ALBANIA	ALGERIA	ANDORRA	ANGOLA	ANGUILLA	ANTIGUA	ARMENIA	ARUBA
1996	2297.99	2291.55	2297.57	2297.79	2297.21	2275.62	2297.41	2295.97
1997	-2298.43	-2310.24	-2311.5	-2300.79	-2303.82	-2307.57	-2300.74	-2302.34
1998	-2302.07	-2306.03	-2301.22	-2299.03	-2291.42	-2299.9	-2298.08	-2296.03
1999	-2308.76	-2306.72	-2311.5	-2311.09	-2321.31	-2293.34	-2312.61	-2313.54
2000	2339.62	2335.98	2339.79	2338.67	2341.25	2296.59	2339.98	2341.61
2001	-2318.62	-2319.57	-2351.01	-2350.06	-2350.23	-2409.16	-2352.06	-2357.34
2002	-2324.55	-2342.09	-2307.72	-2309.9	-2307.75	-2286.6	-2308.27	-2310.1

File Name : \$OUT_DIRECTORY\M2COutput\TimeSeries.xlsx

The Time Series data file provides the coordinate data in a format that is better suited for time series analysis and graphing with line, bar and other graph types.

Each dimension is provided on two spreadsheets (yellow highlight); the first being the raw coordinate values and the second being a log (base e) of the coordinate data (magenta highlight). The column headers are the object labels (green highlight) and each data set name is provided as the row label (blue highlight).

The data can be statistically tested for time series patterns and graphed for visual representation as shown below.



Log (base e) Crd 1 1996 - 2008

Centralities Time Series Data File

	A	B	C	D	E	F	G	H	I	J
1	Object	ADAMS D	ADAMS E	ADAMS F	ADAMS G	ADAMS H	ADAMS I	ADAMS J	ADAMS K	ADAMS L
2	ADAMS D	0								
3	ADAMS E	0								
4	ADAMS F	0	0							0
5	ADAMS G	0	0	0						0
6	ADAMS H	0	0	0	0					0
7	ADAMS I	0	0	0	0	0				0
8	ADAMS J	0	0	0	0	0	0			0
9	ADAMS K	0	0	0	0	0	0	0		0
10	ADAMS L	0	0	0	0	0	0	0	0	0
11	ADAMS M	0	0	0	0	0	0	0	0	0
12	ADAMS N	0	0	0	0	0	0	0	0	0
13	ADAMS O	0	0	0	0	0	0	0	0	0
14	ADAMS P	0	0	0	0	0	0	0	0	0
15	ADAMS Q	0	0	0	0	0	0	0	0	0
16	ADAMS R	0	0	0	0	0	0	0	0	0
17	ADAMS S	0	0	0	0	0	0	0	0	0
18	ADAMS T	0	0	0	0	0	0	0	0	0
19	ADAMS U	0	0	0	0	0	0	0	0	0
20	ADAMS V	0	0	0	0	0	0	0	0	0
21	ADAMS W	0	0	0	0	0	0	0	0	0
22	ADAMS X	0	0	0	0	0	0	0	0	0
23	ADAMS Y	0	0	0	0	0	0	0	0	0
24	ADAMS Z	0	0	0	0	0	0	0	0	0
25	ADAMS AA	0	0	0	0	0	0	0	0	0
26	ADAMS AB	0	0	0	0	0	0	0	0	0
27	ADAMS AC	0	0	0	0	0	0	0	0	0
28	ADAMS AD	0	0	0	0	0	0	0	0	0
29	ADAMS AE	0	0	0	0	0	0	0	0	0
30	ADAMS AF	0	0	0	0	0	0	0	0	0
31	ADAMS AG	0	0	0	0	0	0	0	0	0
32	ADAMS AH	0	0	0	0	0	0	0	0	0

File Name : \$OUT_DIRECTORY\M2COutput\Centralities.TimeSeries.xlsx

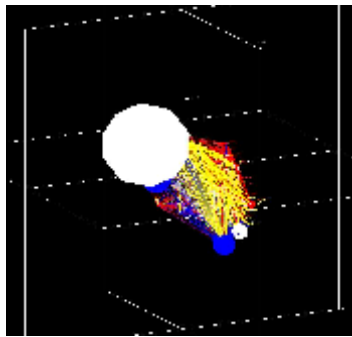
The Centralities Time Series data file provides the centrality calculation data in a format that is better suited for time series analysis and graphing with line, bar and other graph types.

Each centrality measure is provided on a spreadsheet (yellow highlight). The column headers are the data set labels (green highlight) and each object label is provided as the row label (blue highlight).

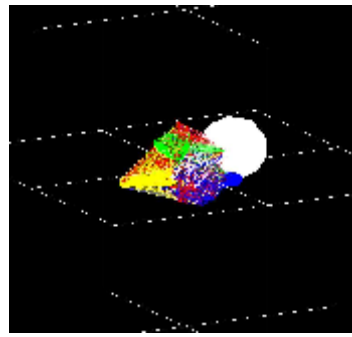
VI Time Series Rotation Algorithm - Evidence of Validity

The following example is provided as evidence of validity for the time series rotation algorithm based on the results of a time-series rotation vs. a non-rotated **M2C 4.0©** execution on the same data set.

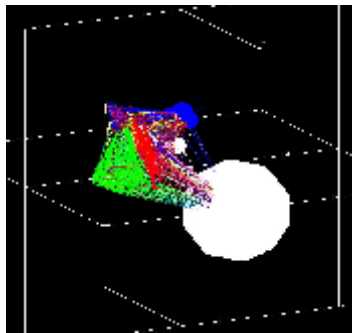
The data (TeleGeography, 2010) contains the number of minutes, in millions, of talk time from one country to the other by year for 13 years 1996 to 2008 inclusive. The number of countries represented varies by year from 189 to 213.



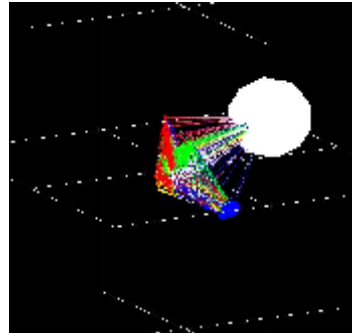
2002 - Unrotated



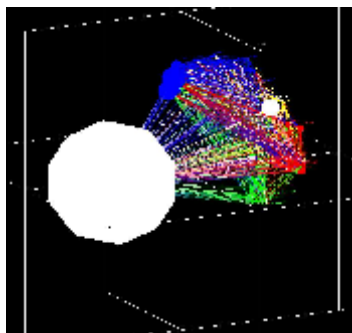
2002 - Rotated



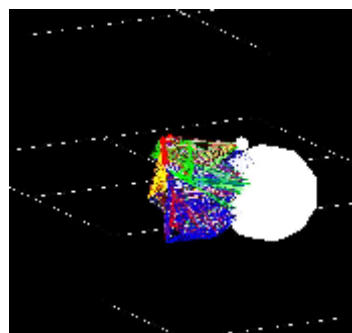
2003 - Unrotated



2003 - Rotated



2004 - Unrotated

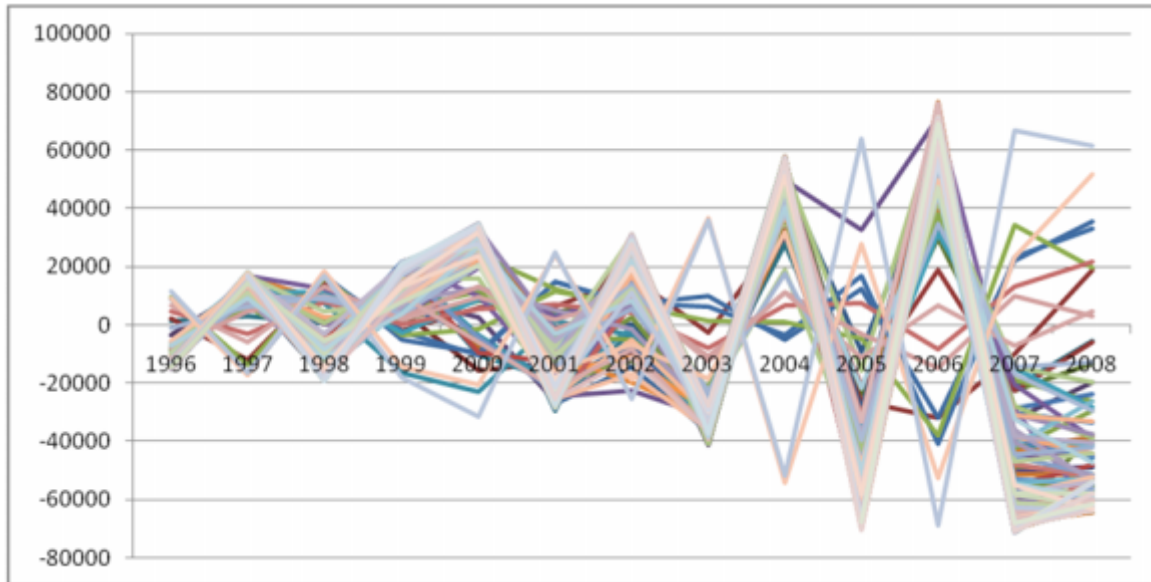


2004 - Rotated

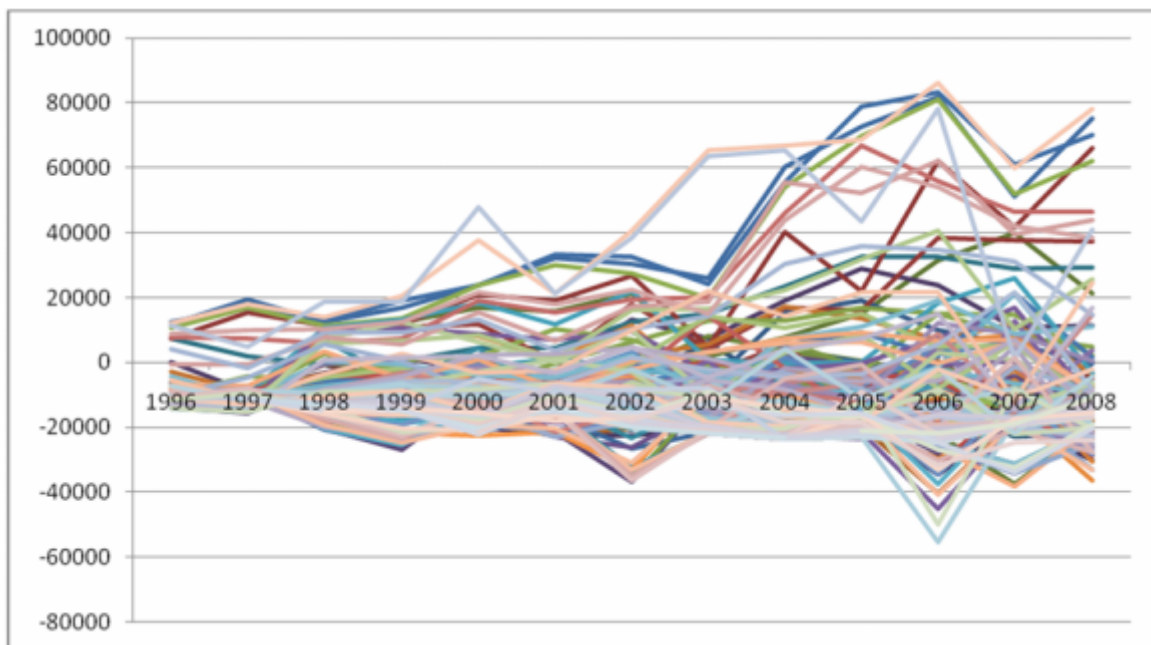
The six graphic panels above (created with **ET-V 1.0©** Screen Capture) illustrate the difference in outputs as a function of rotation. The 3 panels on the left, coordinates

without rotation, shows an obvious orientation change from panel to panel; simply follow the large white sphere (United States).

The 3 panels to the right, coordinates with rotation, show a less erratic rate of change; the large white sphere (United States) seems to move up then down while also moving toward the right (space expansion).



Time Series Data – No Rotation – Coordinate Dimension 1



Time Series Data – Time Series Rotation – Coordinate Dimension 1

The two graph panels above (created with *Microsoft Excel™* and the [Time Series Data File](#)) illustrates the difference of rotation vs. unrotated coordinate sets. The first panel,

no rotation, shows a volatility across the 0 coordinate axis much like the previous visualized graphics show the back and forth motion of the United States (white sphere). The space expansion is present and seems to evenly distribute across the axis (positive and negative) over time suggesting an oscillating pattern.

The second panel, time series rotation, shows a smoother change over time with less 0 coordinate axis volatility while preserving the expansion of the space. The expansion, however, is not evenly distributed across the axis (positive and negative) over time suggesting some pattern other than oscillation.

VII Bibliography

Algorithmics Group. "MDSJ: Java Library for Multidimensional Scaling (Version 0.2)". Available at <http://www.inf.uni-konstanz.de/algo/software/mdsj/>. University of Konstanz, 2009.

Brandes, U. and Pich, C., Eigensolver "Methods for Progressive Multidimensional Scaling of Large Data". Proc. 14th Intl. Symp. * Graph Drawing (GD '06). LNCS 4372, pp. 42-53. © Springer-Verlag, 2007.

J. Madadhain, D. Fisher, P. Smyth, S. White, Y. B. Boey (2005). "Analysis and visualization of network data using JUNG". Journal of Statistical Software: 125.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.101.3015&rep=rep1&type=pdf>.

Schonemann, Peter H., "A generalized solution of the orthogonal Procrustes problem". *Psychometrika*, 1966, 31, 1-10.

TeleGeography. "TeleGeography Report and Database". Available at <http://www.telegeography.com/>. Primetrica Inc., 2010.

Woelfel, J. and E. L. Fink (1980). The Measurement of Communication Processes: Galileo Theory and Method. New York, Academic Press.

Appendix A – Memory Recommendations

The following are memory recommendations based on physical memory available in the computer. These recommendations apply for both 32 and 64 bit operating systems.

System Memory	Minimum Value	Max Value
500 MB	100M	250M
750 MB	250M	375M
1000 MB (1 GB)	250M	500M
1500 MB (1.5 GB)	500M	750M
2000 MB (2 GB)	500M	1000M
3000 MB (3 GB)	1000M	1500M
> 3000 MB (more than 3 GB)	1000M	(½ Total Memory)
Example 3GB:		
6000 MB	1000M	3000M

Appendix B – Error Debugging

How to view the error output to the system command line

Errors that occur when using the short-cut method of application execution are not visible. To view the error messages open a DOS command window. Cut and paste the short-cut *Target* into the DOS window to execute the program and view error messages.

Catastrophic Failure - Application Auto Termination

1. **The application immediately closed when I attempted some action.**

There are many reasons an application will terminate; often the result of a code error or circumstance not foreseen by the developer and performed by the user. The following are a few known causes and solutions.

- a. **Memory Related Error** – Often, the error message provided by the Java Virtual Machine is something like:

Exception in thread "Thread-4" java.lang.OutOfMemoryError: Java heap space

To repair this problem first attempt to [alter the Short-Cut](#) provided with the application to increase the Java Virtual Machine memory allocation. You will need to add more physical memory to your computer if the problem persists and the maximum memory allowed is used by the Java Virtual Machine.

b. **Application Appears to Hang when Processing**

This usually results from a memory availability error. To be sure, open the *Windows Task Manager* and find the *javaw.exe* reference. Check the memory being used and consider the value against the settings used in your [short-cut](#). The number may be very close or larger than the value set; an out-of-memory error.

c. **Platform (32 vs 64 bit) Errors**

Exception in thread "main" java.lang.UnsatisfiedLinkError: C:\Program Files (x86)\Elbirt Technologies\lib\jogl.dll: Can't load AMD 64-bit .dll on a IA 32-bit platform.

This error, or something similar, results from having the wrong Java / OS combination for the installed version. Make sure you are using a 64 bit OS with 64 bit Java on 64 bit systems².

² Apparently the 32 bit version of Java can also be installed on 64 bit systems.

- d. **Data Load File Selection Halt** – I tried to add data sets and the File Selection GUI hangs up when I try to select a large group of files, or some sub-group of files in a directory.

Unfortunately this has occurred a few times during development and the reason is not clear to the developer. Often adding the files in smaller groups will work adding a few extra steps to the process.

Data Output Speeds

1. **The system seems to be processing my output, but it is taking forever!**

Unfortunately the output of the XLSX XML file content can take a long time on large data sets. The progress bar will also update with the current file being output more often to avoid concerns of application lock-up.

The output of the SE and Error files are the longest and largest files. To minimize generation time and memory needed for generation each file generated contains one data set.

You should also check the *javaw.exe* process running in *Windows Task Manager* to see if the Memory and CPU values change. This is an excellent indicator of application hang-up vs. a process that takes a long time.