8-2010

Visualizing the Social Links of Election Monitors with ArcGIS

Gregory C. Couch
University of Redlands

University of Redlands

Visualizing the Social Links of Election Monitors with ArcGIS

A Major Individual Project submitted in partial satisfaction of the requirements for the degree of Master of Science in Geographic Information Systems

by

Gregory C. Couch

Mark Kumler, Ph.D., Committee Chair

Ruijin Ma, Ph.D.

August 2010

Visualizing the Social Links of Election Monitors with ArcGIS

The report of Gregory C. Couch is approved.

Ruijin Ma, Ph.D.

Mark Kumler, Ph.D., Committee Chair

August 2010

Acknowledgements

I cannot thank my wonderful wife Erin enough for her patience, support, and love throughout the past year here. She has supported me through good times and bad and has put up with my frustrations with grace and gentleness.

While all of the staff of the MS GIS program deserve acknowledgment, specific thanks go out to Drs. Mark Kumler and Ruijin Ma for their advice, assistance, and patience on this project.

Debbie and Ruben, your assistance and help at all hours of the day and over weekends was invaluable. This program could not function without you.

Steve and Stephen, the help you provided in coding and debugging was invaluable. Thank you! It prevented minor outbreaks of "code rage".

I'd also like to 'thank' In-n-Out Burger and the Coffee Bean & Tea Leaf for the sustenance and the caffeine that kept me going throughout the year.

"Well, here's what the whales are saying."
…And thus ends eleven months of work.

Abstract

Visualizing the Social Links of Election Monitors with ArcGIS

by

Gregory C. Couch

Spatial analysis of the network of personnel in election monitoring and democratization missions allows users to visualize the impact and interactions these individuals and mission teams have had. This project focuses on the integration of the ArcGIS geographic information system and a private Access database to allow the user to input, edit, and visualize spatial and temporal data during the research of election monitoring networks. These data include the individuals, the missions, the positions held, and the dates of each activity. As these missions and individuals interact through space and time, network analysis will yield nodes of importance. This analysis will assist anthropologists in their study of the spread of democracy and in their study of the individuals supporting election assistance.

Table of Contents

Table of Figures

List of Tables

List of Acronyms and Definitions

EU	European Union
EM	Election monitor, election monitoring
ESRI	Environmental Systems Research Institute
GIS	Geographic Information System
GIScience	Geographic Information Science
i_MMDB	Intermediate Missions Management Database
IPUMSI	Integrated Public Use Microdata Series International
ISO	International Organization for Standardization
LTO	Long Term Observer
MMDB	Missions Management Database
MS	Microsoft
NGO	Non-Governmental Organization
ODIHR	Office for Democratic Institutions and Human Rights
OLE	Object Linking and Embedding
OSCE	Organization for Security and Co-operation in Europe
PC	Personal Computer
SNA	Social Network Analysis
STO	Short Term Observer
UML	Unified Modeling Language
UN	United Nations
VBA	Visual Basic for Applications

Chapter 1 – Introduction

International election monitoring and democratization missions have increased in number and scope since the mid-1990's. During that time, some monitors and democratization experts have participated in multiple short and/or long-term missions. Despite the part-time employment provided by each mission, individuals have strung together multiple missions to create almost full-time careers out of such participation. Naturally, some of these individuals rise in stature or assume greater responsibility on subsequent missions. The individuals involved in these fields are creating a profession out of their participation, and the overall phenomenon of democratization and election monitoring is becoming institutionalized. These fields are becoming their own industry for the participants (K. Coles, personal communication, September 23, 2009). Accordingly, increasing amounts of money and resources are flowing to this fledgling industry. Recalling the old adage of "follow the money," changes such as these often pique academic and political interest. There is a need for a formal method to analyze and visualize the relationships between individuals, nations, and these election monitoring missions.

1.1 Client

Kimberley Coles, Ph.D., an Assistant Professor of Anthropology at the University of Redlands, is the client for this project. Dr. Coles' interests lie in political anthropology, social transformation, and international institutions/global governance. Her work focuses on democratization assistance and international aid as exemplified by past non-governmental organization (NGO) activity in Bosnia and Herzegovina (University of Redlands, n.d.). Dr. Coles has personal experience in international election monitoring and has compiled data for more than 230 democratization and election monitoring missions. Her dataset covers approximately 2,100 individuals participating in more than 230 missions involving 4,000 individual assignments. Currently, the dataset is stored in a Microsoft Access 2003 database (K. Coles, personal communication, October 27, 2009).

1.2 Problem Statement

As part of her research into election monitoring and democratization networks, Dr. Coles identified a need to incorporate spatial and temporal analysis in her work. Her stated desire is a "social network analysis tool and [a visualization of] the spatial and temporal effects on that [network]" (K. Coles, personal communication, September 23, 2009). Dr. Coles holds her election monitoring research data in Microsoft Access database format, in her Missions Management Database (MMDB). In its current format, the dataset cannot be analyzed within a network, either spatially or temporally. Because Dr. Coles desires to understand the relationships amongst the actors in the missions, she must analyze the participants within the context of their networks. These relationships form as election monitors interact during the execution of their monitoring/advising missions, but the dispersion of individuals post-mission creates a network that has a very large spatial extent as well as identifiable bonds between repeat participants. The client needs a way to analyze and visualize who goes where, with whom they interact, and what influence they have. In a way, the program allows in-depth analysis of the 'who,' 'where,' and 'when' of the social networks so that the client can answer 'why.'

1.3 Proposed Solution

This project integrates ArcGIS 9.3, a geographic information system (GIS) with the client's Missions Management Database to allow the client to easily input, manipulate, analyze, and visualize her data associated with election monitoring and democratization missions around the world. It links and displays the participants in the international election monitoring (EM) network both spatially and temporally. It also allows for the export of the network to Pajek and NodeXL, social network analysis programs for non-spatial visualization.

1.3.1 Goals and Objectives

This project was designed to provide the client with a program that improves her visualization and spatial analysis capabilities within her anthropological study of election monitoring and democratization missions. (For the purposes of this document, the terms election monitoring missions and democratization missions will be referred to singly as "EM missions," "EM," or "the missions," as well as interchangeably and without respect to whether they are election monitoring or democratization missions.) Using the program, the client will better understand the worldwide circulation of election monitors and how their past social experiences may influence their current or future performance within their social networks. Visualizing and analyzing the spatial relationship of this network should yield insights into specific nodes of influence and who.

The GIS Tool for Election Monitoring Research ("the project") displays social networks geographically on a base map. Currently, the client has no way to visualize her data in the context of a network or in any spatial aspect. Because of this problem, 'nodes of influence' (people, groups, or areas) maybe be overlooked or undervalued in her research. The client's dataset, held in the MMDB, is robust and thorough in storing and returning her data. However, it yields no contextual information regarding networks of individuals or organizations. Displaying the spatial network was the highest priory of this project. Through the integration of the MMDB, ArcGIS, and other elements the project allows the client to visualize and manipulate her research data in a networked and spatial context.

Furthermore, while achieving the primary goal of displaying the network geographically, the program also displays the linkages as migration flow lines with widths dependent on the user-defined 'strength' of each relationship. The relationship strengths are quantified in two ways. The first, and default, method is simply a function of the cumulative number of interactions between individuals or organizations and is calculated automatically. The second is a user-implemented override of the automatic calculation to account for relationships not recognized in the context of the MMDB. The most basic example of this would be a married couple within the EM network. This hypothetical couple might have only worked together once on a mission, which would lead to a generally weak linkage. However, the relationship strength can be updated manually to reflect the user's knowledge of a stronger linkage. This manual relationship setting addresses external or tangential knowledge of the network that is not or cannot be held within the MMDB. It is not anticipated to be used frequently, as standard / consistent link strengths allow for better quantitative analysis. Third-party scripts have been written to manipulate the width of flow lines in ArcMap and are integrated into the overall system architecture of the program. Thus, the client will be able to use the default relationship-strength function or assign strengths manually to set the width of the flow lines. This will allow quick, visual depiction of the relative strengths of relationships between the displayed entities.

The program also allows manipulation of the views and base maps so that the user can print or display presentation-quality graphics. The exact definition of presentation quality varies from user to user, but for this project, it is defined as printing at a display of 300 dots per inch or better and should be capable of printing to a plotter a 3 foot by 5 foot sheet with no loss of data through pixelation. The client had no way to visualize or share her data in a spatial or networked context, and without any way to print or display the output of this program to an audience or class, the program would have a greatly reduced benefit to the client. The ArcGIS portion of the program handles the output function as ArcMap already has sufficient production capability (print, export to .pdf, etc.). Discussions with the client identified approximately four base maps that are included in the program.

1.3.2 Scope

This project was custom designed for a single user working on a stand-alone personal computer (PC). It was intended to utilize existing programs and licenses to keep costs low for the client. It was built upon agreed-upon parameters. This section lays out those parameters. The responsibilities of the developer and the client are detailed below.

The program was built as a 'mash-up' of MS Access databases, ArcGIS, Pajek, and NodeXL. It was designed to operate as a stable program not adversely affecting any of the client's data or system. A critical requirement was that it not impact the operation of, or data held in, the MMDB.

The project was intended to fulfill Dr. Kimberley Coles' requirements for visualization of a social network in a geographic context. It was designed for use by Dr. Coles, her representatives, or workers in order that they might further explore the data collected by Dr. Coles. Any other use is not supported and is at the user's own risk.

The program design was limited in scope to analyzing the dataset held in the MMDB. It was not intended to accept other database data or other database formats. Although limited in this way, the import of identically formatted tables into ArcGIS would probably work but is not recommended.

Initially, the client was to receive up to eight hours of one-on-one training in the use of the program in addition to a written help document. This one-on-one training proved to be not possible because of the client's departure from the area prior to completion of the program.

The client's responsibilities were to provide a copy of the MMDB, as well as sufficient data to load and test during program build. Upon completion of the project, the client or her representatives became responsible for all subsequent data entry and analysis. She provided the PC hardware for her own use. The University of Redlands already held required software licenses for MS Access and ArcGIS 9.3. Pajek and NodeXL are free programs and were downloaded from their respective sites.

1.3.3 Methods

After identifying the client's needs, research into three aspects of the program was required: social network and link analysis (to include graph theory, to a limited extent), GIS technology, and database design, creation, and modification.

Selection of the GIS program was relatively simple. ArcGIS is the leading program in the field of GIS, and coupled with the unique affordability and accessibility provided by the University's GIS program to the faculty, it was the logical and simple selection for the GIS. ArcMap is currently more capable in adding maps and allowing program development than the only other consideration in this category, Google Earth. Although having a much larger non-professional user base and being the most affordable program (free), Google Earth does not offer anything near the spatial analysis capability of the ArcGIS suite. Furthermore, the ability to export from ArcMap to Google Earth (.kmz/.kml) ensured Google Earth could still be used for other visualizations, if ultimately necessary.

Selection of the social network analysis (SNA) program was much more complicated. The descriptions of various SNA programs considered can be found in Section 2.2.1, but an overview of the selection process is necessary here. i2's Analyst's Notebook was considered because of its widespread use by the US government, but was almost immediately discarded because of its high cost. Pajek and NodeXL met the affordability criteria, as both were free, and Visual Analytics' VisuaLinks appeared to be a viable option for a number of reasons. Initially VisuaLinks was the favored choice because of its ability to export some networks to ArcMap. However, the VisuaLinks' enterprise license required to perform this export proved to be too expensive when compared to the more limited but free programs Pajek and NodeXL. Weighing the options, it became clear that the essentially non-existent development costs of this program, using Access, ArcGIS, and free SNA programs, albeit more limited in capabilities, outweighed the added capabilities and added cost of the higher-end programs. Furthermore, it was found as the project research progressed that the significant portion of the analysis of similar projects was performed in the databases and not in the SNA program. The SNA programs became visualization tools to fill the gaps between the unvisualized database and the two dimensional, but geographical, displays in ArcMap. Interestingly, it was found that to display the network geographically, it was necessary to discard the graph theory principal that the "distance between nodes should express the strength or number of ties as closely as possible" (de Nooy, Mrvar, & Batageli, 2005).

1.3.4 Extent and Resolution of this Project

This project is world-wide in extent. Although the preponderance of mission datasets collected by Dr. Coles are in Eastern Europe and the Caucuses, the globalized nature of the EM network requires the project to be global in scope. It was also agreed upon that although some EM data is at city-level or lower, for this project, country-level analysis was sufficient.

1.4 Overview of the Rest of this Report

The following section discusses some initiatives proposed by leaders in GIS technology and academia who sought to integrate GIS further into anthropology and the social sciences. A discussion of the technology associated with social network analysis and how SNA and GIS are applicable to anthropology follows this section. A sample of prior work involving GIS and social networks and a discussion of the election monitoring industry, including typical team make-up and missions, is included.

Chapter 3 highlights the initial work done to identify the client's needs and the requirements to be included in the program. It also includes the overall architecture for the program, including the MMDB and ArcGIS.

Chapter 4 examines the required datasets and the design models of the geodatabase and the client's database. Chapter 5, Implementation, is the blueprint of the program. It provides documentation of the models and acts as a guide for future duplication, if necessary.

Chapter 6 enumerates the achievements, as well as the shortfalls of the project. It describes the results and leads into Chapter 7, which discusses future work that may enhance this project. It is not limited to this program or its client, but discusses avenues that may further the field of GIS and social network analysis.

Chapter 2 – Background

Within the field of social anthropology, there is a growing sub-field of election observation and consulting (K. Coles, personal communication, September 23, 2009). Although a consistent definition of a “democracy” or a “democratic” government has yet to be agreed upon (Paley, 2002), participants from various entities have created an industry, complete with its own labor force, surrounding the export of, and advice regarding, democratization (Coles, 2009). This industry concerns itself with planning, advising, consulting, or overseeing such political responsibilities as the electoral process, the drafting of constitutions, and oversight. Section 2.1 describes “democratization” for the scope of this project.

An industry is a social relationship, and a social relationship between people in a community constitutes a system (Mitchell, 1967). This system can be analyzed quantitatively, qualitatively, spatially, and as a social network. It is into this system that a GIS and social network analysis (SNA) can be implemented to visualize and analyze portions of the data. “Once the data are organized spatially…it becomes easy to display and to gain insight through more sophisticated concepts of spatial analysis” (Goodchild, 1996). However, despite the ease of display and insight afforded by spatial organization asserted by Goodchild, Wejnert argues that the actual study of spatial flow within a network (in regards to democratization) “has been rarely tested…” (2005, p. 55).

2.1 Election Monitoring and Democratization Missions

As the field of democratization grows, academic researchers have found a specific need for the visual representation of the spatial and temporal components of the social networks that make up the governmental and non-governmental organizations (NGOs) supporting the democratization movements. “Democratization” is often a broad and sometimes poorly defined word, but for the purposes of this project, democratization is a process-oriented approach by which nations, groups, or organizations assist and advise other nations or aspiring nations regarding the implementation of a democratic form of government.

These democratization movements most recently have occurred mainly in post-communist and post-authoritarian states. An example of such a post-authoritarian state is Bosnia and Herzegovina, which is “an emerging federal democratic republic” according to the Central Intelligence Agency’s World Fact Book (2009). Multiple democratization and EM missions have contributed to Bosnia and Herzegovina’s movement to a democratic form of government. The difference between EM missions and democratization missions is that while EM missions can be a critical portion of the democratization process, they are event-centered and more focused than the overall democratization missions are. The United Nations (UN), European Union (EU), and the Organization for Security and Co-operation in Europe (OSCE) have made the most significant contributions of labor to these democratization and EM missions (K. Coles, personal communication, October 27, 2009).

2.1.1 Democratization and Election-Monitoring Team Make-up

Election monitoring mission teams are comprised of individual observers and others who are designated as experts due to their background in a particular field, for example, in holding elections, in constitution composition, etc. These participants make up the assistance and advice teams for each mission. A 2006 OSCE EM mission to Albania is an example of a specific mission. Ten experts and 20 long-term observers (LTOs) began the mission on 13 December 2006 for elections that were scheduled for 20 January 2007. Long-term observers typically serve for one to three months on one assignment. Experts may be in-country for only a few days or weeks. In this instance, fifteen OSCE-participating states contributed members. OSCE requested 400 short-term observers to work in teams of two across the country on election day. The mission assessed the election's compliance with international election standards and national legislation (Organization for Security and Co-operation in Europe, 2006).

The Office for Democratic Institutions and Human Rights (ODIHR, an office of the OSCE), in its Terms of Reference, defines "Long Term Observers" as having at least "two to three years of relevant professional work experience at the national or international level with some experience in organizing and/or observing election processes" (Office for Democratic Institutions and Human Rights, 2006). In these two to three years of relevant professional work, an LTO should typically participate in multiple missions, each having a different make-up of observers, and probably serves these missions in different countries. A method to analyze these LTOs' participation in multiple missions and their interactions in the networks is needed. As an individual's responsibility within a mission increases, one would expect that person to exert more influence or impact within the network.

2.2 Anthropology and Technology

2.2.1 Anthropology and Social Network Analysis

The study of anthropology has been related to the study of social networks for some time. "The main objective of the modern anthropological field worker is to discover principles governing the interaction of the members of the society he is studying" (Marwick, 1967). The primary foundation upon which social anthropology rests is the observation and sampling of individuals. For example, in computing divorce ratios, the "unit of *analysis* is the marriage, but the sampling unit is the individual" (Mitchell, 1967). Mitchell (1967) also states, "in social anthropology…the unit of analysis is not the individual but the social relationships in which he is involved." In this project, it is the individual election monitors who were sampled, but the analysis is on the relationships between monitors. Advances in technology allow anthropologists to record, study, and analyze these networks through databases and social network analysis software.

Many programs are available to the anthropologist for analyzing social networks, ranging from simple, free visualization programs like NodeXL and Pajek to high-end enterprise systems such as VisuaLinks and Analyst's Notebook. A sample of programs is found in Table 2.1.

Table 2.1 – Social Network Analysis Programs

1. **Analyst's Notebook –** from i2. Its strengths come from its ability to gather information into patterns and trends. Capable of accessing multiple databases as well as drag-and-drop importing (i2, 2010).
2. **VisuaLinks –** from Visual Analytics Inc. It is a graphical analysis tool used to find hidden patterns, trends, and associations in multiple datasets (Visual Analytics Inc., 2010)
3. **Pajek –** a Windows-based program for the visualization and analysis of large networks (Pajek Wiki, 2008).
4. **NodeXL –** an Excel 2007 template allowing user-defined edge lists to produce network graphs within MS Excel (Social Media Research Foundation, 2010).

2.2.2 Anthropology and Geographic Information Systems

Knowing specific spatial and temporal locations of the individuals Mitchell wrote of (Section 2.2.1) allows further analysis to enhance the anthropologist's work. Leaders in the field of geographic information science (GIScience) are actively striving for the development of infrastructure to allow social scientists the ability to adopt a spatial approach (Goodchild, Anselin, Appelbaum, & Harthorn, 2000). As such, both anthropologists and GIS professionals are concerned with the "analysis of space and place" (Goodchild, et al., 2000), albeit in different ways at times. These two fields can come together when GIS are used to answer questions about the "location of human activities, the construction of social space, and the relationship between social space and physical environment" (Goodchild, et al., 2000).

Covey (2007) addressed spatial network analysis and GIS in his master's thesis for the University of Redlands MS GIS program. Using online databases holding information pertaining to terrorists and terrorism, he created a program for the United States Army that assists in the "production of consistent and realistic story lines" for training scenarios (Covey, 2007). Social network analysis is quite developed and numerous programs exist to visualize and analyze networks, as evidenced by the International Network for Social Analysis list of various SNA software programs. (International Network For Social Network Analysis, 2008). In Covey's thesis, he used the open-source SNA program Pajek to provide a non-geographic visualization of the network. However, Pajek was only used for non-geographic network visualization and was not necessary for analysis. Many SNA programs use database structures and much of the analysis of simple SNA programs can be duplicated or simulated in a relational database. Microsoft Access (2003 and 2007 releases) and Environmental Systems Research Institute's (ESRI) ArcGIS 9.2 were the backbone of his program. Using ESRI's Network Analyst extension, Covey provided the

Army with a predictive analysis tool that used the shortest paths between terrorist organizations and sympathetic groups, or the shortest paths between the terrorist organizations and potential targets, in order to evaluate links necessary for an attack to be carried out.

2.3 Summary

The use of technology to identify, record, and analyze social networks is becoming easier as programs are developed for this purpose. The spatial display of these networks has been generally omitted from such programs to the detriment of researchers. High-end SNA programs are beginning to address this, but remain out of reach of many analysts. The GIS Tool for Social Network Analysis was designed to fill the gap between SNA programs with no spatial dimension and the high-end SNA programs costing many thousands of dollars. The use of existing programs and licenses available to the client make this possible.

Chapter 3 – Systems Analysis and Design

The overall program is a combination of existing applications: ESRI's ArcGIS 9.3, the client's Microsoft Access database titled "Missions Management Database" (MMDB), an intermediate Access database, and the social network analysis (SNA) visualization programs Pajek and NodeXL. It uses new and existing queries to isolate and export the data from the MMDB through the intermediate database into ArcMap for spatial visualization and analysis. From there, it is available for export to Pajek and NodeXL for network visualization.

3.1 Problem Statement

This program was designed to fulfill the client's need for an affordable program to display a social network spatially and temporally within a geographic information system. It integrated MS Access, ArcGIS 9.3, Pajek, and NodeXL to accomplish this task using the client's pre-existing licenses and free software.

3.2 Requirements Analysis

Requirements were established through verbal and written communication with the client. Three main sets of requirements were identified for this project: functional, technical, and operational. The functional requirements addressed what the program should accomplish in order to meet the client's needs. The technical requirements are specifications for the needed software programs, software licenses, hardware required, and similar issues. The operational requirements are issues that were identified that would arise, or were anticipated to arise once the program was deployed. Transitional requirements will be discussed in Section 6.

3.2.1 Functional Requirements

The main requirement of the GIS Tool for Election Monitoring Research was that it performs analysis on data held in the Missions Management Database, process them into a dataset of social network data, and then display the result in ArcMap, Pajek and NodeXL. It depends on the successful integration of three elements: the MMDB, ArcGIS, and the SNA programs. A list of the functional requirements and their solutions is found in Table 3.1.

Table 3.1 – Functional Requirements

Functional Requirement:	***Solution(s):***
Query MMDB to elicit networks of interest, to include network by mission, network by country, and network by individual.	Custom switchboard added to intermediate database, MS Access queries, filtered and linked to the MMDB.
Process the queries from the MMDB to prepare for export as a network to ArcMap	MS Access make-table functions.
Display selected data from the MMDB in ArcMap.	"QuickMaps" toolset, scripted in Python to produce choropleth maps of elections held by year, countries visited by person, and member states of international organizations.
Geographically display network of election monitors as exported from the MMDB	ArcGIS custom toolbox and Python script.
Allow temporal display, when available.	Show/hide features based on user-input date/time range. Custom model.
Provide print capability of the network.	ArcGIS print functionality.
Export to NodeXL and Pajek for other visualization.	Custom model.

The first requirement was that the client select the data she wants to analyze in a network from the intermediate database. This is carried out from a custom switchboard element in the intermediate database that performs the necessary queries and filters the linked tables in the MMDB to return the appropriate data. The data are then appended to an empty table in Access and are available for use in ArcMap through a Microsoft JET 12 OLE connection. This becomes the network data table. The network data table is also available for full display in ArcMap and is rendered as points and lines using a custom toolbox and Python script. If temporal data are available, they may be used to filter the display of data. The QuickMaps toolset is also available for the client to create limited maps calling data from the MMDB without extensive filtering or querying. Examples of such Quick Maps include choropleth maps of countries belonging to a particular NGO/International Organization, or choropleth maps of countries contributing members to a mission. Detailed descriptions of these functions are all found in Sections 4 and 5.

3.2.2 Technical Requirements

Three main technical requirements were identified for this project: that it run independently of an internet or license connection on the client's own laptop; that it run on Windows XP, Vista, or 7; and that Office 2007 (specifically Excel 2007) be installed. Table 3.2 shows the technical requirements of this program.

Table 3.2 – Technical Requirements

Technical Requirement:	***Solution(s):***
Program runs on client's computer, independent of internet or license connection.	ArcInfo license tested and works on client's machine.
Program runs on Windows XP, Vista, or 7.	Client's new computer runs Macintosh Leopard OS. Install Parallels emulator with Windows OS. (Performed by University Information Technology Services).
Excel 2007 required for NodeXL.	Upgrade client's computer to Office 2007.

The original MMDB ran on the client's laptop PC running Windows XP and Office 2003. A new Apple MacBook Pro laptop was provided for her relocation and it was necessary to ensure a Windows emulator was installed. University of Redlands Information Technology Services installed vmware's® Parallels emulator. Although ESRI does not support its product on Apple's operating system, ArcGIS can generally run through this emulator. The client was notified that this was the case and understood that some limitations may occur. However, because this project does not require intense processing power, it is concluded that it should work in this hybrid environment. The client's computer change happened too late in development to allow time for testing the project on her new MacBook.

Because the client will be using the program away from the University of Redlands campus, it was important to ensure that the program work without internet connection or any license connection that could be interrupted by intermittent or non-existent communication technology. Included in this requirement was that the license be held on the machine.

In order to use NodeXL, the 2007 version of Microsoft Excel must be installed on the client's computer. Microsoft pricing and licensing is such that installing the Office 2007 Professional meets all the technical requirements for this project. NodeXL is an Excel template and add-in and runs only on the 2007 release of Excel. It is free and available at: http://nodexl.codeplex.com/. Figure 3.1 shows a screen capture from NodeXL demonstrating one of its visualization capabilities.

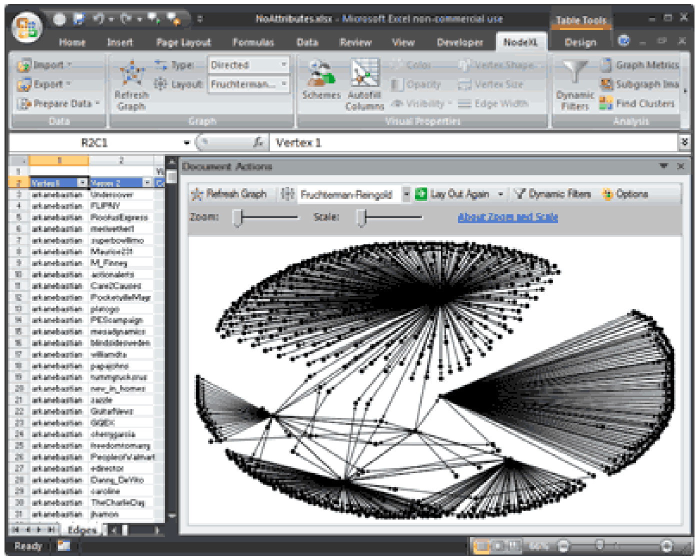

Figure 3.1 – NodeXL screen capture

Pajek (version 1.26) is a free software program available online at: http://pajek.imfm.si/doku.php?id=download. The installation file was included on the DVD that will be sent to the client.

3.2.3 Operational Requirements

This program is primarily a custom program to accomplish the task of analyzing the data held within the client's MMDB. It works with existing as well as new datasets as they are created and entered in the MMDB. It accepts data as-is from the MMDB and under the expectation and assumption that the records and attributes in the MMDB are accurate. It cannot address any errors inherent in the MMDB dataset, nor does it recognize such errors, except in cases where improper data types have been used. It does perform limited error-check any time users enter data in the user interfaces.

Originally, maintenance and support were to be provided while the developer and client remain in Redlands, but the client's departure from the area in summer of 2010 nullified this requirement. No follow-on support is expected to be provided, but may be available on a case-by-case basis via phone or email. Training videos and documentation on the use of all new tools may be provided at final delivery, but some online video chat training is planned as well.

Because the program is not designed for continual use, any modification or maintenance may require the shut-down of the system.

3.3 System Design

The overall project is designed to integrate four different software systems to accomplish the client's requirements. Figure 3.2 depicts the conceptual system design. The four different programs integrated are:

- Missions Management Database & Intermediate MMDB (MS Access 2003)
- ArcMap 9.3
- Pajek 1.26
- NodeXL Template (MS Excel 2007)

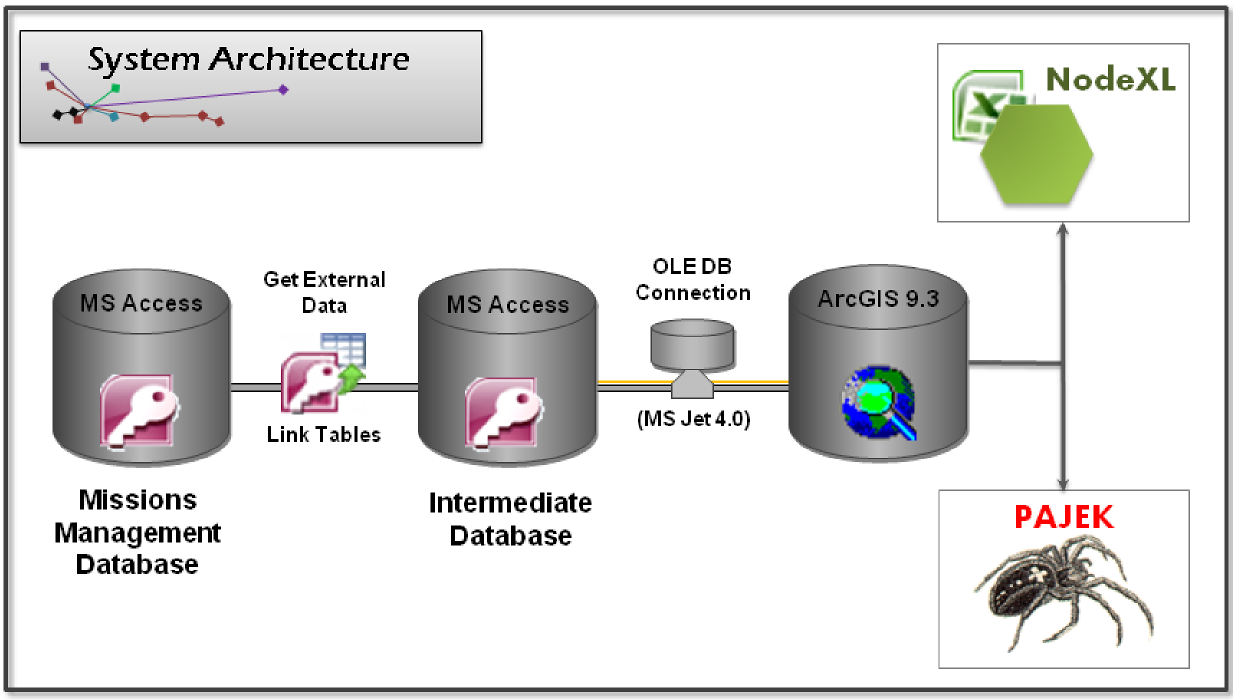

Figure 3.2 – System Architecture

The Missions Management Database, the intermediate database and ArcGIS are the main tools of the application. The client stores all data in the MMDB, but little analysis or querying will be performed from there in this program. That task falls to the intermediate database. The intermediate database is linked to the MMDB through a "Get External Data / Link Tables" linkage in Access. Most of the analysis and querying is performed in that program. The intermediate database is connected to ArcGIS through a Object Linking and Embedding database connection (OLE DB connection). Functionally, with the linked tables, the intermediate database appears to ArcMap to be the MMDB, but is actually an interim layer designed to facilitate the deployment of the program to the client's computer. The link was created in ArcCatalog using the Microsoft Jet 12 provider. OLE DB providers handle the exchange of information between the data source (intermediate database) and the data consumer (ArcMap) that requires access to data (ESRI). The connection allows dynamic communication between ArcGIS and MS

Access. Using filters and queries in both the intermediate database and in ArcMap, users can select data for visualization and further analysis in ArcMap.

For further 3D and SNA-specific (non-geographic) visualization, Pajek and NodeXL are included in the architecture. Export functions in the intermediate MMDB allow data to be formatted correctly for export to the respective programs.

3.4 Project Plan

Initially, the project plan was a straightforward waterfall approach starting at project selection in late September 2009 and concluding in late July 2010. Funding availability and professional obligations created a firm deadline for the project in mid-August 2010 (See Table 3.3 for the proposed verses actual schedules). A literature review would be conducted first, followed by selection of the SNA program, and building and integrating the parts of the project. Documentation, presentation and deployment of the project would commence in mid-summer 2010 and would be in a final state for presentation at the ESRI User Conference in San Diego by 12 July 2010.

The research into, testing, and selection of the SNA program was expected to be the most difficult portion of the project. Initially, coding and modeling were not expected to present as much of a challenge as they did. Writing would be time-consuming but not overwhelming. It was also expected that coursework and training in GIS practice and theory would provide the necessary knowledge for building the GIS segment of the project.

Significant problems delayed and hampered selection of the SNA program. Foremost was the developer’s lack of knowledge of social network analysis or any SNA programs. i2’s Analyst’s Notebook program was appealing because of its use throughout the US government, but its cost was prohibitive. A misunderstanding during the literature review suggested that the SNA software (Pajek) performed a significant portion of the analysis involved in Covey’s SNA/GIS project. This made the SNA portion of the project appear to be the crux of the overall analysis and not simply a visualization tool. This misunderstanding, coupled with the developer’s unfamiliarity with SNA programs, resulted in many weeks learning and evaluating more powerful SNA programs like VisuaLinks and Analyst’s Notebook, when all that was needed was a visualization tool. A fruitful relationship with Visual Analytics (VAI) was established through ESRI, with VAI providing significant technical support and assistance in the analysis of VisuaLinks. From December through April, VisuaLinks was the choice for the SNA portion of this project. Ultimately, the cost of a VisuaLinks license was prohibitive given the limited usage it would get in this project. The relationship with VAI was fruitful, but the need to switch to a more-affordable, but less capable program in April 2010 was a significant change for the project. A more thorough literature review and identification of license prices upfront would have been a better approach.

Because of the significant change in April 2010, other factors and problems arose. Communication between the various programs of the project needed to be established and re-coded. Using VisuaLinks, the connections between the MMDB, ArcGIS and

VisuaLinks would have been set up and hosted in VisuaLinks, vice the actual set-up in ArcGIS. VisuaLinks would maintain the connections and built-in functionality would decrease the coding burden on this project's developer. Ultimately, the cost of VisuaLinks outweighed the time savings and accessibility it provided, as much of the same functionality could be created as necessary. Table 3.3 compares the initial, planned timeline for the project versus the actual timeline.

Table 3.3 – Planned Deadlines and Actual Times

Task:	***Initial Plan:***	***Actual Schedule:***	***Other:***
Literature Review:	Completed by end of October 2009.	Up to 1 June 2010	
Identify SNA programs (followed by testing):	By beginning of November 2009	Mid-November 2009	
Select SNA Program:	18 December 2009	Final Selection: (Pajek/NodeXL) 1 May 2010	Initial Selection of VisuaLinks: 15 January 2010
Integrate SNA/MMDB/ArcMAP	28 May 2010	Mid-June 2010	
Testing:	June 2010	Ongoing throughout build.	
Documentation Complete:	10 June 2010	7 July 2010	
Delivery to client:	2 August 2010	TBD (DVD through mail).	Client departed Redlands mid-June 2010.
ESRI Presentation	12-16 July 2010	15 July 2010	
Defense:	1 August 2010	30 July 2010	

3.5 Summary

Identifying and understanding the client's requirements were crucial to creating the system design. However, misunderstandings of prior projects and the tools necessary to accomplish the client's requirements led to delays in project changes. Understanding the operational and functional requirements assist in planning and initial execution of the project, while when the technical requirements were evaluated, a general system design was formed. Once the overall system was designed, the database and geodatabase design and build phase began.

Chapter 4 – Database Design

As discussed in the previous chapter, this project integrates the Missions Management Database (MMDB), an intermediate database, an ArcGIS project, and the SNA programs Pajek and NodeXL. This chapter discusses the design of the MMMDB, the intermediate database, and the geodatabase that makes up a portion of the ArcGIS project.

4.1 Conceptual Data Model

The client's data model is a network of people, places, events, and organizations. A number of diagrams visualize the concepts needed for this project. The first is the mission composition diagram, followed by the relational model of people and places tracked by the MMDB.

The focus of the entire Missions Management Database is on the missions and their participants. An NGO sponsors a mission to go to a country or region and observe or assist an election. Each mission is organized into three main groups: the Core Team, Long Term Observers (LTOs), and Short Term Observers (STOs). The Core Team is usually seven to twelve people and forms the political and organizational leadership of the mission; it includes the Head of Mission, his or her deputy, and lawyers, logisticians, and others. The nucleus of each mission is the group of Long Term Observers. The LTOs observe the country's preparation for the election as well as lay the groundwork for the arrival of the STOs just prior to the election event. On a typical mission, the LTO group consists of 10 to 30 people and the STO group typically between 100 to 300 people. The members in each group are organized into two-to-four person teams. EU EOM missions have an additional grouping of STO participants. The overall number of STO participants is usually 100 to 300 people, but to better organize and manage the participants, the EU EOM groups them into cohorts. A cohort is a grouping of two to four STO teams. As each STO team has two to four members, a cohort of two to four teams would have anywhere from four to sixteen individuals in it. Figure 4.1 depicts the conceptual make-up of a typical mission. The hierarchy of a mission is shown in Figure 4.2.

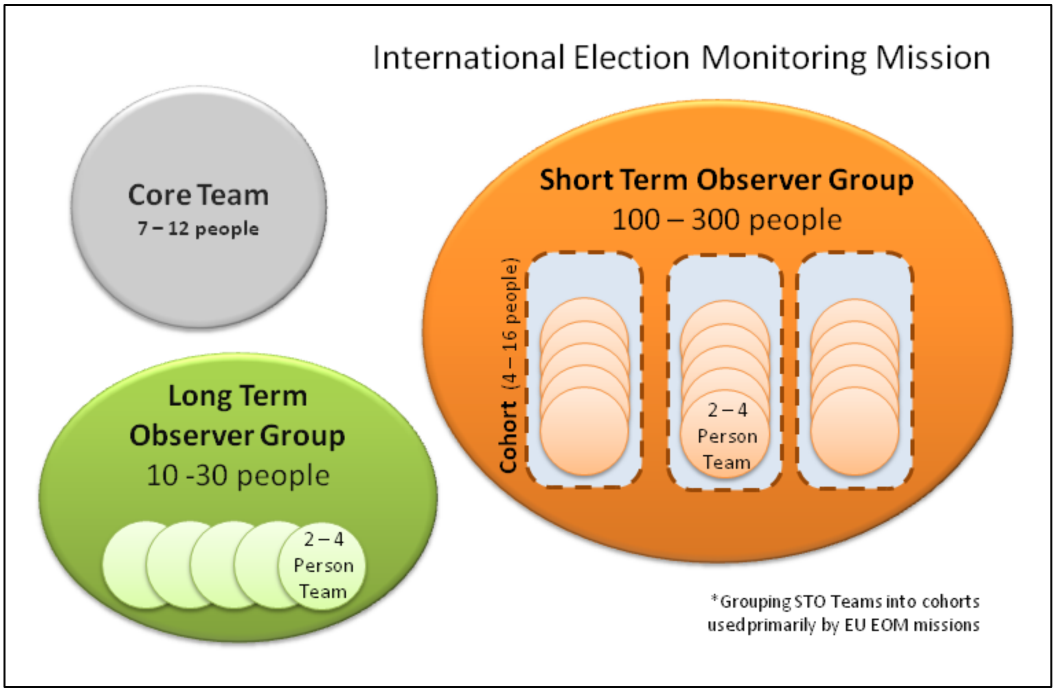

Figure 4.1 – Mission Composition

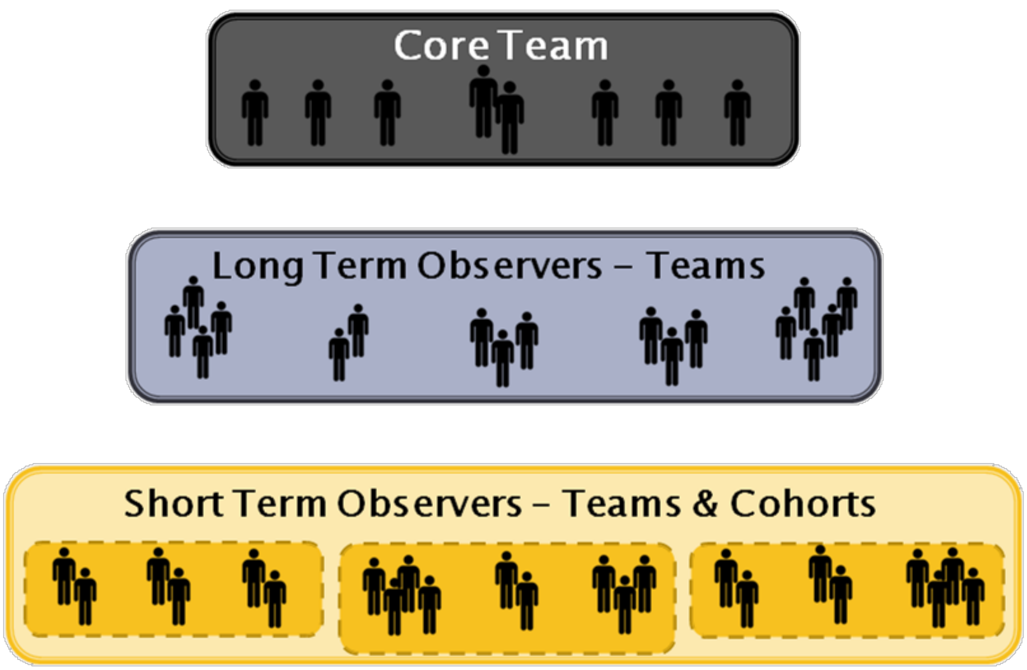

Figure 4.2 – Mission Hierarchy

Within the MMDB, each member of a mission holds one citizenship and one home country (home of record – HoR), though these may not be the same country. The Unified Modeling Language (UML) model in Figure 4.3 diagrams such relationships. In reality, some people hold multiple citizenships and have homes in multiple countries. The MMDB is currently storing only one HoR and one country of citizenship.   The person works in a country on a specific mission, and most work in multiple countries over their career.

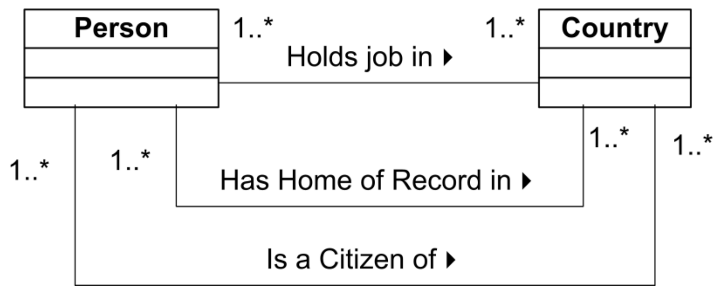

Figure 4.3 – Person - Country Relationship

Other relationships in this model relate the NGOs, the missions, the participants, and their countries. NGOs send missions to countries and, conceptually, the NGOs are non-state actors and do not have a specific location themselves. Their existence is by virtue of the collection of individuals that constitute the NGO. The NGOs and missions are made of individuals who may have relationships with one or more countries.

Finally, this database models the relationships amongst the mission participants. For this model, individuals form relationships by virtue of serving on the same team on each mission. The smaller the team, and the longer the individuals work together, the stronger the relationship. This does not necessarily account for "good" or "bad" relationships, but simply that a link exists between the individuals. The further removed people get from interpersonal relationships built by the teams, the weaker the link becomes. Two short-term observers on the same STO team have a strong relationship link by virtue of working together in geographic proximity over time, but an STO and a Core Team member would not have a strong link, even though they worked on the same mission.

4.2 Logical Data Model

As the program is a combination of three database entities – two instances of Access and one of ArcGIS – it is necessary to document each database structure. To reiterate the overall structure, the MMDB was the client's original Access database for mission data and remains unchanged. The intermediate database was a construction that facilitated building the project without the need to modify the MMDB, as well as easing the delivery of the project to the client at its completion. Only the supplementary data for visualization in ArcMap were built into a geodatabase. This included the oceans, water bodies, and meridians feature classes. The countries shapefile was left outside the geodatabase because of errors that arose in the model-building process.

4.2.1 The Missions Management Database

The Missions Management Database was acquired from the client and included data for over 2,100 individual mission participants, 230 missions, and over 4,000 job assignments. The data are held in ten tables (Figure 4.4). These tables are populated via input forms, but the data are manipulated with twenty custom Make-Table queries instead of the usual Select queries. A Make-Table query is sub-type of the action query group of queries in

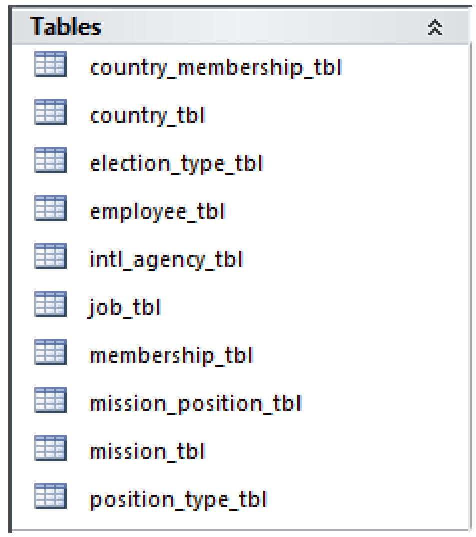

Figure 4.4 – MMDB Tables

MS Access. As the query runs, it creates a new table and saves the results in the new table. The default query type in Access is the select query. A select query returns the query results in a table, but does not save the table when it is closed. Other action queries delete data, update data, or append data. Figure 4.5 illustrates the relational structure of the MMDB.

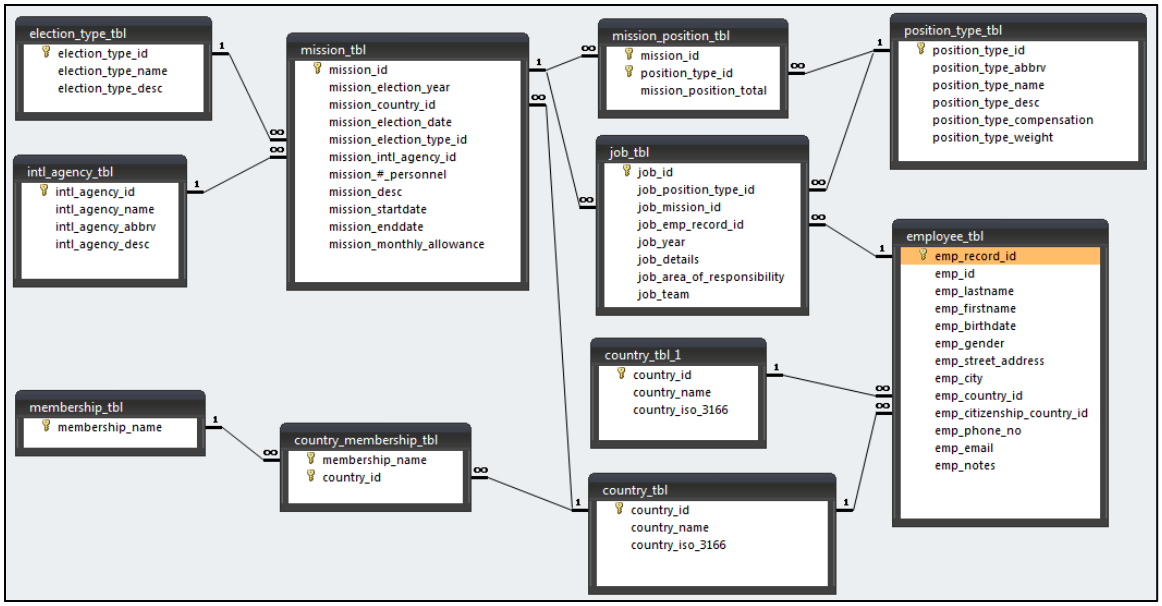

Figure 4.5 – MMDB Relationship Table

4.2.2 The Intermediate Database – i_MMDB

The intermediate 2007 Access database, the i_MMDB, was linked to the MMDB through an "Import" feature included in Access. The Get External Data – Access Database command contains an option to import or link to data sources. Importing the tables, queries, and other data objects from the MMDB would not have allowed the program to be dynamic. The user would have to re-import the tables and other objects every time new data were entered into the MMDB. Selecting the "Link to the data source by creating a linked table" option created the linked tables within a new database. Because the client retained the MMDB and provided a copy to build this project, an intermediate database allowed development of all necessary tools within Access, but when delivered to the client the intermediate database could simply be "re-pointed" to the original MMDB on the client's computer.

The queries used by the project are maintained in the intermediate MMDB. A simple query for a "QuickMaps" function, "Display Missions by Year," is included in Figure 4.6.

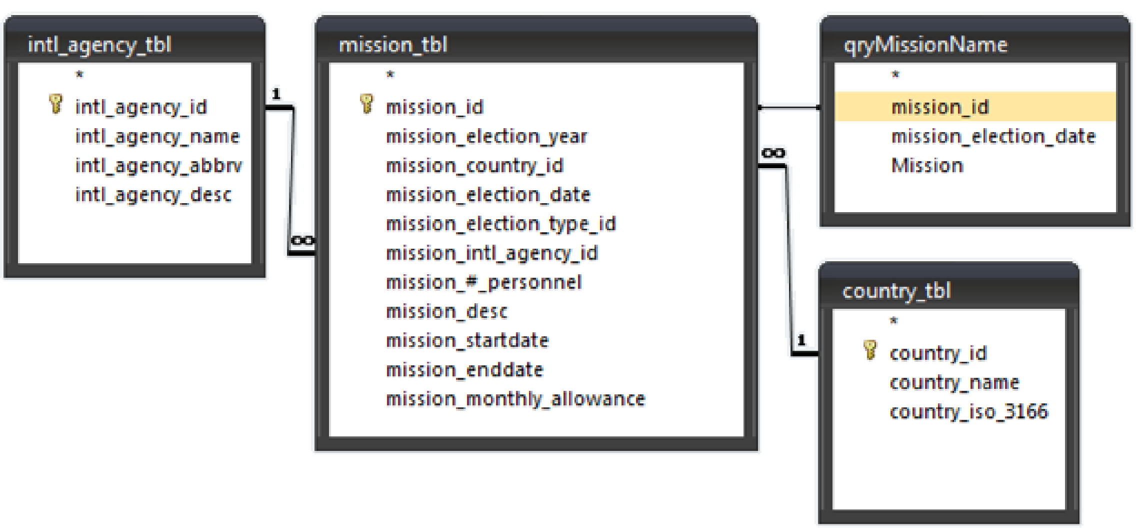

Figure 4.6 – Missions by year query

The field names for the queries were changed because later processing in ArcMap truncated the field names to ten characters or less. When using ModelBuilder to copy and manipulate these tables, the truncated field names could not be identified or predicted automatically within the intermediate data tables, and thus could not be used. Renaming the fields in the Make-Table queries made the later processing possible. This process of renaming fields in the intermediate Access database queries occurred for every query. This also ensured that the original fields from the MMDB and later fields created using queries would not be confused during programming.

4.2.3 ArcGIS – Shapefiles and the Geodatabase

The ArcGIS portion of the project was structured in two forms – one set of data in the geodatabase and one set of shapefiles. The main two datasets were the countries shapefile and the country interior points shapefile (Figure 4.7). Both are discussed in Section 4.3. The datasets were left in shapefile format because unspecific errors occurred while manipulating them as feature classes, yet the errors did not occur when they were in shapefile format. Many of the errors occurred when the feature classes were used within ModelBuilder and were part of the processing. Because there was no need to maintain topology or other feature class unique attributes, there was no pressing reason to convert the files to feature classes.

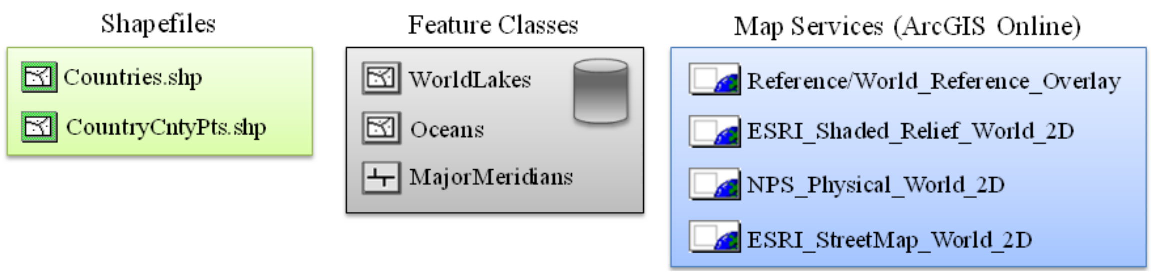

Figure 4.7 – Shapefiles, Feature Classes, and Map Services

Background data were created to add context to the map. Oceans, major water bodies, and major meridians were used as feature classes and built within a file geodatabase to maintain their projection as well to store them within the same file structure.

Links to ArcGIS Online Map Services were also included to give the user other options for basemap visualization. The project requirement remained that the system not require an internet connection, so these basemaps were included only as optional reference layers. The other shapefiles and feature classes already in the project provide sufficient functionality apart from these ArcGIS Online layers. These Map Service files were linked at "arcgis on services.arcgisonline.com".

4.3 Data Sources and Collection Methods

All of the original data for the project were compiled from four places.

- Missions Management Database
- Minnesota Population Center
- ESRI Data & Maps 9.3.1 (2009 Update)
- ArcGIS Online Maps Service

A copy of the MMDB was provided by Dr. Coles via electronic mail. The data from the Minnesota Population Center was downloaded as a .zip file. The ESRI Data & Maps dataset is held by the University and was downloaded from a shared drive. The ArcGIS Map Service data remain held by ESRI and are linked via an ArcGIS-managed connection.

All other data for the project were derived from the original data acquired from these sources.

4.3.1 Data Source: Missions Management Database Data

The MMDB dataset was compiled by Dr. Coles using methods including, but not limited to, personal communications, NGO websites, and her own participation in some missions. The content of the MMDB is her responsibility and data errors within the MMDB do not affect the overall functionality of this project.

The MMDB [country_tbl] table was selected to be the authoritative source for countries (and non-recognized countries) for this project. Although she uses the International Organization for Standardization (ISO) codes for the countries, Dr. Coles does study countries like Kosovo that are not recognized by the entire international community. She also studies areas that are seeking independence and have no international status at all. As a result, the [country_tbl] in the MMDB became the authoritative source for the "country" entities included in the project. Furthermore, it also became the authoritative source for the spelling and abbreviation of each country in the project. Overall, the MMDB became the source for all spelling, formatting, and other standards for this project. That included mission naming conventions, participant names, job titles, and job abbreviations. Using the same format for all data attributes as the MMDB ensured data in the project were consistent and interoperable.

4.3.2 Data Source: Minnesota Population Center – World_GIS Data

The Minnesota Population Center at the University of Minnesota has created the Integrated Public Use Microdata Series International (IPUMSI) website as a project to collect and disseminate census datasets from around the world (Minnesota Population Center, 2010). The Center maintains downloadable GIS files for mapping their data. One of these files, named "World_GIS", was the source for the countries.shp file used in this project. This data was selected because of its format; much of the data was already formatted similarly to the MMDB, and it was produced for visualization at a suitable scale range. It also included detailed metadata. Its major drawback, however, was that it did not account for boundary changes over time (Minnesota Population Center, 2010). Furthermore, because Dr. Coles researches areas holding elections that are emerging or non-recognized countries, these areas had to be created in the shapefile to match the data in the MMDB. Further details are provided in Section 4.4.

The "World_GIS" file was produced by IPUMSI for use in mapping their world-wide census data. It included some of the census variables the Minnesota Population Center uses for their analysis, and those variables not related to this project were deleted. Figure 4.8 shows a sample of the attribute table of the World_GIS file.

FID	Shape *	FIPS_CNTRY	name	bplctry	nation	mgctry1	mgctry2	mgctry3	mgctry5	CNTRY	age_mean
0	Polygon	AA	Aruba	21030	21030	21030	21030	21030	0		0
1	Polygon	AC	Antigua & Barbuda	21020	21020	21020	21020	21020	0		0
2	Polygon	AE	United Arab Emirates	34170	34170	34170	34170	34170	406		0
3	Polygon	AF	Afghanistan	32010	32010	32010	32010	32010	401		0
4	Polygon	AG	Algeria	13010	13010	13010	13010	13010	302		0
5	Polygon	AJ	Azerbaijan	34020	34020	34020	34020	34020	430		0
6	Polygon	AL	Albania	43010	43010	43010	43010	43010	108		0
7	Polygon	AM	Armenia	34010	34010	34010	34010	34010	141	051	0
8	Polygon	AN	Andorra	43020	43020	43020	43020	43020	125		0
9	Polygon	AO	Angola	12010	12010	12010	12010	12010	301		0
10	Polygon	AQ	American Samoa (Eastern Samoa)	54060	54060	0	0	0	0		0
11	Polygon	AR	Argentina	23010	23010	23010	23010	23010	218	032	31
12	Polygon	AS	Australia	51010	51010	51010	51010	51010	434		0
13	Polygon	AU	Austria	44010	44010	44010	44010	44010	109	040	0
14	Polygon	AV	Anguilla	21010	21010	21010	21010	21010	0		0
15	Polygon	AY	Antarctica	0	0	0	0	0	0		0
16	Polygon	BA	Bahrain	34030	34030	34030	34030	34030	0		0
17	Polygon	BB	Barbados	21050	21050	21050	21050	21050	0		0
18	Polygon	BC	Botswana	14010	14010	14010	14010	14010	0		0
19	Polygon	BD	Bermuda	24010	24010	24010	24010	24010	0		0
20	Polygon	BE	Belgium	44020	44020	44020	44020	44020	101		0
21	Polygon	BF	The Bahamas	21040	21040	21040	21040	21040	204		0

Figure 4.8 – World_GIS Attribute Table (Sample)

The “World_GIS” shapefile from the IPUMSI website was renamed to “countries.shp” during post-download processing and will be hereafter referred to as such. This countries.shp file is a major portion of the visualization component of this project. It is overlaid on a background to find variations and changes.

4.3.3 Data Source: ESRI Data & Maps 9.3.1 Media Kit (2009 Update) Data

The requirement that the project function without an internet connection necessitated the inclusion of background data to give the user context and reference. The ESRI Data & Maps Media Kit included the necessary data. The need for context was satisfied with three layers: oceans, major lakes, and major meridians.

The Oceans file was derived from the world30 shapefile to quickly emulate a blue ocean background. The oceans feature class was included so that the client, when exporting graphics, could include a depiction of the oceans if desired. It serves for visual effect only and has no analytic purpose for the project.

The WorldLakes feature class was included to give visual context to the data frame. It is not necessary for analysis but helps define international borders. For example, five nations share borders in the Caspian Sea, and including it and other major world lakes gave countries the shapes with which users are familiar (Figure 4.9).

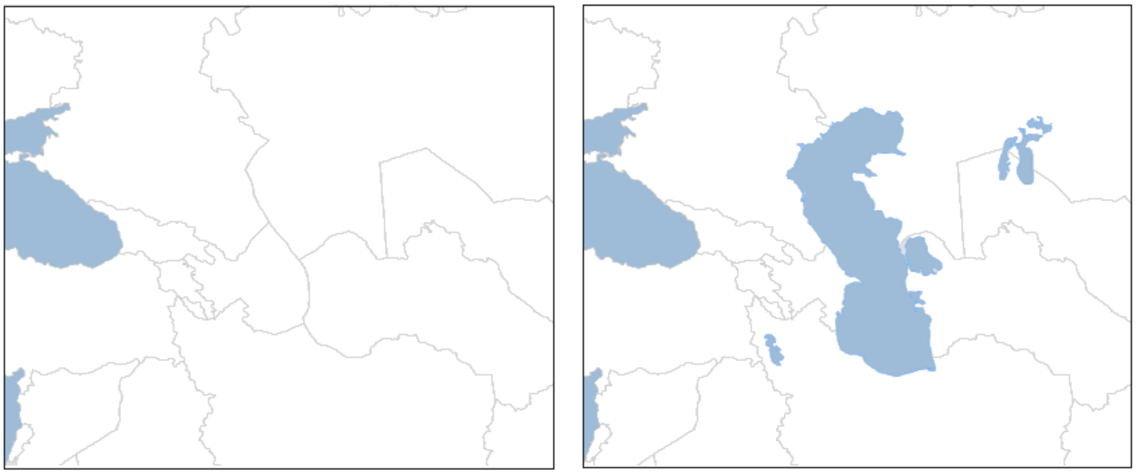

Figure 4.9 – Caspian Sea Area, WorldLakes Feature Class (above, right)

The MajorMeridians feature class provides further visual cues to users, similar to the WorldLakes shapefile.

4.3.4 ArcGIS Online Map Services

A final aesthetic addition to the project was the inclusion of the ArcGIS Online Map Services. Four Map Services were selected for this project:

- Reference/World_Reference_Overlay
- ESRI_Shaded_Relief_World_2D
- NPS_Physical_World_2D
- ESRI_StreetMap_World_2D

All four Map Services are accessed through the ArcGIS Server located at http://services.arcgisonline.com. These four map services were not critical for any analysis and were included only to add further context when the client has access to the internet.

4.4 Data Scrubbing and Loading

The decision to not modify the MMDB meant that most other acquired data needed modification to match those in the MMDB. The first modification to be made was to change the data structure in the World_GIS shapefile to match the MMDB. The original data structure of World_GIS is shown in Table 4.1.

Table 4.1 – World_GIS Original File Structure

Field Name	Data Type
FID	Object ID
Shape	Geometry
FIPS_CNTRY	Text
name	Text
bplctry	Double
nation	Double
mgctry1	Double
mgctry2	Double
mgctry3	Double
mgctry5	Double
CNTRY	Text
age_mean	Short Integer

The first step was to identify the necessary fields and delete the rest. The only necessary fields were FID, Shape, and name. The cntry_ISOa field was added because the MMDB uses the ISO 3166 country code abbreviation. The MMDB's unique ID field for countries, "country_ID", was added to the shapefile as "cntry_ID". Also, because the modified shapefile was designed to be used in join functions in other scripts and models, it was necessary to shorten the field names to ten or fewer characters so that the shapefile could be joined to tables whose field names were automatically produced. Subsequently, the file was renamed to Countries.shp and the data structure looked like Table 4.2.

Table 4.2 – Countries.shp Modified Structure

Field Name	Data Type
FID	Object ID
Shape	Geometry
cntry_Nam	Text
cntry_ID	Short Integer
cntry_ISOa	Text

All other instances of country shapefiles were copied from this file. They received different names but all contain the same fields and data types. They are used as intermediate data in processing.

Once the countries.shp shapefile was prepared, the fields and data types matched the MMDB [country_tbl] table, but the data did not yet match. The first action was to identify all of the discrepancies in the data. First, the country names in the countries file were converted to all uppercase to match the MMDB. The spelling (and accenting, when applicable) of the countries were fixed. Multiple types of changes were made, examples of which are found in Table 4.3. The full list of all changes is found in Appendix A.

Table 4.3 – Sample of Country Name Changes

Country Names Changed:		
Countries.shp	to:	MMDB
American Samoa (Eastern Samoa)		AMERICAN SAMOA
Antigua & Barbuda		ANTIGUA AND BARBUDA
The Bahamas		BAHAMAS
Bouvet I.		BOUVET ISLAND
British Virgin Is.		VIRGIN ISLANDS, BRITISH
Brunei		BRUNEI DARUSSALAM

The second set of changes made to the countries.shp file attributes was to add the countries to the shapefile that were in the MMDB but not depicted in the shapefile (Table 4.4). Reasons for the difference included the Minnesota Population Center's admission that their World_GIS file was not updated to reflect changing borders, the fact that some areas recognized as countries in the MMDB, such as Kosovo and Taiwan, did not have full international country status, and some small islands were simply not included in the shapefile. These areas were digitized from the ESRI Map Service data at a 1:3,000,000 scale, slightly larger than the overall countries.shp reference scale. They do not constitute an official endorsement of any of the areas or countries.

Table 4.4 - Entities added to Countries.shp

Countries or Areas Added	
Countries.shp	**ISO Code**
ÅLAND ISLANDS	AX
EUROPEAN UNION	EU
HONG KONG	HK
KOSOVO	KV
MACAO	MO
MONTENEGRO	ME
PALESTINIAN TERRITORY, OCCUPIED	PS
SAINT BARTHÉLEMY	BL
SAINT MARTIN	MF
TAIWAN, PROVINCE OF CHINA	TW
UNITED STATES MINOR OUTLYING ISLANDS	UM
Unknown	NA

A number of areas, mostly small islands, were included in the World_GIS shapefile and could be removed or merged because they were not recognized within the MMDB. The Gaza Strip was merged with the WEST BANK AND GAZA STRIP as Dr. Coles does not anticipate studying either the West Bank or the Gaza Strip as separate entites for some time. Howland I[sland] is an example of an uninhabited protectorate of the United States, so it was merged into the UNITED STATES record, harmonizing the records with the MMDB, but preserving the geographic feature, rather than simply deleting it. Table 4.5 shows the countries, areas or islands that were deleted or merged into other countries during the conversion from World_GIS to Countries.shp.

Table 4.5 - Countries Deleted or Merged

Countries Deleted or Merged
Countries.shp
Baker I.
Gaza Strip
Glorioso Is.
Howland I.
Jan Mayen
Jarvis I.
Johnston Atoll
Juan De Nova I.
Midway Is.
Wake I.

The cohort and teams used in the MMDB were coded into a single field – the job_team field. In many cases, the code could easily be broken apart and the cohort and team would be identified. This could be used to filter the network and limit the results. However, varied or incomplete record-keeping by the mission organizations hampered this effort. There were 16 different codes in use within the MMDB – 15 codes plus no code. On all missions, the Core team had no code because all Core members were grouped into one unit. The long term observers had no group from time to time, but were typically organized into teams with integer codes as their team number. Given that the EU EOM was the only organization to include STOs in cohorts, attempting to parse and filter the numbers proved to be outside the constraints of this project. Table 4.6 illustrates the various ways the LTO and STO teams were recorded by the missions.

Table 4.6 – MMDB Cohort and Team Codes

Sample Record Number	*Job*	*MMDB Code*	*Cohort Number*	*Team Number*
1	*STO*	*OCP 01*	*N/A*	*OCP 01*
2	*LTO*	*OLP 01*	*N/A*	*OLP 01*
3	*LTO*	*N2*	*N/A*	*N2*
4	*LTO*	*LTO 14*	*N/A*	*14*
5	*LTO*	*C2*	*N/A*	*C2*
6	*STO*	*902*	*9*	*02*
7	*STO*	*9.2*	*9*	*2*
8	*STO*	*092*	*09*	*2*
9	*STO*	*9*	*N/A*	*9*
10	*LTO*	*9*	*N/A*	*9 (different mission from above)*
11	*LTO*	*33*	*N/A*	*33*
12	*STO*	*3-1*	*3*	*1*
13	*STO*	*2202*	*22*	*02*
14	*STO*	*11.1*	*11*	*1*
15	*STO*	*11.01*	*11*	*01*

Parsing the MMDB code usually yields the cohort and team number, but in cases like the sample record numbers six and eight (above) are issues where, without extensive coding, human interaction is needed to determine the actual breakdown of the code. The difference between 902 and 092 is evident when comparing the other codes on the respective missions. On mission 174, the EU EOM mission to Ghana, there are a number of assignments to code 901, then 902, *et cetera*. There are also 801s, 802s, and similar codes. Analyzing the pattern within the codes indicates that the first digit indicates the cohort, and the second and third digits indicate the team. Comparing mission 174 (EU EOM mission to Ghana) to mission 229 (EU EOM mission to El Salvador) reveals differences in their team codes. The comparisons must be done within the context of the overall mission. The lists of all codes used in these two missions are in Appendix B.

4.4.1 Other Data Files

The interior points file was used to provide XY coordinates for the Python script that depicts the network. These XY points served as nodes. The interior points file was derived from the modified IPUMSI country shapefile. The Feature to Point Properties tool (INSIDE) was used to obtain "centroids" of each country. However, the tool incorporates all areas of each country and does not weight those locations according to their area. This skewed some of the points to locations far away from their expected geographic location. For instance, the ArcGIS-calculated "centroid" for the United States was near the Gulf Coast of Texas. Lacking a technical explanation, it appears that the tool calculates the center of the polygon's bounding rectangle and then locates the point in the rectangle's center. As an example, the United States has protectorates and outlying islands in lower latitudes than the contiguous states. These small outliers pull the calculated point away from the visual center of the country's land mass. The point near the Gulf Coast was possibly calculated because of this.   If the (INSIDE) selection is checked, the point is moved to the nearest point inside the polygon. For these reasons, and for aesthetic reasons, some of the points were manually shifted nearer to the visual center of their respective countries.   Figure 4.10 shows the original position of the ArcGIS-calculated "centroid" near Houston, Texas and the modified location of the point closer towards the visual center of the country's land mass.

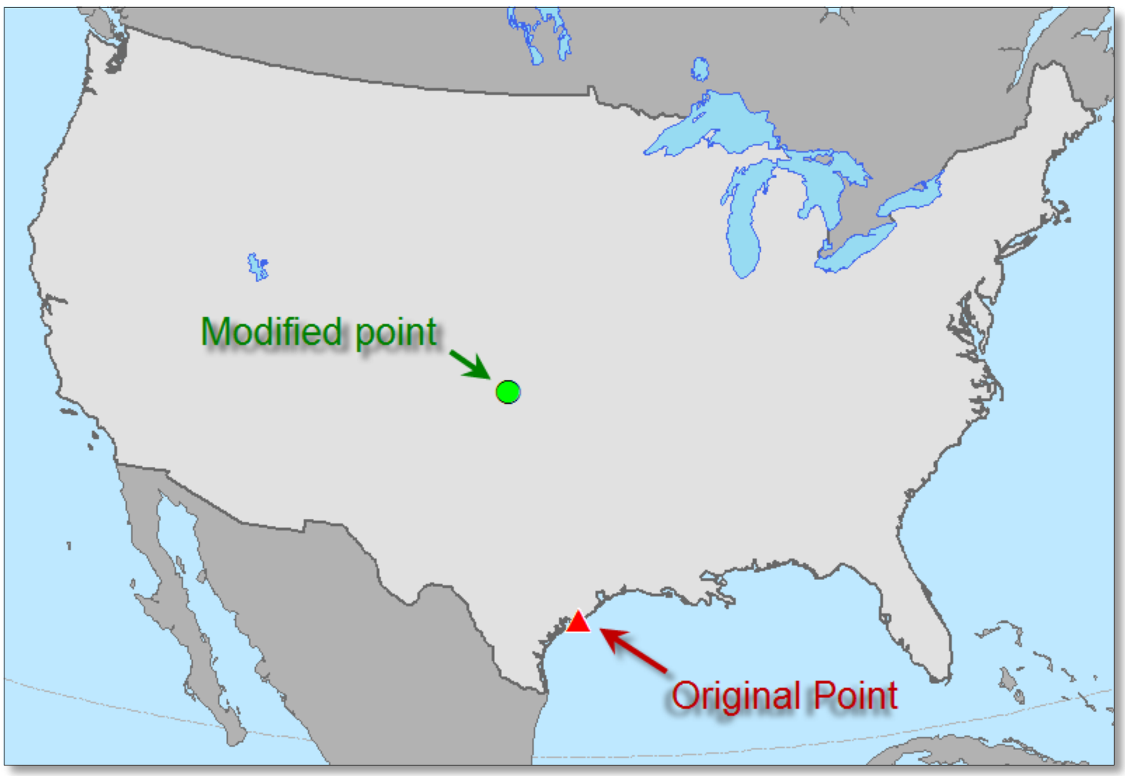

Figure 4.10 – Calculated versus Modified Interior Points

4.4.2 Data Frame Projection

A modified Times projection (itself a modification of the Gall's Stereographic projection) was selected for this project for a number of reasons (Figure 4.11). The first reason was that the project's geographic scope potentially includes the whole world, with the exception of Antarctica. The second reason for a modified Times projection was that it is a compromise projection and visually pleasing to many users. It is neither equal-area nor conformal, but neatly displays most of the world in a way that makes the output of this project more visible than most other projections. Finally, because of the Times' secant lines at 45° north and 45° south, eastern Europe was subject to less distortion than other similar projections, such as Miller Cylindrical. As many of the client's data are located in Eastern Europe, this was a benefit over the Miller Cylindrical (Figure 4.12).

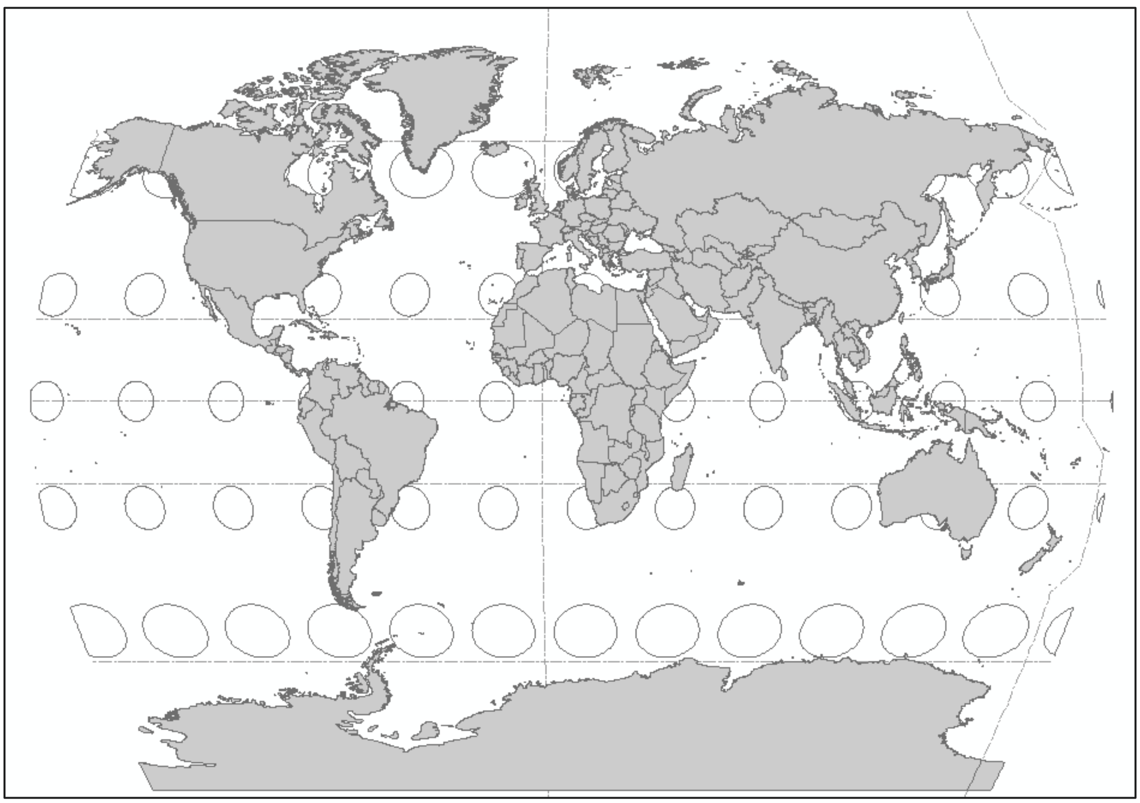

Figure 4.11 – Times Projection (with Tissot's Indicatrix)

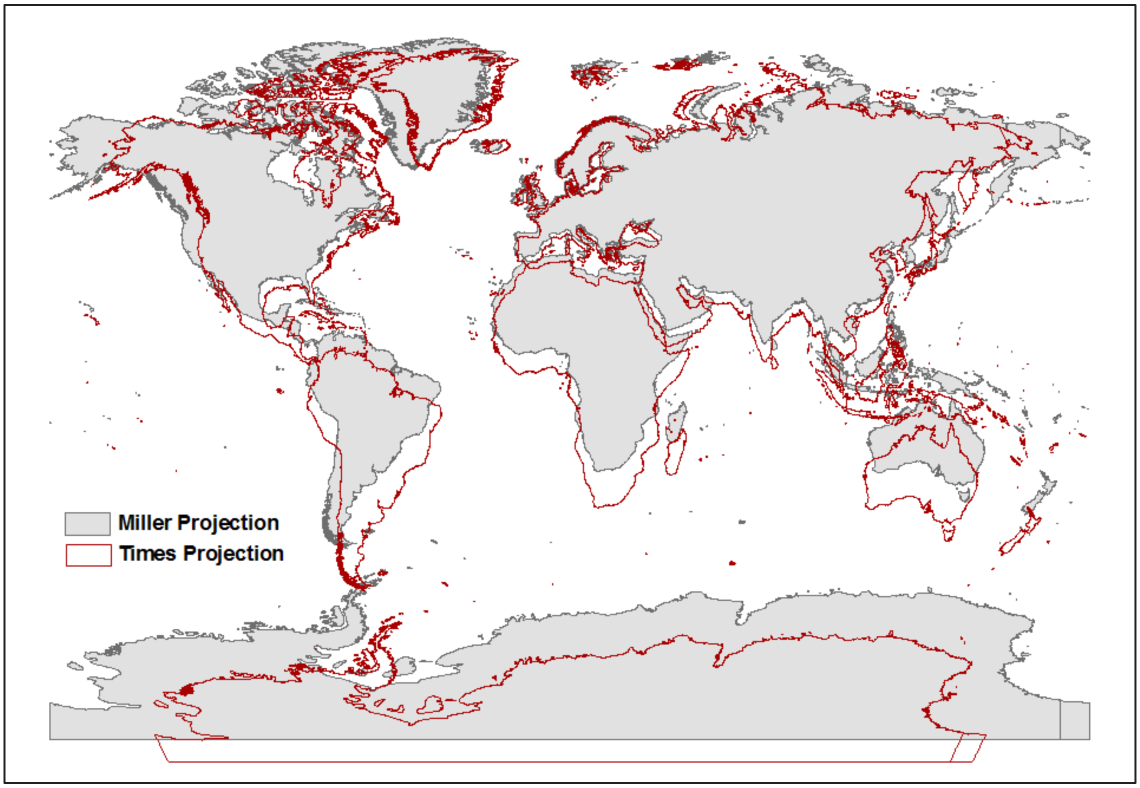

Figure 4.12 – Miller Cylindrical and Times Projections

The first modification to the Times projection was to eliminate the portions of the poles that were included in the projection. The second modification was to move the central meridian from 0° to 10° east so that all of eastern Russia would appear in the east half of the map.

The second modification to the data frame and its Times projection was to eliminate the southern polar region that is so distorted by this projection. Because the client did not have a current or future need for Antarctica, the data frame was clipped to exclude Antarctica. It was also clipped to eliminate all areas north of Svalbard and Jan Mayen islands. The area eliminated included portions of Greenland, Canada, and Russian islands (Figure 4.13).

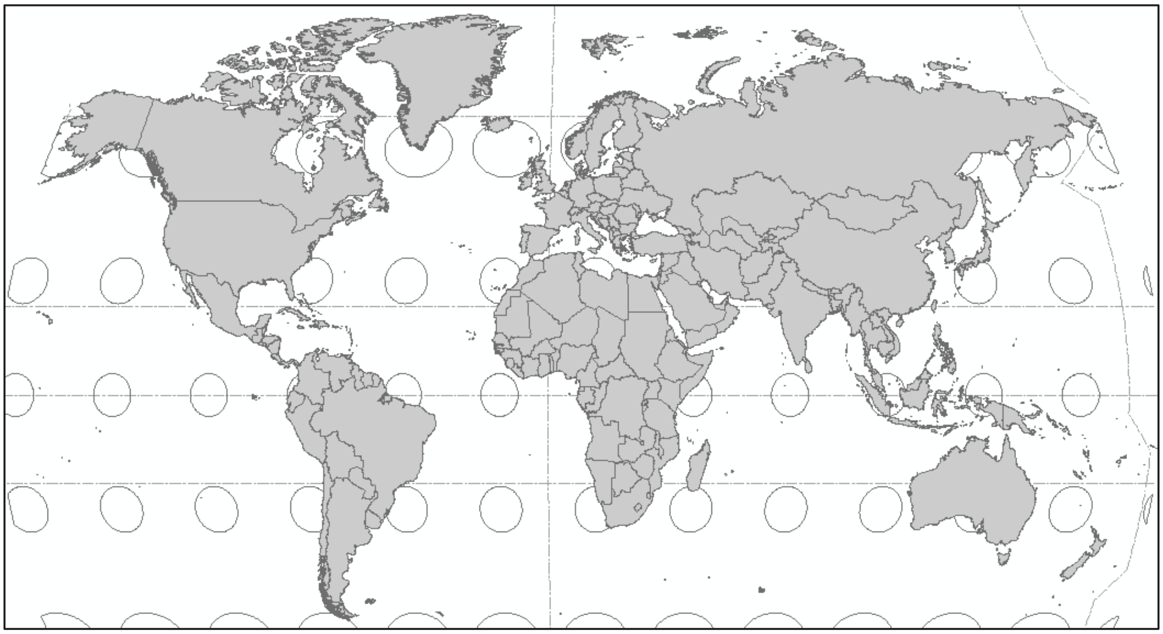

Figure 4.13 – Modified Times Projection

The perspective view of the edge of the map (such as Alaska) does distort the area of the map, but it was a reasonable trade-off for the benefits of the Times depiction of Europe. The manner in which data are used in this project means that Alaska is not used for any analysis; everything is located at the country level and the interior points used will fall in the contiguous United States. Some Southern Pacific islands are emphasized less in the Times projection than in the Miller Cylindrical, but again, the trade-off was acceptable.

4.5 Summary

Building the databases for the project to match the MMDB assured interoperability of the various components of the program. Some data had to be edited and modified to ensure they worked with the MMDB. Data could be collected from different sources to meet different requirements, so optional data such as the Map Services could be used once the functional requirements of the project had been met by using specific shapefiles and feature classes.

Because of the visualization component of the project, care was taken to select the Times projection for its worldwide scope and specific secant lines at 45° North and South. Modifying the projection further enhanced the aesthetics and usability of the project.

With the foundation of the databases set, the data loaded, and the parameters set, the analysis could begin.

Chapter 5 – Implementation

This chapter is divided into two parts – the implementation of the databases and the implementation of the ArcGIS portion of the project. Section 5.1 discusses the original Missions Management Database (MMDB) and the intermediate database (i_MMDB). Section 5.2 discusses the ArcGIS application. Portions of the integration of all the components are found in the first two sections, but detailed descriptions of the Pajek/NodeXL component of the project is found in 5.3.

5.1 Implementation of MS Access

An overall goal of the project was to integrate the overall structure without significant modification to the client's original database. When her career took her out of the country and the MMDB could not be accessed, it was essential to create a solution that could be implemented remotely. Furthermore, this solution ensured that updates to the data within the MMDB would be reflected in the final project without compromising the updates nor hindering the project. The Intermediate Database (i_MMDB) solved this problem.

5.1.1 Missions Management Database

The client provided the Missions Management Database version 9 for this project early in the planning and development stage. She retained the original database for her use during that time. While this project progressed, data were added to the original database, but no structural changes occurred. It was essential that the structure remained the same so that the i_MMDB interacted as planned upon implementation. The MMDB version 9 acted as a placeholder for the actual database. No modifications or data entry occurred to the MMDB version 9 in this project.

5.1.2 Intermediate Missions Management Database

The i_MMDB was created to act as a bridge between the dynamic MMDB and ArcGIS and the SNA programs. It is the stable link that allowed the client to depart the country prior to project completion but still allow modification to the database portion of the project. In essence, the i_MMDB could be pointed at any database similarly structured to the MMDB and still function.

The first step in building the i_MMDB was to copy the structure of the MMDB to ensure complete integration. An empty MS Access 2007 database schema was created without modification. Using the "Import>>Access >>Link to data…" command in Access 2007, all eleven tables in the MMDB were linked into the i_MMDB. The first benefit of this was that it preserved the structure of the tables. Secondly, it prevented unintentional modification or corruption of the data by the i_MMDB. Furthermore, it appeared and functioned just as importing the actual tables would have. Finally, it preserved the integrity of the data and their primary keys. Incidentally, all of the tables in the MMDB use AutoNumbers to create their ID / Primary Key field. Without the i_MMDB intermediary, any data added to the tables, even if deleted, during the creation of the project, would have conflicted with the data and their automatically-numbered primary keys added by the client.

Subsequent to linking to the MMDB tables, all queries, forms, reports and macros were imported from the MMDB into the i_MMDB using the Import>>Access command. At this point, the i_MMDB looked and functioned identically to the original MMDB, with the only difference being that the tables and their data were stored in the MMDB. A result of this is that any changes to the data in the MMDB are reflected in the i_MMDB, but the converse cannot occur, because no new data can be created in the i_MMDB without saving it to a new table within that database. At the same time though, the i_MMDB performed the same analysis on the data and simulated the MMDB exactly. There was no functional difference to the user at this point.

New tables did need to be added to the database. The first was the XY coordinates table. This table could have been created and given to the client for import into her original MMDB, but because it is not essential to the original intent of the MMDB, it could be saved in the i_MMDB.

The XY coordinate table was created from the interior points file discussed in section 4.4.1. The table was normalized so that the only fields necessary were the cntry_ID (the primary key), and the Point_X and Point_Y fields. During use, the table is joined to the linked country_tbl table when necessary. The process for creating this table was somewhat convoluted because the interior points shapefile did not use the same FID/primary key as the country_tbl. The MMDB country_tbl structure had three fields: country_id (an AutoNumber field and primary key), country_name, and country_ISO_3166. The interior points shapefile was modified to match the country_name and country_ISO_3166, but already contained the FID field. As discussed in section 4.4.1, the names and ISO codes in the shapefile were modified to match the MMDB records exactly. Once this was done, adding the country_id field from the MMDB to the shapefile could take place. It was this file that was exported from ArcGIS as a dBase table, with the FID and Shape fields being deleted. The dBase table was then imported into the i_MMDB and converted to Access format and named countryXY_Times_tbl. The country_name and ISO code fields remained in this new table while it was checked against the country_tbl for accuracy. When this was complete, the name and ISO code fields were deleted from the countryXY_Times_tbl, leaving the cntry_ID, Point_X, and Point_Y fields. (See Figure 5.1)

Figure 5.1 – Creating new XY table for i_MMDB

Once the countryXY_Times_tbl table was created, five sets of queries were created:

1. Find the extended network of a mission – that is, identify all of the participants on the mission, and all of their other missions, including all of the participants of those missions. This mission was limited to one iteration because the dataset returned on subsequent iterations was too vast for Access to handle stably.

2. Find the network of a mission, limited to prior missions. This query returns all of the participants of the selected missions, but limits those participants' other missions to only the missions that occurred prior to the selected mission.

3. Find the limited network of a mission. This query functions similarly to the extended network of a mission query, but on the second iteration, where participants are linked to their other missions, it only returns participants that were on the same team or in the same cohort as the participants of the selected mission. This drastically limits the network and improves performance. It is also a more realistic network than the extended network.

4. Find the extended network of a person. This query returns the extended network of all people that the selected person has ever worked with on the same mission. This was also limited in scope due to the quick growth of the dataset in subsequent iterations.

5. Find the limited network of a person. Similar to the limited network of a mission, this query filters the individuals returned to only those that had close contact with the selected participant, and documented through their work on the same team or cohort of a mission.

The first set of queries included the queries for selecting data from missions. Using them, the user would select a mission and the queries would return the participants, their roles within the mission, and their country of citizenship. Further iterations identified those participants' other missions and the participants of those other missions.

A form was created to control the filters and actions for the queries. This was the Network_and_GIS_Form (Figure 5.2). This form needed to be created because the switchboard used in the MMDB was imported with the tables, and functions in conjunction with them. Thus, it could not be modified and still be used within the original MMDB. The Network and GIS form operates independently of the MMDB and within the i_MMDB.

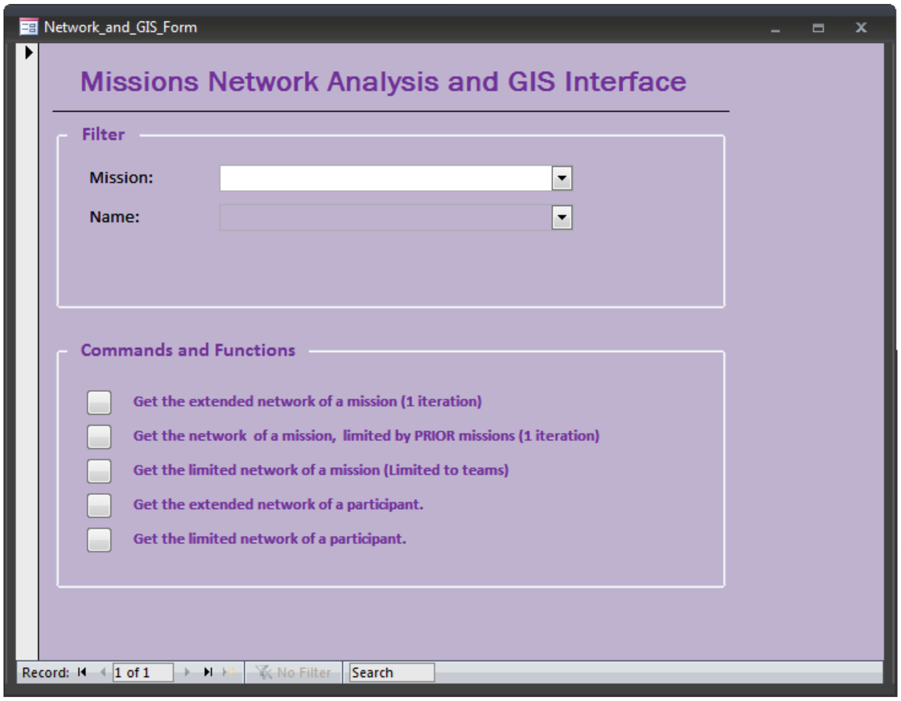

Figure 5.2 – Network_And_GIS_Form

As the queries were built and run, various errors were encountered on subsequent iterations of the queries. Access could not successfully complete a filtered query after the second iteration of querying. This will be discussed as “Results” in Chapter 6, but is mentioned here because it necessitated the creation of more tables in the i_MMDB to function properly. The original scheme for data flow looked like Figure 5.3, where one query would query the results of the prior query, which was the result of the initial query. In effect, the queries were querying the tables of each prior select query. The reason for the failures was unknown, but a solution was devised.

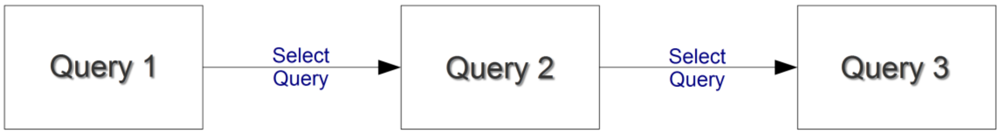

Figure 5.3 – Original data flow

That solution turned out to be changing the type of query performed. Append queries were used instead of select queries. The results of the queries were then appended into empty tables that could be used for queries that followed. In effect, the data flowed from Query 1, which queried the original data tables and appended the results, to Table 1. Query 2 then queried Table 1 instead of the Query 1's results table. Query 2 appended to Table 2, which was queried by Query 3, and so on (Figure 5.4). This stabilized the program and greatly increased the speed at which it functioned.

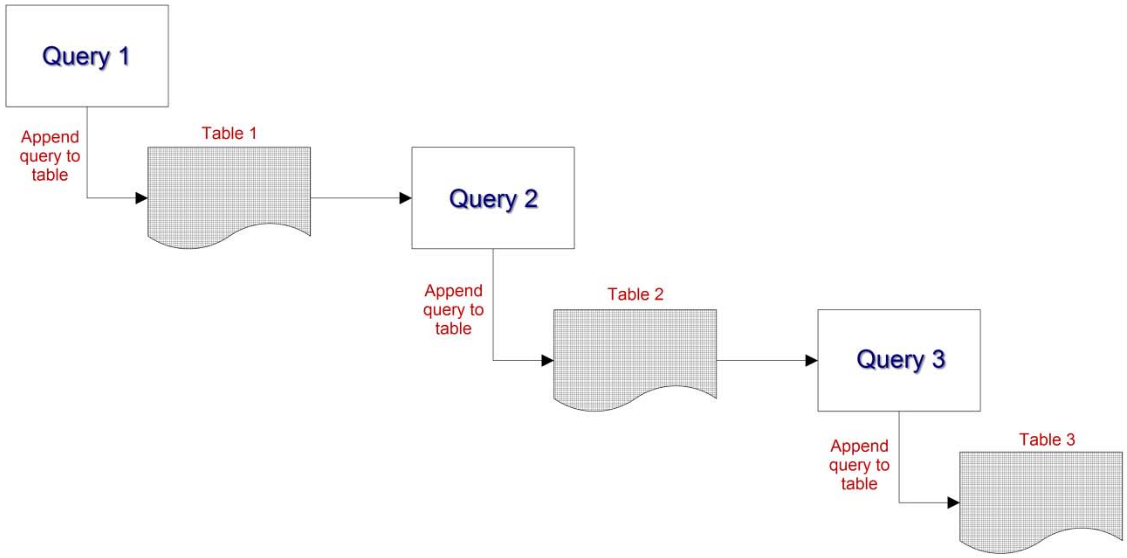

Figure 5.4 – Actual query data flow

To ensure all the tables were empty when appending data to them, delete queries were built into the beginning of each macro that called these append queries.

This turned out to be more beneficial than expected because the tables were available for access through ArcMap and the OLE connection. This enabled the user to display intermediate data results in ArcMap rather than only displaying the final result. It also greatly increased the speed at which results were returned from the queries.

The append queries were built into macros and run from the buttons on the Network_and_GIS_Form. Once the macros run, the tables are ready for use by ArcMap.

5.2 ArcGIS & ArcMap

ArcMap 9.3 is the main visualization tool for the project. It provides the spatial context and display of the networks. To do this, it uses ModelBuilder processes and Python scripts to manipulate and transform the data. ModelBuilder and Python scripts are primarily used to perform the display of the "QuickMaps" and Python scripts manipulate and display the network information.

5.2.1 QuickMaps Functions and Map Output

Certain functions could be quickly accomplished to display data pertaining to the missions but not within any networked context. This provided the client extra visualization of her data. These functions were grouped into a Toolset that was named "QuickMaps" under the custom-made "SNA Tools" toolbox (Figure 5.5). These QuickMaps are scripts that require little user input other than a few parameters and produce choropleth maps of the output.

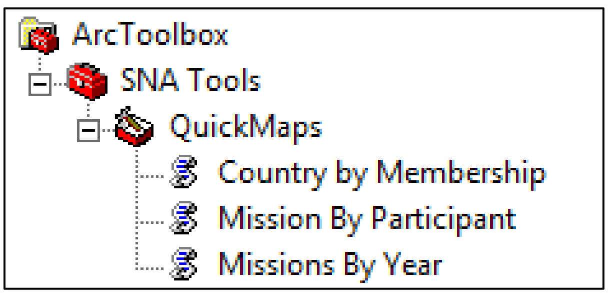

Figure 5.5 – QuickMaps Toolset

The three QuickMaps were generally created the same way in ModelBuilder and Python. They were built in ModelBuilder, exported as a Python script, and then modified to include the proper parameters necessary for each tool. The Missions by Year tool serves as a following example.

For each tool there were two concurrent paths built in ModelBuilder. One was for the QueryMissionByYear table in Access, and the other was for a .dbf table used as the output table. Using a Table Select function, with the year (or years) as the parameters for the expression, the QueryMissionByYear was filtered to include only the years in question. To ensure the .dbf table was empty and free of previous data, a Delete Rows command was run to empty it of all rows. The output of the QueryMissionByYear expression was then appended into the .dbf table. This table was then joined to a working copy of the countries shapefile and limited by un-checking the “Keep All” toggle. This returned only the countries that had missions in a particular year or range of years. An example of the Missions by Year model is found in Figure 5.6.

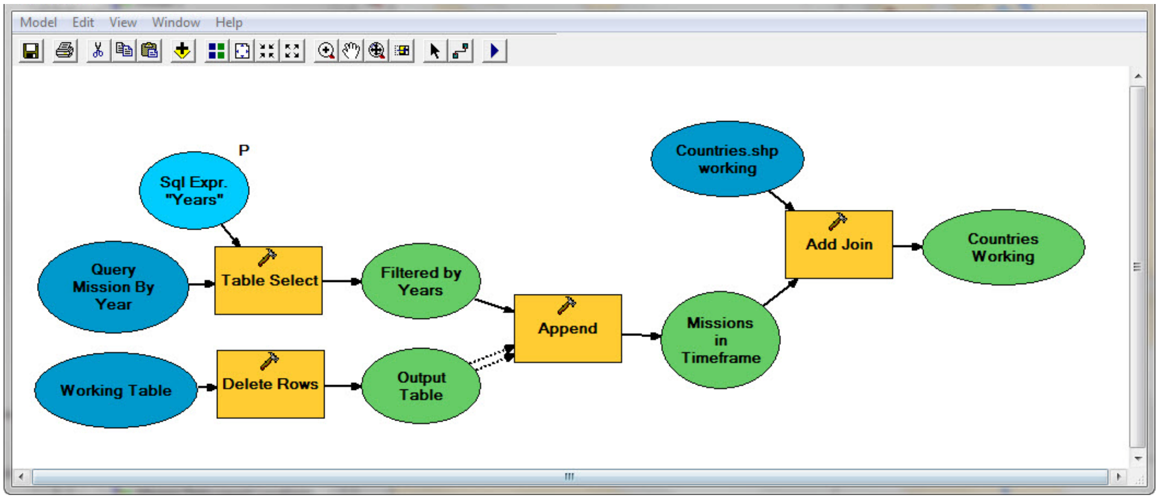

Figure 5.6 – Initial Model of the Missions by Year Tool

The preceding model was the first step in building the Missions by Year(s) QuickMap tool. In order to allow two input parameters, the model was exported to a Python script and modified to accept the two parameters. Specifically, this tool allows the user to input a range of years and returns the countries in which missions participated during that time frame. The user input form has two entry boxes: Start Year and End Year (Figure 5.8). The user can limit the range to one year by entering the same start and end year. In the case of a range of years, it includes all missions in the first and last years, as well as all those in between. Figure 5.7 depicts the output of this QuickMaps script. In this case, it shows countries that held elections between 1999 and 2003 (inclusive) and in which election monitors were present.

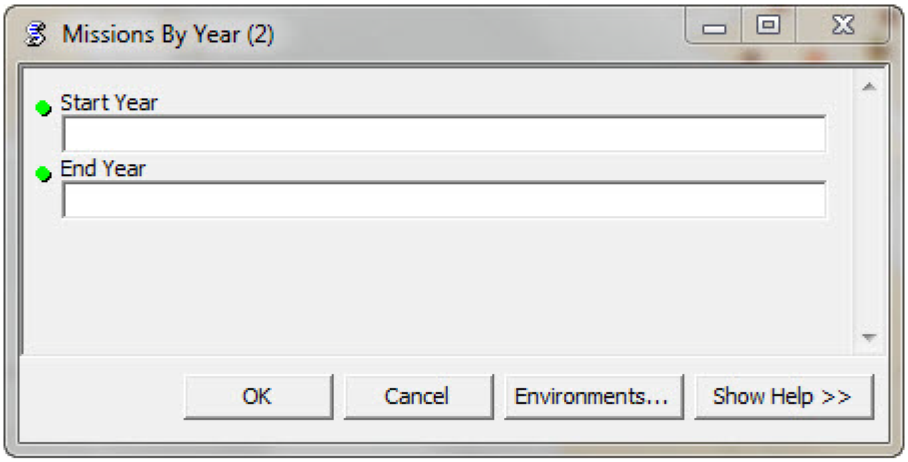

Figure 5.8 – Missions By Year(s) Entry Form

Figure 5.7 – Missions by Year, 1999 - 2003

The second QuickMap tool is Country by Membership. It was built in ModelBuilder, exported as a Python script, then modified to include the proper parameters. This script identifies and displays the countries belonging to a user-selected group. Dr. Coles defines the groups in her "Membership_tbl" table in the MMDB. It includes such groups as "Africa", "post-socialist", and "G8". These groups are then populated with country identification numbers in the Membership_tbl table. The values of the groups were hard-coded into the script. This is the only of the three QuickMaps scripts that has hard-coded inputs. The 'Groups' dropdown menu allows the user to select which group to display. The output of this script, however, is dependent on the data held in the MMDB. The full members of each group were not populated in the MMDB so results can be incomplete. Because the MMDB was not modified, this could not be corrected in this project, but can be correctly quickly with some data entry by the user of the original MMDB. The script on which this tool was built can be found in Appendix C. Figure 5.9 shows a depiction of the Country by Membership command menu as well as the output behind it. Ten countries were identified by Dr. Coles and recorded in the MMDB as belonging to the "post-socialist" group, and are returned by this script and displayed in red.

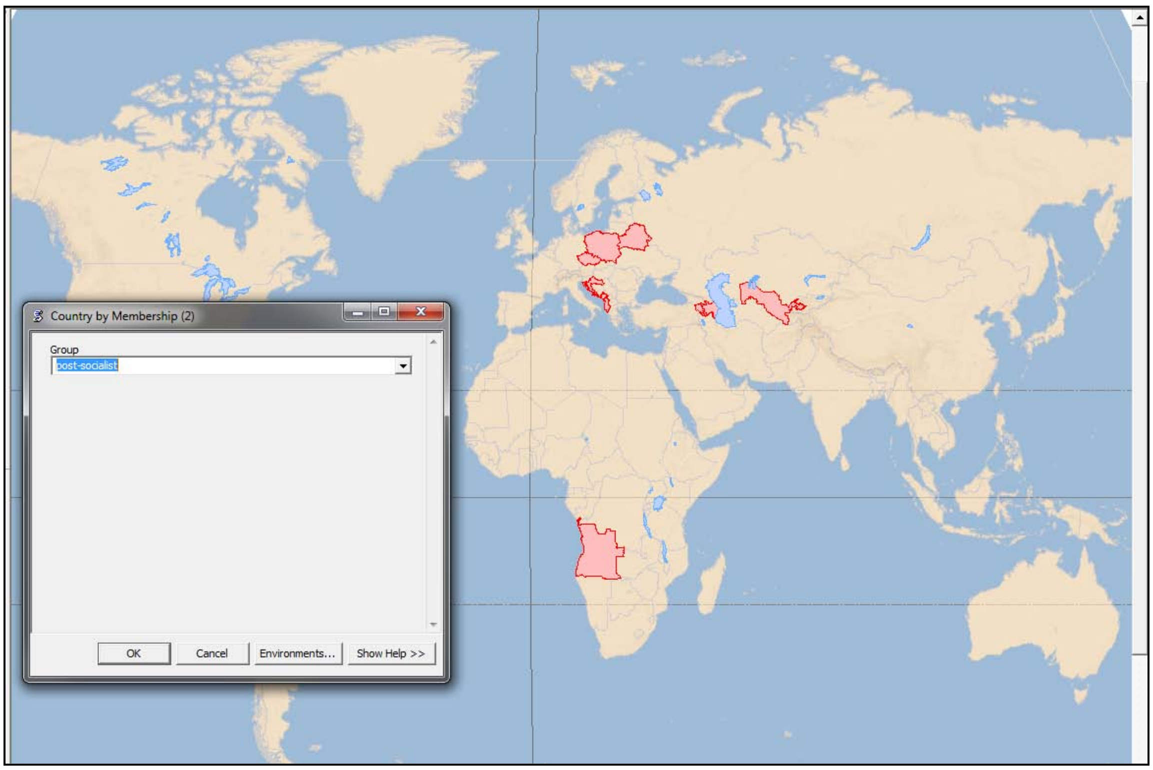

Figure 5.9 – Country by Membership Command and Output

The available values for this script's parameters were taken from the MMDB "Membership_tbl" table and hard-coded into the script using the properties of the script. The "Filter" property was set to "Value List" and the list of values were entered that way. This list can be added or changed in this screen (Figure 5.10). The user changes the membership group(s) of a country in the MMDB and the new values will transfer through the iMMDB and be reflected in the output when this script is run.

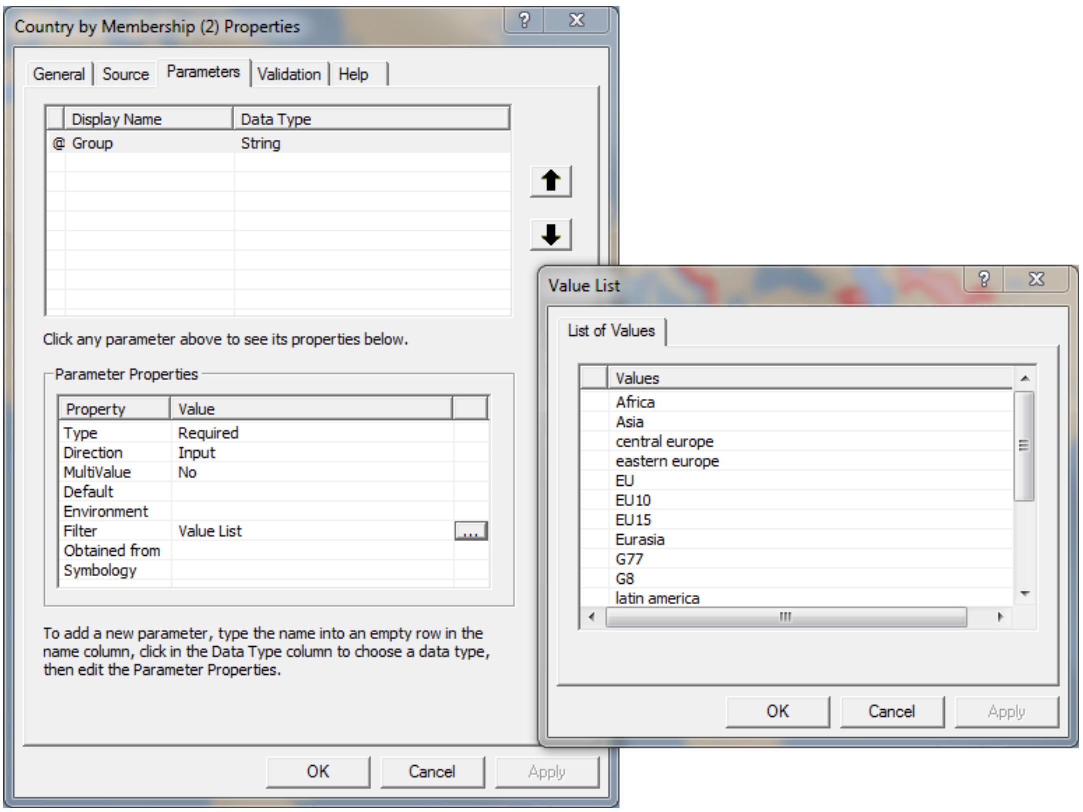

Figure 5.10 – Country by Membership Filter Property Value List

The third QuickMap is the Missions by Participant QuickMap. This tool takes the name of a participant (in "Last Name, First Name" format) and displays the person's country of citizenship and all countries they have served in on missions. Because it would have been too time consuming to type all of the participants into a Value List filter (such as in the Country by Membership QuickMap), a "Display Table" button calls the attribute table of the linked "qryEmpFullname" query of the MMDB. This table displays all 2094 employees' names in last name, first name format. The user has but to copy and paste from the table into the script to load the form. For the following example, the employee selected was Nicholas Aarons of the United Kingdom. He participated as a Long Term Observer in four Office for Democratic Institution and Human Rights (ODIHR) missions: Albania, Belarus, Ukraine, and the United States. In Figure 5.11, the four countries he worked in are returned in red, and his country of citizenship is returned in green.

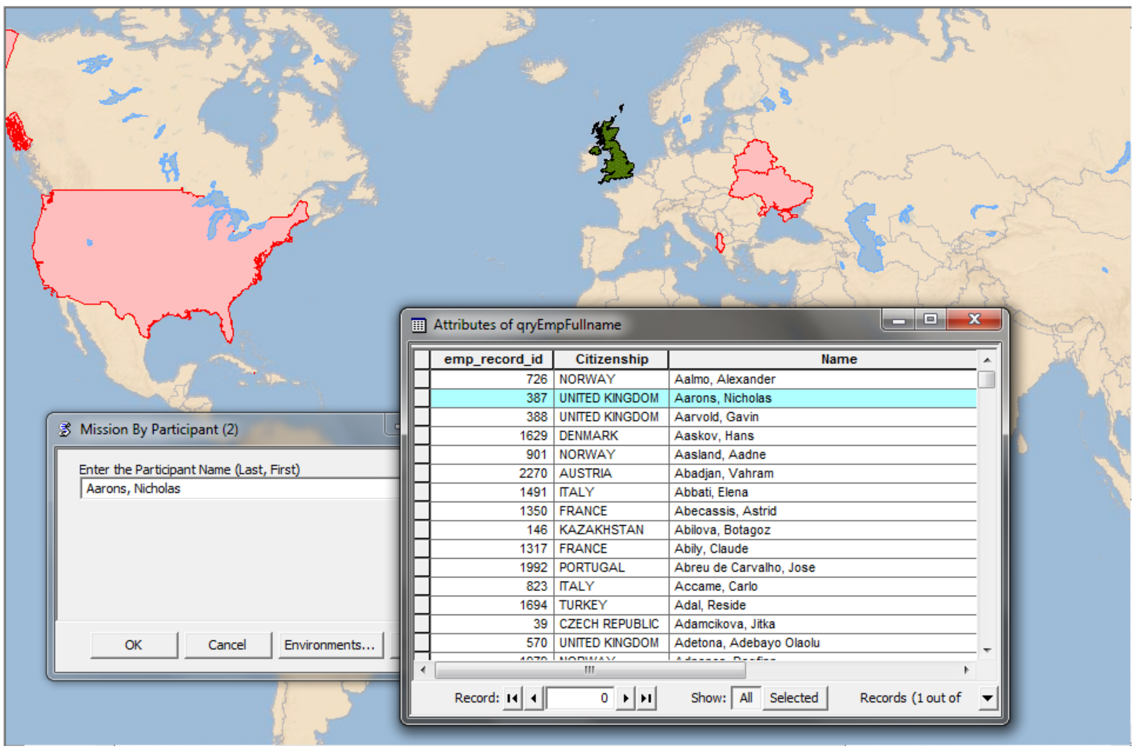

Figure 5.11 – Missions by Participant Results

5.2.2 Points to Lines Script and Network Display in ArcMap

The points to lines script is a Python script that was written to take the locations selected as network nodes in the i_MMDB and connect them via lines. In all cases for this project, those points are the interior point of a country. The method in which these points were selected was discussed in Section 5.1. It is sufficient to say here that multiple tables of events were created to select the network of interest and that this Python script makes the shapefile lines that are displayed.

The Python script written for this program was adapted from two sources. One was Nathan Strout, Technology Manager at the Redlands Institute. The other source was a script downloaded from the ArcGIS Resource Center. Elements of both these scripts were modified and then included in this custom script. The entire script can be found in Appendix C.

When using this script, the user must select the "from" and "to" coordinates – the origin of the line and the destination of the line. The "from" and "to" points were built as drop-down box selections when identifying the parameters of the script. The relevant field names in the shapefile are f_LonX, f_LatY, t_LonX, and t_LatY. The "f" indicates that the point should usually be considered the "from" (origin) point and the and the "t" indicates "to" (destination). The "Lat" and "Lon" designations were used because the preponderance of users are expected to be familiar with the terms "latitude" and "longitude" and thus recognize these points as coordinates, even though the actual data are in meters.

Within the Points to Lines user-interface dialogue box (Figure 5.12), the expected selections are: inTable – the table of points to be converted into lines; OriginX – the user should select the origin X coordinate, usually "f_LonX"; OriginY – the same point's Y coordinate column, usually "f_LatY"; DestinationX – usually "t_LonX"; DestinationY – usually "t_LatY", OutlineFC – the name and location where the output should be stored, and spref – the projected coordinate system in which the output is set.

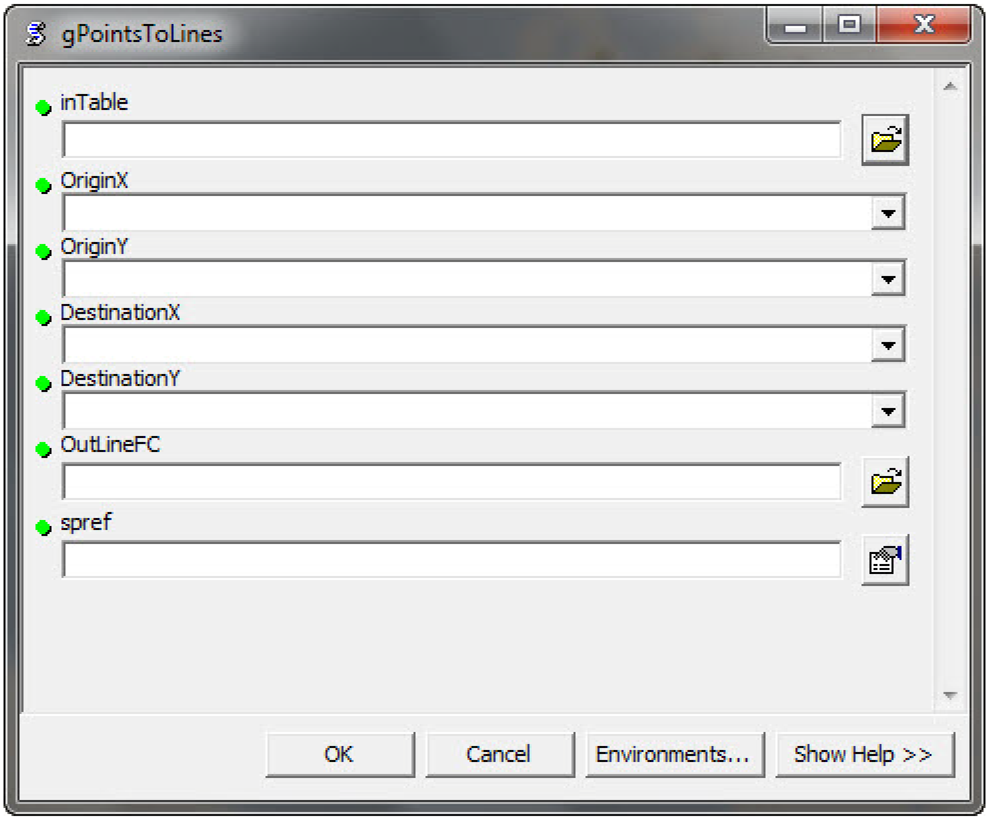

Figure 5.12 – Points to Lines Script Dialogue Box

Once the user selects the four fields to be the XY points for the “from” and “to” points, the output shapefile must be named and assigned a projection.

To ensure the proper fields are available for the new shapefile, a dummy empty shapefile was created to use as a template. It contains all the same fields and formats as the origin and destination shapefiles, but is not populated with any records. The points to lines script uses this dummy shapefile as a template for the fields when it creates the new shapefile. This is hard coded into the script. Figure 5.13 depicts the output of running the Points to Lines script on the EU EOM Mission to Angola in 2009.

Figure 5.13 – Points to Lines Script Output: EU EOM Mission to Angola, 2009

The points for this event are still held in the Access table and are available for display using the “Display XY Data” tool. This tool in ArcMap allows the user to select a table, chose the fields representing XY attributes, and then displays points at those locations. Using this tool to display the destination points of this event allows the user to symbolize the points. Symbolizing by category further enhances the usefulness of this tool. The participants are categorized into Core, LTO, and STO groups. The groups are then assigned different symbols – diamonds for Core, boxes for LTOs, and dots for STOs. Symbolizing the output assists in the visualization of the results but can also depict trending amongst participant assignments, if it exists. An example of this is found in Figure 5.14.

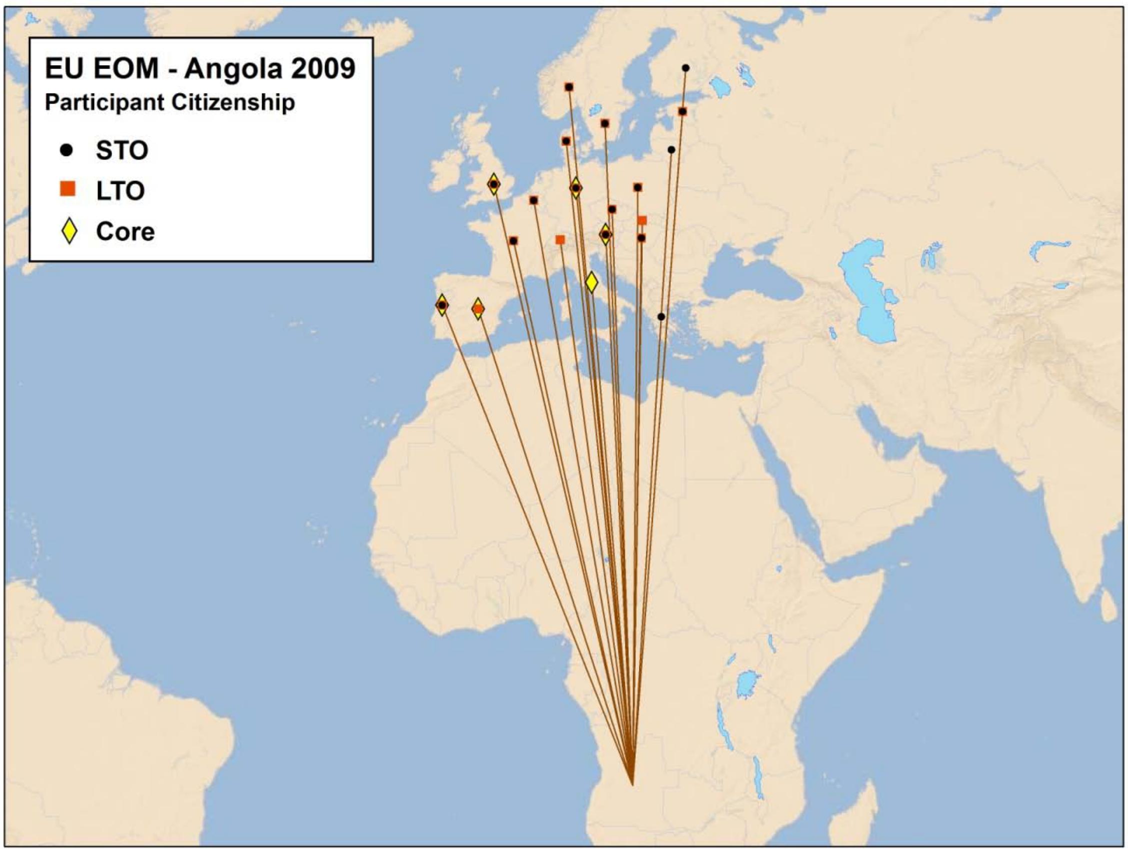

Figure 5.14 – Symbolized Output of a Mission

From the example in Figure 5.14, it is clear that the leadership of the election monitoring mission came from western Europe while the long term and short term observers are spread throughout Europe. Seeing these results can assist anthropologists in their study of this mission and its make-up.

Chapter 6 – Results

The Social Network Analysis with GIS program displays the social networks found in the Missions Management Database. It performs this function fairly well, but a number of problems hindered its development into a more robust program. This chapter discusses the results of the program development as well as the results of the analysis the program performs.

6.1 Results Within MS Access

The most significant results in the project were found in the MS Access portion. The first result was that using queries, filters, and other tools within Access, data could be taken from one Access database and manipulated in order to identify links in the data and to create a table with those links in another Access database. Running the tool sequentially on the output data returned datasets described as iterations.

Two types of queries were designed within Access to find the links. One query started with an individual election monitor and found his or her network, and the other started with an election monitoring mission to begin the analysis.

The first query, using an individual's associations as a starting point, finds all the missions in which that individual participated and then returns that list. Figure 6.1 illustrates that Nicholas Aarons, from the United Kingdom, held assignments on four missions– to Albania, Ukraine, Belarus, and the United States that were recorded in the MMDB.

f_ID	f_Nam	f_Cntry	t_ID	t_Nam	t_Cntry	t_Job
387	Aarons, Nicholas	UNITED KINGDOM	148	ODIHR-UNITED STATES-2008-11	UNITED STATES	LTO
387	Aarons, Nicholas	UNITED KINGDOM	28	ODIHR-BELARUS-2008-9	BELARUS	LTO
387	Aarons, Nicholas	UNITED KINGDOM	141	ODIHR-UKRAINE-2006-3	UKRAINE	LTO
387	Aarons, Nicholas	UNITED KINGDOM	8	ODIHR-ALBANIA-2005-7	ALBANIA	LTO

Figure 6.1 – Nicholas Aarons' Mission Participation (selected columns)

He was assigned as a Long Term Observer on all four missions. The second step in the process is to identify with whom he worked on these four missions. The process in Access uses the results in the first query to begin the second query. The query finds all the participants on the four missions and then filters that list to only include those election monitors who held the same job or were on the same team as the subject. It also identifies those participants' homes of record (Figure 6.2).

ex_Fld	f_Nam	f_Job	t_ID	t_Nam	t_Cntry	t_
Aarons, Nicholas	ODIHR-UNITED STATES-20	LTO	388	Aarvold, Gavin	UNITED KINGDOM	LTO
Aarons, Nicholas	ODIHR-BELARUS-2008-9	LTO	475	Amisulashvili, Valeri	GEORGIA	LTO
Aarons, Nicholas	ODIHR-UKRAINE-2006-3	LTO	1381	Arakelyan, Hayk	ARMENIA	LTO
Aarons, Nicholas	ODIHR-ALBANIA-2005-7	LTO	504	Aulas-Faure, Marie-Christi	FRANCE	LTO
Aarons, Nicholas	ODIHR-UNITED STATES-20	LTO	1380	Babic, Maria	SERBIA	LTO
Aarons, Nicholas	ODIHR-UNITED STATES-20	LTO	149	Balganova, Aida	KAZAKHSTAN	LTO
Aarons, Nicholas	ODIHR-ALBANIA-2005-7	LTO	1676	Ball, Nicholas	UNITED KINGDOM	LTO
Aarons, Nicholas	ODIHR-UNITED STATES-20	LTO	386	Bedritsky, Alexandr	RUSSIAN FEDERATION	LTO
Aarons, Nicholas	ODIHR-UNITED STATES-20	LTO	385	Bedzhanyan, Yury	RUSSIAN FEDERATION	LTO
Aarons, Nicholas	ODIHR-UNITED STATES-20	LTO	156	Beisenbayeva, Assem	KAZAKHSTAN	LTO
Aarons, Nicholas	ODIHR-ALBANIA-2005-7	LTO	1677	Bergman, Christina	SWEDEN	LTO
Aarons, Nicholas	ODIHR-UKRAINE-2006-3	LTO	1285	Best, Franziska	GERMANY	LTO
Aarons, Nicholas	ODIHR-BELARUS-2008-9	LTO	452	Birchler, Hans	SWITZERLAND	LTO
Aarons, Nicholas	ODIHR-BELARUS-2008-9	LTO	490	Bischoff, Paul Henri	SWITZERLAND	LTO
Aarons, Nicholas	ODIHR-UKRAINE-2006-3	LTO	1098	Blanck, Fredrik	SWEDEN	LTO
Aarons, Nicholas	ODIHR-BELARUS-2008-9	LTO	465	Bouchard, Andre	CANADA	LTO
Aarons, Nicholas	ODIHR-BELARUS-2008-9	LTO	487	Brandin, Roger	SWEDEN	LTO
Aarons, Nicholas	ODIHR-BELARUS-2008-9	LTO	491	Bryant, Roger	UNITED KINGDOM	LTO
Aarons, Nicholas	ODIHR-BELARUS-2008-9	LTO	470	Buure-Hagglund, Ritva-Kaa	FINLAND	LTO
Aarons, Nicholas	ODIHR-UNITED STATES-20	LTO	373	Buurman, Hendrik	GERMANY	LTO
Aarons, Nicholas	ODIHR-UKRAINE-2006-3	LTO	410	Cairns, Rosemary	CANADA	LTO
Aarons, Nicholas	ODIHR-UKRAINE-2006-3	LTO	903	Chamorro Benito, Andres	SPAIN	LTO
Aarons, Nicholas	ODIHR-UNITED STATES-20	LTO	389	Clayton, John	UNITED KINGDOM	LTO
Aarons, Nicholas	ODIHR-UKRAINE-2006-3	LTO	278	Condur, Georgeta	ROMANIA	LTO

Figure 6.2 – Nicholas Aarons' Associations Through Missions (selected columns)

The process can be repeated twice more in this fashion so that the user can find a wider network, but the network grows very rapidly.

The second method for creating the networks begins the query with a user-selected mission instead of an election monitor. The user selects the mission of interest from the Mission Selection combo box on the Network_and_GIS_Form in Access (see Figure 5.2 for example). The first time the query runs it returns all of the participants on the selected mission. In Figure 6.3, the European Union Election Observation Mission (EU EOM) to Afghanistan in September 2005 was selected. This query is potentially more expansive than the election monitor query because it involves querying all of the participants on the mission of interest rather than the missions of a single individual.

f_ID	f_Nam	f_Dat	t_ID	t_Nam	t_Cntry	t_Job	t_Tm	ltr
161	EU EOM-AFGHANISTAN-2005-9	9/18/2005	692	Freire, Martim	PORTUGAL	Core		1
161	EU EOM-AFGHANISTAN-2005-9	9/18/2005	245	Domanski, Jaroslaw	POLAND	LTO		1
161	EU EOM-AFGHANISTAN-2005-9	9/18/2005	220	Wojtan, Mariusz	POLAND	LTO		1

Figure 6.3 – EU EOM Mission to Afghanistan Recorded Participants

Three participants of this mission were recorded in the MMDB and returned by this query in the intermediate database. One member was assigned to the Core Team and two were assigned as Long Term Observers. The second step of this query returns all of the other assignments on which these election monitors have participated (Figure 6.4). The three participants of the 2005 EU EOM mission to Afghanistan participated on 31 other assignments as well. The assignments are also sorted in the “Status” column as to whether the mission took place before the initial, selected mission (in this case, EU EOM Afghanistan in September 2005) or after. This table also includes team information on the participants for filtering in the next step.

f_ID	f_Nam	t_ID	t_Nam	t_Job	t_Tm	ltr	Status
245	Domanski, Jaroslaw	27	ODIHR-BELARUS-2006-3	LTO		2	Post
245	Domanski, Jaroslaw	28	ODIHR-BELARUS-2008-9	Core		2	Post
245	Domanski, Jaroslaw	140	ODIHR-UKRAINE-2004-10	STO		2	Pre
245	Domanski, Jaroslaw	161	EU EOM-AFGHANISTAN-2005-9	LTO		2	Pre
245	Domanski, Jaroslaw	168	EU EOM-CAMBODIA-2008-7	STO	15.1	2	Post
245	Domanski, Jaroslaw	172	EU EOM-ECUADOR-2007-9	LTO	OLP 03	2	Post
245	Domanski, Jaroslaw	174	EU EOM-GHANA-2008-12	Core		2	Post
245	Domanski, Jaroslaw	194	EU EOM-BOLIVIA-2006-7	LTO		2	Post
245	Domanski, Jaroslaw	209	EU EOM-MAURITANIA-2006-11	STO		2	Post
245	Domanski, Jaroslaw	223	EU EOM-VENEZUELA-2005-12	STO		2	Post
692	Freire, Martim	161	EU EOM-AFGHANISTAN-2005-9	Core		2	Pre
692	Freire, Martim	173	EU EOM-ETHIOPIA-2005-5	Core		2	Pre
692	Freire, Martim	180	EU EOM-KENYA-2002-12	Core		2	Pre
692	Freire, Martim	188	EU EOM-PAKISTAN-2008-2	Core		2	Post
692	Freire, Martim	189	EU EOM-RWANDA-2008-9	Core		2	Post
692	Freire, Martim	194	EU EOM-BOLIVIA-2006-7	Core		2	Post
692	Freire, Martim	199	EU EOM-TIMOR-LESTE-2001-8	Core		2	Pre
692	Freire, Martim	211	EU EOM-NICARAGUA-2006-11	Core		2	Post
692	Freire, Martim	216	EU EOM-SIERRA LEONE-2007-8	Core		2	Post
692	Freire, Martim	222	EU EOM-VENEZUELA-2006-12	LTO		2	Post
692	Freire, Martim	223	EU EOM-VENEZUELA-2005-12	Core		2	Post
692	Freire, Martim	251	EU EOM-KENYA-2007-12	Core		2	Post
692	Freire, Martim	263	EU EOM-SIERRA LEONE-2002-5	STO	2	2	Pre
220	Wojtan, Mariusz	26	ODIHR-BELARUS-2004-10	LTO		2	Pre
220	Wojtan, Mariusz	151	ODIHR-UZBEKISTAN-2007-12	Core		2	Post
220	Wojtan, Mariusz	161	EU EOM-AFGHANISTAN-2005-9	LTO		2	Pre
220	Wojtan, Mariusz	173	EU EOM-ETHIOPIA-2005-5	STO		2	Pre
220	Wojtan, Mariusz	174	EU EOM-GHANA-2008-12	LTO	5	2	Post
220	Wojtan, Mariusz	186	EU EOM-NEPAL-2008-4	LTO	18	2	Post
220	Wojtan, Mariusz	188	EU EOM-PAKISTAN-2008-2	LTO		2	Post
220	Wojtan, Mariusz	206	EU EOM-INDONESIA-2006-12	LTO	01	2	Post
220	Wojtan, Mariusz	224	EU EOM-WEST BANK & GAZA ST	LTO		2	Post
220	Wojtan, Mariusz	225	EU EOM-YEMEN-2006-9	LTO		2	Post
220	Wojtan, Mariusz	269	EU EOM-LEBANON-2009-6	Core		2	Post

Figure 6.4 – All Other Assignments of the Three EU EOM Afghanistan 2005 Participants (selected columns)

The list in Figure 6.4 displays all of the missions and assignments on which the three participants of the Afghanistan example were members. The next step queries and returns all of the co-participants of the 34 assignments listed in Figure 6.4 according to their links to the original three participants. These links are established by finding the other election monitors who participated in the mission, their jobs, and their team information; they are linked by virtue of holding the same assignments. They are then only selected if the job and team fields (plus cohort, in the case of EU EOM STOs) match the original subjects (in this case, Domanski, Freire, and Wojtan). Often, no information is recorded in the team field so they are linked on the basis of having the same role (core, LTO, STO) only. Figure 6.5 is a selection of records from the 202 links returned. These

202 links are the individuals who were linked by virtue of holding the same assignments from the previous table of 34 assignments. In only 12 cases is there team information to filter the results further than just the Core, LTO, or STO levels. It can also be found that on the EU EOM mission to Sierra Leone, in May 2002, there were eight STOs on Team 2.

ex_Fld	f_ID	f_Nam	f_Job	f_Tm	t_ID	t_Nam	t_Cntry	t_Job	t_Tm	Itr
Domanski, Jaroslaw	172	EU EOM-ECUADOR-2007-9	LTO	OLP 03	654	Finelli, Liza	FRANCE	LTO	OLP 03	3
Wojtan, Mariusz	174	EU EOM-GHANA-2008-12	LTO	5	623	Leal, Rita S.	PORTUGAL	LTO	5	3
Freire, Martim	263	EU EOM-SIERRA LEONE-2002-5	STO	2	1850	Bjorklund, Lars	SWEDEN	STO	2	3
Freire, Martim	263	EU EOM-SIERRA LEONE-2002-5	STO	2	1848	Dal Borgo, Antonio	ITALY	STO	2	3
Freire, Martim	263	EU EOM-SIERRA LEONE-2002-5	STO	2	1847	Fourman, Michel	UNITED KINGDOM	STO	2	3
Freire, Martim	263	EU EOM-SIERRA LEONE-2002-5	STO	2	764	Geny, Emmanuel	FRANCE	STO	2	3
Freire, Martim	263	EU EOM-SIERRA LEONE-2002-5	STO	2	1295	Kuijper, Jelske	NETHERLANDS	STO	2	3
Freire, Martim	263	EU EOM-SIERRA LEONE-2002-5	STO	2	1259	Saraliotou, Despina	GREECE	STO	2	3
Freire, Martim	263	EU EOM-SIERRA LEONE-2002-5	STO	2	1849	Schumer, Tanja	UNITED KINGDOM	STO	2	3
Wojtan, Mariusz	186	EU EOM-NEPAL-2008-4	LTO	18	359	Ioannou, Dimitra	GREECE	LTO	18	3
Domanski, Jaroslaw	168	EU EOM-CAMBODIA-2008-7	STO	15.1	734	Klopp, Laurence	LUXEMBOURG	STO	15.1	3
Wojtan, Mariusz	206	EU EOM-INDONESIA-2006-12	LTO	01	1124	Marques, Tania	PORTUGAL	LTO	01	3
Domanski, Jaroslaw	194	EU EOM-BOLIVIA-2006-7	LTO		628	Afonso, Ana	PORTUGAL	LTO		3
Domanski, Jaroslaw	28	ODIHR-BELARUS-2008-9	Core		458	Ahrens, Geert-Hinrich	GERMANY	Core		3
Domanski, Jaroslaw	27	ODIHR-BELARUS-2006-3	LTO		412	Alken, Ib	DENMARK	LTO		3
Domanski, Jaroslaw	194	EU EOM-BOLIVIA-2006-7	LTO		1494	Ansede, Lola	SPAIN	LTO		3
Freire, Martim	251	EU EOM-KENYA-2007-12	Core		18	Apostolov, Simeon	BULGARIA	Core		3
Domanski, Jaroslaw	28	ODIHR-BELARUS-2008-9	Core		18	Apostolov, Simeon	BULGARIA	Core		3
Freire, Martim	211	EU EOM-NICARAGUA-2006-11	Core		751	Arranz, Miguel	BELGIUM	Core		3
Domanski, Jaroslaw	27	ODIHR-BELARUS-2006-3	LTO		1421	Ashworth, Stephen	UNITED KINGDOM	LTO		3
Domanski, Jaroslaw	27	ODIHR-BELARUS-2006-3	LTO		877	Backlund, Jorgen	SWEDEN	LTO		3
Freire, Martim	222	EU EOM-VENEZUELA-2006-12	LTO		603	Bangel, Jutta	GERMANY	LTO		3
Domanski, Jaroslaw	194	EU EOM-BOLIVIA-2006-7	LTO		603	Bangel, Jutta	GERMANY	LTO		3
Freire, Martim	188	EU EOM-PAKISTAN-2008-2	Core		691	Bernhard, Manfred	AUSTRIA	Core		3
Freire, Martim	222	EU EOM-VENEZUELA-2006-12	LTO		1219	Bessieres, Andre	FRANCE	LTO		3
Freire, Martim	189	EU EOM-RWANDA-2008-9	Core		573	Binder, Christina	AUSTRIA	Core		3
Freire, Martim	222	EU EOM-VENEZUELA-2006-12	LTO		1225	Blake, Donald P.	IRELAND	LTO		3
Freire, Martim	216	EU EOM-SIERRA LEONE-2007-8	Core		1098	Blanck, Fredrik	SWEDEN	Core		3
Wojtan, Mariusz	269	EU EOM-LEBANON-2009-6	Core		1098	Blanck, Fredrik	SWEDEN	Core		3
Freire, Martim	211	EU EOM-NICARAGUA-2006-11	Core		41	Boldi, Milan	CZECH REPUBLIC	Core		3
Freire, Martim	222	EU EOM-VENEZUELA-2006-12	LTO		41	Boldi, Milan	CZECH REPUBLIC	LTO		3
Freire, Martim	194	EU EOM-BOLIVIA-2006-7	Core		41	Boldi, Milan	CZECH REPUBLIC	Core		3
Freire, Martim	222	EU EOM-VENEZUELA-2006-12	LTO		1796	Borello, Rossana	ITALY	LTO		3
Freire, Martim	222	EU EOM-VENEZUELA-2006-12	LTO		1476	Bosco, Caroline	BELGIUM	LTO		3
Freire, Martim	222	EU EOM-VENEZUELA-2006-12	LTO		1220	Bouyssou, Benoit	FRANCE	LTO		3
Domanski, Jaroslaw	194	EU EOM-BOLIVIA-2006-7	LTO		487	Brandin, Roger	SWEDEN	LTO		3
Domanski, Jaroslaw	27	ODIHR-BELARUS-2006-3	LTO		491	Bryant, Roger	UNITED KINGDOM	LTO		3
Freire, Martim	222	EU EOM-VENEZUELA-2006-12	LTO		1538	Butler, Georgiana Susa	UNITED KINGDOM	LTO		3
Freire, Martim	222	EU EOM-VENEZUELA-2006-12	LTO		1165	Caligiuri, Giovanni	ITALY	LTO		3
Freire, Martim	189	EU EOM-RWANDA-2008-9	Core		2107	Cashman, Michael	UNITED KINGDOM	Core		3
Freire, Martim	189	EU EOM-RWANDA-2008-9	Core		530	Castanias, Alexandre	GREECE	Core		3
Domanski, Jaroslaw	194	EU EOM-BOLIVIA-2006-7	LTO		682	Castillo Blanco, Hermi	SPAIN	LTO		3
Freire, Martim	251	EU EOM-KENYA-2007-12	Core		1023	Chatenet, Scipion du	FRANCE	Core		3
Freire, Martim	216	EU EOM-SIERRA LEONE-2007-8	Core		1023	Chatenet, Scipion du	FRANCE	Core		3
Freire, Martim	180	EU EOM-KENYA-2002-12	Core		1028	Clayton, Peter Eduard	UNITED KINGDOM	Core		3
Freire, Martim	211	EU EOM-NICARAGUA-2006-11	Core		1768	Cobos Flores, Francisc	SPAIN	Core		3
Wojtan, Mariusz	151	ODIHR-UZBEKISTAN-2007-12	Core		830	Colombi, Alice	ITALY	Core		3
Freire, Martim	211	EU EOM-NICARAGUA-2006-11	Core		690	Cox, Rebecca	UNITED KINGDOM	Core		3
Wojtan, Mariusz	269	EU EOM-LEBANON-2009-6	Core		1760	de Gabriel Perez, Jose	SPAIN	Core		3
Freire, Martim	251	EU EOM-KENYA-2007-12	Core		1020	De Herdt, Vincent	BELGIUM	Core		3
Freire, Martim	251	EU EOM-KENYA-2007-12	Core		1021	Demauge, Aurelie	FRANCE	Core		3
Freire, Martim	222	EU EOM-VENEZUELA-2006-12	LTO		1778	Diaz Rojas, Javier Alexi	SPAIN	LTO		3
Wojtan, Mariusz	151	ODIHR-UZBEKISTAN-2007-12	Core		25	Dimitrov, Dimitar	BULGARIA	Core		3
Wojtan, Mariusz	151	ODIHR-UZBEKISTAN-2007-12	Core		500	Dokter, Kerstin	GERMANY	Core		3
Wojtan, Mariusz	161	EU EOM-AFGHANISTAN-2005-9	LTO		245	Domanski, Jaroslaw M	POLAND	LTO		3
Freire, Martim	189	EU EOM-RWANDA-2008-9	Core		754	Eick, Mathias	GERMANY	Core		3
Freire, Martim	188	EU EOM-PAKISTAN-2008-2	Core		754	Eick, Mathias	GERMANY	Core		3
Freire, Martim	211	EU EOM-NICARAGUA-2006-11	Core		1766	Fava, Claudio	ITALY	Core		3
Freire, Martim	222	EU EOM-VENEZUELA-2006-12	LTO		654	Finelli, Liza	FRANCE	LTO		3
Domanski, Jaroslaw	194	EU EOM-BOLIVIA-2006-7	LTO		654	Finelli, Liza	FRANCE	LTO		3
Domanski, Jaroslaw	194	EU EOM-BOLIVIA-2006-7	LTO		1554	Flores, Sandra	FRANCE	LTO		3
[illegible]	[illegible]	[illegible]	[illegible]		[illegible]	[illegible]	[illegible]	[illegible]		[illegible]

Figure 6.5 – Teammates of Afghanistan Mission Participants' Other Missions (selected rows and columns)

The next step in the networking process involves finding all of the other assignments the election monitors found in the previous step have held. The result is not depicted here, but there are 715 links. Thus, from a single mission (Afghanistan 2005) there are three participants who have held 31 other assignments. From those assignments, the program returns 202 direct links through job and team associations. The next iteration yielded 715 people, the number of other assignments that those 202 EMs have held. The final available step is to find the number of election monitors that have been associated with those 715 assignments. At this point, the network gets inexplicably large at 1,993,942 links. This means that Access has computed that there are 2,788 associations with each of the 715 assignments. This is clearly erroneous and it is not known why this occurs, but it is believed to be an error in Access' handling of the data.

Another significant error in Access arose early in the program building process. Access is unable to perform a three-iteration select query. This problem of the multiple-iteration select query was introduced in section 5.1.2. The original intent of the program was that it be built as a cascade of select queries, where one query would filter and build on the results of previous queries. On the third running of a query, Access would freeze, report an error, or crash entirely. It was unclear what was causing this error, but it was avoided by using append queries rather than select queries and thus appending the data to established tables. One explanation may be that the errors were the result of the tables created by the select queries had neither unique IDs nor indices. The size of the third select query did not matter to the result of the queries – queries expected to return a small number or a large number of results failed the same way until the append queries were used.

The need to create tables to accommodate the append queries changed the structure of the database slightly. The tables were saved in the i_MMDB because they were not necessary to the original functionality of the MMDB.

Two significant notes must be made here, however. The first is that when using the append queries, Access successfully calculated all of the iterations before producing erroneous results (1.9 million links) at the fifth iteration. Iterating this one step further produced over 50 million results, and at this point Access reached its two gigabyte limit and quit processing. This is a significant improvement over Access being unable to make three consecutive select queries of any size. Although it is calculating data the user cannot use on the final step, the fact that Access runs without crashing or freezing demonstrates that using append queries in lieu of select queries is a more reliable method of querying very large datasets.

The second note is that this reinforces the concept that network queries must be targeted and filtered in the early iterations and limited to no more than a few iterations. Furthermore, additional analysis should be performed to ensure the validity of the results.

Other results that came from Access were that not all select query tables were accessible to ArcGIS through the Jet 12 OLE connection. Queries using parameters or queries that relied on other dynamic links were not available to ArcGIS for use in any scripts or displays. Fortunately, this did not significantly hinder the overall project.

6.2 Results in ArcGIS

Many of the results of analysis using the MMDB and i_MMDB became apparent when visualized in ArcMap. Other results were only observed in ArcMap and influenced the design of the program.

The first results observed in ArcMap were the QuickMaps. ModelBuilder models and Python scripts were built that allow the user to select a few input parameters and produce a choropleth map. Three QuickMap tools were built: Missions By Year(s), Missions by Participant, and Countries by Organization. Details about the three tools are found in Section 5, but selected examples are presented here.

The Missions by Year(s) QuickMap allows user-entered parameters to select the year or range of years to query. The script then uses a select by attributes SQL query to return a table of the countries in which elections were monitored during the entered timeframe. This table is joined to a countries layer and displayed (Figure 6.6).

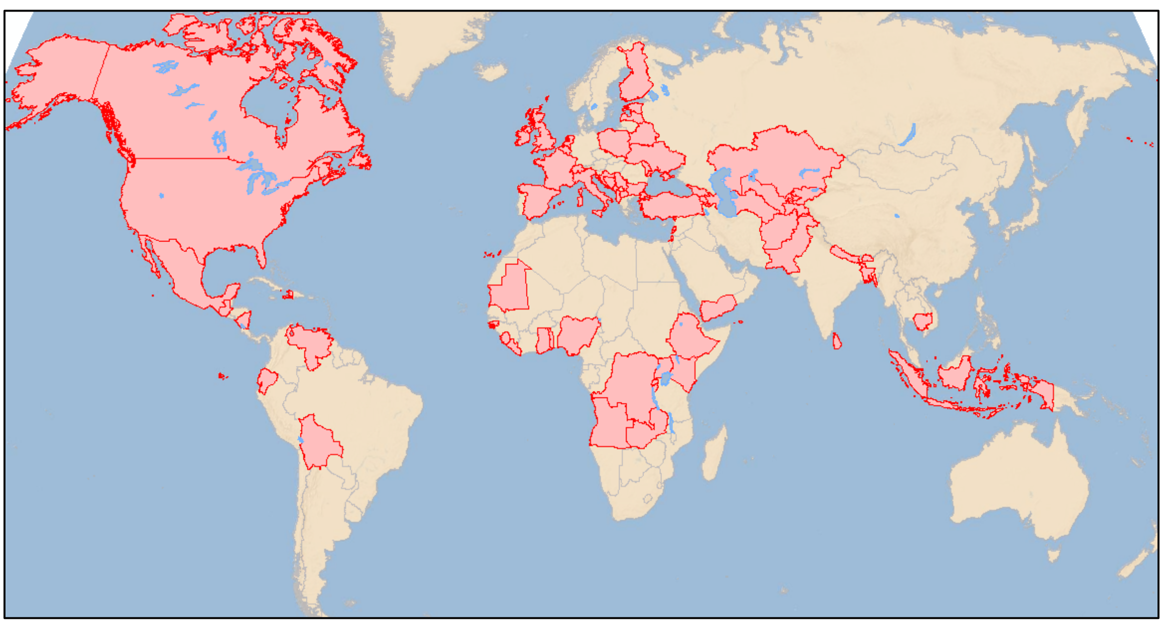

Figure 6.6 – Missions by Years Output (2005 – 2008)

The second QuickMap produces a map that identifies a selected election monitor's home of record and the countries in which he or she has monitored elections. Manually changing the symbology to reflect categories displays the different assignments the observer has held in the missions. For instance, Monica Reintjes participated in 12 missions to 11 countries and was a citizen of the Netherlands. Her home of record was symbolized with a red fill, the countries in which she was an LTO are symbolized in green and her STO positions are shown in orange (Figure 6.7).

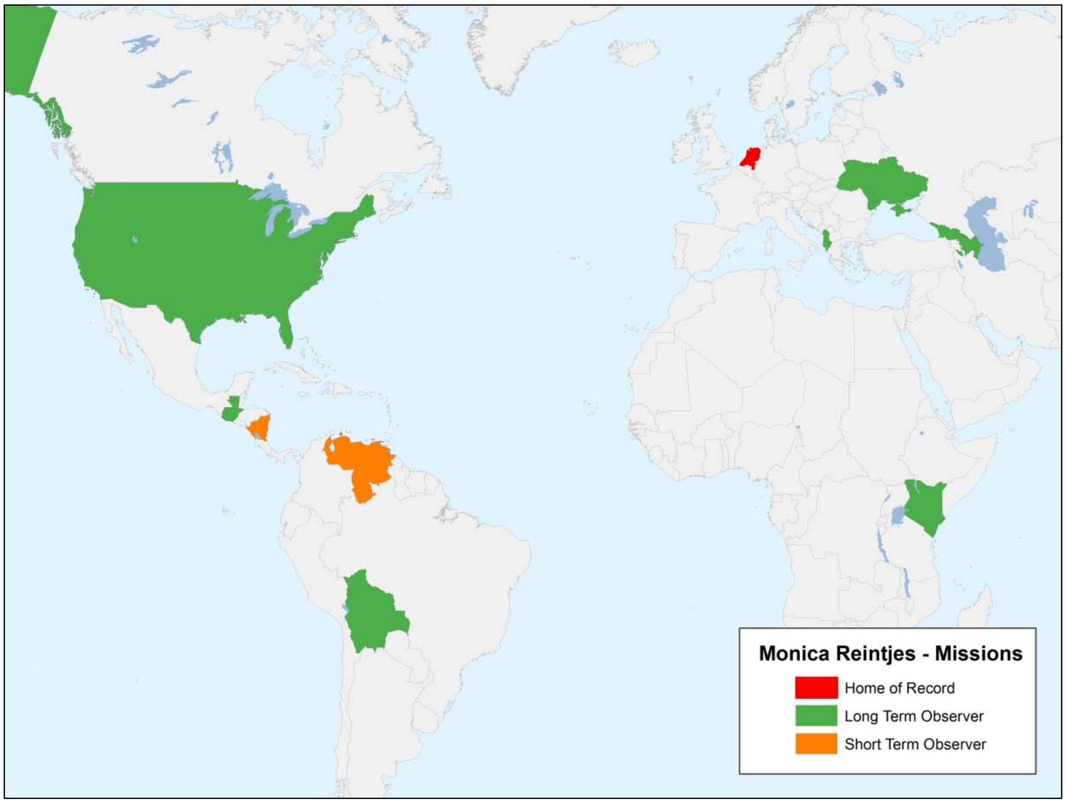

Figure 6.7 – Missions by Employee Output with Manual Symbology

The final QuickMap product is the Countries by Membership tool. This tool functions correctly but was limited because the membership dataset within the MMDB was not complete. The intent for the membership table was a list of countries that belonged to any particular grouping the user may chose – whether the countries within an NGO or a loose grouping such as "post-socialist" countries (Figure 6.8).

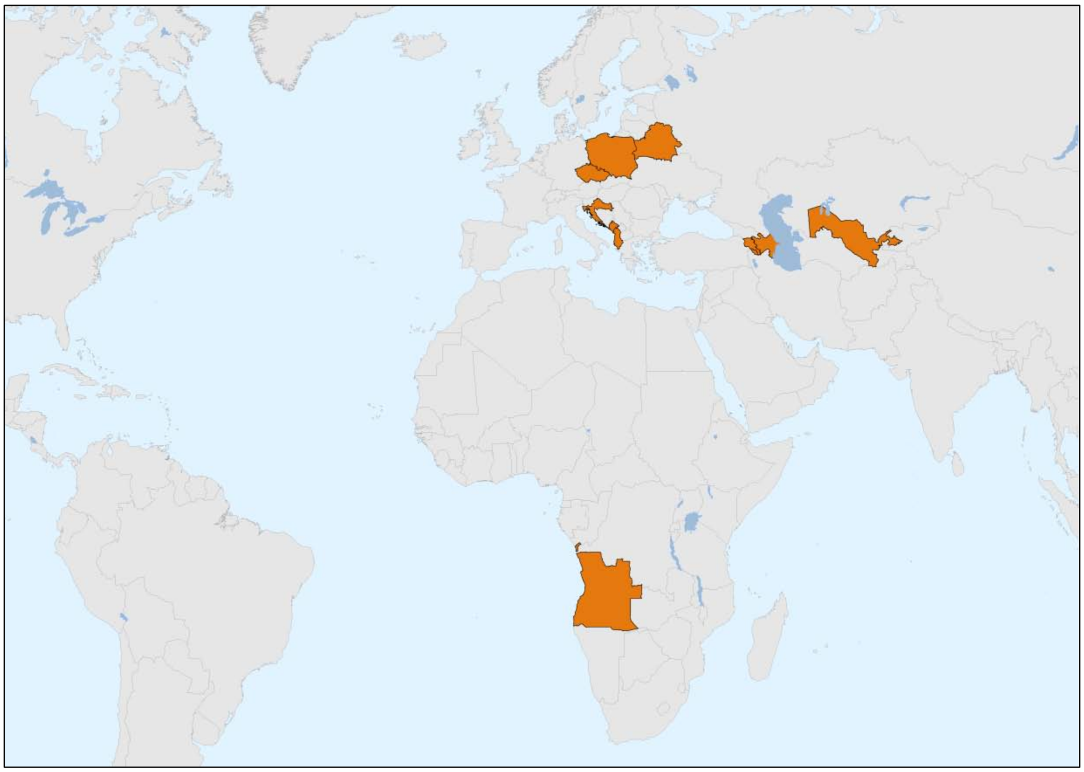

Figure 6.8 – Countries by Membership Output - Post-socialist Group (detail)

The majority of the network visualization was performed using the custom Points to Lines tool. It uses coordinates stored in the i_MMDB to find the start and end nodes for each line and create a line shapefile between them. That was followed by using the "Display XY Data…" tool to symbolize the points according to their role – either election event, home of record, or job (Figure 6.9).

Figure 6.9 – First Links of the EU EOM Mission to Afghanistan, 2005

Following the identification of all participants of the mission in interest came the second level analysis. In this case, the three participants on the 2005 mission to Afghanistan held 31 other assignments. Their links and their jobs are depicted in Figure 6.10. These links and jobs were again manually symbolized to reflect the different positions the participants held in their different missions and to make the visualization easier.

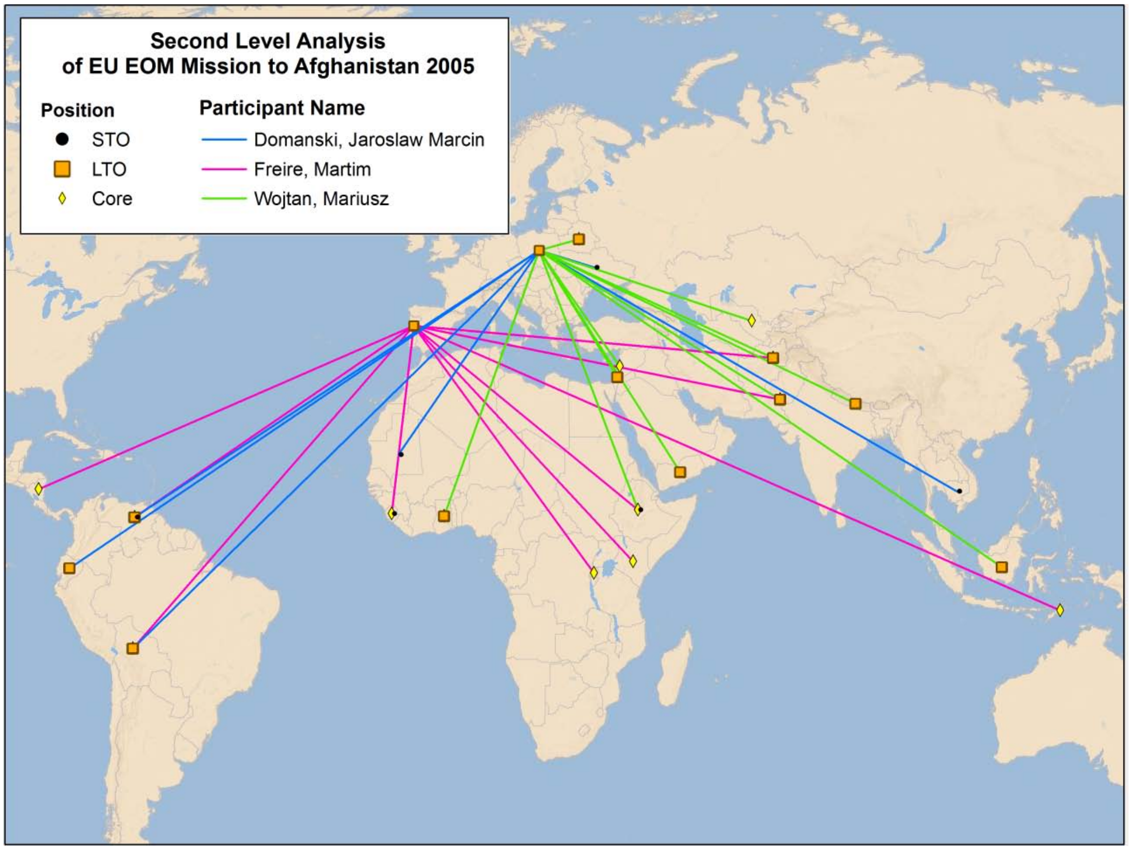

Figure 6.10 – Other Missions of the Three Afghanistan 2005 Participants (detail)

The next step was to find the participants of all the other missions on which the three original participants were assigned. A partial list of these participants is in Figure 6.5. The visualization becomes much more cluttered at this level (Figure 6.11).

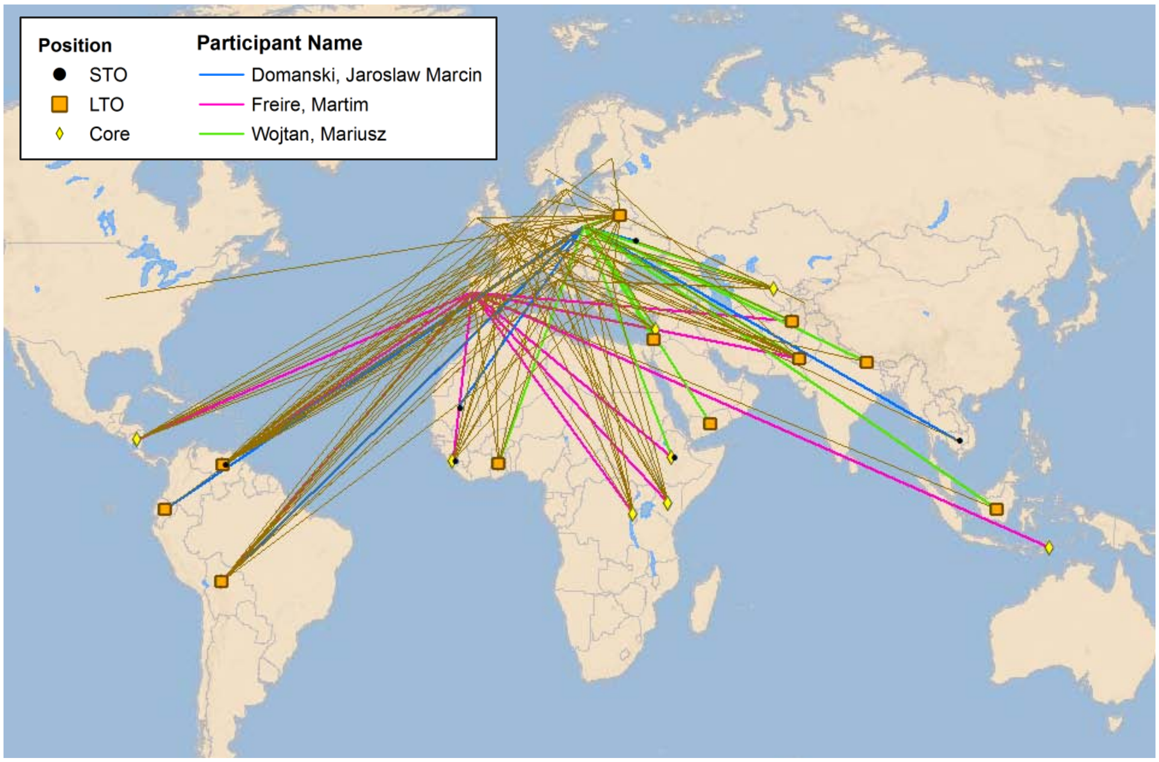

Figure 6.11 – Third-level Analysis of Participants

After the third level of analysis, the visualization was too cluttered to extract any meaningful information. Manual filtering, selection, or other manipulation mitigates some of this, but it requires significant user intervention.

The total number of assignments within the MMDB is 4158. This is the same as the number of first-level links – that is, each link from an election monitor's home of record to the country of each one of their missions (Figure 6.12).

Figure 6.12 – All Assignments. First Level Analysis

Visualization in NodeXL and Pajek is similar. Both offer multiple ways to energize, or draw, a link chart. This is set by the user at the outset of each chart creation. For this project, the Harel-Koren Fast Multiscale layout was selected because of its speed in rendering. The first rendering in NodeXL is that of the EU EOM Afghanistan 2005-9 mission with three participants. The colors and labels were added to improve visualization. The nodes were also manually moved so that the labels were visible (Figure 6.13).

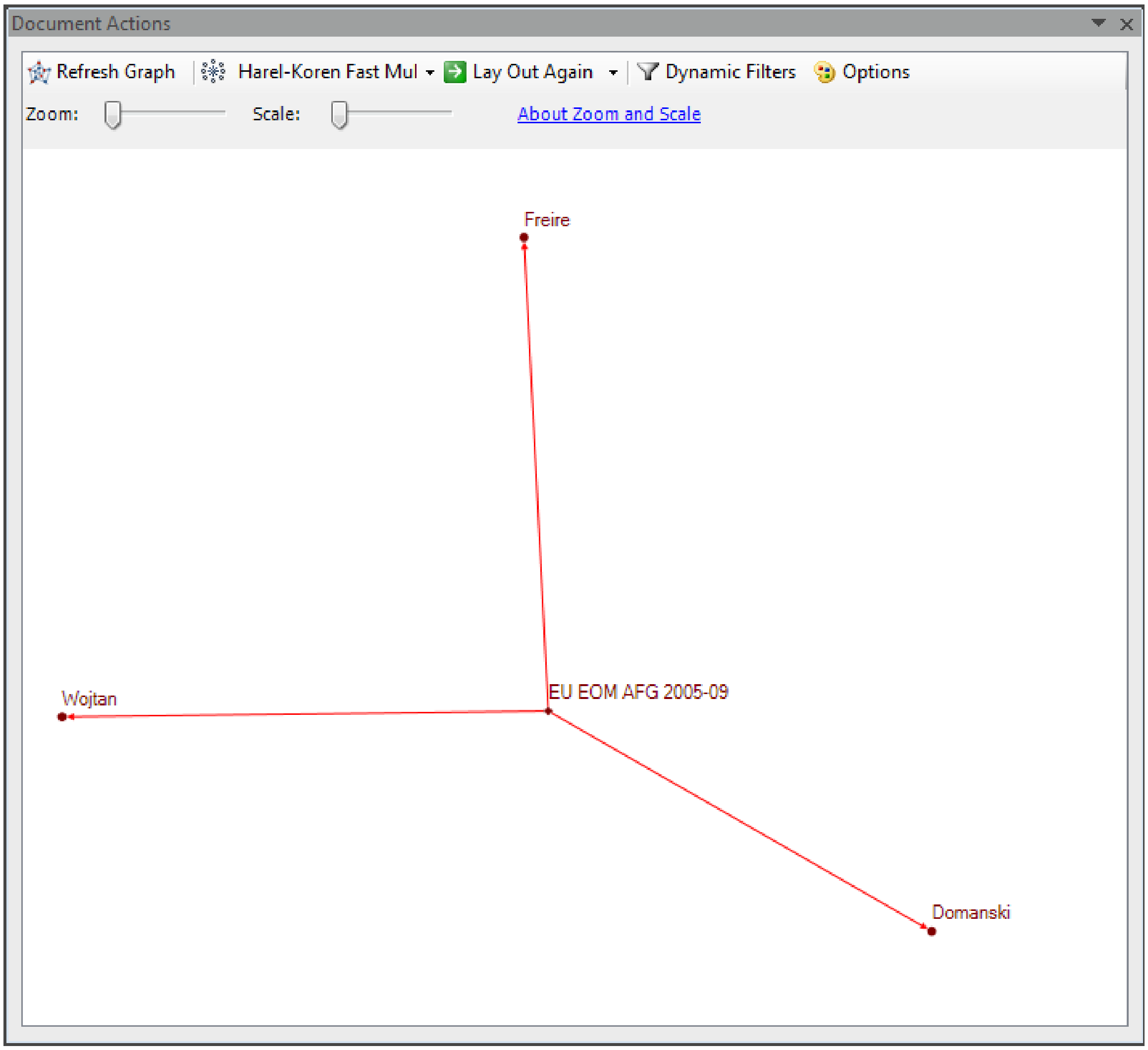

Figure 6.13 – EU EOM Afghanistan 2005-09 Mission Participants (NodeXL)

Exporting the tables from Access to Excel assists in the visualization within NodeXL. For the second level of analysis – all of the other missions the three participants worked on – the table was exported to Excel and then imported to NodeXL (Figure 6.14). Labels were added to the missions in which two or more of the election monitors participated. Neither NodeXL nor Pajek have good label placement engines, so only the most important labels were added.

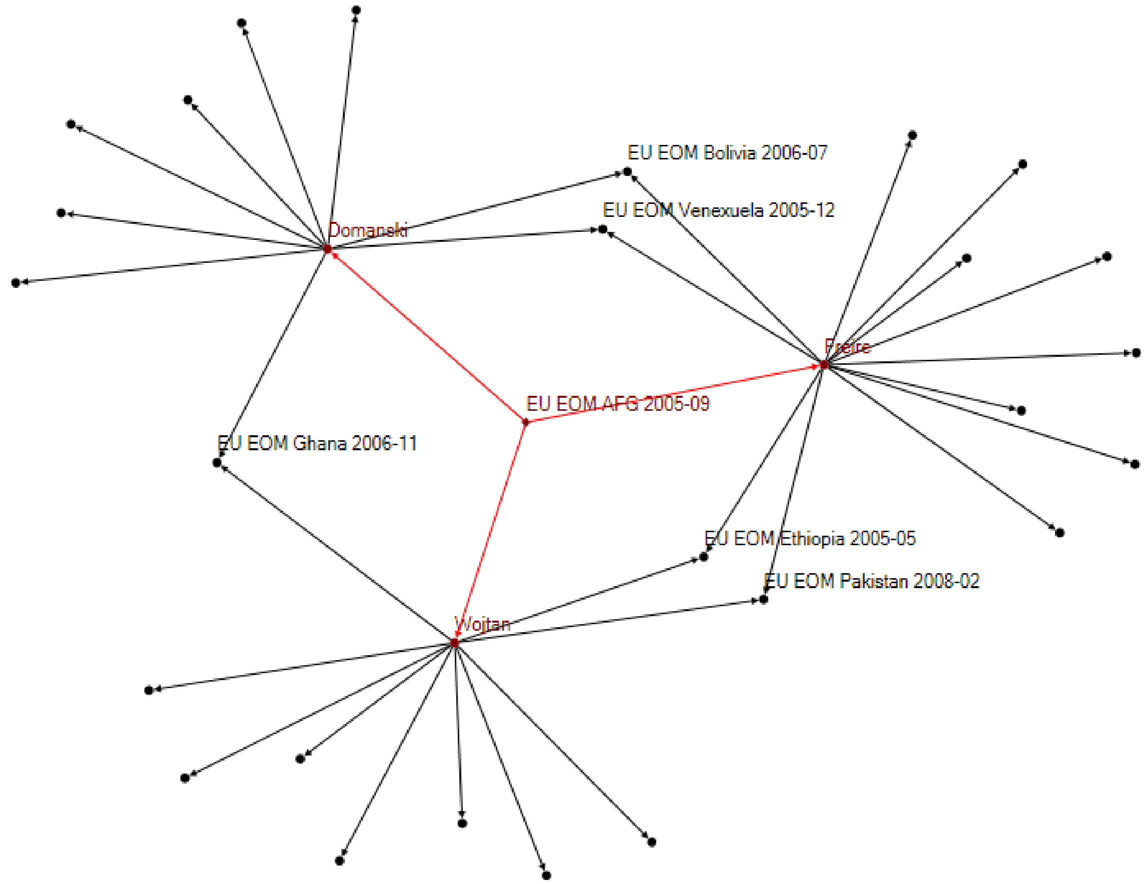

Figure 6.14 – Second-level Link Analysis (NodeXL, Harel-Koren Fast Multiscale)

An example of a different layout is in Figure 6.15. The Sugiyama layout resembles a hierarchy or family-tree in design and is useful for seeing descendents of a node of interest. Labeling the nodes, however, is not very easy without manually moving each labeled node. Because each of the levels of analysis fall on a horizontal plane, the labels overlap and obscure each other. In Figure 6.15 the third level labels have been removed.

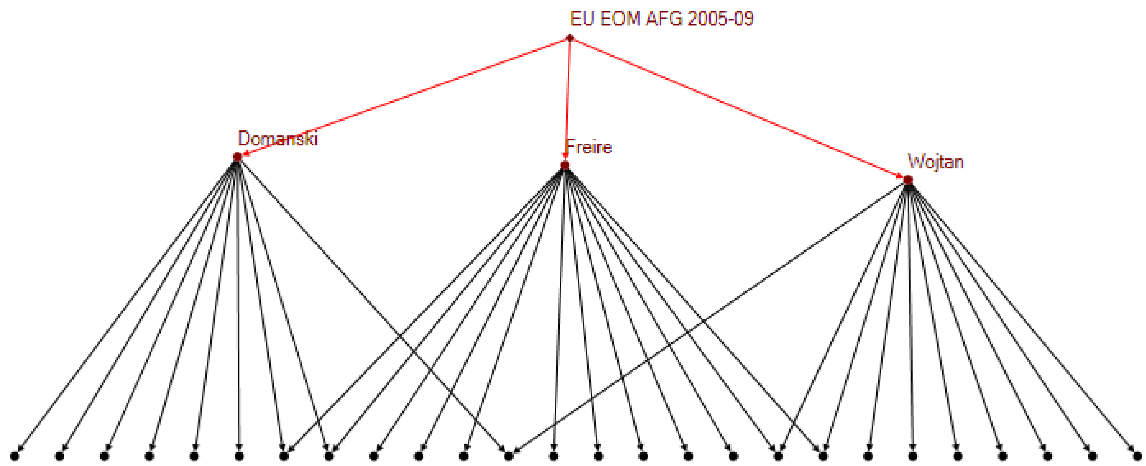

Figure 6.15 – Sugiyama Visualization

The third level of analysis becomes complicated in the SNA programs as well. The number of nodes and edges, as well as the lack of labels, makes analysis of a printed or hardcopy layout difficult. The analysis must be done within the programs. In the case of Figure 6.16, the mission nodes were manually colored blue and the participant nodes were changed to green. This gives the user more context in which to understand the output.

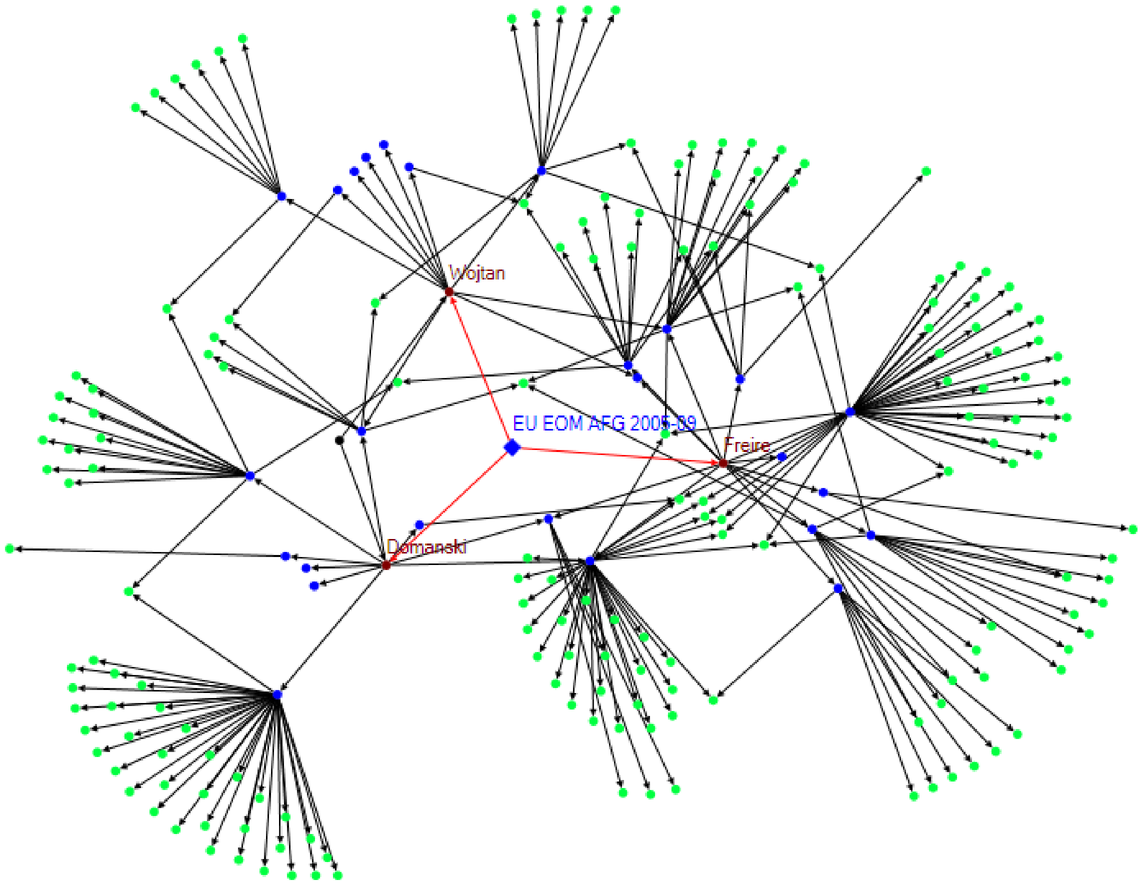

Figure 6.16 – Third Level Analysis

Pajek is a more powerful SNA program than NodeXL, but it is less intuitive and less user-friendly to those familiar with Windows and MS Office interfaces. As such, analysis is better within Pajek but visualization is less efficient. It is up to the individual user to decide which program to use. Figure 6.17 depicts the third level analysis of the EU EOM mission to Afghanistan in September 2005. It is the same data and was energized the same way (Harel-Koren) as in Figure 6.16, but has a slightly different layout. The label for the 2005 Afghanistan mission was added to assist the reader.

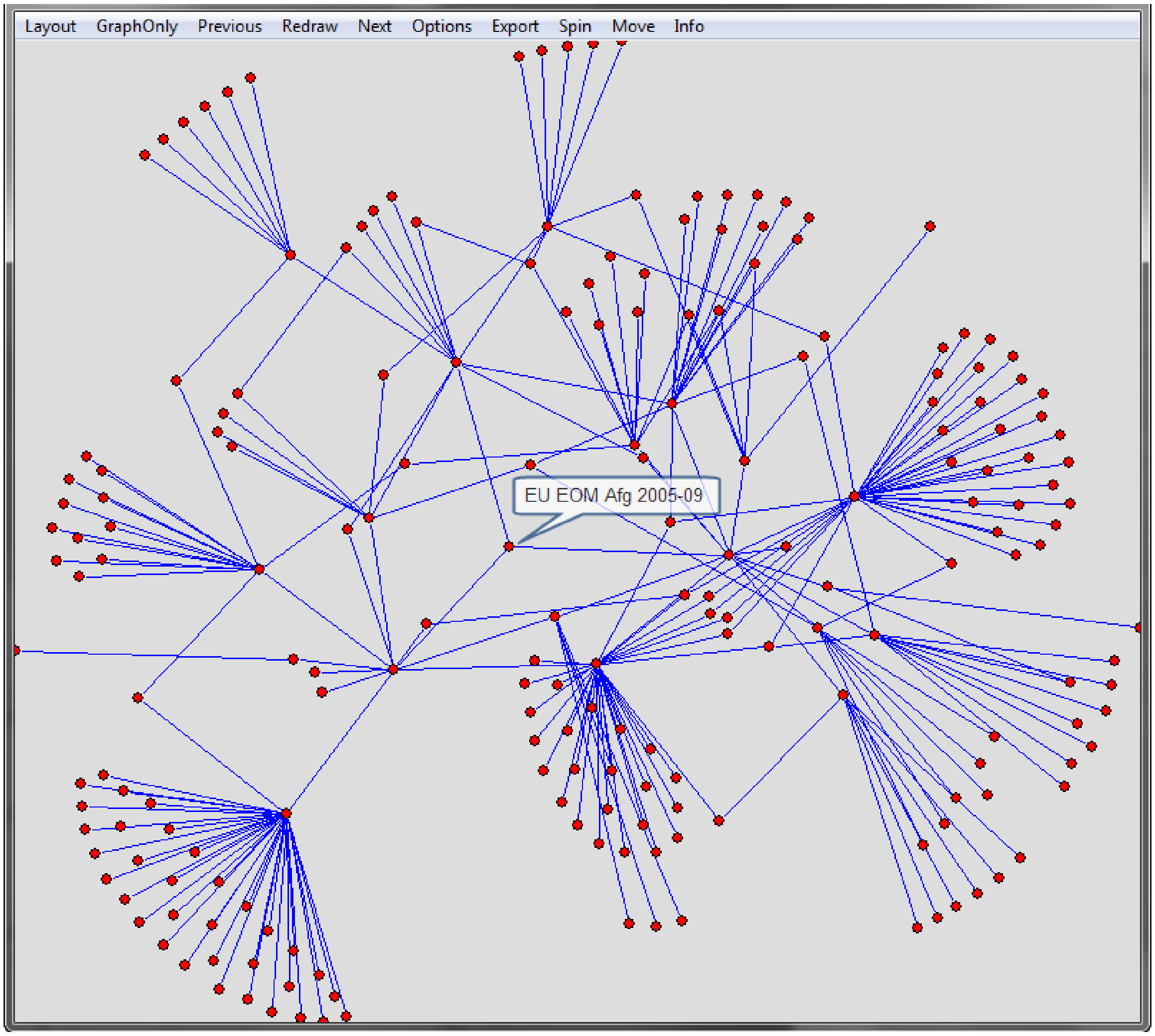

Figure 6.17 – Third Level Analysis in Pajek (Label Added)

Overall, the project met or can meet the client's functional requirements, albeit in a more limited way or with manual manipulation of the data. The visualization of the network of international election monitors is possible using ArcGIS, but any spatial analysis is not practical. The best results come when the user includes aggressive filtering of the results as well as only visualizing limited portions of the network. Some aspects can be improved through future work on the project; other aspects can only be solved through the use of high-end, purpose-built social network analysis software.

Chapter 7 – Conclusions and Future Work

The GIS Tool for Election Monitoring Research is a GIS/SNA application that was custom-built to assist Dr. Kimberley Coles in her research and visualization of the networks of international election monitors and democratization advisors. Through the use of an intermediate database, ArcGIS, and the availability of export to two SNA programs, the program successfully accomplished the geographic visualization of the network. The visualization showed, however, that the networks for such groups can and do grow quickly through each iteration, to the extent that they can be rendered almost meaningless after a few iterations.

After identifying the client's functional and non-functional requirements, the decision was made to not modify her original data storage program, the Missions Management Database (MMDB). This necessitated the construction of an intermediate Access 2007 database (the i_MMDB) that simulated the original MMDB in every way except that the data tables were linked to the MMDB rather than duplicating the tables. Furthermore, this intermediate database and its linkages meant that the client could continue data entry to the original MMDB during the construction of this project and would not encounter update errors when this project and its version of the MMDB was turned over to her. The i_MMDB was constructed to take the data from the MMDB, query, sort, and return it as a social network, and then make it available to ArcMap for spatial visualization. It also formatted the data for proper export to Pajek and NodeXL when desired.

With the data available to ArcMap through a Microsoft Jet 12 OLE connection, it was manipulated and displayed spatially. Through the use of Python scripts and ModelBuilder, the data were geo-located to the countries desired. A modified points-to-lines Python script then connected the nodes in the network as a shapefile of lines. Finally, a "Display XY events" command provided the necessary points to symbolize the attributes of the individuals in the network (Figure 7.1).

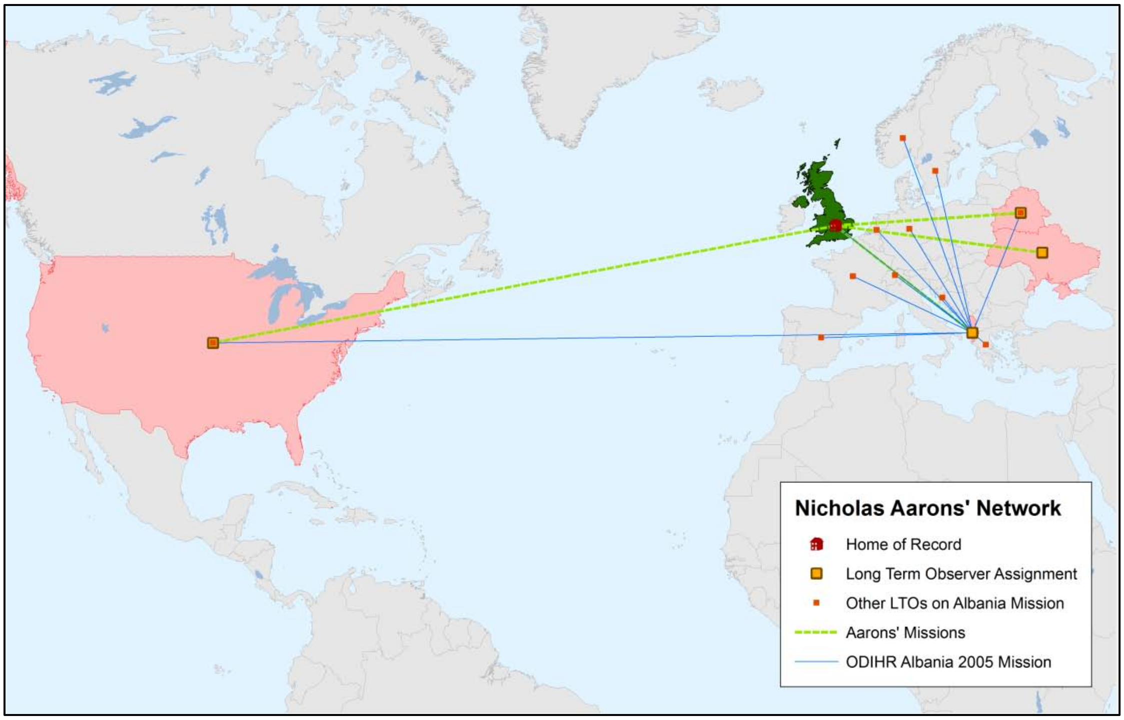

Figure 7.1 – Links visualized in ArcMap

As with any time-constrained project, there is more work that can be done to this program to expand the functionality and audience of this project. Because the project was custom-built to specifically solve the problem of geographic visualization of international election monitors, its construction limited it to interaction with the MMDB. Some of the queries are hard-coded into the program and do not work with other datasets without modification. Further work could alter the program to include user-defined variables that would link it to any relational database. Doing this would make it act as a framework for the geographic visualization of many other types of networks.

As mentioned previously, aggressive filtering is crucial for obtaining good results for visualization. Numerous methods exist that could be included in future work, but one of the more promising methods is presented here. To begin with, it should be noted that high-end social network analysis and link analysis programs are available that could perform functions similar to this. The cost of those programs makes them unviable as solutions in this case and a lower-cost system is needed. The high-end SNA programs are purpose-built from the beginning to include such analysis; building such a feature in Access and Excel – while possible – is time consuming and inefficient and thus not included in this program at this time. The following sections discuss how that could be accomplished if desired. It re-examines the relationships defined in this project and proposes different methods to identify links.

The first step in this process would be to change the way in which relationships are tracked. Because the project attempts to track relationships through space and time, election monitors who have only participated in one mission can be excluded at the outset of any query. With only one assignment, it is unlikely they have had as much influence over the network as someone who has participated in many missions and held many assignments. This also fits better with the evolution of the election monitoring field mentioned in chapter 2: *ad hoc* participation has given way to pseudo-careerism. The individuals holding only one assignment are effectively *ad hoc* participants in the network until multiple assignments build their experiences into careers worth studying.

While eliminating one-time participants is beneficial, eliminating one-time partnerships further limits the network. Multiple occurrences of partnerships suggests some sort of affinity between members or some other bond atypical of the norm. While these multiple partnerships may be coincidental, identifying them allows the user further opportunity to research the pairings. Running the query in this manner eliminates some intermediate steps that are present in the current application. Rather than recording the links as occurring through their assignments, this way directly links the participants and eliminates mention of the mission in the final results. An example of that would be Person 1 and Person 2 (P_1 and P_2, respectively) holding the same team and cohort assignments on two different missions, M_1 and M_2. In the current database, the relationships would be recorded as P_1 linked to P_2 through M_1, and P_1 linked to P_2 through M_2. It could be transcribed as P_1-M_1-P_2 and P_1-M_2-P_2. Using the proposed method of linking, the mission, while initially used to define the link, would be dropped from the final result. The result would be that P_1 and P_2 are linked two times – symbolized as P_1-P_2 (2). The difference is subtle, but focuses the result on the relationship between the participants, not the assignments on which the relationships were formed.

Once the query that finds the multiple occurring partnerships runs, it would append the results to a table with a field for the number of times the partnership took place. This field is used to determine part of the visualization in the link chart. Graph theory typically places nodes with higher bonds closer together than nodes with weaker bonds. Some of the social network analysis programs also allow the manipulation of the widths of the arcs connecting the nodes. The number of occurrences would be used to dictate these parameters.

A final item for future consideration would allow the user to select a number of specific elections and then analyze the election monitors assigned to these missions. A significant application of this would be to study recurring suspected fraud, collusion, or other inappropriate activity by election monitors. Using outside knowledge, the user could form a list of suspect elections or suspect missions. The program would compare the lists of election monitors and their assignments on all those suspect missions and return any instances of similarity. The user could then decide if any of those recurring partnerships on suspect missions warrant further investigation or are in any way significant.

Taking the above tool one step further would yield a form of a predictive-analysis tool. Using the lists of suspect elections and suspect election monitors, the program could then identify suspect partnerships in the records of new entries. The system could run every proposed mission assignment roster against the existing suspect list and flag the new mission as being vulnerable to compromise if a user-defined threshold number of suspected partnerships existed.

These tools described above all build on each other, but are a result of the re-definition of how relationships are defined and linked in the program.

Secondly, the data for the MMDB was collected from the websites of the international election monitoring organizations. This project could be redesigned into a web GIS program that would allow sharing with a limited-access or unlimited web audience. Sharing the work online with other interested researchers enhances the robustness of the overall results. Furthermore, allowing the international NGOs permission to use the results on their websites would provide graphical depictions of the diversity, or lack thereof, of participants and leaders of the respective missions.

Specific refinements to this project could include the use of Network Analyst to improve the trace functionality of the networks. Other refinements would be tools to limit the networks more effectively. Because of variances or incompleteness of mission record-keeping, the data structure of many of the missions vary widely. These variances could not always be addressed in this project and contributed to the large size of later network query iterations. Finding a method to decrease the size of these networks would be beneficial in that it would likely yield more informative results.

Finally, the future work could include the dynamic visualization of the missions over time. The idea is that the missions display as an animation and links are drawn as time progresses through a time-slider. Currently this technology is not well developed in ArcGIS 9.3 or ArcGIS Explorer and cannot be implemented efficiently. Improvements are expected in ArcGIS 10 and may make this future work achievable.

Overall, this project was designed to take data from the Missions Management Database, manipulate it, and display it in a geographic and spatial context. It performs this function and provides the client with such visualization. Although a similar project has been completed, this was the first project known to undertake the concurrent visualization of such large networks of individual actors from a system not designed specifically to undertake such large networks. It used existing programs and tools to visualize data originally held in a system not designed to address spatial or geographic visualization.

Works Cited

Brown, D. S., Brown, J. C., & Desposato, S. W. (2008). Who Gives, Who Receives, and Who Wins? *Comparative Political Studies, 41*, 24-47.

Central Intelligence Agency. (2009, November 30). *Bosnia and Herzegovina*. Retrieved October 27, 2009, from The World Factbook: https://www.cia.gov/library/publications/the-world-factbook/geos/bk.html

Coles, K. A. (2004). Election Day: The Construction of Democracy through Technique. *Cultural Anthropology, 19* (4), 551-580.

Covey, M. E. (2007, July). *Social Network Analysis for Army Experimentation Events (Master's Thesis).* Retrieved September 29, 2009, from http://vwe.spatial.redlands.edu/MSGIS/MIP%20Library/Forms/AllNoAbstract.aspx

de Nooy, W., Mrvar, A., & Batageli, V. (2005). *Exploratory Social Network Analysis with Pajek.* Cambridge: Cambridge University Press.

ESRI. (n.d.). *Adding OLE Database Connections*. Retrieved June 1, 2010, from ArcGIS Desktop 9.3 Help: http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=Adding%20OLE%20DB%20connections

Goodchild, M. F. (1996). Geographic Information Systems and Spatial Analysis in Social Software. In M. Aldenderfer, & H. D. Maschner (Eds.), *Anthropology, Space, and Geographic Information Systems* (pp. 241-250). New York: Oxford University Press.

Goodchild, M. F., Anselin, L., Appelbaum, R. P., & Harthorn, B. H. (2000). Toward Spatially Integrated Social Science. *International Regional Science Review, 23* (2), 139-159.

Goodchild, M. F., Egenhofer, M. J., Kemp, K. K., Mark, D. M., & Sheppard, E. (1999). Introduction to the Varenius Project. *International Journal of Geographical Information Science, 13* (8), 731-745.

i2. (2010). *www.i2group.com.* Retrieved May 19, 2010, from Analyst's Notebook 8 Update: http://www.i2group.com

International Network For Social Network Analysis. (2008). Retrieved October 27, 2009, from http://www.insna.org

Marwick, M. G. (1967). The Study of Witchcraft. In A. Epstein (Ed.), *The Craft of Social Anthropology* (pp. 231-244). Delhi: Hindustan Publishing Corporation.

Minnesota Population Center. (2010). Integrated Public Use Microdata Series, International: Version 6.0 [Machine-readable database]. Minneapolis: University of Minnesota.

Mitchell, J. C. (1967). On Quantification in Social Anthropology. In A. Epstein (Ed.), *The Craft of Social Anthropology* (pp. 17-45). Delhi: Hindustan Publishing Corporation.

Office for Democratic Institutions and Human Rights. (2006, June). *Terms of Reference, Long-Term Observer Coordinator, Limited Election Montoring Mission.* Retrieved October 27, 2009, from http://www.osce.org/documents/odirh/2009/06/38072_en.pdf

Organization for Security and Co-operation in Europe. (2006, December 13). Retrieved October 27, 2009, from Press Release - OSCE democratization body begins election observation mission to Albania: http://www.osce.org/item/22692.html

Pajek Wiki. (2008, May 23). Retrieved December 1, 2009, from Pajek Wiki: http://pajek.imfm.si/doku.php

Paley, J. (2002). Toward an Anthropology of Democracy. *Annual Review of Anthropology, 31*, 469-496.

Social Media Research Foundation. (2010, May 28). *NodeXL Overview*. Retrieved June 1, 2010, from NodeXL: http://nodexl.codeplex.com

University of Redlands. (n.d.). *Faculty Page: Kimberley Coles*. Retrieved October 1, 2009, from University of Redlands: http://www.redlands.edu/kimberley_coles.asp

Visual Analytics Inc. (2010). *VisuaLinks*. Retrieved December 1, 2009, from Visual Analytics Inc Web site: http://www.visualanalytics.com/default.cfm?Tier=2&MenuItemName=VisuaLinkVisuaLin&URL=%2Fproducts%2Fvisualinks%2Findex.cfm

Visual Analytics, Inc. (n.d.). *VisuaLinks*. Retrieved from Visual Analytics: http://www.visualanalytics.com/default.cfm?Tier=2&MenuItemName=VisuaLinksProduct&URL=%2Fproducts%2Fvisualinks%2Findex.cfm

Wejnert, B. (2005). Diffusion, Development, and Democracy. *American Sociological Review, 70*, 53-81.

Appendix A. List of Country Entries Modified

These modifications were made during the conversion of the World_GIS shapefile to the Countries.shp shapefile using the country_tbl table in the MMDB as a guide.

Countries or entities added to Countries.shp

Name	Country Code
ÅLAND ISLANDS	AX
HONG KONG	HK
KOSOVO	KV
MACAO	MO
MONTENEGRO	ME
PALESTINIAN TERRITORY, OCCUPIED	PS
SAINT BARTHÉLEMY	BL
SAINT MARTIN	MF
TAIWAN, PROVINCE OF CHINA	TW
UNITED STATES MINOR OUTLYING ISLANDS	UM
Unknown	NA

Countries or entities deleted or merged:

Baker I.
Gaza Strip
Glorioso Is.
Howland I.
Jan Mayen
Jarvis I.
Johnston Atoll
Juan De Nova I.
Midway Is.
Wake I.

Names in World_GIS Shapefile	Changed to MMDB Name Format
American Samoa (Eastern Samoa)	AMERICAN SAMOA
Antigua & Barbuda	ANTIGUA AND BARBUDA
The Bahamas	BAHAMAS
Bosnia & Herzegovina	BOSNIA AND HERZEGOVINA
Bouvet I.	BOUVET ISLAND
British Virgin Is.	VIRGIN ISLANDS, BRITISH
Brunei	BRUNEI DARUSSALAM
Cayman Is.	CAYMAN ISLANDS
Christmas I.	CHRISTMAS ISLAND
Cocos Is.	COCOS (KEELING) ISLANDS
Congo, DRC	CONGO, THE DEMOCRATIC REPUBLIC OF THE
Cook Is.	COOK ISLANDS
Cote d'Ivoire	CÔTE D'IVOIRE
Falkland Is.	FALKLAND ISLANDS (MALVINAS)
Faroe Is.	FAROE ISLANDS
French Southern & Antarctic Lands	FRENCH SOUTHERN TERRITORIES
The Gambia	GAMBIA
Heard I. & McDonald Is.	HEARD ISLAND AND MCDONALD ISLANDS
Vatican City	HOLY SEE (VATICAN CITY STATE)
Iran	IRAN, ISLAMIC REPUBLIC OF
North Korea	KOREA, DEMOCRATIC PEOPLE'S REPUBLIC OF
South Korea	KOREA, REPUBLIC OF
Laos	LAO PEOPLE'S DEMOCRATIC REPUBLIC
Libya	LIBYAN ARAB JAMAHIRIYA
Macedonia	MACEDONIA, THE FORMER YUGOSLAV REPUBLIC OF
Marshall Is.	MARSHALL ISLANDS
Micronesia	MICRONESIA, FEDERATED STATES OF
Moldova	MOLDOVA, REPUBLIC OF
Pitcairn Is.	PITCAIRN
Russia	RUSSIAN FEDERATION
St. Helena	SAINT HELENA
St. Kitts & Nevis	SAINT KITTS AND NEVIS
St. Lucia	SAINT LUCIA
St. Pierre & Miquelon	SAINT PIERRE AND MIQUELON
St. Vincent & the Grenadines	SAINT VINCENT AND THE GRENADINES
Sao Tome & Principe	SAO TOME AND PRINCIPE
Serbia & Montenegro	SERBIA
Solomon Is.	SOLOMON ISLANDS
South Georgia & the South Sandwich Is.	SOUTH GEORGIA AND THE SOUTH SANDWICH ISLANDS
Svalbard	SVALBARD AND JAN MAYEN
Syria	SYRIAN ARAB REPUBLIC
Tanzania	TANZANIA, UNITED REPUBLIC OF
Trinidad & Tobago	TRINIDAD AND TOBAGO
Turks & Caicos Is.	TURKS AND CAICOS ISLANDS
Virgin Is.	VIRGIN ISLANDS, U.S.
Wallis & Futuna	WALLIS AND FUTUNA
West Bank	WEST BANK & GAZA STRIP

Appendix B. List of Team Codes from Missions 174 and 229

Mission 174: European Union Election Observation Mission to Ghana, December 2008

Name	Country	Job	Team
Leemet, Liina	ESTONIA	LTO	4
Herrmann, Ron	GERMANY	Core	
Domanski, Jaroslaw Marcin	POLAND	Core	
Mladenov, Nickolay	BULGARIA	HoM	
Ward, David	UNITED KINGDOM	DCO	
Marques, Tania	PORTUGAL	Core	
Marchese, Mirella	ITALY	Core	
Eick, Mathias	GERMANY	Core	
Huidobro Foncillas, Alberto	SPAIN	Core	
Bruinsma, Sikke	NETHERLANDS	LTO	1
Cernouskova, Marketa	CZECH REPUBLIC	LTO	1
O'Shea, Aidan	IRELAND	LTO	2
Skopa, Marianna	GREECE	LTO	2
Luksaite, Jurga	LITHUANIA	LTO	3
Pschikal, Alexander	AUSTRIA	LTO	3
Geny, Emmanuel	FRANCE	LTO	4
Leal, Rita S.	PORTUGAL	LTO	5
Jockers, Heinz	GERMANY	LTO	6
Ganne, Joseph	BELGIUM	LTO	6
Alborghetti, Maria Serena	ITALY	LTO	7
Day, John	UNITED KINGDOM	LTO	7
Merino, Mathieu	FRANCE	LTO	8
de Felix, Silvia	SPAIN	LTO	8
Somogyi, Annamaria	HUNGARY	LTO	9
Charanas, Iordanis	GREECE	LTO	9
Forno, Enrico Pier Patrizio	ITALY	LTO	10
Kooijmans, Cornelis	NETHERLANDS	LTO	10
Felch, Peter	AUSTRIA	LTO	11
Popova, Deliana	BULGARIA	LTO	11
Braconnier, Luc	LUXEMBOURG	STO	201
Jurecko, Sasa	SLOVENIA	STO	201
Ioannou, Dimitra	GREECE	STO	301
Botnen, Trond	NORWAY	STO	301
Lecheva, Lydia	BULGARIA	STO	501
Somers, James	IRELAND	STO	501
Yordanov, Nikola	BULGARIA	STO	601
Pedersen, Flemming	DENMARK	STO	701
Tumba, Irene	BELGIUM	STO	701
Rainer, Dieter	AUSTRIA	STO	702
Adnanes, Dagfinn	NORWAY	STO	702
Savina, Giulio	ITALY	STO	801
Wiik, Oystein	NORWAY	STO	801
Lourenco, Vera	PORTUGAL	STO	901
Agintas, Rolandas	LITHUANIA	STO	901
Ciganikova, Martina	SLOVAKIA	STO	1201

Name	Country	Job	Team
Jodal, Morten	NORWAY	STO	1201
Sola Martin, Andreu	SPAIN	LTO	12
Vennala, Antti	FINLAND	LTO	12
Bernard, Claire-Emmanuelle	FRANCE	STO	601
Wojtan, Mariusz	POLAND	LTO	5

Mission 229: European Union Election Observation Mission to El Salvador, March 2009

t_Nam	t_Cntry	t_Job	t_Tm
de Gabriel Perez, Jose Antonio	SPAIN	Core	
Alves, Cristina	PORTUGAL	Core	
Jordan, Andreas	AUSTRIA	Core	
Arranz, Miguel	BELGIUM	Core	
Pappalardo, Carlo	ITALY	Core	
Garcia, Juan Pedro	SPAIN	Core	
Gutierrez, Javier	SPAIN	Core	
Yanez-Barnuevo, Luis Garcia	SPAIN	Core	
Caligiuri, Giovanni	ITALY	LTO	1
Vincze, Zsuzsa	HUNGARY	LTO	1
Galea, Sebastian	FRANCE	LTO	2
Gloyer, Gilian	UNITED KINGDOM	LTO	2
Wieschiolek, Heike	GERMANY	LTO	3
Kretzschmar, Jan	GERMANY	LTO	4
Chalmeta, Juan Ribo	SPAIN	LTO	5
Majasaari, Timo J.	FINLAND	LTO	6
Silva, Ruth	PORTUGAL	LTO	7
Bonnannini, Maximo	ITALY	LTO	7
Damiani, Cesira	ITALY	LTO	8
Roiss, Manuel	AUSTRIA	LTO	8
Ileva, Slavena	BULGARIA	LTO	9
Michalik, Peter	SLOVAKIA	LTO	9
Kjelstrup, Bjarne	DENMARK	LTO	10
Carlsson, Ulla	SWEDEN	LTO	10
Contreras de Passos, Madalena	PORTUGAL	LTO	11
Gonzalez Diaz, Dulce Maria	SPAIN	LTO	11
Cayzac, Hugues	FRANCE	LTO	3
Mier Sainz, Javier	SPAIN	LTO	4
Carrillo, Maria Bertha	ITALY	LTO	5
Allegret, Marie	FRANCE	LTO	6
Jandura, Dalibor	CZECH REPUBLIC	STO	011
Tirone, Letizia	ITALY	STO	012
Bukovska, Erika	SLOVAKIA	STO	021
Pesti, Mele	ESTONIA	STO	022
Castanias, Marie-Helene	GREECE	STO	031
Hansen, Krister	SWEDEN	STO	032
Szlefarska, Barbara	POLAND	STO	032

t_Nam	t_Cntry	t_Job	t_Tm
Rodriguez, Maria Victoria	SPAIN	STO	041
Petraskaite, Vida	LITHUANIA	STO	042
Jensterle, Marko	SLOVENIA	STO	042
Peavoy, Diarmuid	IRELAND	STO	051
Gaywood, Annabel	UNITED KINGDOM	STO	061
Dolata, Piotr	POLAND	STO	061
Ditteova, Lenka	SLOVAKIA	STO	062
Tammsaar, Rein	ESTONIA	STO	071
Iwanicka, Anna	POLAND	STO	072
De Dycker, Karen	BELGIUM	STO	081
Wlaschutz, Christian	AUSTRIA	STO	081
Schnitzer, Karl	GERMANY	STO	082
Henriques Pereira, Marta	PORTUGAL	STO	091
Satkauskas, Rytis	LITHUANIA	STO	092
George, Isabelle	BELGIUM	STO	101
Dunne, Mary Helena	IRELAND	STO	101
Cox, Rebecca	UNITED KINGDOM	STO	111
Szabo, Levente	HUNGARY	STO	112
Giongo, Alessandro	ITALY	STO	113
Makihannu, Mery Tuuli	FINLAND	STO	113
Gross, Toomas	ESTONIA	STO	114
Nezmeskalova, Olga	CZECH REPUBLIC	STO	114
Chinchon Alvarez, Javier	SPAIN	STO	111
Andrade Thuau, Maria Joao	PORTUGAL	STO	112
Kervinen, Aino Inkeri	FINLAND	STO	082
Zikevicius, Dainius	LITHUANIA	STO	072
Huck, Alexandra	GERMANY	STO	071
Borok, Gyorgy	HUNGARY	STO	062
Barre, Marie-Chantal	FRANCE	STO	051
Roudiere, Francis	FRANCE	STO	031
Garcia de la Rosa, Philippe	FRANCE	STO	041
Borkowska, Desislava	BULGARIA	STO	022
Clausse, Germaine	LUXEMBOURG	STO	021
Lopez-Medel Bascones, Jesus	SPAIN	STO	012
Poschl, Caroline	AUSTRIA	STO	011
Suvadova, Gabriela	SLOVAKIA	STO	091
Radmore, Rebecca	UNITED KINGDOM	STO	092

Appendix C. QuickMap Python Scripts

```
# ---------------------------------------------------------------------------
# MsnByPaxScript.py
# Created on: Thu Jul 29 2010 02:05:35 PM
#   (generated by ArcGIS/ModelBuilder)
# Usage: MsnByPaxScript <Expression>
# ---------------------------------------------------------------------------

# Import system modules
import sys, string, os, arcgisscripting

# Create the Geoprocessor object
gp = arcgisscripting.create(9.3)
gp.OverWriteOutput = 1

# Load required toolboxes...
gp.AddToolbox("C:/Program Files (x86)/ArcGIS/ArcToolbox/Toolboxes/Data Management
Tools.tbx")
gp.AddToolbox("C:/Program Files (x86)/ArcGIS/ArcToolbox/Toolboxes/Analysis Tools.tbx")

# Script arguments...
msnPaxName = gp.GetParameterAsText(0)
Expression = "e_Nam = '" + msnPaxName + "'"


# Local variables...
gQryMsnByPax_TableSelect_dbf = "C:\\MIP\\SCRATCH\\gQryMsnByPax_TableSelect.dbf"
gQryMsnByPax = "gQryMsnByPax"
gQryMsnByPax_Wrking_dbf_2_ = "C:\\MIP\\DATA\\WorkingTables\\gQryMsnByPax_Wrking.dbf"
gQryMsnByPax_Wrking_dbf = "C:\\MIP\\DATA\\WorkingTables\\gQryMsnByPax_Wrking.dbf"
gQryMsnByPax_Wrking_dbf__3_ = "C:\\MIP\\DATA\\WorkingTables\\gQryMsnByPax_Wrking.dbf"
Countries_HOR = "Countries_HOR"
CountriesWorking__2_ = "CountriesWorking"
Countries_HOR__3_ = "Countries_HOR"
CountriesWorking = "CountriesWorking"

# Process: Table Select...
gp.TableSelect_analysis(gQryMsnByPax, gQryMsnByPax_TableSelect_dbf, Expression)

# Process: Delete Rows...
gp.DeleteRows_management(gQryMsnByPax_Wrking_dbf)

# Process: Append...
gp.Append_management("C:\\MIP\\SCRATCH\\gQryMsnByPax_TableSelect.dbf",
gQryMsnByPax_Wrking_dbf_2_, "NO_TEST", "e_Nam 'e_Nam' true false false 254 Text 0 0
,First,#,C:\\MIP\\SCRATCH\\gQryMsnByPax_TableSelect.dbf,e_Nam,-1,-1;e_Citzn 'e_Citzn'
true false false 100 Text 0 0
,First,#,C:\\MIP\\SCRATCH\\gQryMsnByPax_TableSelect.dbf,e_Citzn,-1,-1;m_Pos 'm_Pos' true
false false 50 Text 0 0 ,First,#,C:\\MIP\\SCRATCH\\gQryMsnByPax_TableSelect.dbf,m_Pos,-
1,-1;m_Job_ID 'm_Job_ID' true false false 10 Double 0 10
,First,#,C:\\MIP\\SCRATCH\\gQryMsnByPax_TableSelect.dbf,m_Job_ID,-1,-1;m_Nam 'm_Nam' true
false false 254 Text 0 0 ,First,#,C:\\MIP\\SCRATCH\\gQryMsnByPax_TableSelect.dbf,m_Nam,-
1,-1;m_Cnty 'm_Cnty' true false false 100 Text 0 0
,First,#,C:\\MIP\\SCRATCH\\gQryMsnByPax_TableSelect.dbf,m_Cnty,-1,-1;m_Yr 'm_Yr' true
false false 4 Text 0 0 ,First,#,C:\\MIP\\SCRATCH\\gQryMsnByPax_TableSelect.dbf,m_Yr,-1,-
1;m_NGO 'm_NGO' true false false 20 Text 0 0
,First,#,C:\\MIP\\SCRATCH\\gQryMsnByPax_TableSelect.dbf,m_NGO,-1,-1", "")

# Process: Add Join...
gp.AddJoin_management(Countries_HOR__3_, "cntry_Nam", gQryMsnByPax_Wrking_dbf__3_,
"e_Citzn", "KEEP_COMMON")

# Process: Add Join (2)...
gp.AddJoin_management(CountriesWorking, "cntry_Nam", gQryMsnByPax_Wrking_dbf__3_,
"m_Cnty", "KEEP_COMMON")
```

```
# ---------------------------------------------------------------------------
# MsnByYrScript.py
# Created on: Thu Jul 29 2010 01:33:59 PM
#   (generated by ArcGIS/ModelBuilder)
# Used to identify missions in a user input year or range of  years
# Uses data from Missions Management Database
# ---------------------------------------------------------------------------

# Import system modules
import sys, string, os, arcgisscripting

# Create the Geoprocessor object
gp = arcgisscripting.create()

# Load required toolboxes...
gp.AddToolbox("C:/Program Files (x86)/ArcGIS/ArcToolbox/Toolboxes/Data Management
Tools.tbx")
gp.AddToolbox("C:/Program Files (x86)/ArcGIS/ArcToolbox/Toolboxes/Analysis Tools.tbx")

# Script arguments...
StartYear = gp.GetParameterAsText(0)
EndYear = gp.GetParameterAsText(1)
Expression = "m_Yr >= '" + StartYear + "'AND m_Yr <= '" + EndYear + "'"


# Local variables...
gQryMsnByYr =
"C:\\Users\\gregory_couch\\AppData\\Roaming\\ESRI\\ArcCatalog\\iMMDBconnection.odc\\gQryM
snByYr"
gQryMsnByYr_Wrking_dbf = "C:\\MIP\\DATA\\WorkingTables\\gQryMsnByYr_Wrking.dbf"
gQryMsnByYr_TblSel = "C:\\MIP\\SCRATCH\\gQryMsnByYr_TblSel"
Output_Table = "C:\\MIP\\DATA\\WorkingTables\\gQryMsnByYr_Wrking.dbf"
gQryMsnByYr_Wrking_dbf__2_ = "C:\\MIP\\DATA\\WorkingTables\\gQryMsnByYr_Wrking.dbf"
CountriesWorking = "CountriesWorking"
CountriesWorking__2_ = "CountriesWorking"

# Process: Table Select...
gp.TableSelect_analysis(gQryMsnByYr, gQryMsnByYr_TblSel, Expression)

# Process: Delete Rows...
gp.DeleteRows_management(gQryMsnByYr_Wrking_dbf)

# Process: Append...
gp.Append_management("C:\\MIP\\SCRATCH\\gQryMsnByYr_TblSel", Output_Table, "NO_TEST",
"m_ID 'm_ID' true false false 10 Double 0 10
,First,#,C:\\MIP\\SCRATCH\\gQryMsnByYr_TblSel,m_ID,-1,-1;m_Yr 'm_Yr' true false false 4
Text 0 0 ,First,#,C:\\MIP\\SCRATCH\\gQryMsnByYr_TblSel,m_Yr,-1,-1;m_Ctry 'm_Ctry' true
false false 100 Text 0 0 ,First,#,C:\\MIP\\SCRATCH\\gQryMsnByYr_TblSel,m_Ctry,-1,-1;m_Nam
'm_Nam' true false false 254 Text 0 0
,First,#,C:\\MIP\\SCRATCH\\gQryMsnByYr_TblSel,m_Nam,-1,-1;m_Pax 'm_Pax' true false false
10 Double 0 10 ,First,#,C:\\MIP\\SCRATCH\\gQryMsnByYr_TblSel,m_Pax,-1,-1;m_NGO 'm_NGO'
true false false 20 Text 0 0 ,First,#,C:\\MIP\\SCRATCH\\gQryMsnByYr_TblSel,m_NGO,-1,-1",
"")

# Process: Add Join...
gp.AddJoin_management(CountriesWorking__2_, "cntry_Nam", gQryMsnByYr_Wrking_dbf__2_,
"m_Ctry", "KEEP_COMMON")
```

```
# ---------------------------------------------------------------------------
# MembershipOrgs.py
# Created on: Tue Jun 08 2010 12:52:27 PM
#   (generated by ArcGIS/ModelBuilder)
# Usage: MembershipOrgs <Expression>
# ---------------------------------------------------------------------------

# Import system modules
import sys, string, os, arcgisscripting

# Create the Geoprocessor object
gp = arcgisscripting.create()

# Load required toolboxes...
gp.AddToolbox("C:/Program Files (x86)/ArcGIS/ArcToolbox/Toolboxes/Data Management
Tools.tbx")
gp.AddToolbox("C:/Program Files (x86)/ArcGIS/ArcToolbox/Toolboxes/Analysis Tools.tbx")

# Script arguments...
MembershipGroup = gp.GetParameterAsText(0)
Expression = "mbr_Grp = '" + MembershipGroup + "'"

#Expression = sys.argv[1]
#if Expression == '#':
# Expression = "mbr_Grp =" # provide a default value if unspecified

# Local variables...
gQryMembershipCountries =
"C:\\Users\\gregory_couch\\AppData\\Roaming\\ESRI\\ArcCatalog\\MMDBi_2007_Working_Connect
ion.odc\\gQryMembershipCountries"
gQryMbrCtry_Working_dbf__2_ = "C:\\_MIP-
Working\\Data\\WorkingTables\\gQryMbrCtry_Working.dbf"
Output_Table = "C:\\_MIP-Working\\Data\\WorkingTables\\gQryMbrCtry_Working.dbf"
gQryMbrCntry_TableSelect_dbf = "C:\\_MIP-Working\\SCRATCH\\gQryMbrCntry_TableSelect.dbf"
gQryMbrCtry_Working_dbf = "C:\\_MIP-
Working\\Data\\WorkingTables\\gQryMbrCtry_Working.dbf"
Cntry_Miller_Working = "Cntry_Miller_Working"
Cntry_Miller_Working__2_ = "Cntry_Miller_Working"

# Process: Table Select...
gp.TableSelect_analysis(gQryMembershipCountries, gQryMbrCntry_TableSelect_dbf,
Expression)

# Process: Delete Rows...
gp.DeleteRows_management(gQryMbrCtry_Working_dbf__2_)

# Process: Append...
gp.Append_management("C:\\_MIP-Working\\SCRATCH\\gQryMbrCntry_TableSelect.dbf",
Output_Table, "NO_TEST", "mbr_Grp 'mbr_Grp' true false false 50 Text 0 0
,First,#,C:\\_MIP-Working\\SCRATCH\\gQryMbrCntry_TableSelect.dbf,mbr_Grp,-1,-1;mbr_Cntry
'mbr_Cntry' true false false 100 Text 0 0 ,First,#,C:\\_MIP-
Working\\SCRATCH\\gQryMbrCntry_TableSelect.dbf,mbr_Cntry,-1,-1", "")

# Process: Add Join...
gp.AddJoin_management(Cntry_Miller_Working__2_, "cntry_Nam", gQryMbrCtry_Working_dbf,
"mbr_Cntry", "KEEP_COMMON")
```

```
# ----------------------------------------------------
# Points to lines tool. Version 0.2
# Version adapted/written by G. Couch
# 20100621
# ----------------------------------------------------

# Script uses input table with start and end coordinates and creates
# a polyline FC.

import sys, string, os, arcgisscripting

# Create the geoprocessor object
gp = arcgisscripting.create(9.3)
gp.OverWriteOutput = 1

# Get the user-defined parameters:
inTable = gp.GetParameterAsText(0)
originX = gp.GetParameterAsText(1)
originY = gp.GetParameterAsText(2)
destX = gp.getparameterastext(3)
destY = gp.GetParameterAsText(4)
outLineFC = gp.GetParameterAsText(5)
spref = gp.GetParameterAsText(6)

# Get an input feature class as a template to use later:
gp.addmessage("Getting template...")
tempFile = "C:/_MIP-Working/Data/EmptyShapeTmplt.shp"

# Make an empty line feature class using the temp template
outWorkspace = os.path.dirname(outLineFC)
gp.addmessage(outWorkspace)
gp.addmessage("Found the output directory...")
outFcName = os.path.basename(outLineFC)
gp.addmessage("Creating FC...")
gp.CreateFeatureClass(outWorkspace, outFcName, "POLYLINE", tempFile, "DISABLED",
"DISABLED", spref)
gp.addmessage("Empty Feature Class created...")

gp.addmessage("About to insert rows")
# Open a cursor to INSERT ROWS into the shapefile
#cur = gp.insertcursor(outFcName)
cur = gp.InsertCursor(outLineFC)
gp.addmessage("Rows Inserted")

gp.addmessage("About to Create array and point object")
# Create an Array and Point Object.
lineArray = gp.createobject("Array")
pnt = gp.createobject("Point")
gp.addmessage("array and point objects created...")

# Open a cursor on the table of XY coords to read from.
rows = gp.searchcursor(inTable)

# Reset the cursor to the top
row = rows.Next()

#e List the fields in the table so we can copy them later
fields = gp.ListFields(inTable)

# Loop through each record in the XY table.
gp.addmessage("Starting the while row")
while row:
    # Get the X and Y coords for the origin vertex.
    pnt.x = row.GetValue(originX)
    pnt.y = row.GetValue(originY)

    # Insert that point into the line array
    lineArray.add(pnt)

    # Get the X and Y coords for the destination vertex.
    pnt.x = row.GetValue(destX)
```

```
    pnt.y = row.GetValue(destY)

    # Insert that point into the line array
    lineArray.add(pnt)

    # Go to the next row in the table
    row = rows.Next()

    # Insert the new line into the feature class
    feat = cur.NewRow()
    feat.shape = lineArray

    cur.InsertRow(feat)

    gp.addmessage("Finished cur.insertRow...")
    lineArray.RemoveAll()

del cur, row, rows
```

Appendix D. Other Coding

To create the Overview Window:

```
-------------------------------------
Private Sub CreateOVwindow_Click()
    Call CreateOverviewWindow
End Sub

------------------------------------
Public Sub CreateOverviewWindow()

    'Declare variables
    Dim pOverview As IOverview
    Dim pOverviewWindow As IOverviewWindow
    Dim pDataWindowFactory As IDataWindowFactory
    Dim pFillSymbol As ISimpleFillSymbol
    Dim pLineSymbol As ISimpleLineSymbol
    Dim pRgbColor As IRgbColor


    'Set Variables
    Set pDataWindowFactory = New OverviewWindowFactory
    If Not pDataWindowFactory.CanCreate(Application) Then Exit Sub
    'Create a new overview window
    Set pOverviewWindow = pDataWindowFactory.Create(Application)
    'Change the area of interest fill symbol
    'to a hollow fill with a blue border
    Set pOverview = pOverviewWindow.Overview
    Set pFillSymbol = New SimpleFillSymbol
    Set pLineSymbol = New SimpleLineSymbol
    Set pRgbColor = New RgbColor


    'Set object properties
    pRgbColor.Blue = 255
    pLineSymbol.Color = pRgbColor
    pFillSymbol.Style = esriSFSNull
    pFillSymbol.Outline = pLineSymbol
    pOverview.AoiFillSymbol = pFillSymbol

End Sub
```

To create the Remove Join from Selected Layer button:

```
Private Sub RemoveJoin_Click()
        Call RemoveJoinMod
End Sub
     ------------------------------
Public Sub RemoveJoinMod()
     On Error GoTo EH

       Dim pDoc As IMxDocument
       Set pDoc = ThisDocument
       Dim pMap As IMap
       Set pMap = pDoc.FocusMap

       Dim thisLayer As ILayer
       Dim pDispRC As IDisplayRelationshipClass
       For i = 0 To (pDoc.FocusMap.LayerCount - 1) Step 1
         Set thisLayer = pDoc.FocusMap.Layer(i)

         If TypeOf thisLayer Is IFeatureLayer Then

             Set pDispRC = thisLayer

             ' Remove all joins
             If Not pDispRC.RelationshipClass Is Nothing Then
                 pDispRC.DisplayRelationshipClass Nothing, esriLeftInnerJoin
             End If

         End If

       Next i

       Exit Sub
     EH:

       MsgBox Err.Number & "  " & Err.Description
End Sub
```