Carnegie Mellon University

CARNEGIE INSTITUTE OF TECHNOLOGY

THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF Doctor of Philosophy

TITLE **Computational Models of Nuclear Proliferation**

PRESENTED BY **William Frankenstein**

ACCEPTED BY THE DEPARTMENT OF

Engineering and Public Policy

Kathleen M. Carley — May 4, 2016
ADVISOR, MAJOR PROFESSOR — DATE

Douglas Sicker — May 4, 2016
DEPARTMENT HEAD — DATE

APPROVED BY THE COLLEGE COUNCIL

Vijayakumar Bhagavatula — May 5, 2016
DEAN — DATE

Computational Models of Nuclear Proliferation

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Engineering and Public Policy

William Frankenstein

B.S., Mathematics, Hns. International Security, Stanford University
M.S., Mathematics & Statistics, Georgetown University

Carnegie Mellon University
Pittsburgh, PA

May 2016

Keywords: Nuclear Proliferation, International Security, Computational Modeling, Computational Social Science, Large Scale Data Analysis, Social Media Analysis, Network Analysis, Sentiment Analysis

Acknowledgements

This work is supported in part by the Defense Threat Reduction Agency (DTRA) under grant HDTRA11010102, Office of Naval Research (ONR) MURI (N000140811186), Department of Education Graduate Assistance in Areas of National Need (GAANN) (P200A090055), the Carnegie Mellon Crosswalk – Graduate Student Small Project Help (GuSH) fund and the center for Computational Analysis of Social and Organizational Systems (CASOS). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of DTRA, ONR, Dept. of Ed. or the US Government.

This thesis would not have been possible without the support of an incredible number of individuals. I would first like to thank my committee for their critical input in refining and improving several key arguments and aspects of each chapter. I am indebted to Kathleen M. Carley's ongoing support for my research and academic freedom during my time with EPP. I have grown as a teacher thanks to Baruch Fischhoff's mentorship while at EPP. And I am thankful for both Kiron Skinner and Paul Davis's willingness to lend their expertise and take on a naïve policy analyst combining disparate disciplines in this field.

I'd like to thank my mentors outside of CMU, who encouraged me to continue with the PhD: Michael May, who has always encouraged broader thinking; Scott Savitz, who has always counseled perseverance; David Sallach, who introduced me to the world at the national labs; Greg Treverton, who introduced me to the world of intelligence studies; and Henry Willis, who has consistently challenged me to be a better policy analyst.

I would also like to thank my family for their unwavering support from across the globe, as well as insistence on 加油 (presumably because of the health hazards associated with 加铀). At CMU, I would like to thank my fellow CASOS students, both past and present: Ghita Mezzour, a delight, for tolerating me in Wean 4215 as well as being a fantastic officemate, as well as Geoff Morgan, Kenny Joseph, Wei Wei, and Mike Lanham for their ongoing feedback, support, and powering through Summer Institutes. Laurie Jones and Pete Landwehr for neverending discussions on computational social science and its far-reaching implications. And to newer CASOS students Matt Benigni, Binxuan Huang, Geoff Dobson– thanks for the encouragement. And to other past occupants of Wean 4215: Miray Kas and Aman Ahuja – the space is in better (i.e. Sumeet Kumar's) hands now. To the COS and SE students in ISR who opened up their offices to an engineer without windows in his office – thank you. The CASOS programmers – Jeff Reminga, Mike Kowalchuck, Jon Storrick, and Neil Altman. And last but certainty not least, the CASOS staff: Hope Armstrong, Phil Curley, Maria Lauro, and Sienna Watkins.

In Engineering and Public Policy, I'd like to especially thank Ahmed Abdulla, for continuing to listen to me discuss nuclear policy, as well as Parth Vaishnav, Nathaniel Horner, and John Helveston for joining the ride through a quick tour of international security literature and taking my co-taught class with Ahmed. I also want to thank Rebecca Balebako, Casey Canfield, Travis Carless, Daniel Gingerich, Serban Mogos, Kim Mullins, Manar Sarie and Brinda Thomas for their work in helping make EPP a more welcoming and inclusive environment through being a GSA rep. Mitch Small, Granger Morgan, and Benoit Morel for introducing me to Kathleen while a first year graduate student in the department. All of the EPP staff for making the department a friendly home during my entire time at Carnegie Mellon. And to the rest of my cohort for being so efficient, productive, and amazing around the world that I realized I had to hurry up writing in order to graduate.

And finally, I don't think I could have made it without my Pittsburgh family. Whether in GSA Exec – Carolyn Commer, Mary Glavan, Hassan Khan, Carolyn Norwood, Sarah Tedrow-Azizi, Mark Whiting, Brittani Mckenna, and Nancy Stiger - through NAGPS, through Sam Adhikari , Dena Asta, Wayne Chuang, Mark Fuhs, Jack Kurutz, Ashwini Rao, and Mili-Ann Tamayao, Ben Tu, or through Meg Evans, Ashley Grice, Jeannene Jones , Jonathan Porras, Rich Truncellito, Laura Posner, Christian Köhler – thank you for what you do.

Abstract

This thesis utilizes social influence theory and computational tools to examine the disparate impact of positive and negative ties in nuclear weapons proliferation. The thesis is broadly in two sections: a simulation section, which focuses on government stakeholders, and a large-scale data analysis section, which focuses on the public and domestic actor stakeholders. In the simulation section, it demonstrates that the nonproliferation norm is an emergent behavior from political alliance and hostility networks, and that alliances play a role in current day nuclear proliferation. This model is robust and contains second-order effects of extended hostility and alliance relations. In the large-scale data analysis section, the thesis demonstrates the role that context plays in sentiment evaluation and highlights how Twitter collection can provide useful input to policy processes. It first highlights the results of an on-campus study where users demonstrated that context plays a role in sentiment assessment. Then, in an analysis of a Twitter dataset of over 7.5 million messages, it assesses the role of 'noise' and biases in online data collection. In a deep dive analyzing the Iranian nuclear agreement, we demonstrate that the middle east is not facing a nuclear arms race, and show that there is a structural hole in online discussion surrounding nuclear proliferation. By combining both approaches, policy analysts have a complete and generalizable set of computational tools to assess and analyze disparate stakeholder roles in nuclear proliferation.

Table of Contents

Table of Figures

Table of Tables

Chapter 1: Introduction

This thesis introduces computational methodologies to the field of policy analysis in nuclear proliferation, focusing on the role of social influence and positive and negative ties at both the national and sub-national level. There are two distinct components to this thesis: a simulation section that explores social science theories in nuclear proliferation, and a large-scale data analysis section that utilizes social media data. These two approaches speak directly to two sets of stakeholders in nuclear proliferation: governments and domestic actors. In the simulation section, the thesis utilizes historical political networks – alliances (positive ties) and hostilities (negative ties) to explore changes in behavior related to nuclear proliferation, and then draws out insights related to the current situation in Iran utilizing and expanding the historical model, focusing on governments and state actors. In the data analysis section, the thesis explores a dataset collected specifically to study nuclear discussion on Twitter and looks specifically at discussion related to the Iranian nuclear negotiations. This section addresses domestic stakeholders by directly assessing public opinion as well as the public account of domestic political figures. The section also explores the impact of context on sentiment analysis (positive and negative ties) on the Twitter platform; unlike standard Twitter collections, the nuclear Twitter dataset contains a high number of messages that are responses to messages. By combining these two approaches, policy analysts have a complete set of computational tools to assess and analyze disparate stakeholder roles in nuclear proliferation.

Table 1. Thesis chapter overview.

Chapter	Contribution	Approach	Data
2	Implementation of security model to nuclear proliferation and sensitivity testing	Adapted Friedkin model sensitivity analysis	Political Alliance and Hostility networks
3	Extension of security model to explore policy implications of security model	Adapted Friedkin model	Political Alliance and Hostility networks
4	Social media evaluation changes with context	In-person study and evaluation of social media messages with and without context	In-person study with 124 users utilizing Twitter data
5	In-depth Twitter analysis	Sentiment analysis, Latent Dirchlet Analysis, Social Network Analysis	Nuclear twitter dataset

Countries develop nuclear weapons for a variety of reasons, from concerns arising from security deficits to a commitment to norms and prestige surrounding nuclear weapons. This work broadly comes out of two literatures: nuclear deterrence and nuclear proliferation. Nuclear deterrence is traditionally contrasted with compellence – threats, as opposed to actions – and arguments for nuclear deterrence focus on relations between two nuclear states as opposed to actions between non-nuclear and nuclear states. Nuclear proliferation is concerned with the spread of nuclear material and ultimately, nuclear weapons, outside of the existing international regime outlined by the non-proliferation treaty (NPT). Unlike deterrence, the proliferation literature examines both internal domestic motivations for developing weapons in addition to motivations driven by external actors. We review the "demand" literature – exploring what motivates countries to develop nuclear weapons – as well as the "supply" literature for insights to develop policy.

The literature on nuclear deterrence emerges out of doctrine and policy developed in response to a bipolar nuclear world. Deterrence is commonly accepted to have evolved over at least three “waves” [1]: an initial wave, which explored the impact of nuclear weapons on world politics, to a second wave, which combined policy and theory, to a third wave, which highlighted empirical work. The second wave, which incorporated game theory models, such as the ‘Chicken Game,’ led to important insights about the nature of international relations, but did not contribute to direct policy implications: while it explained superpower relations, and framed broad strategic issues, it did not significantly contribute to smaller diplomatic and military efforts [2], [3] .The lack of empirical evidence made it difficult to evaluate claims made in the deterrence literature[4], which helped lead to the third wave’s emphasis of empirical work on risk taking, rewards, misperceptions, and bureaucratic politics[1], [5].

More recent work has expanded deterrence to other types of weapons of mass destruction (WMD). Traditionally, deterrence requires actors who are rational, resolute, and credible – all traits that rogue actors, such as North Korea, may not consistently demonstrate [6]. An alternate perspective is that the presence and threat of any type of WMD make deterrence easier [7]. The idea is that because the risks surrounding other WMD types is well understood, enough actors will behave rationally that we can observe deterrent behavior – even with a rogue actor in the system[8], [9]. Others have argued that the WMD threat makes it easier for rogue states to deter other actors, such as the United States, from involvement [10], [11].

These threat-based motivations for developing nuclear weapons are most commonly associated with a realist, or security-based motivation approach to developing nuclear weapons – one that focuses on nations as actors, and discounts the role of international institutions and internationally held norms[12]. There are two other major schools, which include domestic politics and constructivism [13]. The domestic politics school, which focuses on the roles of domestic actors, argues that the nuclear capability of a country can emerge from disparate actor politics, including domestic actors reacting to international actors differently, political economic ambitions, and nuclear ambivalence [14], [15] [16]. The constructivist school, which focuses on norms, argues that national leaders and identity play a major role in developing domestic nuclear policy agendas, including navigating a country's relationship to international norms [17]-[19].

Past work on quantitative models of nuclear proliferation has focused on generalized dynamic models incorporating the entirety of the historical data available, finding and identifying coefficients of global variables utilizing hazard models and generalized logistic regression models [20]-[22]. As good models, these approaches incorporate variables reflecting qualitative political science theories, including economic, institutional, and prestige indicator variables. However, recent work has found that the sign on estimated model coefficients relating to security concerns are not robust to small changes in data, such as mislabeled dyadic data, missing conflicts, and new information about past nuclear weapons programs [23], [24] [25]. Specifically, this indicates that we need a deeper understanding of the dynamics of security-driven motivations to pursue nuclear weapons.

We will focus on the mechanics of the security model, which has broad support, in developing our model for nuclear weapons [12]. In the security model, a country that has a nuclear enemy perceives a security deficit, and is thus motivated to acquire nuclear weapons [26]. The country may also seek an alliance with a nuclear power that promises to retaliate in case the country is attacked [27]-[29]. Such alliances provide reassurance for the country and reduce its need for developing indigenous nuclear weapons. With this approach, we find evidence of emergent international institutional behavior captured by this system coincident with the introduction of the Non-Proliferation Treaty (NPT), with the security model paradigm having greater explanatory power for nuclear proliferation in a world with the NPT than a world without the NPT.

On the “supply” side of nuclear proliferation, researchers have turned to assessing a country’s latent capability—the self-sufficiency of its nuclear industry[30]. Many of these assessments focus on broad industrial capacity [20], [31], [32]. Examples of such indicators are uranium deposits, steel production, and vehicle and radio production—all commercial signals. More recent work has expanded the definition of latent capability to signals in basic research and policy in addition to commercial activity; Kroenig finds that states receiving specialized nuclear assistance are more likely to develop nuclear weapons, and Fuhrmann finds that any type of nuclear assistance increases the probability of a country developing nuclear weapons [33], [34]. These more comprehensive latent capability indices, which place additional emphasis on access to highly enriched uranium and plutonium, allow for a clearer understanding of the role that deterrence plays in modern conflict [35], [36].

Three major technical reports—Swords from Plowshares, Harney, and a technical report from the Pacific Northwest National Laboratory (PNNL)—focus on key steps needed to develop an indigenous weapons program and the length of time involved in achieving full capability [37]; [38];[39]. By focusing on timelines, these papers implicitly highlight the organizational challenges involved in developing the necessary national institutions involved in nuclear weapons production, but instead explicitly only focus on institutional outcomes, the nuclear weapons technology. Furthermore, the focus on timelines obscures the difficulty of obtaining sufficient fissile material and the broader question of defining a ‘‘full capability’’[40]. Looking at weaponization exclusively, as these studies did, does not take into account the policy and commercial contexts that would arise that significantly impede further progress in developing the associated technology.

Understanding the interplay between supply and demand is crucial to developing more effective nuclear nonproliferation policy today. One of the more challenging aspects of studying this field, especially utilizing quantitative and analytical tools, is that nonproliferation policy has had some undeniable success over the years: despite the proliferation of nuclear energy technologies and material, relatively few states have made the decision to become a nuclear weapons state[41]-[44]. There is something of a catch-22 here: factors that contribute to proliferation, in being addressed by ongoing nonproliferation policy, make it difficult to study the quantitative relationship between that factor and nonproliferation.

The goal of this thesis is to connect the ongoing body of policy-oriented political science [45], [46] with policy-focused explorations of strategic issues [47], [48] [49]. The field is understandably, and correctly, worried about the impact of statistical "small-n" issues on interpretations of models, which make uncertainty analysis difficult [50]. This thesis introduces "large-n" approaches to the field to shift the focus to understanding changes in behavior in systems – away from focusing on statistically significant inference on sparse data, a feature of historical nonproliferation data. It introduces a dynamic Friedkin model of social influence utilizing political networks to infer security motivations to develop a nuclear weapons capability. This model is then extended to the current security context surrounding the middle east, and explores implications of this model surrounding Iran. The thesis then shifts focus to Iran and social media discussion. One of the features of the twitter discussion surrounding Iran is the fact that many posts in this dataset are reactive and reflective – a feature distinct from standard online social media. The thesis includes an analysis of some of the challenges in assessing sentiment of those types of social media posts.

These chapters are intended to be self-contained; there is some overlap in the background section between them. A technical report, containing the demographic information as well as the training materials utilized for the study in Chapter 4, is contained in this thesis as the first appendix. The second appendix is a listing of ego-networks for all countries in 2015 utilizing Correlates of War alliance data and International Crisis Behavior hostility data.

Chapter 2: Simulating Nuclear Proliferation Motivations

In this chapter we develop and test an adapted Friedkin model of opinion dynamics to evaluate security scenarios and country-level motivation to develop nuclear weapons. In doing so, we utilize political networks of alliances and hostilities, drawing on both non-proliferation policy and international relations literature to explore changes in behavior over time. We perform a large-scale sensitivity analysis of the adapted Friedkin model to observe changes in system behavior over time. We then consider three hypotheses to explore the implications of the adapted Friedkin model and future counter-proliferation work.

Background

Countries develop nuclear weapons for a variety of reasons, from concerns arising from security deficits to a commitment to norms and prestige surrounding nuclear weapons. This work broadly comes out of two literatures: nuclear deterrence and nuclear proliferation. Nuclear deterrence is traditionally contrasted with compellence – threats, as opposed to actions – and arguments for deterrence have focused on actions between two nuclear states as opposed to actions between non-nuclear and nuclear states. Nuclear proliferation is concerned with the spread of nuclear material and ultimately, nuclear weapons, outside of the existing international regime outlined by the non-proliferation treaty (NPT). Unlike deterrence, the proliferation literature examines both internal domestic motivations for developing weapons in addition to motivations driven by external actors.

The historical literature on nuclear deterrence emerges out of doctrine and policy developed in response to a bipolar nuclear world. Deterrence is commonly accepted to have evolved over at least three "waves" [1]: an initial wave, which explored the impact of nuclear weapons on world politics, to a second wave, which combined policy and theory, to a third wave, which highlighted empirical work.   The second wave, which incorporated game theory models, such as the ‘Chicken Game’, led to important insights about the nature of international relations, but did not contribute to direct policy implications: while it explained superpower relations, and framed broad strategic issues, it did not significantly contribute to smaller diplomatic and military efforts [2], [3] .The lack of empirical evidence made it difficult to evaluate claims made in deterrence literature[4],which helped lead to the third wave’s emphasis of empirical work on risk taking, rewards, misperceptions, and bureaucratic politics[1], [5].

Traditionally, deterrence requires actors who are rational, resolute, and credible – all traits that rogue actors, such as North Korea, may not consistently demonstrate [6]. An alternate angle, however, is that the presence and threat of any type of weapon of mass destruction make deterrence easier [7] – threatening an actor, combined with the crystallization of the risk posed by a WMD to other actors, can make it easier to respond to threats [8], [9]. Others have argued that the WMD threat makes it easier for rogue states to deter other actors, such as the United States, from involvement [10], [11].

These external motivations for developing nuclear weapons are most commonly associated with a realist, or security based motivation approach to developing nuclear weapons – one that focuses on nations as actors in a state of anarchy [12]. There are two other major schools, which include domestic politics and constructivism [13]. The domestic politics school, which focuses on the role of different domestic actors, argues that the nuclear capability of a country can emerge from disparate actor politics, including responding to international institutions, political economic ambitions, and nuclear ambivalence [14], [15] [16]. The constructivist school, which focuses on norms, argues that national leaders and identity play a major role in country motivations [17]-[19].

Past work on quantitative models and nuclear proliferation have focused on generalized dynamic models incorporating the entirety of the historical data available, finding and identifying coefficients of global variables utilizing hazard models and generalized logistic regression models [20]-[22]. As good models, these approaches incorporate different variables reflecting these theories, incorporating economic, institutional, and prestige indicator variables.in However, recent work has found that these model coefficients relating to security concerns are not robust due to small changes in data, from mislabeled dyadic data to missing conflicts [23], [24]. Specifically, this indicates that we need a deeper understanding of the dynamics of security-driven motivations to pursue nuclear weapons.

Maoz has found some limited support for balanced triads – namely, that states with common enemies tend to ally with each other, states that are indirect enemies also tend to conflict with each other. Since not all of these may be directly recorded in the immediate data that we utilize, we have to infer some additional links to test these hypotheses.

The two primary quantitative analyses of the historical data, Jo & Gartzke and Singh & Way, take two different approaches to tracking national nuclear weapons programs. Jo & Gartzke distinguish only between having nuclear weapons programs and possession of nuclear weapons. Singh & Way allow for additional nuance, allowing for countries to transition from no nuclear weapons program to exploring, from exploring to pursuit, and from pursuit to acquisition.

Table 2. Table of timelines of nuclear weapons programs in Jo & Gartzke and Singh & Way, adapted from Montgomery & Sagan. Bolded entries in Jo & Gartzke come from additional sources in Montgomery & Sagan.

	Jo & Gartzke (1941-2002)		**Singh & Way (1945-2000)**		
Acquired States	**Programs**	**Possession**	**Explore**	**Pursue**	**Acquire**
USA	1942-	1945-	*	*	1945-
Russia	1943-	1949-	*	1945-	1949-
UK	1941-	1952-	1945-	1947-	1952-
France	1954-	1960-	1946-	1954-	1960-
China	1956-	1964-	1955-	1955-	1964-
Israel	1955-	1966-	1949-	1958-	1972-
India (1)	1964-65		1954-	1964-	1974-74
India (2)	1972-	1988-	1975-	1980-	1988-
South Africa	1971-90	1979-91	1969-	1974-	1979-93
Pakistan	1972-	1987-	1987-	1972-	1990-
Exploring or Pursued States					
South Korea	1971-75		1959-	1970-78	
Libya	**1970-03**		1970-	1970-	
Brazil	1978-90		1953-	1978-90	
North Korea	1982-		1965-	1980-	
Iraq	1973-02		1976-	1982-	
Iran (1)	1974-48				
Iran (2)	1984-		1984-	1985-	
Argentina	1976-90		1968-	1978-90	
Germany	1941-45				
Japan	1943-45				
Switzerland			1956-69		
Sweden	1946-69		1954-69		
Yugoslavia (1)	1953-63		1954-65		
Yugoslavia (2)	1982-87		1974-88		
Australia	**1956-72**		1956-73		
Taiwan (1)	1967-76		1967-77		
Taiwan (2)			1986-88		
Algeria			1983-		
Romania	1981-89		1985-93		
Egypt	**1960-67**				

Model structure

Traditionally, Friedkin has been used to show the opinion dynamics of a small group [51]. This approach has been used to model opinion changes among political stakeholders in major processes such as voting in the EU and extremist behavior [52] . The basic concept is simple: individual opinions change after interacting with others through a social network. Operationalizing it initially seems intimidating, but the fundamental approach is still fairly simple.

Unlike past applications of the Friedkin model where only one type of network is considered, we modify it to reflect the different levels of motivation enforced by different networks. We are influenced not only by individuals in our immediate social friend network, but also by the actions of those we consider to be competitors or those hostile to us. In an international relations setting, we have two sets of networks, which we might expect to have very different influences on our behavior: an alliance network and a hostility network.

At a high level, the Friedkin equation (equation 1) consists of 3 parts. Wy_{t-1} represents actors' extrinsic attitudes resulting from external influence, y_1 represents intrinsic attitudes that reflect actors' own characteristics and constraints, and A represents the relative weight that actors place on extrinsic attitudes. The Friedkin model can apply to any attitude. In this work, we consider the attitude to be the motivation to develop nuclear weapons.

Equation 1. Friedkin model equation

$$y_t = AWy_{t-1} + (I - A)y_1$$

In the equation, y_t is an $N \times 1$ vector that represents country attitudes at time t. The attitude of each country follows scaling given in Figure 1, where 0.5 represents an indifferent attitude, larger values represent a positive attitude and smaller values represent a negative attitude. $A = diag(a_{11}, \dots, a_{ii}, \dots, a_{NN}), 0 \leq a_{ii} \leq 1$ is a $N \times N$ diagonal matrix with diagonal weights indicating the level of influence that each actor puts on outside actors. $W = [w_{ij}], (0 \leq w_{ij} \leq$

$1, \sum_{j}^{N} w_{ij} = 1)$ is an $N \times N$ matrix that represent inter-actor influence. More specifically, w_{ij} represents the extent to which actor j has on actor i. W is computed using the formula $W = AC + I - A$ where $C = [c_{ij}]$ is a $N \times N$ matrix of relative interpersonal influence such that ($c_{ii} = 0, 0 \leq c_{ij} \leq 1, \sum_{j=1}^{N} c_{ij} = 1$). Finally, y_1 is a $N \times 1$ vector representing actors' initial attitudes.

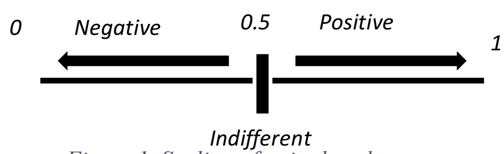


Figure 1. Scaling of attitude values

The extrinsic attitude term Wy_{t-1} can only capture influence from one class of tie between two actors. We modify this to represent two kinds of extrinsic influence for two kinds of networks: alliances and hostilities. In order to simplify the discussion, we initially consider a single country that has a single enemy and a single ally, and derive a new term for the extrinsic motivation. Subsequently, we modify that expression into a vectorial expression that captures the extrinsic motivation of all countries. Finally, we include that vectorial expression into Equation 1, obtaining the modified equation model.

Based on past work examining historical trends in nuclear proliferation, we find coefficients α_1, α_2 and α_3 in Equation 2 that reflect current behavior of countries interested in developing nuclear weapons. At the country level, we describe motivation as the linear sum of the effect of hostilities, alliances, and the interaction of alliances and hostilities.

Equation 2. Extrinsic motivation coefficient equation

$$m_t = \alpha_0 + \alpha_1 e_{t-1} + \alpha_2 f_{t-1} + \alpha_3 e_{t-1} f_{t-1}$$

This motivation applies to only one country. We are interested in obtaining a vectorial expression that simultaneously captures the motivation of all states. We now assume that M_t, E_{t-1} and F_{t-1} are Nx1 vectors. To preserve these vectors, we now place values of F along a diagonal matrix.

Equation 3. Vectorial form of extrinsic motivation

$$M_t = \alpha_1 E_{t-1} + \alpha_2 F_{t-1} + \alpha_3 (E_{t-1})\ diag(F_{t-1})$$

Given that E_{t-1} captures whether countries' enemies have nuclear weapons at time t-1, and F_{t-1} captures whether allies have nuclear weapons, we can write $E_{t-1} = W_H y_{(t-1)}$ and $F_{t-1} = W_F y_{(t-1)}$, where W_H captures international hostilities and W_F captures international alliances. Finally, we obtain

Equation 4. Adapted Friedkin Model Equation

$$y_t = A\,[\,\alpha_1 W_H y_{t-1} + \alpha_2 W_F y_{t-1} + \alpha_3 W_H\, y_{t-1}\, diag(W_F y_{t-1})] + (I - A) y_1$$

Comparing the initial Friedkin model and the adapted Friedkin model, the primary change has been allowing for two types of networks – as well as a third interaction term – to exist within the W term.

We utilize the model in the following way:

1. Obtain GDP of all countries for the year being studied, and calculate 1-log(GDP_i)/log(GDP_{max}). These are the values of diagonal matrix A.
2. Union the past 10 years of alliances and hostilities networks. These are the alliance and hostility networks W_F, W_H.
3. Determine the coefficients $\alpha_1, \alpha_2, \alpha_3$ that you will utilize for this run.
4. Identify the countries that have nuclear weapons and set their values in the y_1 vector to 1; set all other values of y_1 to 0.5
5. Iterate the model. The first few iterations will show some dramatic changes in opinion; however, these changes will settle after the first few iterations, with increasingly smaller marginal differences between each iteration after the 10th iteration. For consistency, the results in this study utilize the 30th iteration.
6. Evaluate the 30th iteration: if a country is 'Pursuing' nuclear weapons, or is listed as having 'Acquired' nuclear weapons at any point over the following 5 years, it should

have a score greater than 0.5 – this contributes to 'True Positives'. If a country has a score greater than 0.5 but is not in either category, it is a 'False Positive'. If a country has a score below 0.5 but is 'Pursuing' nuclear weapons or is a nuclear weapons state, it is scored as being a 'False Negative'.

7. Repeat for all possible combinations of coefficients $\alpha_1, \alpha_2, \alpha_3$
8. Utilizing True Positive, False Positive, and False Negative counts, calculate F1 scores and identify the range of coefficients for the highest-scoring F1 runs.

The next section goes into further detail about the rationale for utilizing these inputs to the adapted Friedkin model.

Table 3. Summary table of data utilized for Friedkin model

Model Variable	Dimensions	Description	Modification	Source
A	*n x n*	Diagonal matrix indicates how much country *i* weighs opinions of others in its network	1-log(GDP_i)/log(GDP_{max})	National GDP from World Bank Historical GDP Data
W_H	*n x n*	Network of hostilities	Normalized by column sum	International Crisis Behavior, Uppsala Conflict Data Program
W_F	*n x n*	Network of alliances		Correlates of War
y_1	*n x 1*	Country initial motivation to develop nuclear weapons	Set to 0.5 if country is a non-nuclear weapons state; 1 if country is a nuclear weapons state	Acquired

Data and Implementation

There are four primary sources of data utilized for the adapted Friedkin model: world GDP data, for matrix A, the political alliance and hostility networks, and the initial set of countries with nuclear capabilities. These are summarized in the table below.

Influence ceded to others

To model the weight given to other individuals in the network, A, we utilize a normalized version of the country's GDP. We utilize the ratio of the log of country GDPs; we anticipate that countries with smaller GDPs are more sensitive to pressures from alliances and hostilities than countries with larger GDPs. The log of the country's GDP, rather than the strict ratio of a country's GDP, is used to spread out the magnitude of different levels of GDP; otherwise, most countries are simply dwarfed by the scale of America's overall GDP. This allows for larger regional economic powerhouses, such as Germany and Japan, to not simply cede everything to external influence – for example, consider Germany's GDP (roughly 3.8 trillion USD) against the US GDP (roughly 17 trillion USD). Without the log normalization, the model would attribute 80% of Germany's opinion on nuclear proliferation to external networks; with the log normalization, the model assumes that roughly 50% of Germany's opinion on nuclear proliferation is determined by its external networks.

A country with a larger GDP anticipating economic retaliation for changing its position on weapons of mass destruction will be more insulated from an anticipated marginal shock than a country with a smaller GDP, where the equivalent nominal economic shock will have a much higher marginal value. In our runs, we utilize historical GDP data from the World Bank[53]. For countries where there is missing World Bank data (i.e. parts of the USSR, African states as they declare independence), the median GDP for the time period is used instead.

Alliance and Hostility Networks

The political alliance network is based on the formal alliance network from the Correlates of War (COW) project, using the interstate alliance data set v4.1[54]. The data set distinguishes between four kinds of treaties: a defense pact, a neutrality pact, a non-aggression pact, and ententes.

For our analysis, we consider all forms of alliance to indicate an alliance. For the historical models, active alliances over the past ten years are considered; i.e. for the 1960 run, all alliances in effect from 1950-1960 are considered.

For the hostilities network, we use two sources: the International Crisis Behavior (ICB) project at the University of Maryland and the Uppsala Conflict Data Program (UCDP) at the University of Uppsala, Sweden[55], [56]. The ICB project data covers violent and non-violent conflicts during the period 1918-2013. The UCDP data covers violent conflicts that caused at least 25 deaths in a calendar year during the period 1993-2014. We consider all state-to-state hostilities covered in both datasets over the past ten years; i.e. for the 1960 run, all hostilities from 1950-1960 are considered.

In the Friedkin model, we modify both the Alliance and Hostility networks so that the column sums of the matrix are equal to 1. This is counterintuitive, as typically the influence that country i exerts over country j is normalized by row i. However, in the Friedkin influence equation, these networks are used as inputs to describe the amount of influence that country j exerts over country i – so the column sum is used. This is especially important given the relative centrality of key actors such as the United States. For many countries, the US is the only country in its hostility or alliance network. If these networks were normalized by row, instead of by column, because the US is involved in so many alliances and hostilities, the model would show that the US exerts relatively little influence in the system.

For the extended networks in each time period, we utilize matrix multiplication to identify friends-of-friends, a common approach in social network analysis. This approach utilizes a trick to combine the two distinct networks first; as a result it differs slightly from the approach used in Maoz [57] as we are primarily utilizing symmetric alliance and hostility matrices. We first illustrate the method in pseudocode and then demonstrate its efficacy in a three key examples. *A* indicates the Alliance network; *H* indicates the Hostilities network; they have the same size, are symmetric, and all elements are positive. *EA* and *EH* represent the Extended Alliance and Extended Hostilities networks respectively.

```
[Combined] = [A]-[H]
[Extended] = [Combined]^2
[EA] = [A] ; [EH] = [H]
For element (i,j) in [Extended]:
If element (i,j) > 0: [EA]_ij = 1
If element (i,j) < 0: [EH]_ij = 1
Set diagonals of [EA], [EH] = 0
```

Example 1:

For example, suppose there are three countries A, B, and C. An alliance exists between A and B, as well as B and C. In a network, we would describe this in a matrix as:

$$Alliance = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

With countries A, B, and C corresponding respectively to rows and columns 1, 2, and 3.

Taking the square of this, we obtain:

$$Alliance^2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Going through this element by element, and adding positive elements to the existing alliance, we get the transitive alliance between A and C.

$$\therefore Extended\ Alliance = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Example 2:

Instead of the ties between A and B and B and C being alliances, consider them as hostilities.

$$Hostility = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Combining this with the (empty) alliance matrix, we obtain:

$$Combined = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Taking the square of this matrix, we obtain:

$$Combined^2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Going through this element by element, we now have the following extended alliance and extended hostility networks:

$$Extended\ Alliance\ = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}\ Extended\ Hostility\ = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The new tie between A and C can be considered an extension of "the enemy of my enemy is my friend".

Example 3:

Finally, we examine an unbalanced triad: assume an alliance exists between A and B, and a conflict exists between B and C.

$$Alliance\ = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\ Hostility\ = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Then:

$$Combined = Alliance - Hostility = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$\therefore Combined^2 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

Leading to:

$$Extended\ Alliance\ = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\ Extended\ Hostility\ = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

This can be interpreted as extending the hostility network; "the enemy of an ally is also an enemy".

Dynamic implementation and hypotheses

A primary concern of past work in analyzing past proliferation trends focuses on regression models, left-censoring, and over-representing risks of proliferation. While these models are useful analyses of historical events, they also implicitly assume that behavior is identical over time. An alternative approach is to directly simulate the security context that each state finds itself in over time, compare relevant simulations to the historical record, observe emergent behavior, and draw policy conclusions to determine if behavior changes over time.

In the simulation, we model the security context for countries over the past ten years and use it to try and predict the motivation of countries to develop nuclear weapons over the next five years. For each simulation run, we consider the initial motivation of a country at time y_1 to pursue nuclear weapons as 1 if the country possesses nuclear weapons and 0.5 otherwise as in Figure 1 above. As the Friedkin model runs towards consensus, we run the simulation 30 times to identify the outcomes of each time period's security context on country motivations.

As we are working with a very sparse dataset of countries that have in fact proliferated we have to perform an exhaustive search over the range of coefficient space to identify optimal models that have the "best fit" for historical data. The table below specifies the range and grid size of coefficients used.

Equation 5. Adapted Friedkin Model Equation

$$y_t = A\,[\,\alpha_1 W_H y_{t-1} + \alpha_2 W_F y_{t-1} + \alpha_3 W_H\, y_{t-1}\, diag(W_F y_{t-1}\,)] + (I - A)y_1$$

Table 4. Table of coefficients and grid size used for each model run

Coefficient	Description	Range, Grid Size
$\boldsymbol{\alpha_1}$	Hostility Coefficient	[-1,1], by 0.1
$\boldsymbol{\alpha_2}$	Alliance Coefficient	[-1,1], by 0.1
$\boldsymbol{\alpha_3}$	Interaction Coefficient	[-.5,.5], by 0.1
$\boldsymbol{W_H, W_F}$	Hostility, Alliance networks	Taken from past 10 years of alliance, hostility data

Runs are evaluated based on F1 scores, which place equal weight on precision and recall. At the end of each run, the model is evaluated with countries having motivation above 0.5 as being countries pursuing nuclear weapons, and countries with motivation below 0.5 as not being motivated to pursue nuclear weapons.

To evaluate the effectiveness of this model, in addition to the existing nuclear powers at the time, the model had to identify countries that were "Pursuing" (when using Singh & Way data) at any point or had "Programs" (when using Jo & Gartzke data). These countries had to be effectively pursuing nuclear weapons at any point over the next five years. For example, the model in 2000 was evaluated based on whether it detected, in addition to existing nuclear powers, North Korea, Libya, and Iran, all countries that according to Singh & Way were pursuing nuclear weapons from 2000-2005. If a country ended its program in that year, it was not considered as part of the evaluation criteria; for example, while the 1985 model was evaluated based on whether it detected Argentina and Brazil's programs, since both countries ended their programs in 1990, the 1990 model is not evaluated based on whether Argentina and Brazil are detected.

We consider three hypotheses about these models:

1. Do these models hold up in robustness when using different sources of data?
2. Do these models provide evidence for seeing extended alliance and deterrence activity in nuclear proliferation?
3. Do these models perform better when trying to predict countries that are interested in exploring nuclear weapons as opposed to detecting just countries that are actively pursuing nuclear weapons?

To explore the first hypothesis, we run the model using Singh and Way data and contrast it with a modified version of Jo & Gartzke's data, considering Libya, Australia, and Egypt's programs [58] [59] [60].

To explore the second question, we run Singh and Way data and explore the impact on the coefficients when using extended alliance, extended hostility, and extended alliance and hostility networks as inputs.

To explore the third question, we explore comparing inputs and outputs in two ways: we first expand the set of countries that we evaluate models on, to see if the models can identify countries that are both "pursuing" and "exploring" nuclear weapons. We also examine how well the model can identify countries "exploring" nuclear weapons if we consider "acquisition" and "pursuing" countries to have an initial motivation of 1 in the models.

Results

Each summary table of results contains the following statistics: the maximum F1 score of the model that ran for that time period, the median coefficients for Hostility, Alliance, and Interaction for models with the highest F1 score, a count of the number of models with the maximum F1 score, and a percentage of the overall gridded coefficient space that the "best fit" model matches with.

The first question speaks to the robustness of the overall model. It should be noted that the two datasets are most distinct at the beginning of the timeline; while the best fit models in 1960 do not overlap between Singh and Way and Jo & Gartzke, the later dates do have significant overlap.

Table 5. Summary table of models utilizing Singh & Way data

	Max F1	Hostility	Alliance	Interaction	Count	% of Grid
1960	0.6153846	-0.2	0	-0.1	3893	80%
1965	0.625	-0.1	-0.1	-0	4380	90%
1970	0.5714286	1	-0.3	0.5	3	0%
1975	0.6956522	-0.1	-0.1	0	4233	87%
1980	0.64	0.4	0.8	0.4	55	1%
1985	0.6428571	1	1	-0.2	3	0%
1990	0.8148148	0.7	0.7	0.3	197	4%
1995	0.8148148	1	1	0.3	5	0%
2000	0.7619048	-0.4	0	-0.1	2797	58%

Table 6. Summary table of model utilizing Jo & Gartzke corrected data

Year	Max F1	Hostility	Alliance	Interaction	Count	% Of Grid
1960	0.5263158	0.7	0.6	0.5	61	1%
1965	0.6315789	0.8	0.5	0.4	141	3%
1970	0.5454545	-0.2	0	-0.1	4020	83%
1975	0.5	-0.1	-0.1	0	4235	87%
1980	0.5925926	0.4	0.8	0.4	55	1%
1985	0.6	1	1	-0.2	3	0%
1990	0.8148148	0.7	0.7	0.3	197	4%
1995	0.8148148	1	1	0.3	5	0%
2000	0.7826087	0.2	-0.1	0.1	553	11%

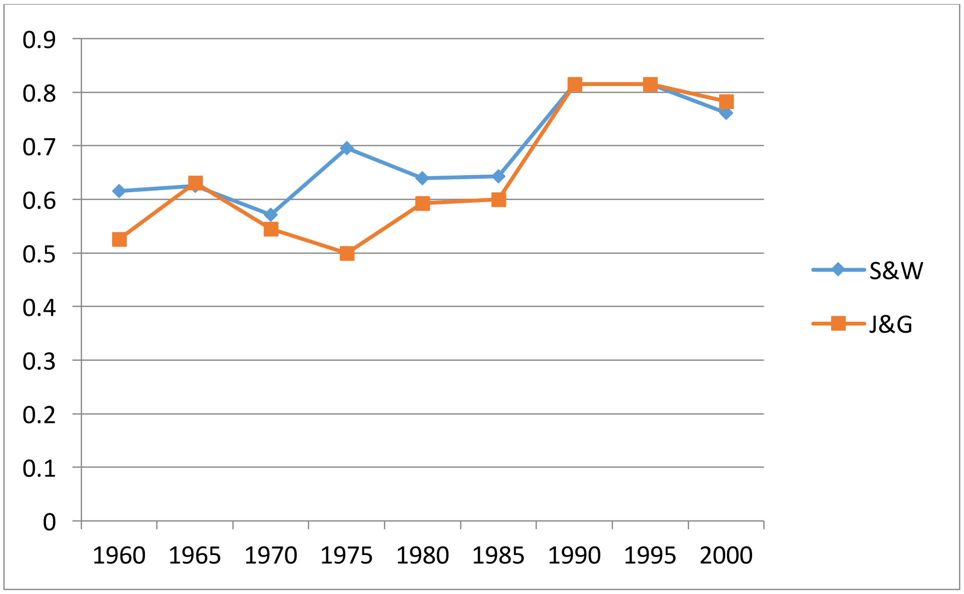


Figure 2. Graph comparing maximum F1 scores of Singh and Way (S&W) and Jo & Gartzke (J&G) data

We find there is weak evidence supporting looking at the extended networks. Strictly looking at the maximum F1 scores, the model that uses both extended networks does a slightly better job of predicting nuclear programs compared to using either extended network alone or the baseline Singh & Way data. However, this very small level of increased accuracy suggests that the second order effects of extended hostilities are already demonstrated in the results of the Friedkin model.

In considering the improved accuracy of extended networks, the majority of the increased accuracy appears to come from utilizing the extended hostility network. This is probably due to the way the hostility network is transformed in looking at extended hostilities; the original hostility network is relatively sparse, while the alliance network has relatively dense components (as it takes into account large military alliances such as NATO and the Arab League). The extended hostility network therefore is changed much more significantly when compared to the original hostility network. In contrast, the extended alliance network is not significantly different from the original alliance network, even though it considers both “enemy of my enemy is my friend" and "friends of friends are friends" types of connections; the former set of new links is small as the hostility network is sparse, and the latter set of new links is relatively small as the alliance network has distinct components.

A significant difference in considering the results of extended networks compared to the baseline is the changed value of alliances. Unlike the baseline Singh and Way data, which clearly highlights the role that alliances can play in promoting proliferation in a post NPT world, the extended networks tell a slightly different story, with alliances only sometimes contributing to nuclear proliferation. While there is a similar emergent NPT effect in a smaller number of coefficients matching the top F1 score, the onset of the smaller set of coefficients takes place in 1985, not in 1980 – suggesting that while the NPT was successful in promoting a nonproliferation norm, it might not have spread as quickly as suggested in the baseline model.

Table 7. Summary table of Model utilizing Extended Alliance network

	Max F1	Hostility	Alliance	Interaction	Count	% Of Grid
1960	0.6153846	-0.2	-0.1	-0.1	3820	79%
1965	0.625	-0.1	-0.1	0	4369	90%
1970	0.5555556	-0.2	-0.1	-0.1	3936	81%
1975	0.6956522	-0.2	-0.1	-0.1	4145	85%
1980	0.6153846	0.8	0	0.3	216	4%
1985	0.6428571	1.0	0.9	0.2	9	0%
1990	0.8148148	0.8	0.6	0.4	58	1%
1995	0.7826087	0.6	0	0.1	2051	42%
2000	0.7619048	-0.4	0	-0.1	2800	58%

Table 8. Summary table utilizing Extended Hostility network

	Max F1	Hostility	Alliance	Interaction	Count	% Of Grid
1960	0.7272727	-0.8	-0.2	-0.3	590	12%
1965	0.7142857	-0.7	-0.1	-0.2	1384	29%
1970	0.625	-0.7	0.1	-0.1	925	19%
1975	0.7619048	-0.8	-0.25	-0.3	588	12%
1980	0.6363636	-0.8	-0.2	-0.3	699	14%
1985	0.6428571	0.9	-0.4	-0.4	8	0%
1990	0.8461538	0.7	0.6	-0.2	40	1%
1995	0.88	0.5	0.4	0	426	9%
2000	0.7619048	-0.5	0	-0.1	2761	57%

Table 9. Summary table utilizing Extended Alliance and Hostility networks

	Max F1	Hostility	Alliance	Interaction	Count	% Of Grid
1960	0.7272727	-0.8	-0.2	-0.3	635	13%
1965	0.7142857	-0.7	-0.1	-0.2	1406	29%
1970	0.625	-0.6	-0.1	-0.2	1008	21%
1975	0.7619048	-0.8	-0.3	-0.3	644	13%
1980	0.6363636	-0.8	-0.4	-0.3	769	16%
1985	0.6666667	0.25	-0.75	0.35	6	0%
1990	0.8461538	0.7	0.8	-0.2	24	0%
1995	0.88	0.5	0.4	0.1	421	9%
2000	0.8421053	-1	-0.9	-0.5	21	0%

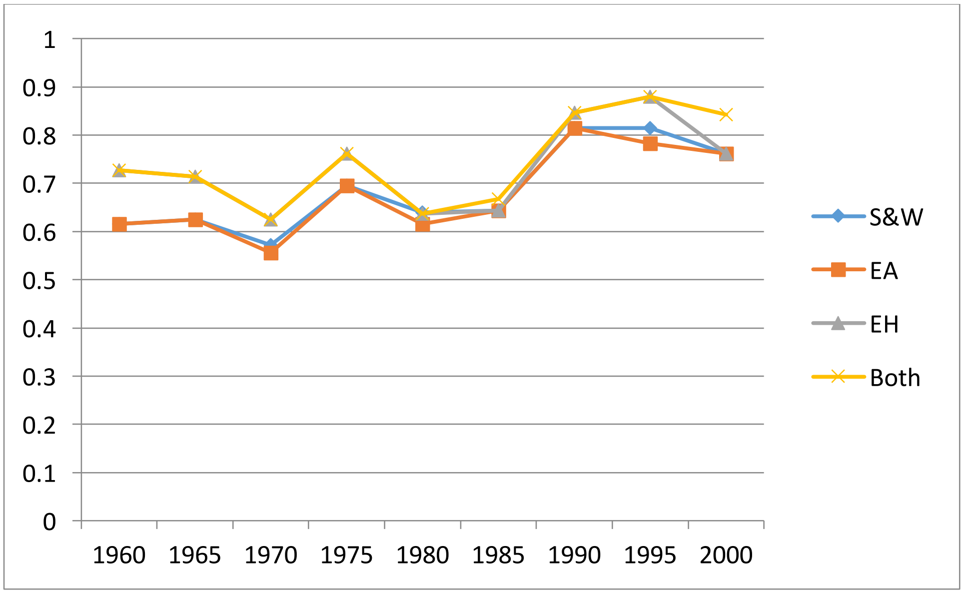


Figure 3. Graph comparing maximum F1 scores for models using Singh & Way, Extended Alliance (EA) network input, Extended Hostility (EH) network input, and both EA & EH networks input.

Overall, trying to predict both "Exploring" and "Pursuing" countries significantly decreases model F1 scores. This suggests that countries may enter the "Exploring" phase of starting to consider nuclear weapons for reasons beyond their security context. We see this again when trying to predict "Exploring" countries directly, in the next set of models.

Table 10. Summary table of model looking at predicting "Seeking": both "Exploring" and "Pursuing"

	Max F1	Hostility	Alliance	Interaction	Count	% Of Grid
1960	0.5	0.7	-0.1	0.2	811	17%
1965	0.4347826	-0.1	-0.1	0	4380	90%
1970	0.4444444	1	-0.3	0.5	3	0%
1975	0.5517241	-0.1	-0.1	0	4233	87%
1980	0.5714286	0.4	0.8	0.4	55	1%
1985	0.6956522	-0.1	1	0.3	7	0%
1990	0.7586207	0.7	0.7	0.3	197	4%
1995	0.7857143	1	1	0.3	5	0%
2000	0.7619048	-0.4	0	-0.1	2797	58%

Here it should be noted that while the overall F1 scores are higher for predicting "Exploring" while setting initial values to 1 for countries that are both "Pursuing" and have "Acquired" nuclear weapons, the overall number of cases is much higher. This suggests that the security model does not do a good job explaining these cases overall.

Table 11. Summary table of model looking at predicting Singh and Way "Exploring"

	Max F1	Hostility	Alliance	Interaction	Count	% Of Grid
1960	0.6666667	0.7	-0.2	0.3	793	16%
1965	0.6428571	-0.1	-0.1	0	4320	89%
1970	0.7333333	-0.2	-0.1	-0.1	3911	81%
1975	0.8	-0.2	0	0	4149	86%
1980	0.8571429	-0.2	0	-0.1	3902	80%
1985	0.8823529	-0.3	-0.1	-0.1	3437	71%
1990	0.8666667	-0.3	-0.1	-0.1	3505	72%
1995	0.8888889	0.3	0	0.1	3306	68%
2000	0.88	-0.4	0	-0.1	2796	58%

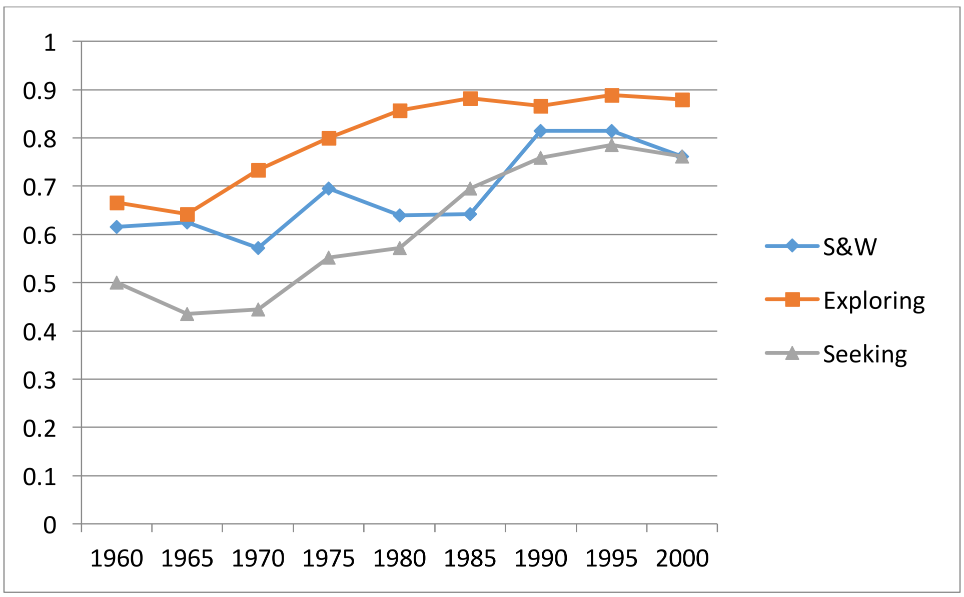


Figure 4. Graph comparing maximum F1 scores for Singh & Way, Exploring, and Seeking.

Discussion

We find some emergent effects of the NPT – signed in 1969 by major nuclear powers and adopted by countries around the world largely through the 1970s. In all of these models, focus on the models in 1980 and beyond; these model runs incorporate networks of alliances and hostilities in a world after the NPT was signed, as the models incorporate all alliances and hostilities looking back over 10 years.

Looking at the baseline Singh and Way models in 1980 and beyond, we find that alliances tend to be positive, instead of negative, contributors to a country's motivation to develop nuclear weapons. This suggests that countries receiving sensitive nuclear assistance, as suggested by Koenig and Fuhrmann, are driving modern proliferation efforts.

This is not to discount security concerns and motivations – in cases and models where alliances contribute to motivation, it should be noted that security concerns and hostilities also contribute to country motivations to develop nuclear capabilities.

However, it should be noted that before the NPT, it looks like there were several additional motivations besides a security model for developing nuclear weapons. Consider that if a country is strictly motivated by its security context, the time period's best run should occupy a relatively small percent of the overall gridded coefficient space.

Table 12. Percentage of overall grid space occupied by "best fit" runs.

	S&W	J&G	EA	EH	EA & EH	Seeking	Exploring
1960	80%	1%	79%	12%	13%	17%	16%
1965	90%	3%	90%	29%	29%	90%	89%
1970	0%	83%	81%	19%	21%	0%	81%
1975	87%	87%	85%	12%	13%	87%	86%
1980	1%	1%	4%	14%	16%	1%	80%
1985	0%	0%	0%	0%	0%	0%	71%
1990	4%	4%	1%	1%	0%	4%	72%
1995	0%	0%	42%	9%	9%	0%	68%
2000	58%	11%	58%	57%	0%	58%	58%

Before 1980, both Singh & Way and Jo & Gartzke data suggest that there are alternate explanations for why countries are developing nuclear weapons – the 'best fit' models capture a majority of the entire gridded space. After 1980 – and in a world with the NPT – we see better evidence for countries wanting to develop nuclear weapons strictly due to the security model.

This analysis does limit itself to states as unitary actors. It also focuses only on states, when we know that there are domestic actors, as well as sub-state actors, that drive security considerations and proliferation decisions. It does not consider the terror associated with a 'dirty' bomb, or the technical capabilities required for missile delivery systems. But it shows that the historical motivation for developing nuclear weapons cannot only be described by the security context.

Another limitation of the analysis is its focus on a general nuclear weapons capability, which does not distinguish between a broad second-strike capability (such as Trident), having a hydrogen bomb (significantly more powerful), or having tactical nuclear weapons ('micro-nukes'). To extend the analysis to each of these different types of weapons, a very similar approach could be utilized; instead of evaluating and initializing the Friedkin model on having a nuclear weapons program, the focus should be on delivery systems and platforms. The focus on distinct platforms and actual delivery capability distinguishes some of the different nuclear proliferation timelines available, such as Singh and Way and Jo and Gartzke.

The analysis does a thorough sensitivity analysis of different coefficients utilized in the Friedkin model. However, it also takes as a given existent political alliance and hostility networks. There is the potential for 'noise' in these input networks, especially in the alliance network – as alliances may exist on paper but may not address cultural hostilities. The effect of these 'paper' alliances is mitigated by the hostility network and the interaction coefficient in the

Friedkin equation. These alliances also may not reflect actual operational alliances, such as ongoing intelligence sharing between the US, Canada, the UK, Australia, and New Zealand, or backchannels of communication used between the United States and rogue states such as North Korea.

The primary impact of these missing types of alliances is to highlight the different degrees to which states are embedded in broader security alliances – missing operational alliances may play a role in overstating some country motivations to develop a nuclear capability. However, missing backchannel alliances do not play a significant role in the analysis as we are interested in assessing and simulating country security contexts – while backchannels provide insight into favorable diplomatic alliances, they do not consistently reflect security reassurances.

It is important to highlight that this is a global model, highlighting global trends in nuclear proliferation. For using this model in a policy application, I recommend focusing on a regional set of actors and changing the behavior of different networks to determine how those changes affect the subset of regional actors – keeping in mind that the entirety of the global system is being simulated. A sample of this type of analysis is presented in the next chapter, which focuses on the Middle East and Iran in particular. One of the major questions to explore in a policy analysis is the impact of alliances on motivations to develop a nuclear capability to determine if a nuclear 'arms race' appears imminent in the region.

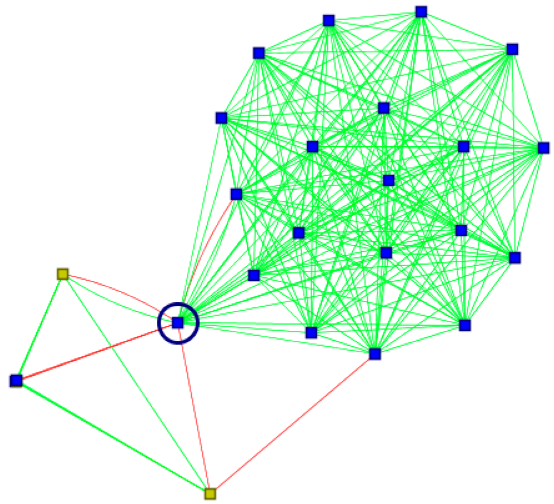


Figure 5. Stylized ego network of state (circled) motivated to develop nuclear capability. Yellow node indicates nuclear weapons state.

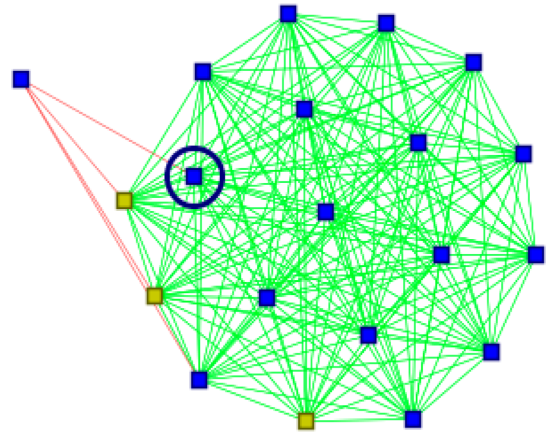


Figure 6. Stylized ego network of country (circled) not motivated to develop nuclear capability. Yellow node indicates nuclear weapons state.

One of the conclusions reached in this analysis is that alliances play a significant role in developing nuclear weapons. To provide some additional stylized examples, the first figure above shows a country motivated to develop nuclear weapons due to both its alliances and its hostilities with nuclear weapons states, while the second figure above shows a country not motivated to develop nuclear weapons despite being embedded in an alliance with nuclear weapons states. In the former, the country at risk has both an alliance and a hostility with a nuclear weapons state, as well as a broad alliance with non-nuclear countries and hostilities with two other nuclear weapons states and conventional weapons states. It faces both an alliance as well as hostility from nuclear weapons, which in the model contributes to its motivation to develop a nuclear capability. In contrast, the second example shows a country very well embedded in a defense alliance with some nuclear weapons states. It does not face a threat from a nuclear weapons state; despite having nuclear weapons powers as part of its alliance network, because it does not also face a nuclear security hostility, the model does not show the country as being motivated to develop a nuclear capability.

There are very few countries that have a similar match to the former type of country currently – one that is well positioned and motivated to develop a nuclear weapon. The countries that currently closest fit this profile are Syria, Libya, Iran, Ukraine, and South Korea – all simply because they are closely tied to countries that have nuclear weapons in both their alliance and hostility networks. Most countries do not appear in the hostility network data, and the ones that appear in alliance network data are primarily in regional military alliances. In the most recent set of ICB conflict data, for example, India and Pakistan are primarily concerned with each other.

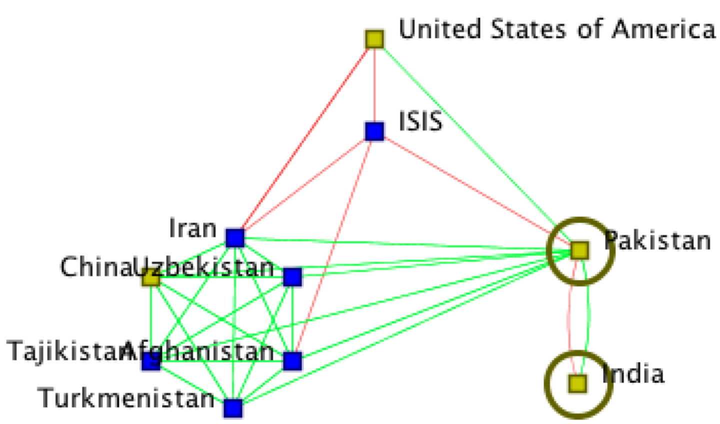


Figure 7. India-Pakistan dyad ego-net. Red indicates hostile tie; green indicates alliance.

And for the opposite scenario, Denmark appears firmly ensconced in the NATO alliance; highly unmotivated to develop a nuclear capability despite being in an alliance with three major nuclear powers. As the majority of other countries in the alliance do not have nuclear powers, combined with the fact that there is no hostility with a nuclear weapons state, other members of NATO in the Friedkin model are not motivated to develop a nuclear capability.

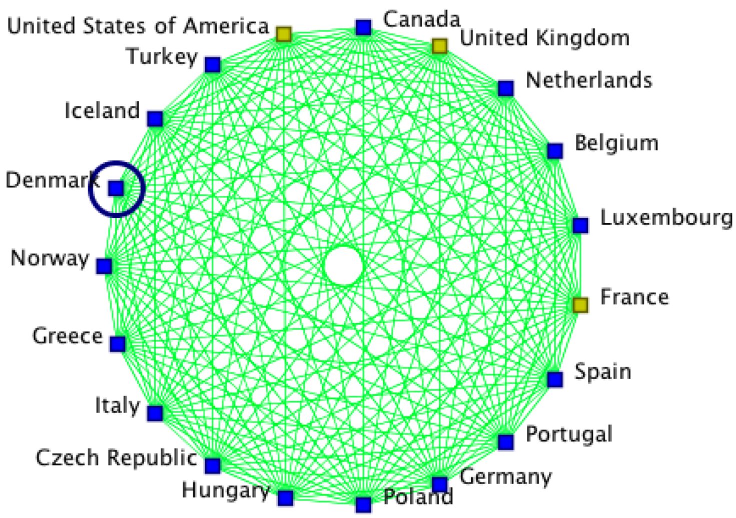


Figure 8. Denmark's ego-net. While heavily involved in NATO missions, there are no direct hostility ties.

This work suggests that current counter-proliferation efforts should focus on supplies, alliance networks, and special assistance. Next in-depth analyses should focus on the spread of nuclear scientific and technical capabilities, regional and sub-state actor contributions to motivations, as well as discerning potential signals from broader populations. A later chapter aims to do this by analyzing a large Twitter data set focused on discussions of nuclear technologies.

Chapter 3: Extending Historical Friedkin Model to Iran

We investigate the broader strategic context surrounding a nuclear capable Iran given current hostilities and alliances in the region. We emphasize that these results are part of a broader ongoing discussion regarding proliferation in the region. We first discuss some background on nuclear proliferation, describe the model and data being used, list the hypotheses used, review results, and then conclude.

Background

Countries develop nuclear weapons for a variety of reasons, from concerns arising from security deficits to a commitment to norms and prestige surrounding nuclear weapons. This work broadly comes out of two literatures: nuclear deterrence and nuclear proliferation. Nuclear deterrence is traditionally contrasted with compellence – threats, as opposed to actions – and arguments for deterrence have focused on actions between two nuclear states as opposed to actions between non-nuclear and nuclear states. Nuclear proliferation is concerned with the spread of nuclear material and ultimately, nuclear weapons, outside of the existing international regime outlined by the non-proliferation treaty (NPT). Unlike deterrence, the proliferation literature examines both internal domestic motivations for developing weapons in addition to motivations driven by external actors.

The literature on nuclear deterrence remains broadly based in Cold War thinking as doctrine and policy developed in response to a bipolar nuclear world. Deterrence is commonly accepted to have evolved over at least three "waves" [1]: an initial wave, which explored the impact of nuclear weapons on world politics, to a second wave, which combined policy and theory, to a third wave, which highlighted empirical work. The second wave, which incorporated game theory models, such as the 'Chicken Game', led to important insights about the nature of international relations, but did not contribute to direct policy implications: while it explained superpower relations, and framed broad strategic issues, it did not significantly contribute to smaller diplomatic and military efforts [2], [3] .The lack of empirical evidence made it difficult to evaluate claims made in deterrence literature[4],which helped lead to the third wave's emphasis of empirical work on risk taking, rewards, misperceptions, and bureaucratic politics[1], [5].

Traditionally, deterrence requires actors who are rational, resolute, and credible – all traits that rogue actors, such as North Korea, may not consistently demonstrate [6]. An alternate angle, however, is that the presence and threat of any type of weapon of mass destruction make deterrence easier [7] – threatening an actor, combined with the crystallization of the risk posed by a WMD to other actors, can make it easier to respond to threats [8], [9]. Others have argued that the WMD threat makes it easier for rogue states to deter other actors, such as the United States, from involvement [10], [11].

These external motivations for developing nuclear weapons are most commonly associated with a realist, or security based motivation approach to developing nuclear weapons – one that focuses on nations as actors in a state of anarchy [12]. There are two other major schools, which include domestic politics and constructivism [13]. The domestic politics school, which focuses on the role of different domestic actors, argues that the nuclear capability of a country can emerge from disparate actor politics, including responding to international institutions, political economic ambitions, and nuclear ambivalence [14], [15] [16]. The constructivist school, which focuses on norms, argues that national leaders and identity play a major role in country motivations [17]-[19].

Past work on quantitative models and nuclear proliferation have focused on generalized dynamic models incorporating the entirety of the historical data available, finding and identifying coefficients of global variables utilizing hazard models and generalized logistic regression models

[20]-[22]. As good models, these approaches incorporate different variables reflecting these theories, incorporating economic, institutional, and prestige indicator variables.in However, recent work has found that these model coefficients relating to security concerns are not robust due to small changes in data, from mislabeled dyadic data to missing conflicts [23], [24]. Specifically, this indicates that we need a deeper understanding of the dynamics of security-driven motivations to pursue nuclear weapons.

The goal of this paper is to connect the ongoing body of policy-oriented political science [45] with policy-focused explorations of strategic issues [47], [48] [49]. The field is understandably- and correctly- worried about the impact of statistical "small-n" issues on interpretations of models, which make uncertainty analysis difficult. We propose a new approach that aims to follow overall changes in behavior – as opposed to model a static system – to capture changes in international system norms and behaviors over time. This paper is an application and extension of this work.

We will focus on the mechanics of the security model in developing our model for nuclear weapons, which has broad support [12]. In the security model, a country that has a nuclear enemy perceives a security deficit, and is thus motivated to acquire nuclear weapons[26]. The country may also seek an alliance with a nuclear power that promises retaliation in case the country is attacked [27]-[29]. Such alliances provides reassurance for the country and reduces its need for developing indigenous nuclear weapons. With this approach, we find evidence of emergent international institutional behavior captured by this system coincident with the introduction of the Non-Proliferation Treaty.

In this paper, we specifically explore the impact of existing networks of hostilities and alliances on actors in the middle east, utilizing a modified version of Friedkin's opinion dynamics model to simulate the effects of these networks on individual actors. While there has been a concerted effort to remove other kinds of WMD from the region (i.e. removing chemical weapons from Syria), one of the core policy goals of the Iranian negotiations was to remove Iranian nuclear capability. We explore this impact, as well as the sensitivity of the model to changes in network structure, and implications of the model.

Model structure

Traditionally, Friedkin has been used to show the opinion dynamics of a small group [51]. This approach has been used to model opinion changes among political stakeholders in major processes such as voting in the EU and extremist behavior [52] .

Unlike past applications of the Friedkin model where group consensus is reached, we modify it to reflect the different levels of motivation enforced by different networks. We are influenced not only by individuals in our immediate social friend network, but also by the actions of those we consider to be competitors or those hostile to us. In an international relations setting, we consider the networks of countries in alliances and networks of countries with hostilities. We first introduce the formal model and then demonstrate how the model is modified.

At a high level, the Friedkin equation (equation 1) consists of 3 parts. Wy_{t-1} represents actors' extrinsic attitudes resulting from external influence, y_1 represents intrinsic attitudes that reflect actors' own characteristics and constraints, and A represents the relative weight that actors place on extrinsic and intrinsic attitudes. The Friedkin model can apply to any attitude. In this work, we consider the attitude to be the motivation to develop nuclear weapons.

Equation 6. Friedkin model equation

$$y_t = AWy_{t-1} + (I - A)y_1$$

In the equation, y_t is an $N \times 1$ vector that represents country attitudes at time t. The attitude of each country follows scaling given in Figure 1, where 0.5 represents an indifferent attitude, larger values represent a positive attitude and smaller values represent a negative attitude. $A = diag(a_{11}, \dots, a_{ii}, \dots, a_{NN}), 0 \leq a_{ii} \leq 1$ is a $N \times N$ diagonal matrix with diagonal weights indicating the level of influence that each actor puts on outside actors. $W = [w_{ij}], (0 \leq w_{ij} \leq 1, \sum_j^N w_{ij} = 1)$ is an $N \times N$ matrix that represent inter-actor influence. More specifically, w_{ij} represents the extent to which actor j has on actor i. W is computed using the formula $W = AC + I - A$ where $C = [c_{ij}]$ is a $N \times N$ matrix of relative interpersonal influence such that $(c_{ii} = 0, 0 \leq c_{ij} \leq 1, \sum_{j=1}^N c_{ij} = 1)$. Finally, y_1 is a $N \times 1$ vector representing actors' initial attitudes.

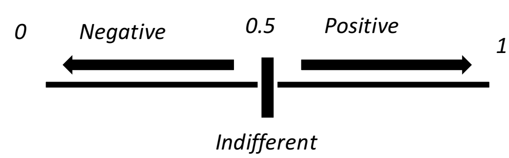


Figure 9. Scaling of attitude values

The extrinsic attitude term Wy_{t-1} can only capture influence from one class of tie between two actors. We modify this to represent two kinds of extrinsic influence for two kinds of networks: alliances and hostilities.

We find a new term that captures nuclear deterrence, nuclear reassurance and the interaction between the deterrence and reassurance. We then substitute Wy_{t-1} in Equation 1 by the new term in order to find the modified equation model. In order to simplify the discussion, we initially consider a single country that has a single enemy and a single ally, and derive a new term for the extrinsic motivation. Subsequently, we modify that expression into a vectorial expression that captures the extrinsic motivation of all countries. Finally, we include that vectorial expression into Equation 1, obtaining the modified equation model.

Based on past work examining historical trends in nuclear proliferation, we find coefficients α_1, α_2 and α_3 in Equation 2 that reflect current behavior of countries interested in developing nuclear weapons. At the country level, we describe motivation as the linear sum of the effect of hostilities, alliances, and the interaction of alliances and hostilities.

Equation 7. Extrinsic motivation coefficient equation

$$m_t = \alpha_0 + \alpha_1 e_{t-1} + \alpha_2 f_{t-1} + \alpha_3 e_{t-1} f_{t-1}$$

This motivation applies to only one country. We are interested in obtaining a vectorial expression that simultaneously captures the motivation of all states. We now assume that M_t, E_{t-1} and F_{t-1} are Nx1 vectors. To preserve these vectors, we now place values of F along a diagonal matrix.

Equation 8. Vectorial form of extrinsic motivation

$$M_t = \alpha_1 E_{t-1} + \alpha_2 F_{t-1} + \alpha_3 (E_{t-1})\, diag(F_{t-1})$$

Given that E_{t-1} captures whether countries' enemies have nuclear weapons at time t-1, and F_{t-1} captures whether allies have nuclear weapons, we can write $E_{t-1} = W_H y_{(t-1)}$ and $F_{t-1} = W_F y_{(t-1)}$, where W_H captures international hostilities and W_F captures international alliances. Finally, we obtain

Equation 9. Adapted Friedkin Model Equation

$$y_t = A\,[\,\alpha_1 W_H y_{t-1} + \alpha_2 W_F y_{t-1} + \alpha_3 W_H\, y_{t-1}\, diag(W_F y_{t-1})] + (I - A) y_1$$

We discuss the different coefficients used in this model in the following section.

Data

There are four primary sources of data utilized for the adapted Friedkin model: world GDP data, for matrix A, the political alliance and hostility networks, and the initial set of countries with nuclear capabilities. These are summarized in the table below.

Table 13. Summary table of data utilized for Friedkin model

Model Variable	Description	Modification	Source
A	Diagonal matrix indicates how much country i weighs opinions of others in its network	1-log(GDP_i)/log(GDP_{max})	National GDP in 2010 from World Bank Data
W_H	Network of hostilities	Normalized by column sum	International Crisis Behavior, Uppsala Conflict Data Program
W_F	Network of alliances		Correlates of War
y_1	Country initial motivation to develop nuclear weapons	Set to 0.5 if country is a non-nuclear weapons state; 1 if country is a nuclear weapons state	Recognized nuclear weapons states: United States, Russia, China, United Kingdom, France, Israel, Pakistan, India, North Korea

To model the weight given to other individuals in the network, A, we utilize a normalized version of the country's GDP. We utilize the ratio of the log of country GDPs; we anticipate that countries with smaller GDPs are more sensitive to pressures from alliances and hostilities than countries with larger GDPs. A country with a larger GDP anticipating economic retaliation for changing its position on weapons of mass destruction will be more insulated from an anticipated marginal shock than a country with a smaller GDP, where the equivalent nominal economic shock will have a much higher marginal value. In our simulations, we utilize 2010 GDP data from the World Bank[53].

The political alliance network is based on the formal alliance network from the Correlates of War (COW) project, using the interstate alliance data set v4.1[54]. The data is first filtered to reflect dyadic alliances in force from 1995-2012. The data set distinguishes between four kinds of treaties: a defense pact, a neutrality pact, a non-aggression pact, and ententes. For our analysis, we consider all forms of alliance to indicate an alliance. We utilize alliances formed from 1995-2005 for the Alliance network in the 2005 simulation, and alliances formed from 2005-2012 for the Alliance network in the 2015 simulation. In the 2015 simulation, we explicitly add the alliance between Russia and Syria.

For the hostilities network, we use two sources: the International Crisis Behavior (ICB) project at the University of Maryland and the Uppsala Conflict Data Program (UCDP) at the University of Uppsala, Sweden[55], [56]. The ICB project data covers violent and non-violent conflicts during the period 1918-2013. The UCDP data covers violent conflicts that caused at least 25 deaths in a calendar year during the period 1993-2014.

The last used networks to fit the historical model are from 2000 – however, there have been several major international events, such as the attacks on September 11th, 2001, as well as the invasion of Afghanistan and Iraq, since then. We contrast the hostilities and alliance networks from 1995- 2005 with the hostilities and alliance networks from 2005-2015. The most recent dyadic version of these conflicts, ICB 11, was released earlier in February 2016 and only goes up through 2013 hostilities and events. These updated hostility networks include NATO involvement in the Libyan civil war, as well as Russian aggression in Georgia and Ukraine.

In the Friedkin model, we modify both the Alliance and Hostility networks so that the column sums of the matrix are equal to 1. This way, an element in row *i*, column *k* reflects normalized weight that country *k* has on country *i*. For example, in the network picture below of hostilities in 2005 and 2015, Syria has hostilities with Israel, the United States, and Turkey in 2015, while the United States has conflicts with Syria, Iran, Libya, and North Korea. The value of

$Hostile_{Syria,\ US}$ is 0.33 while $Hostile_{US,\ Syria}$ *is* 0.25, reflecting the different normalizations based on the number of hostilities that the country has.

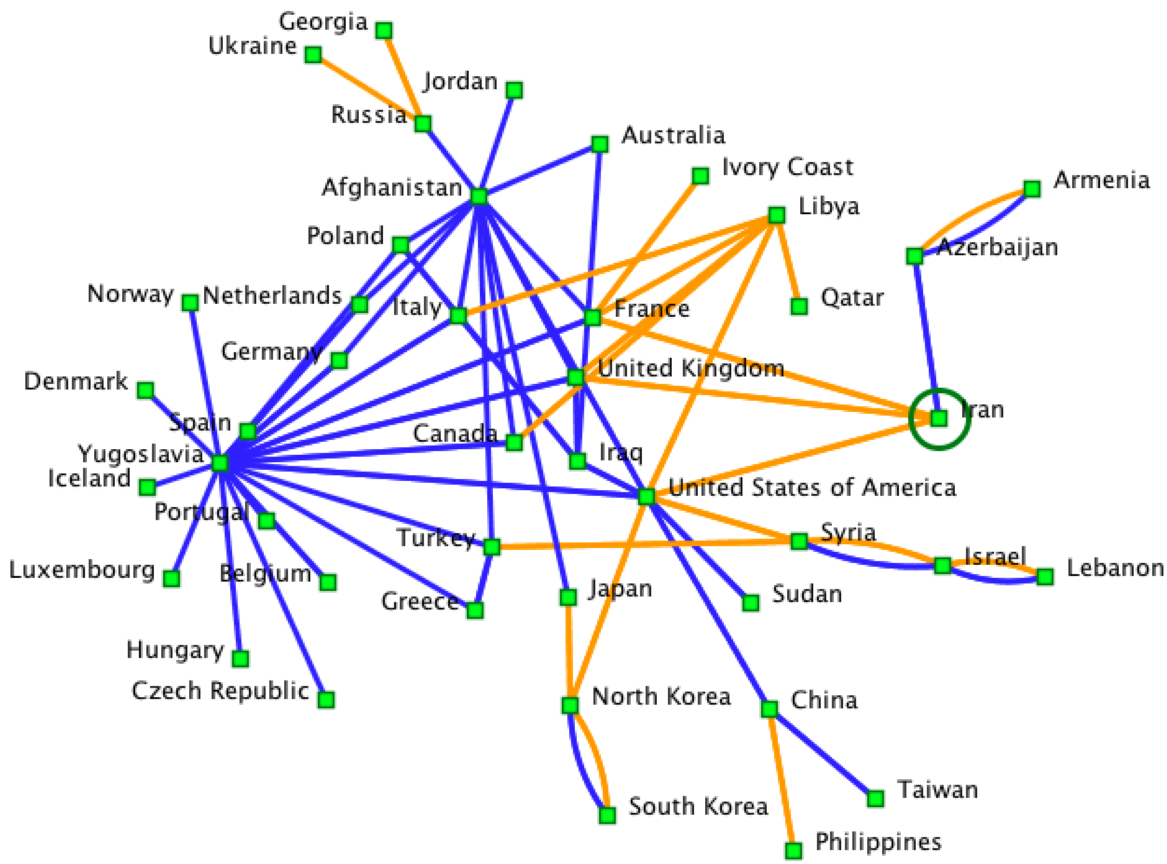


Figure 10. Sphere of influence of Iran, two degrees out. Hostilities 1995-2005 in blue; Hostilities 2005-2015 in orange.

For this analysis we consider the following countries to have nuclear weapons: the United States, Russia, China, the United Kingdom, France, Israel, Pakistan, India, and North Korea. Their initial values are set to 1; all other countries are set to have an initial motivation level of 0.5.

Historical model and hypotheses

The historical model considers nuclear powers in the given year to have a high motivation to develop nuclear weapons, and utilizes the adapted Friedkin equation to explore a gridded coefficient space to identify the countries over the next five years that will also want to develop nuclear weapons. It considers motivations over 0.5 as motivated to develop nuclear weapons. We find that the Non-Proliferation Treaty does seem to affect behavior – after ratification, it appears that the coefficients that most accurately model the spread of nuclear weapons indicate a role for the alliance network. The summary table below highlights the coefficients that contribute to versions of the model with the maximum F1 score for the predicted five years.

The first thing to note about the coefficients that we use as a baseline for the Iranian scenario is that unlike the previous four time periods, which have small numbers of cases matching, there are a wide set of potential coefficients that explain the current situation. Additionally, unlike the previous four time periods, where hostilities and security seem to be significant drivers of motivation, as evidenced by the fact that the alliance coefficient is smaller than the hostility coefficient, hostilities seem to drive down motivation in the median case for 2000.

This allows us to consider three hypotheses about the drivers of proliferation in a world after 9/11:

- Countries continue to be significantly motivated by security concerns to develop nuclear weapons (large hostility coefficient)
- Countries distrust extended security guarantees (large alliance coefficient)
- Alliances discourage countries from developing nuclear weapons (negative alliance coefficient)

Table 14. Summary table of historical model[1].

Year	Max F1	Median Hostility Coefficient	Median Alliance Coefficient	Median Interaction Coefficient	No. Matching Cases	% of grid
1960	0.612	-0.2	0	0.1	3893	80%
1965	0.625	-0.1	-0.1	0	4380	90%
1970	0.57	1	-0.3	-0.5	3	0%
1975	0.70	-0.1	-0.1	0	4233	87%
1980	0.64	0.4	0.8	-0.4	55	1%
1985	0.64	1	1	0.2	3	0%
1990	0.81	0.7	0.7	-0.3	197	4%
1995	0.81	1	1	-0.3	5	0%
2000	0.76	-0.4	0	0.1	2797	57%

We then use these hypotheses to explore the implications for the security context in the middle east. A table of the coefficients used in the two hypotheses is shown below. All three sets of coefficients have the same F1 score of 0.76. For the security hypothesis, the smallest interaction coefficient that was also part of the set had a value of -0.5. Just as in the historical model, the Friedkin models are run for 30 iterations; as we are not randomizing effects, we do not utilize Monte Carlo or average out results.

[1] The historical model is explored using the historical proliferation data from Singh & Way as a baseline.

We also run a parallel set of simulations exploring the impact of ISIS in the region. For the purposes of the simulation, we assume ISIS is influenced by other nations in its network at the same rate as Syria; we do not give it any neighbors in the Alliance network, but highlight its hostilities with the following countries, either through direct military operations, such as France, Iran, Iraq, Syria, Russia, Turkey, and United States, or through countries that have official ISIS affiliates, including Afghanistan, Algeria, Egypt, Nigeria, Pakistan, Saudi Arabia, and Yemen [61]-[63].

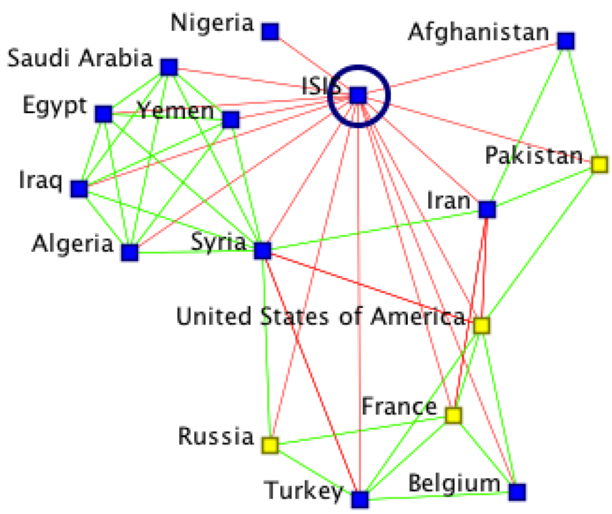


Figure 11. ISIS Egonet. Red ties indicate hostilities; green indicate alliances. Yellow nodes indicate nuclear weapons state.

Table 15. Table of coefficients used in hypotheses

Hypotheses	Hypothesis	Hostility Coefficient	Alliance Coefficient	Interaction Coefficient
Baseline		-0.4	0	-0.1
Security	A	0.6	0	-0.5
Alliance distrust	B	-0.4	0.6	-0.1
Alliance discourage	C	-0.4	-0.6	-0.1

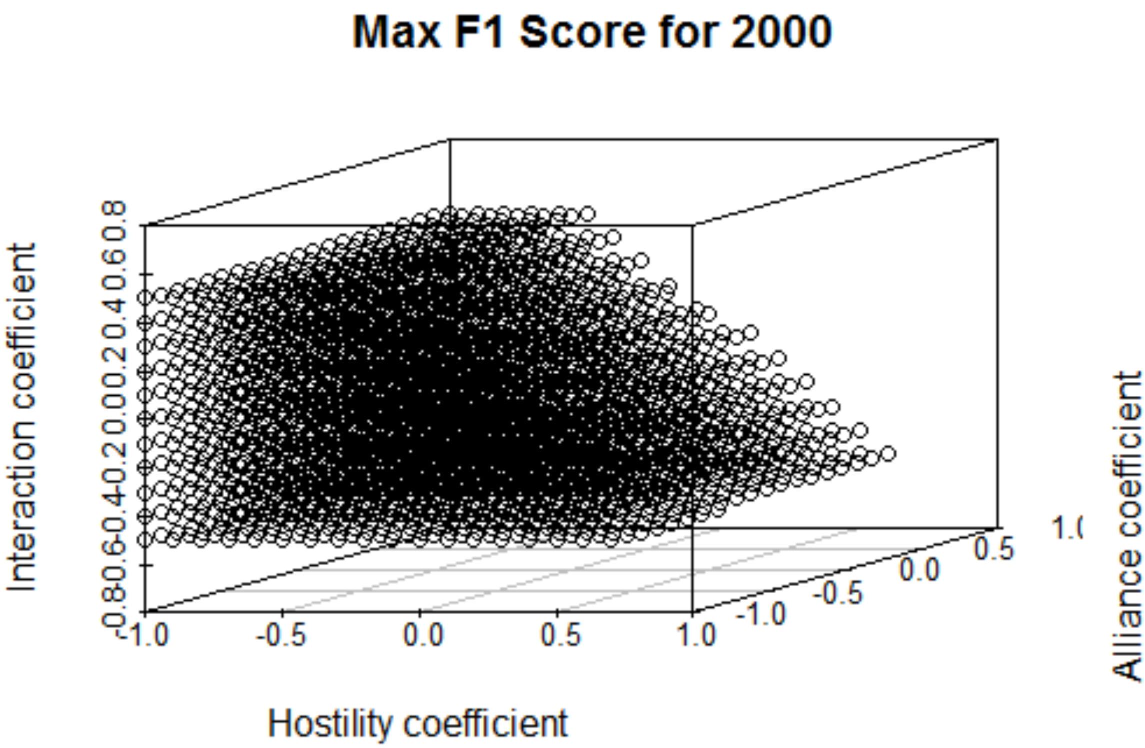


Figure 12. Plot of coefficients with identical F1 scores at F1 = 0.76.

Results

The first thing to note about the model – as well as the different hypotheses used – is that while we focus on Iran, Syria emerges in all time periods as being a country that is significantly motivated to develop a nuclear capability. This suggests that even as the country continues to face civil war, it is a country at a unique crossroads in terms of alliances and hostilities. We now examine the three hypotheses, focusing first on results without considering the role of ISIS, and conclude with an examination of how ISIS has impacted other country motivations in this model.

Under the security hypothesis, that hostilities increase motivations to develop nuclear weapons, we find that Iran would find today's current security context to be more of a hostile environment – and should increase its motivation to develop nuclear weapons. As a landmark security event in 2015 has been Iran's negotiating away its major nuclear capabilities for the immediate future, this suggests that Iran views investments in conventional weapons and warfare sufficient for its current security needs.

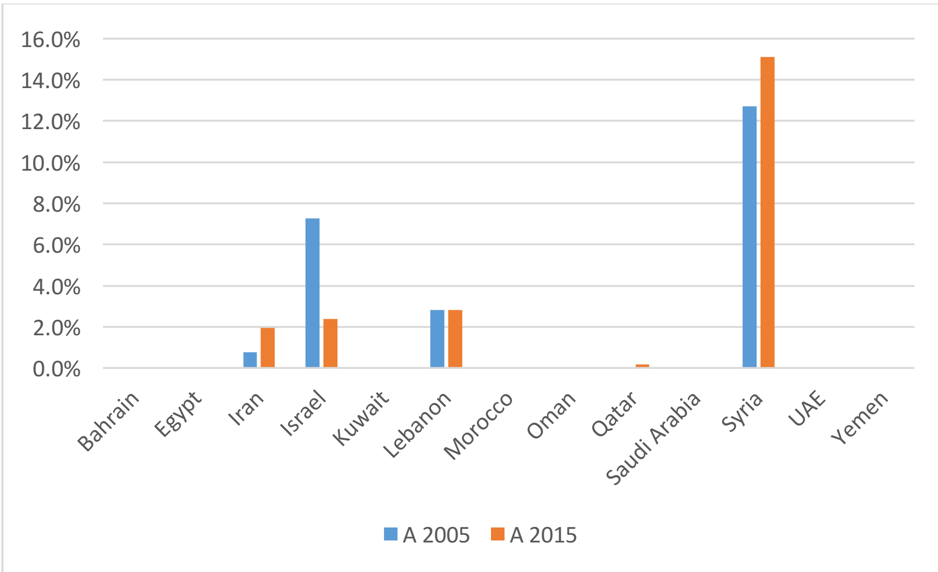


Figure 13. % Change from baselines under Hypothesis A in 2005 (blue) and 2015 (orange)

In hypothesis B, we find little evidence of a nuclear distrust of alliances – there would have been a greater push for pursuit of nuclear weapons back in 2005. The network has changed and distrust would only marginally increase motivation to develop a capability. It should be noted here that in the alliance network, many of the majority Sunni states are structurally equivalent; they embed themselves in the Arab League and as a result of the weighting normalization, exert relatively little influence on each other due to the sheer number of member states.

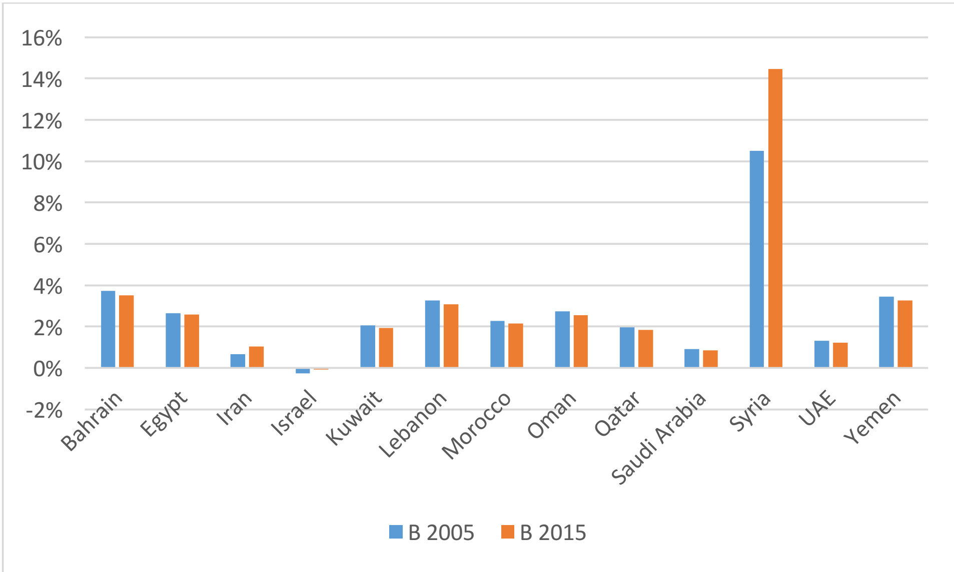


Figure 14. % Change in baseline under hypothesis B, 2005 (blue) and 2015 (orange)

In examining the alliances hypothesis we find that Israel, as a relatively isolated country in the region, slightly increases its motivation to hold on to its nuclear capability. Many of the other countries face a mirror image of what we observe in Hypothesis B, suggesting existing alliances are providing sufficient support for the region at the national level.

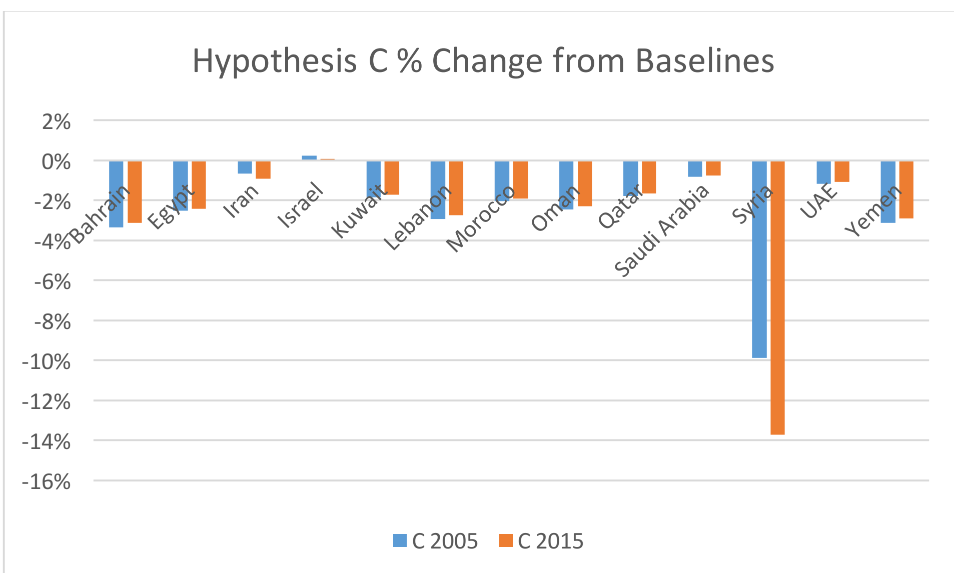


Figure 15. % Change in baseline under Hypothesis C, 2005 (blue) and 2015 (orange)

Comparing the three hypotheses in 2005 and 2015, we see that despite major structural changes in the hostilities network from 2005-2015, we see relatively few major swings in overall country-level motivation to develop nuclear weapons. Both increasing tensions and distrusting alliances only moderately increase motivations to develop an indigenous nuclear capability. Due to the relatively few direct interactions that Iran has with many Arab states in this model, the model also finds that directly simulating an Iranian nuclear capability exerts a relatively small change in the overall behavior of other actors in the region.

Unlike the previous three presented hypotheses, when incorporating ISIS, we only compare hypotheses to the 2015 baseline, re-calculated with ISIS as an agent with hostilities listed in the previous section. We then re-run the coefficients listed to test hypotheses A, B, and C.

Table 16. Summary table of % change results incorporating ISIS, which has a negative motivation value in the baseline calculation. It was compared against a value of 0.5 for the 2015 baseline. It has a positive actual value under Hypothesis A and negative actual value under Hypotheses B and C.*

	% Change from 2015 Baseline (without ISIS)	**% Change Hypothesis A**	**% Change Hypothesis B**	**% Change Hypothesis C**
Bahrain	0%	0%	3%	-3%
Egypt	0%	0%	3%	-2%
Iran	0%	2%	1%	-1%
ISIS*	-232%	-335%	18%	-17%
Israel	0%	1%	0%	0%
Kuwait	0%	0%	2%	-2%
Lebanon	-1%	4%	3%	-3%
Morocco	0%	0%	2%	-2%
Oman	0%	0%	2%	-2%
Qatar	0%	0%	2%	-2%
Saudi Arabia	0%	0%	1%	-1%
Syria	26%	-2%	16%	-15%
UAE	0%	0%	1%	-1%
Yemen	0%	0%	3%	-3%

In modeling the motivations of ISIS, we find that strictly incorporating its hostilities lead to an overall decrease in motivation for its utilizing a nuclear weapon. The only hypothesis where it shows an increased motivation is in the security hypothesis.

Discussion

This modeling approach has some limitations, which are related to broader trends and the reality of conflict in this region. Despite these limitations, this approach and framework are useful first steps towards identifying and quantifying non-proliferation policy goals.

The first major limitation of this approach is that it only takes into account state actors. While it can be easily expanded to account for non-state actors, as we have done by incorporating hostilities such as ISIS, in doing so it can provide a false sense of security by showing that relations are relatively stable in the region. Utilizing the adapted Freidkin model to model state actors was a deliberate choice: developing a fully-fledged indigenous nuclear capability historically has only been successfully achieved by nation-states. Many non-state actors have attempted – and ultimately, failed – to pursue a fully fledged program. While our historical model finds evidence of emergent institutional behavior thanks to the introduction of the Non-Proliferation Treaty, we do not directly model institutional or economic effects related to the pursuit of nuclear weapons, such as the mitigating effect of an institution such as the International Atomic Energy Agency or the economic possibility of purchasing nuclear material and materiel through covert networks.

The reality of conflict in this region is dominated by non-state actors playing proxy roles for major regional states. We do not directly observe the increasingly important role played by non-state actors in the region that contribute to the regional conflict between Shiite and Sunni factions. In particular, two conflicts have escalated and have made modeling more difficult as they involve significant non-state actors: the civil wars in Yemen and Syria. In Yemen, Shiite Houthi rebels, thought to be aided by state support from Iran, are fighting against Yemeni government forces, which are aided by a Saudi-led coalition of predominantly Sunni forces[64]. In Syria, the government has recently been aided by direct Russian military support nominally focusing on attacking ISIS but in reality providing significant cover for the existing regime [65]. In both cases, there are major non-state actors involved – Houthi rebels and ISIS.

It is appropriate to use ISIS in the Friedkin model as it has shown itself to be a substantial and persistent actor that has presented a security threat with a wide range of state actors. It has also demonstrated its willingness and ability to use other types of WMDs, giving it additional credibility. It presents a different type of security risk than, for example, the Yemeni Houthi rebels, who primarily present a threat to Yemen.

Despite these limitations, Friedkin remains a useful tool to measure and simulate network effects. It is important to remember that it only considers strategic structural motivations related to nuclear weapons. It does not, for example, directly relate to the extremist narratives espoused by some of these sub-state actors. Day-to-day policy needs to not only rely on these ‘demand’ side modeling approaches but also keep in check ‘supply’ side considerations for all kinds of WMD[48]. These supply side considerations have led to cooperation, such as US-Syrian-Russian cooperation in removing chemical weapons from Syria [66], as well as broader concern over ISIS access to enriched uranium material sufficient for a ‘dirty bomb’ [67].

Consider the following ego nets of four key players in the Iran case: Iran, Israel, Saudi Arabia, and Syria. In Iran’s case we see that Iran is lightly embedded in an alliance with Azerbaijan, Afghanistan, Tajikistan, Turkmenistan, Uzbekistan, Syria, China, and Pakistan, with security pressures from the US, UK, France, and ISIS. It appears to be relatively balanced in terms of having nuclear powers both facing it down in hostility as well as nuclear powers that support it militarily. In the COW data, Russia is not listed as providing a military alliance to Iran.

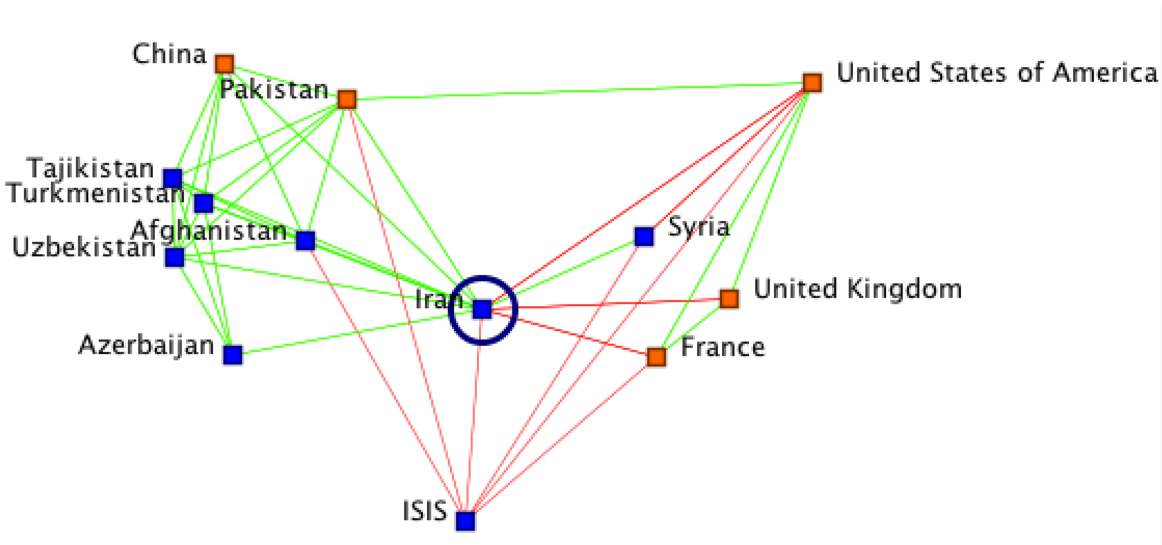


Figure 16. Iran Egonet in 2015. Orange node indicates nuclear weapons state; red ties indicate hostility, green ties indicate alliance.

In Israel's case, we see that it has direct hostilities with Lebanon and Syria, and alliances with Egypt and Jordan. The network data does not capture ongoing hostilities that Israel has with the rest of the region, and as can be seen in the appendix, Egypt and Jordan take on major brokerage roles in connecting Israel to the rest of the middle east.

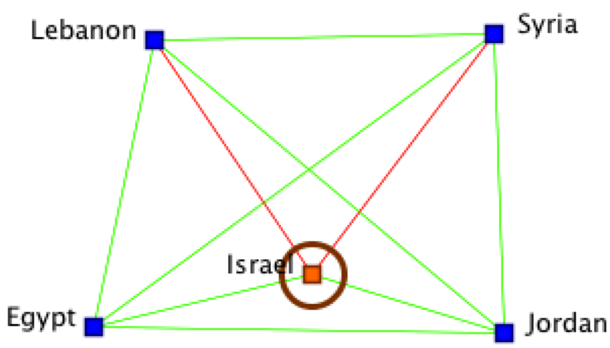


Figure 17. Israel egonet in 2015. Orange node indicates nuclear weapons state; red ties indicate hostility, green ties indicate alliance.

We see that Saudi Arabia is tightly embedded in the Arab League states; its only listed hostility is with ISIS, which we have added. Saudi Arabia has provided extensive military assistance to several Arab League members, including Yemen and Bahrain, to stop their own domestic disturbances.

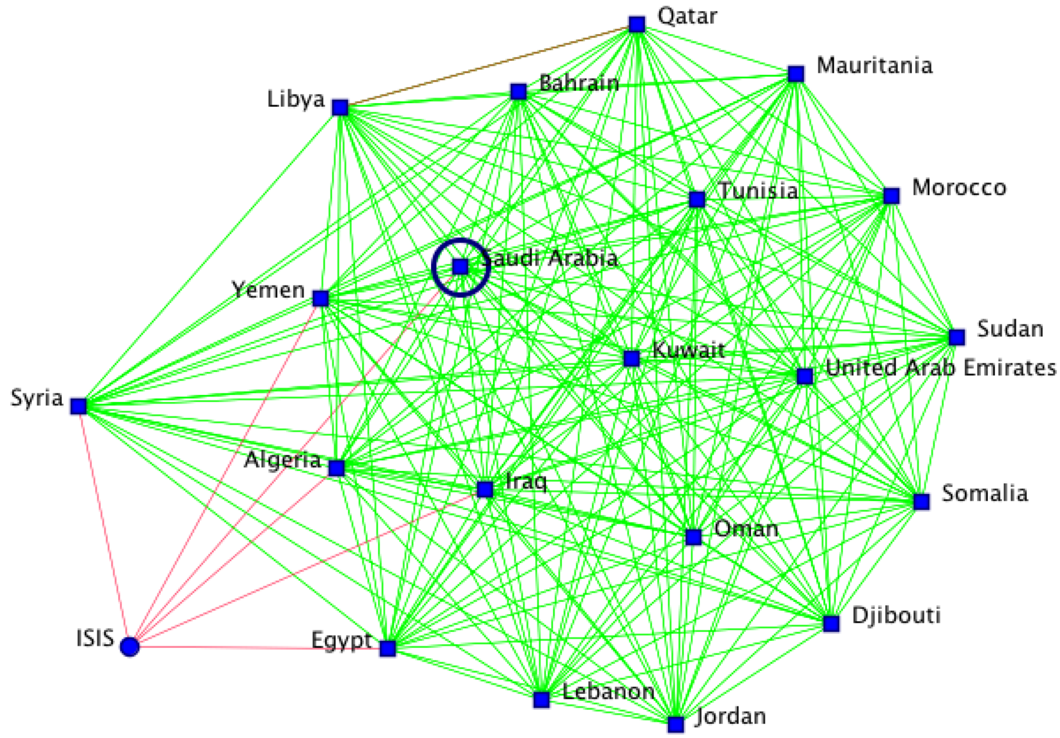


Figure 18. Saudi Arabia egonet in 2015. Orange node indicates nuclear weapons state; red ties indicate hostility, green ties indicate alliance.

Syria is an outlier state – we see that it is embedded in the Arab League, but that it also has positive ties with Iran and hostilities with the US, Turkey, ISIS, and Israel.

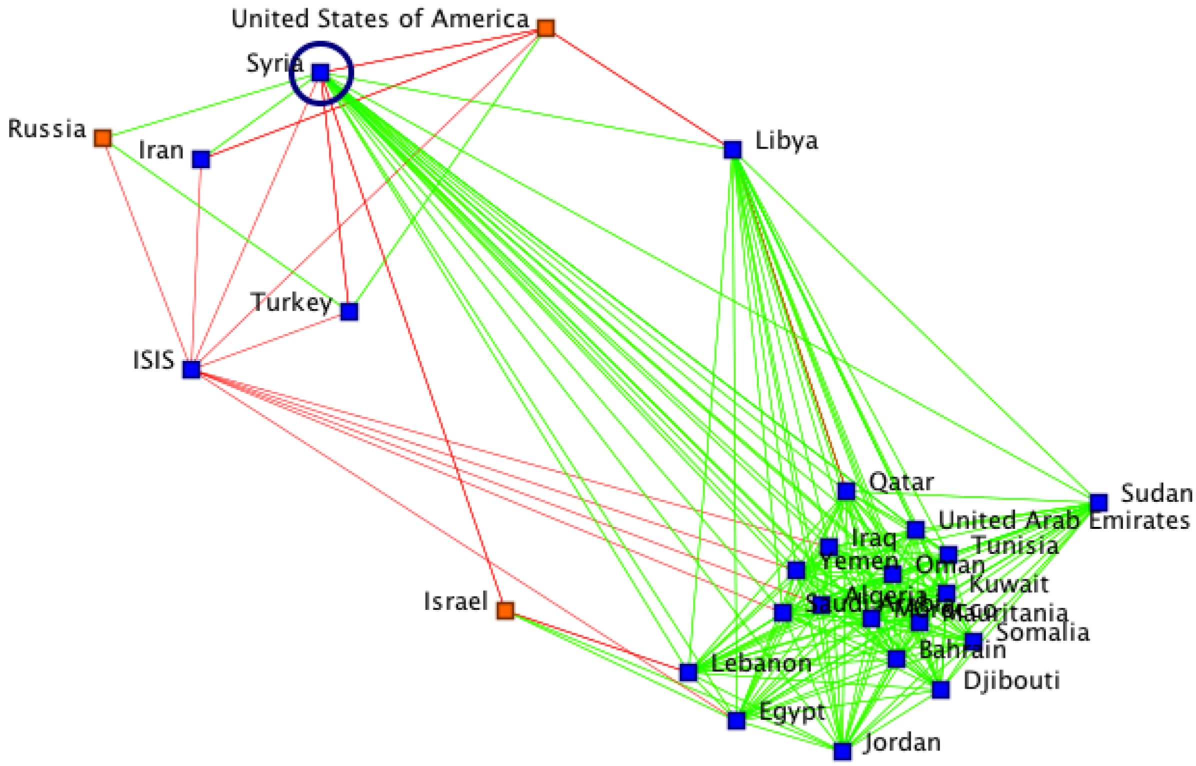


Figure 19. Syria egonet in 2015. Orange node indicates nuclear weapons state; red ties indicate hostility, green ties indicate alliance.

In assessing the different results from this what-if scenario, one of the most interesting things that comes out is Syria's emergence as being a critical country from a nuclear proliferation perspective over Iran's much more prominent nuclear program. Like Saudi Arabia, Syria is embedded in the Arab League, which appears to have a moderating effect on most of its members. Unlike Saudi Arabia, however, Syria has explicit military alliances with a nuclear power (Russia), as well as hostilities with two nuclear powers (the US and Israel). The addition of these two nuclear threats, combined with a nuclear power ally, puts Syria over the top in its motivation to develop a nuclear capability.

Iran, by contrast, while maintaining alliances with two nuclear power states (China and Pakistan) and hostilities with three (the US, France, and the UK), appears to balance out its overall motivation to develop a nuclear capability. Iran is embedded in a far less dense alliance among the Central Asian states – its only alliance outside of that nexus is to Syria, a non-nuclear state. Despite a worsening of its security context – as measured here, through hostility networks – Iran appears to be less motivated to develop a nuclear capability when both its alliance and hostility networks are considered in the Friedkin model.

The last two hypotheses in this chapter's experiments explored the impact of alliances on the pursuit of nuclear weapons at a regional level. In finding that many of the Arab League states are not motivated to develop a nuclear capability, combined with Iran's willingness to give up its most direct path to a nuclear weapon, we find that the region is unlikely to pursue a nuclear arms race. Saudi Arabia's embeddedness in a political network that overwhelmingly pursues conventional weapons, as well as its lack of a direct hostility with nuclear weapons states, indicates that it is on a track to continue pursuing conventional weapons to satisfy its security requirements – even with Israel's nuclear capability in the region. As Israel is relatively isolated in the political networks, if it continues to stay isolated its effect will be relatively small.

Conclusion

We have demonstrated an application of computational modeling to explore different political science hypotheses in a highly contested region. We have demonstrated it is relatively easy to extend this computational model to take into account non-state actors, but caution that it can be difficult to see emergent regional trends without regional expertise to validate certain linkages in the network data. By combining this computational model with supply-side approaches, we will be better able to highlight risks and impending changes to existing networks of alliances and hostilities.

Future work in this area should address other stakeholders in the nuclear proliferation process, from sub-state actors (including domestic political actors) to direct measurement of public stakeholders. Some stakeholders can be assessed through developing richer network data to more effectively model sub-state actor alliances and hostilities. This can be done through more extensive text mining of extremist narratives, as well as funding network information, which can be obtained at a regional level. It should also marry motivation and demand-side simulation models with supply-side data, such as the presence of CBRN materiel and delivery mechanisms, to develop a more accurate profile of the immediate risks in the region.

Another approach to identifying stakeholders is through monitoring social media data. For example, in monitoring social media tweets related to nuclear proliferation, we find that while Iran was not a major source of tweets in the lead-up to the Iran Deal being announced, when the joint framework was announced, the majority of nuclear-related tweets coming out of the country were highly positive. Twitter in particular is a particularly rich vein for monitoring and understanding the PMOI Iranian government in exile, which came out as being both strongly negative on the Iran deal as well as strongly negative on nuclear topics. Further exploring this area will prove to be a rich data source for future researchers.

Data and Figure Appendices

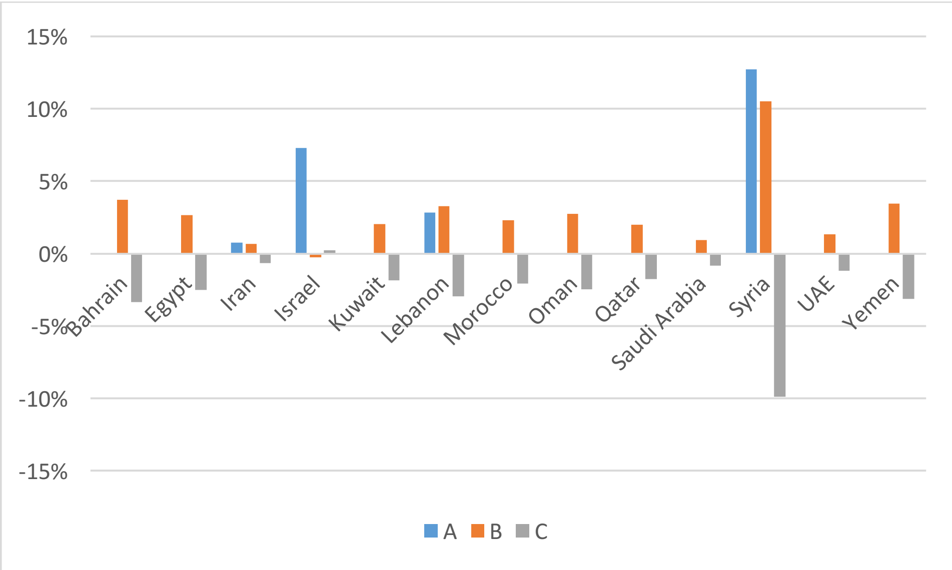


Figure 20. % compared to 2005 baseline; Hypothesis A (blue), Hypothesis B (Orange), Hypothesis C (Grey)

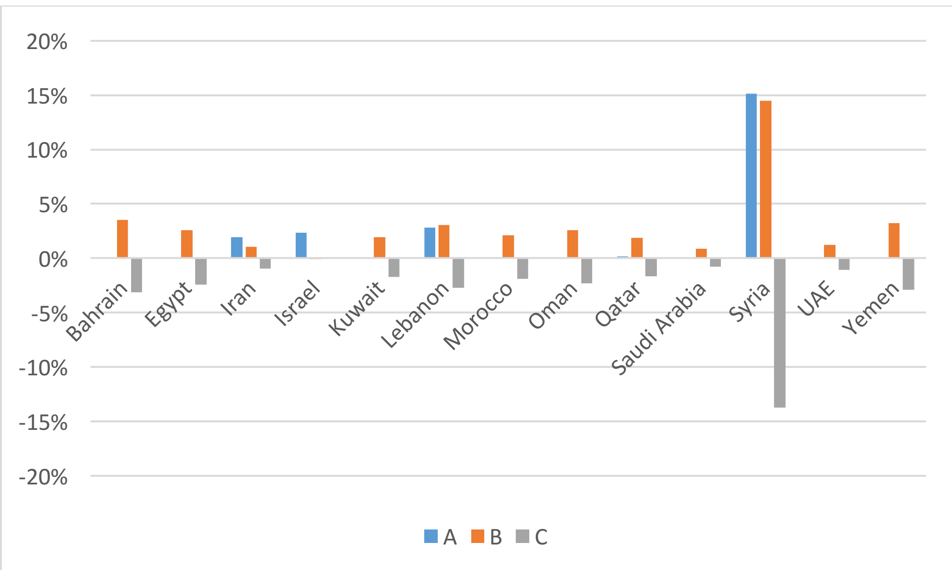


Figure 21. % compared to 2015 baseline; Hypothesis A (blue), Hypothesis B (Orange), Hypothesis C (Grey)

Table 17. Normalized Total Degree Centrality of Binarized Alliance and Hostility Networks

	Bahrain	Egypt	Iran	Israel	Kuwait	Lebanon	Morocco
Alliances 2005	0.10	0.11	0.04	0.02	0.10	0.10	0.10
Alliances 2015	0.10	0.10	0.04	0.01	0.10	0.10	0.10
Hostilities 2005	0.00	0.00	0.01	0.01	0.00	0.01	0.00
Hostilities 2015	0.00	0.00	0.02	0.01	0.00	0.01	0.00

	Oman	Qatar	Saudi Arabia	Syria	UAE	Yemen
Alliances 2005	0.10	0.10	0.10	0.10	0.10	0.10
Alliances 2015	0.10	0.10	0.10	0.11	0.10	0.10
Hostilities 2005	0.00	0.00	0.00	0.01	0.00	0.00
Hostilities 2015	0.00	0.01	0.00	0.02	0.00	0.00

Chapter 4: Impact of Context on Social Media Posts

Traditional approaches to sentiment analysis have three problems: the approaches were originally developed to analyze larger bodies of text, they ignore the social context of social media, and they are primarily focused on only one dimension of sentiment. As social media text can be extremely short, and due to the expense associated with obtaining labeled data necessary to train machine learning algorithms, most approaches to sentiment analysis today rely on extensive lexicons with the goal of having some text match words that we know map to generally positive or negative sentiment [68]-[70].

Most approaches to sentiment analysis in social media focus exclusively on the content of the message, ignoring the metadata and subsequent social context that the message comes out of [71]-[74]. For example, a user posting she is ill will receive positive, supportive posts on social media. Analyzing the social network associated with the flow of those messages would result in an incorrectly classified positive association with that sickness. While some analyses of social network sentiment incorporate analysis of a user's social media ties, these studies rely on aggregated posts and do not consider individual responses to news, topics, or events[75] [76]. Finally, sentiment is typically analyzed along a single dimension: positive and negative, with a minority of research considering objectivity [71] [77]. However, there are other dimensions to emotions, informed by cultures, which affect how individuals respond to events. Affect control theory (ACT) formalizes the way that individuals respond to events by classifying evaluation, potency, and action, allowing for cross-cultural comparisons of events [78], [79] [80]. Evaluation is the most similar dimension to most sentiment tools today: it is a spectrum from unpleasant and negative to pleasant and positive. Power reflects the social and external relations individuals have, going from weak and powerless to strong and powerful. Activity, in contrast to power, reflects internal relationships to emotion, going from unexciting and inactive to exciting and active. In this study, we utilize a recent dictionary consisting of over 2,000 terms to populate lexicons to identify messages along potency and activity[81].

The paper seeks to explore three key areas: how affect control theory can inform sentiment analysis, how individuals perceive messages seen without context differently from messages with context, and finally, the implications of context for existing tools. We examine the impact of context along all three dimensions of affect control theory, compare evaluations of messages with and without context, and compare individual ratings with automated scores given by sentiment analysis tools.

Data

We utilize a subset of a study where 96 individuals collectively rated 5,780 Twitter posts [82]. In the broader study, individuals were given a brief 5-minute training on the three dimensions of ACT, which can be viewed in the technical report [82]. Individuals then each rated 120 Twitter posts three times, once for each dimension of ACT. The 120 Twitter posts evaluated fall into four categories: A) individual Twitter posts, B) responses to Twitter posts, C) the original post that response posts were made to, and D) the same responses seen in category B) – presented this time with the context of the original post. This paper focuses on the changes in response that individuals had from rating category B) tweets to category D) tweets.

Each set of 120 Twitter posts were evaluated twice. We only considered Twitter conversations where the original post was not a response itself. To ensure a broad diversity of topics, we chose Twitter posts from four broad areas, as outlined in the table below.

Table 18. Topic categories for data used.

	Nuclear	Arab Spring	General	Haiyan
Dates	Sep 2014 – Oct 2014	Oct 2009-Nov 2013	Sep 2013 – Aug 2014	Nov 2013 – Dec 2013
Sample Keywords	Nuclear proliferation, heavy water, uranium	Tahrir Square, Arab Spring	n/a	Haiyan, Typhoon Yolanda
Number of Twitter Posts	720	720	720	720

For "General" posts we randomly selected English-language posts from the "Gardenhose", or 10% of the total Twitter firehose, so we did not utilize keywords to select the topics.

Comparing responses with and without context

We first explore the data by displaying the distribution of ratings across message categories. We then perform a deeper dive into the different topics making up the dataset and show that we see the same behavior in changed evaluations across all topics. This allows us to make generalizations about the data as a whole and not limited to a subset of our data.

In the histograms below we plot the overall ratings that individuals recorded. Ratings are on a five point Likert scale from negative to positive for Evaluation, weak to strong for Power, and active to passive for Activity. We see that within Power and Activity, the overall profile of responses is consistent whether the post is the original post, the response, or the response viewed with context. The most variation appears to be within Evaluation, which sees slightly more negative posts in responses.

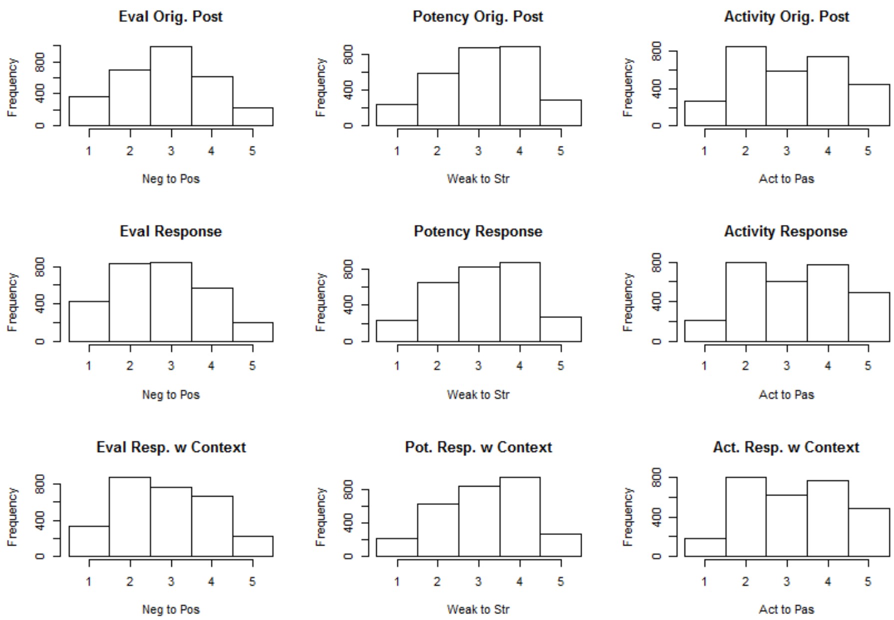


Figure 22. Histogram of responses across ACT dimensions and post category.

There is some minor variation across topic categories, but there is significant robustness when comparing differences in the evaluation of responses with and without context.

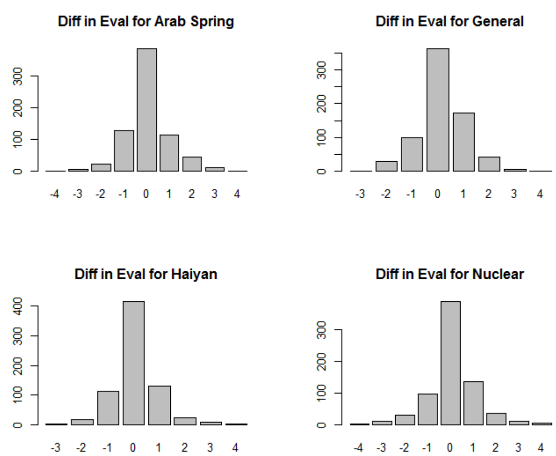


Figure 23. Difference in evaluation ratings of responses with and without context

We see that in all four categories, we see substantially similar distributions of differences in evaluation across the four categories. The largest bin of changes across all four topics is no change. There is a slightly larger number of individuals changing their evaluations to more negative in Arab Spring tweets.

In repeating this analysis for the other two dimensions of ACT, we see a similar pattern unfold – that regardless of the source of the data, there is a significant amount of change occurring across all three dimensions of Affect Control Theory. We now describe these changes more quantitatively and show that a similar analysis on simulated data does not yield the same result.

Features of responses with context

While the histograms give the appearance that the most common change in ratings after seeing context is no change - half the time, individuals are, in fact, changing their ratings. 46% of Evaluations were changed upon seeing context, 50% of Potency ratings were changed, and 52% of Activity ratings were changed.

Table 19. Table of features of changed ratings. Changed Total and Changed Valence percentages are based on all responses; other percentages are based on the number of responses that changed valence.

	Evaluation	Potency	Activity
Changed Total	1,329 (46%)	1,439 (50%)	1497 (52%)
Changed Valence	905 (31%)	1140 (40%)	1138 (40%)
Changed to Neutral	316 (35%)	391 (34%)	360 (32%)
Changed to Pos./Str./Act.	341 (38%)	430 (38%)	375 (33%)
Changed to Neg./Weak/Pas.	267 (30%)	329 (29%)	419 (37%)

In fact, at least 30% of post ratings changed valence after seeing context – 40% for Potency and Activity ratings. Since all ratings were made on a five point Likert scale, we considered all ratings to be one of 3 valences: Negative, Neutral, or Positive for Evaluation; Weak, Neutral or Strong for Potency; and Passive, Neutral, or Active for Activity.

We find that of the posts which changed valence, changes were made relatively uniformly – to either positive/strong/active, neutral, or negative/weak/passive – in overall similar numbers, with about one third of the posts that changed valence going to each category.

We investigated whether viewing context made it more likely to make a post be perceived as being more extreme or whether it largely attenuated ratings. Of posts that changed ratings, 22%, 18%, and 23% of ratings respectively for Evaluation, Potency, and Activity changed to extreme positions. It appears that it is more likely to attenuate an overall rating – while there are larger numbers of neutral ratings in general, a larger proportion of those posts that changed valence across all dimensions of ACT changed to neutral as opposed to changing to a more "extreme" position on the Likert scale.

Validation

To validate these findings, we created two simulated datasets with similar summary properties as our data to highlight how the results we obtain are not simply due to data manipulation. Two simulated datasets were used because of uncertainty in the underlying distribution of responses. Each simulated dataset replicates one third of the responses for a given topic area, so there are 12 paired sets of 90 draws.

The first simulated dataset is drawn from a binomial distribution with four draws and a probability of success of 50%. The second simulated dataset is drawn from a multinomial distribution with five bins with probabilities matching the distribution of categories in the Evaluative dataset. As in the original experiment, where we had two individuals evaluate the same data, we duplicated this data and randomly replaced half of the simulated data to ensure our data had a weighted Cohen's kappa around 0.60, again in line with the data collected from the in-person study.

Table 20. Table of summary statistics comparing binomial and multinomial simulated data

	Eval.	Potency	Activity	Binom.	Multi.
1st Quartile	2	2	2	2	2
Median	3	3	3	3	3
Mean	2.8	3.1	3.2	3.0	2.9
3rd Quartile	4	4	4	4	4
Std. Dev.	1.1	1.1	1.2	0.98	1.1

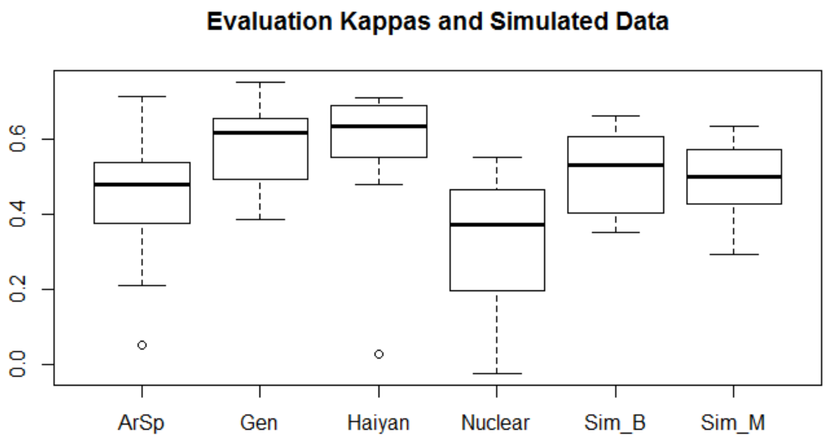


Figure 24. Distribution of Kappas across topic areas and for simulated data; 'Sim_B' indicates data drawn from the binomial distribution, 'Sim_M' indicates data drawn from the multinomial distribution.

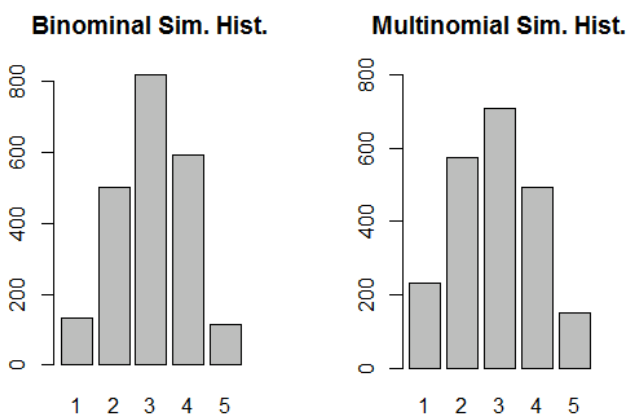


Figure 25. Histogram of binomial and multinomial simulated data sets.

We find that when comparing our simulated data with difference ratings seen with and without context, the simulated data has a considerably larger variance. In addition to this larger variance, significantly more respondents choose not to change their rating when compared with our randomly generated data.

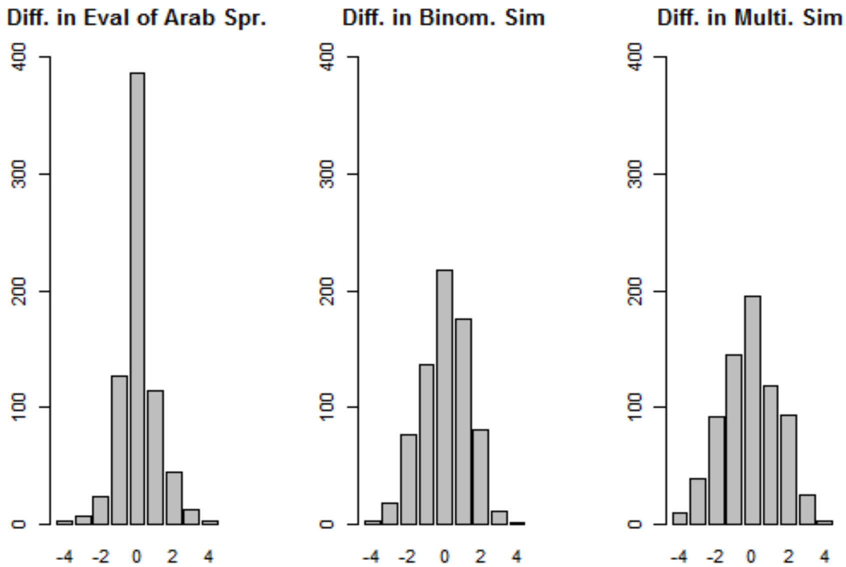


Figure 26. Histogram of difference in evaluation ratings for Arab Spring contrasted with difference in ratings taken from simulated data.

These results show that a key finding of our original study – that about 50% of all ratings change after re-evaluating the message with context – is not simply an artifact of data manipulation.

Table 21. Table of difference statistics, compared with binomial and multinomial simulated data.

	Eval.	Pot.	Act.	Binom.	Multin.
Mean	0.10	0.04	0.02	0.03	-0.13
Variance	0.94	1.3	1.5	1.7	2.4
Mean count of '0' ratings across topic areas	388	360	346	217	196

Implications for current tools

We evaluated the implications of these findings for current sentiment analysis tools in use. We used VADER [83], as well as the most recent ACT lexicon [81] and the CASOS Universal

Thesaurus to create a simple sentiment analysis tool that matched n-gram expressions within the Twitter messages – all dictionary methods that are the current standard approach for sentiment analysis tools due to the problem of sparse training data given the short length of Twitter messages [3]. Due to the relatively low kappa ratings, a larger machine learning model was not created – as that would have compounded the problem of fragile machine learning algorithms overtraining on incorrect data.

We found through sensitivity analysis that changing the window of what was considered a "neutral" message to being a score from (-0.1,0.1), to (-0.05, 0.05), to (-0.01, 0.01) did not significantly change overall accuracy rates of the sentiment analysis tools used. We set 0.05 as the window for neutral messages for both of the following tables.

Table 22. Sentiment Analysis Tool Matching Rates for Evaluation with neutral score window of 0.05

	VADER	**Universal Thesaurus**	**ACT**
Original Message	51%	35%	39%
Response	52%	33%	34%
Response with Context	50%	35%	35%

Table 23. Sentiment Analysis Match rates for Power and Activity using ACT Lexicon, neutral score window of 0.05

	Power	**Activity**
Original Message	39%	34%
Response	37%	29%
Response with Context	38%	29%

We see that overall sentiment analysis tool ratings appear to match response ratings – as well as original message ratings – at relatively low rates. While our data shows that individuals do change their perceptions of social media messages once they view the message in context, it is harder to draw a connection between automated evaluations of sentiment and what these perceptions are. Future work should further examine the role of size of neutral-rated messages and see if this significantly impacts overall accuracy ratings of sentiment miners.

We take a closer look at match ratings by identifying datasets that had high kappa and datasets that had low kappa. We isolated the ten highest and ten lowest kappa ratings for each axis of ACT; in taking our study, raters had different agreement rates for each axis. All subsets incorporated datasets from each topic group. The table below shows the ranges of the kappas for the data analyzed.

Table 24. Ranges of 10 highest and 10 lowest weighted kappas for each ACT axis.

Evaluation		**Potency**		**Activity**	
Low	High	Low	High	Low	High
-0.023-0.37	0.66-0.75	-0.33-0.007	0.33-0.49	-0.13-0.042	0.27-0.34

While we would expect a higher match rate for the subset with higher kappas, we find that overall match rates are identical to the overall population. These rates are not significantly improved by looking at the average rating provided by both raters.

Table 25. Match rates for Evaluation tools, contrasting 10 highest and 10 lowest kappa datasets

	Highest Kappas			**Lowest Kappas**		
	VADER	UT	ACT	VADER	UT	ACT
Original Message	47%	36%	35%	46%	38%	38%
Response	42%	41%	42%	47%	34%	32%
Response with Context	40%	44%	40%	47%	33%	34%

Table 26. Match ratings for Power with neutral window of 0.05

	Highest Kappas	**Lowest Kappas**
Original Message	40%	38%
Response	38%	33%
Response with Context	40%	35%

Table 27. Match rates for Activity with neutral window of 0.05

	Highest Kappas	Lowest Kappas
Original Message	36%	35%
Response	28%	27%
Response with Context	29%	29%

Discussion

Social media is a dynamic communication medium – useful for a variety of policy applications, from tracking extremist groups to guiding soft power efforts internationally to raising social awareness. Social media messages are inherently social – they are messages that are meant to be shared and disseminated across platforms. In this study, we have limited our analysis to short conversation snippets on Twitter, and we have only examined the text messages contained in those social media posts. However, many platforms also allow embedding more dynamic media – from GIFs to memes to YouTube videos.

Understanding social contagion and the dynamics of social movements requires understanding the context that these movements come out of. Messages are always viewed in context: for example, a popular online hashtag, #NetflixAndChill, while sounding innocuous, refers to a casual sexual encounter – and quickly served as a shibboleth for 'hip' internet users. Understanding the context surrounding the hashtag requires readers to be aware of considerably more than the current 140 characters Twitter allows in messages. If we are going to quantitatively assess these movements and understand how this change is proliferating across social media, we need to develop better tools that can capture and reflect the ratings of individuals reading and responding to these messages.

The implications of this finding on measuring soft power sentiment: additional structural considerations need to be taken when measuring and observing online discussion of topics. While it is useful to aggregate and distinguish social media posts by their immediate sentiment, additional consideration must be taken to couch posts in the structure of online conversation. If there are several unique posts about a topic, it is going to be more informative to do an analysis of the original posts instead of simply analyzing and aggregating responses to the posts, many of which may be a simple endorsement of the original message. While different social media platforms are able to provide different levels of access to their underlying social network structure, future researchers utilizing social media should try and utilize and incorporate that structure into their sentiment analysis and overall assessment of the platform.

A condensed version of this chapter can be found in the 2016 proceedings of the International Conference on Social Computing, Behavioral-Cultural Modeling & Prediction and Behavior Representation in Modeling and Simulation (SBP-BRiMS), as part of volume 9708 of the Springer Lecture Notes in Computer Science.

Data Appendix

VADER, unlike the other sentiment analysis tools, has a persistent bimodal distribution with modes at -.5 and -.5, so changing the neutral match rates does not significantly impact match rates. However, expanding the neutral match window does negatively impact ACT potency and activity ratings.

Table 28. Sentiment analysis match rates with neutral window of 0.01

	Evaluation			**Pot.**	**Act.**
	VADER	UT	ACT	ACT	ACT
Original Message	51%	35%	40%	42%	38%
Response	52%	32%	40%	40%	34%
Response with Context	50%	35%	39%	40%	33%

Table 29. Sentiment analysis match rates with neutral window of 0.10

	Evaluation			**Pot.**	**Act.**
	VADER	UT	ACT	ACT	ACT
Original Message	51%	35%	37%	35%	26%
Response	52%	33%	34%	33%	24%
Response with Context	50%	33%	33%	34%	25%

The highest correlation across sentiment analysis tools was between Potency and Activity at 0.5 – followed by Evaluation and Potency at 0.42. Potency and activity did not seem to have significant correlation with VADER or the universal thesaurus.

Table 30. Table of correlations across VADER, Universal Thesaurus, and ACT Lexicon scores.

Rating	**VADER**	**UT**	**Eval.**	**Pot.**	**Act.**
VADER	X	.33	.37	.08	.09
UT		X	.23	.11	.01
Eval			X	.42	.36
Pot.				X	.50

The standard statistic used to measure inter-rater reliability is Cohen's kappa, with 1 indicating "perfect" agreement and 0 indicating no agreement. Cohen's kappa considers the potential of errors and inter-rater agreement due to chance. For these boxplots, we show the weighted Cohen's kappa, which penalizes more heavily for disparate ratings and is appropriate for the ordinal scale used in the SOLO study.

Overall, we see higher Cohen's kappa for Evaluation and the lowest kappa for Activity. The Nuclear topic has the highest amount of inter-rater disagreement, while Haiyan has the highest rate of agreement.

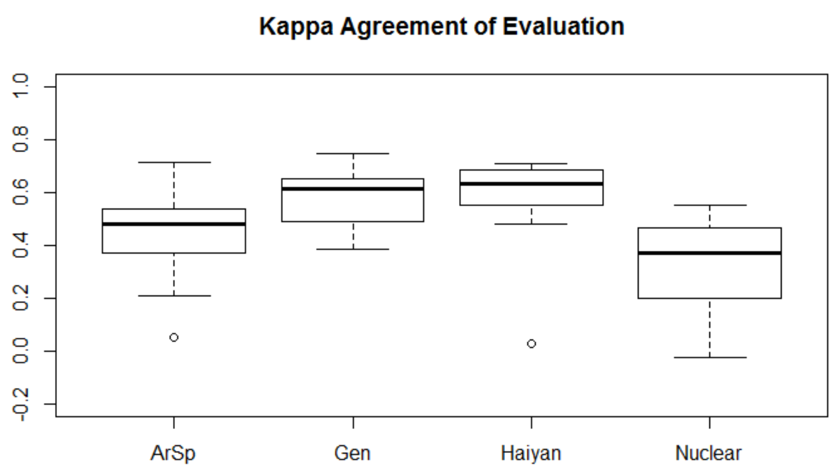


Figure 27. Weighted Cohen's Kappa of Evaluation broken out by category

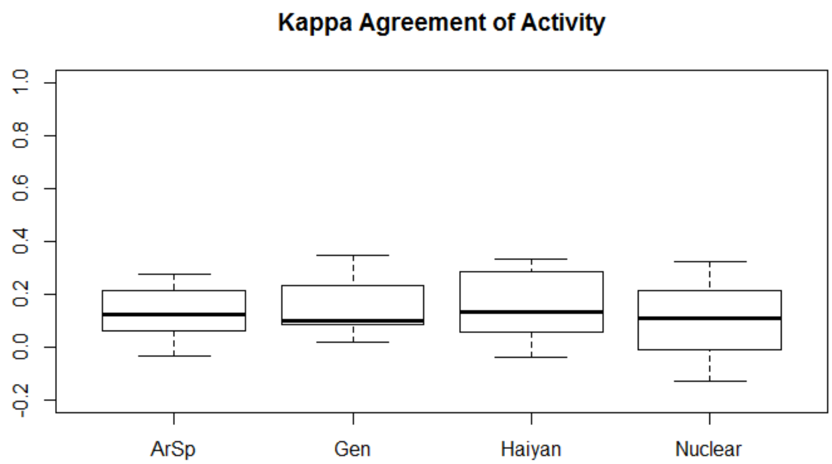


Figure 28. Weighted Cohen's Kappa of Potency broken out by category

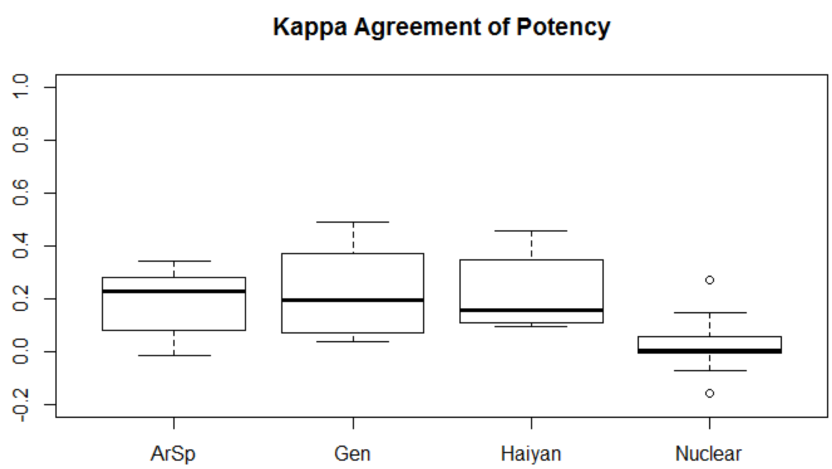


Figure 29. Weighted Cohen's Kappa of Activity broken out by category

Chapter 5: Nuclear Tweets

One of the major challenges of modern intelligence is the changing role of the analyst – traditionally the go-between for source collectors and policymakers, and now an individual responsible for increased interaction with policymakers[84], [85]. These analysts are required to provide intelligence products that address increasing levels of complexity – while past intelligence products provided concrete solutions, current products address scenarios and attempt to identify key actors with unclear specific objectives for policymakers[86].

In this case, a new data source available to analysts is social media. While the Obama administration made social media a central part of its engagement, including creating a new Twitter account specifically to highlight issues and questions the public had about the nuclear agreement with Iran, it is unclear what was the ongoing discussion surrounding the Iranian nuclear agreement. This chapter introduces the Twitter dataset, highlights some basic descriptions, and discusses policy implications of having a technology-focused social media search. We provide a quick overview of the entire dataset, highlight sentiment trends, and then do a deeper analysis of tweets related to the Iranian negotiations. We contrast this with topics discussed by known nuclear experts, identified using snowball sampling based on Washington, DC focused nonproliferation think tanks.

Background

The Nuclear Twitter dataset is a collection of over 7.5 million tweets made from September 2014 through November 2015 based on generic terms related to nuclear technology and policy (this is available as an appendix at the end of this chapter). We focus on a subset of tweets discussing

Iran – the framework of the nuclear agreement was announced at the beginning of April 2015 and brought with it considerable online discussion.

There are two major kinds of 'noise' analysts are concerned with in social media – the primary concern is identifying relevant messages, and the secondary concern is identifying similar but fundamentally identical ideas and concepts. This section will discuss the first concern – 'noise' related to identifying relevant messages – while the second concern will be discussed in the section discussing hashtags and parsing social media messages. While the primary keywords utilized to populate this dataset relate to nuclear weapons technologies and nuclear deterrence keywords, we focus on a subset of the data to specifically filter out 'noise' tweets – tweets not relevant to the overarching purpose of this paper. For example, earlier iterations of this data collection included a high proportion of tweets South-East Asia – a contestant act on "Indonesia's Got Talent" included the word "Uranium" in their stage name. In early testing of keywords, I removed "Heavy water" as a keyword due to the high number of messages that included both terms – but used in non-nuclear contexts, such as "Heavy classwork today – and soaked by the rain. Water everywhere." After removing this keyword (and the translation of this keyword) from the collection, the number of tweets dropped dramatically – but the proportion that appeared to be relevant increased dramatically. By utilizing a primary collection focused on one set of technologies, and then utilizing a second set of keywords to further identify relevant messages, we are able to dramatically increase the overall proportion of tweets in the collection that are relevant to the topic at hand. It does not completely eliminate the problem of spam or malicious bots, but it does provide researchers with added confidence that the messages being studied are relevant and not simply messages that utilize the same keywords in unforeseen contexts.

One of the advantages of utilizing large datasets is that while this concern about noise remains relevant, the analysis focuses on broad trends and averages. While there may be ‘noise’ in the sense of extraneous tweets not actually related to the topic at hand (e.g. “I got so angry at work today I wanted to nuke Tehran!”), these types of messages generally do not accumulate, or generate sufficient interest for us to be concerned about. If the message was made by a key decision maker, such as a US Senator, it’s likely the message would generate a lot of attention – and it would be something worth assessing and analyzing. Similarly, if a variant of the message appeared in multiple user posts, we should be able to detect similar words appearing through topic modeling and would then be able to do a deeper analysis of how that message has spread.

Since we are utilizing two sets of keywords to identify the tweets for further study, it is reassuring that the overall proportion of tweets from the larger nuclear tweet dataset is fairly consistent at around 20%. In identifying broader topic clusters, as well as sentiment trends, noise does not have a significant impact on our research findings. It is most apparent in the section that links geotagged Twitter messages and sentiment scores – primarily because there are relatively few geotagged messages at all. Tweets appearing in countries where Twitter is banned (e.g. Iran, China, North Korea) are particularly suspect and may only be there due to geotag spoofing (or user VPN use).

Table 31. Breakdown of Nuclear tweets dataset and Iran subset

	Total count of tweets	Iran Tweets	Percentage of total
September 2014	780,403	148,466	19%
October 2014	796,216	150,516	19%
November 2014	33,159	4,965	15%
January 2015	558,364	115,432	21%
February 2015	974,554	341,527	35%
March 2015	1,028,835	227,808	22%
April 2015	3,064,151	658,535	21%
May 2015	23,653	2,411	10%
September 2015	9,053	86	1%
October 2015	252,886	9409	4%
November 2015	22,740	347	2%

Due to some problems with collection (including equipment failure), collection was not continuous. We focus on six months in particular: September and October of 2014 and January-April 2015. A distinguishing feature of this dataset, compared to most collections of social media data, is the high proportion of retweets.

Table 32. Engagement statistics for Nuclear dataset and Iran tweet subset

	Responses	Retweets	Iran - Responses	Iran - Retweets
Sep 2014	6%	42%	1%	41%
Oct 2014	6%	43%	2%	45%
Jan 2015	6%	51%	2%	61%
Feb 2015	5%	51%	2%	60%
Mar 2015	4%	56%	2%	46%
Apr 2015	5%	52%	2%	51%

A major concern surrounding the use of social media data is the use of different platforms by malicious actors or automated bots to hijack ongoing conversations. The engagement statistics reported in the table above, with over 50% of tweets counting as retweets, or re-broadcasts of existing messages, as well as 99% of all messages in the dataset containing a URL or being a retweet, does raise immediate concerns about the veracity of overall user engagement with the topic being discussed. While an imperfect science, there are a wide range of different heuristics utilized by analysts to detect fake posts in similar keyword-based Twitter message collections: detecting spam URLs, unrelated hashtag use, and unique characteristics of both the mention and social networks of spam accounts [87]-[91]. In analyzing the Iran subset of data, I did not find any highly suspicious accounts, URLs, hashtags, or mention or social networks that would indicate hijacking of the dataset by malicious actors. While sufficient for this chapter's analysis, future work analyzing the nuclear Twitter dataset as a whole, or alternative subsets of the dataset, should be analyzed to check against these sorts of actors.

A breakdown of their sentiment values, utilizing the multilingual sentiment tool developed by CASOS and built incorporating rules developed by VADER[83], [92]. Messages were considered "neutral" if their rating was between -0.05 and 0.05.

Table 33. Sentiment profile of overall nuclear tweet dataset over time.

	Positive	Neutral	Negative	% in English	Total
Sep 2014	34%	20%	45%	79%	780,403
Oct 2014	35%	23%	42%	78%	796,216
Jan 2015	34%	25%	40%	83%	558,364
Feb 2015	34%	22%	43%	80%	974,554
Mar 2015	31%	13%	55%	40%	1,028,835
Apr 2015	33%	18%	49%	42%	3,064,151

Noticeable here – English takes up the majority of tweets until March and April. The second largest collection comes from Spanish – the languages are broken out below. These languages are tagged according to Twitter's built-in language detector. It is noticeable that Arabic surges ahead to 5% in April, while there are almost no Persian tweets. While Twitter may be popular for an international tool of discussion, as well as for Iranian political leaders, the platform remains banned from within the country.

Table 34. Additional languages in March and April 2015 providing at least 90% of language cover for these months

	March	April
Portuguese	4%	4%
Arabic	0%	5%
Turkish	8%	15%
Spanish	45%	27%

Table 35. Sentiment profile for Iran tweets and English language breakdown

	Positive	Neutral	Negative	% in English	Total
Sep 2014	32%	16%	51%	97%	148,466
Oct 2014	28%	23%	49%	93%	150,516
Jan 2015	38%	21%	41%	96%	115,432
Feb 2015	37%	22%	41%	97%	341,527
Mar 2015	41%	16%	42%	95%	227,808
Apr 2015	53%	12%	34%	91%	658,535

Top users in Iran dataset

Twitter has three forms of social network: a retweet network, the mention network, and the follow network. The retweet network is a simple re-broadcasting network: if two users in a network see the message posted by the other, they can choose to re-tweet and broadcast the other's message. The mention network is similar to a conversation on Twitter, or tagging users to alert them to a story they might be interested. The follow network is the list of individuals that Twitter users can immediately "see" in their social network. The follow network is directed – for example, User A can 'follow' User B – but that does not guarantee that User B 'follows' User A back. Many official accounts, such as the accounts of news organizations, are followed by many users, but they themselves have a much smaller 'follow' network.

Unlike traditional social networks, where measures of centrality can be utilized to identify key users and actors, Twitter as a data source reduces the number of links in a way that hides the true social network when assessing Twitter dataset.   For example, in a retweet network, User C might broadcast an interesting message that is seen and re-tweeted by User B. Suppose User A follows User B and sees User C's message and also decides to re-tweet it – the brokerage activity of User B, who provided the actual link between User A and User C, is not captured in this data, and the data simply captures that both Users A and B retweeted User C.

In this dataset, we are interested in two features of users: who is frequently re-tweeted, and who has potential reach. To do this we identify accounts with high out-degrees in the retweet network, as well as users with high k-closeness using weighted data, including follower count (the order of the user's followed network) [93], [94].

We measure "top" users in two ways: identifying Twitter accounts with high out-degrees in the retweet network, and identifying potential reach – calculated as k-closeness using weighted data from both the retweet network as well as the mention network.

Table 36. Twitter accounts with High Out-degree in retweet network over time in Iran dataset

Sept 2014	Oct 2014	January 2015
Reuters	Reuters	SenTedCruz
HassanRouhani	AdamMilstein	SenatorKirk
IsraeliPM	Maryam_Rajavi	Peymaneh123
FoxNews	ALNAQ33B	Kredo0
AP	Al3AMEED_70	JavadDabiran
UN	Nasrinfoiran	paydaran
Iran_policy	Mojahedineng	IranNewsUpdate1
BBCBreaking	4FreedominIran	BehzadMoezi
YahooNews	IranNewsUpdate1	mdubowitz
Mojahedineng	Iran_policy	France24
February 2015	**March 2015**	**April 2015**
Netanyahu	AmyMek	Nytimes
SenTedCruz	AP	WhiteHouse
Cnnbrk	WSJ	Thehill
AJEnglish	SenTomCotton	Netanyahu
Ggrenwald	DineshDSouza	Reuters
AdamMilstein	GovMikeHuckabee	RT_com
RT_com	Netanyahu	TheEconomist
Iran_policy	Gerfingerpoken	AdamMilstein
PatDollard	Cnnbrk	Democracynow
haaretzcom	weknowwhatsbest	iranianaffairs

In the retweet network, we note that as news of the Iranian negotiation comes closer, the most retweeted accounts switch from news accounts and Iran policy enthusiasts to being dominated by American politicians and media figures, such as Senator Ted Cruz, Senator Kirk, Senator Tom Cotton, and Governor Mike Huckabee.

In the k-betweenness measure, we find a slight shift towards individual reporters, instead of direct official news accounts. These reporters are, however, drowned out by the news in April surrounding the framework agreement in April, where official news accounts and individual politicans note the occasion.

Table 37. High k-betweenness outreach potential

Sept 2014	Oct 2014	January 2015
FoxNews	FoxNews	FoxNews
CNN	Lrozen	WSJ
WSJ	Rezaaslan	Lrozen
BBCWorld	AlMonitor	BBCKasraNaji
Lrozen	Tparsi	AlMonitor
HamzeiAnalytics	Ginamcnaughton	ABC
CBSNews	AP	McFaul
AlMonitor	BrookingsFP	Whitehouse
Ceydak	IranNewsNow	Jerusalem_post
BreakingNews	mehdirhasan	greta
February 2015	**March 2015**	**April 2015**
Cnnbrk	AP	JZarif
CNN	Gordonnyt	JohnKerry
FoxNews	bklapperAP	Nytimesworld
Lrozen	Reuters	Nytopinion
LOLGOP	NikiBlasina	BarackObama
AlMonitor	Cnnbrk	Netanyahu
IranNewsNow	Adamschreck	HassanRouhani
Reuters	Perry_dan	Nytpolitics
Nprpolitics	Laurnorman	WhiteHouse
BBCWorld	WSJ	nytgraphics

Topics

We assess topics in two ways: we first examine the distribution of sentiment across tweet keywords, and then run Latent Dirichlet Analysis to identify clusters of topic groups and overlap between different topic groups [95], [96].

We first assess the distribution of hashtags being used to determine whether there are any particular hashtags that could be significantly influencing sentiment. We find that while there are a few clear cases – e.g. using #No2Rouhani and #tcot (Top Conservatives on Twitter, used widely by right-wing groups) – for the most part, the most widely used hashtags for the dataset broadly seem to be the most widely used hashtags regardless of the tweet's rated sentiment.

The tables below highlight the top hashtags used by tweets for each month; key terms that were also used to identify the tweets (either as part of the nuclear twitter dataset or used to filter tweets relative to Iran) were removed. The table reports absolute counts, and the percentage sign represents the percentage of overall sentiment class for that month. For example, No2Rouhani appeared in September 2014 in 15,924 tweets in the dataset, in 21% of all negative tweets that month. The hashtag appeared in 13% of all tweets that month. Hashtag sentiment tables for the remaining months are available at the end of this chapter as an appendix.

Table 38. Hashtag sentiment profile table for September 2014.

	Negative	Neutral	Positive	Total
No2Rouhani	15924 (20.9%)	1762 (7.2%)	1467 (3.1%)	19153 (12.9%)
IRANTALKS	8607 (11.3%)	1287 (5.3%)	1612 (3.4%)	11506 (7.7%)
news	3400 (4.5%)	2583 (10.6%)	1109 (2.4%)	7092 (4.8%)
IranTalksNYC	5229 (6.9%)	607 (2.5%)	1178 (2.5%)	7014 (4.7%)
usa	3295 (4.3%)	452 (1.9%)	469 (1%)	4216 (2.8%)
Iraq	3220 (4.2%)	273 (1.1%)	340 (0.7%)	3833 (2.6%)
ISIS	1226 (1.6%)	2078 (8.5%)	485 (1%)	3789 (2.6%)
IAEA	2925 (3.8%)	34 (0.1%)	513 (1.1%)	3472 (2.3%)
Syria	2606 (3.4%)	269 (1.1%)	360 (0.8%)	3235 (2.2%)
EU	2840 (3.7%)	113 (0.5%)	69 (0.1%)	3022 (2%)

Both LDA and the hashtag sentiment profile tables shows that online discussion of the Iranian negotiations and nuclear deal that have been going on for a persistent time. However, the fact that many of the recurring hashtags in the LDA topic clusters don't overlap indicates many of these discussions are happening in isolation. While the fact that individuals are predominantly using the same hashtags regardless of sentiment indicates that posts can be found using the same language, there is still a barrier when it comes to engagement across different topics.

For consistency, LDA was run to identify five clusters of topic groups for each month. A complete list of LDA tables is available in the appendix. LDA topic clusters were run for both hashtags as well as parsed Twitter messages. While hashtags are more easily identified in the

Twitter metadata, and broadly sufficient given media dominance in this collection, utilizing parsed Twitter messages makes it easier to identify common topics across different clusters. It also clearly identifies key decision makers.

Parsing Twitter messages brings its own concern surrounding generating 'noise' in the data. To try and reduce the amount of 'noise' generated by parsing messages, the CASOS Universal Thesaurus was utilized, which allows disparate words such as "Barack" and "Obama" to both be matched with the concept "Barack Obama". For each month, concepts found in over 100 messages were reviewed and reduced to ensure that the LDA topic clustering algorithm did not cluster on an incorrectly parsed acronym or common typo (e.g. 'nuclear' and 'nuklear').

Table 39. LDA Table for September 2014 Iran Data.

Topic A	Topic B	Topic C	Topic D	Topic F
Irantalks	Iran	Iran	Iran	Iran
Iran	nuclear	nuclear	nuclear	World
IAEA	Israel	news	United.Nations	nuclear
nuclear	Netanyahu	ISIS	tehran	Iraq
No2Rouhani	bad	Zarif	threat	news
MaryamRajavi	nuclear_weapons	Kerry	Islamic	Politics
Rouhani	Barack_Obama	deadline	Site	LONDON
NYC	regime	progress	Secret	Syria
EU	cease	United_States_of_America	uncertain	FOX
de	warn	Reuters_Ltd	amp	United_States_of_America

One major takeaway of running this analysis was that there was persistent concern over the secrecy of past Iranian projects to develop a nuclear capability, as well as a persistent topic involving the People's Mujahedin of Iran (PMOI). However, there was no significant discussion of the technical nuclear issues at hand in the negotiation, such as stopping centrifuges or shutting down the plutonium-producing reactor at Arak. To further explore this, I took a quick sample of major nonproliferation policy think tanks in Washington, DC, as shown in the table below, and identified the top twitter individuals associated with these centers to extract individual discussions.

Table 40. Short List of Nonproliferation Policy Organizations

Abbreviation	Organization Name	Twitter Handle
ACA	Arms Control Association	armscontrolnow
ANS	American Nuclear Society	ans_org
AC	Atlantic Council	ACScowcroft
Belfer	Belfer Center	BelferCenter
CACNP	Center for Arms Control and Nuclear Proliferation	Nukes_of_Hazard
CNPP	Carnegie Endowment for International Pace	carnegienpp
CISAC	Center for International Security and Cooperation	StanfordCISAC
CNS	Center for Nonproliferation Studies	CNS_Updates
CSIS	Center for Strategic and International Studies	csisponi
FAS	Federation of American Scientists	FAScientists
UCS	Union of Concerned Scientists	UCSUSA

In reviewing tweets by these organizations, many posts were for official events or promotional messages about the larger organization these offices belonged to. To mitigate the effects of these curated Twitter accounts, I focused on users affiliated with those organizations to identify topics of discussion. In addition to users referenced by these organizations, I included the official twitter accounts of prominent American and Iranian diplomats and officials involved in the Iranian negotiations. In the case of Russian Forces, as the project is operated by a primary researcher with no other official Twitter account, it was included as a data source for topic modeling.

Table 41. Table of Twitter handles used for "Expert" topic modeling.

Affiliation	Twitter Handle
ACA	KingstonAReif
ACA	DarylGKimball
ACA	GregThielmann
CACNP	nukes_of_hazard
CNPP	toby_dalton
CISAC	SecDef19
CNS	ArmsControlWonk
CNS	mhanham
CNS	atomic_pickles
CNS	CNS_MasakoToki
CNS	sclieggi
FAS	cdfergusonII
FAS	saftergood
Iran Official	khamenei_ir
Iran Official	HassanRouhani
Iran Official	JZarif
Russian Forces	russianforces
UCS	StephenUCS
US Government	Gottemoeller
US Government	ernestmoniz
US Government	johnkerry
US Government	potus
US Government	theirandeal

After running LDA topic clustering, I find that many of these experts are also not explicitly discussing the technical issues surrounding the Iranian negotiations. However, there are clear clusters of discussion surrounding distinct countries – Iran, Russia, and North Korea – as well as some unique nuclear proliferation conferences that are discussed. Tables for three months: April 1-14 of 2016, as well as March 15 and April 2015 – are available in the appendix.

Geographic distribution of sentiment

Finally, we assess the global distribution of sentiment. To assess these, we further screen tweets to only focus on tweets that have a geographic tag – typically provided by a user's phone if used to tweet. We then extrapolate from the GPS coordinates the country of origin for the sentiment expressed.

The tables below highlight the classification of tweets and show the percentage of tweets that come from the United States. In this case, the table has the absolute count of tweets, followed by the percentage of the number directly below it. For example, there are 73 negative sentiment geotagged tweets from the United States – which is 50% of all negative geotagged tweets in the data in September 2014. Altogether, these 352 geotagged tweets represent 0.2% of all tweets in the data related to Iran for the month.

The reason for calculating and showing percentages in this way is to highlight whether countries carry the primary negative, neutral, or positive voice in the global conversation surrounding the Iranian nuclear negotiations. For example, in March 2015, the United States directly contributed to 62% of the negative discussion surrounding the negotiation – as well as just 37% of the neutral discussion surrounding the topic.

Table 42. Geotagged count of tweets, and percent of global discussion, in Iran dataset.

		Negative	Neutral	Positive	Total
Sept 2014	United States	73 (50%)	20 (33.3%)	58 (39.7%)	151 (42.9%)
	Geotagged	146 (0.3%)	60 (0.2%)	146 (0.2%)	352 (0.2%)
	Total	47159	24416	76255	147830
Oct 2014	United States	93 (52.8%)	31 (59.6%)	53 (43.8%)	177 (50.7%)
	Geotagged	176 (0.4%)	52 (0.2%)	121 (0.2%)	349 (0.2%)
	Total	42210	34634	73299	150143
Jan 2015	United States	74 (67.3%)	16 (53.3%)	65 (55.1%)	155 (60.1%)
	Geotagged	110 (0.3%)	30 (0.1%)	118 (0.3%)	258 (0.2%)
	Total	43968	23904	47099	114971
Feb 2015	United States	204 (68%)	42 (54.5%)	114 (49.1%)	360 (58.9%)
	Geotagged	300 (0.2%)	77 (0.1%)	232 (0.2%)	611 (0.2%)
	Total	125844	74859	139753	340456
Mar 2015	United States	205 (61.7%)	38 (36.5%)	185 (53.8%)	428 (54.1%)
	Geotagged	332 (0.4%)	104 (0.3%)	344 (0.4%)	791 (0.3%)
	Total	92659	37245	96459	226363
Apr 2015	United States	343 (53%)	110 (45.5%)	437 (43.3%)	890 (46.9%)
	Geotagged	647 (0.2%)	242 (0.3%)	1009 (0.5%)	1898 (0.3%)
	Total	351938	80666	220797	653401

The United States dominates the geotagged set of tweets for Iran. However, in the broader dataset, while it remains the country with the largest number of geotagged tweets, there are other countries that also have significant numbers of posts. For contrast, here are the top five country sentiment profiles for September 2014 and April 2015. While some countries have relatively stable overall sentiment profiles – the United States and the United Kingdom have similar overall sentiment profiles related to nuclear tweets – Brazil and Argentina's profiles swing significantly. In September 2014, its sentiment is overall slightly negative at 42% of its tweets – which swings into being strongly negative by April 2015, at 72% of its tweets, mirroring Argentina's overall sentiment profile.

Table 43. Geotagged tweets in nuclear twitter dataset for September 2014 and April 2015; percentages refer to within country sentiment profiles except n the total column, where they represent the country's total share of geotagged tweets for that month.

		Negative	Neutral	Positive	Total
Sept 2014	United States	743 (42%)	382 (22%)	634 (36%)	1760 (29%)
	United Kingdom	513 (52%)	142 (14%)	326 (33%)	982 (16%)
	France	265 (84%)	18 (6%)	34 (11%)	317 (5%)
	Brazil	114 (43%)	71 (27%)	79 (30%)	264 (4%)
	Spain	109 (51%)	60 (28%)	46 (21%)	215 (4%)
	Argentina	78 (57%)	29 (21%)	29 (21%)	136 (2.3%)
April 2015	United States	1251 (41%)	662 (21%)	1173 (38%)	3086 (17%)
	Turkey	1344 (48%)	747 (27%)	718 (26%)	2809 (15%)
	Venezuela	789 (54%)	88 (6%)	574 (40%)	1451 (8%)
	United Kingdom	661 (50%)	184 (14%)	485 (36%)	1330 (7%)
	Argentina	793 (73%)	119 (11%)	174 (16%)	1086 (6%)
	Brazil	768 (72%)	141 (13%)	161 (15%)	1070 (6%)

Conclusion

Recent work in demand-side nuclear proliferation has focused on psychological profiles of national leaders [17]and government agency [16]. However, despite the fact that current research focuses on how political structure prevents countries from perusing nuclear weapons [97] [98], whether through audience costs or transparency measures, we are able to measure public reaction directly through social media today.

There are significant limits to this. Coverage, for one; like many analyses of social media, this data is heavily over-represented by media organizations, professional reporters, and politicians aiming to get their message out. In addition to this targeted data collection, a search of the Gardenhose data (10% of the complete Twitter firehose) housed at CMU did not yield any useful tweets related to nuclear proliferation. Utilizing geotagged tweets reflects a significant bias in favor of individuals who access Twitter through their smartphones – disproportionately rich, and disproportionally American, even when this data is captured overseas.

This work can potentially be utilized as inputs to the adapted Friedkin model in two ways: in identifying initial attitudes towards nuclear weapons, and identifying flows of positive and negative reaction to news events related to nuclear weapons. In the current adapted Friedkin model, all countries are initialized as being indifferent towards nuclear weapons; however, we can measure country-level sentiment towards nuclear weapons by identifying influential country-level accounts and assessing their attitudes towards nuclear weapons. This would allow us to initialize the model with higher accuracy.

Another output of the Twitter analysis would be to identify flows and reactions to other countries. By measuring the intensity and uniqueness of country to country sentiment, we can infer new potential hostilities, as well as country alliances and sympathies, that may not be apparent in the political network data currently available to researchers. This could be taken as an input to the Friedkin model to assess how these additional relationships change country behavior and motivation towards developing a nuclear capability.

However, this is only the start. Even with this initial start, we can see trends and changes in country-level sentiment towards certain topics; this is how we will measure the future effectiveness of diplomacy.

Appendix: Search Terms Used

To populate the nuclear twitter dataset, the following terms were translated into Spanish, Chinese, Japanese, Arabic, Persian, Portuguese, French, German, and Russian.

Nuclear
Nuclear energy
uranium enrichment
plutonium enrichment
plutonium
uranium
nuclear weapons

tritium
strategic deterrence
deterrence
deterrent capability

To identify tweets related to Iran:

Using regular expressions, any tweets with the following fragments were considered relevant to Iran.

Table 44. Table of terms used to identify tweets related to Iran

Fragment (case ignored)	Full Expression (Matched with)	Rationale
Iran*	Iran, Iranian	Country of interest; full name used to match country and avoid confusion with Iraq
Tehran	Tehran	Capital city of Iran
Khamen	Khmeni, @khamenei_ir, @khanmenei_fra	Iranian Supreme Leader
Rouh	Rouhani, @HassanRouhani	Current Iranian President
Zarif	Zarif, @JZarif	Current Iranian Minister of Foreign Affairs
Salehi	Ali Akbar Salehi	Head of Atomic Energy Organzation of Iran
JCPOA	Joint Comprehensive Plan of Action	Final official title of latest round of Iranian negotiation
JPOA	Joint Plan of Action	Initial title of working Iranian negotiations

Appendix: Hashtag Sentiment Profile Tables

Hashtag sentiment profiles for Iran twitter dataset. Cell value is an absolute count; percentage refers to the percent of all classed sentiment tweets that month with that hashtag.

Table 45. Hashtag sentiment profile for September 2014.

	Negative	Neutral	Positive	Total
No2Rouhani	15924 (20.9%)	1762 (7.2%)	1467 (3.1%)	19153 (12.9%)
IRANTALKS	8607 (11.3%)	1287 (5.3%)	1612 (3.4%)	11506 (7.7%)
news	3400 (4.5%)	2583 (10.6%)	1109 (2.4%)	7092 (4.8%)
IranTalksNYC	5229 (6.9%)	607 (2.5%)	1178 (2.5%)	7014 (4.7%)
usa	3295 (4.3%)	452 (1.9%)	469 (1%)	4216 (2.8%)
Iraq	3220 (4.2%)	273 (1.1%)	340 (0.7%)	3833 (2.6%)
ISIS	1226 (1.6%)	2078 (8.5%)	485 (1%)	3789 (2.6%)
IAEA	2925 (3.8%)	34 (0.1%)	513 (1.1%)	3472 (2.3%)

Syria	2606 (3.4%)	269 (1.1%)	360 (0.8%)	3235 (2.2%)
EU	2840 (3.7%)	113 (0.5%)	69 (0.1%)	3022 (2%)

Table 46. Hashtag sentiment profile for October 2014

	Negative	Neutral	Positive	Total
IranTalks	6038 (8.2%)	3961 (11.4%)	2647 (6.3%)	12646 (8.4%)
No2Rouhani	7487 (10.2%)	1048 (3%)	1218 (2.9%)	9753 (6.5%)
IranTalksVienna	4865 (6.6%)	2719 (7.9%)	1787 (4.2%)	9371 (6.2%)
UN	3457 (4.7%)	2732 (7.9%)	269 (0.6%)	6458 (4.3%)
NoNuclearIran	2100 (2.9%)	2576 (7.4%)	1039 (2.5%)	5715 (3.8%)
Iraq	2448 (3.3%)	1722 (5%)	648 (1.5%)	4818 (3.2%)
news	2045 (2.8%)	1221 (3.5%)	814 (1.9%)	4080 (2.7%)
USA	1863 (2.5%)	1034 (3%)	641 (1.5%)	3538 (2.4%)
humanrights	2355 (3.2%)	888 (2.6%)	180 (0.4%)	3423 (2.3%)
syria	1587 (2.2%)	890 (2.6%)	256 (0.6%)	2733 (1.8%)

Table 47. Hashtag sentiment profile for January 2015.

	Negative	Neutral	Positive	Total
IranTalks	7841 (16.6%)	3304 (13.8%)	7235 (16.5%)	18380 (16%)
usa	1491 (3.2%)	356 (1.5%)	2358 (5.4%)	4205 (3.7%)
sotu	1476 (3.1%)	399 (1.7%)	1838 (4.2%)	3713 (3.2%)
NEWS	584 (1.2%)	1943 (8.1%)	987 (2.2%)	3514 (3.1%)
IranTalksVienna	689 (1.5%)	376 (1.6%)	2315 (5.3%)	3380 (2.9%)
UN	57 (0.1%)	247 (1%)	1753 (4%)	2057 (1.8%)
nonuclearIran	240 (0.5%)	1501 (6.3%)	225 (0.5%)	1966 (1.7%)
Iraq	1337 (2.8%)	292 (1.2%)	310 (0.7%)	1939 (1.7%)
JohnKerry	2 (0%)	1833 (7.7%)	19 (0%)	1854 (1.6%)
HumanRights	43 (0.1%)	1325 (5.5%)	353 (0.8%)	1721 (1.5%)

Table 48. Hashtag sentiment profile for February 2015.

	Negative	Neutral	Positive	Total
Irantalks	24738 (17.7%)	8669 (11.6%)	20274 (16.1%)	53681 (15.8%)
News	4849 (3.5%)	2543 (3.4%)	5948 (4.7%)	13340 (3.9%)
IranTalksGeneva	5626 (4%)	1781 (2.4%)	3797 (3%)	11204 (3.3%)
US	4400 (3.1%)	893 (1.2%)	2946 (2.3%)	8239 (2.4%)
Iraq	2115 (1.5%)	2209 (3%)	3894 (3.1%)	8218 (2.4%)
UK	4203 (3%)	658 (0.9%)	2813 (2.2%)	7674 (2.3%)
PMOI	1257 (0.9%)	3168 (4.2%)	3134 (2.5%)	7559 (2.2%)
cnn	2547 (1.8%)	1541 (2.1%)	2515 (2%)	6603 (1.9%)
Phoenix	2625 (1.9%)	2891 (3.9%)	1070 (0.9%)	6586 (1.9%)
usa	2089 (1.5%)	1577 (2.1%)	2814 (2.2%)	6480 (1.9%)

Table 49. Hashtag sentiment profile for March 2015.

	Negative	Neutral	Positive	Total
IranTalks	5650 (5.9%)	1818 (4.9%)	4315 (4.7%)	11783 (5.2%)
Bahrain	1412 (1.5%)	1475 (4%)	2896 (3.1%)	5783 (2.6%)
tcot	3349 (3.5%)	369 (1%)	1715 (1.9%)	5433 (2.4%)
news	2812 (2.9%)	834 (2.2%)	1439 (1.6%)	5085 (2.2%)
US	2334 (2.4%)	458 (1.2%)	1485 (1.6%)	4277 (1.9%)
Israel	1609 (1.7%)	588 (1.6%)	738 (0.8%)	2935 (1.3%)
WakeUpAmerica	1147 (1.2%)	209 (0.6%)	1000 (1.1%)	2356 (1%)
PMOI	1344 (1.4%)	171 (0.5%)	616 (0.7%)	2131 (0.9%)
Obama	721 (0.7%)	420 (1.1%)	912 (1%)	2053 (0.9%)
UN	1214 (1.3%)	202 (0.5%)	529 (0.6%)	1945 (0.9%)

Table 50. Hashtag sentiment profile for April 2015.

	Negative	Neutral	Positive	Total
Irantalks	17739 (8%)	5130 (6.4%)	23115 (6.6%)	45984 (7%)
IranDeal	14934 (6.8%)	4753 (5.9%)	26075 (7.4%)	45762 (7%)
news	8706 (3.9%)	2337 (2.9%)	12822 (3.6%)	23865 (3.7%)
usa	5544 (2.5%)	1212 (1.5%)	12583 (3.6%)	19339 (3%)
Bahrain	4224 (1.9%)	1217 (1.5%)	8490 (2.4%)	13931 (2.1%)
UK	4267 (1.9%)	770 (1%)	6317 (1.8%)	11354 (1.7%)
US	5738 (2.6%)	1172 (1.5%)	3246 (0.9%)	10156 (1.6%)
tcot	3780 (1.7%)	1006 (1.2%)	4663 (1.3%)	9449 (1.4%)
MaryamRajavi	6351 (2.9%)	171 (0.2%)	1805 (0.5%)	8327 (1.3%)
Israel	3374 (1.5%)	1861 (2.3%)	2765 (0.8%)	8000 (1.2%)

Appendix: LDA Tables

The LDA Tables are presented in chronological order by type: the topic groups as determined by Twitter hashtags, followed by the topic groups as determined by parsed Twitter messages, followed by the three topic groups as determined by parsed expert Twitter messages for March 2015, April 2015, and April 1-14 2016.

Table 51. Hashtag September 2014 LDA Topics.

Topic A	Topic B	Topic C	Topic D	Topic F
World	Iran	Irantalks	Iran	Irantalks
Iraq	nuclear	Iran	nuclear	Iran

Politics	Israel	nuclear	news	IAEA
LONDON	US	No2Rouhan i	uk	No2Rouhan i
google	Tweetzup	NonuclearIr an	Spain	NYC
Syria	MaryamRaja vi	Rouhani	CampLiberty	EU
LeMonde	PMOI	newyork	ISIS	UNSC
Euronews	tcot	NuclearDeal	Senat	timetoact
FOX	UNGA	pictures	United.Nations	Facts
United_States_of_Ame rica	USAHeadlin es	negotiation	United_States_of_Ame rica	STATEDE PT

Table 52. Hashtag October 2014 LDA Topics.

Topic A	Topic B	Topic C	Topic D	Topic F
Iran	world	Iran	Iran	Iran
nuclear	CNN	US	Irantalks	nuclear
Irantalks	politics	Israel	IAEA	Tweetzup
No2Rouhani	bbc	news	UNGA	tcot
Iraq	london	Iraq	No2Rouhani	IranDeal
BREAKING	AP	PMOI	middleeast	Israel
humanrights	abc	EU	UNSC	ISIS
MaryamRajavi	Woman	syria	deathpenalty	Russia
NoNuclearIran	United.Nations	NoNuclearIran	Facts	PMOI
Commons	United_States_of_America	Barack_Obama	DirectAN	Parchin

Table 53. Hashtag January 2015 LDA Topics.

Topic A	Topic B	Topic C	Topic D	Topic F
Iran	Iran	Iran	Iran	Iran
ISIS	Nuclear	IranTalks	IranTalks	Opp4All
bbc	US	IranTalksVienna	IranTalksGeneva	JohnKerry
UN	nucleartalks	usa	MiddleEast	NEWS
Syria	Obama	sanctions	nonuclearIran	Tweetzup
Iraq	Israel	Paris	BREAKING	tcot
Reuters	HumanRights	Senate	IAEA	Israel
FOX	MaryamRajavi	world	StateDept	sotu
Mumbai	EU	UK	ZDF	StopTheClock
newsfeed	CoE	FNPolitics	TeamIranYouth	FoxNewsCHAT

Table 54. Hashtag February 2015 LDA Topics.

Topic A	Topic B	Topic C	Topic D	Topic F
Iran	Irantalks	Irantalks	Irantalks	Iran
Euronews	Iran	Iran	Iran	Woman
usa	Syria	NUCLEAR	Iranians	Phoenix
tcot	PMOI	breakingnews	ISIS	Arizona
cnn	paris	Iranian	IranTalksVienna	UN
Reuters	Obama	BREAKING	ZDF	politics
Israel	Iraq	IAEA	IranNuclearTalks	FOX
AP	News	Bahrain	IranTalksGeneva	health
News	UK	Ø.Ù.Ø.Ø.Ø.ÙŠÙ.	Facts	egypt
world	US	NCRI	events	humanrights

Table 55. Hashtag March 2015 LDA Topics.

Topic A	Topic B	Topic C	Topic D	Topic F
World	Iran	Irantalks	Iran	Irantalks
Iraq	nuclear	Iran	nuclear	Iran
Politics	Israel	nuclear	news	IAEA
LONDON	US	No2Rouhan i	uk	No2Rouhan i
google	Tweetzup	NonuclearIr an	Spain	NYC
Syria	MaryamRaja vi	Rouhani	CampLiberty	EU
LeMonde	PMOI	newyork	ISIS	UNSC
Euronews	tcot	NuclearDeal	Senat	timetoact
FOX	UNGA	pictures	United.Nations	Facts
United_States_of_Ame rica	USAHeadlin es	negotiation	United_States_of_Ame rica	STATEDE PT

Table 56. Hashtag April 2015 LDA Topics.

Topic A	Topic B	Topic C	Topic D	Topic F
nuclear	news	Iran	nuclear	Iran
news	Irantalks	Israel	Iran	World
Iran	IranDeal	NETANYAHU	Irantalks	politics
Iraq	US	UniteBlue	IranDeal	FOX
CNN	Iranians	WakeUpAmeric a	Congress	Syria

MiddleEast	FNPolitics	tcot	MaryamRajavi	Reuters
PMOI	NuclearTalks	pjnet	UK	LeMonde
ZDF	UK	NoNuclearIran	Tehran	egypt
sverige	NYC	nomoreGOPwar	Paris	Bahrain (in Arabic)
United.Nations	United_States_of_America	Barack_Obama	IranFreedom	Bahrain

Table 57. Parsed September 2014 LDA topics.

Topic A	Topic B	Topic C	Topic D	Topic F
Irantalks	Iran	Iran	Iran	Iran
Iran	nuclear	nuclear	nuclear	World
IAEA	Israel	news	United.Nations	nuclear
nuclear	Netanyahu	ISIS	tehran	Iraq
No2Rouhani	bad	Zarif	threat	news
MaryamRajavi	nuclear_weapons	Kerry	Islamic	Politics
Rouhani	Barack_Obama	deadline	Site	LONDON
NYC	regime	progress	Secret	Syria
EU	cease	United_States_of_America	uncertain	FOX
de	warn	Reuters_Ltd	amp	United_States_of_America

Table 58. Parsed October 2014 LDA Topics

Topic A	Topic B	Topic C	Topic D	Topic F
Iran	Iran	Iran	Iran	Iran
nuclear	nuclear	nuclear	nuclear	nuclear
Irantalks	news	Tehran	Netanyahu	Rouhani
IAEA	Iraq	Site	Israel	ISIS
No2Rouhani	Progress	deadline	leader	PMOI
humanrights	threat	secret	cease	regime
NoNuclearIran	bad	United_States_of_America	spirits_and_gods	Khamenei
Kerry	amp	Reuters_Ltd	nuclear_weapons	continue
Zarif	uncertain	concern	Warn	weapon
de	United.Nations	United.Nations	Barack_Obama	sign

Table 59. Parsed January 2015 LDA Topics.

Topic A	**Topic B**	**Topic C**	**Topic D**	**Topic F**
Iran	Iran	Iran	Iran	Iran
Nuclear	Kerry	Irantalks	Irantalks	Nuclear
Netanyahu	Irantalks	Nuclear	Nuclear	JohnKerry
Israel	Nuclear	Tehran	IAEA	NEWS
Iraq	nonuclearIran	Israel	sotu	Progress
Threat	Zarif	leader	regime	Weapon
amp	Deadline	secret	nuclear_weapons	warn
uncertain	EU	United_States_of_America	Reuters_Ltd	United_States_of_America
bad	CoE	site	cease	Butah
Barack_Obama	de	X_policy	United.Nations	seek

Table 60. Parsed Feburary 2015 LDA Topics.

Topic A	**Topic B**	**Topic C**	**Topic D**	**Topic F**
Irantalks	Iran	Iran	Iran	Iran
Iran	Euronews	NUCLEAR	NUCLEAR	Phoenix
PMOI	NUCLEAR	tcot	Israel	Arizona
NUCLEAR	cnn	Zarif	netanyahu	NUCLEAR
Iraq	AP	Kerry	bad	health
News	NCRI	IAEA	Threat	world
UK	secret	deadline	uncertain	Rouhani
US	site	Progress	Barack_Obama	sign
IranTalksGeneva	United_States_of_America	Reuters_Ltd	nuclear_weapons	reach
United.Nations	cease	spirits_and_gods	amp	Senator

Table 61. Parsed March 2015 LDA Topics.

Topic A	**Topic B**	**Topic C**	**Topic D**	**Topic F**
Iran	Israel	Iran	Iran	tcot
nuclear	nuclear	nuclear	nuclear	Israel
Irantalks	Netanyahu	Bahrain	Tehran	Iran
US	Kerry	threat	FM	nuclear
news	Zarif	weapon	Reveal	WakeUpAmerica

IAEA	Progress	Leader	SECRET	Senate
Deadline	bad	nuclear_weapons	site	de
regime	United_States_of_America	amp	extend	sign
Barack_Obama	spirits_and_gods	uncertain	Newsmedia	Los_Angeles_Ca_Usa
United.Nations	Reuters_Ltd	cease	X_policy	El

Table 62. Parsed April 2015 LDA Topics.

Topic A	Topic B	Topic C	Topic D	Topic F
nuclear	nuclear	nuclear	nuclear	nuclear
Iran	Iran	Tehran	news	KERRY
World	Israel	threat	Iran	Zarif
politics	NETANYAHU	nuclear_weapons	Irantalks	Progress
CNN	tcot	site	IranDeal	Bahrain
Rouhani	Israeli	Regime	US	deadline
Khamenei	Senate	secret	UK	Butah
De	Barack_Obama	cease	BAD	Reuters_Ltd
El	warn	amp	Leader	spirits_and_gods
Los_Angeles_Ca_Usa	Sign	uncertain	United_States_of_America	United.Nations

Table 63. Expert March 2015 LDA Topics.

Topic A	Topic B	Topic C	Topic D	Topic F
nukefest2015	Irantalks	Irantalks	Russia	Irantalks
nuclear	Iran	Iran	Iran	nukefest2015
security	nukefest2015	IranDeal	nukefest 2015	NPT
amp	nuclear	nuclear	nuclear	Korea
Today_Online	amp	security	amp	amp
Hawaii_Usa	Today_Online	amp	read	good
Butah	good	good	Butah	nuclear_weapons
North	nuclear_weapons	nuclear_weapons	launch	North
Learn	United_States_of_America	listen	weapon	United_States_of_America

United_States_of_America	thank_you	United_States_of_America	policy	weapon

Table 64. Expert April 2015 LDA Topics.

Topic A	Topic B	Topic C	Topic D	Topic F
NPTRevCon	IranDeal	IranDeal	IranDeal	nuclear
NPT	nuclear	nuclear	nuclear	NPT2015
Iran	NPT2015	Iran	NPT2015	NPT
Irantalks	NPTRevCon	Russia	NPTRevCon	Iran
United_States_of_America	Iran	United_States_of_America	Korea	Security
good	Irantalks	good	United_States_of_America	effort
amp	policy	amp	good	nuclear_weapons
nuclear_weapons	good	Butah	amp	Today_Online
Butah	amp	launch	North	X.s
X.s	Butah	weapon	Today_Online	Read

Table 65. Expert April 1-14 2016 LDA Topics.

Topic A	Topic B	Topic C	Topic D	Topic F
DPRK	DPRK	Korea	NSS2016	DPRK
InnovationForum	nuclear	nuclear	Iran	InnovationForum
NSS2016	Russia	Russia	United_States_of_America	NSS2016
MITnukes	Iran	nuclear_weapons	nuclear_weapons	MITnukes
United_States_of_America	global	Butah	good	United_States_of_America
good	good	amp	Butah	nuclear_weapons
Butah	Newsmedia	North	Today_Online	good
Read	Today_Online	weapon	amp	Butah
Hawaii_Usa	amp	Launch	weapon	Today_Online
X.s	security	security	X.s	foreign

Appendix: Sentiment Maps

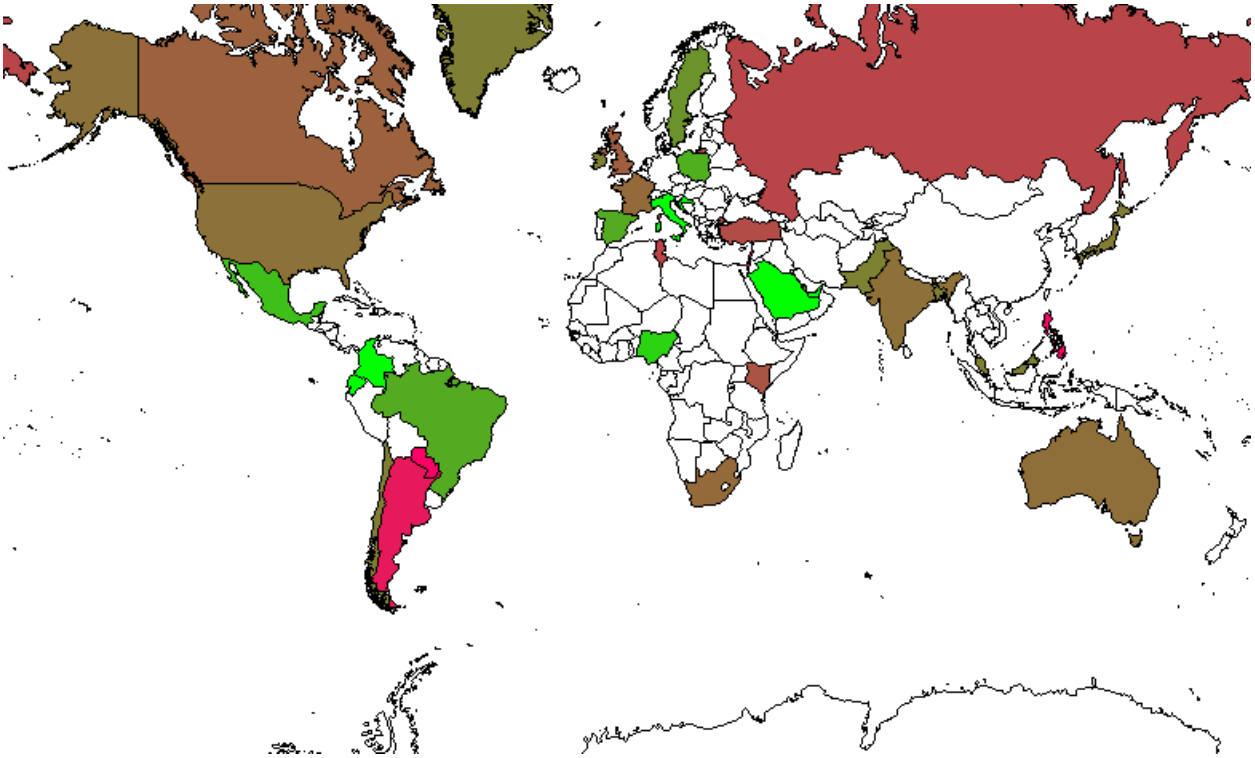


Figure 30. Sentiment map for entire nuclear Twitter dataset in September 2014. Green indicates positive average sentiment for the country; red indicates negative average sentiment..

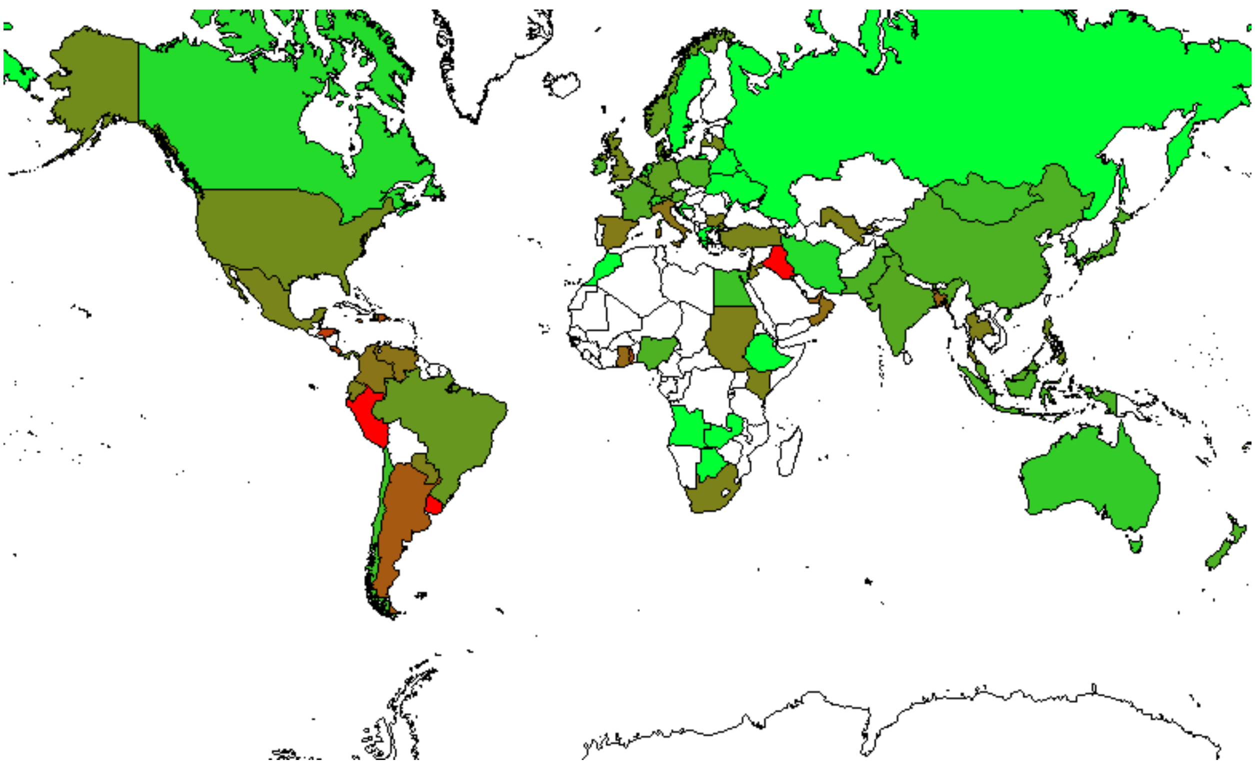


Figure 31. Sentiment map for entire nuclear Twitter dataset in April 2015. Green indicates positive average sentiment; red indicates negative average sentiment.

Chapter 6: Conclusions

In this thesis, I demonstrate the utility of computational modeling and approaches to the field of nuclear nonproliferation studies. This is shown both through simulation as well as large scale data analysis of social media. These contributions focus on new methodologies as well as data that can be utilized by this field to develop policy in this field.

The adapted Friedkin model extends the Friedkin approach to understanding and modeling social influence utilizing political networks as inputs. It also extends the Friedkin model by considering the distinct influence of two different types of ties between political actors. In the historical analysis utilizing this approach, there is evidence that the Nonproliferation Treaty (NPT) had an effect on international relations by reducing motivations to develop a nuclear weapon to security concerns – taking away cultural prestige and economic motivations to develop a nuclear weapon. However, in analyzing current political networks, this security driven model does not sufficiently explain motivations to develop nuclear capability.

In analyzing current political networks, comparing Iran's current security context to Iran's security context in 2005 shows that Iran is in a slightly worse current political situation. In analyzing the political networks, Syria appears at a unique crossroads in having a high motivation to develop nuclear weapons due to its alliances and hostilities. A broader analysis of different security model coefficients shows that conventional weapons will play a more significant role in the region than nuclear capability.

To analyze online discussion of nuclear topics on Twitter, we first find that discussion related to nuclear topics is frequently embedded in conversations – replies to links and central actors. In order to perform sentiment analysis, we find that most users do in fact change their sentiment rating of social media messages when viewed in context. In a majority of cases, this leads to an attenuation of sentiment. In a policy context, this implies that more attention should be given to the structure of a network's discussion: if many individual people are posting about a link or a topic, the sentiment measured on those posts is going to be more representative than a simple repost of a media or organization post.

In assessing tweets related to the Iranian nuclear agreement, we find there is a major distinction between online discussion obtained by utilizing Twitter's keyword search and focusing on the Twitter accounts of nonproliferation experts. There is some evidence of changes in global sentiment at a country level, but the relevant data is very small.

Each of these contributions has their own drawbacks. The Friedkin models do not account for bureaucratic obfuscation, or for the presence of sub-national and regional movements that contribute to each country's nuclear weapons development decision. They do not account for economic approaches of analysis that contribute to decisions, or take into account nuclear latency issues. In analyzing social media data, a major limitation is that this only assesses the Twitter platform; other platforms and multimedia, which provide a clarifying context, aren't assessed. In the Iranian twitter data, incomplete coverage, combined with significant English-language bias and self-promotion makes it difficult to fully integrate the analysis into a coherent policy brief.

However, each of these contributions in computational modeling help the modern analyst. The Friedkin model can be clearly extended to account for sub-national and regional movements. In intelligence, analysts traditionally provided the link between intelligence collectors and policymakers. But in today's counterterror world, that role is being questioned. Intelligence is now being asked to address broader scenarios, and help policymakers understand complex problems involving a range of agents without any clear pattern. This 'sensemaking' approach is vital to understanding new sources of intelligence for future analysts – an approach that is clearly demonstrated in this thesis.

References

[1] R. Jervis, “Deterrence Theory Revisited,” *World Politics*, vol. 31, no. 2, pp. 1–37, Jan. 1979.

[2] F. M. Kaplan, *The Wizards of Armageddon*. Stanford University Press, 1983.

[3] M. Trachtenberg, *History and Strategy*. Princeton University Press, 1991.

[4] E. Adler, “The Emergence of Cooperation: National Epistemic Communities and the International Evolution of the Idea of Nuclear Arms Control,” *International Organization*, vol. 46, no. 1, pp. 101–145, 1992.

[5] P. K. Huth, “Deterrence and international conflict: Empirical findings and theoretical debates,” *Annu. Rev. Polit. Sci.*, vol. 2, no. 1, pp. 25–48, 1999.

[6] D. D. Smith, *Deterring America*. Cambridge University Press, 2006.

[7] J. H. Lebovic, *Deterring International Terrorism and Rogue States*. Routledge, 2007.

[8] P. M. Morgan, *Deterrence Now*. Cambridge University Press, 2003.

[9] J. Nolan and M. Strauss, “The Rogues' Gallery, Stealing the Fire,” *Brown Journal of World Affairs*, vol. 4, no. 1, pp. 21–38, 1997.

[10] R. Jervis, “The confrontation between Iraq and the US: Implications for the theory and practice of deterrence,” *European Journal of International Relations*, vol. 9, no. 2, pp. 315–337, Jun. 2003.

[11] R. S. Litwak, *Regime Change: US Strategy through the Prism of 9/11*. Washington: Woodrow Wilson Center Press, 2007.

[12] K. Waltz and S. Sagan, *The Spread of Nuclear Weapons*, 2nd ed., no. 2002. WW Norton & Company, 2002.

[13] S. Sagan, “Why do states build nuclear weapons?: three models in search of a bomb,” *International Security*, vol. 21, no. 3, pp. 54–86, 1996.

[14] X. Dai, *International Institutions and National Policies*. Cambridge University Press, 2007.

[15] E. Solingen, *Nuclear logics: contrasting paths in East Asia and the Middle East*. Princeton University Press, 2007.

[16] I. Abraham, “The Ambivalence of Nuclear Histories,” *Osiris*, vol. 21, no. 1, pp. 49–65, 2006.

[17] J. E. C. Hymans, *The Psychology of Nuclear Proliferation: Identity, Emotions and Foreign Policy*. Cambridge University Press, 2006.

[18] N. Tannenwald, “The Nuclear Taboo: The United States and the Normative Basis of Nuclear Non-Use,” *International Organization*, vol. 53, no. 3, pp. 433–468, Jan. 1999.

[19] M. R. Rublee, *Nonproliferation Norms*. University of Georgia Press, 2009.

[20] D. Jo and E. Gartzke, “Determinants of nuclear weapons proliferation,” *Journal of Conflict Resolution*, vol. 51, no. 1, pp. 167–194, 2007.

[21] S. Singh and C. Way, “The Correlates of Nuclear Proliferation: A Quantitative Test,” *Journal of Conflict Resolution*, vol. 48, no. 6, pp. 859–885, Dec. 2004.

[22] R. Rauchhaus, “Evaluating the Nuclear Peace Hypothesis: A Quantitative Approach,” *Journal of Conflict Resolution*, vol. 53, no. 2, pp. 258–277, Jan. 2009.

[23] M. S. Bell and N. L. Miller, “Questioning the Effect of Nuclear Weapons on Conflict,” *Journal of Conflict Resolution*, vol. 59, no. 1, pp. 74–92, Jan. 2015.

[24] A. Montgomery and S. Sagan, “The Perils of Predicting Proliferation,” in *Causes and Consequences of Nuclear Proliferation*, no. 14, R. Rauchhaus, M. Kroenig, and E.

Gartzke, Eds. New York: Routledge, 2011, pp. 294–329.
[25] M. S. Bell, “Examining Explanations for Nuclear Proliferation,” *International Studies Quarterly*, pp. sqv007–10, Dec. 2015.
[26] K. N. Waltz, *Theory of International Politics*. Waveland Press, 2010.
[27] R. K. Betts, “Paranoids, pygmies, pariahs and nonproliferation revisited,” in *The Proliferation Puzzle: Why nuclear weapons spread (and what results)*, Z. S. Davis and B. Frankel, Eds. Portland: Frank Cass, 1993.
[28] Z. S. Davis, “The Realist Nuclear Regime,” in *The Proliferation Puzzle: Why nuclear weapons spread (and what results)*, Z. S. Davis and B. Frankel, Eds. Routledge, 1993.
[29] B. A. Thayer, “The Causes of Nuclear Proliferation and the Utility of the Nuclear Non-Proliferation Regime,” *Security Studies*, 1995.
[30] S. Sagan, “Nuclear Latency and Nuclear Proliferation,” in *Forecasting Nuclear Proliferation in the 21st Century*, 1st ed., vol. 1, no. 5, W. C. potter and G. Mukhatzhanova, Eds. Stanford: Stanford University Press, 2010, pp. 58–79.
[31] S. Meyer, *The Dynamics of Nuclear Proliferation*. Chicago: University of Chicago Press, 1984.
[32] J. Stoll, “World production of latent nuclear capacity,” 1996.
[33] M. Kroenig, *Exporting the Bomb: Technology Transfer and the Spread of Nuclear Weapons*. Cornell Univ Pr, 2010.
[34] M. Fuhrmann, “Taking a Walk on the Supply Side The Determinants of Civilian Nuclear Cooperation,” *Journal of Conflict Resolution*, 2009.
[35] M. Fuhrmann and B. Tkach, “Almost nuclear: Introducing the Nuclear Latency dataset,” *Conflict Management and Peace Science*, vol. 32, no. 4, pp. 443–461, Sep. 2015.
[36] B. C. Smith and W. Spaniel, “Introducing ν-CLEAR: A Latent Variable Approach to Measuring Nuclear Capability,” presented at the Midwest Political Science Association, 2016, pp. 1–34.
[37] A. J. Wohlstetter, *Swords From Plowshares*. University Of Chicago Press, 1979.
[38] R. Harney, G. Brown, and M. Carlyle, “Anatomy of a Project to Produce a First Nuclear Weapon,” *Science and Global Security*, vol. 14, pp. 163–182, 2006.
[39] M. D. Talbert, G. Zentner, and R. Coles, “Nuclear Proliferation Technology Trends Analysis,” Office of Scientific and Technical Information, DOE, PNNL-14480, Oct. 2005.
[40] J. Hymans, “When Does a State Become a ‘Nuclear Weapon State’? An Exercise in Measurement Validation,” *Nonproliferation Review*, 2010.
[41] N. L. Miller, “Nuclear Dominoes: A Self-Defeating Prophecy?,” *Security Studies*, vol. 23, no. 1, pp. 33–73, Feb. 2014.
[42] G. Gerzhoy, “Alliance Coercion and Nuclear Restraint: How the United States Thwarted West Germany's Nuclear Ambitions,” *International Security*, vol. 39, no. 4, pp. 91–129, May 2015.
[43] O. Rabinowitz, “Keeping the Bombs in the Basement: U.S. Nonproliferation Policy toward Israel, South Africa, and Pakistan,” *International Security*, vol. 40, no. 1, pp. 47–86, Aug. 2015.
[44] F. J. Gavin, “Strategies of Inhibition: U.S. Grand Strategy, the Nuclear Revolution, and Nonproliferation,” *International Security*, vol. 40, no. 1, pp. 9–46, Aug. 2015.
[45] S. D. Sagan and K. N. Waltz, “Political Scientists and Historians in Search of the

Bomb," *Journal of Strategic Studies*, vol. 36, no. 1, pp. 143–151, Feb. 2013.
[46] F. J. Gavin, "Politics, History and the Ivory Tower-Policy Gap in the Nuclear Proliferation Debate," *Journal of Strategic Studies*, vol. 35, no. 4, pp. 573–600, Aug. 2012.
[47] D. J. Caswell, R. A. Howard, and M. E. Paté-Cornell, "Analysis of National Strategies to Counter a Country's Nuclear Weapons Program," *Decision Analysis*, vol. 8, no. 1, pp. 30–45, Mar. 2011.
[48] W. Frankenstein, G. Mezzour, K. M. Carley, and L. R. Carley, "Remote assessment of countries' nuclear, biological, and cyber capabilities: joint motivation and latent capability approach," *Soc. Netw. Anal. Min.*, vol. 5, no. 1, pp. 1–21, Jan. 2015.
[49] J. Arquilla and P. K. Davis, *Modeling Decisionmaking of Potential Proliferators as Part of Developing Counterproliferation Strategies*. RAND, 1994, pp. 1–60.
[50] E. Gartzke and M. Kroenig, "A strategic approach to nuclear proliferation," *Journal of Conflict Resolution*, 2009.
[51] N. E. Friedkin and E. C. Johnsen, "Social influence networks and opinion change," *Advances in Group Processes*, vol. 16, pp. 1–29, 1999.
[52] B. Bueno de Mesquita, "A Decision Making Model: Its structure and form," *International Interactions*, vol. 23, no. 3, pp. 235–266, Jan. 1997.
[53] World Bank, *Indicators*.
[54] D. M. Gibler, *International military alliances, 1648-2008*. CQ Press, 2009.
[55] J. Wilkenfeld, M. Brecher, J. Hewitt, K. Beardsley, and P. Eralp, *International Crisis Behavior Dataset*. 2010.
[56] U. E. Daxecker, "Rivalry, Instability, and the Probability of International Conflict," *Conflict Management and Peace Science*, vol. 28, no. 5, pp. 543–565, Nov. 2011.
[57] Z. Maoz, L. G. Terris, R. D. Kuperman, and I. Talmud, "What is the enemy of my enemy? Causes and consequences of imbalanced international relations, 1816–2001," *Journal of Politics*, vol. 69, no. 1, pp. 100–115, 2007.
[58] W. Q. Bowen, "Libya and Nuclear Proliferation: Stepping Back from the Brink," 2006.
[59] J. Walsh, "Surprise down under: The secret history of Australia's nuclear ambitions," *The Nonproliferation Review*, 1997.
[60] M. R. Rublee, *Nonproliferation Norms: Why States Choose Nuclear Restraint*.
[61] D. Kilcullen, "The deadly efficiency of Isis and how it grew on the global stage," *The Guardian*, 11-Mar-2016.
[62] "Mercurial Putin keeps the west guessing in Syria - FT.com," Mar. 2016.
[63] J. Stavridis, "It's All About the Base," *Foreign Policy*, 29-Jul-2015.
[64] S. Almosawa, "Yemen: Houthi Rebels Enter Peace Talks With Saudi Arabia," *New York Times*, 08-Mar-2016.
[65] N. V. Osipova, "Russian Withdrawal from Syria: Takeaways," *New York Times*.
[66] J. Cirincione and P. Walker, "Is ISIS Using Chemical Weapons?," *Defense One*, Oct-2014. [Online]. Available: http://scholar.google.com/scholar?q=related:2TKVQez_1RoJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5. [Accessed: 16-Mar-2016].
[67] A. Acharya, "ISIS' Chemical Weapons," Mar. 2016.
[68] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis," *FNT in Information Retrieval*, vol. 2, no. 1, pp. 1–135, 2008.
[69] J. Grimmer and B. M. Stewart, "Text as Data: The Promise and Pitfalls of Automatic

Content Analysis Methods for Political Texts," *Political Analysis*, vol. 21, no. 3, pp. 267–297, Jul. 2013.
[70] M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas, "Sentiment strength detection in short informal text," *Journal of the American Society for Information Science and Technology*, vol. 61, no. 12, pp. 2544–2558, Dec. 2010.
[71] A. Esuli and F. Sebastiani, "Sentiwordnet: A publicly available lexical resource for opinion mining," presented at the Proceedings of LREC, 2006.
[72] J. W. Pennebaker, C. K. Chung, and M. Ireland, "The development and psychometric properties of LIWC2007," LIWC.net, Austin, TX, USA, LIWC2007, 2007.
[73] P. J. Stone, *User's Manual for The General Inquirer*. MIT Press (MA), 1968.
[74] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, "Sentiment analysis of twitter data," pp. 30–38, 2011.
[75] D. Davidov, O. Tsur, and A. Rappoport, "Enhanced sentiment learning using twitter hashtags and smileys," *Proceedings of the rd international conference on computational linguistics posters*, pp. 241–249, 2010.
[76] A. Bermingham, M. Conway, L. McInerney, N. O'Hare, and A. F. Smeaton, *Combining Social Network Analysis and Sentiment Analysis to Explore the Potential for Online Radicalisation*. IEEE, 2009, pp. 231–236.
[77] M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment strength detection for the social web," *Journal of the American Society for Information Science and Technology*, vol. 63, no. 1, pp. 163–173, Jan. 2012.
[78] L. S. Lovin, "Affect control theory: An assessment*," *Journal of Mathematical Sociology*, vol. 13, no. 1, pp. 171–192, Jan. 1987.
[79] D. R. Heise, "Affect control theory: Concepts and model," *Journal of Mathematical Sociology*, vol. 13, no. 1, pp. 1–33, Jan. 1987.
[80] C. E. Osgood, W. H. May, and M. S. Miron, *Cross-cultural Universals of Affective Meaning*. University of Illinois Press, 1975.
[81] L. Smith-Lovin and D. T. Robinson, "Interpreting and Responding to Events in Arabic Culture," Office of Naval Research, Grant N00014-09-1-0556, Aug. 2015.
[82] W. Frankenstein, K. Joseph, and K. M. Carley, "Social Media ACTion: SOLO Data Description," CMU-ISR-16-103, Feb. 2016.
[83] C. J. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," presented at the … AAAI Conference on Weblogs and Social Media, 2014.
[84] J. Davis, "Paul Wolfowitz on Intelligence Policy-Relations," 1996.
[85] J. A. Barry, J. Davis, D. D. Gries, and J. Sullivan, "Bridging the Intelligence-Policy Divide," *Studies in Intelligence*, 1994.
[86] G. F. Treverton, *Intelligence for an Age of Terror*. Cambridge University Press, 2011.
[87] C. Grier, K. Thomas, V. Paxson, and M. Zhang, "@ spam: the underground on 140 characters or less," pp. 27–37, 2010.
[88] K. Thomas, C. Grier, and V. Paxson, "Adapting social spam infrastructure for political censorship," pp. 13–13, 2012.
[89] Z. Miller, B. Dickinson, W. Deitrick, W. Hu, and A. H. Wang, "Twitter spammer detection using data stream clustering," *Information Sciences*, vol. 260, no. 0, pp. 64–73, 2014.
[90] A. A. Amleshwaram, N. Reddy, S. Yadav, G. Gu, and C. Yang, "CATS:

Characterizing automation of Twitter spammers," pp. 1–10, 2013.
[91] I. Santos, I. Miñambres-Marcos, C. Laorden, P. Galán-García, A. Santamaría-Ibirika, and P. G. Bringas, "Twitter Content-Based Spam Filtering," in *Advances in Intelligent Systems and Computing*, vol. 239, no. 46, Cham: Springer International Publishing, 2014, pp. 449–458.
[92] K. M. Carley, J. Pfeffer, J. Reminga, J. Storrick, and D. Columbus, "ORA User's Guide 2013," Jun. 2013.
[93] D. Ediger, K. Jiang, J. Riedy, D. A. Bader, and C. Corley, "Massive Social Network Analysis: Mining Twitter for Social Good," presented at the 2010 39th International Conference on Parallel Processing (ICPP), 2010, pp. 583–593.
[94] D. E. A. D. A. B. K Jiang, "Generalizing k- Betweenness centrality using short paths and a parallel multithreaded implementation," presented at the The 38th International Conference on Parallel Processing (ICPP 2009), 2009.
[95] W. Wei, K. Joseph, W. Lo, and K. M. Carley, "A Bayesian Graphical Model to Discover Latent Events from Twitter," *... International AAAI Conference on ...*, 2015.
[96] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *The Journal of Machine Learning Research*, vol. 3, pp. 993–1022, Mar. 2003.
[97] J. E. C. Hymans, "The Study of Nuclear Proliferation and Nonproliferation: Toward a New Consensus?," in *Forecasting Nuclear Proliferation in the 21st Century*, 1st ed., vol. 1, no. 2, W. C. potter and G. Mukhatzhanova, Eds. Stanford, CA: Stanford University Press, 2010, pp. 13–17.
[98] S. Sagan, "The Causes of Nuclear Weapons Proliferation," *Annual Review of Political Science*, vol. 14, no. June, pp. 225–244, 2011.
[99] M. M. Benham-Hutchins, B. B. Brewer, K. M. Carley, M. Kowalchuk, and J. A. Effken, "Development of a Social Network Analysis Data Collection Application," *Under Review*, Jan. 2016.
[100] G. P. Morgan and J. H. Morgan, "Surveyor 3.0: A Note on an Open Source Application for Sentiment Analysis," *Sociological Methods & Research*, pp. 1–12, Oct. 2014.
[101] "Import and Export Surveys - Qualtrics."

Appendix A: SOLO Study Data Technical Report

Abstract

This technical report summarizes the demographics and Socially Observed Linked Opinion (SOLO) dataset, which came out of the Social Media ACTion study that took place at Carnegie Mellon during the summer of 2015. 124 individuals rated 4,320 social media posts and 1,680 news clips along the three dimensions used in Affect Control Theory. The report includes a description of the data, the training materials provided, and the consent form used for this study.

Introduction

This report describes the SOLO dataset, which is available to all researchers at Carnegie Mellon University. The report has the following sections: study background, study setup and interface, demographics, and data format description. This study was completed under IRB code HS14-670.

Study background

"State-of-the-art" research in sentiment analysis has three problems: the approaches were developed to analyze large bodies of text, it ignores the social context of social media, and it does not consider social media's international dimension. Social media text can be extremely short – making traditional machine learning approaches difficult, as the data to be classified has features not included in the training set. It is inherently social – frequently responding to individuals or events.

Most approaches focus exclusively on content[71]-[73]. For example, a user posting she is ill will receive positive, supportive posts on social media. The illness would be misclassified as a positive event due to the positive words in their responses. Finally, posts contain international content – cultures affect how individuals respond to events. Affect control theory formalizes the way that individuals respond to events by classifying evaluation, potency, and action, allowing for cross-cultural comparisons of events [78], [79].

To address these problems, the Social Media ACTion study had three primary goals:

1. To examine the role of context in evaluating valence of social media posts
2. To expand the lexicons available for Affect Control Theory
3. To develop a gold standard sentiment dataset of hand-labeled social media and news posts

To achieve the first goal, participants were asked to evaluate a set of Twitter posts twice: once, seeing a Twitter response post before seeing the original post, and the second time, seeing the response post directly beneath the original post. The second goal is product of analysis done on this dataset. The third goal is the primary focus of this technical report.

Study Participants

Individuals were recruited for 45 minute sessions to evaluate 90 social media posts and received $8 compensation in the form of an Amazon Gift Card. Individuals were recruited from both the CMU Center for Behavioral and Decision Research (CBDR) as well as from flyers posted around the Oakland neighborhood of Pittsburgh. To qualify for the study participants had to be over 18 years of age and native English speakers to ensure that participants understood all social media posts. Individuals who did not finish the study were compensated at the rate of 6 cents per social media post. More information about participant demographics is available in a later section of this technical report.

Study interface

Initially, participants were asked to attend in-person sessions and input answers using an internal CASOS server using a modified version of a survey designed for collecting medical informatics [99]. To facilitate collection over the course of the summer, however, we switched to using Qualtrics after 6 individuals had taken our study. We have updated the data collected from these participants so that all data is comparable, regardless of which platform the data was collected from. In particular, following best practice in Affect Control Theory coding there are three features of the interface that we manipulated to reduce framing and anchoring heuristics: double labeled axes, changing the lateral direction of intensity for "Activity" evaluations, and having an individual axis on each page seen by the participant [100].

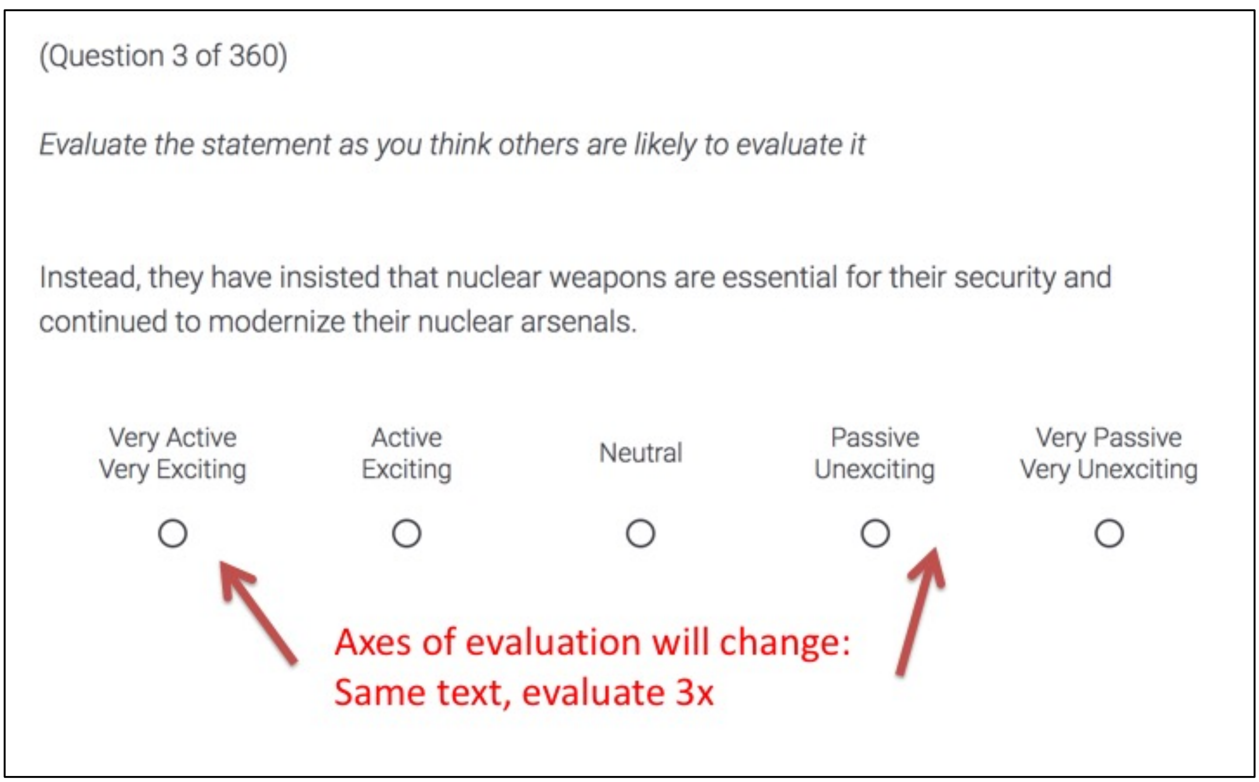


Figure 32. Sample evaluation screenshot taken from training slides.

In particular, note that:

1. We asked participants to evaluate the statement from a "general" perspective – which increases overall inter-rater agreement rates [83]
2. Participants rated statements on a 5-point Likert scale
3. Participants rated the same statement along Evaluation, Potency, and Activity scales immediately, on separate screens
4. Axes were given two reference points:
    a. Evaluation: Negative/Unpleasant to Positive/Pleasant
    b. Potency: Weak/Powerless to Strong/Powerful
    c. Activity: Active/Exciting to Passive/Unexciting
5. Activity was evaluated with "Active" on the left hand side and "Passive" on the right hand side to emphasize its distinction from potency

Participants each underwent a five minute training session to familiarize themselves with the ACT concepts of "Evaluation", "Potency", and "Activity". These slides, as well as the accompanying script, are available as an appendix to this technical report.

Questionnaire Structure

Participants who encoded Twitter posts rated posts in the following order (keep in mind ACT required participants to evaluate each post three times):

I. 30 "standalone" Twitter posts (90 evaluations)
II. 30 "conversation" Twitter posts (180 evaluations)
    a. Response seen first
    b. Original post seen second
III. 30 Response Twitter posts (90 evaluations)

The posts in section III were identical to the posts in section IIA; however, while in Section II they were presented in isolation, in section III they were presented together with the original post they responded to.

Participants who encoded news clips simply rated 120 sentences or sentence pairs.

Data Source

To ensure a diverse set of evaluations, we utilized four distinct topics across two platforms – Twitter and news articles pulled from Lexis Nexis. We ensured to the best of our ability that all tweets and news articles evaluated were in English; if non-English words were utilized, participants were instructed to mark the message as “Neutral”.

For “general” topics – we utilized the “Gardenhose” Twitter dataset. It is part of a larger dataset available at CMU that is composed of 10% of the total Twitter firehose. We selected random tweets from this dataset to represent commonly used English on Twitter. For news articles, we utilized sentences from the 2014 New York Times set of editorials.

Table 66. Table of different topics used

	Nuclear	Arab Spring	General	Typhoon Haiyan
Dates Covered	Sep 2014 – Oct 2014	Oct 2009-Nov 2013	Sep 2013 – Aug 2014	November – December 2013
Sample Keywords	Nuclear proliferation, heavy water, uranium	Tahrir Square, Arab Spring	n/a for Gardenhose; New York Times editorials	Typhoon Haiyan, Typhoon Yolanda
Number of Twitter Posts	1,080	1,080	1,080	1,080
Number of News Clips	420	420	420	420

Demographics

The primary constraints on participants involved being able to attend an in-person coding session in Pittsburgh and being over 18 years old. We expected to have a body of participants that matched closely with the undergraduate population at Carnegie Mellon; while this was largely the case, we also had participants from the local Pittsburgh community. We had a total of 124 participants. All demographics questions were asked at the end of the study and were completely voluntary.

Table 67. Gender of participants

Female	77 (62%)
Male	47 (38%)
Other / Decline to state	0 (0%)

Table 68. Age distribution of participants

Under 25	73 (59%)
25-30	25 (20%)
31-40	10 (8%)
41-50	5 (4%)
Over 50	11 (9%)

While we screened for native English speakers, we asked participants to rate their own English ability. 4 individuals (3%) self-identified as speaking English "well" as opposed to 120 individuals (97%) identifying their English proficiency as native speakers. We also asked individuals to identify other languages spoken at home. 88 (71%) of participants said that they only spoke English at home.

Table 69. Count of other languages spoken at home by participants

No other languages spoken at home	88
American Sign Language	1
Chinese (Mandarin)	4
Czech	1
Spanish	6
French	3
Guajarati	2
Hindi	6
Tamil	5
Marathi	4
Telugu	1
Kannada	1
Taiwanese	1
Punjabi	1
Russian	2
Urdu	2
Vietnamese	1

We asked participants to identify their race and ethnicity. Participants were able to select more than one category of race. 4 individuals (3%) self-identified as being Hispanic, Latino, or Spanish origin.

Table 70. Participant ethnic and racial distribution. Individuals could select multiple categories.

American Indian or Alaska Native	0 (0%)
Asian or Pacific Islander	43 (35%)
Black or African American	12 (10%)
White	72 (58%)

Other	4 (3%)

Data Format

The data is available on the CASOS Megadon server, which can be accessed at megadon.casos.cs.cmu.edu. The data is located on the D:// drive under "Public SOLO Data". An additional folder containing the questionnaires uploaded to Qualtrics, for researchers interested in replicating the study, is available upon request.

Eval_Tweets and Eval_News contain Evaluative ratings of tweets and news clips respectively; Power_Tweets and Power_News contain Power ratings, and Active_Tweets and Active_News contain Activity tweets.

Tweet row names have the format: "X[[NUMBER1]]_[R/S]_[[NUMBER2]].[[NUMBER3]].

News row names have the format: "[[LETTER1]]_X[[NUMBER1]]_S_[[NUMBER2]].

[NUMBER1] refers to the Tweet ID - this is located in either ArabSpring, Garden, Haiyan, or NukeTweets.tsv. [LETTER1] identifies the news topic – "A" for Arab Spring, "G" for General, "N" for nuclear, and "T" for Typhoon.

S indicates that the tweet was evaluated in isolation.

R indicates that the tweet was evaluated in context. You can view the mapping of what this tweet responded to in the XX_Pairs.tsv text file.

[NUMBER2] refers to the EPA rating, where 0 = evaluative, 1 = power, 2 = activity.

[NUMBER3] occasionally some tweets were selected twice. If this is the case, there will be a .1 or .2 after the main string identifying the tweet.

For Evaluation, the 5 point Likert goes from Most negative = 1 to most positive = 5.

For Power, the 5 point Likert goes from Weakest = 1 to Strongest = 5.

For Activity, the 5 point Likert goes from Most Active = 1 to Most Passive = 5.

Lessons Learned

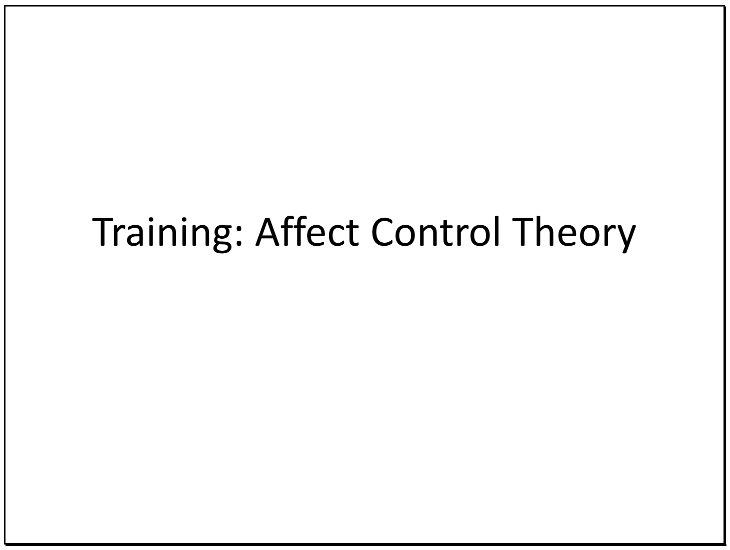
Over the course of conducting this study, several lessons were learned. I hope this document can serve to improve future studies.

For Carnegie Mellon researchers – I highly recommend advertising the study on the Center for Behavioral and Decision Making Research (CBDR) website (http://cbdr.cmu.edu). Studies performed online (such as those which would normally be done through Amazon Mechanical Turk) can also be posted to that website. I had significantly more success recruiting participants through CBDR than through flyers.

Carnegie Mellon also maintains a subscription to Qualtrics. This is particularly useful as Qualtrics allows for individuals to create relatively customized questionnaires very easily, as outlined in their technical support [101]. What isn't mentioned in their reference notes is that you can utilize basic HTML formatting – which significantly improves the presentation of the questionnaire.

Appendix: Training Slides

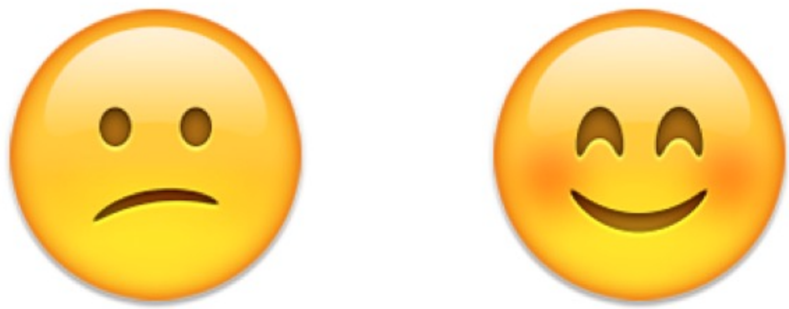
Slide 1



Slide 2

Study Goal: Expand Concepts of Sentiment

- Goal of study is to develop understanding of sentiment based on different feelings through a framework called Affect Control Theory



Slide 3

Why are we trying to do this?

- Most tools currently developed only look at positive and negative sentiment:

- We know, however, there are lots of other aspects of sentiment, as well as a social theory behind sentiment

Slide 4

Affect Control Theory (ACT)

- We perceive social identities through dimensions of sentiment
- Social events change sentiment and evoke emotion within us
- Structuring sentiment along three specific dimensions allows for cross-cultural comparisons of emotion

Heise, "Social action as the control of affect". *Systems Research and Behavioral Science* 1977 (22).
Heise, *Expressive Order: Confirming Sentiments in Social Actions*. (2007).

Slide 5

ACT Captures Multiple Dimensions of Sentiment

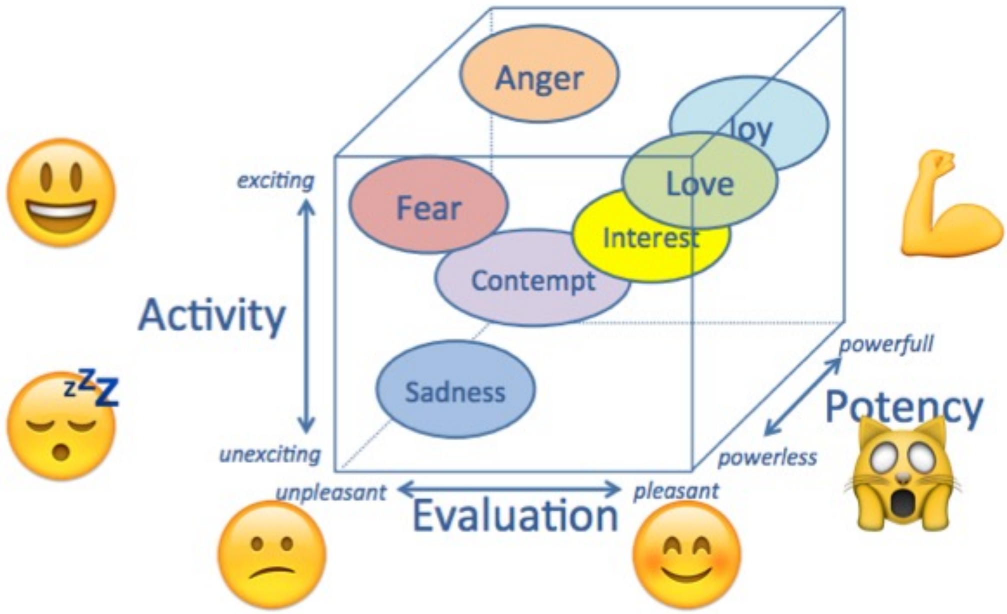


Illustration from Tobias Schröder's ACT lectures:
https://cs.uwaterloo.ca/~jhoey/research/bayesact/lectures/lecture1/lecture1Slides.pdf

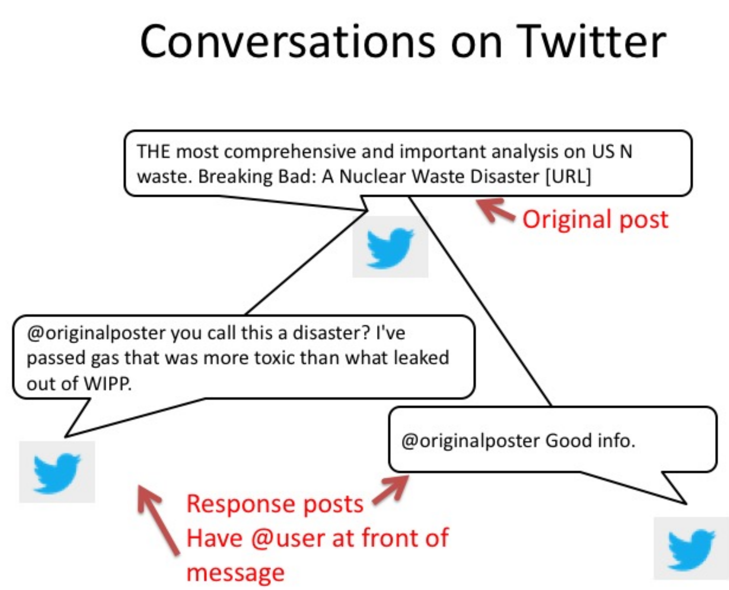
Slide 6

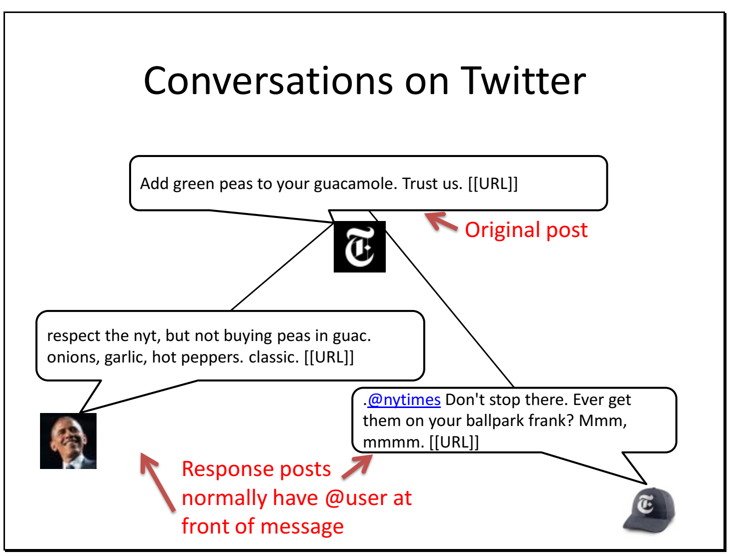
Dimensions of Sentiment (cont.)

- ACT's three dimensions of sentiment are:
  - **Evaluation** – how "good" or "bad", "positive" or "negative", "pleasant" or "unpleasant" something is
  - **Potency** – how "powerful" or "powerless" something is, the degree of status something or someone exhibits
  - **Activity** – how "active" or "passive" something is, the level to which it provokes excitement

Slide 7 – Two versions provided. Initially used "Breaking Bad" example to illustrate how Twitter users had conversations on the platform, switched to President Obama's "peas in guacamole" comment after it was made during the summer of 2015.

Conversations example 1:



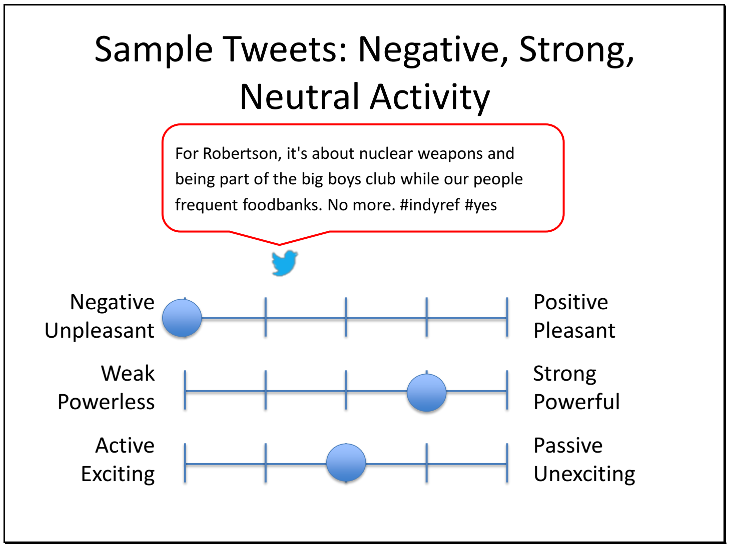


Slide 8

Twitter Peculiars

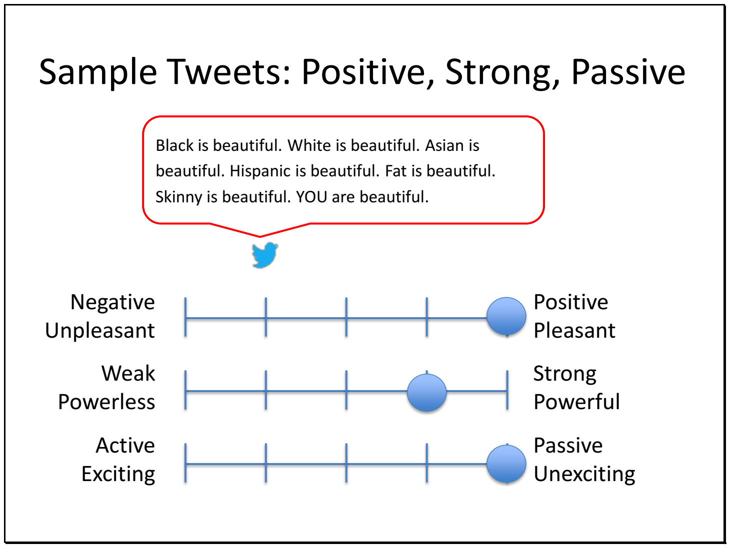
- Retweets: how information is spread across twitter
  - Sometimes prefaced with RT, or quotes around the tweet
  - E.g. “@potus: respect the nyt, but not buying peas in guac. onions, garlic, hot peppers. classic. [[URL]]”
  - MT: ‘Modified Tweet’, effectively the same
- Hashtags: #whatsupwiththat #topic #trending
  - Use # to indicate topics
  - Some tweets have multiple hashtags
- Responding to others: @user1
  - Sometimes posts have @user at beginning of tweet
  - Messages with @user later in message used to alert others

Slide 9



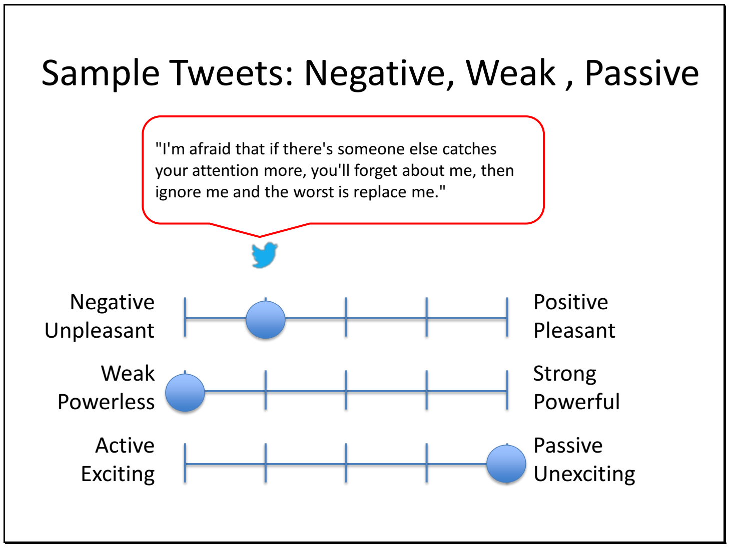
Message clearly negative on Robertson. Robertson a powerful individual, "part of the big boys club"; neither active nor passive as it's unclear what action he is taking.

Slide 10



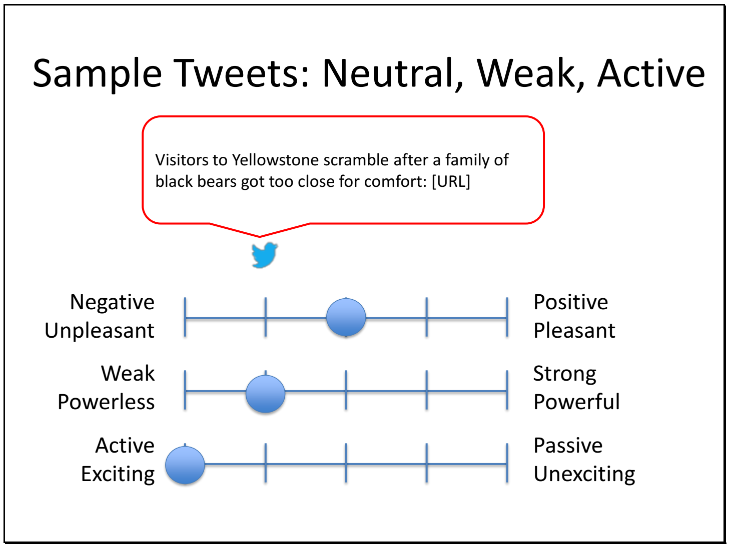
All positive statements. All empowering statements – however, also all passive. Unclear what action is being taken here.

Slide 11



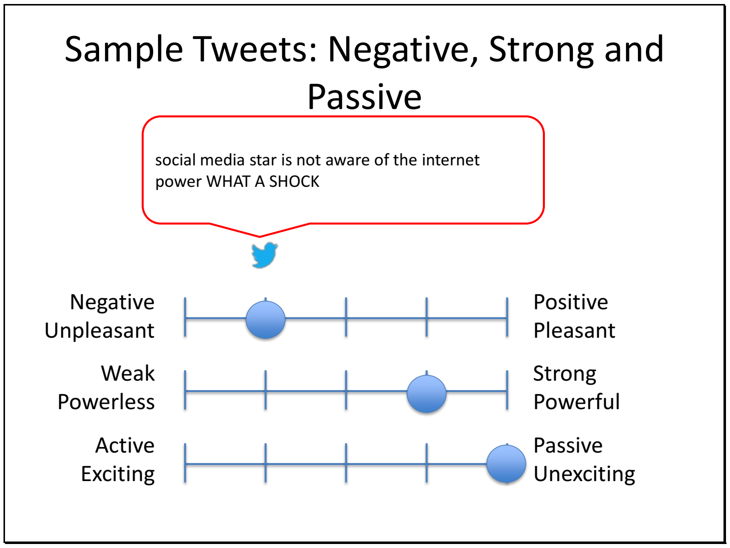
Message of a 14-year-old angsty teenager that you want to reach out and hug – clearly, a negative message, they feel weak, and they feel inactive.

Slide 12



Here's an example of a neutral tweet – neither positive or negative; it's just reporting the news. However, it's weak as the people in the sentence had to run away – which is itself an active process.

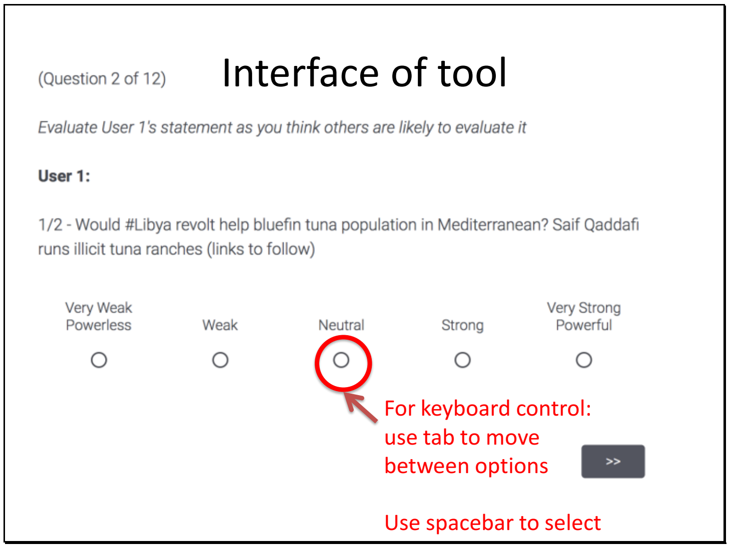
Slide 13



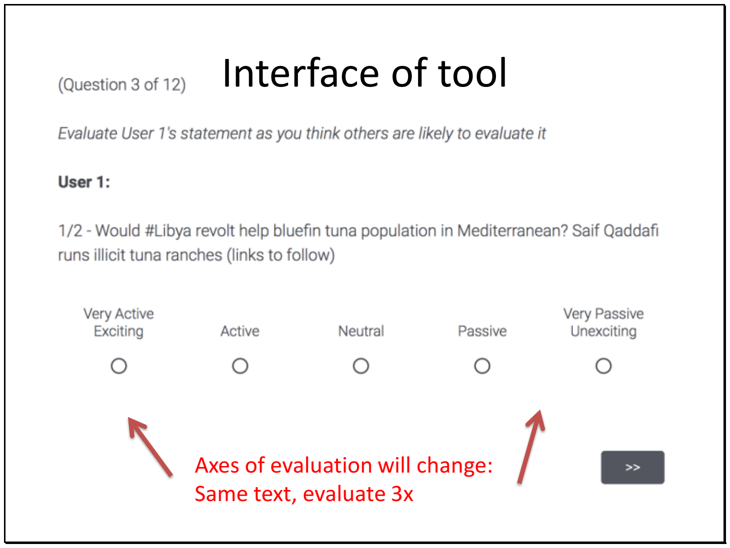
Finally, a sarcastic tweet. A bit negative, mocking a powerful “social media star” – someone with status. But also passive; the message mocks the “star” for their inaction.

Twitter evaluators shown the following three slides, which highlight "User 1" and "User 2"

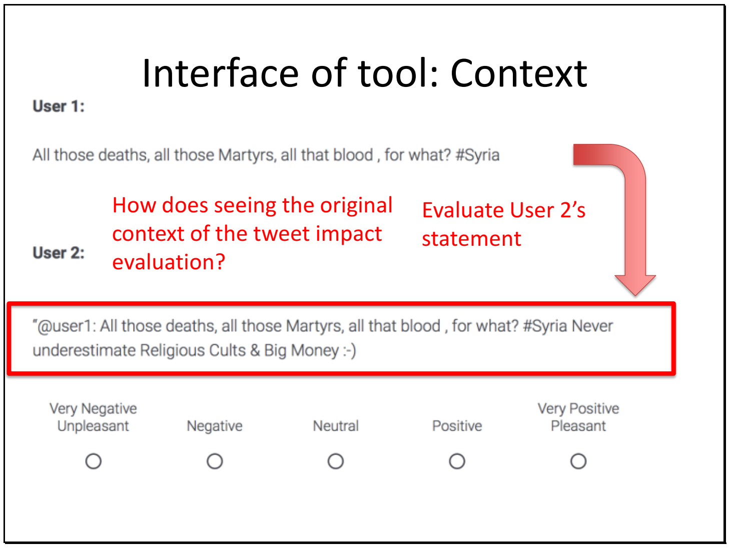
Slide 14



Slide 15

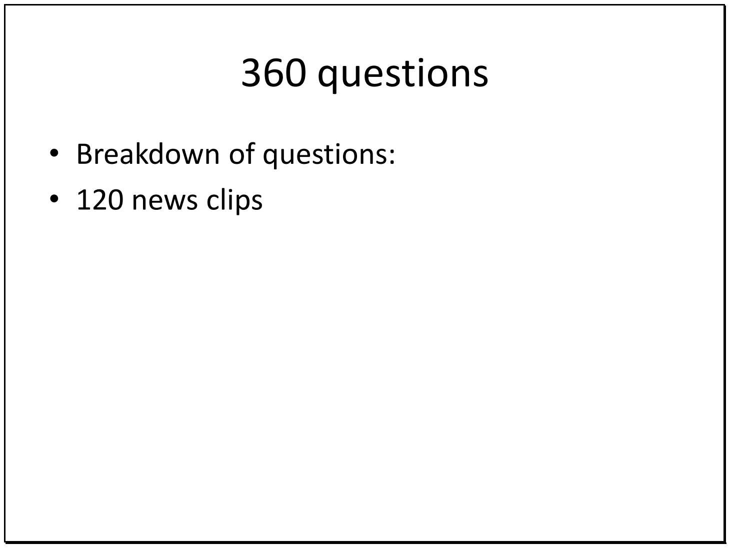

(Question 3 of 12)
Interface of tool
Evaluate User 1's statement as you think others are likely to evaluate it
User 1:
1/2 - Would #Libya revolt help bluefin tuna population in Mediterranean? Saif Qaddafi runs illicit tuna ranches (links to follow)
Very Active Exciting
Active
Neutral
Passive
Very Passive Unexciting
Axes of evaluation will change:
Same text, evaluate 3x
>>

Slide 16

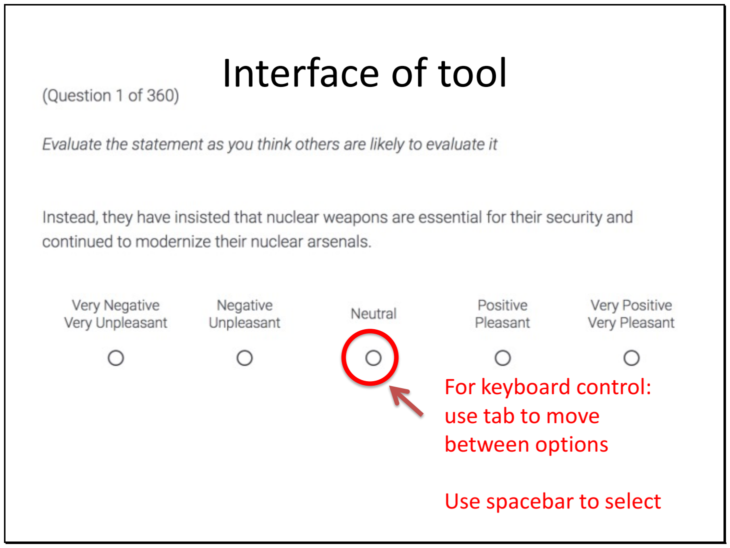

Interface of tool: Context
User 1:
All those deaths, all those Martyrs, all that blood , for what? #Syria
How does seeing the original context of the tweet impact evaluation?
Evaluate User 2's statement
User 2:
"@user1: All those deaths, all those Martyrs, all that blood , for what? #Syria Never underestimate Religious Cults & Big Money :-)
Very Negative Unpleasant
Negative
Neutral
Positive
Very Positive Pleasant

Individuals evaluating news clips first saw the news clip examples, followed by screenshots of the interface.

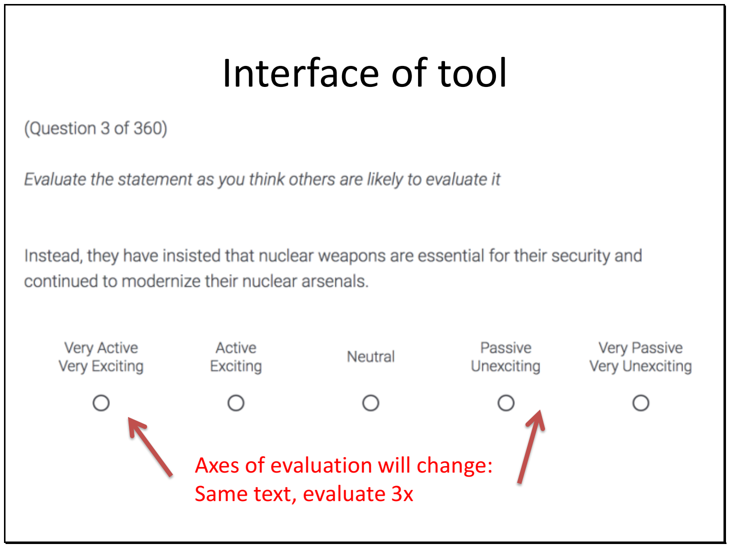
Slide 17



Slide 18

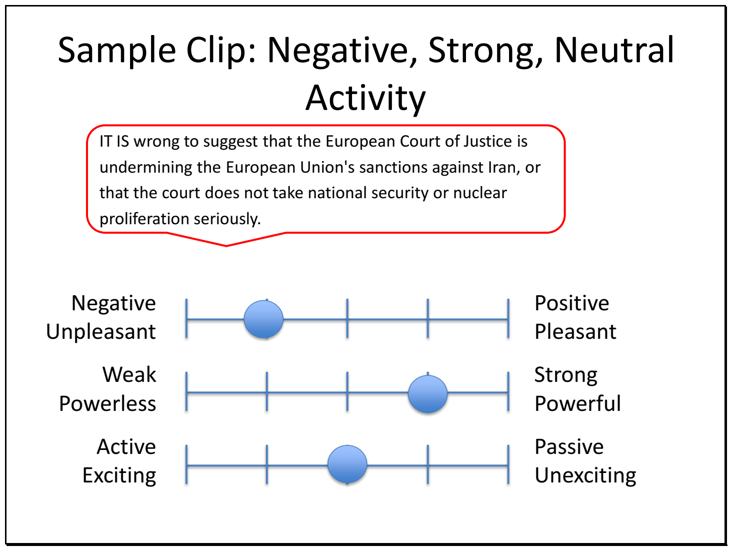

Interface of tool
(Question 1 of 360)
Evaluate the statement as you think others are likely to evaluate it
Instead, they have insisted that nuclear weapons are essential for their security and continued to modernize their nuclear arsenals.
Very Negative
Very Unpleasant
Negative
Unpleasant
Neutral
Positive
Pleasant
Very Positive
Very Pleasant
For keyboard control: use tab to move between options
Use spacebar to select

Slide 19


Interface of tool
(Question 3 of 360)
Evaluate the statement as you think others are likely to evaluate it
Instead, they have insisted that nuclear weapons are essential for their security and continued to modernize their nuclear arsenals.
Very Active
Very Exciting
Active
Exciting
Neutral
Passive
Unexciting
Very Passive
Very Unexciting
Axes of evaluation will change:
Same text, evaluate 3x

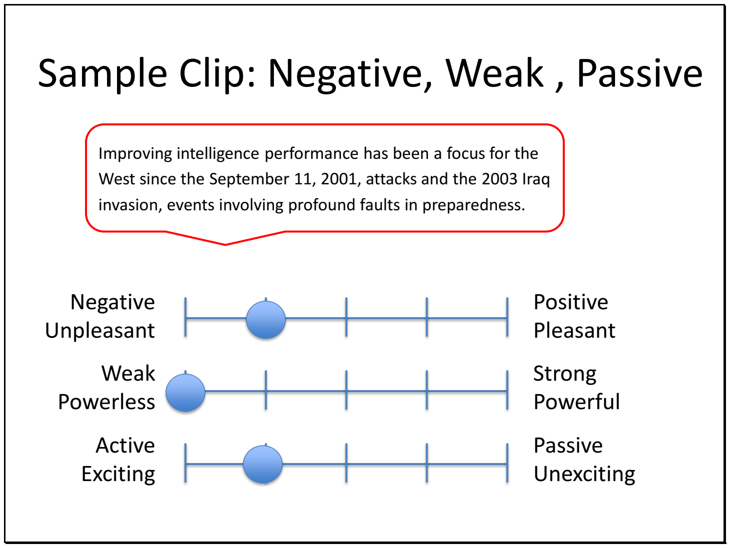
The following news clips examples were provided to individuals rating news statements.

Slide 20



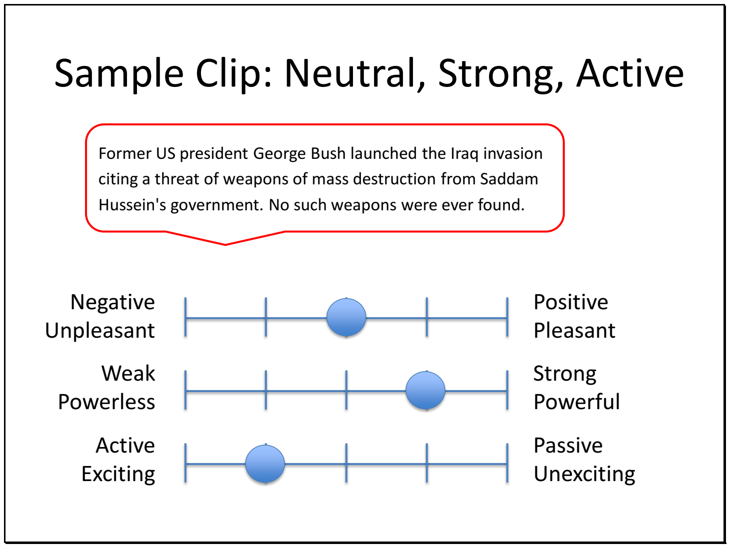
Slightly negative – the Court of Justice is clearly missing something here. These are powerful institutions. However, since we don't know what action is taking place, neither active nor passive.

Slide 21



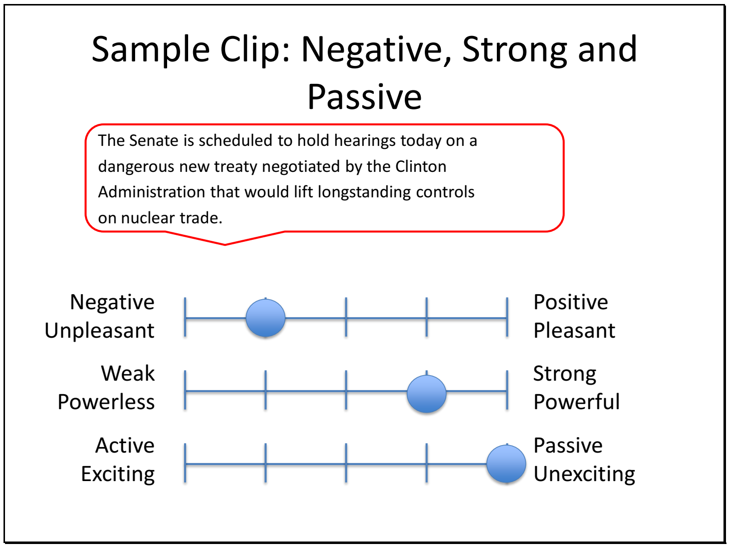
A negative statement – ”faults in preparedness”. The institution is being brought up in the context of weakness – 9/11 – so weak. But very active – the statement is focusing on how to improve intelligence.

Slide 22



Neutral statement – while written in a slightly negative tone, alone these sentences are neutral. Strong institutions referenced. Active positions taken.

Slide 23



Slightly negative here – a "dangerous" new treaty. "Senate", "Clinton Administration" powerful institutions. However, passive – it's not that the Senate is holding hearings – they're "scheduled to hold" hearings.

Appendix: Consent Form

Online Consent Form: Social Media ACTion

This social media coding is part of a research study conducted by Will Frankenstein and Kenneth Joseph at Carnegie Mellon University and is funded by Crosswalk- Graduate Student Small Project.

The purpose of the research is to develop a 'gold standard' of social media posts encoded by affect control theory. The three dimensions measured in affect control theory are: emotion (positive to negative), powerful (weak to strong), and action (lively to quiet).

Procedures

Participants will view a variety of short, anonymous social media posts, and rate them along a 5 point Likert scale for each of the three dimensions of affect control theory. In some cases, participants will see a social media post twice; this will be done to provide more context for the original post.

The study is expected to take 45 minutes to complete.

Participant Requirements

Participation in this study is limited to individuals age 18 and older. Participants must be native English speakers.

Risks

The risks and discomfort associated with participation in this study are no greater than those ordinarily encountered in daily life or during other online activities. The primary risk to participants is boredom or fatigue from reading several social media posts in one sitting.

Benefits

There may be no personal benefit from your participation in the study but the knowledge received may be of value to humanity.

Compensation & Costs

Participants will be paid $8 in Amazon gift cards for completion of the study. Individuals who do not complete the study will be compensated at a rate of 6 cents per social media post viewed in Amazon gift cards.

There will be no cost to you if you participate in this study.

Confidentiality

The data captured for the research does not include any personally identifiable information about you. We will capture some summary demographic information about you, but it will not be linked to yourself or the data provided.

Your data and consent form will be kept separate. Your consent form will be stored in a locked location on Carnegie Mellon property and will not be disclosed to third parties. By participating, you understand and agree that the data and information gathered during this study may be used by Carnegie Mellon and published and/or disclosed by Carnegie Mellon to others outside of Carnegie Mellon. However, your name, address, contact information and other direct personal identifiers in your consent form will not be mentioned in any such publication or dissemination of the research data and/or results by Carnegie Mellon.

Right to Ask Questions & Contact Information

If you have any questions about this study, you should feel free to ask them by contacting the Principal Investigator, Will Frankenstein, PhD Candidate in Department of Engineering & Public Policy, Baker Hall 129, 5000 Forbes Avenue, Pittsburgh, PA 15213 / frankenstein@cmu.edu / 412-589-9788. If you have questions later, desire additional information, or wish to withdraw your participation please contact the Principal Investigator by mail, phone or e-mail in accordance with the contact information listed above.

If you have questions pertaining to your rights as a research participant; or to report objections to this study, you should contact the Office of Research integrity and Compliance at Carnegie Mellon University. Email: irb-review@andrew.cmu.edu . Phone: 412-268-1901 or 412-268-5460.

Voluntary Participation

Your participation in this research is voluntary. You may discontinue participation at any time during the research activity. However, not completing the study will mean that you will not be compensated for your time.

[Design the web page so that the following questions must be answered appropriately before the individual can proceed to the study task.]

I am age 18 or older. ☐ Yes ☐ No

I have read and understand the information above. ☐ Yes ☐ No

I want to participate in this research and continue with the coding ☐ Yes ☐ No

[if the answer is no to any of the above questions, the individual cannot participate and should not be allowed to proceed to the next question.]

Appendix B: Ego Nets of 2015

This appendix contains the ego nets of ally and hostility networks in 2015. Blue nodes indicate countries without nuclear weapons; yellow indicates countries with nuclear weapons. Green ties indicate an alliance; red indicates a hostility. Alliance information is taken from COW; Hostility information is taken from the International Crisis Behavior dyadic dataset. Both networks reflect a 10-year span, from 2005-2015.

Table 71. Table of network isolates

Alliance Network Isolates	Hostility Network Isolates
Andorra, Austria, Bangladesh, Bhutan, Botswana, Brunei, Burma, Cambodia, Comoros, Cuba, Cyprus, East Timor, Fiji, Indonesia, Ireland, ISIS, Kiribati, Kosovo, Laos, Lesotho, Liechtenstein, Macedonia, Madagascar, Malawi, Malaysia, Maldives, Malta, Marshall Islands, Mauritius, Micronesia, Monaco, Montenegro, Mozambique, Nauru, Nepal, New Zealand, Palau, Papua New Guinea, Samoa, San Marino, Seychelles, Singapore, Slovenia, Solomon Islands, Sri Lanka, Sweden, Switzerland, Taiwan, Thailand, Tonga, Tuvalu, Vanuatu, Vietnam	Afghanistan, Albania, Algeria, Andorra, Angola, Antigua Barbuda, Argentina, Armenia, Australia, Austria, Bahamas, Bahrain, Bangladesh, Barbados, Belarus, Belgium, Belize, Benin, Bhutan, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Brunei, Bulgaria, Burkina Faso, Burma, Burundi, Cameroon, Cape Verde, Central African Republic, Chile, Colombia, Comoros, Congo, Costa Rica, Croatia, Cuba, Cyprus, Czech Republic, Denmark, Dominica, Dominican Republic, East Timor, Ecuador, Egypt, El Salvador, Equatorial Guinea, Estonia, Federated States of Micronesia, Fiji, Finland, Gabon, Gambia, Germany, Ghana, Greece, Grenada, Guatemala, Guinea, Guinea Bissau, Guyana, Haiti, Honduras, Hungary, Iceland, Indonesia, Iraq, Ireland, Jamaica, Jordan, Kazakhstan, Kenya, Kiribati, Kosovo, Kuwait, Kyrgyzstan, Laos, Latvia, Lesotho, Liberia, Liechtenstein, Lithuania, Luxembourg, Macedonia, Madagascar, Malawi, Malaysia, Maldives, Mali, Malta, Marshall Islands, Mauritania, Mauritius, Mexico, Moldova, Monaco, Mongolia, Montenegro, Morocco, Mozambique, Namibia, Nauru, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Oman, Pakistan, Palau, Panama, Papua New Guinea, Paraguay, Peru, Poland, Portugal, Romania, Rwanda, Samoa, San Marino, Sao Tome and Principe, Saudi Arabia, Senegal, Seychelles, Sierra Leone, Singapore, Slovakia, Slovenia, Solomon Islands, South Africa, Spain, Sri Lanka, St Kitts and Nevis, St Lucia, St Vincent, Suriname, Swaziland, Sweden, Switzerland, Taiwan, Tajikistan, Tanzania, Togo, Tonga, Trinidad and Tobago, Tunisia, Turkmenistan, Tuvalu, Uganda, United Arab Emirates, Uruguay, Uzbekistan, Vanuatu, Venezuela, Vietnam, Yemen, Yugoslavia, Zambia, Zimbabwe
Egonet Isolates (not connected in either Alliance or Hostility Networks): Andorra, Austria, Bangladesh, Bhutan, Botswana, Brunei, Burma, Comoros, Cuba, Cyprus, East Timor, Fiji, Indonesia, Ireland, Kiribati, Laos, Lesotho, Lichtenstein, Macedonia, Madagascar, Malawi, Malaysia, Maldives, Malta, Marshall Islands, Mauritius, Micronesia, Monaco, Montenegro, Mozambique, Nauru, Nepal, New Zealand, Palau, Papua New Guinea, Samoa, San Marino, Seychelles, Singapore, Slovenia, Solomon Islands, Sri Lanka, Sweden, Switzerland, Taiwan, Tonga, Tuvalu, Vanuatu, Vietnam	

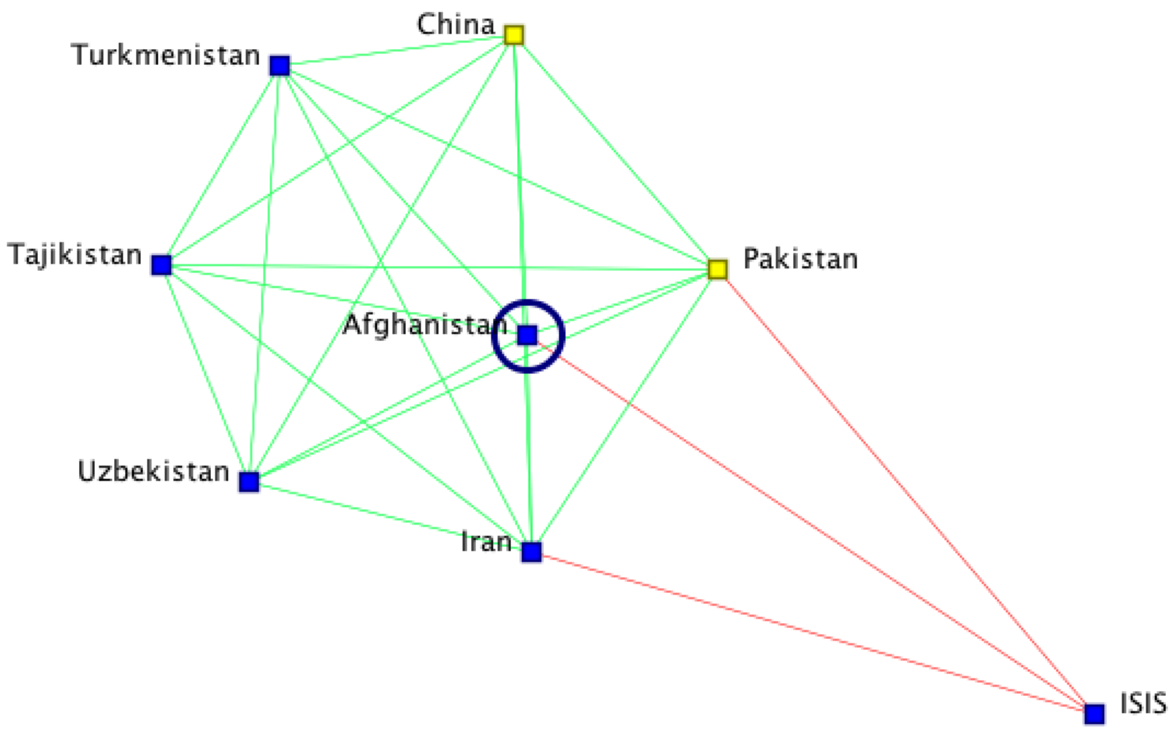


Figure 33. Egonet of Afghanistan

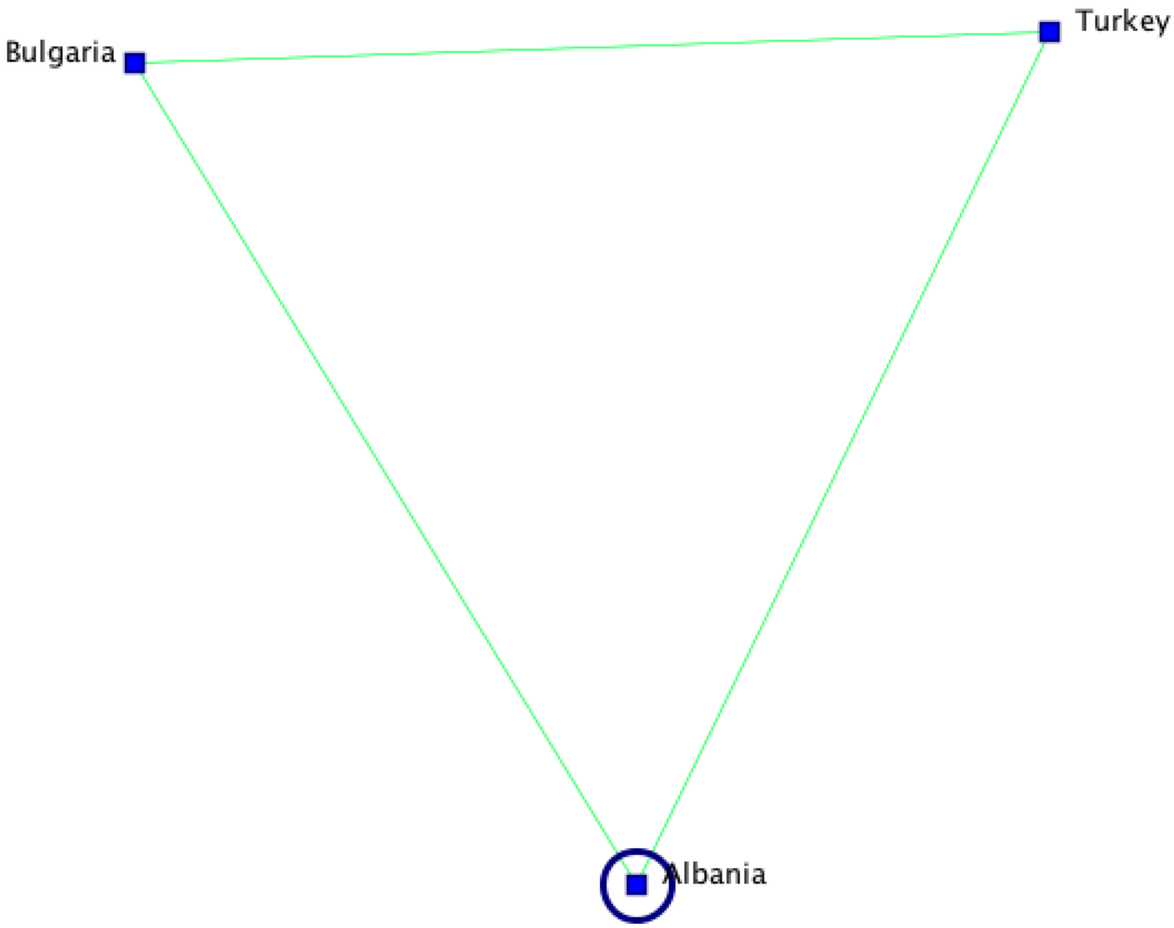


Figure 34. Egonet of Albania

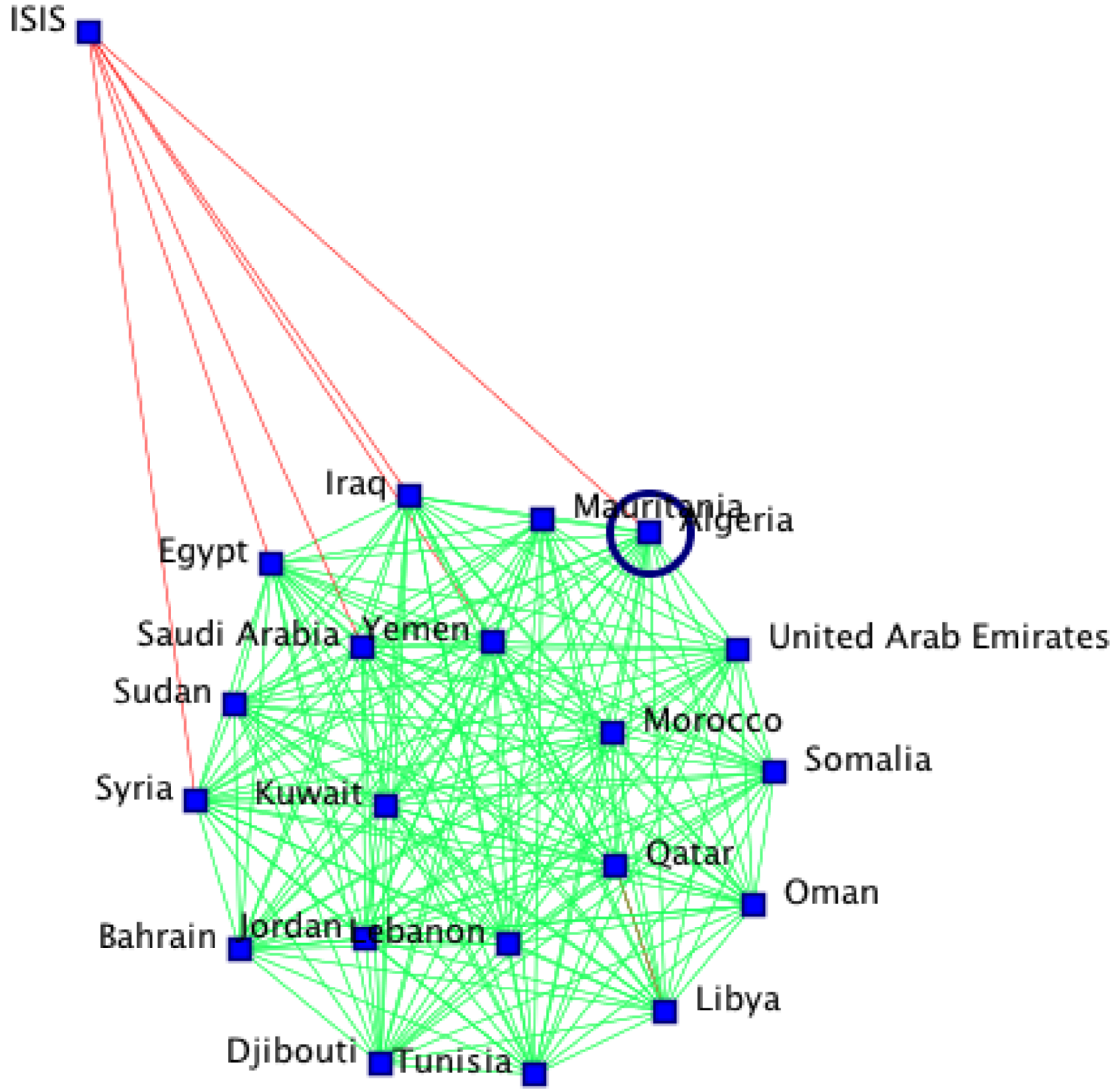


Figure 35. Egonet of Algeria

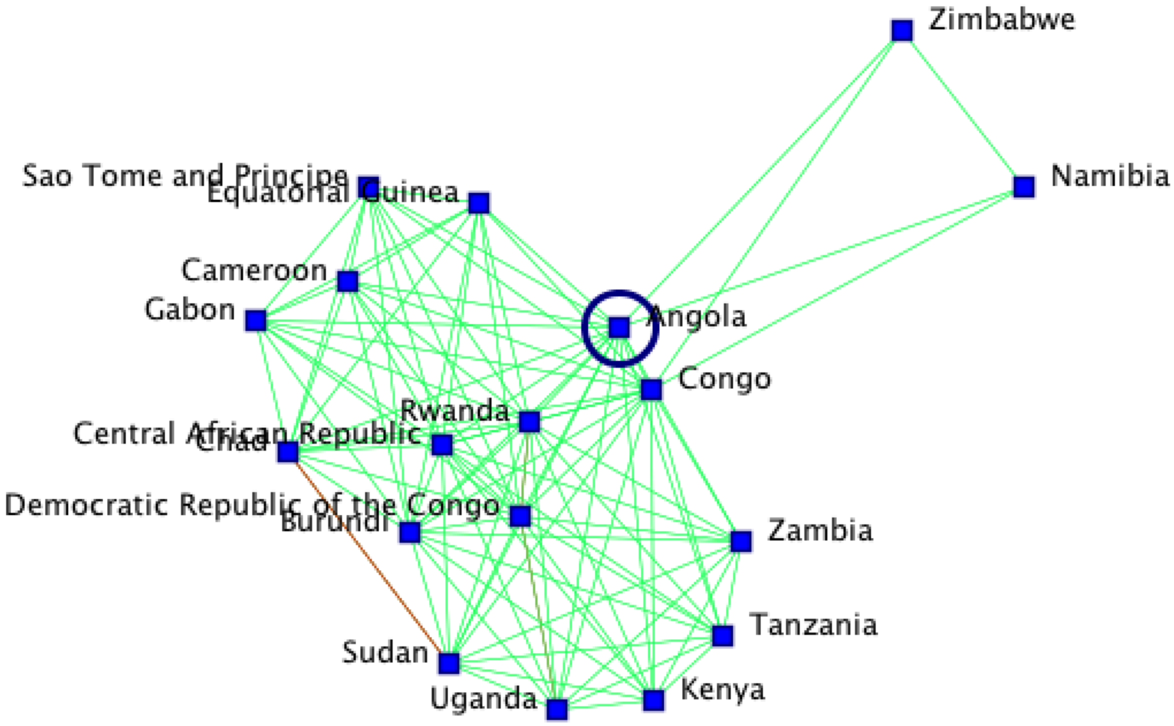


Figure 36. Egonet of Angola

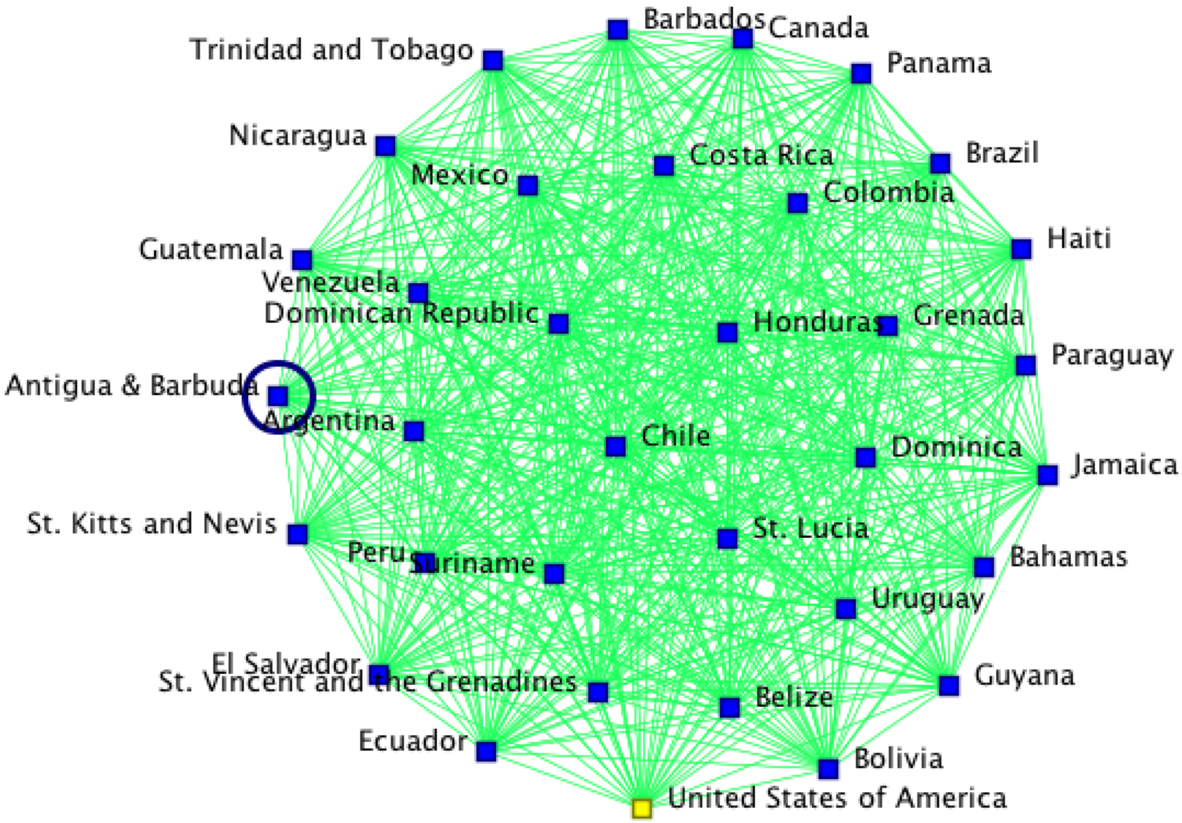


Figure 37. Egonet of Antigua

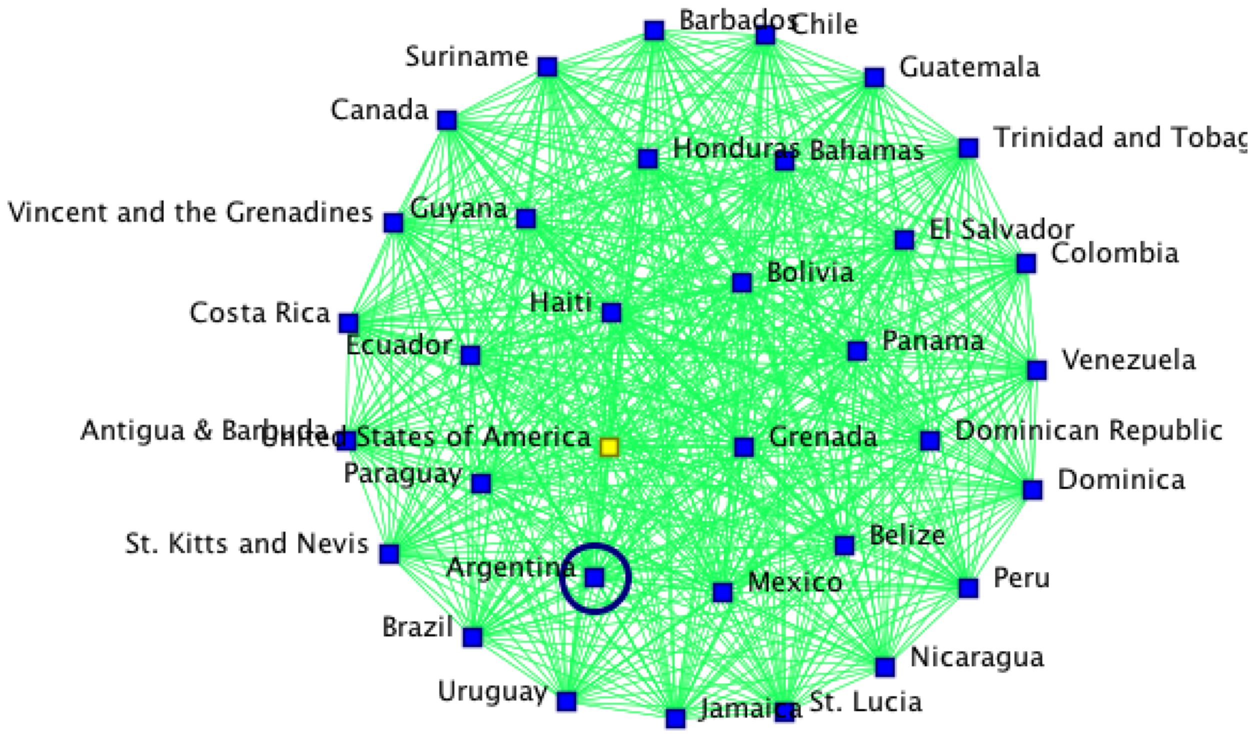


Figure 38. Egonet of Argentina

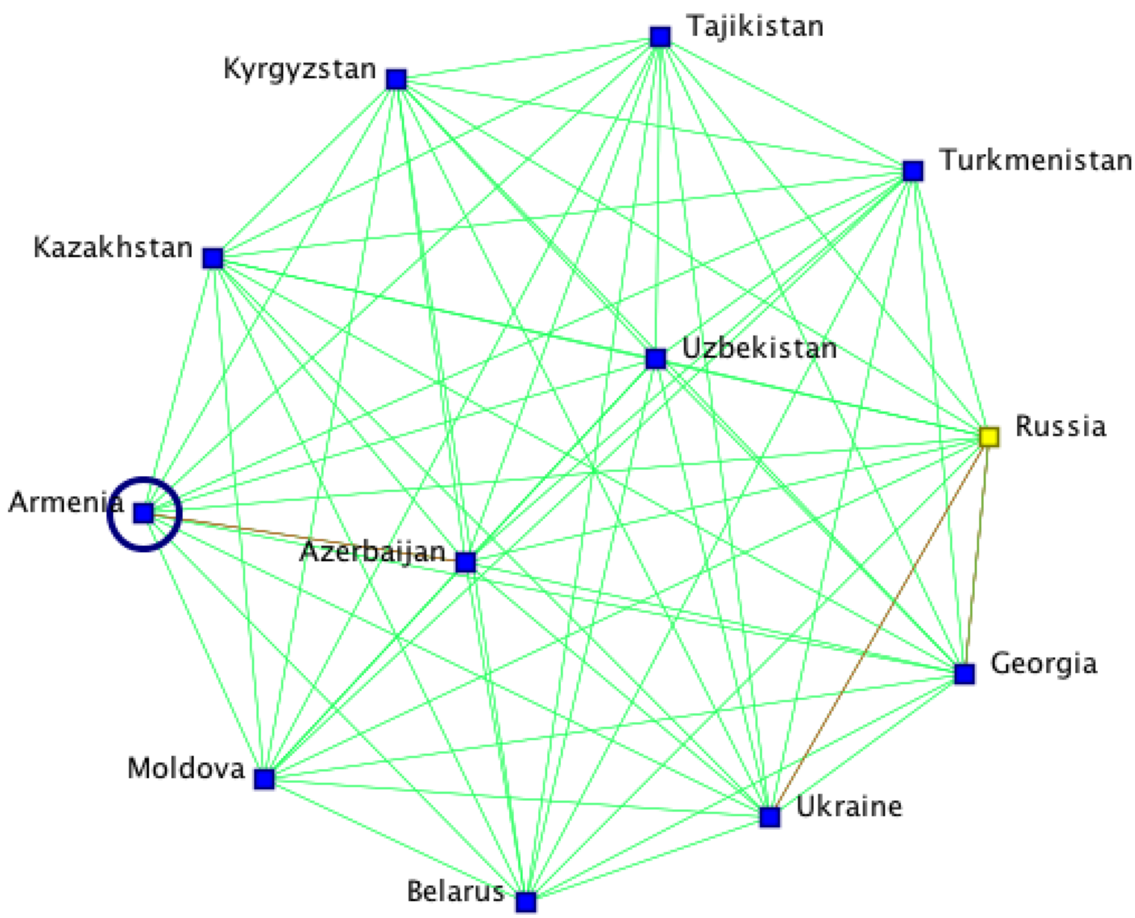


Figure 39.Egonet of Armenia

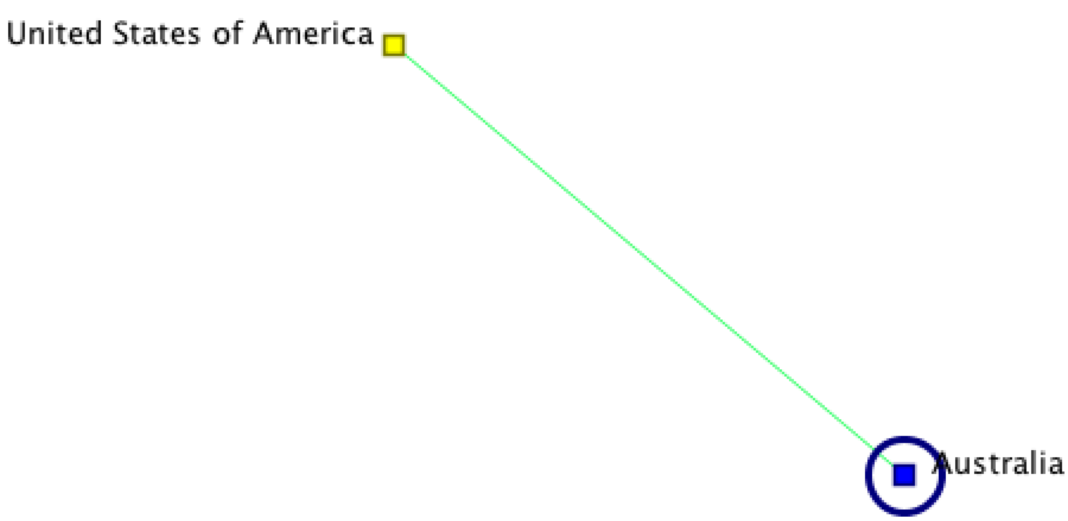


Figure 40. Egonet of Australia

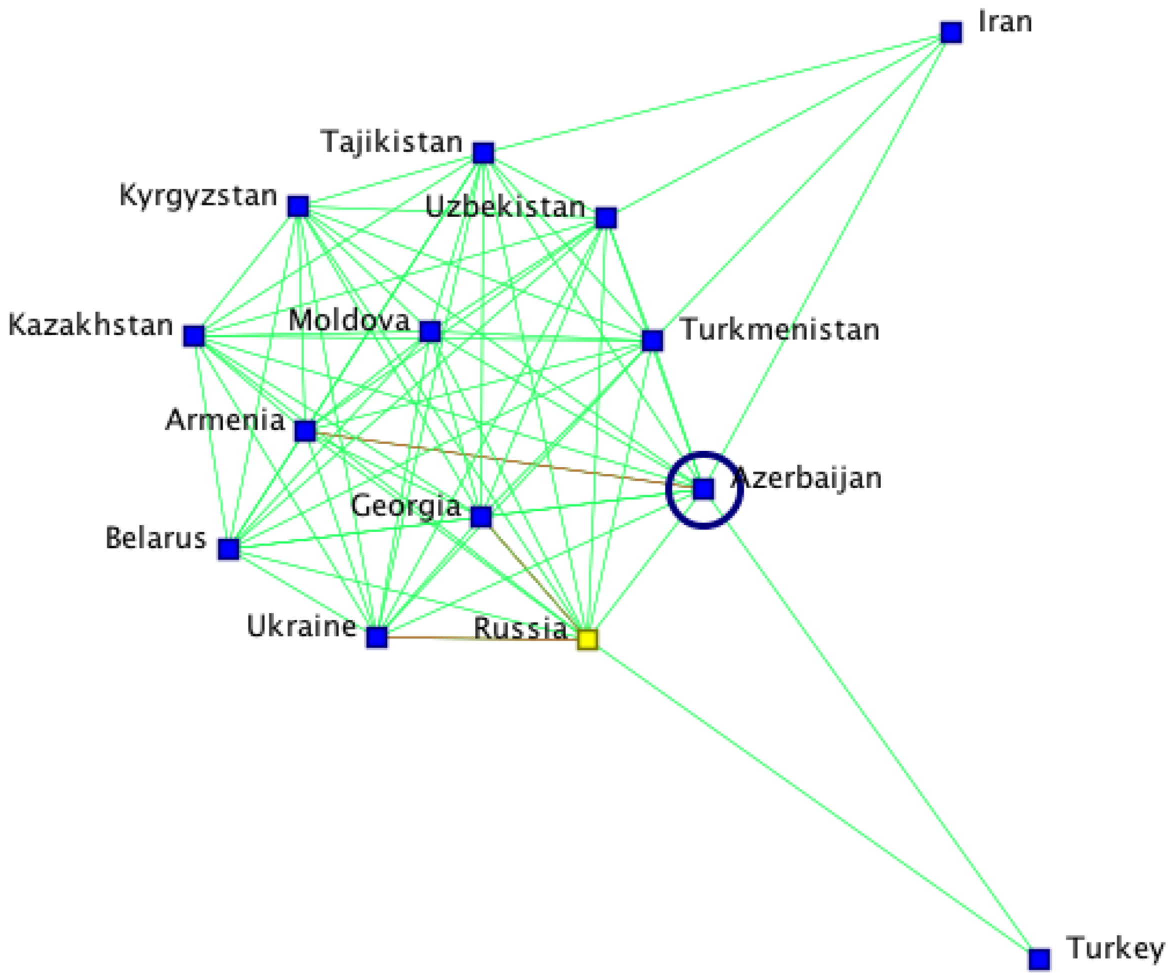


Figure 41. Egonet of Azerbaijan

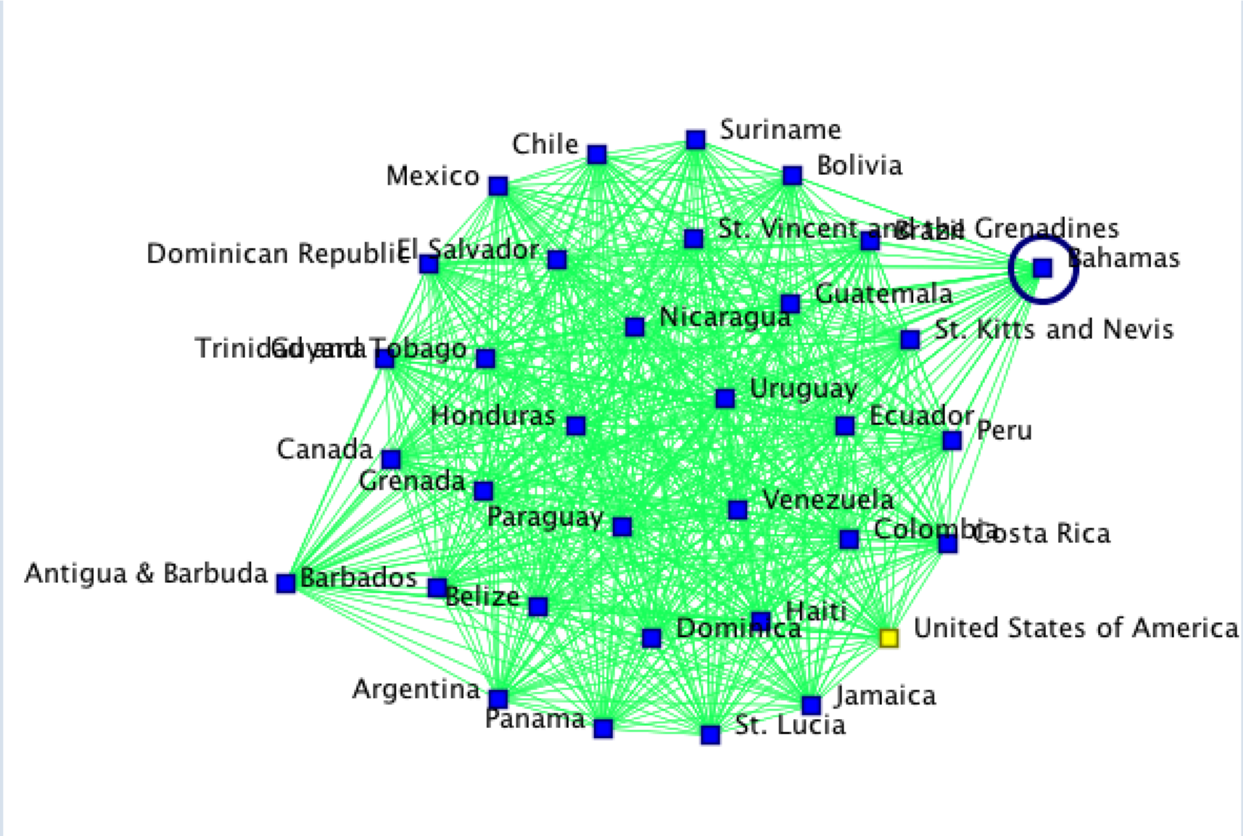


Figure 42 .Egonet of Bahamas

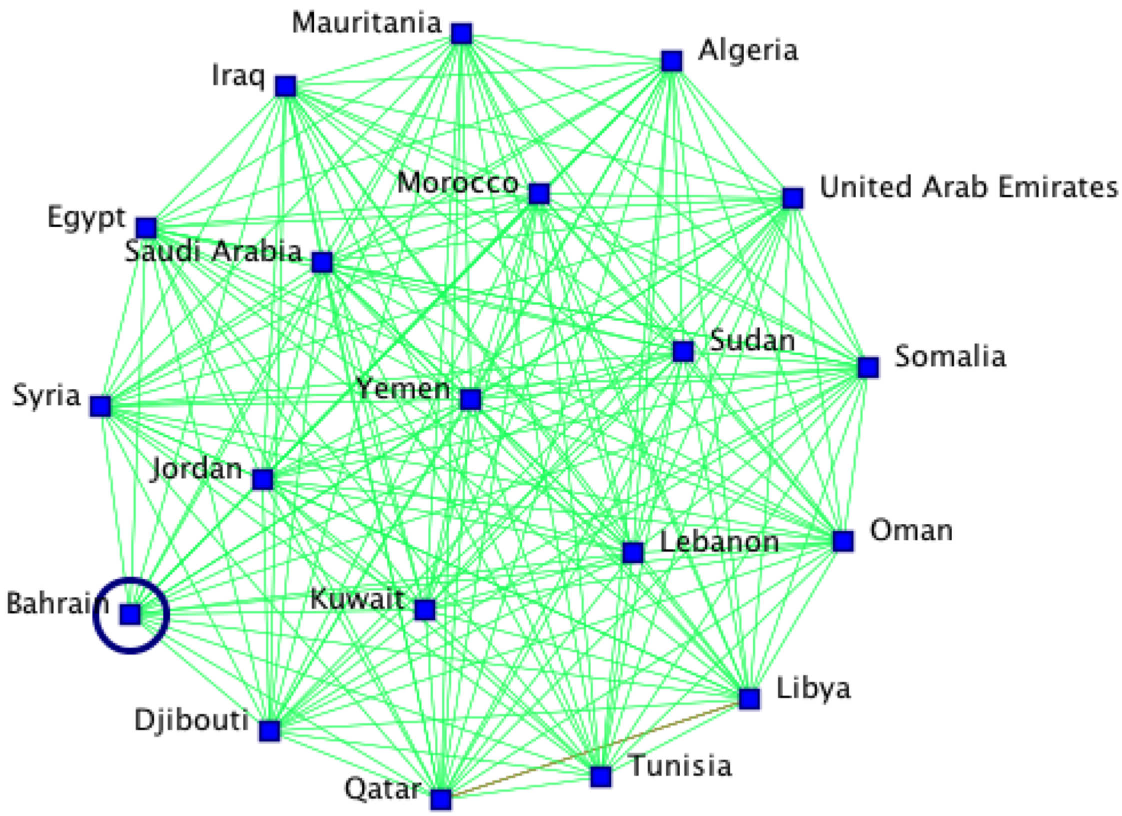


Figure 43. Egonet of Bahrain

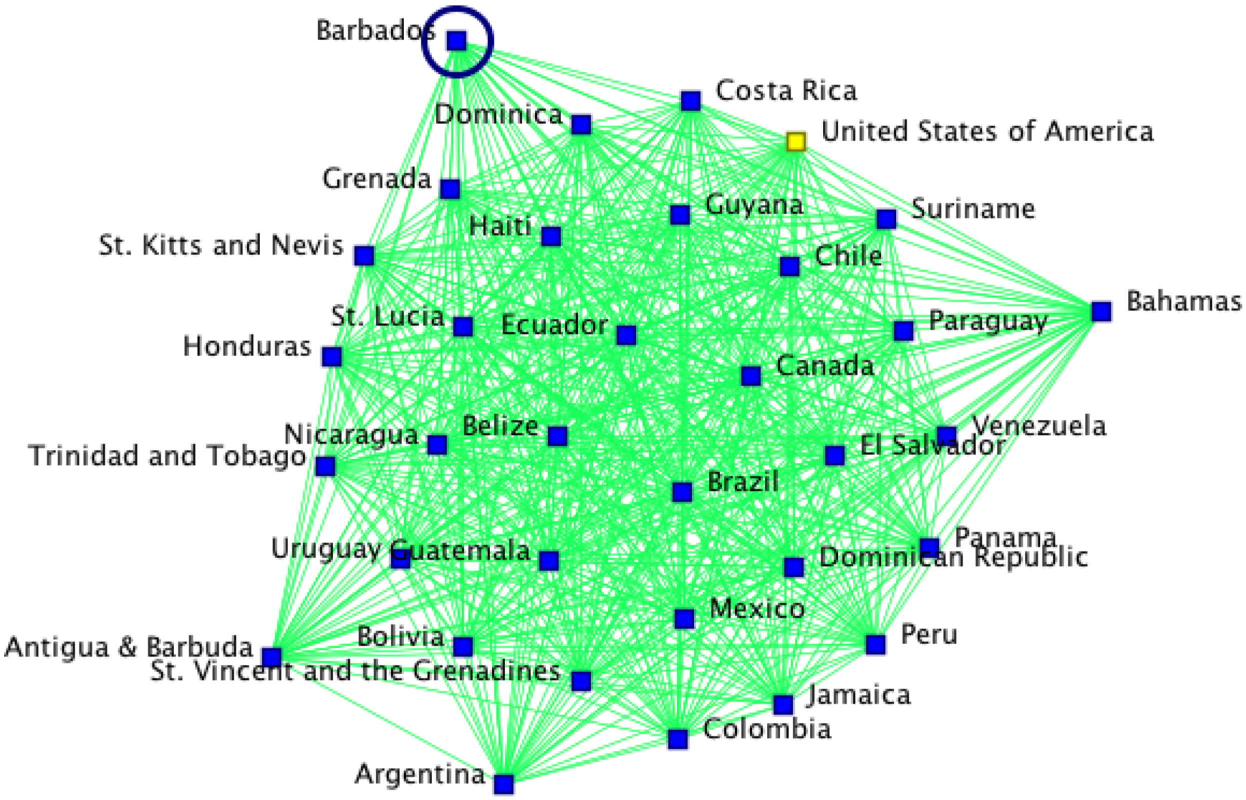


Figure 44. Egonet of Barbados

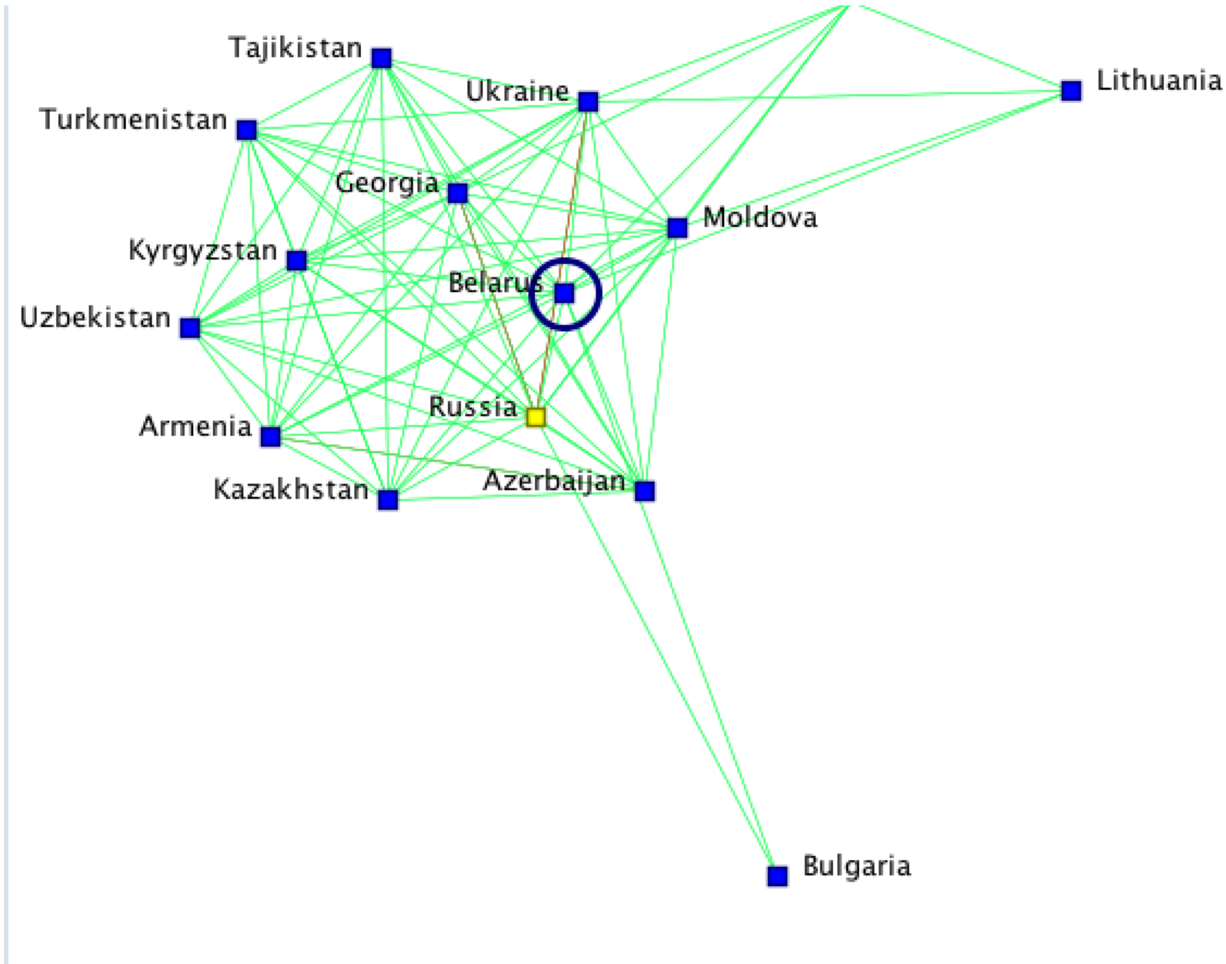


Figure 45. Egonet of Belarus

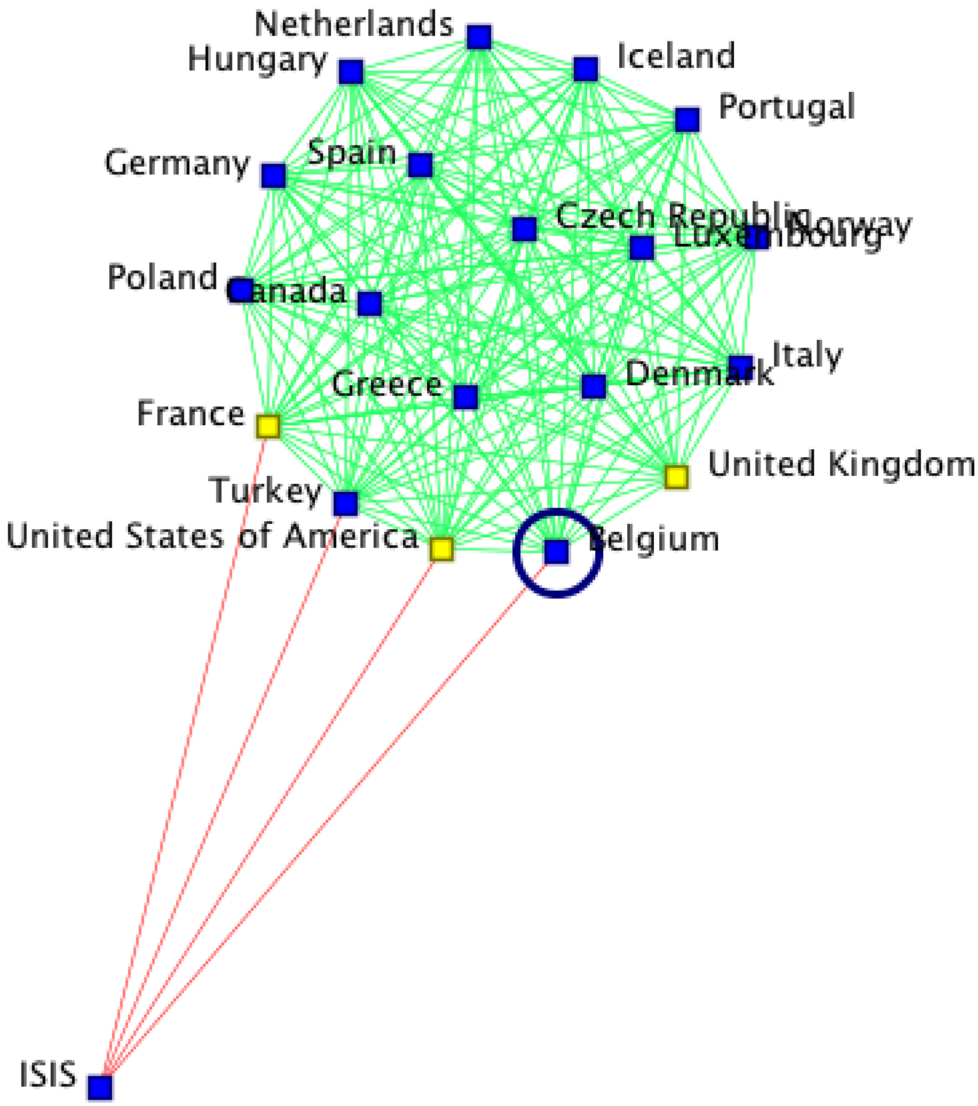


Figure 46. Egonet of Belgium

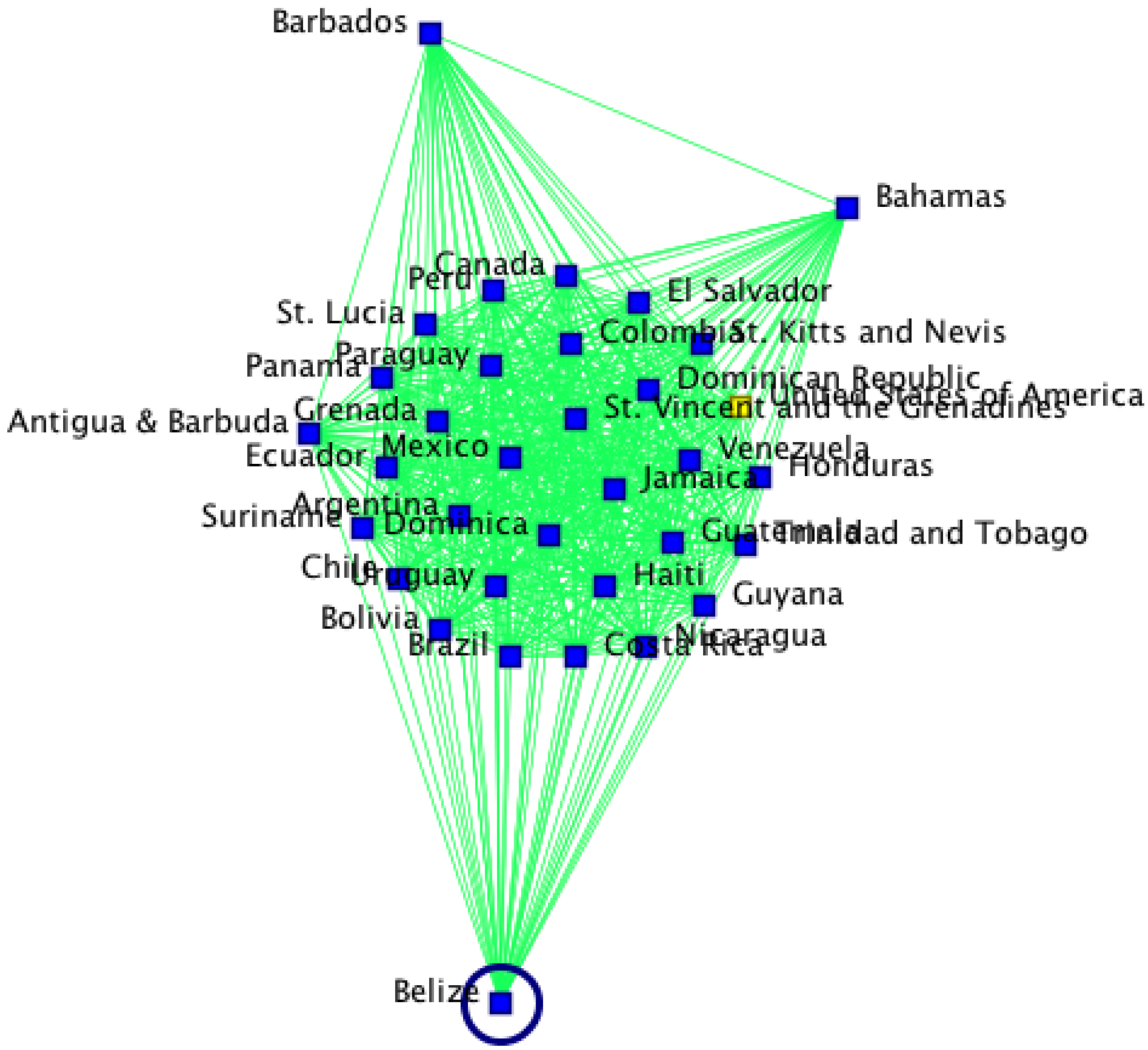


Figure 47. Egonet of Belize

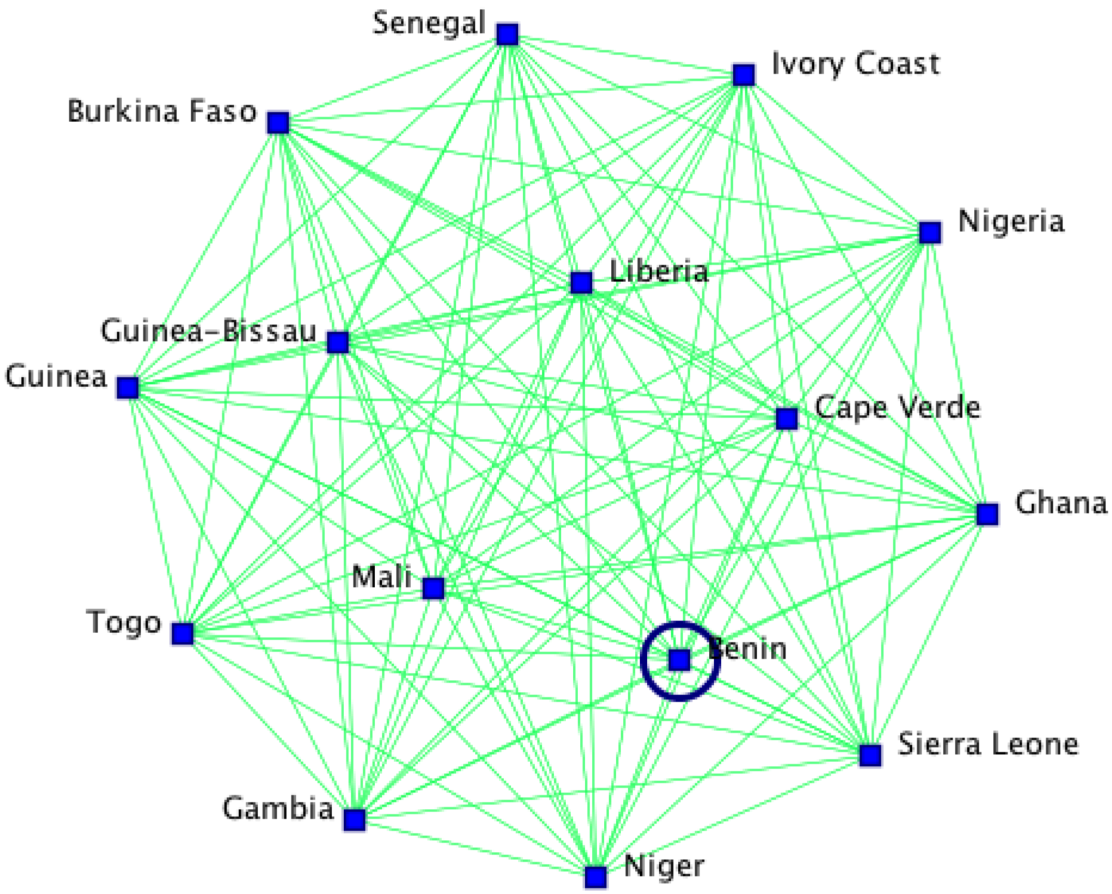


Figure 48. Egonet of Benin

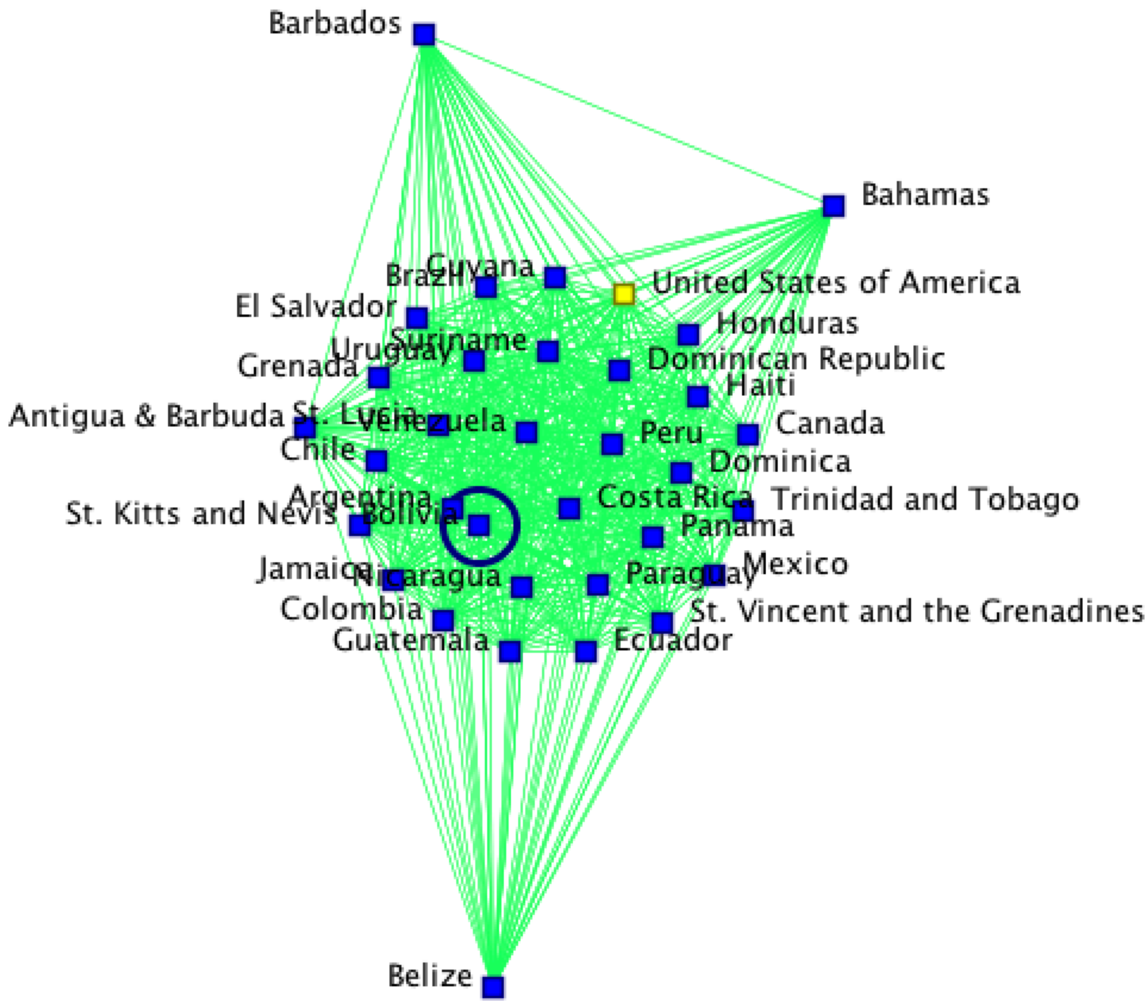


Figure 49. Egonet of Bolivia

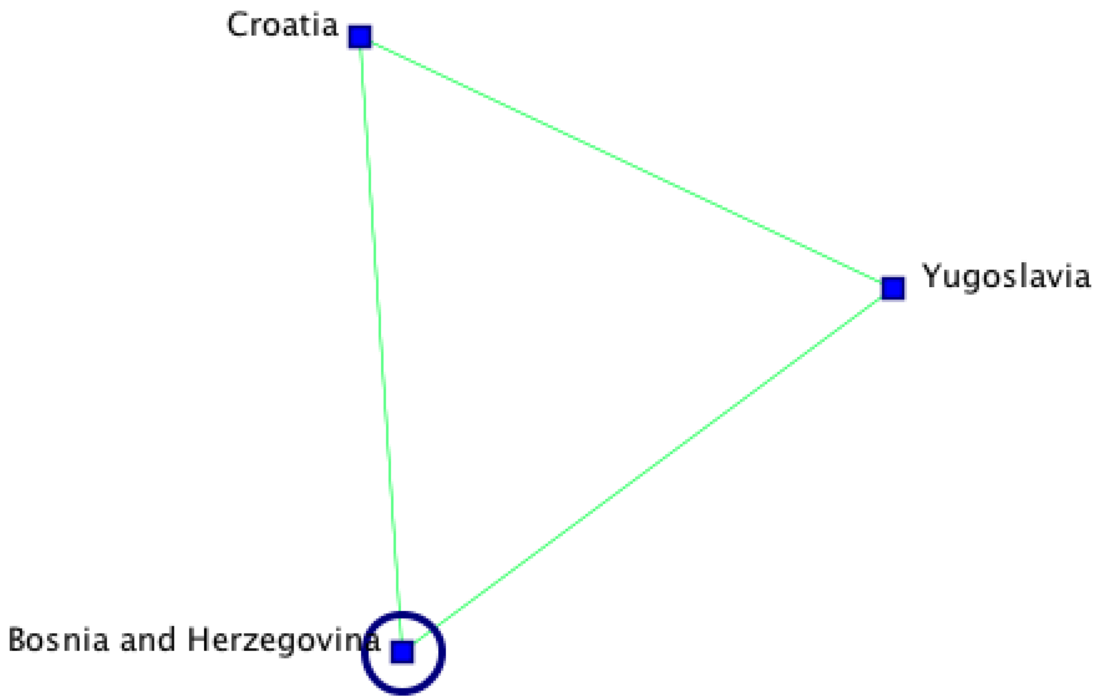


Figure 50. Egonet of Bosnia.

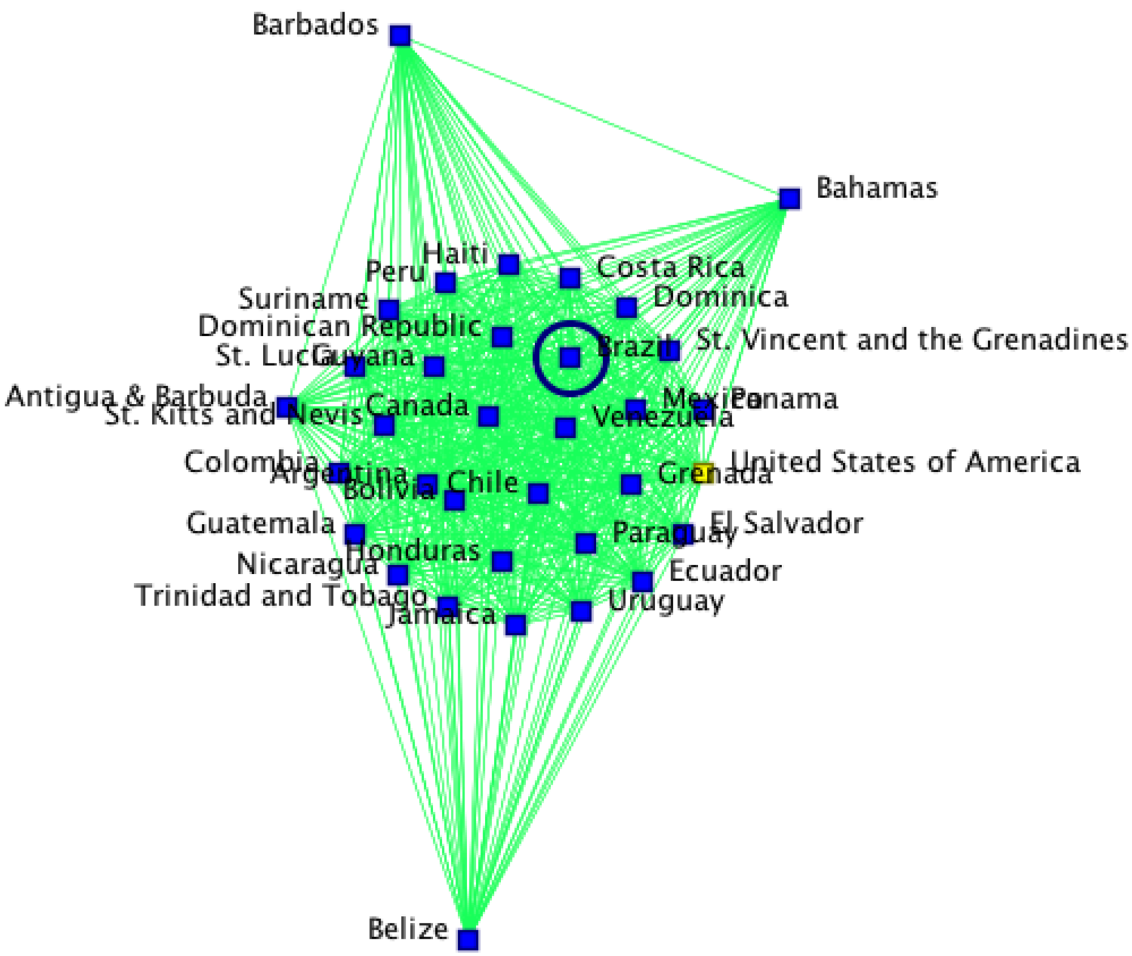


Figure 51. Egonet of Brazil

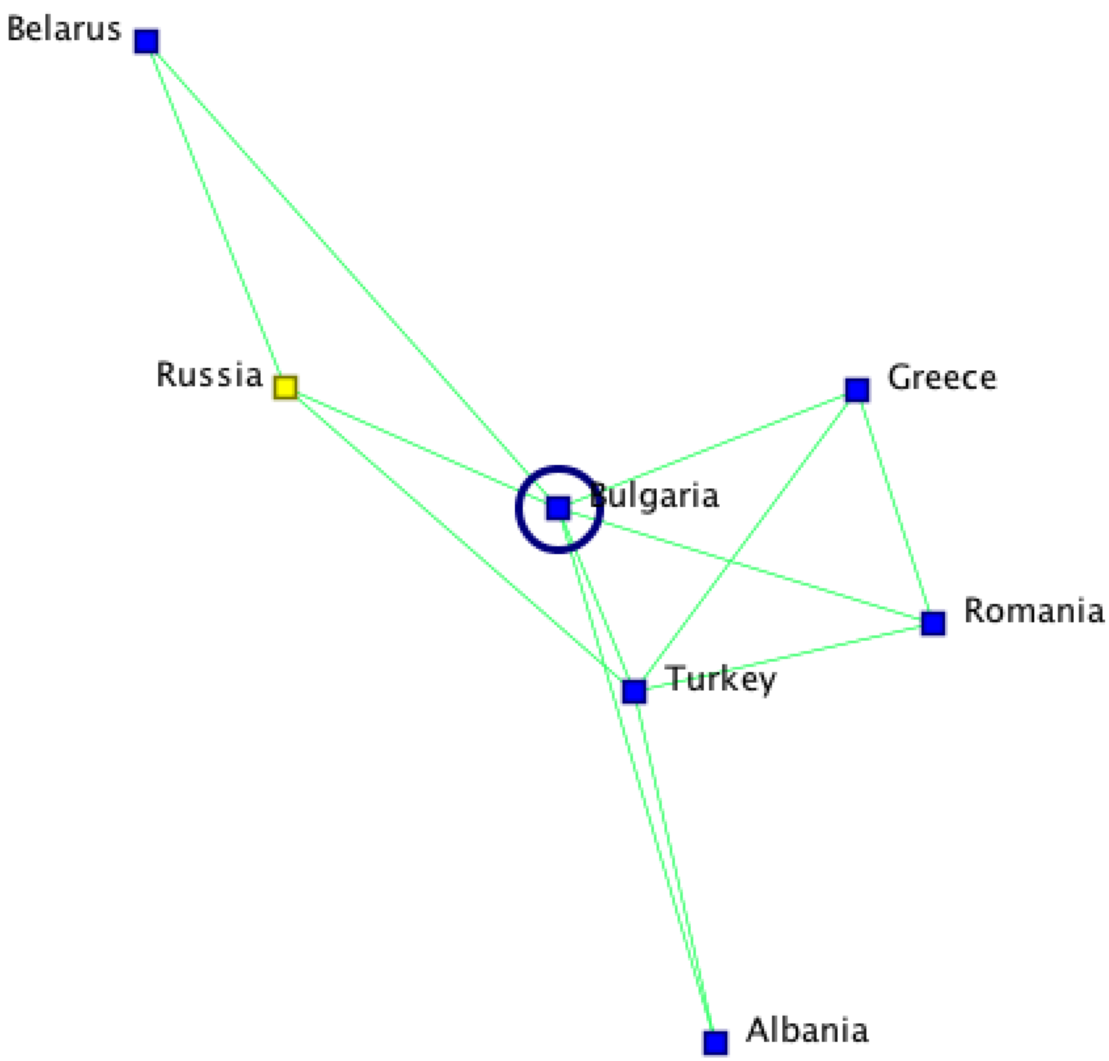


Figure 52. Egonet of Bulgaria

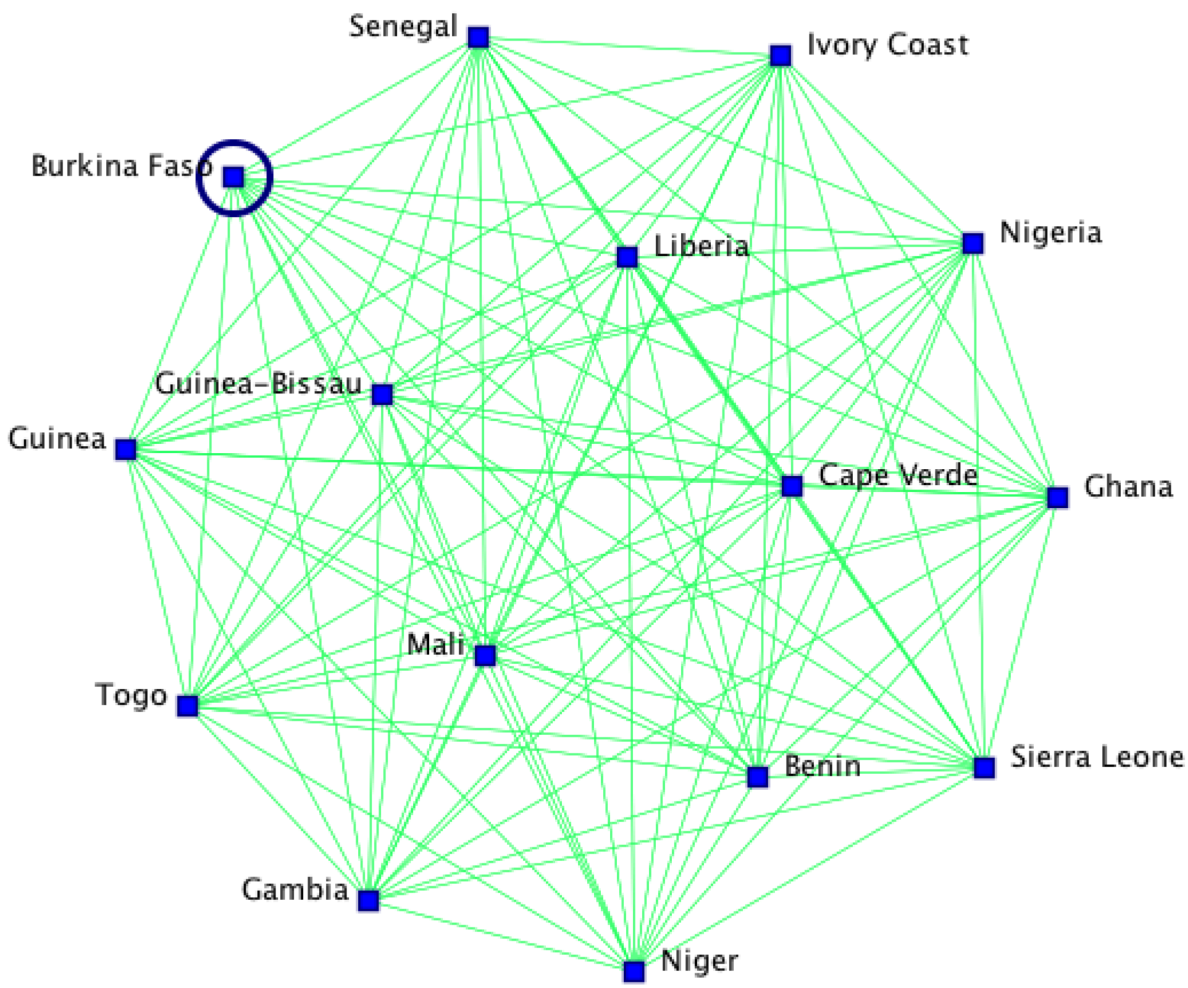


Figure 53. Egonet of Burkina Faso

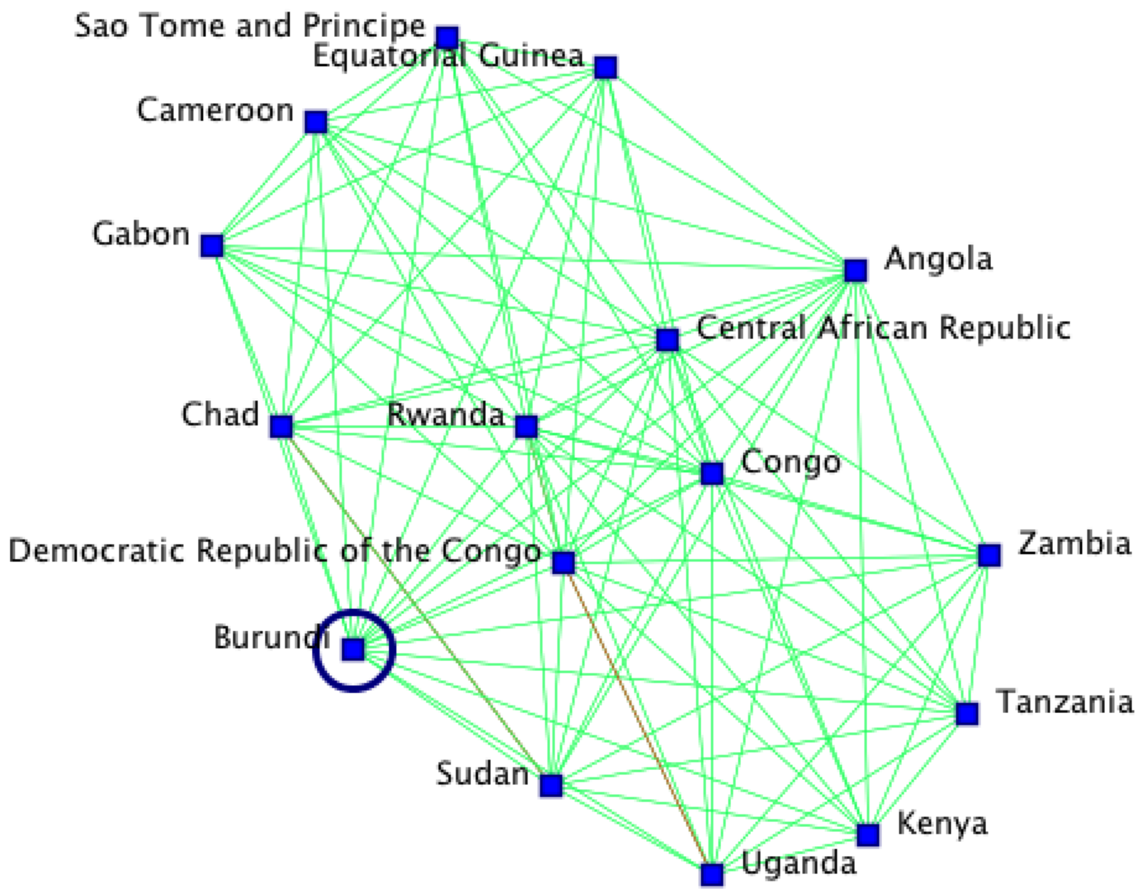


Figure 54. Egonet of Burundi

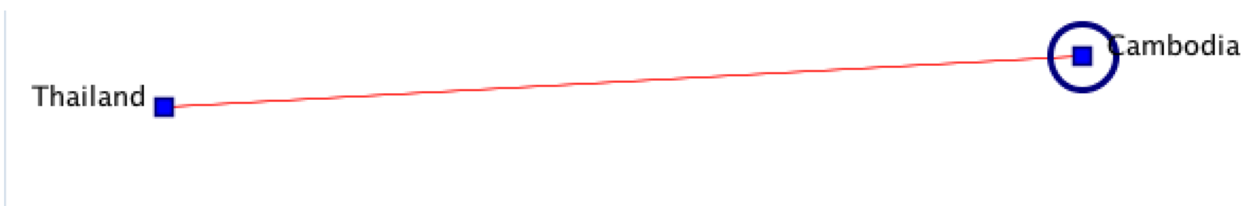


Figure 55. Egonet of Cambodia

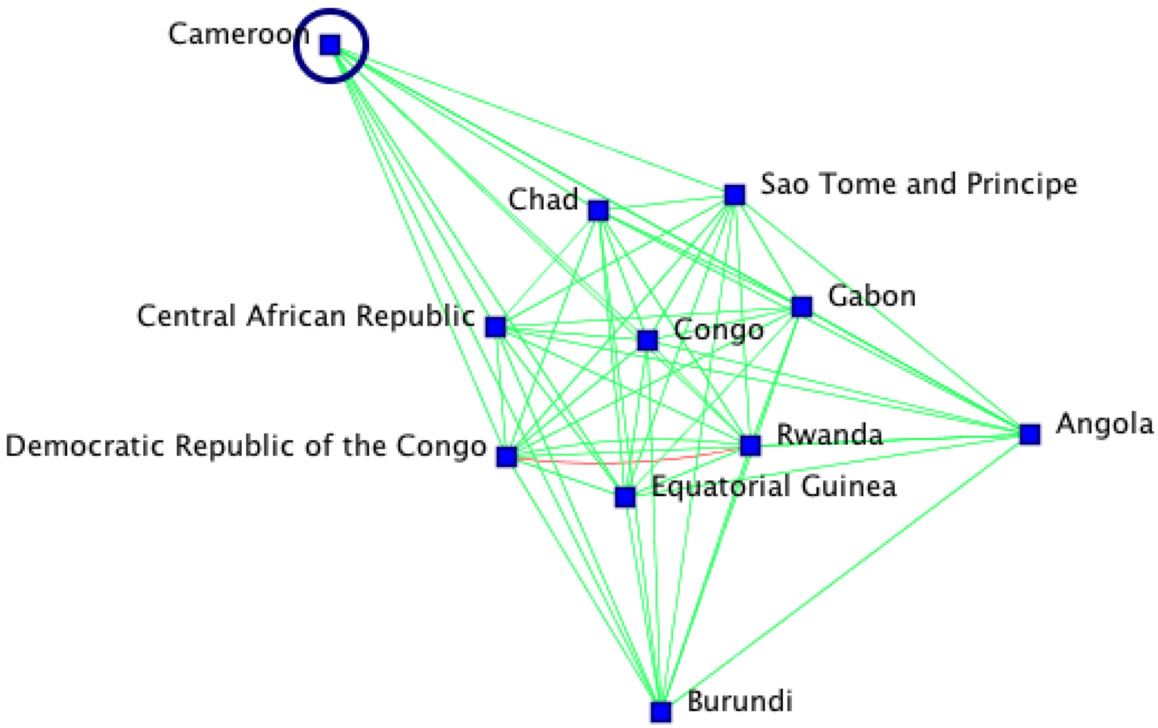


Figure 56. Egonet of Cameroon

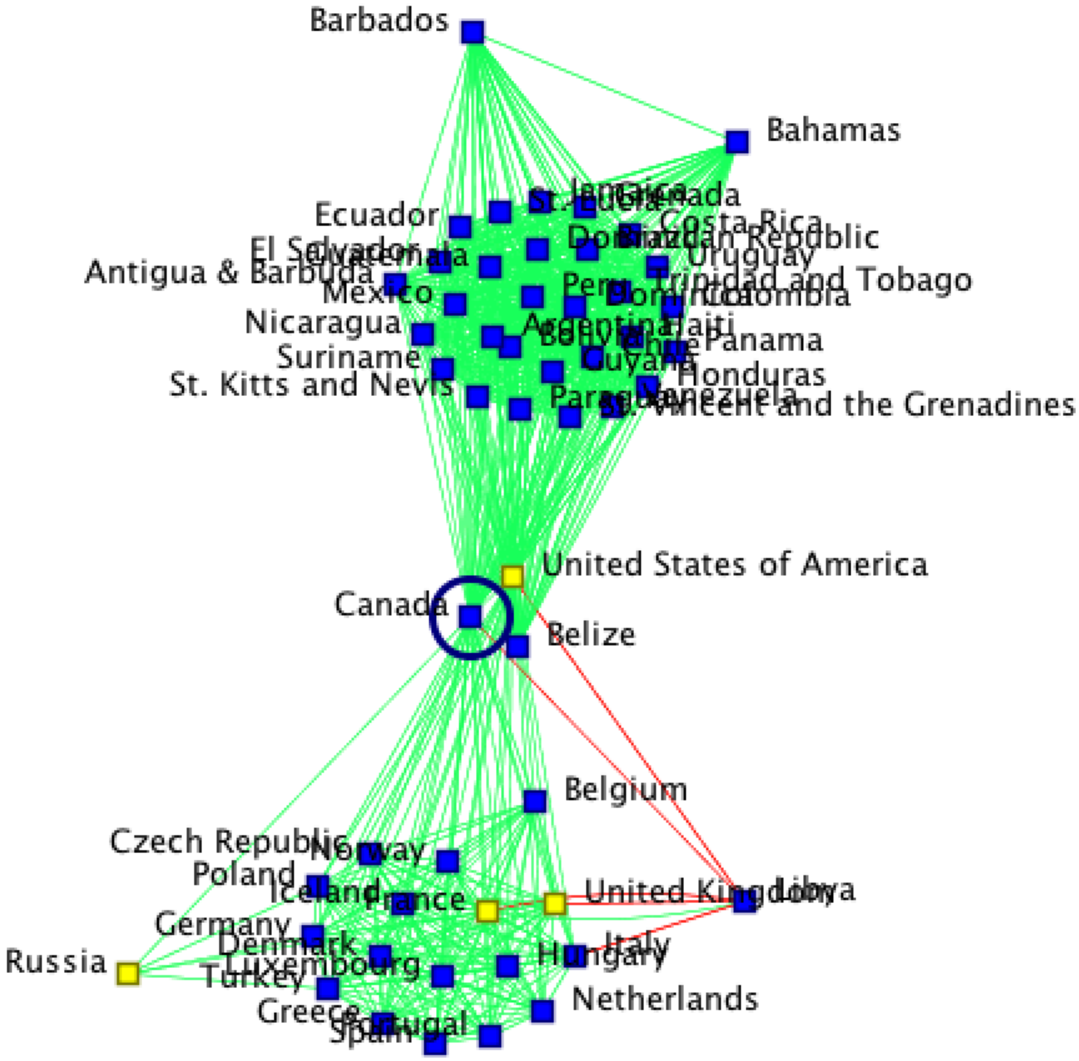


Figure 57. Egonet of Canada

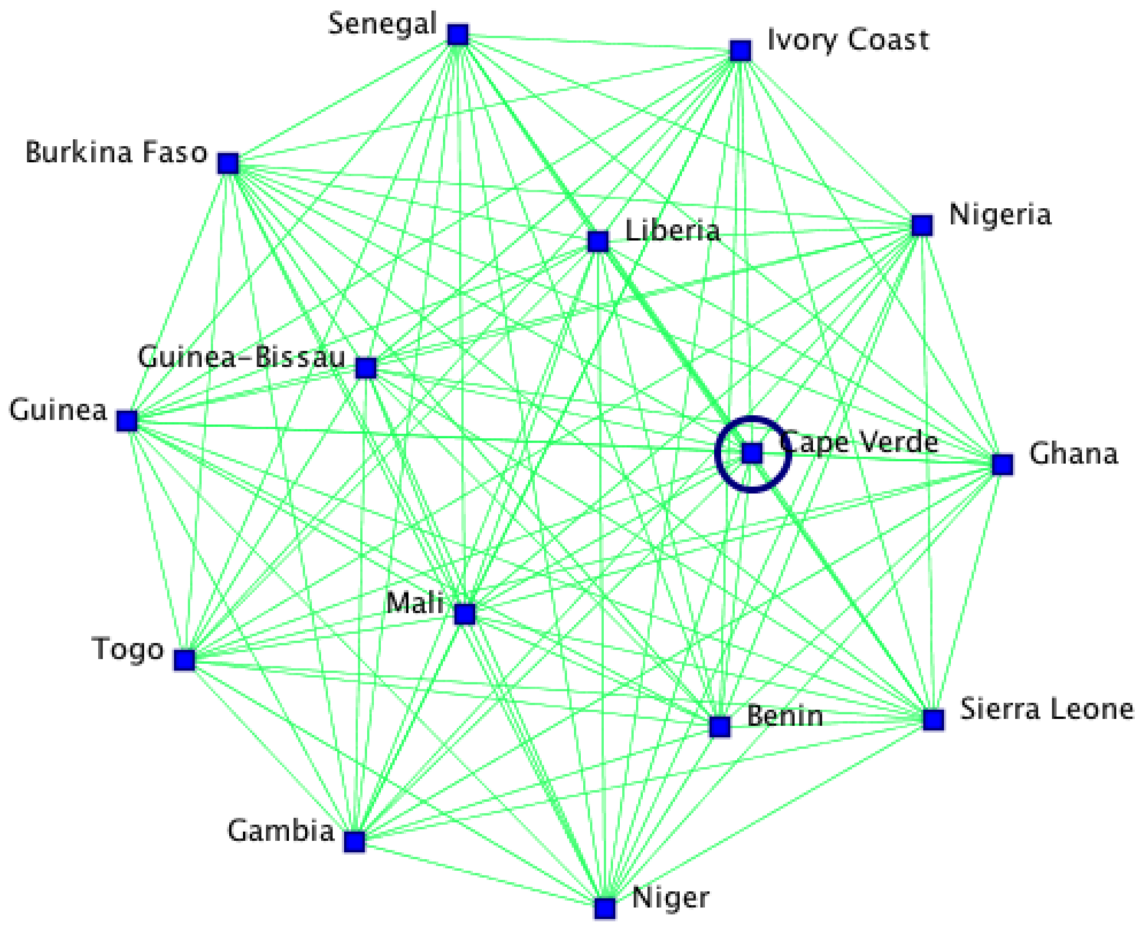


Figure 58. Egonet of Cape Verde

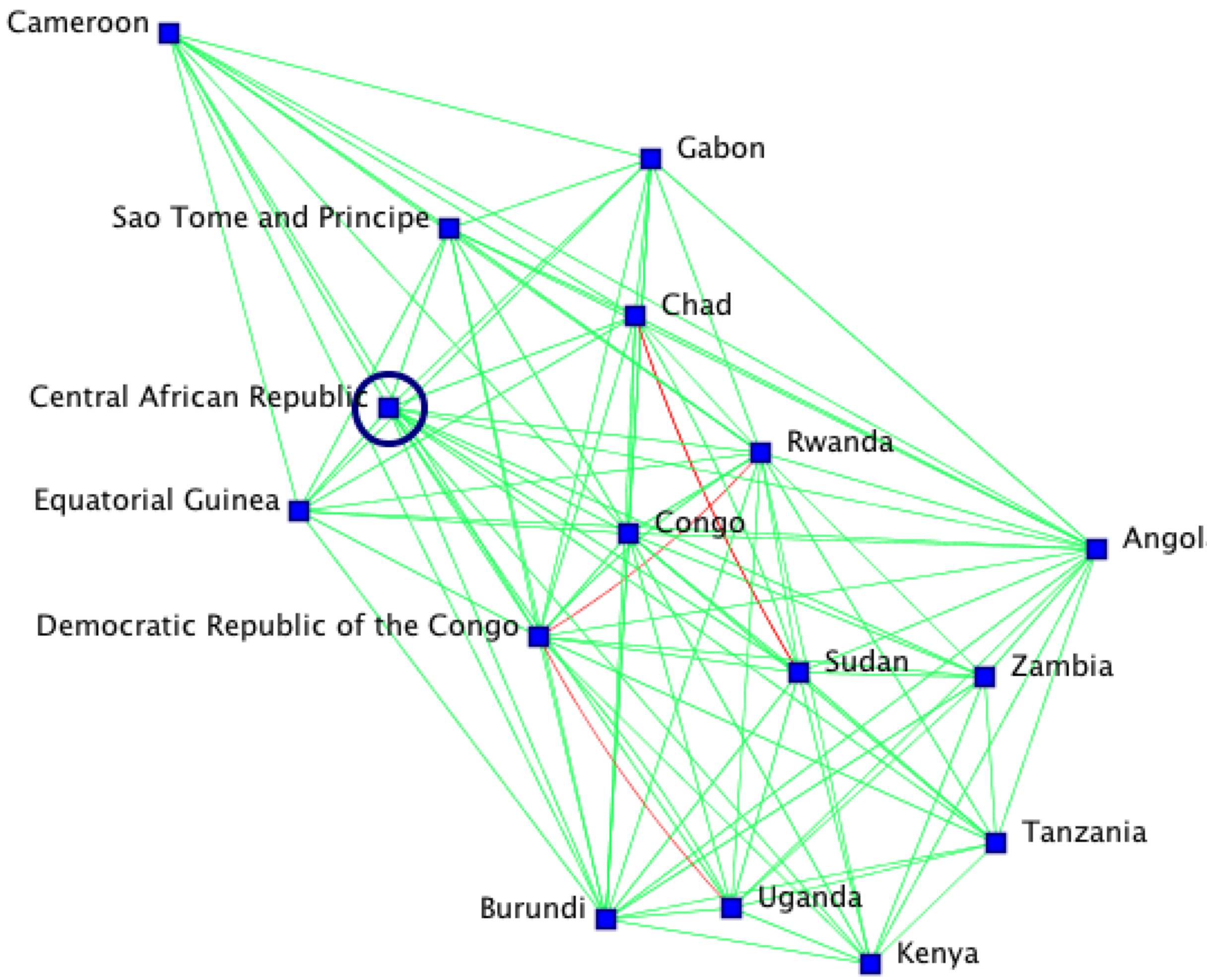


Figure 59. Egonet of CAR

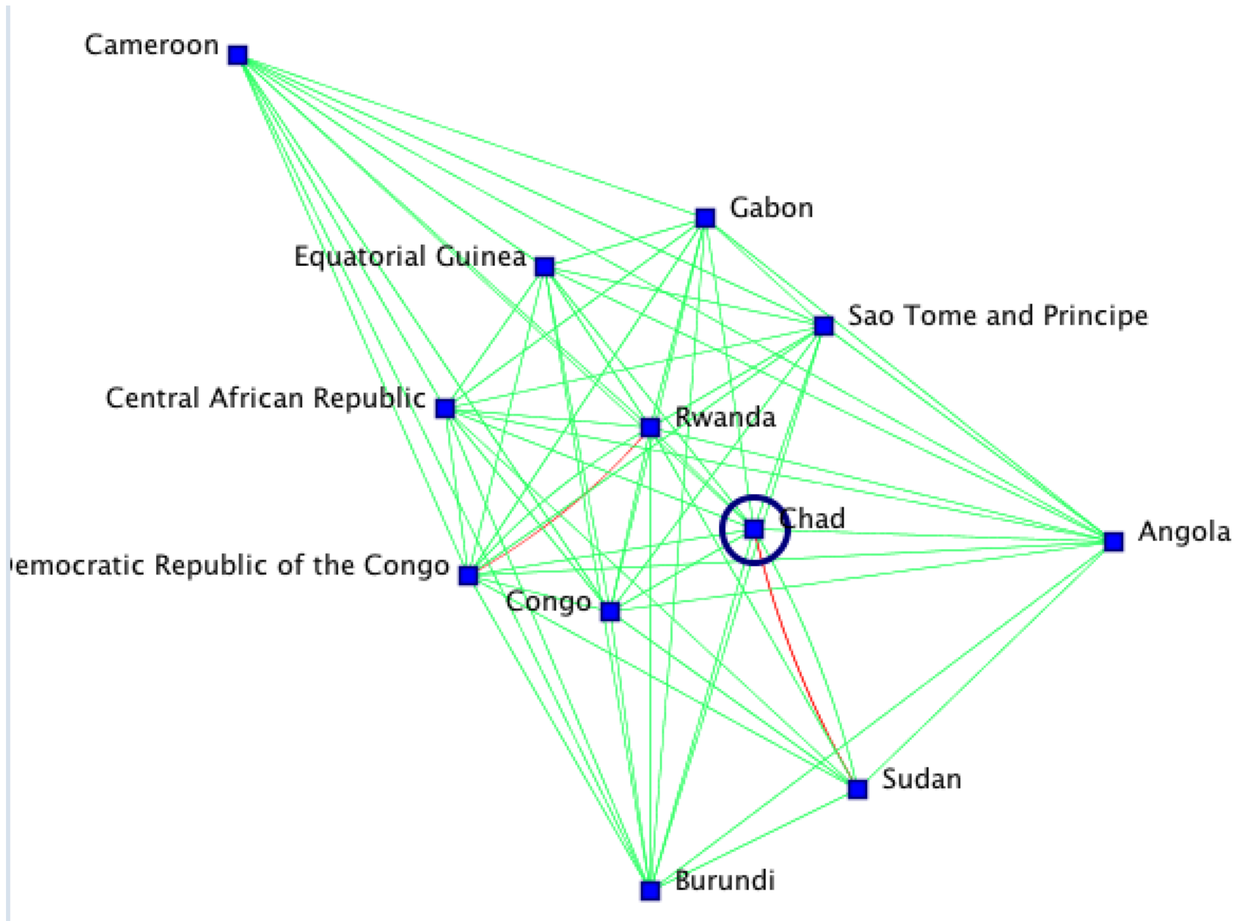


Figure 60. Egonet of Chad

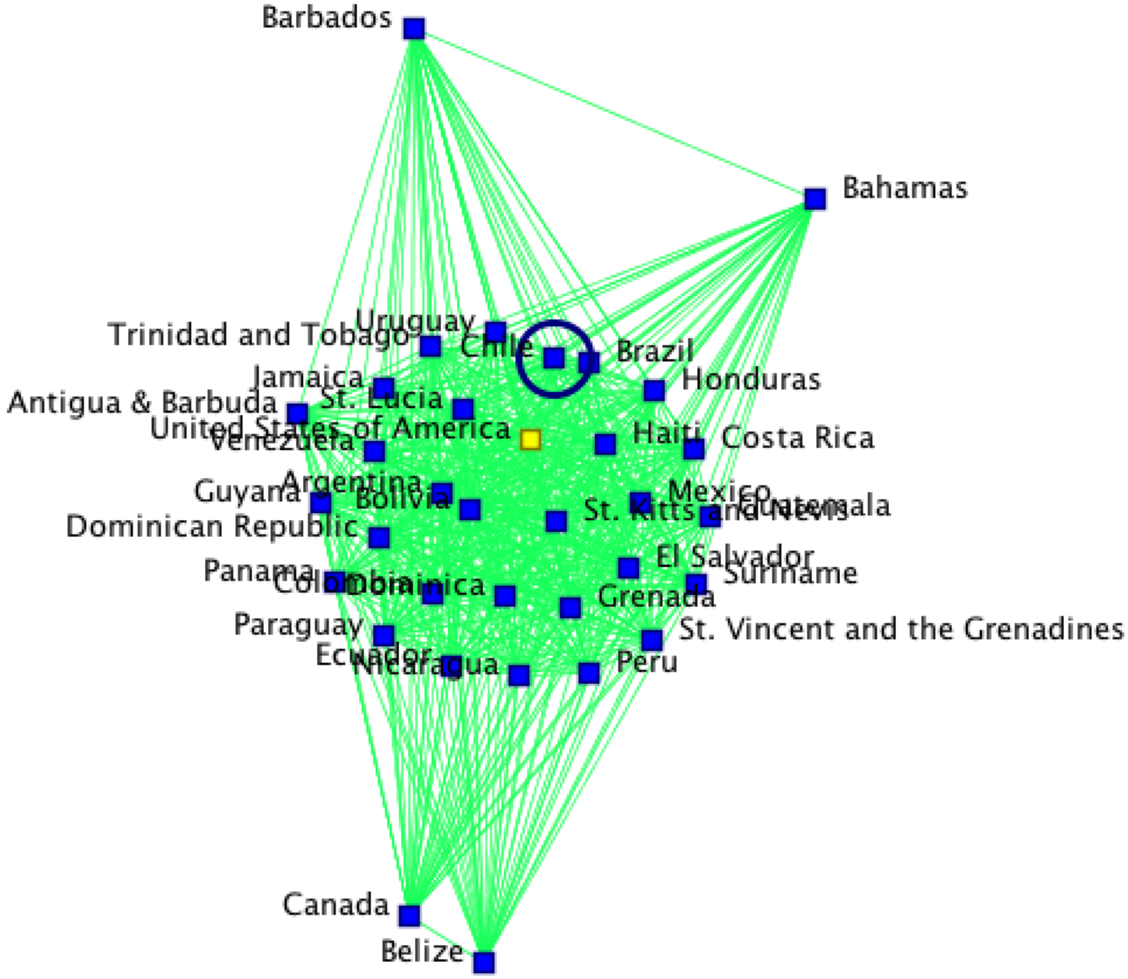


Figure 61. Egonet of Chile

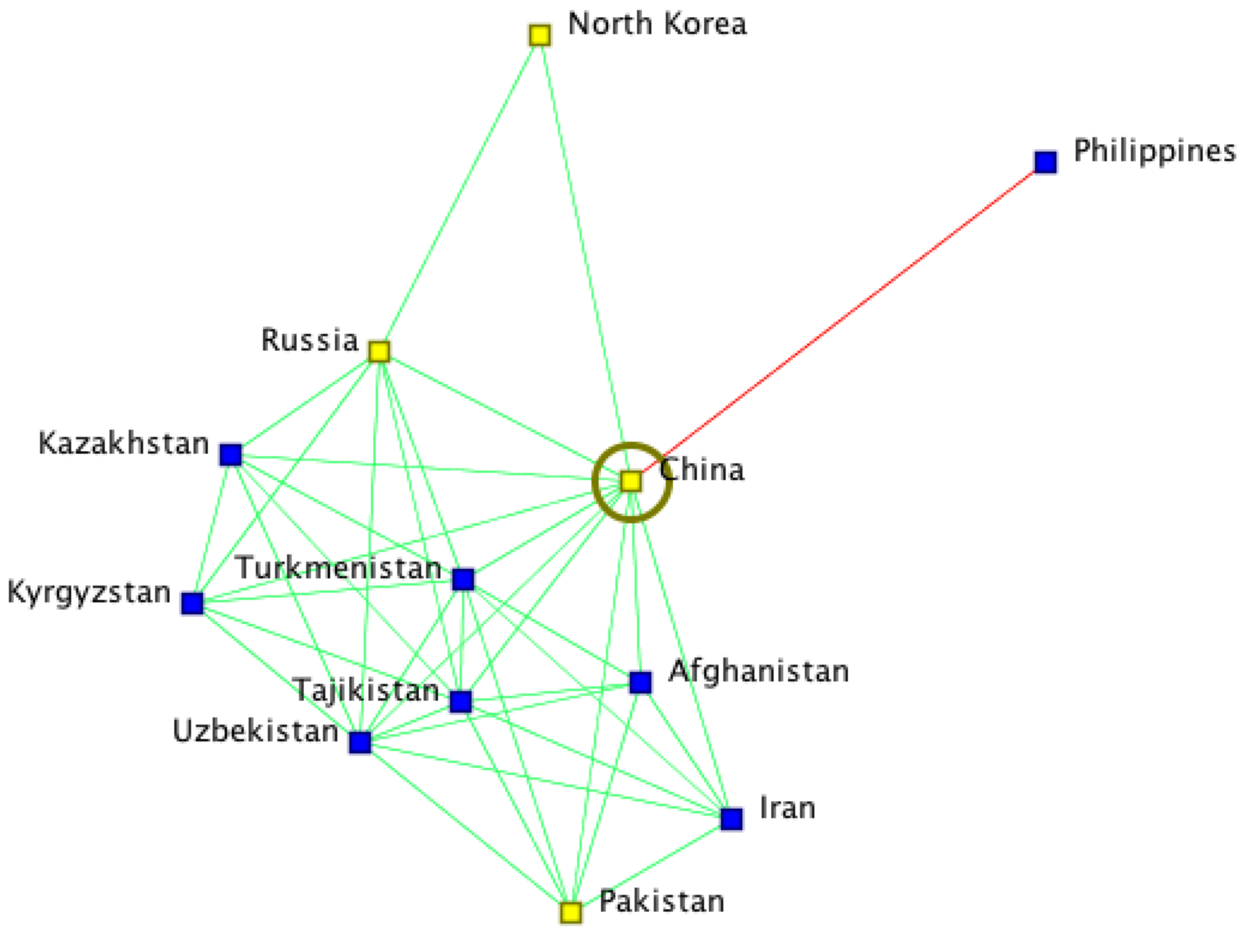


Figure 62. Egonet of China

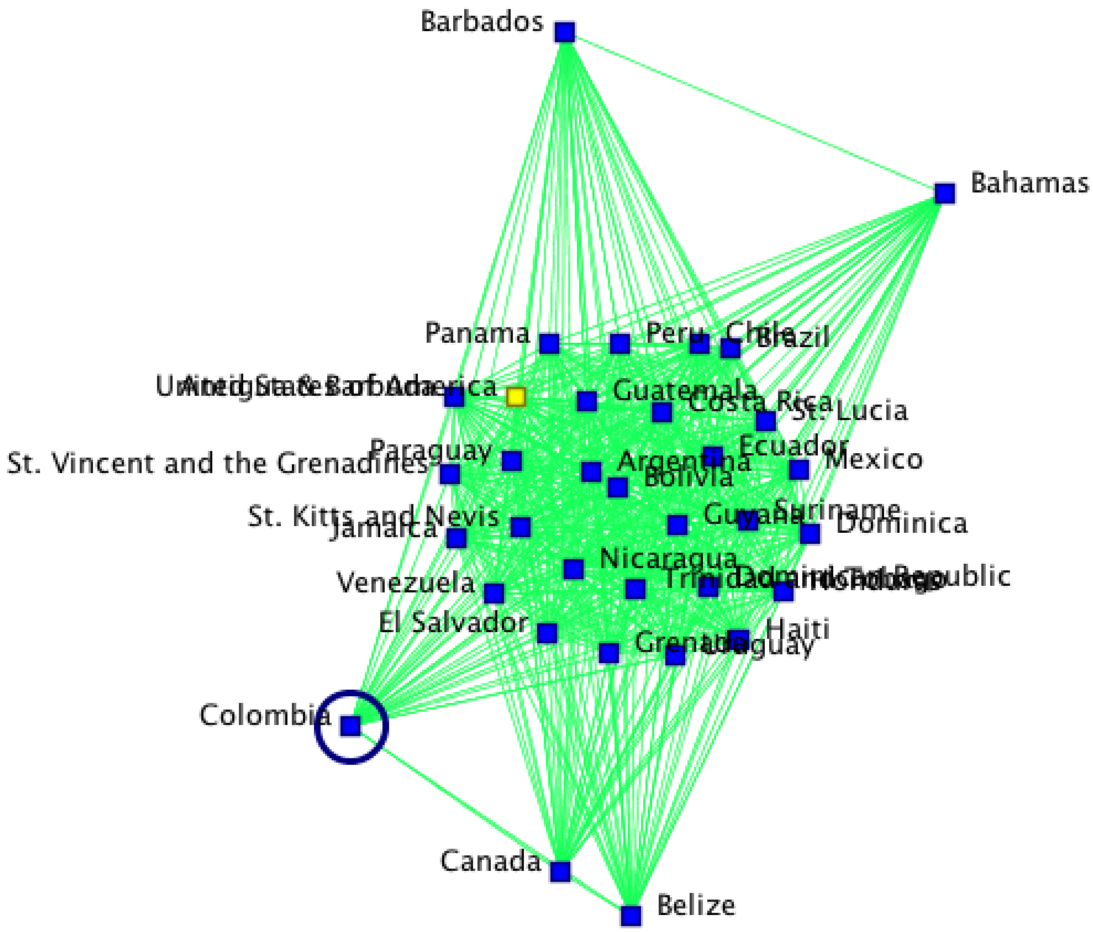


Figure 63. Egonet of Columbia

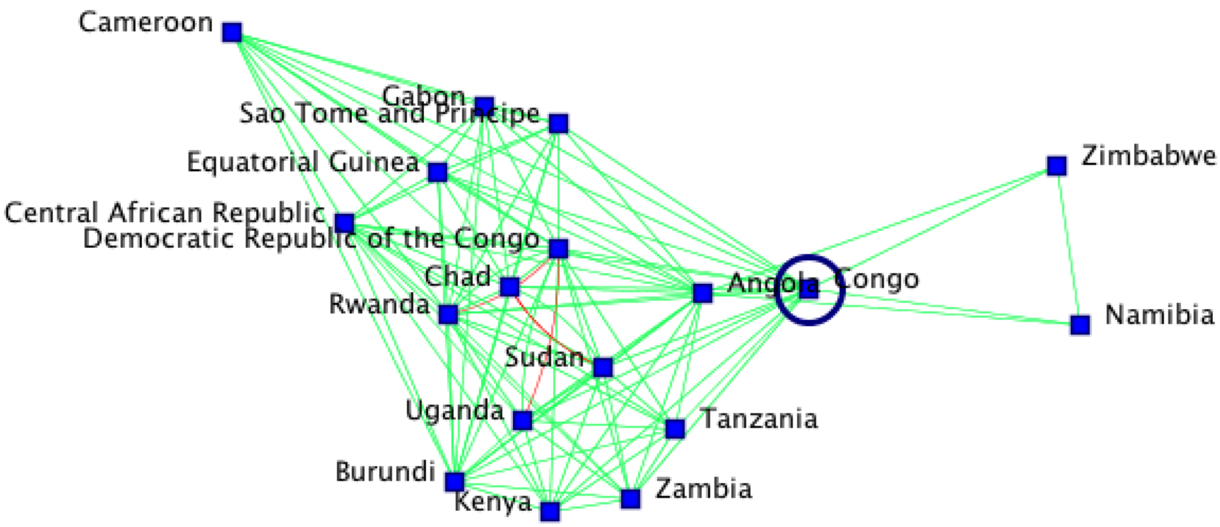


Figure 64. Egonet of Congo

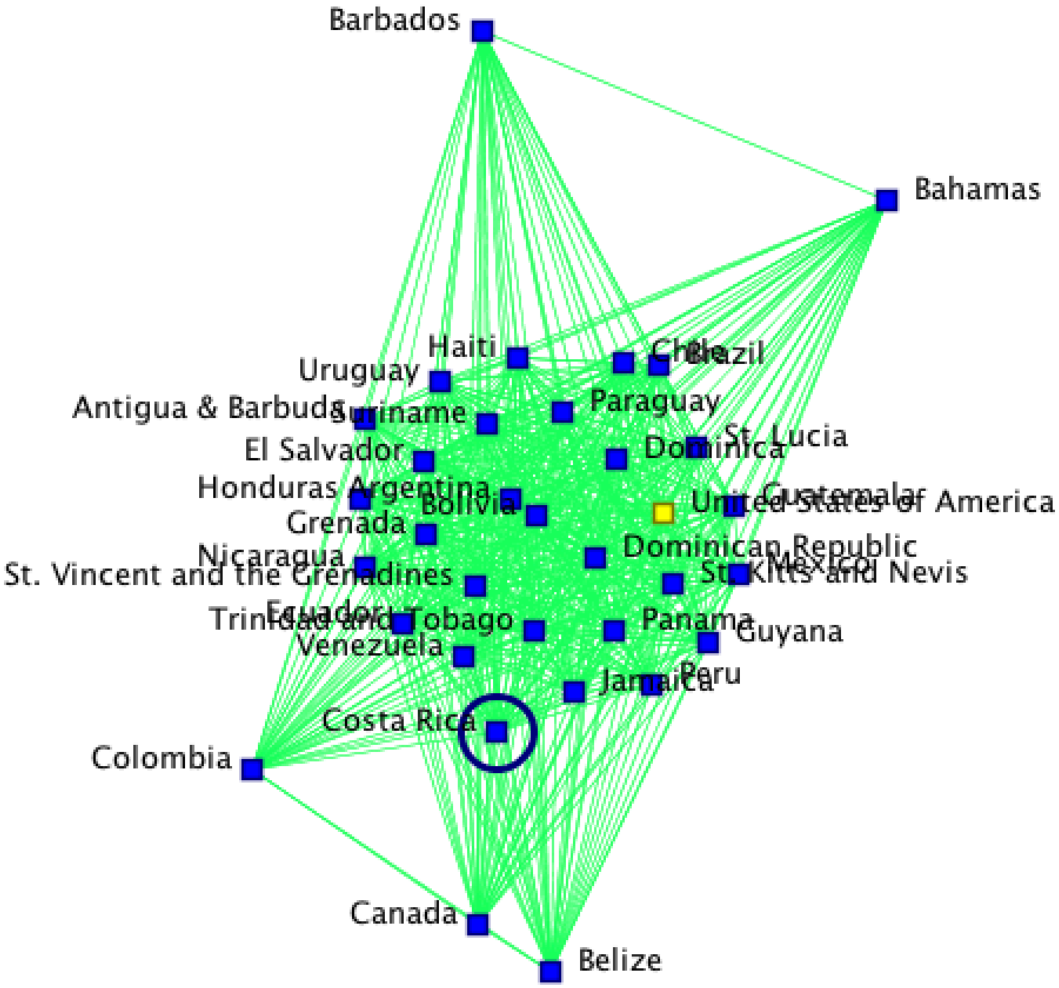


Figure 65. Egonet of Costa Rica

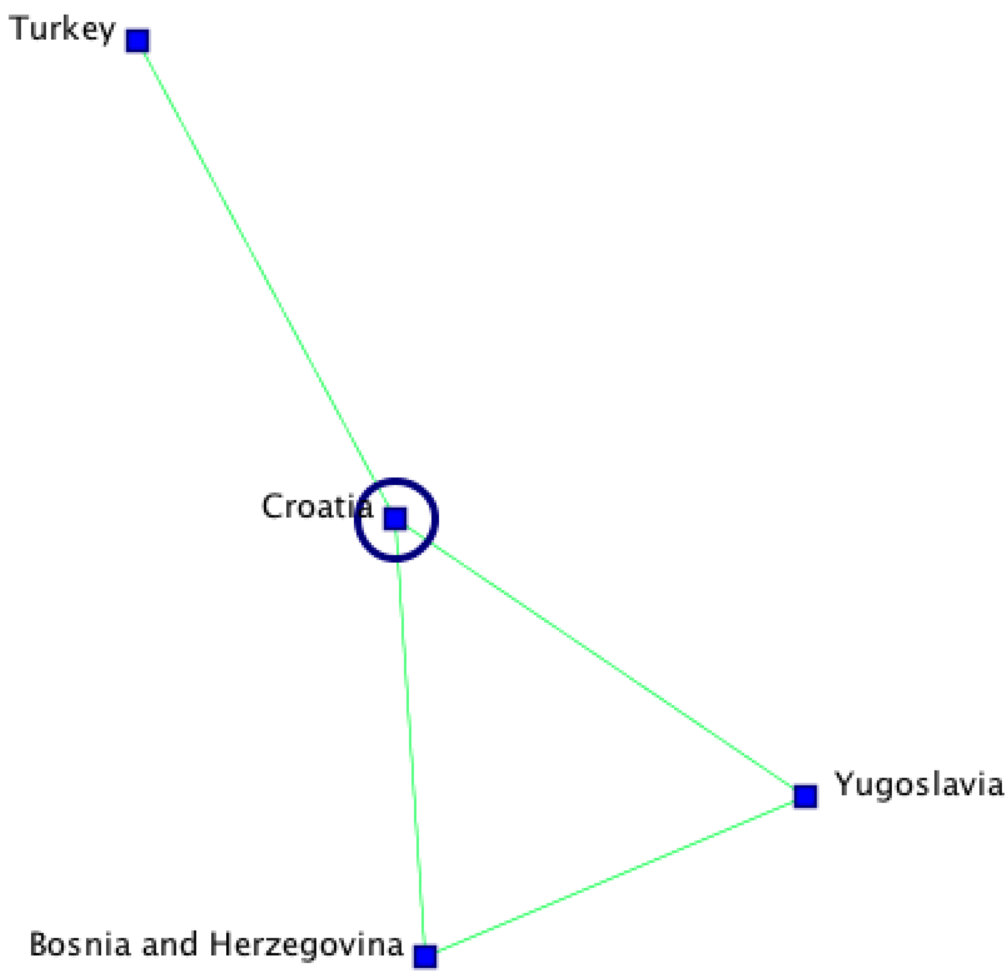


Figure 66. Egonet of Croatia

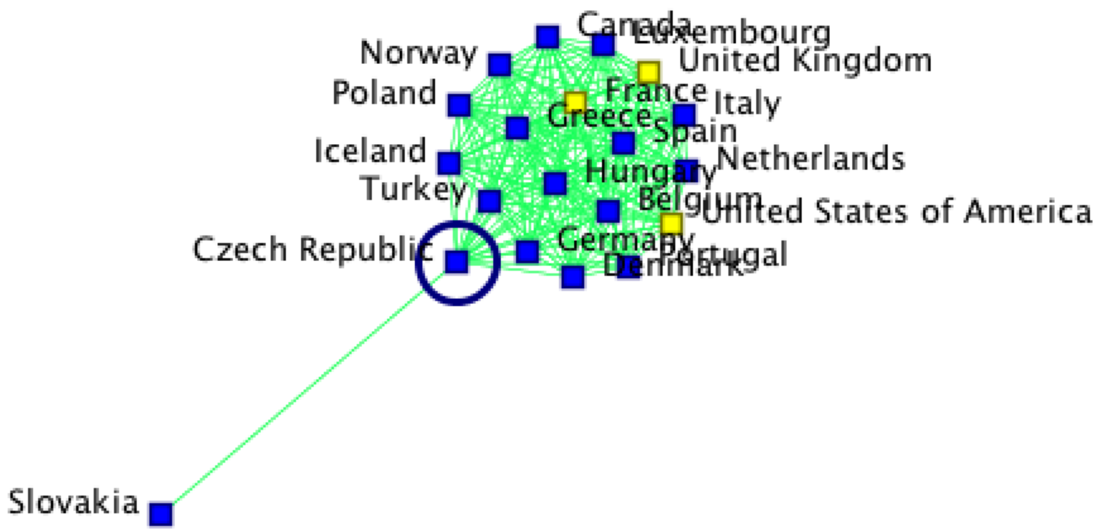


Figure 67. Egonet of Czech Republic

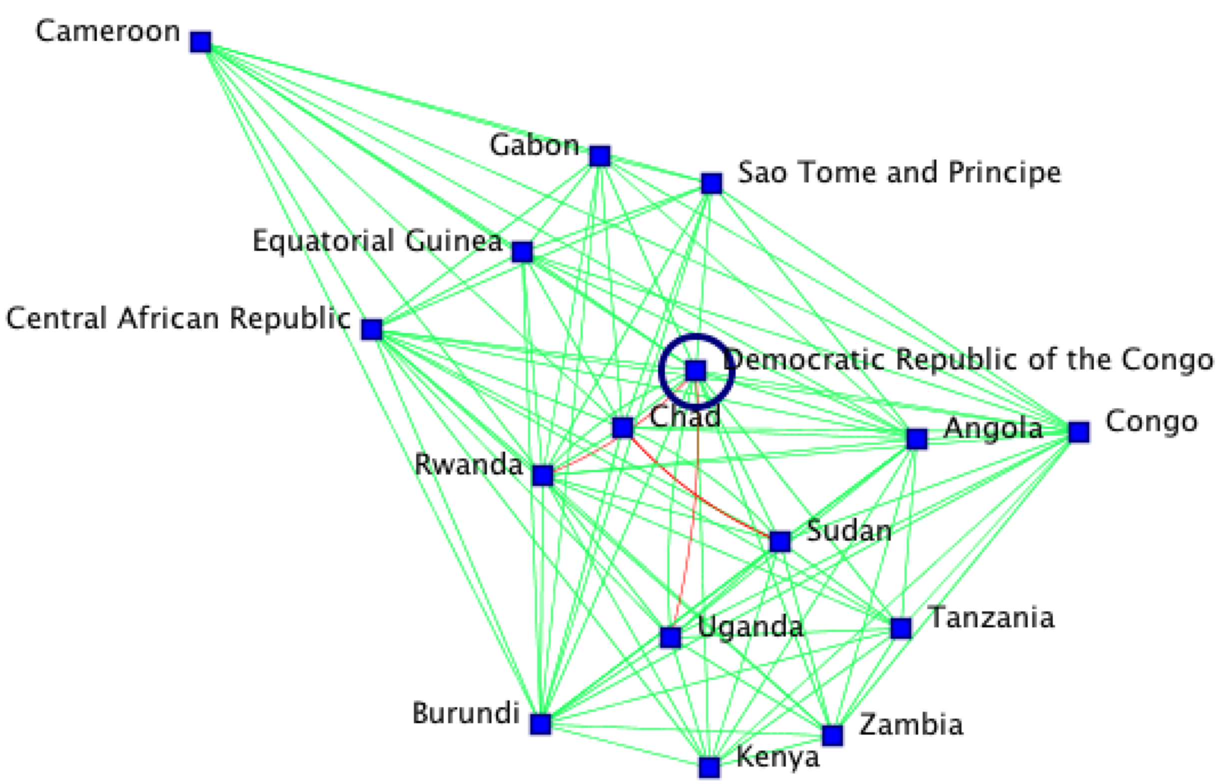


Figure 68. Egonet of Democratic Republic of the Congo

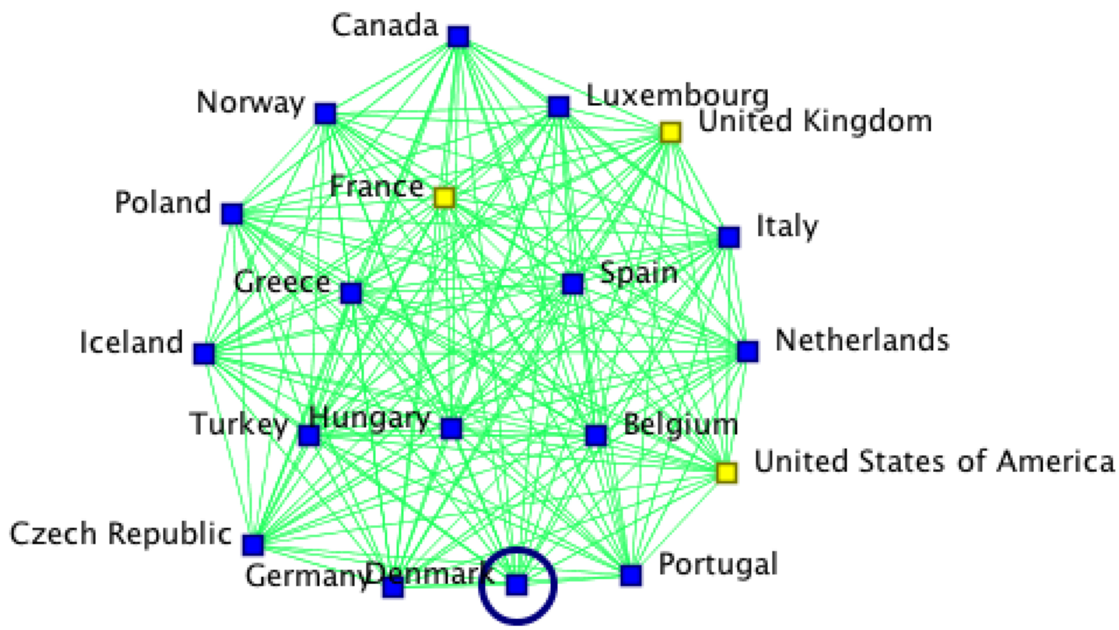


Figure 69. Egonet of Denmark

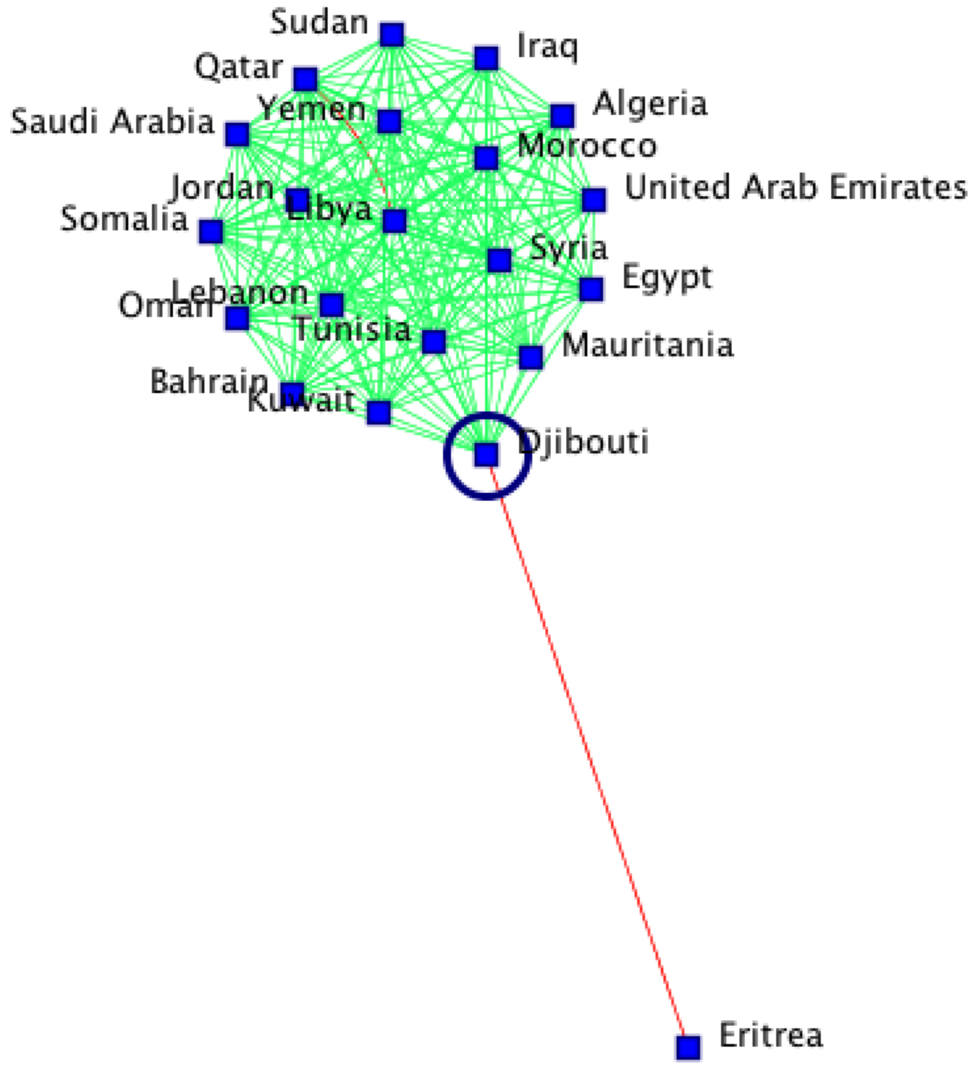


Figure 70. Egonet of Djibouti

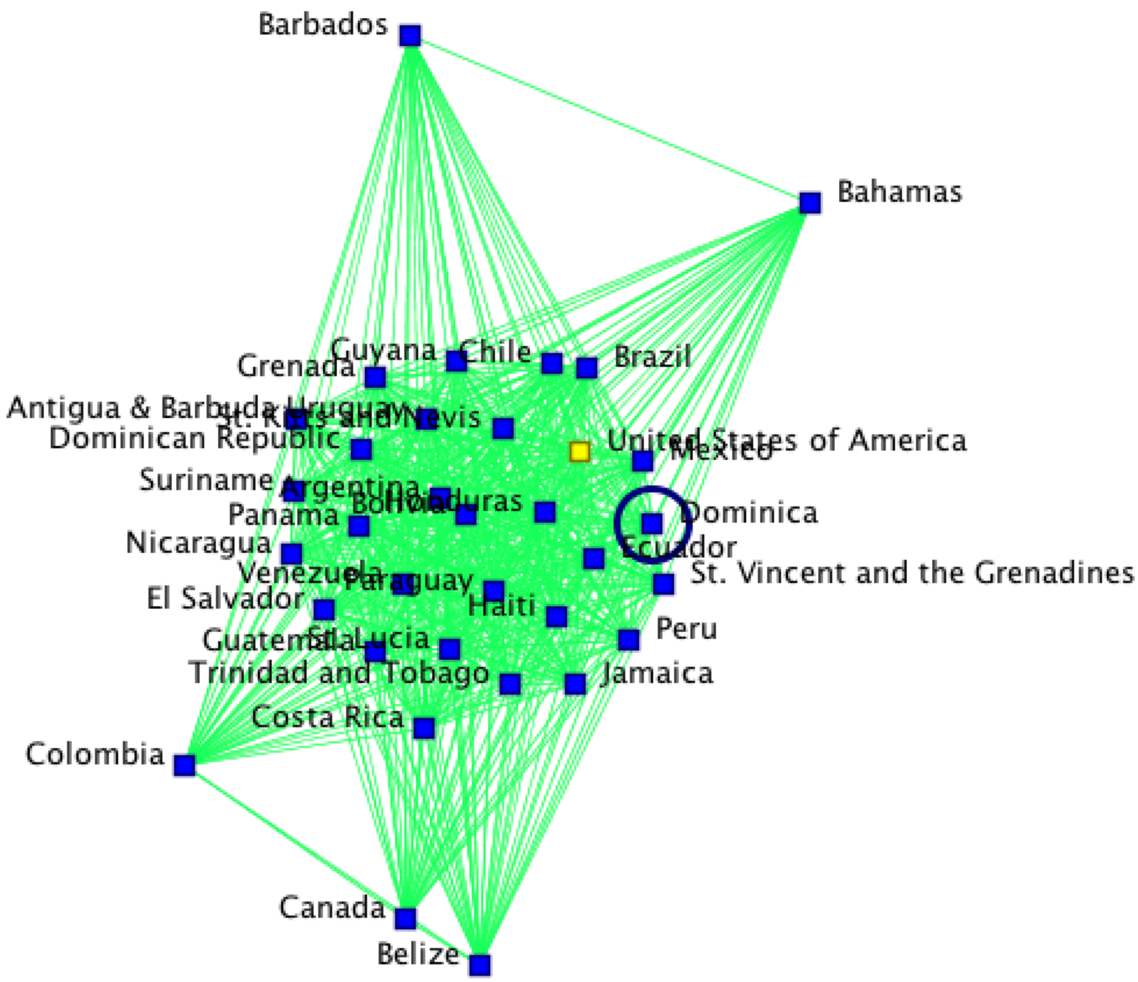


Figure 71. Egonet of Dominica

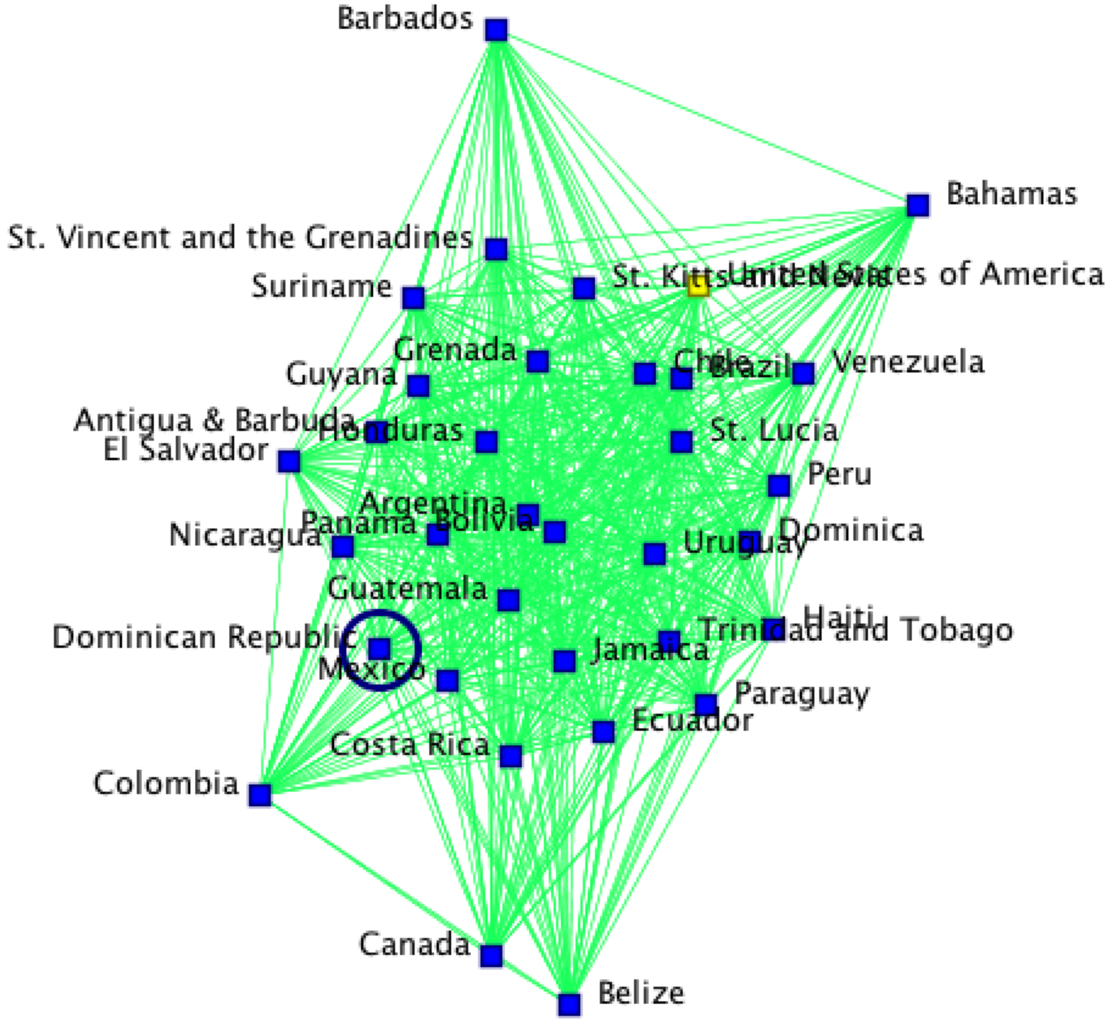


Figure 72. Egonet of Dominican Republic

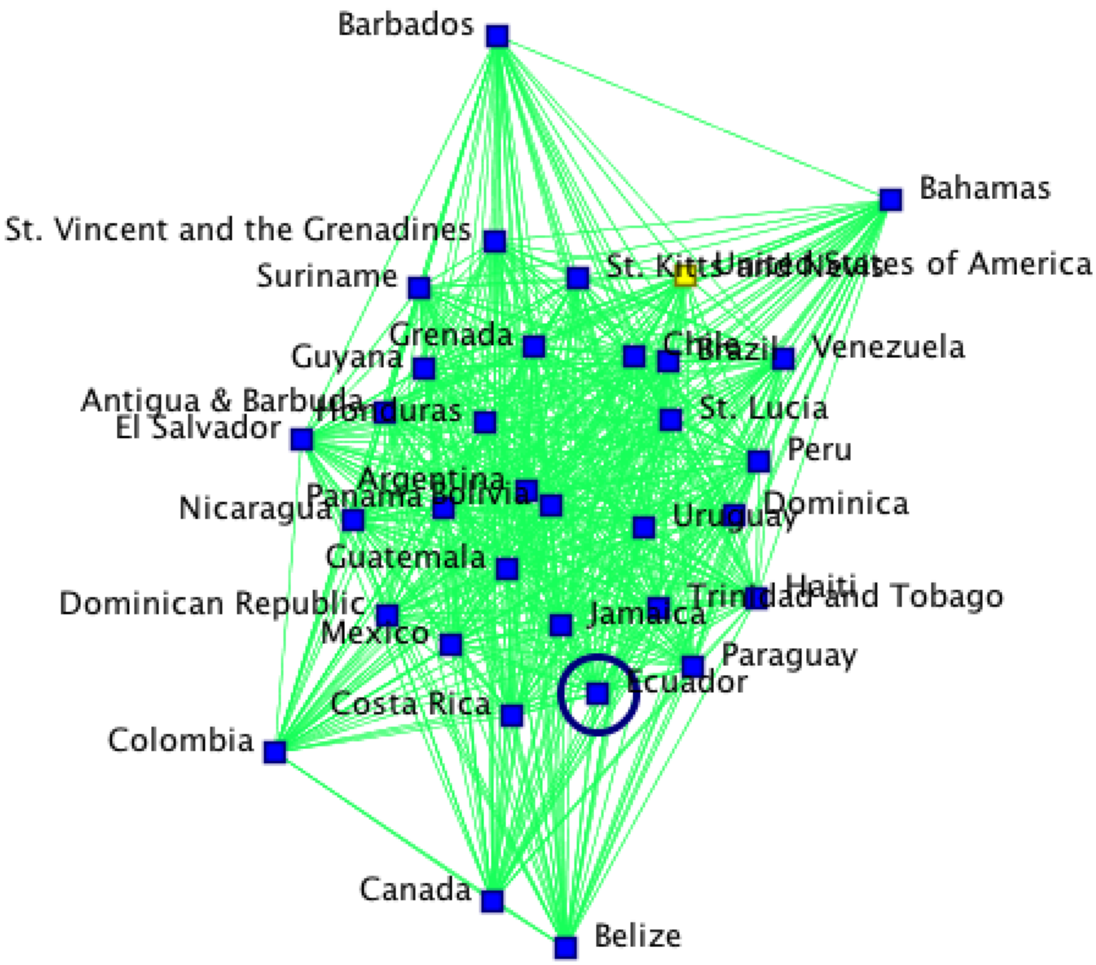


Figure 73. Egonet of Ecuador

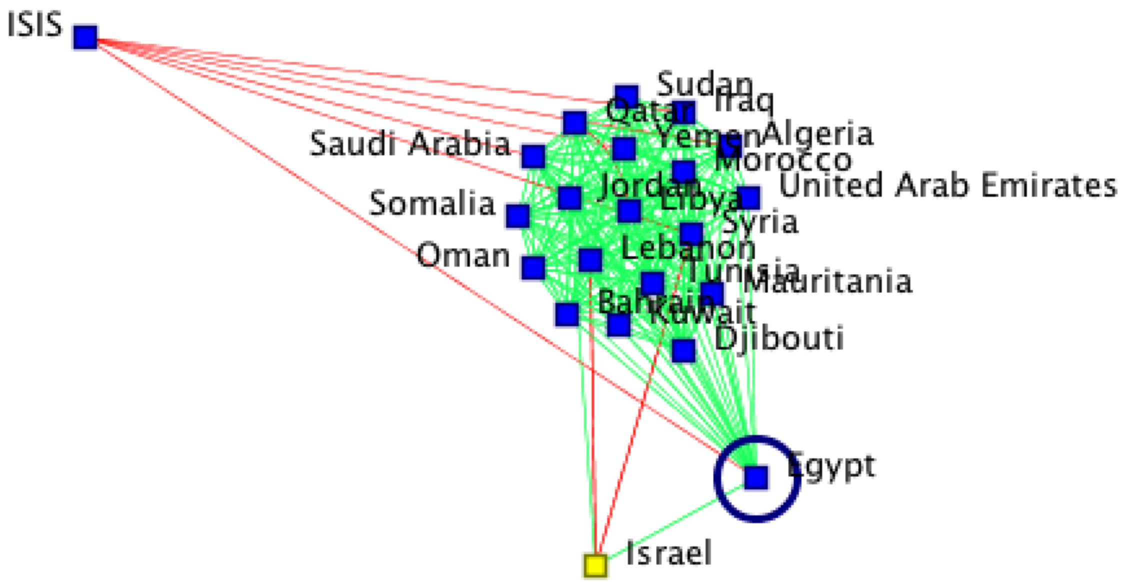


Figure 74. Egonet of Egypt

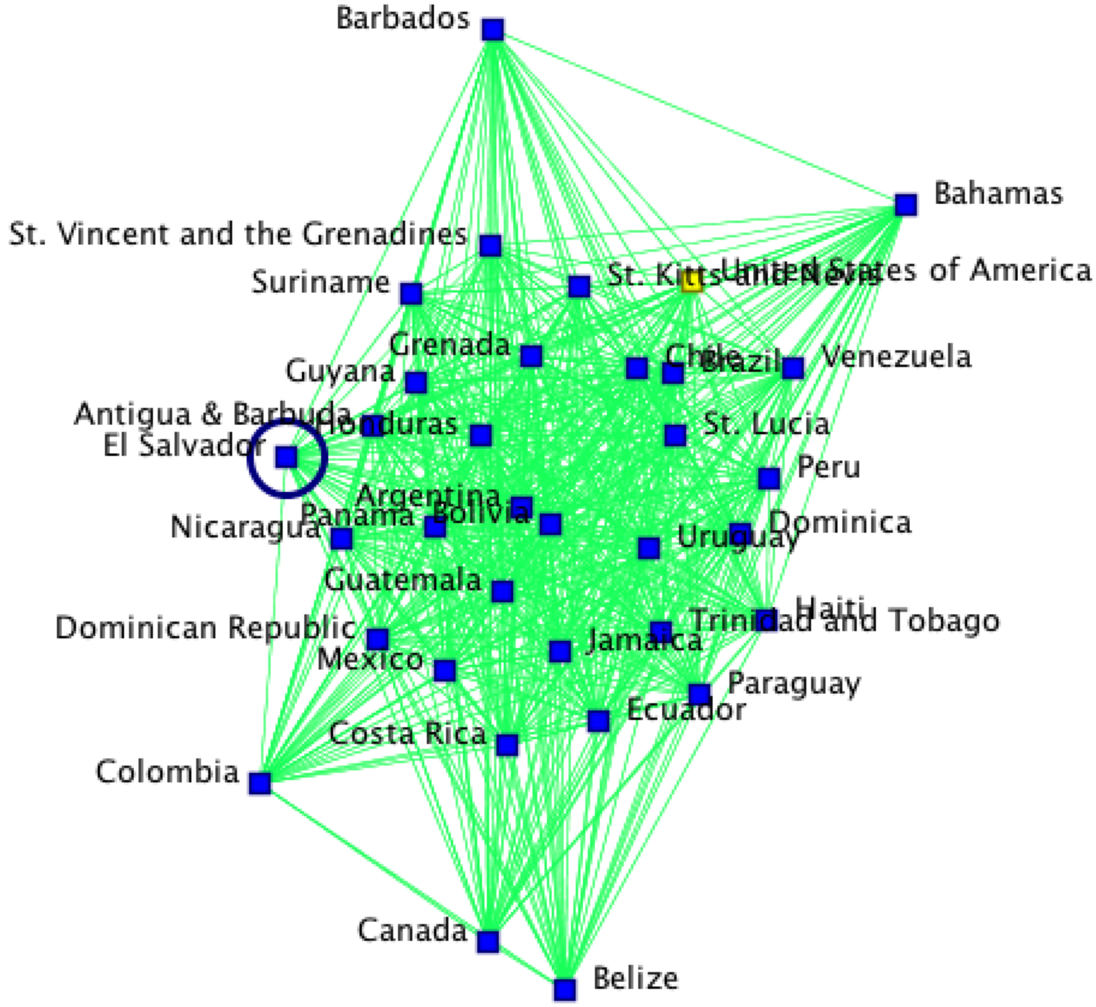


Figure 75. Egonet of El Salvador

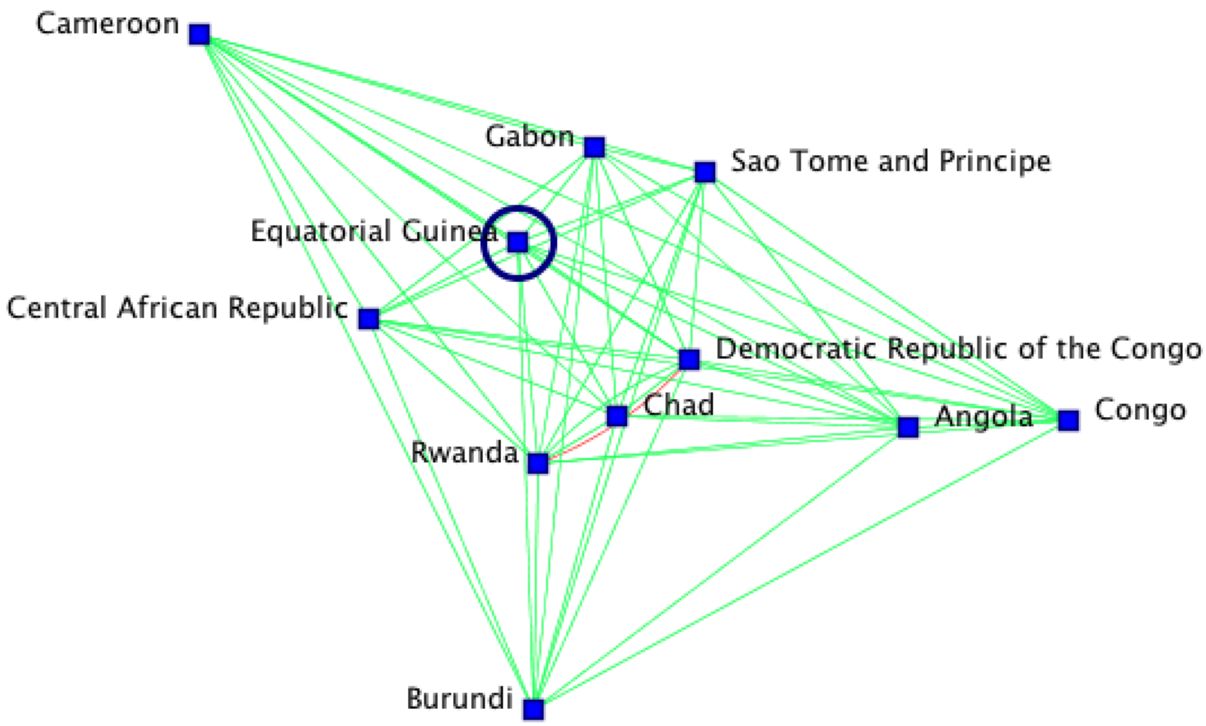


Figure 76. Egonet of Equatorial Guinea

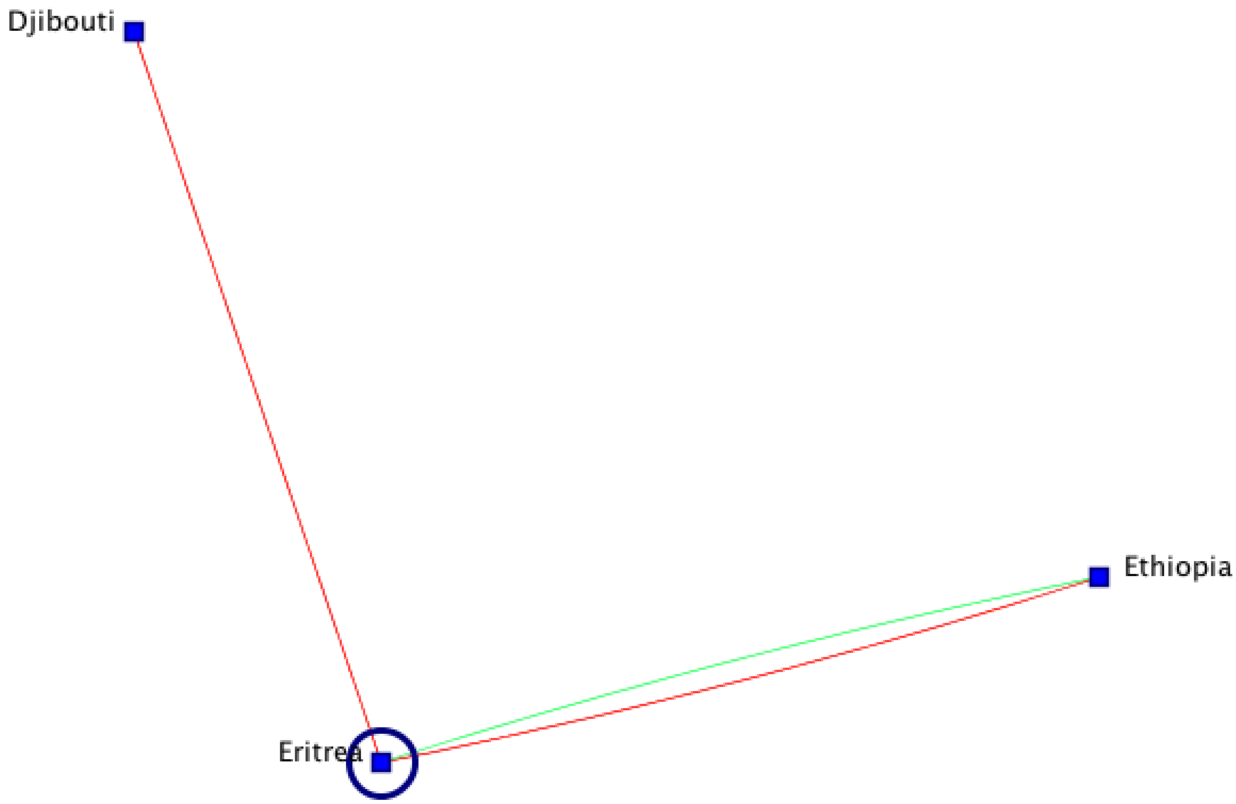


Figure 77. Egonet of Eritrea

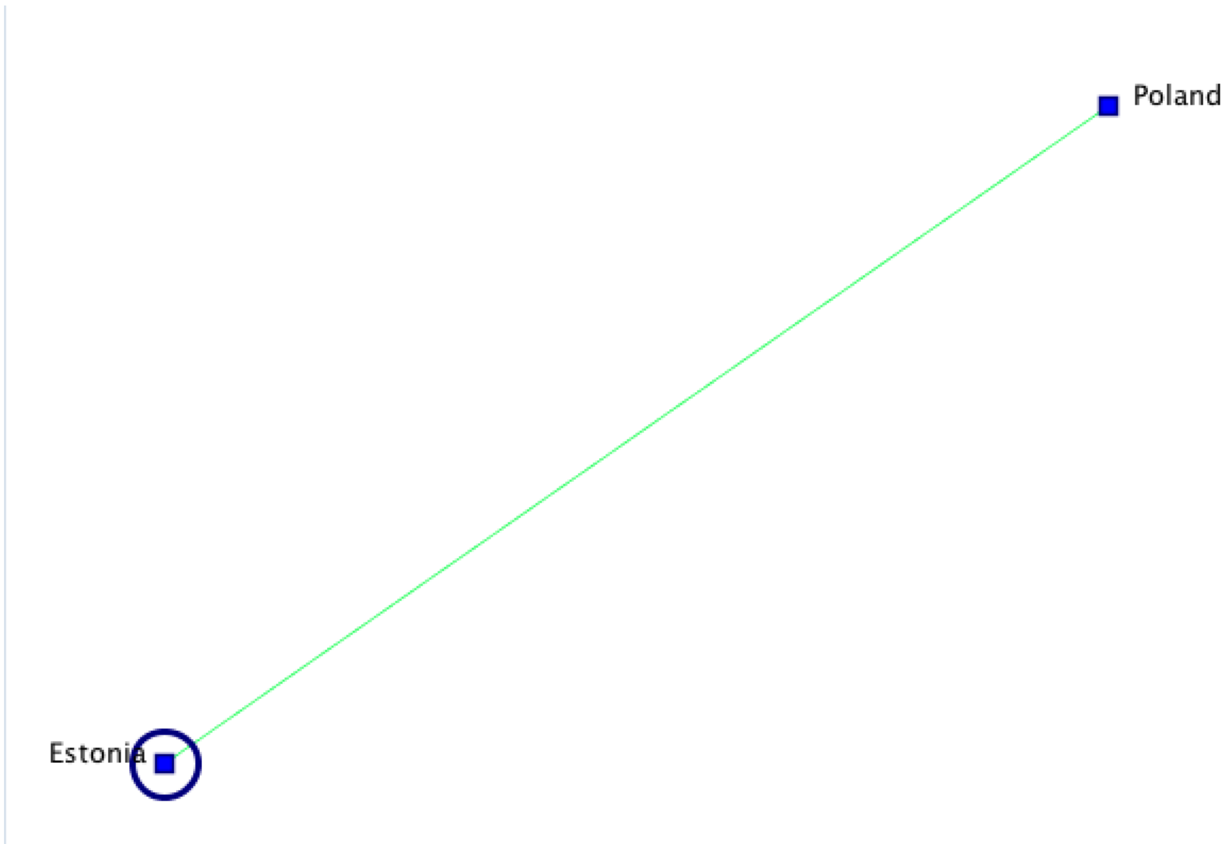


Figure 78. Egonet of Estonia

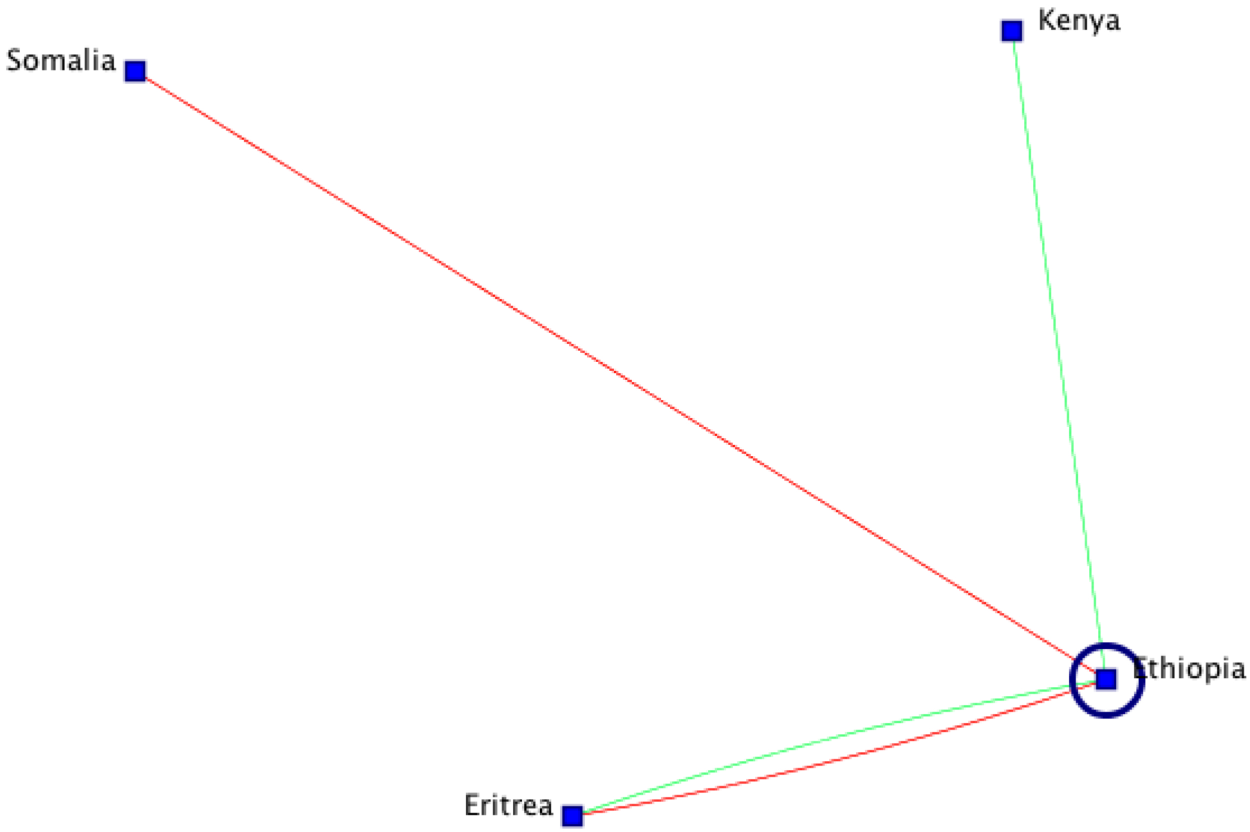


Figure 79. Egonet of Ethiopia

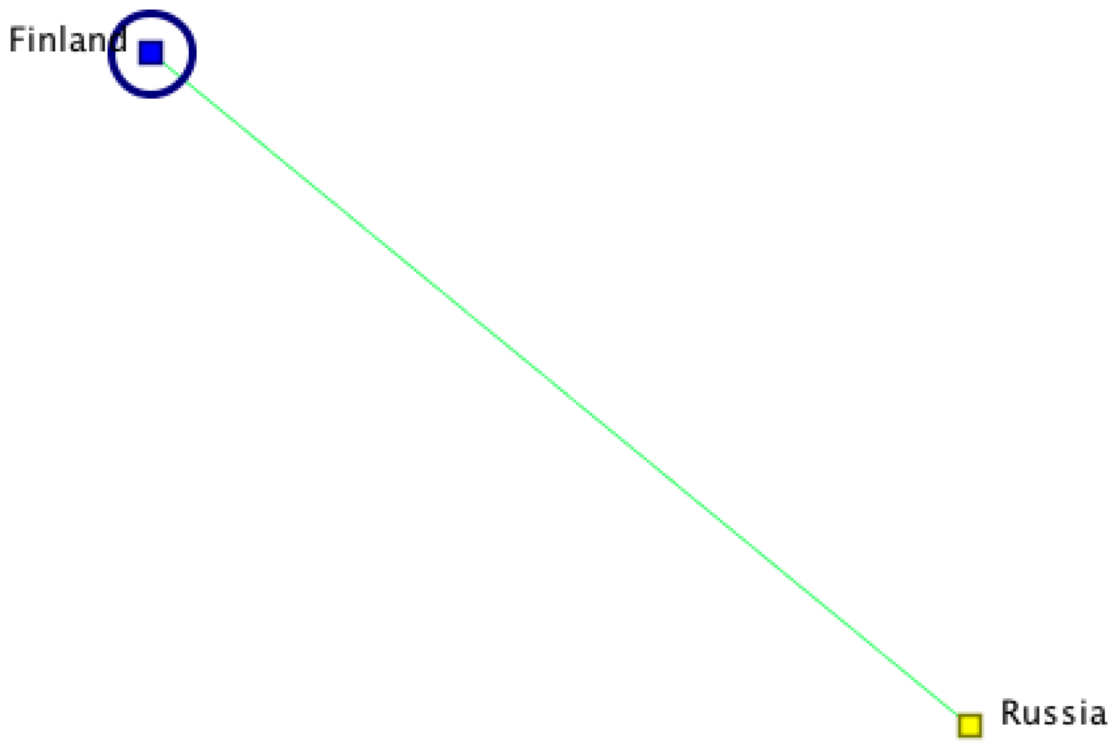


Figure 80. Egonet of Finland

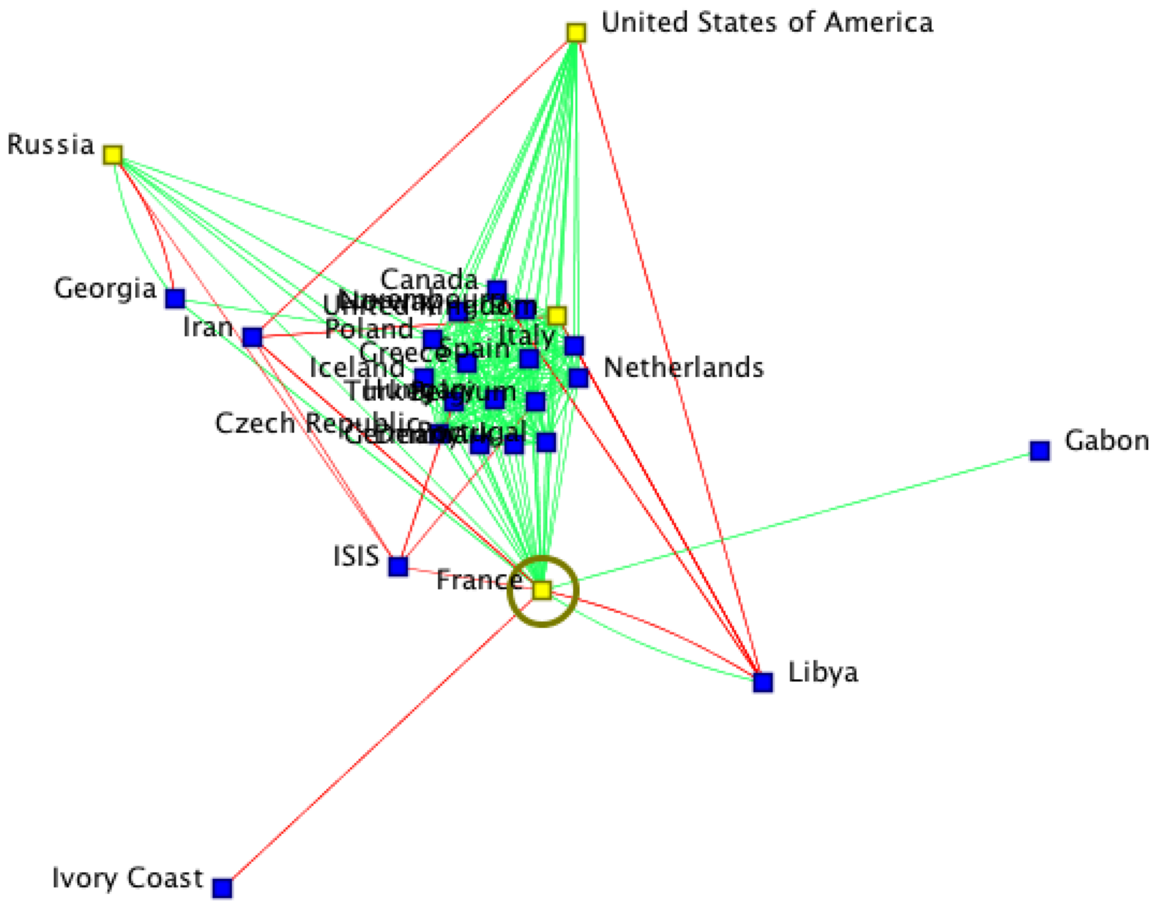


Figure 81. Egonet of France

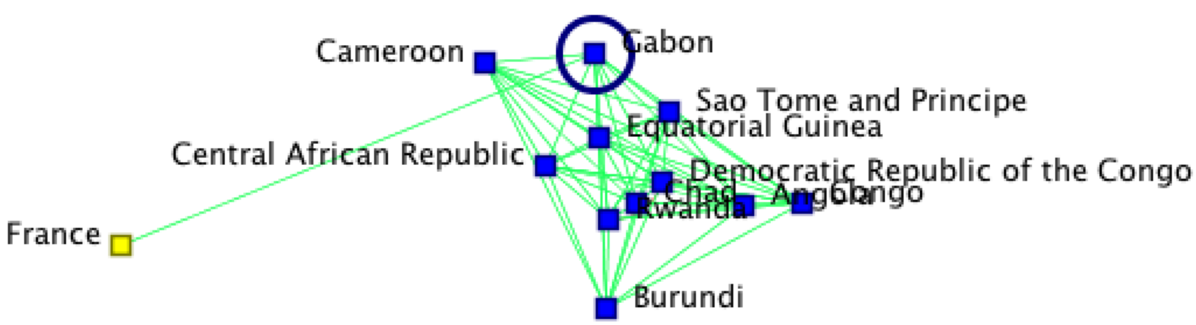


Figure 82. Egonet of Gabon

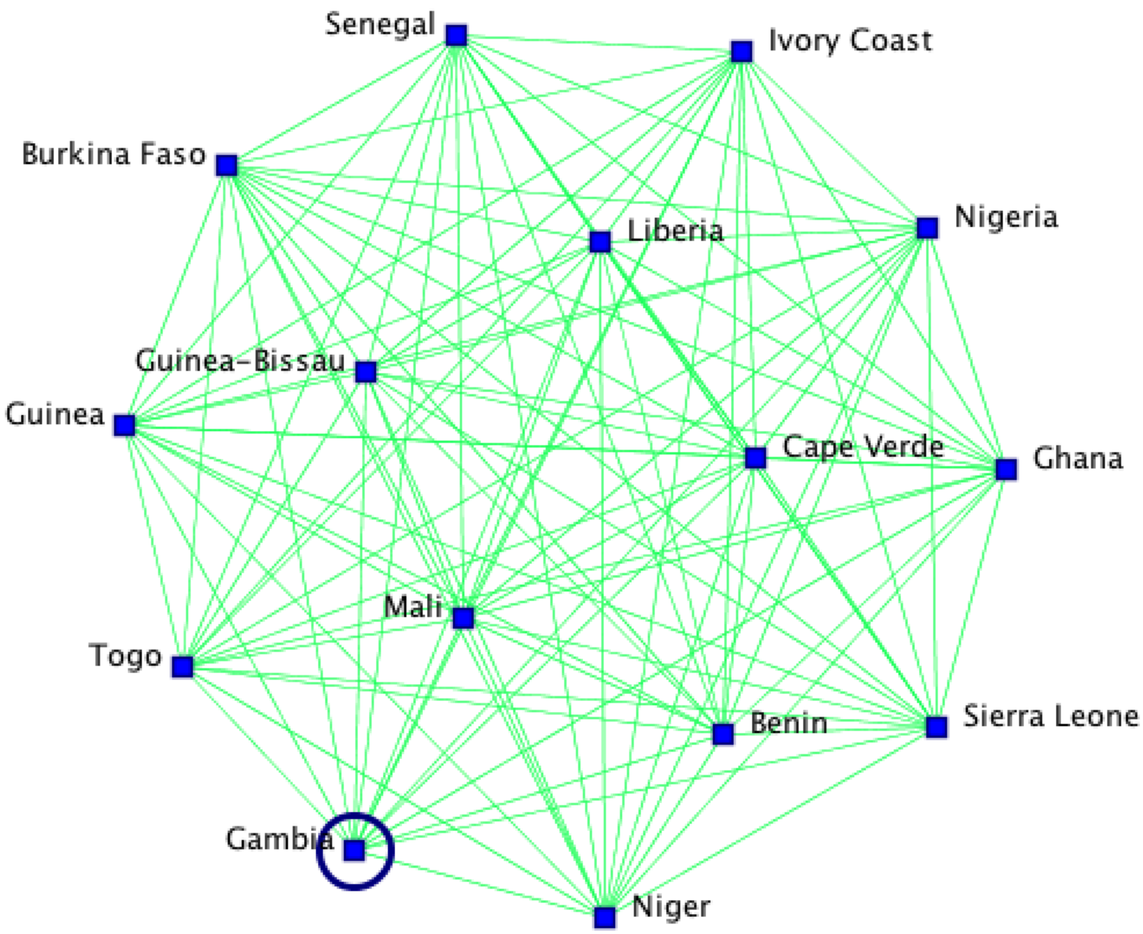


Figure 83. Egonet of Gambia

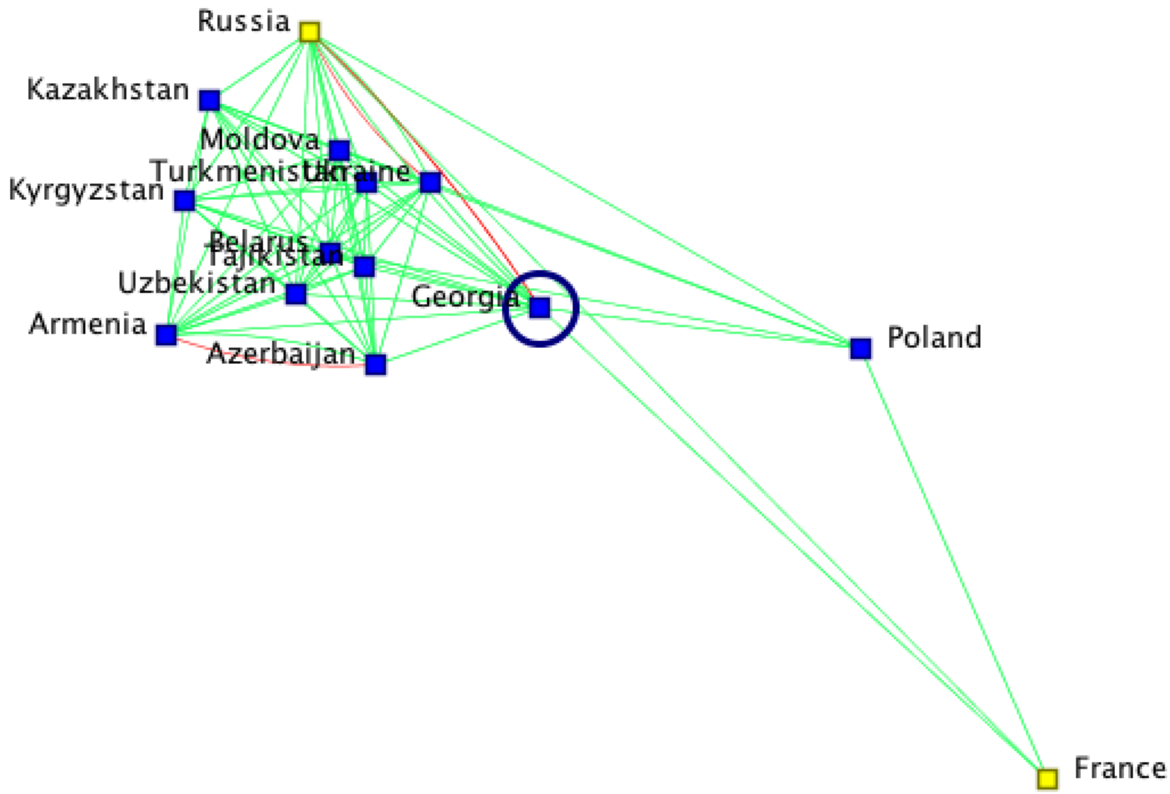


Figure 84. Egonet of Georgia

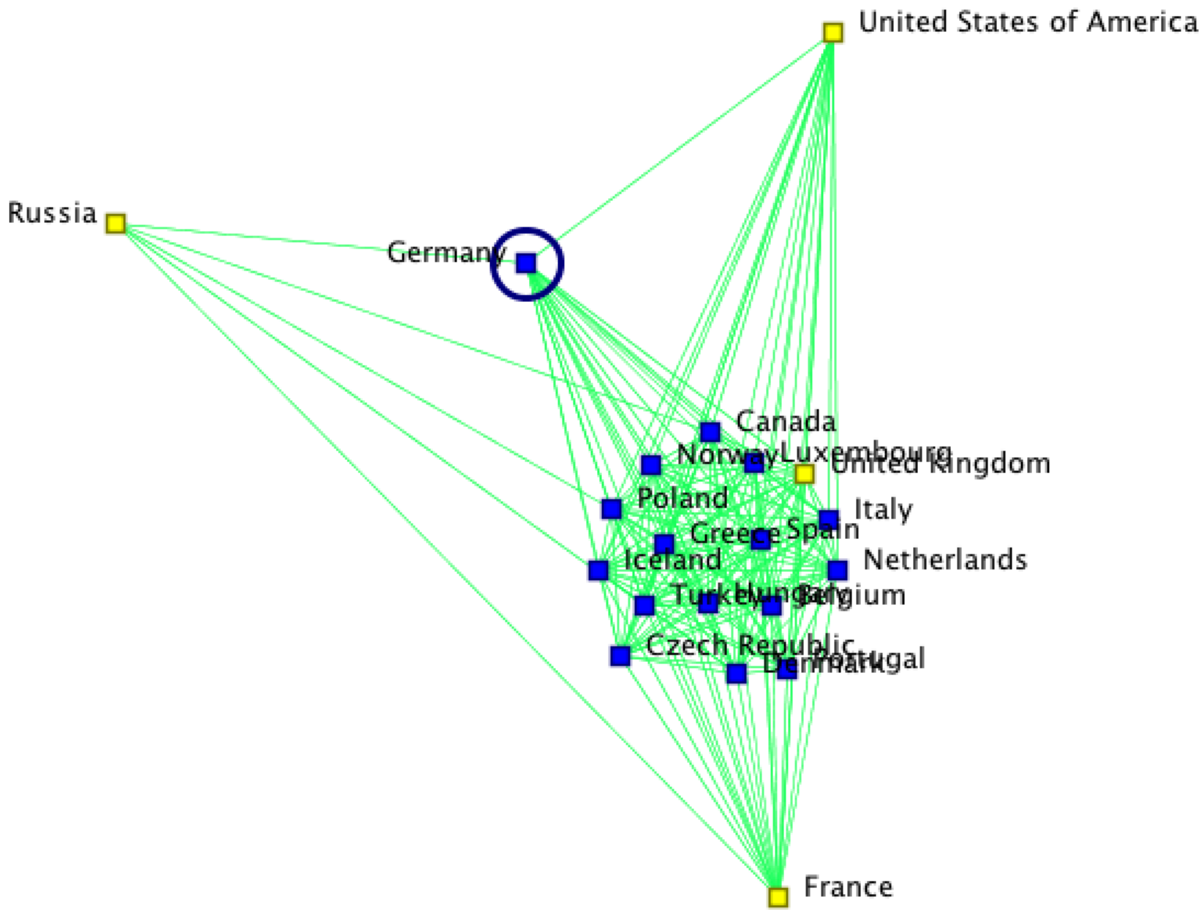


Figure 85. Egonet of Germany

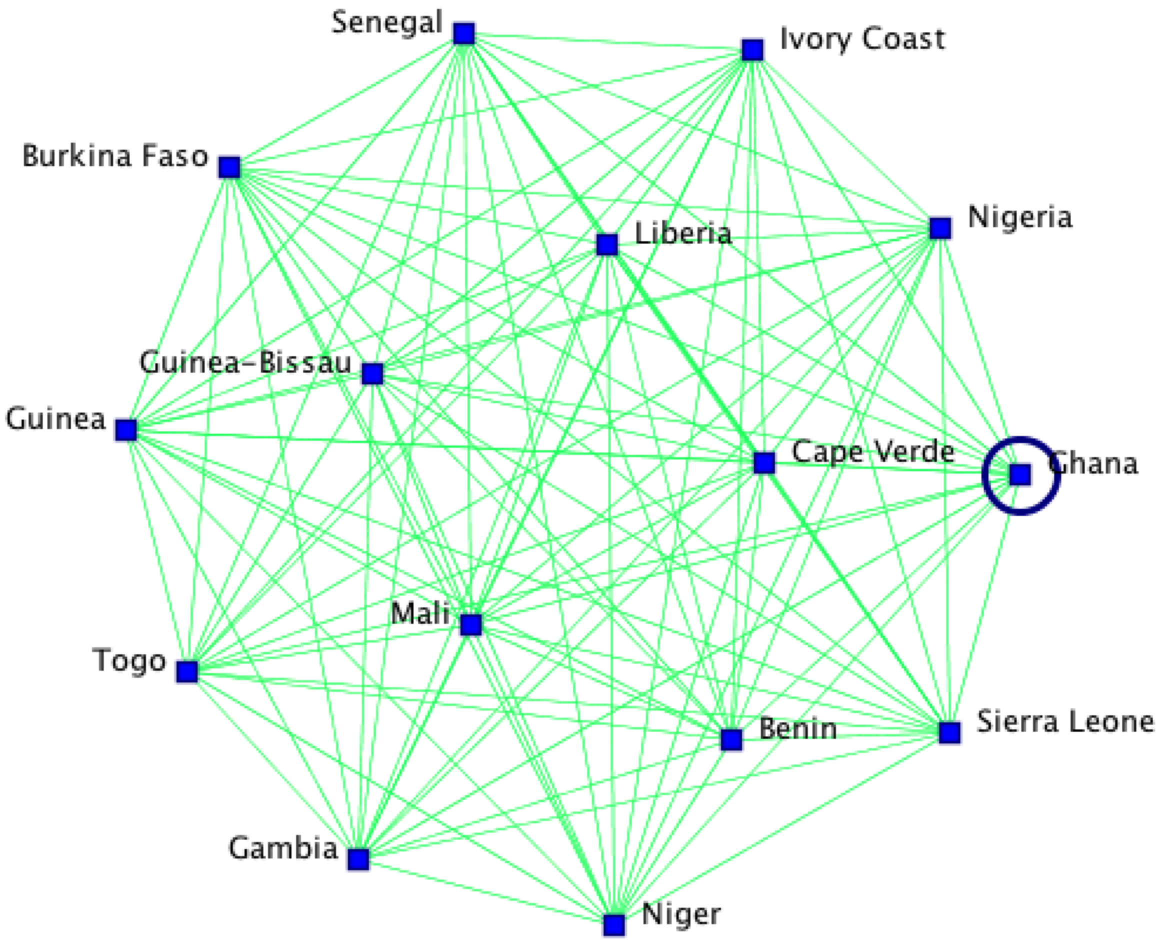


Figure 86. Egonet of Ghana

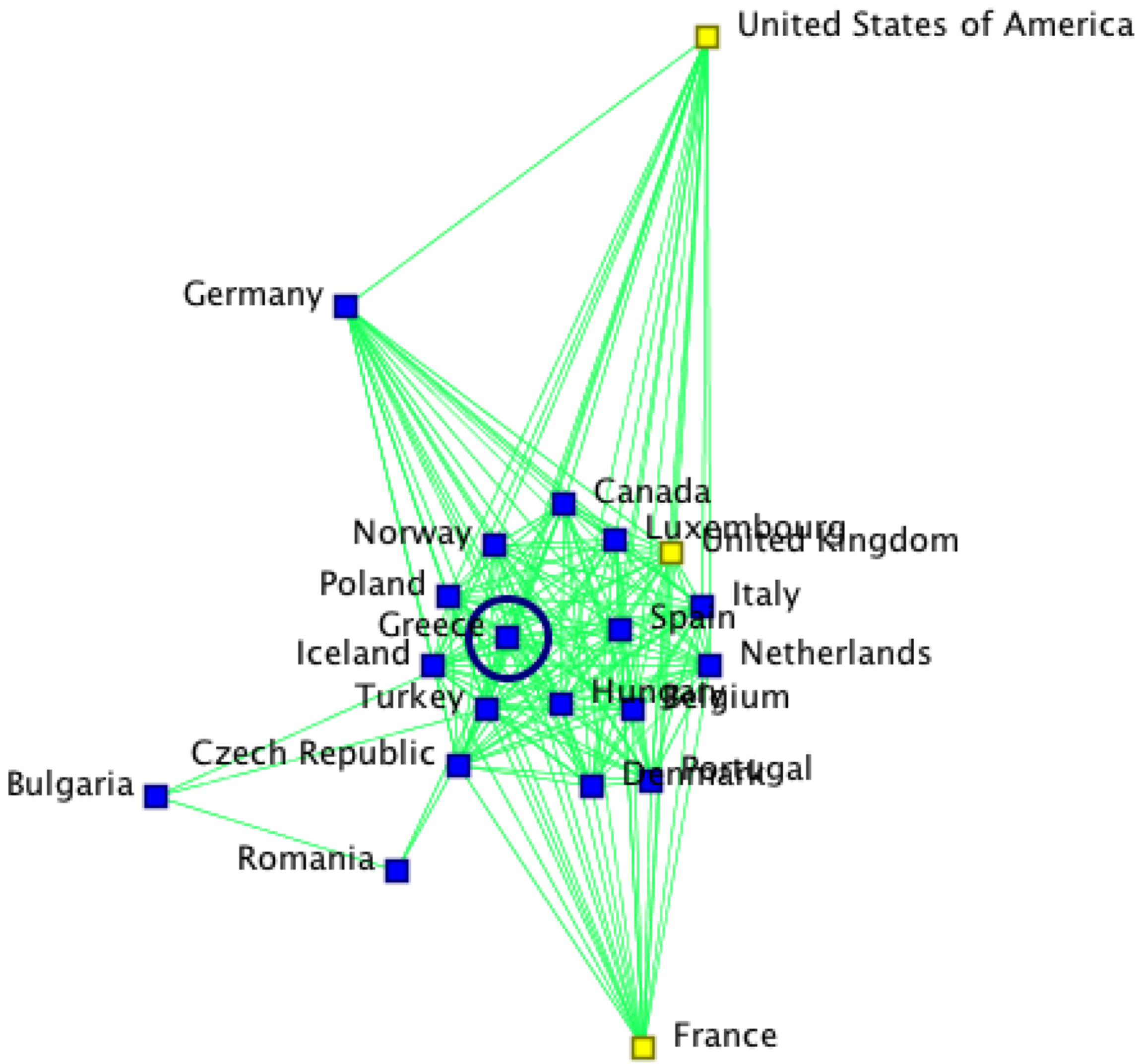


Figure 87. Egonet of Greece

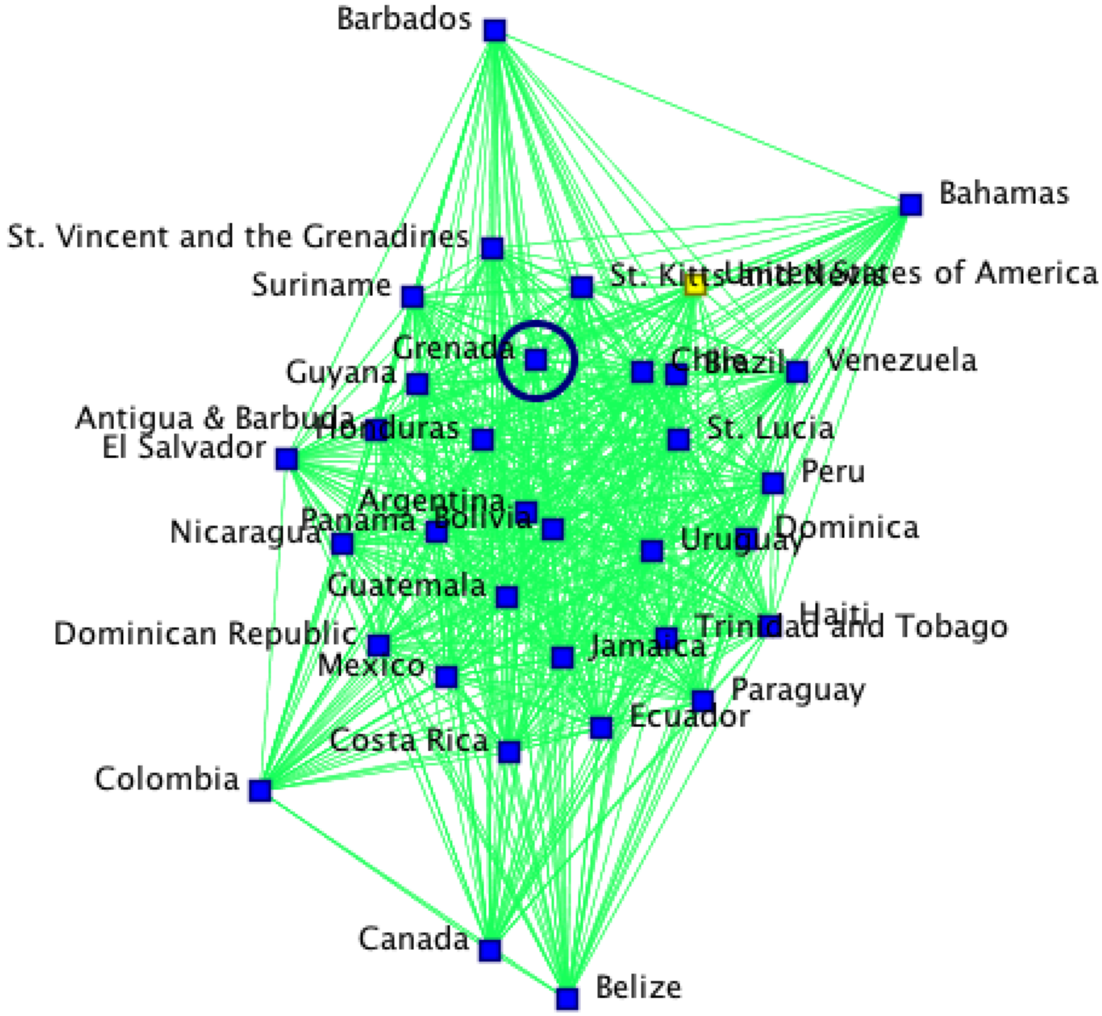


Figure 88. Egonet of Grenada

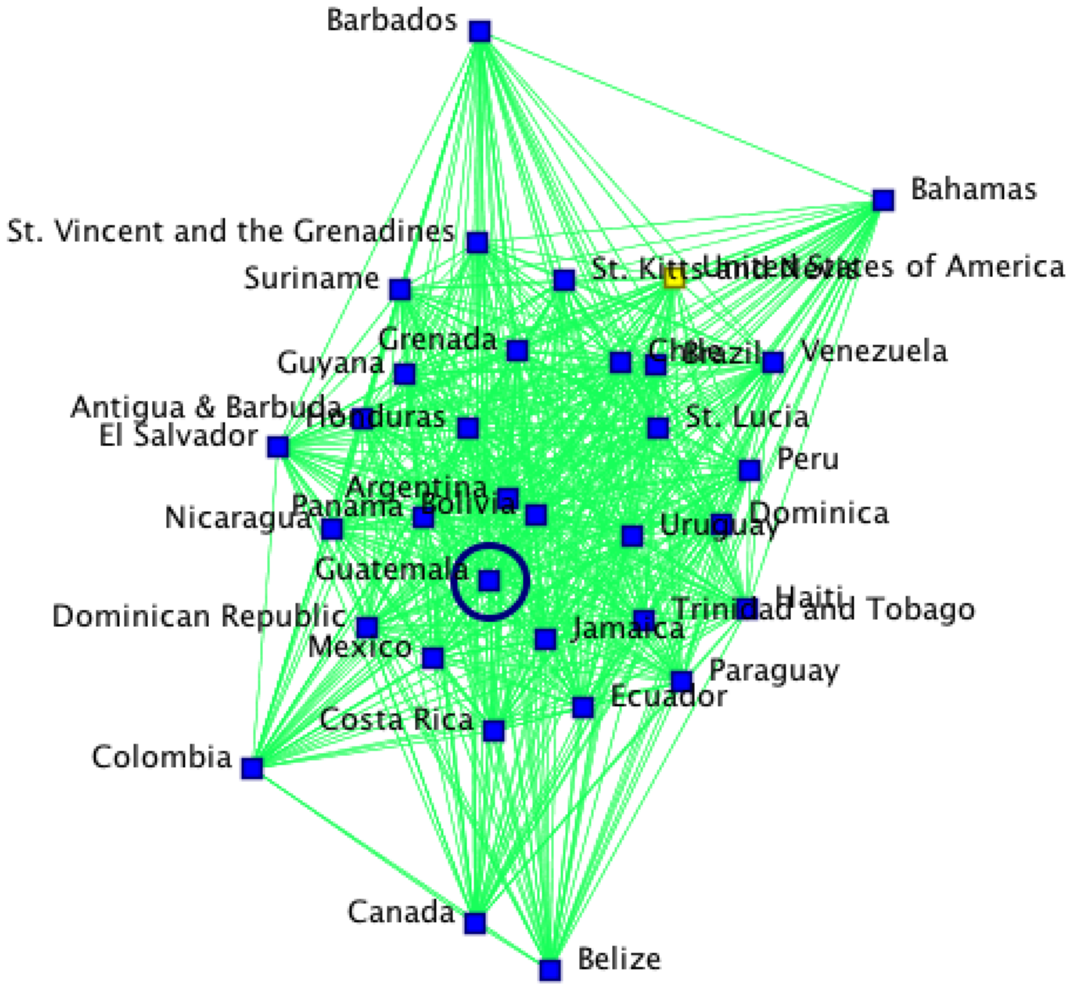


Figure 89. Egonet of Guatemala

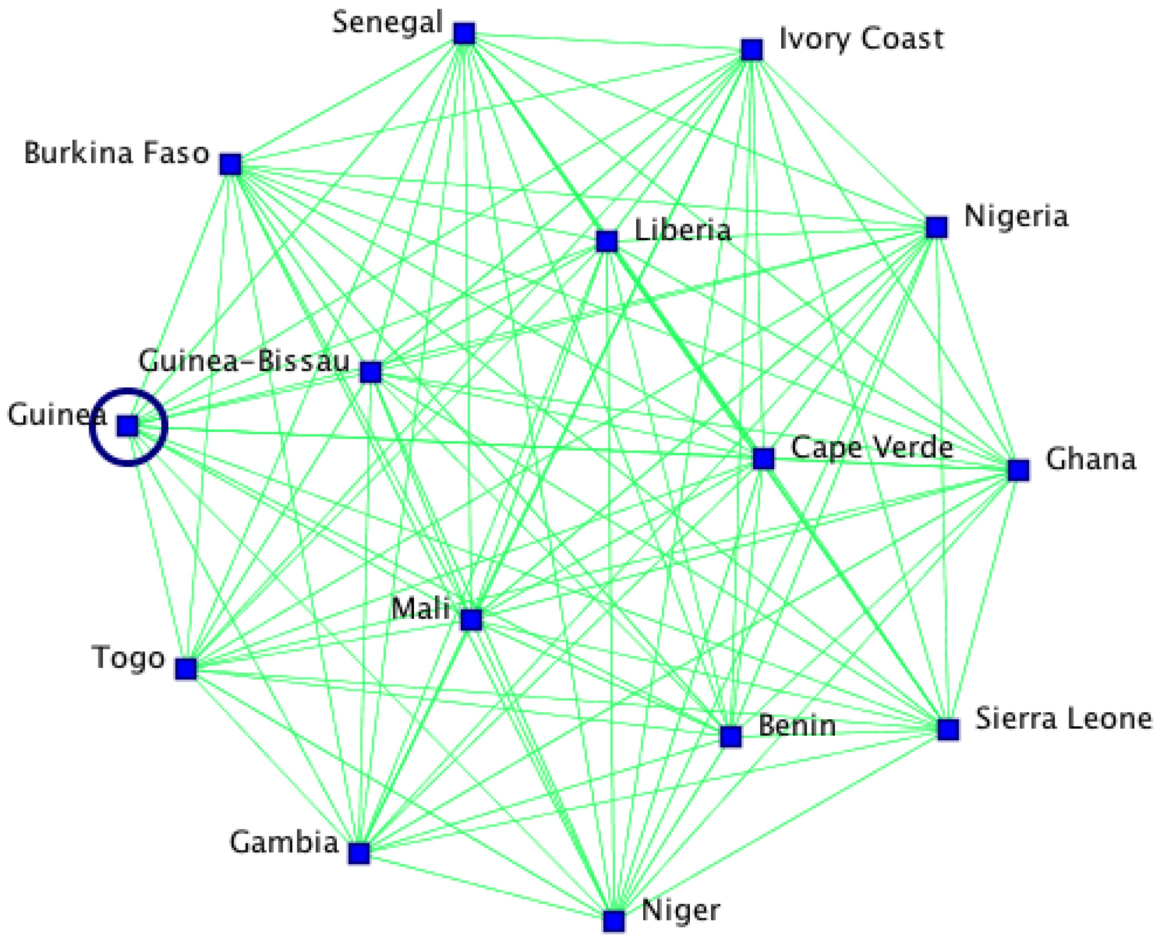


Figure 90. Egonet of Guinea

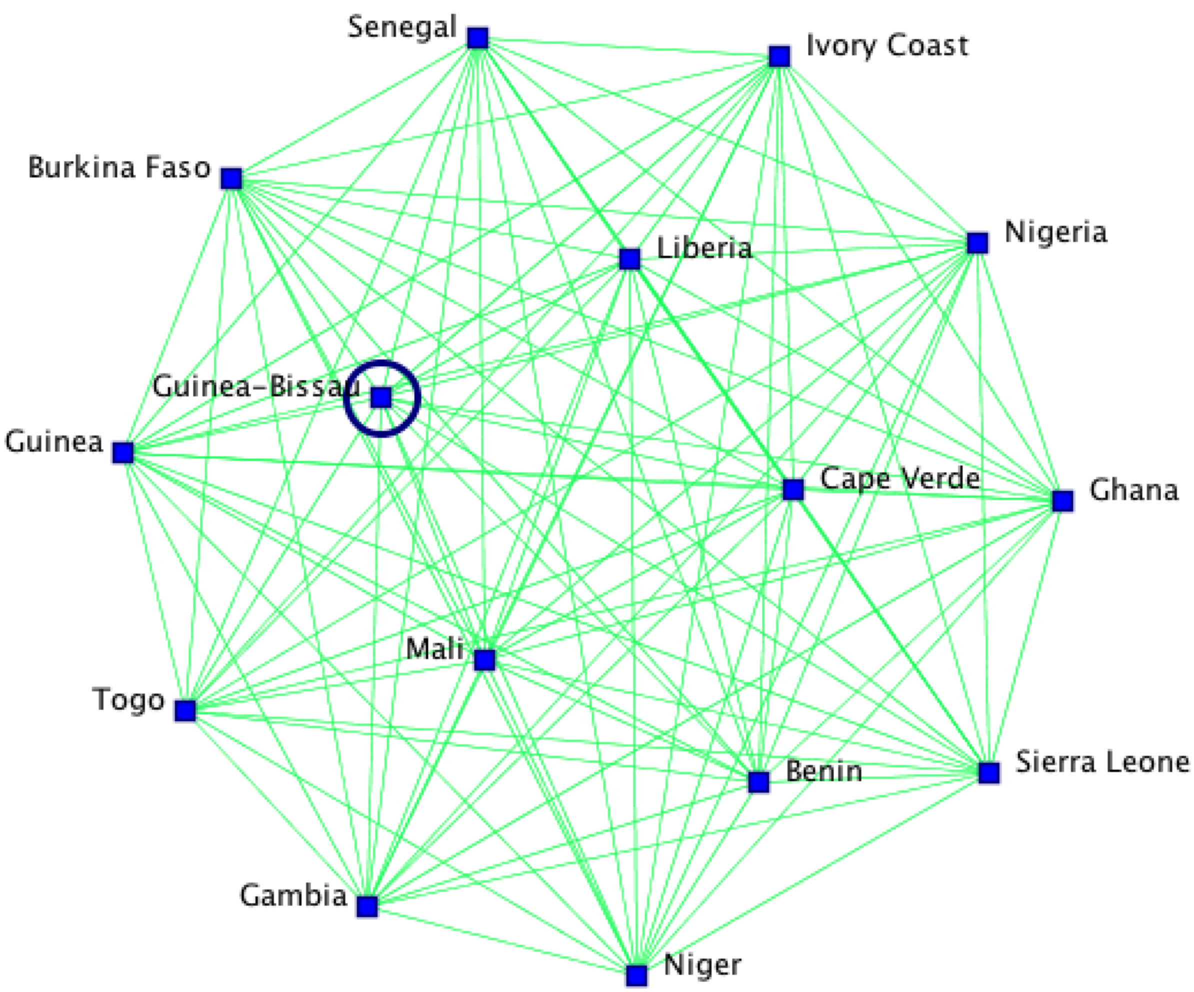


Figure 91. Egonet of Guinea-Bissau

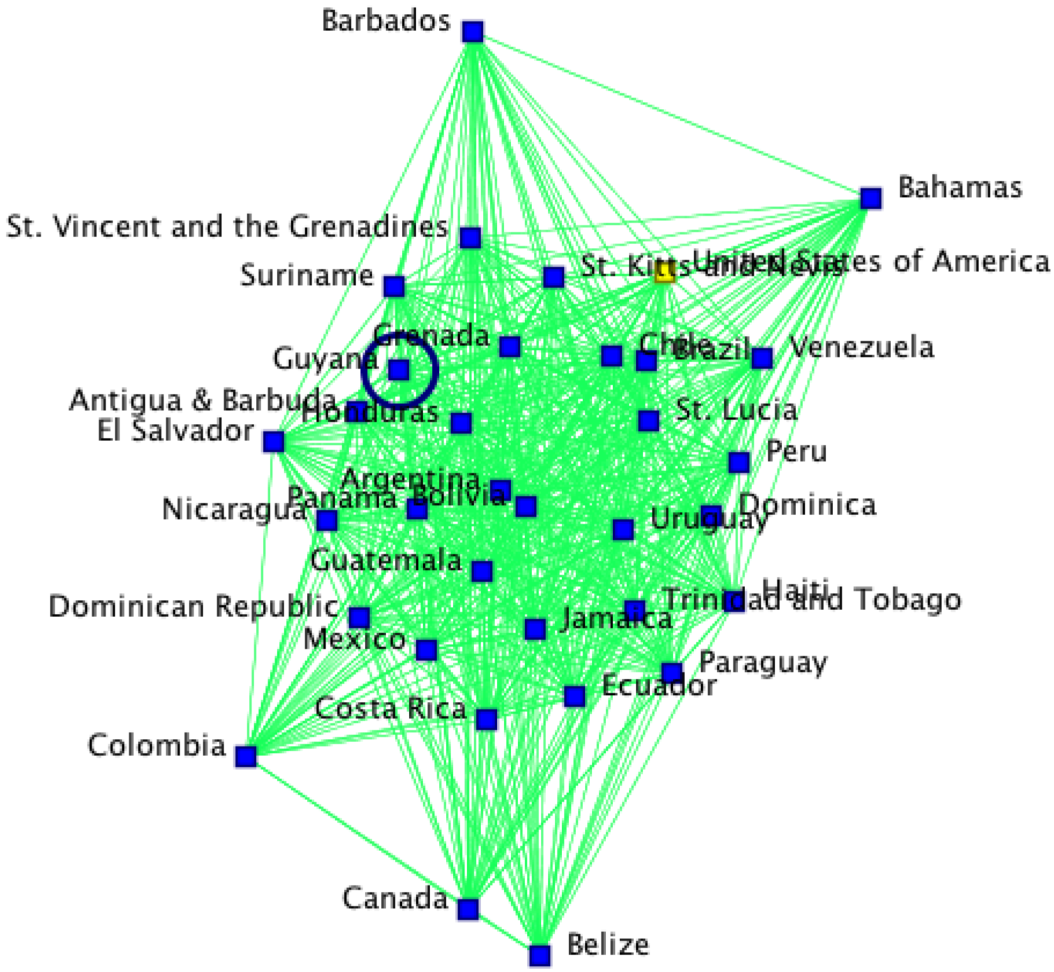


Figure 92. Egonet of Guyana

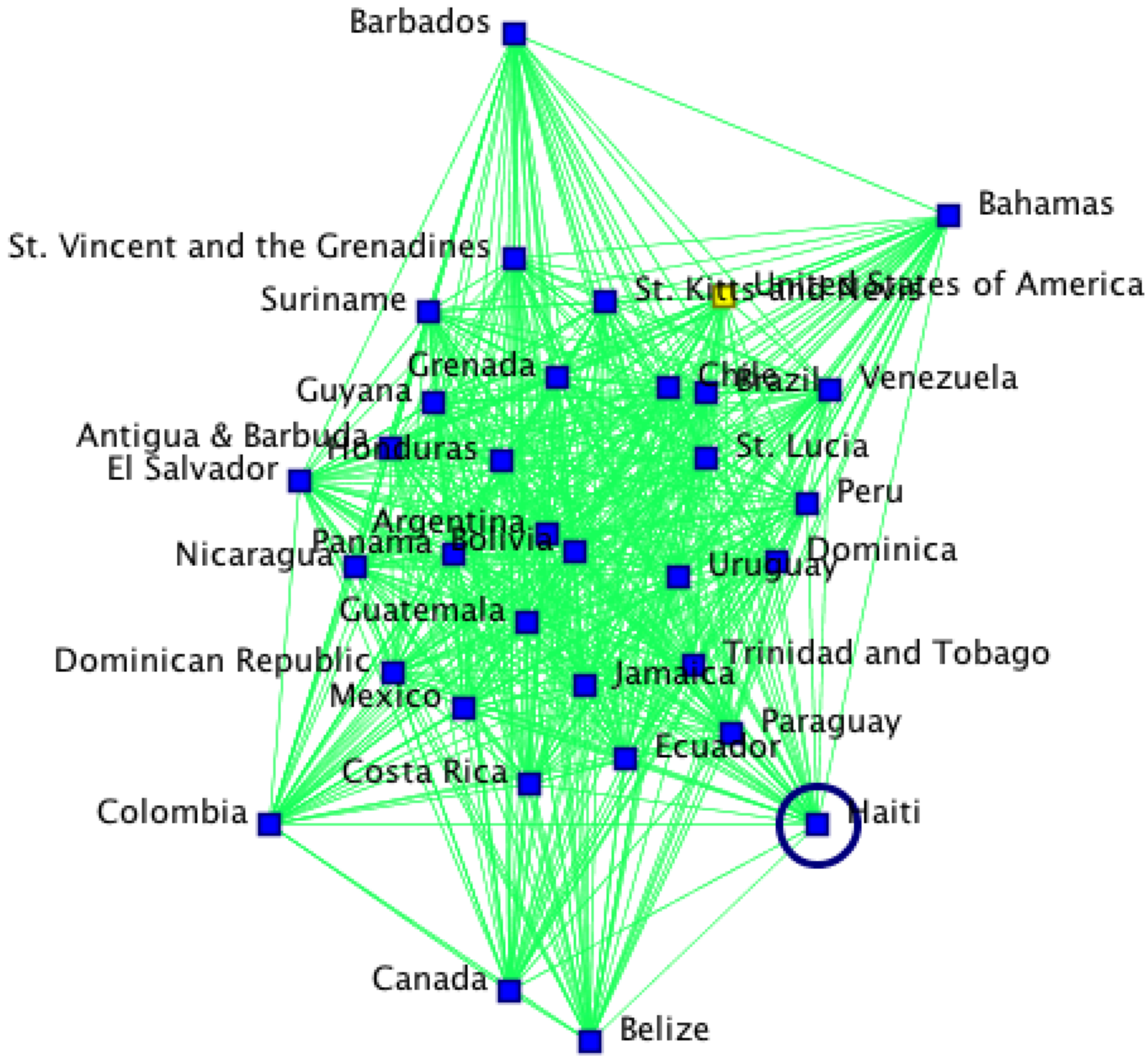


Figure 93. Egonet of Haiti

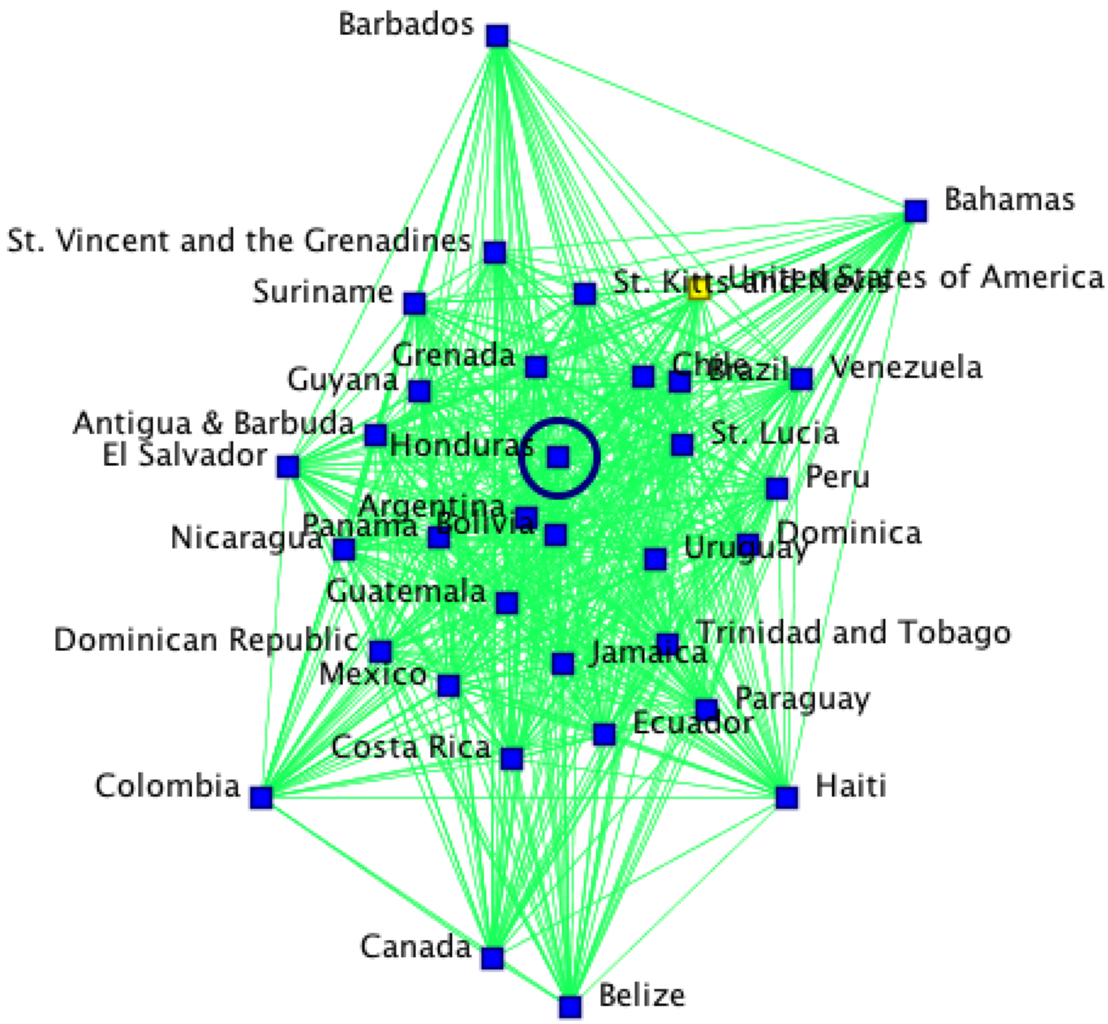


Figure 94. Egonet of Honduras

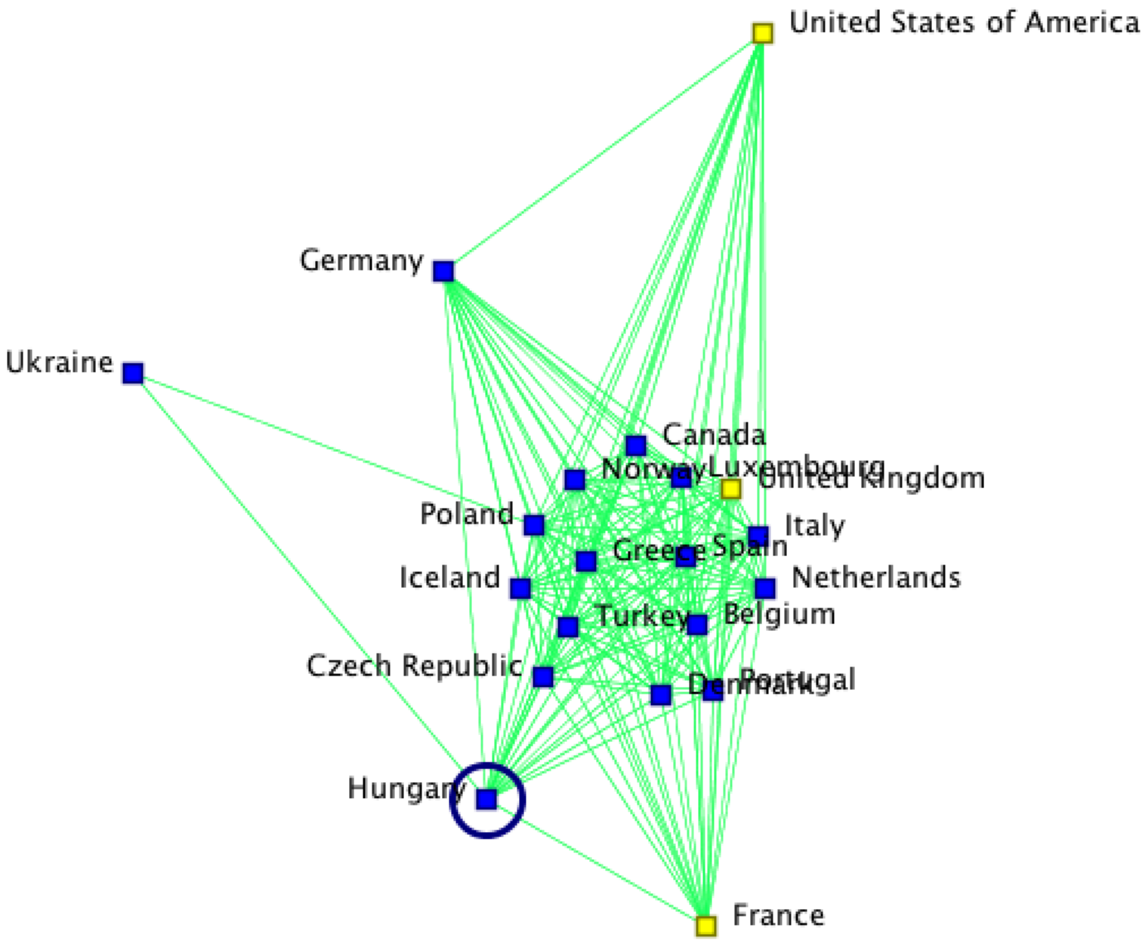


Figure 95. Egonet of Hungary

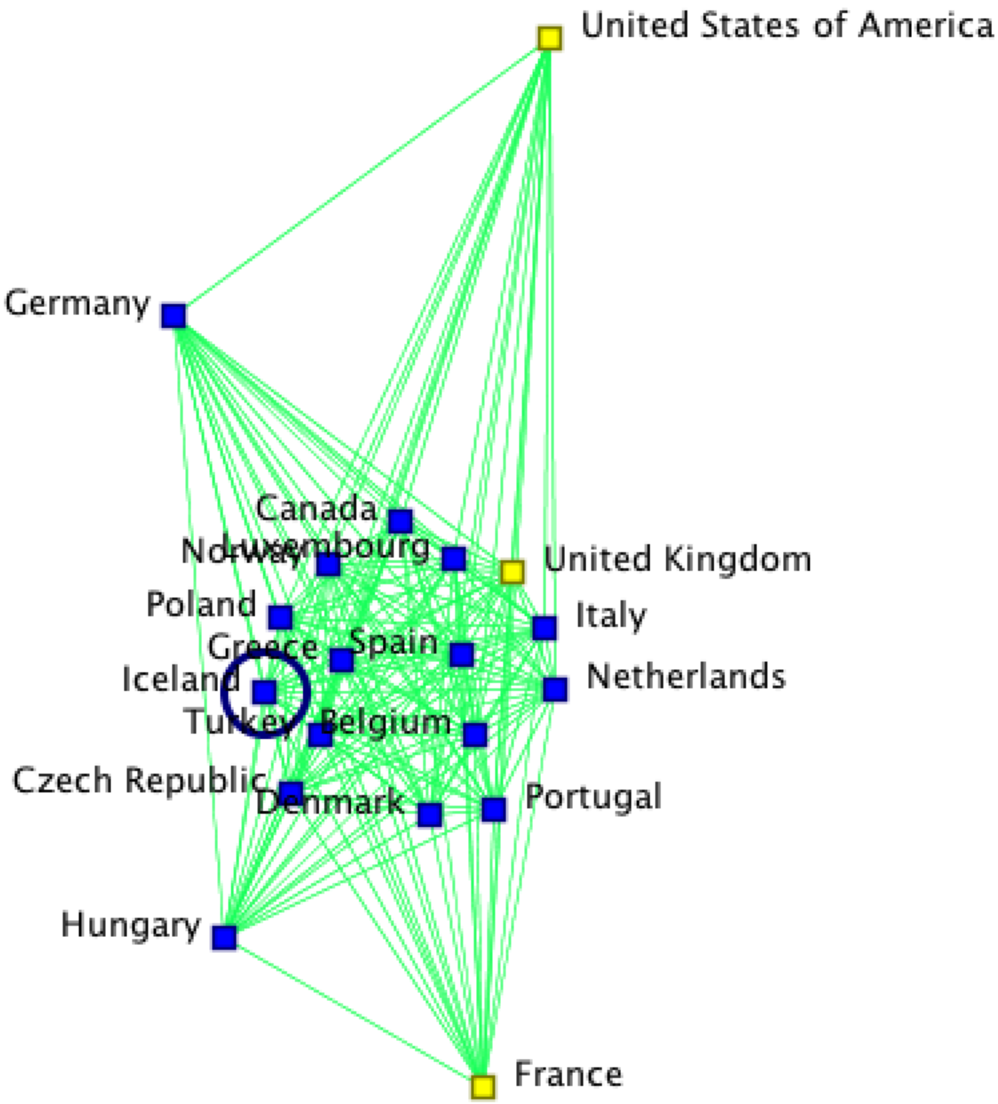


Figure 96. Egonet of Iceland

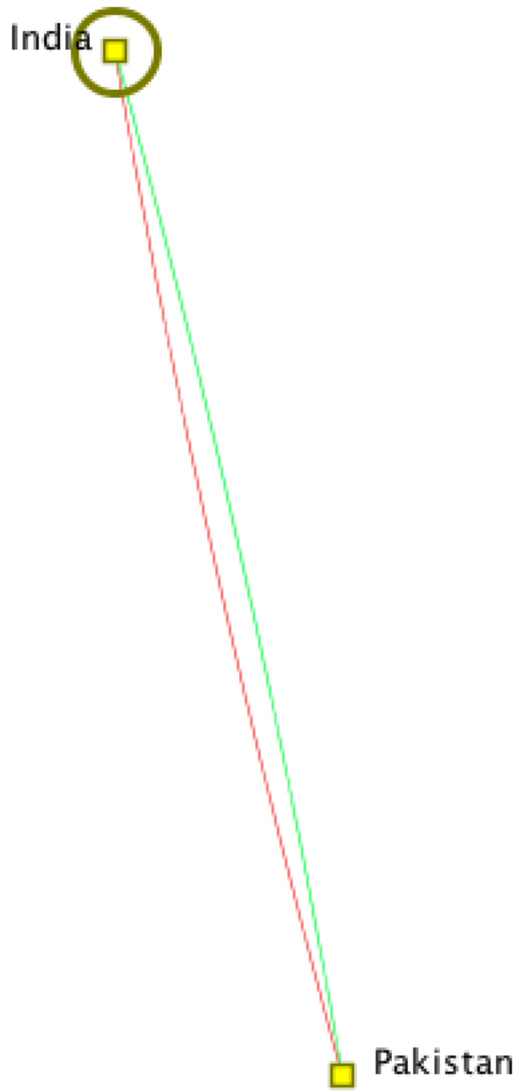


Figure 97. Egonet of India

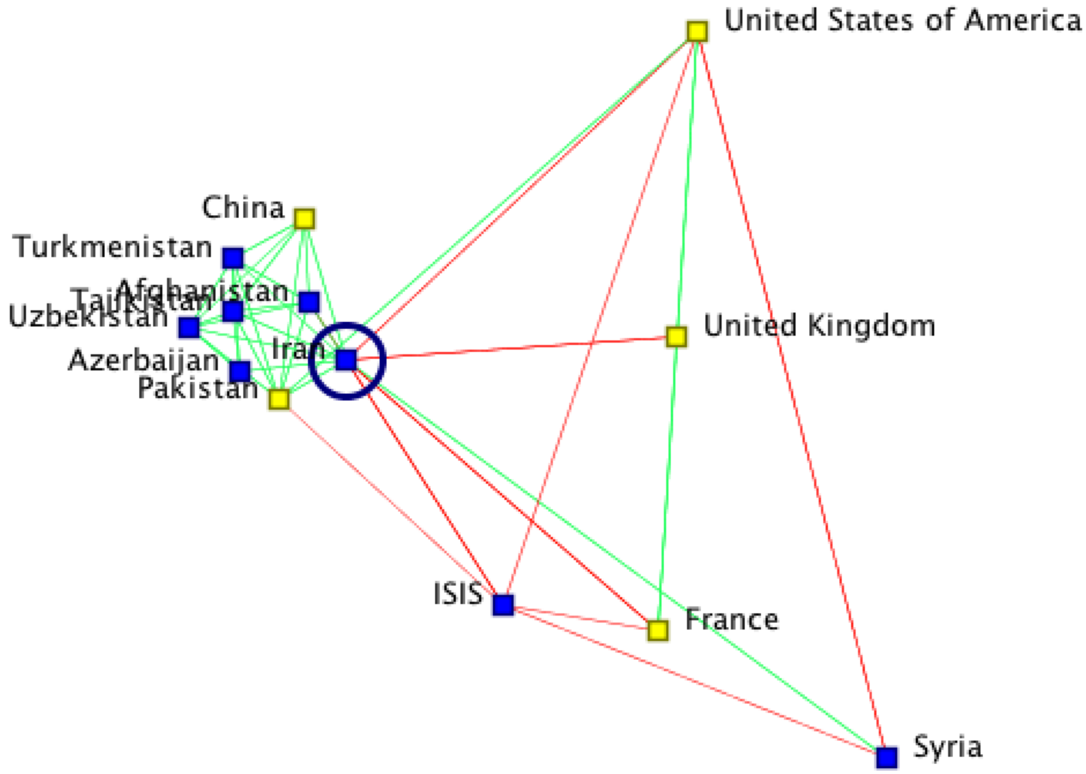


Figure 98. Egonet of Iran

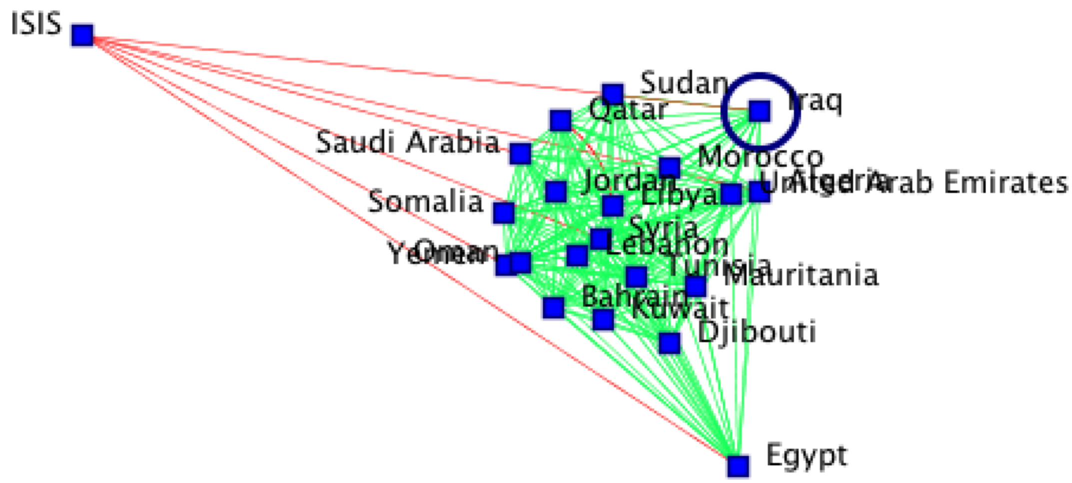


Figure 99. Egonet of Iraq

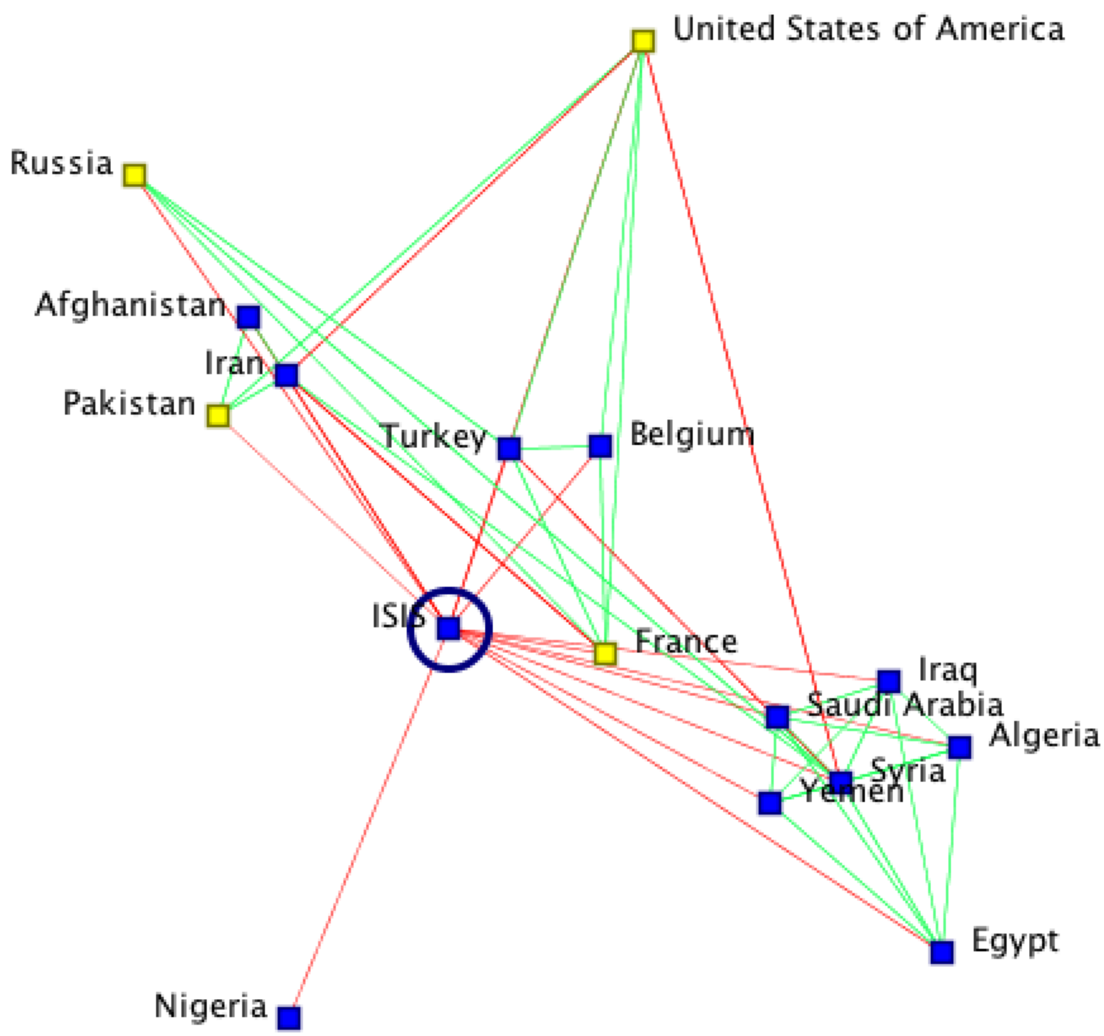


Figure 100. Egonet of ISIS

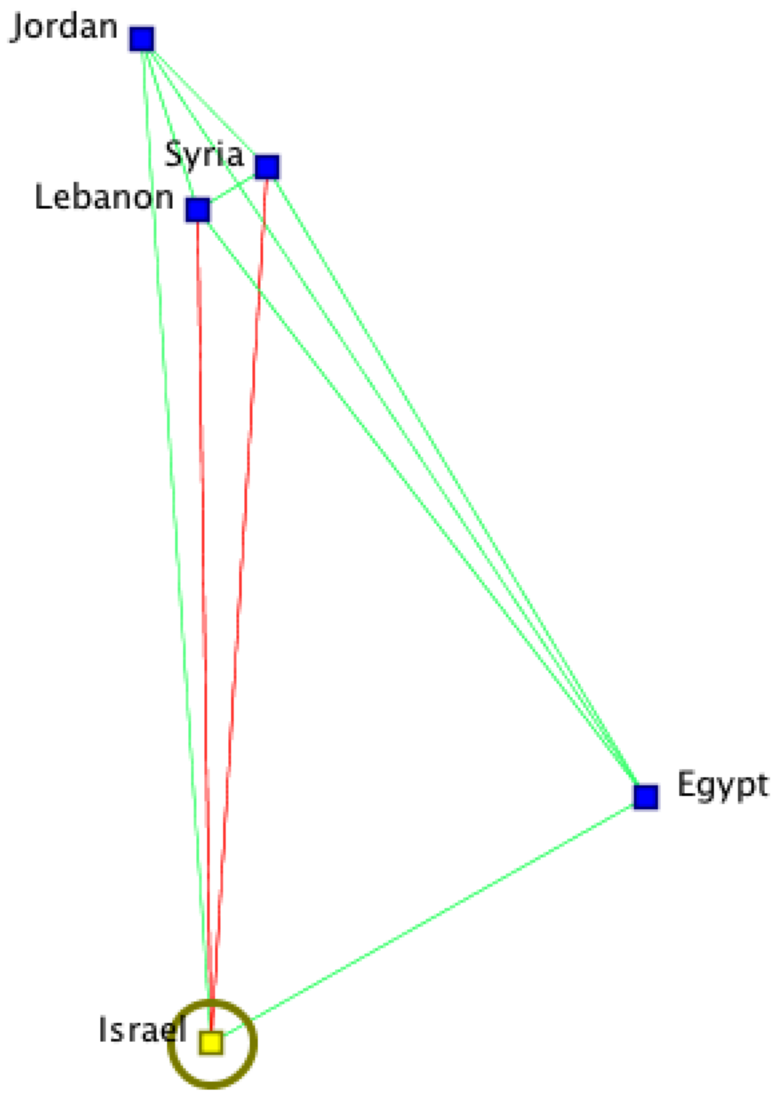


Figure 101. Egonet of Israel

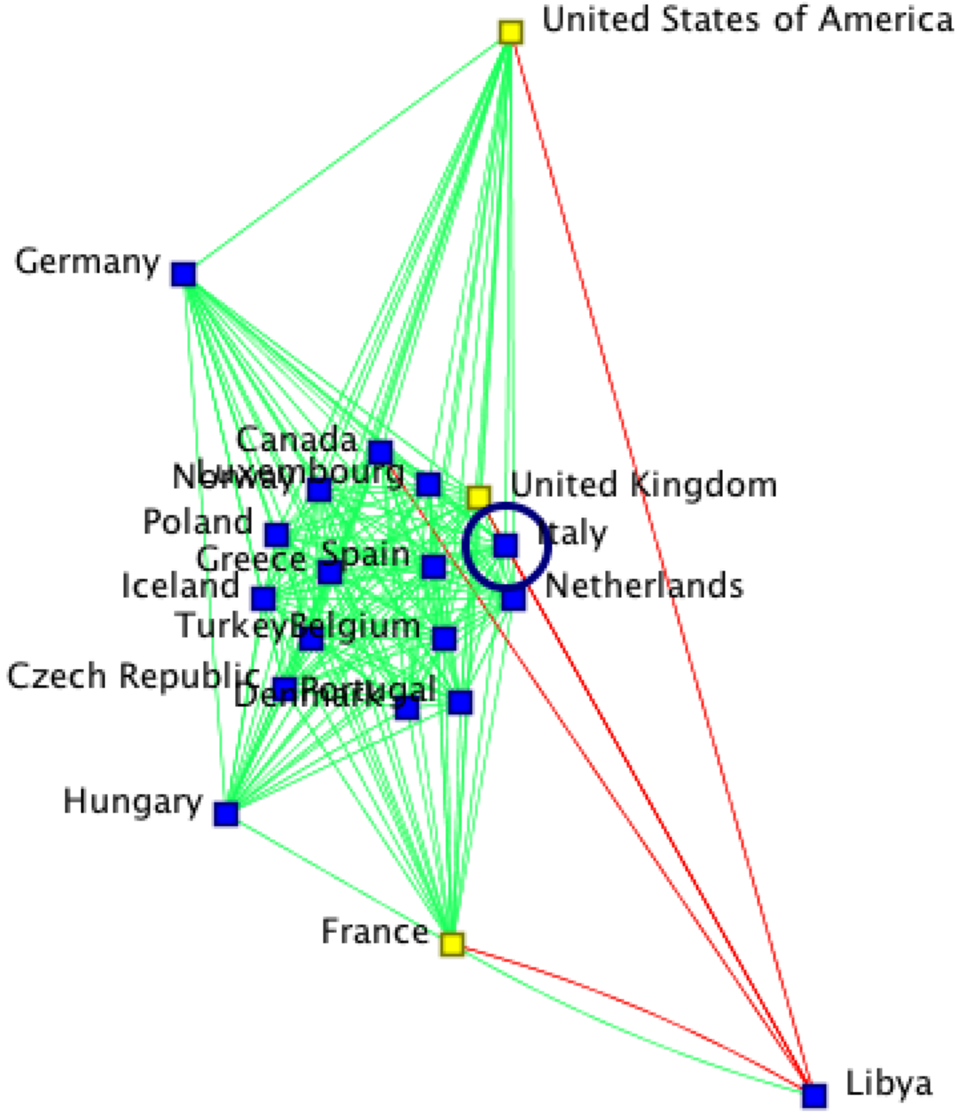


Figure 102. Egonet of Italy

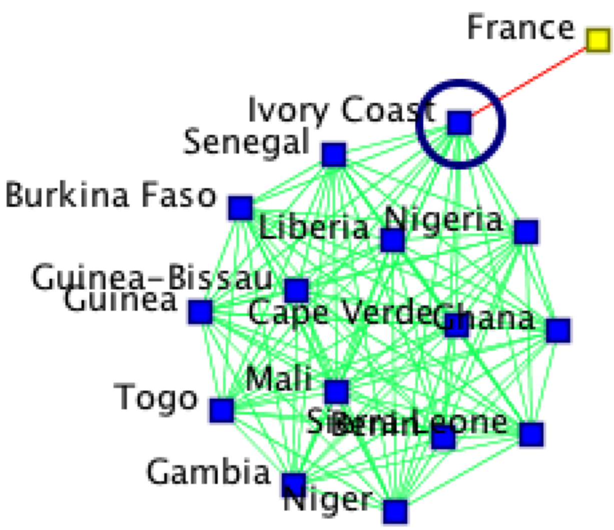


Figure 103. Egonet of Ivory Coast

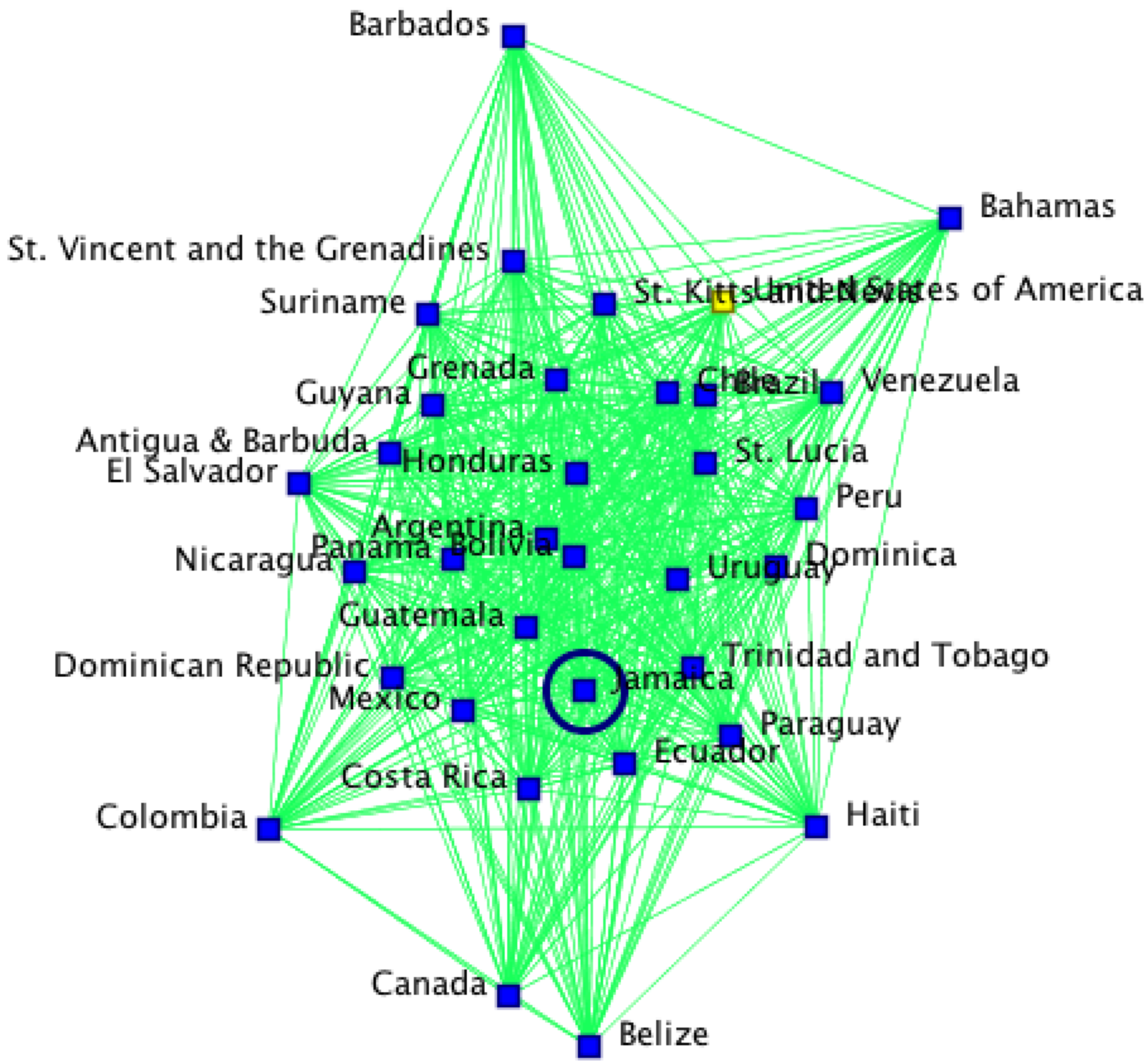


Figure 104. Egonet of Jamaica

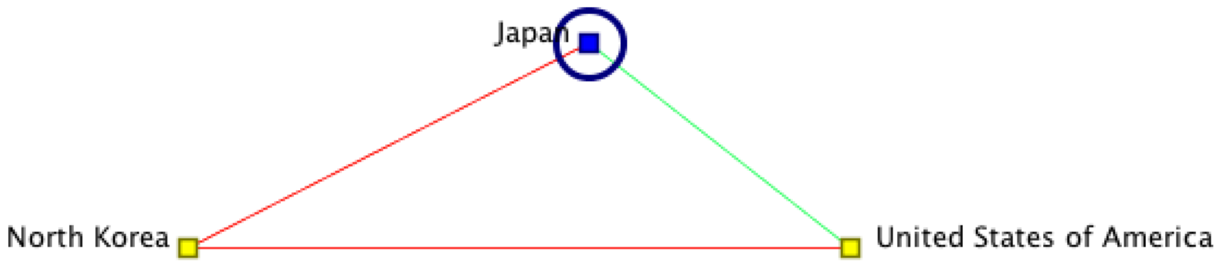


Figure 105. Egonet of Japan

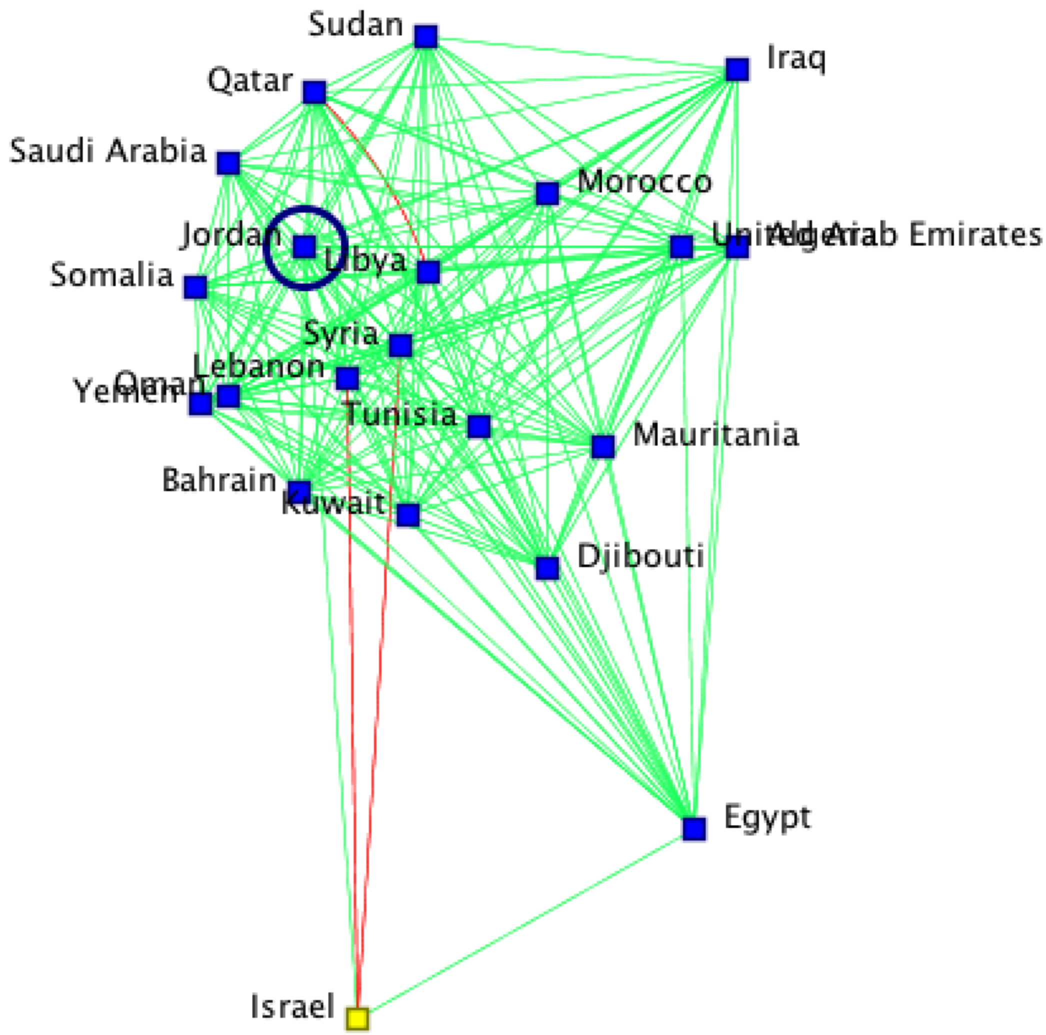


Figure 106. Egonet of Jordan

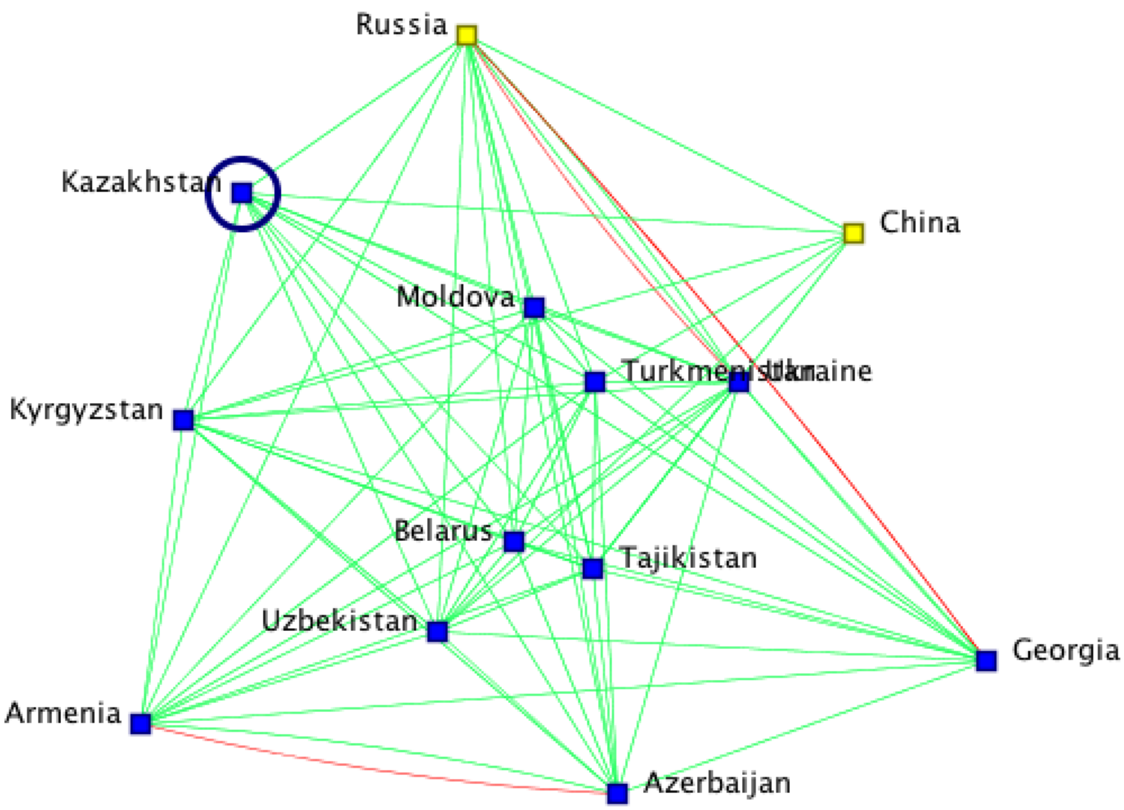


Figure 107. Egonet of Kazakhstan

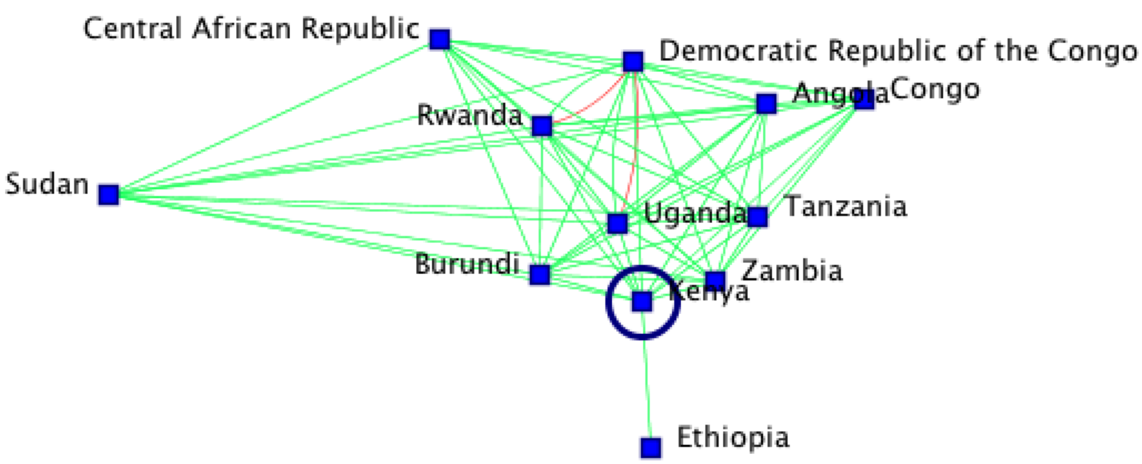


Figure 108. Egonet of Kenya

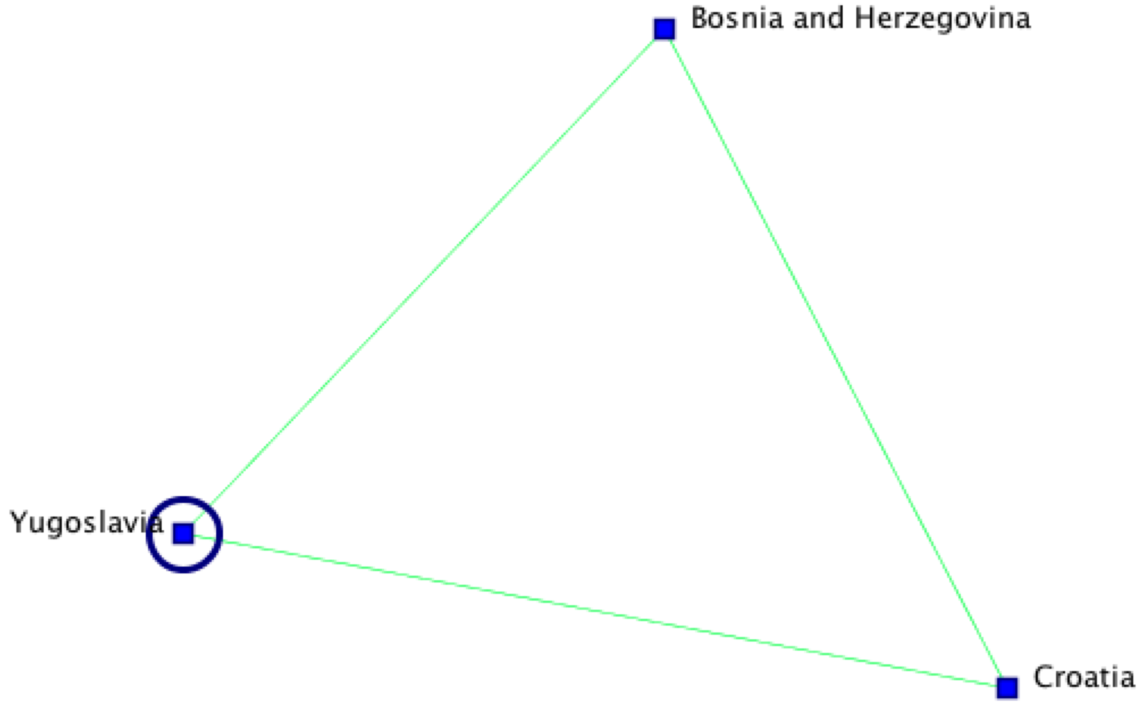


Figure 109. Egonet of Kosovo (Yugoslavia)

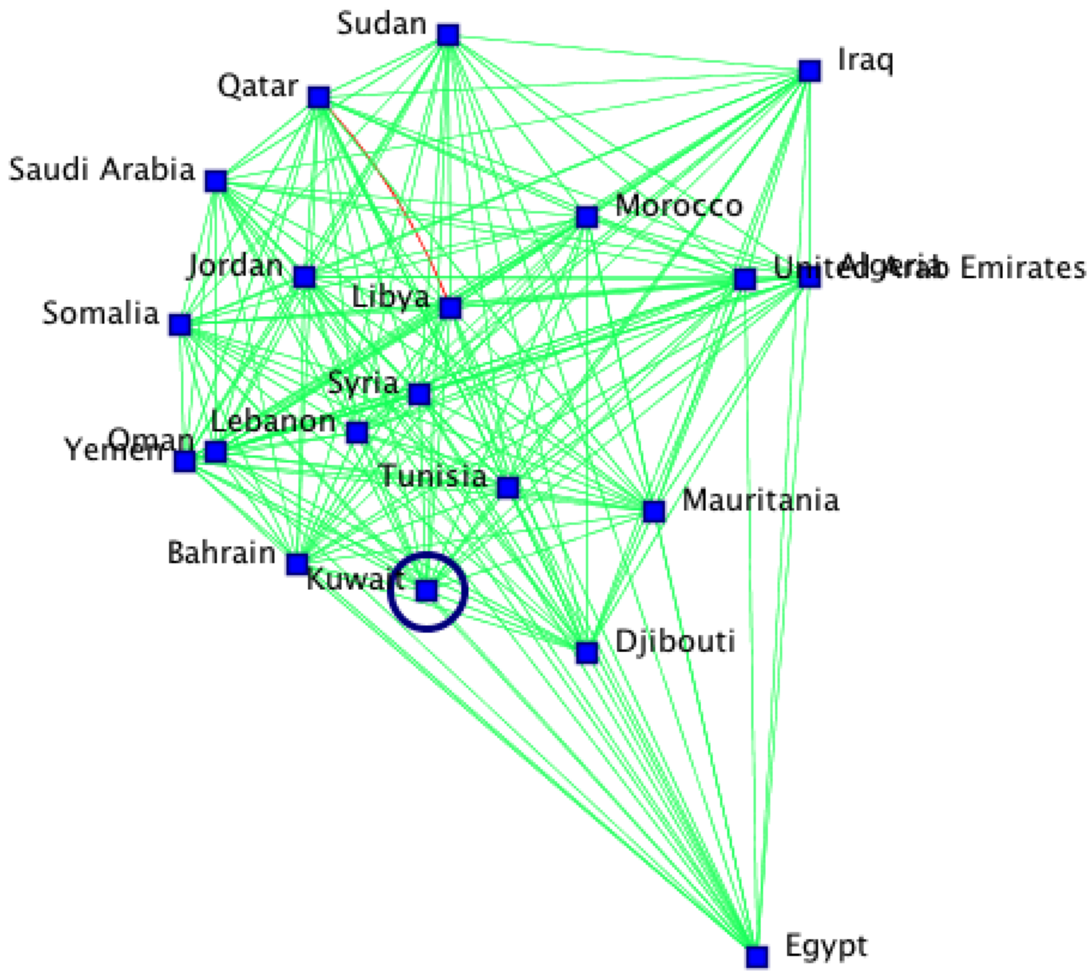


Figure 110. Egonet of Kuwait

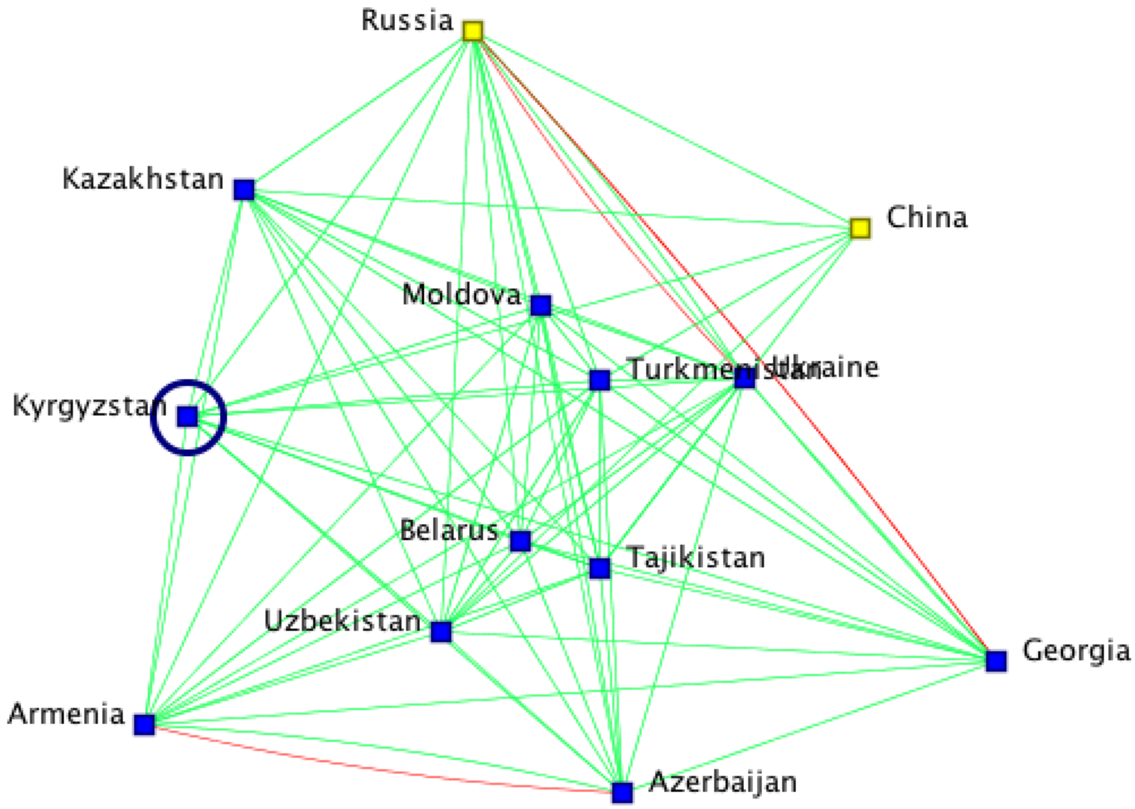


Figure 111. Egonet of Kyrgyzstan

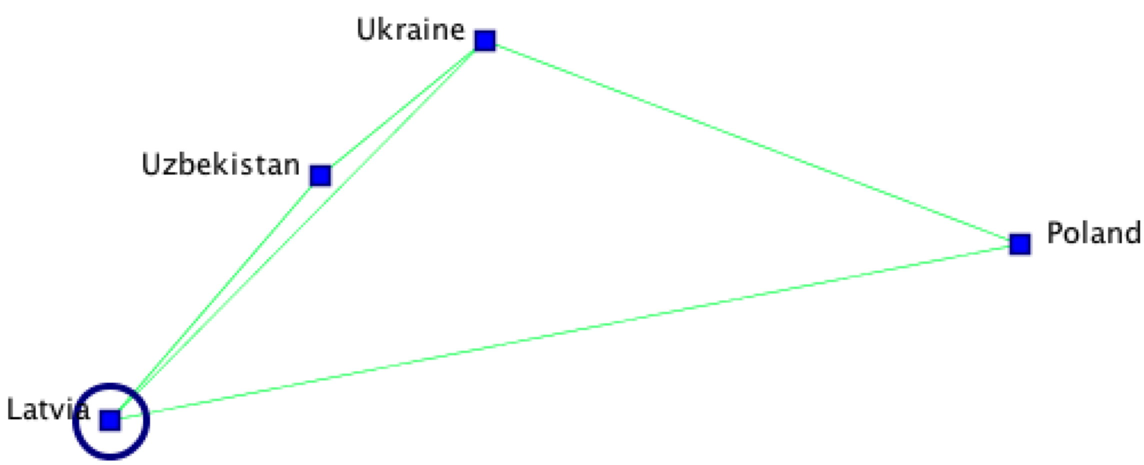


Figure 112. Egonet of Latvia

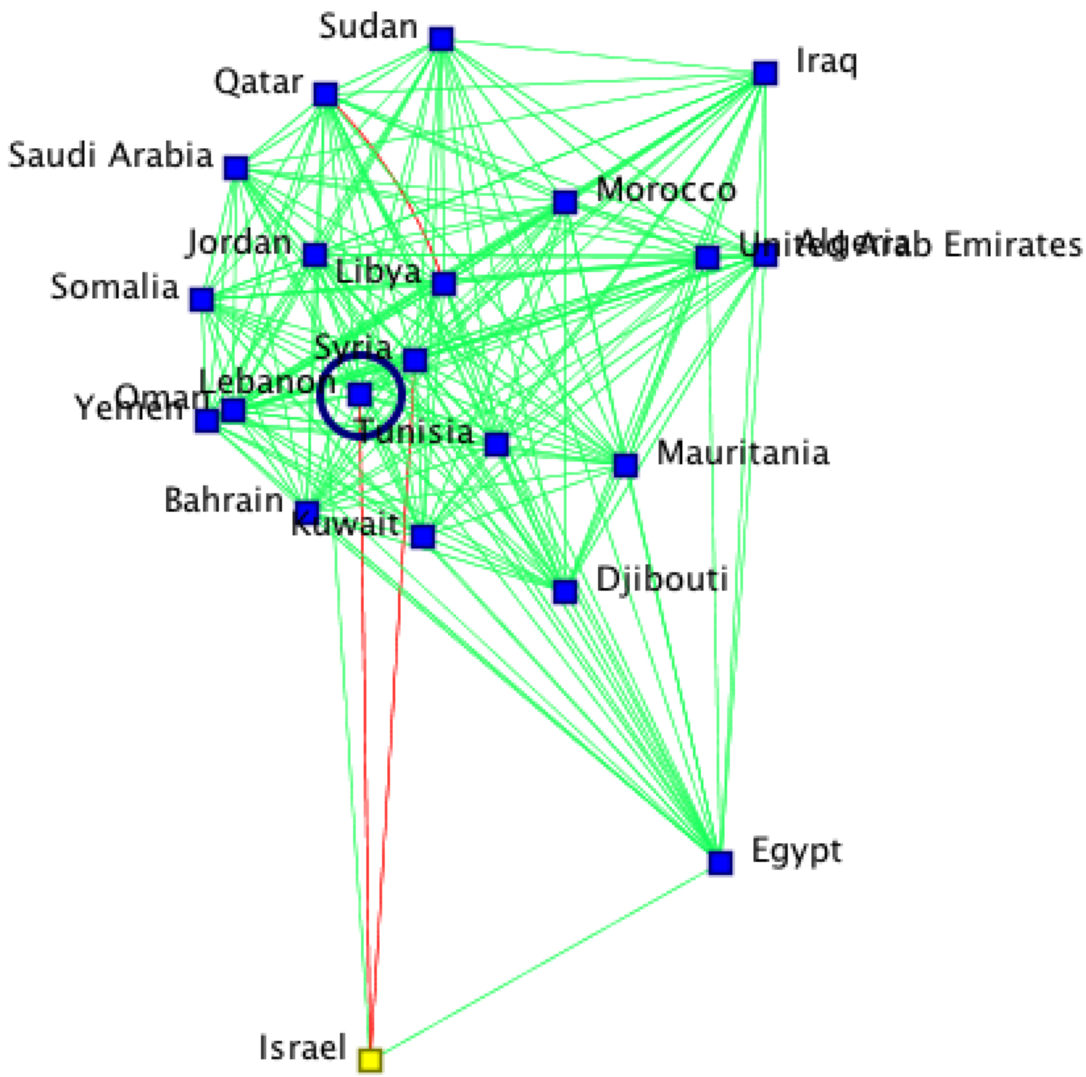


Figure 113. Egonet of Lebanon

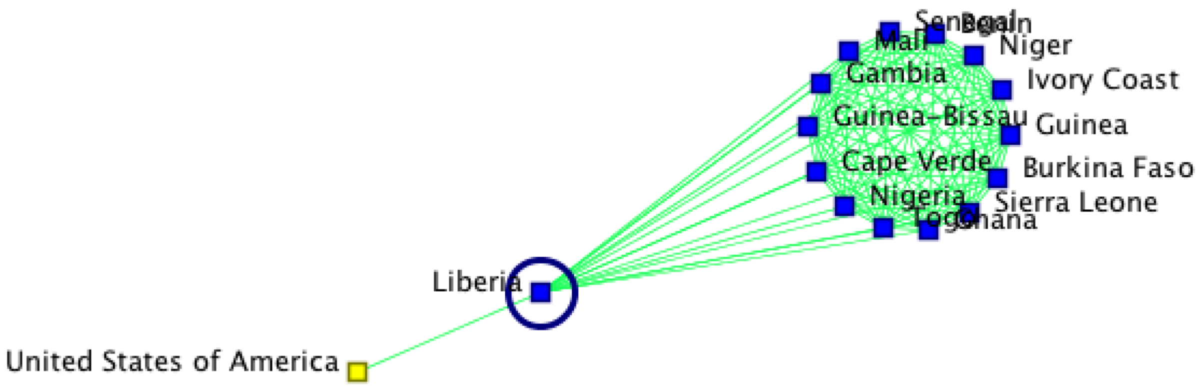


Figure 114. Egonet of Liberia

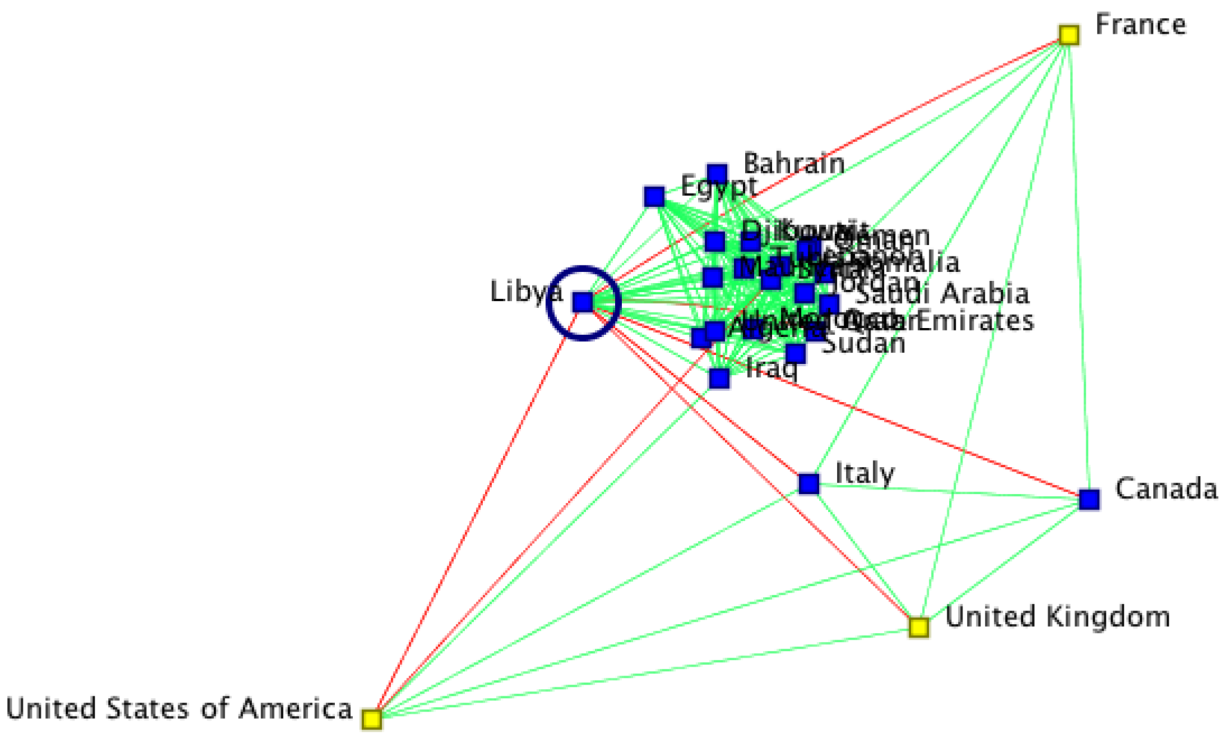


Figure 115. Egonet of Libya

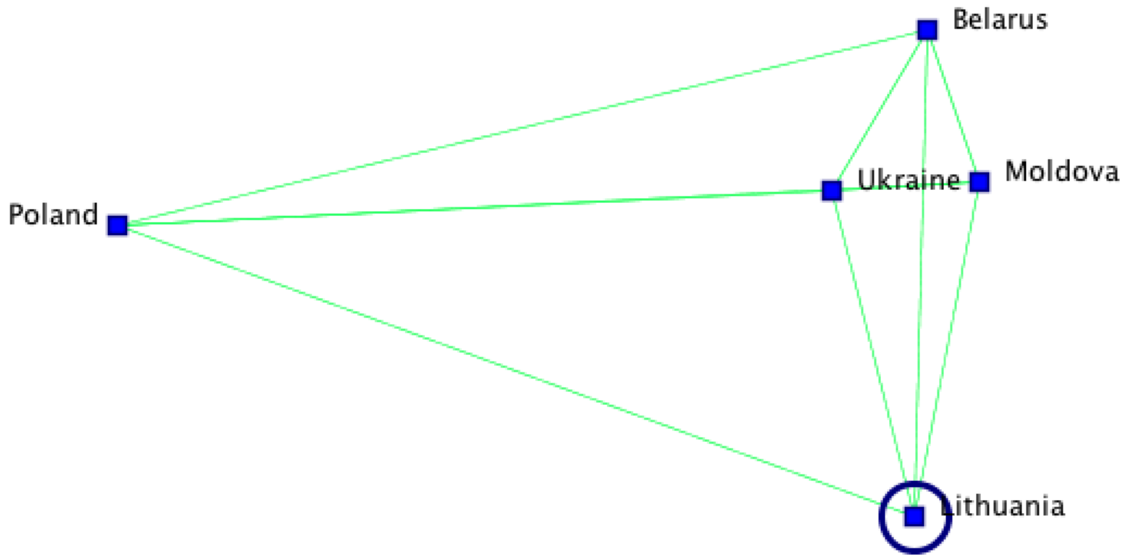


Figure 116. Egonet of Lithuania

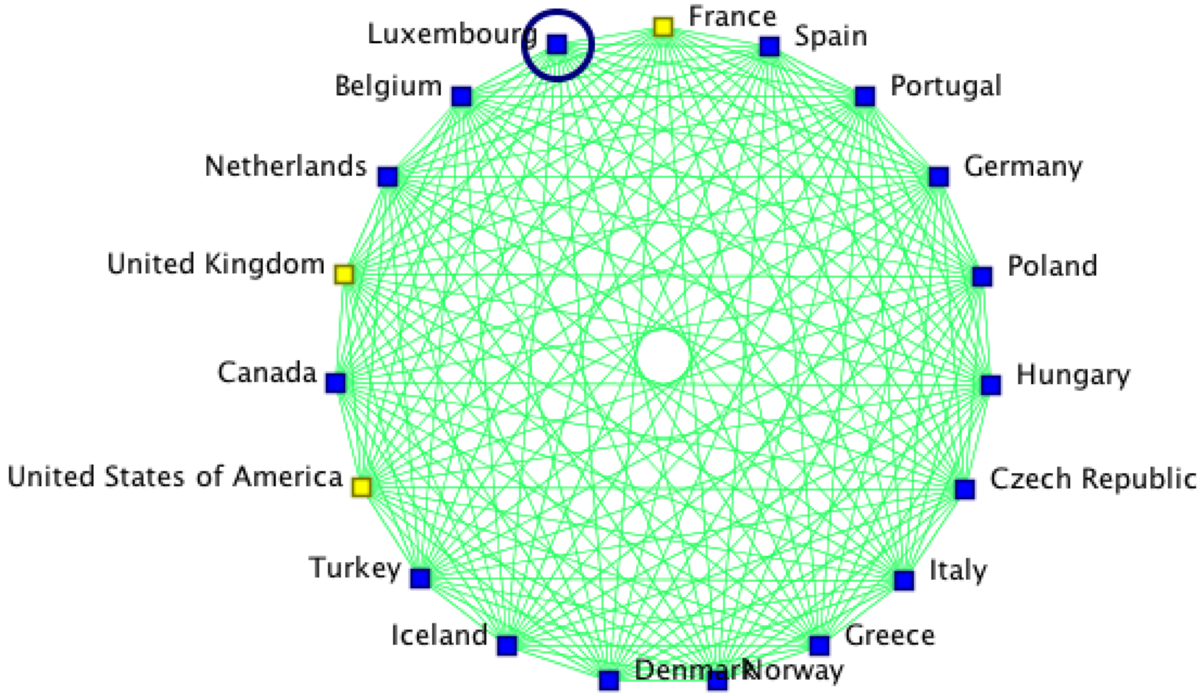


Figure 117. Egonet of Luxembourg

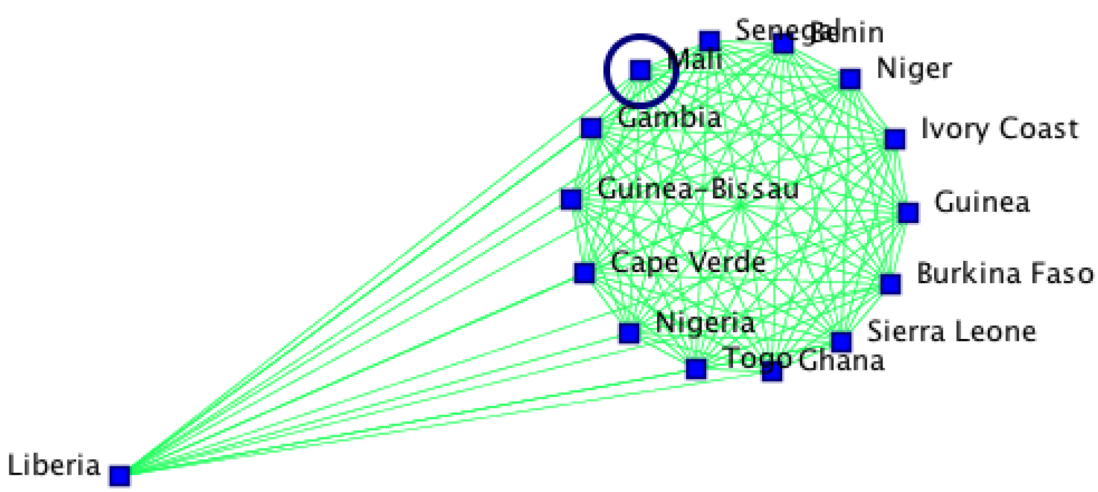


Figure 118. Egonet of Mali

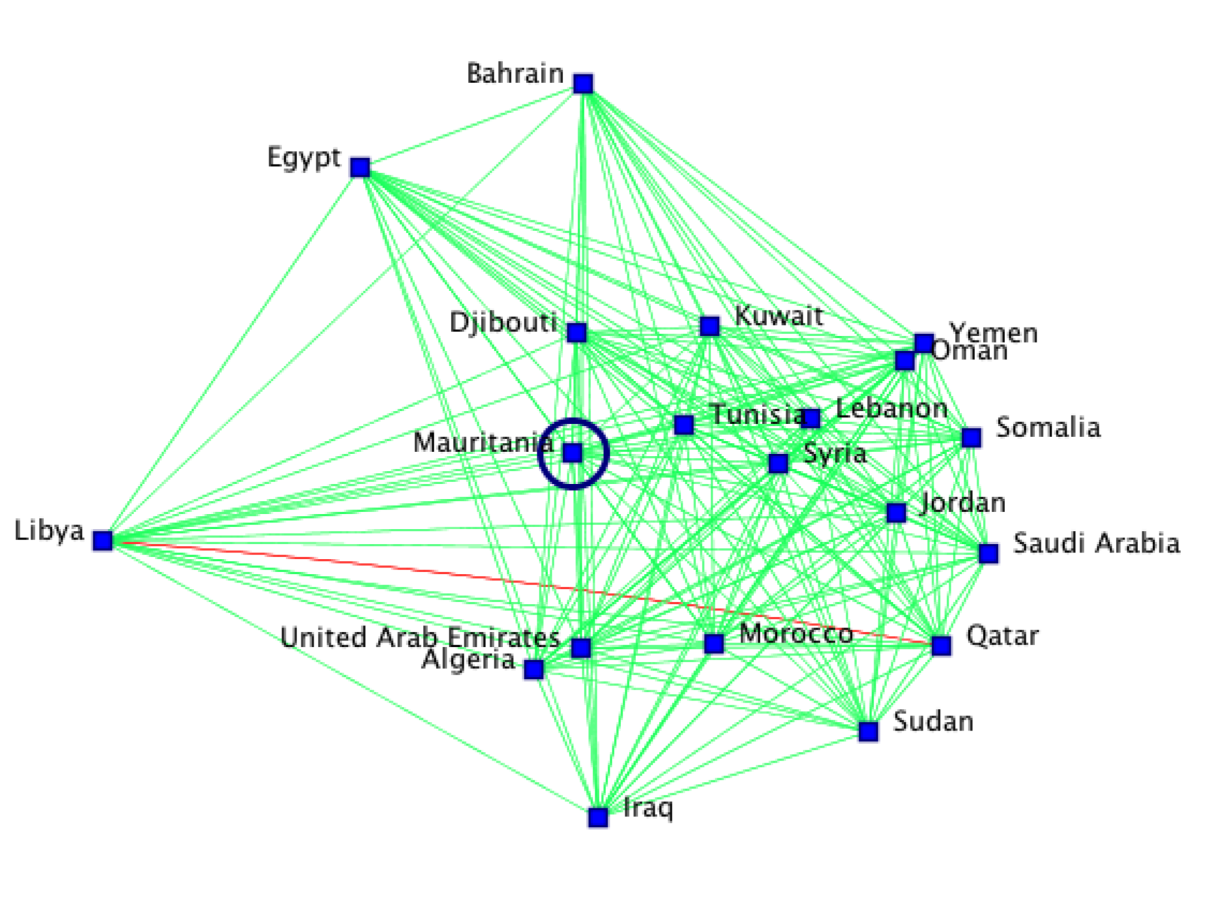


Figure 119. Egonet of Mauritania

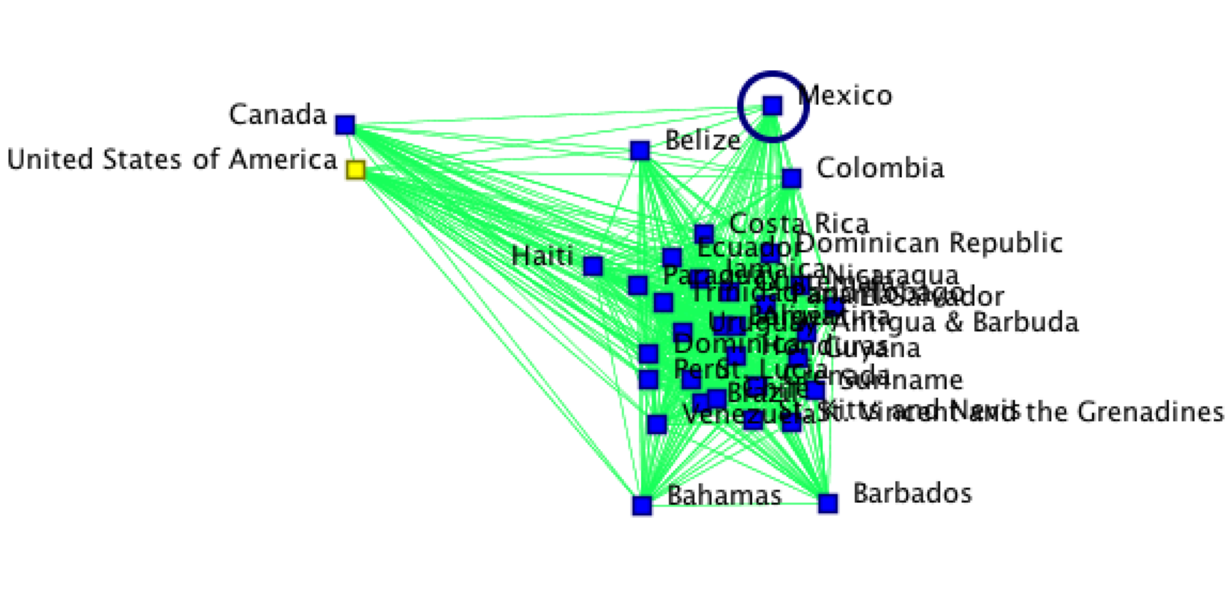


Figure 120. Egonet of Mexico

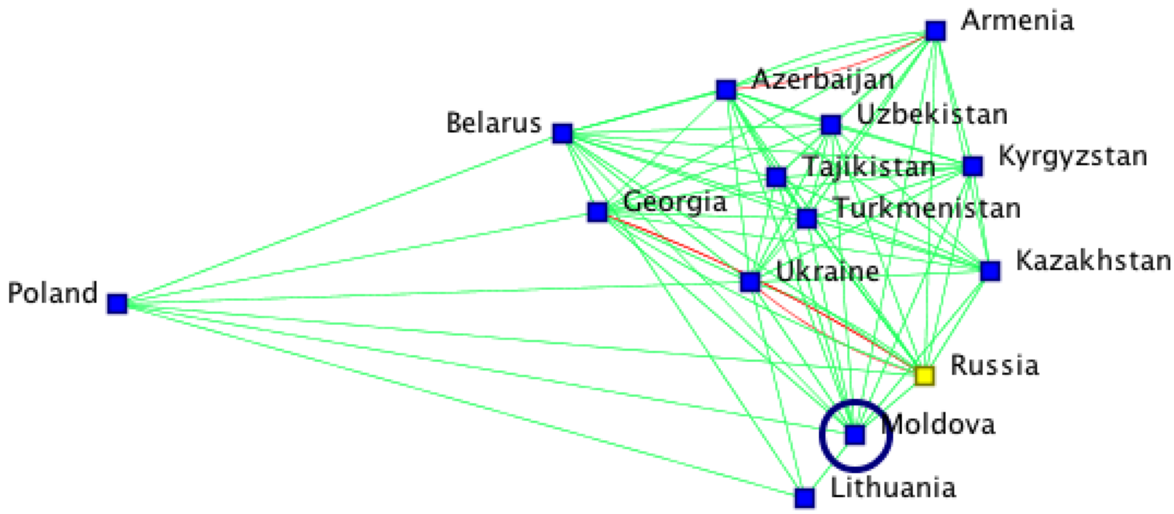


Figure 121. Egonet of Moldova

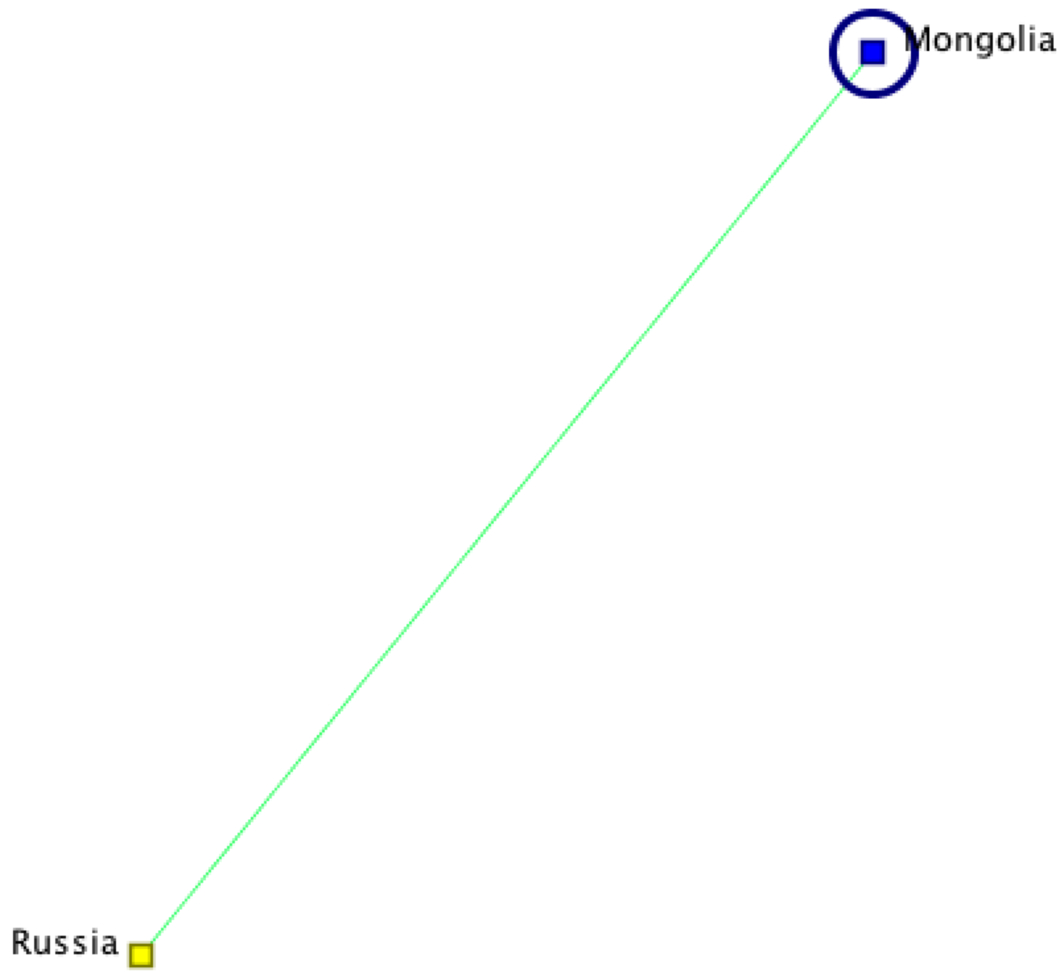


Figure 122. Egonet of Mongolia

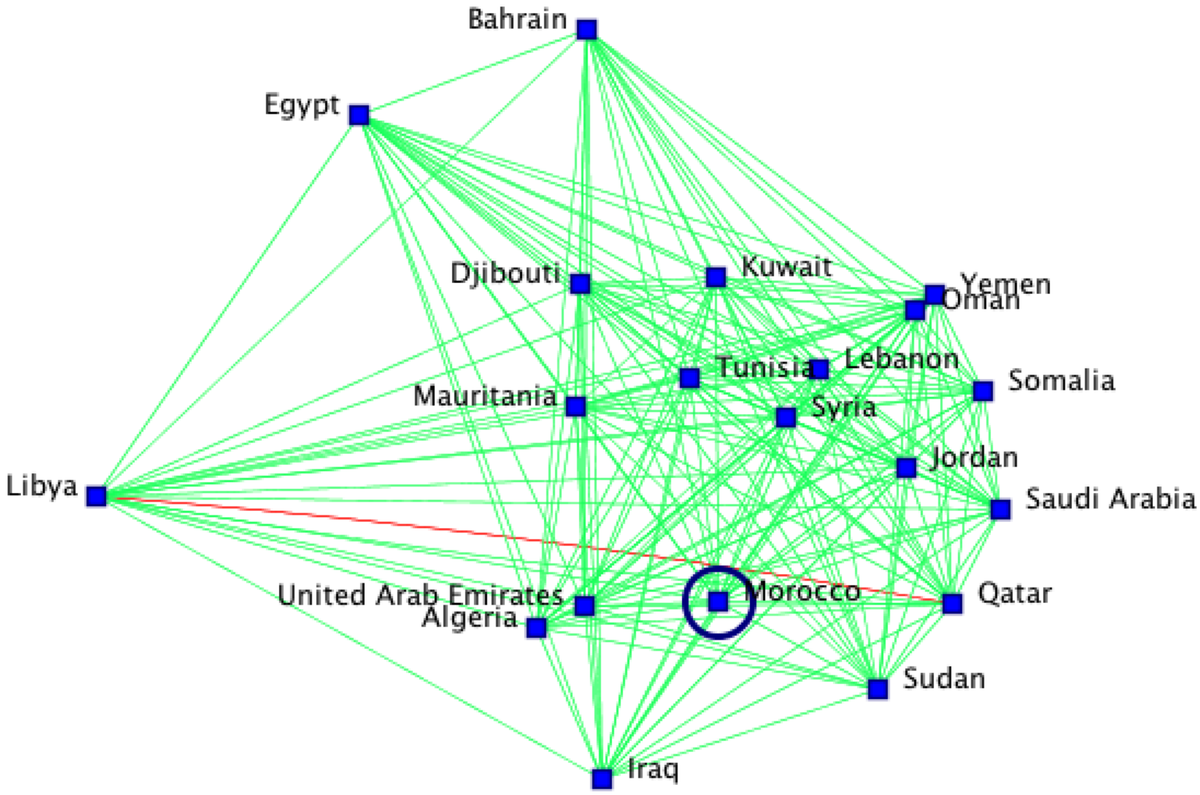


Figure 123. Egonet of Morocco.

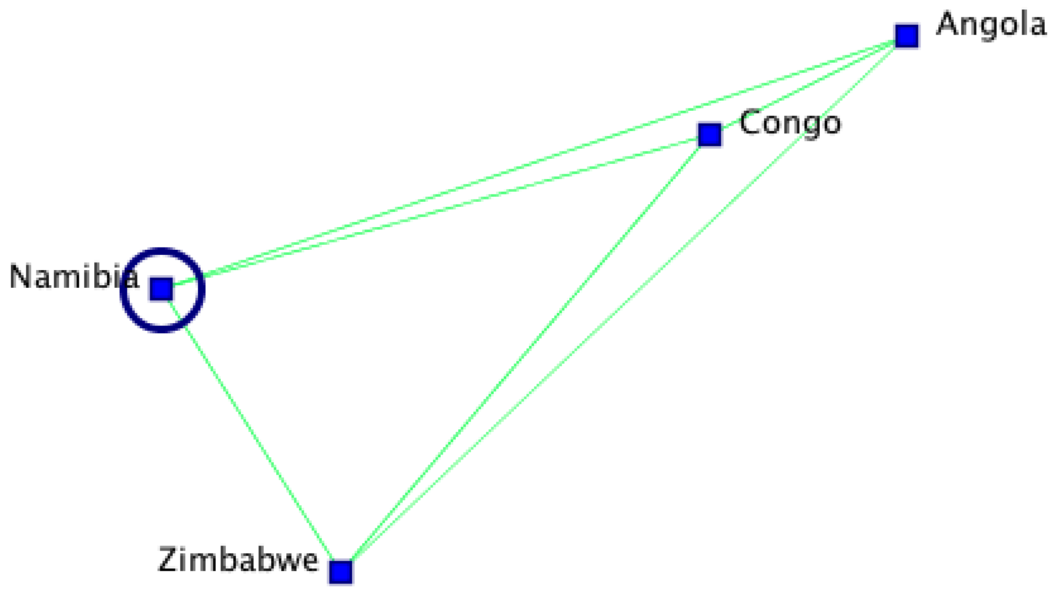


Figure 124. Egonet of Namibia

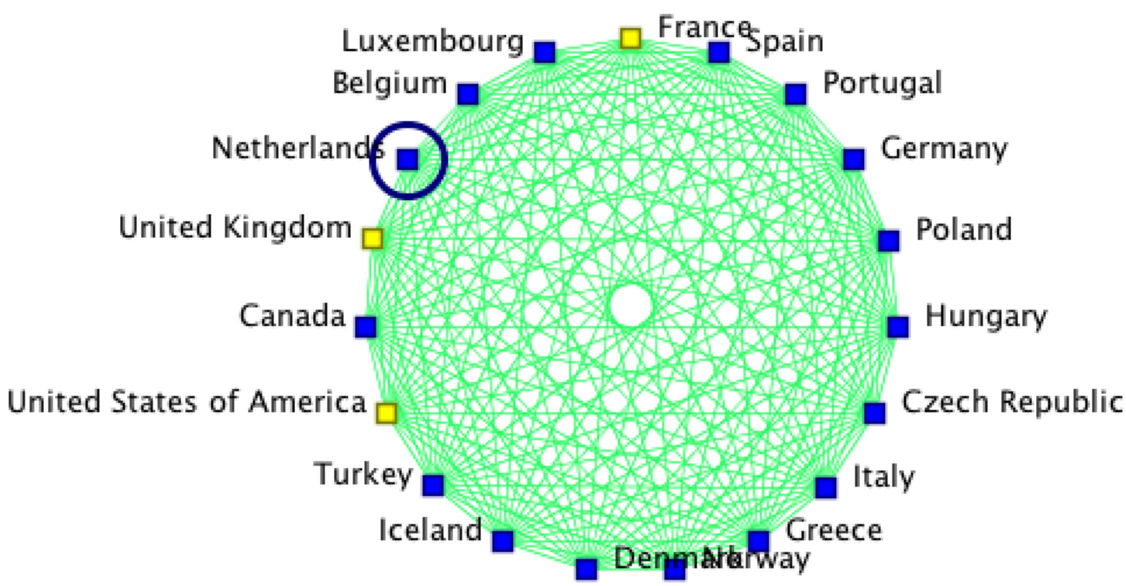


Figure 125. Egonet of Netherlands

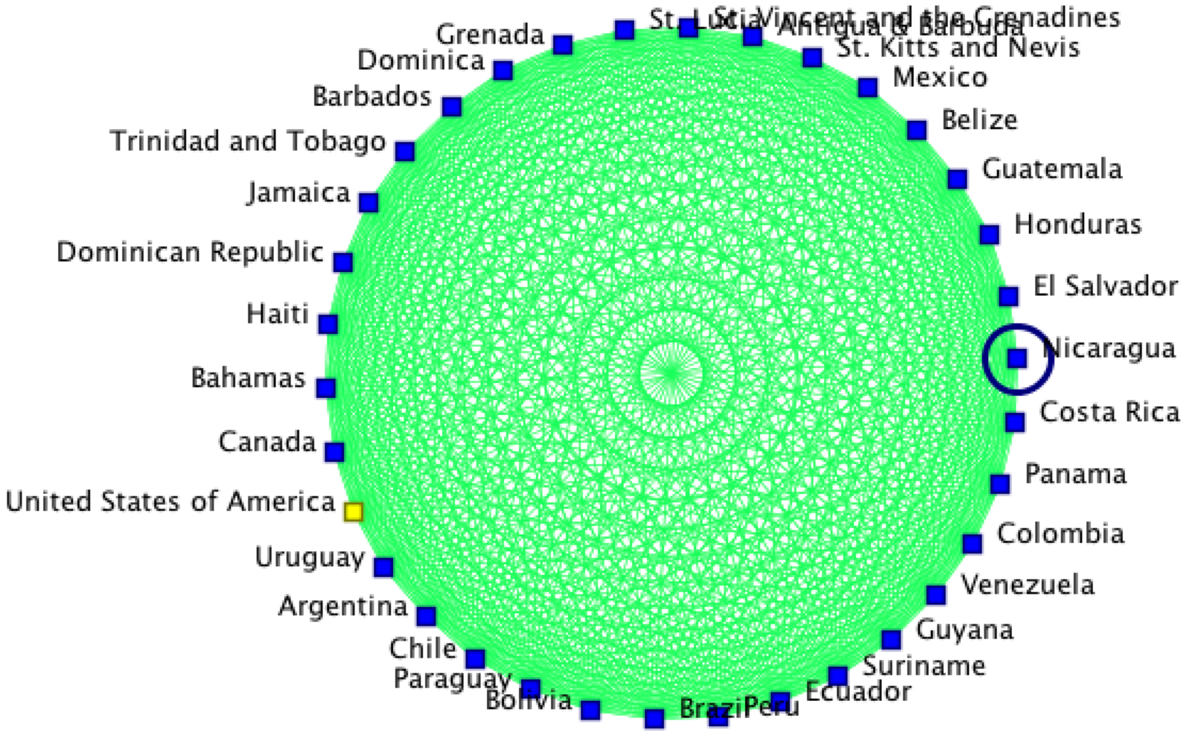


Figure 126. Egonet of Nicaragua

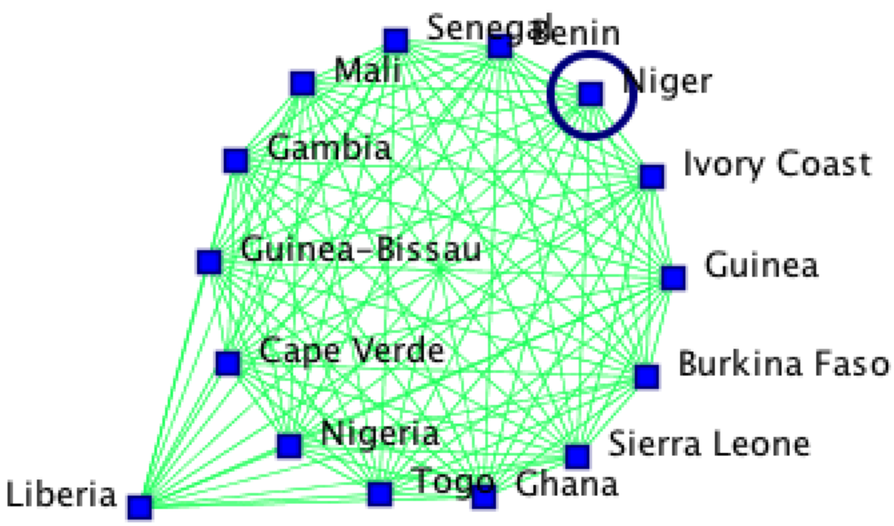


Figure 127. Egonet of Niger

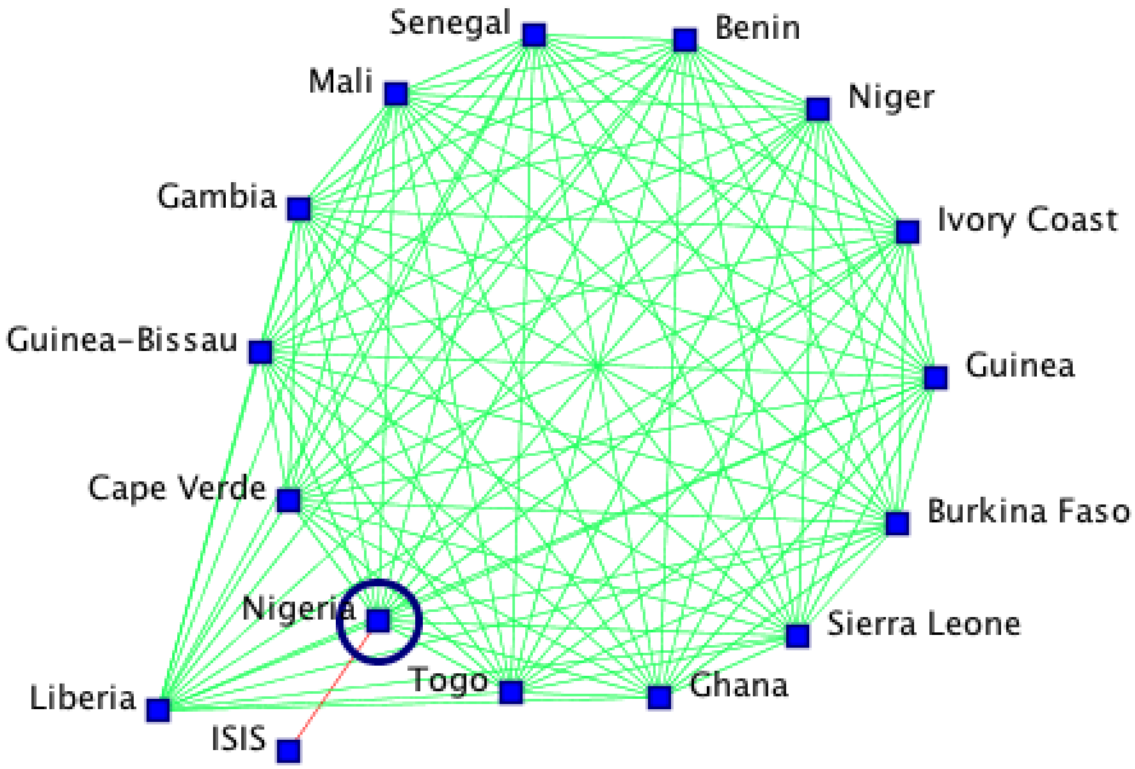


Figure 128. Egonet of Nigeria

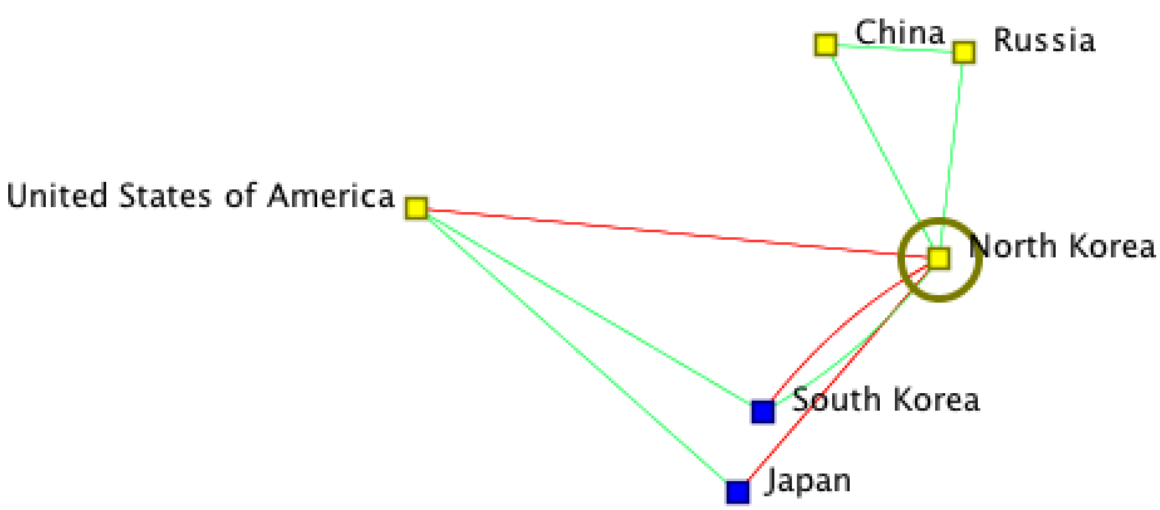


Figure 129. Egonet of North Korea

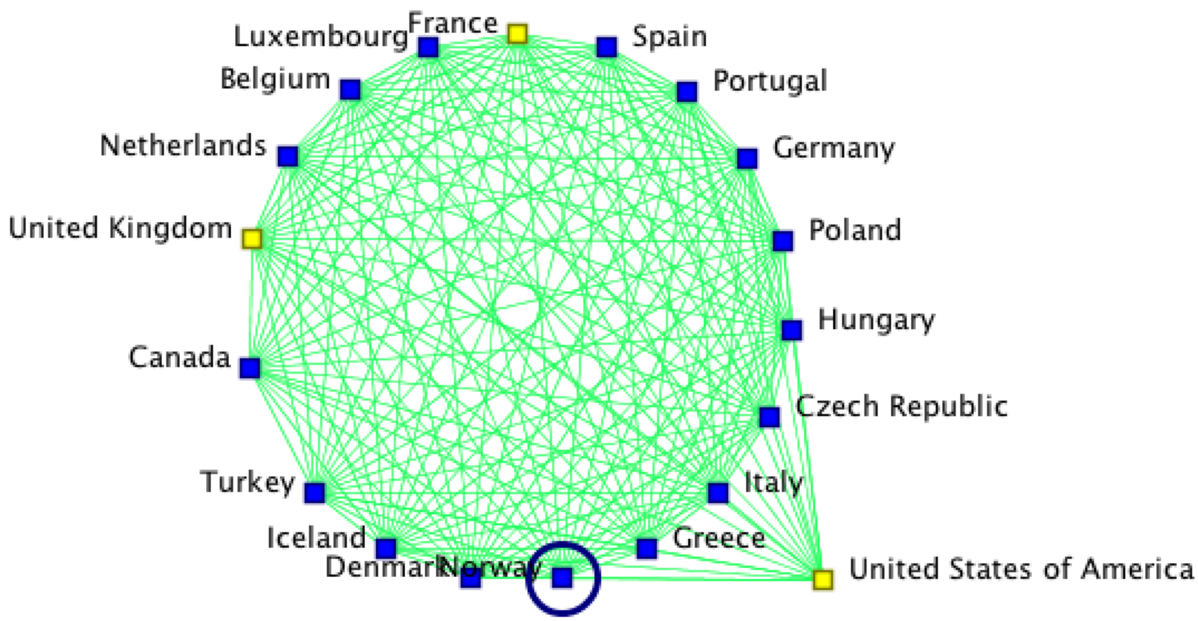


Figure 130. Egonet of Norway

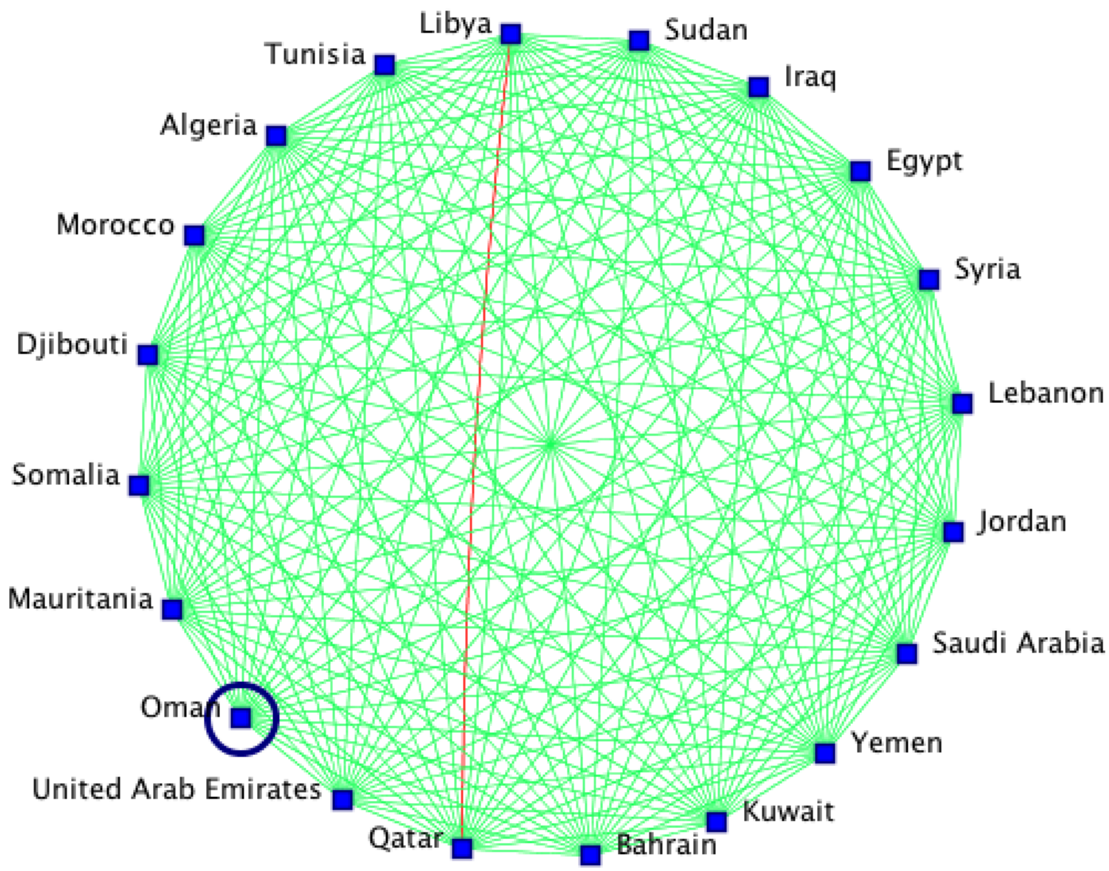


Figure 131. Egonet of Oman

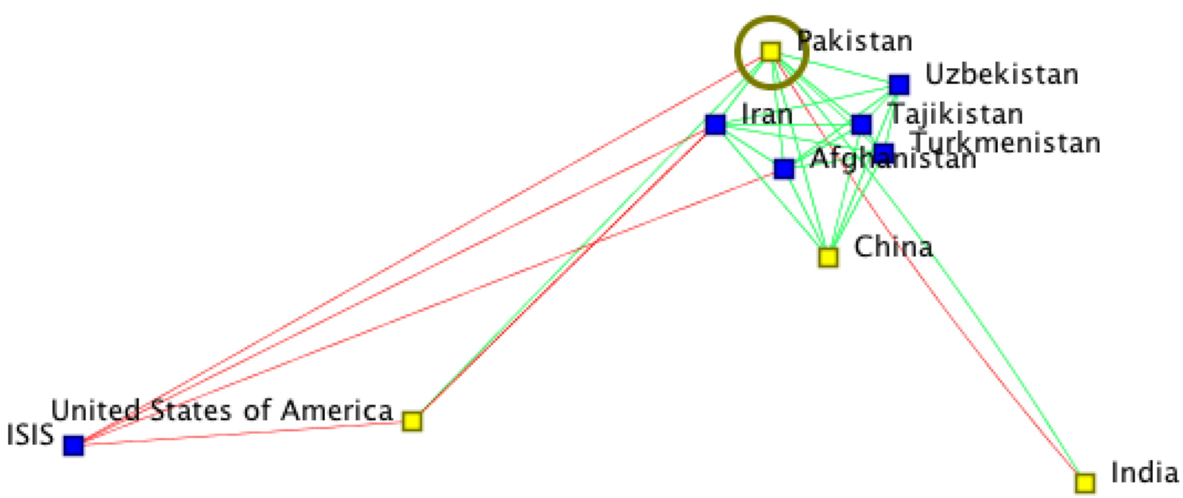


Figure 132. Egonet of Pakistan

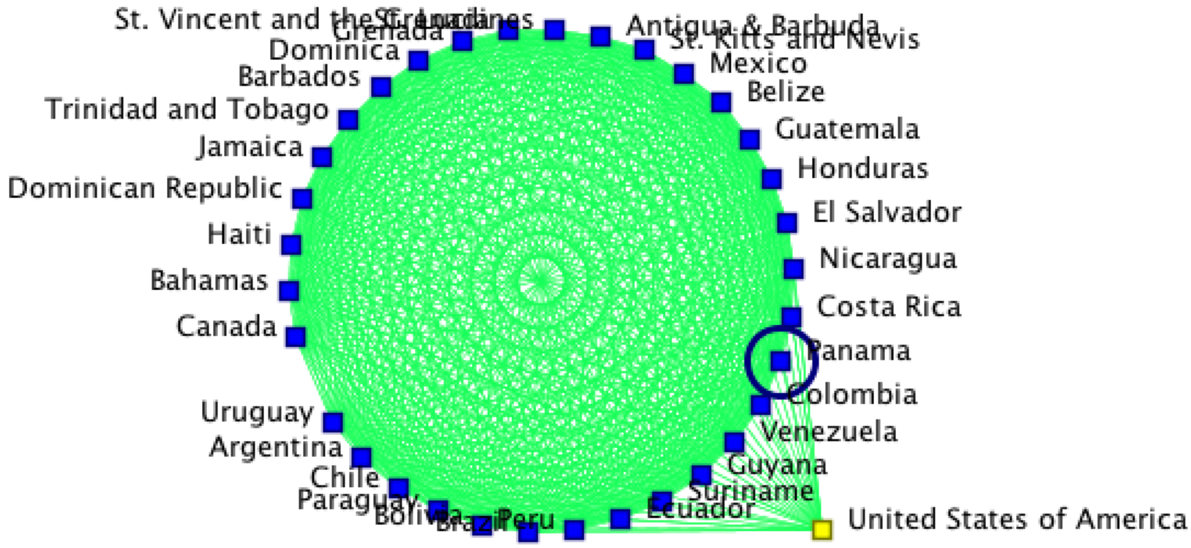


Figure 133. Egonet of Panama

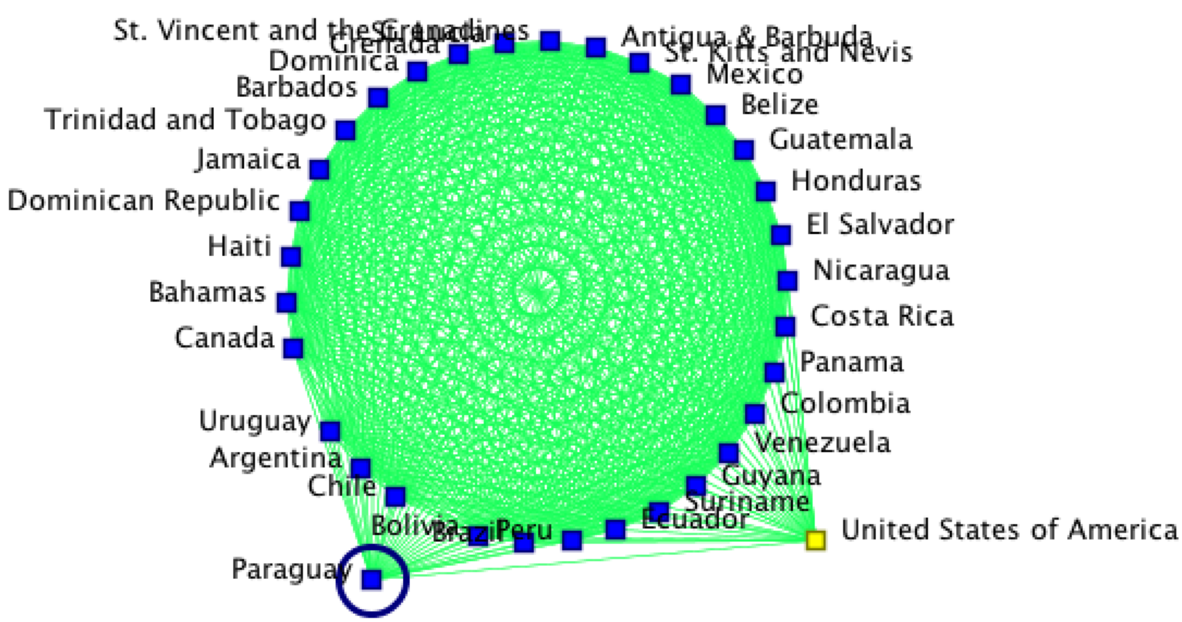


Figure 134. Egonet of Paraguay

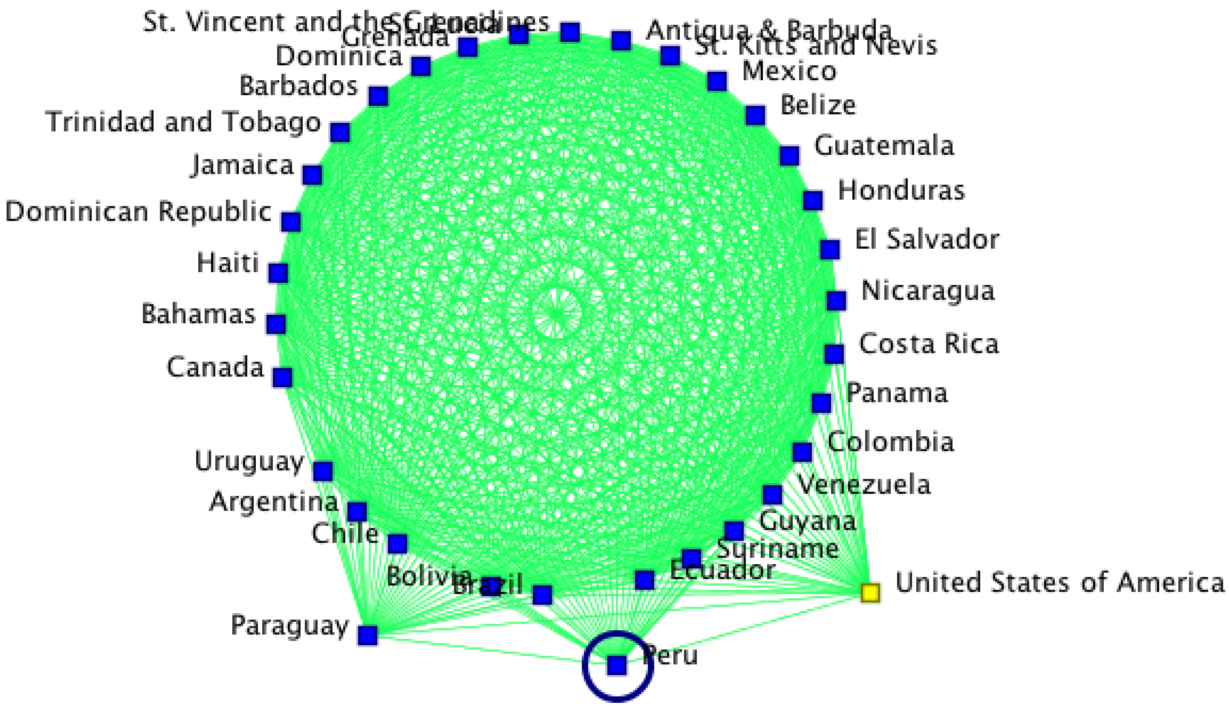


Figure 135. Egonet of Peru

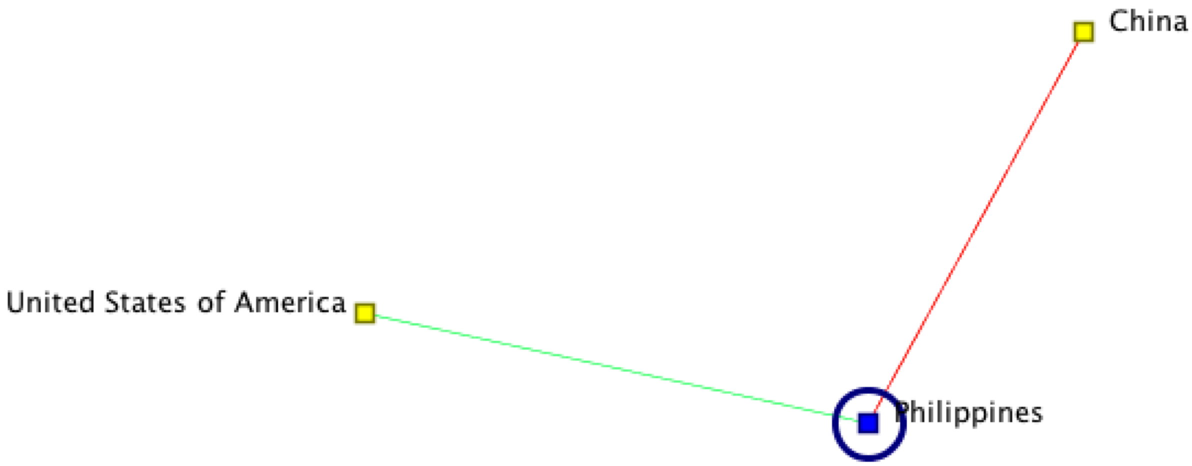


Figure 136. Egonet of Philippines

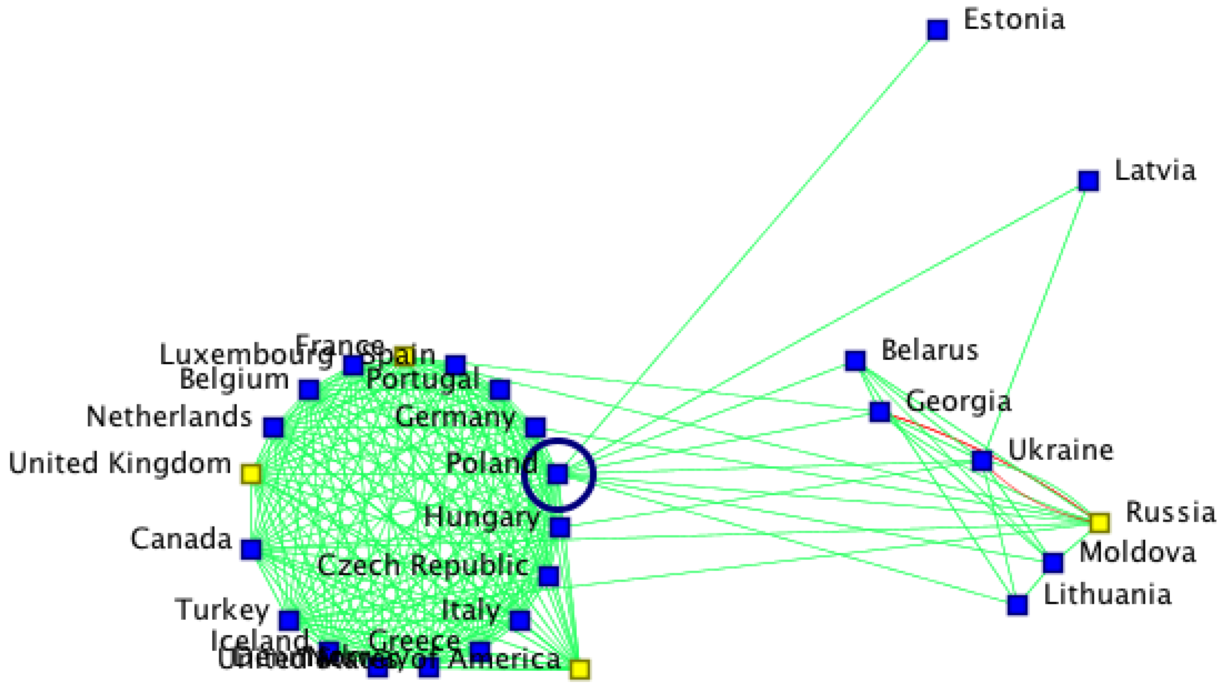


Figure 137. Egonet of Poland

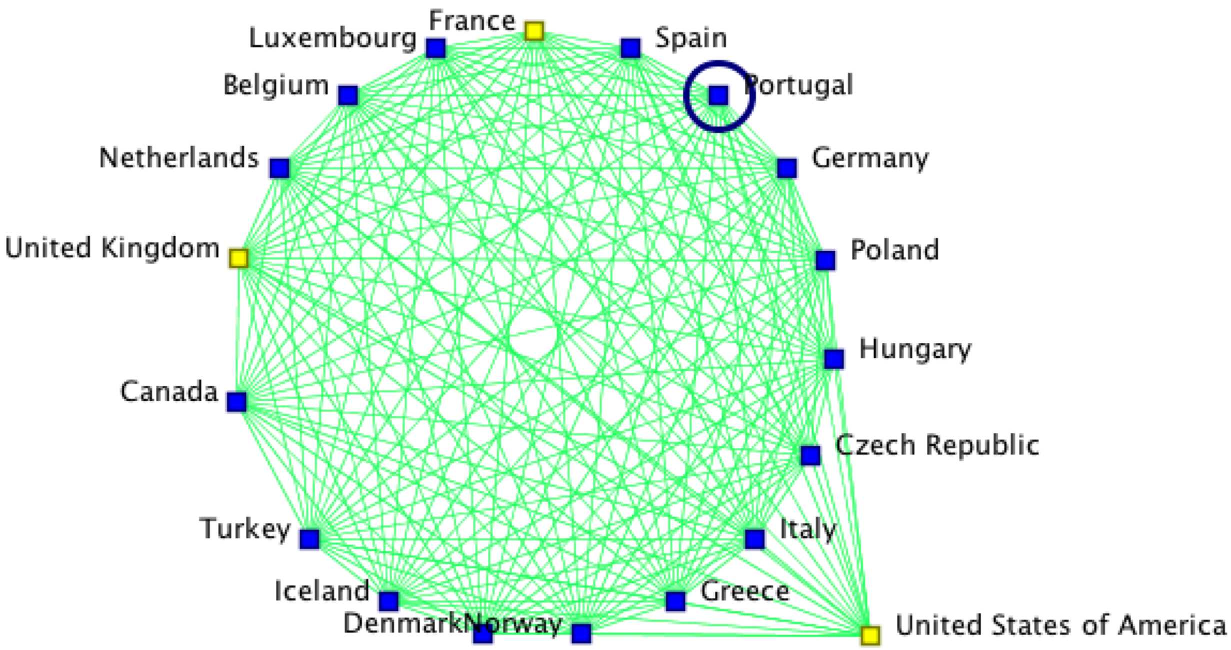


Figure 138. Egonet of Portugal

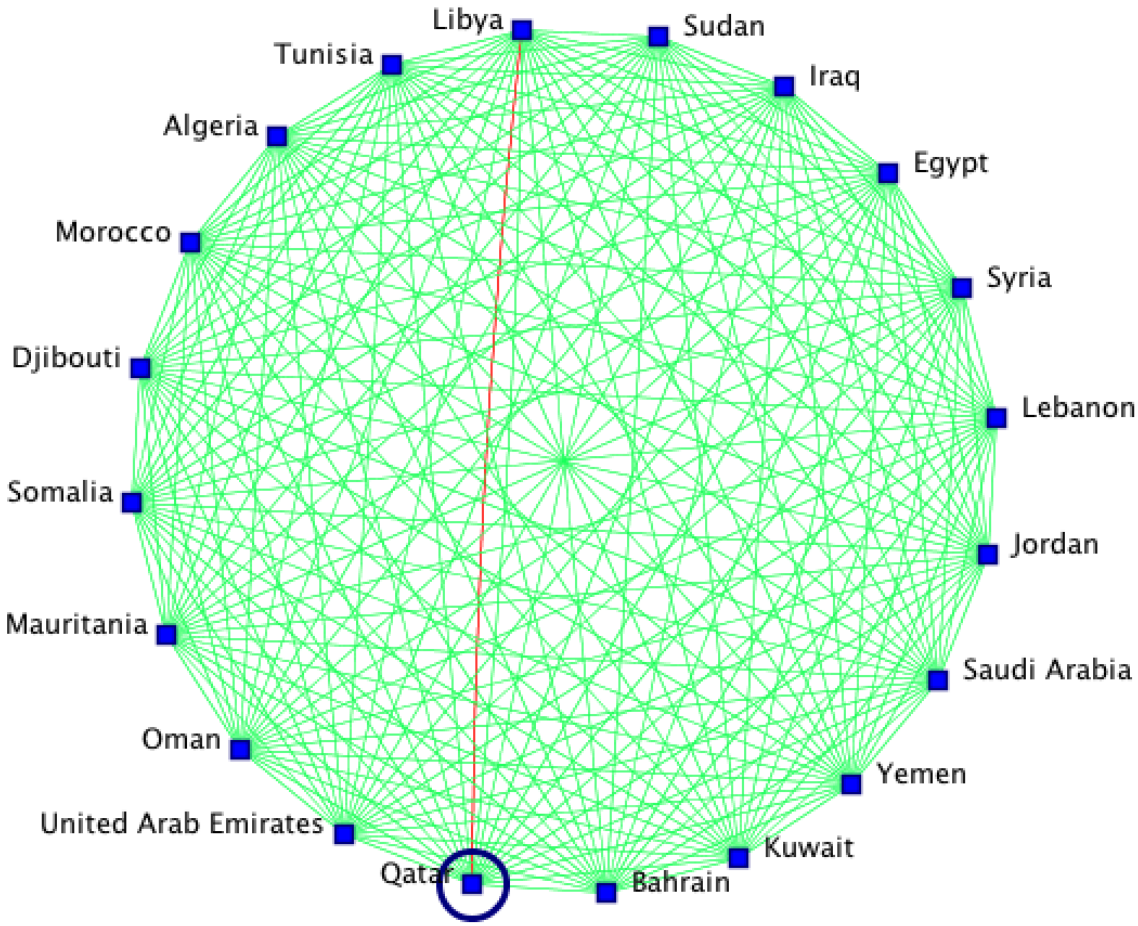


Figure 139. Egonet of Qatar

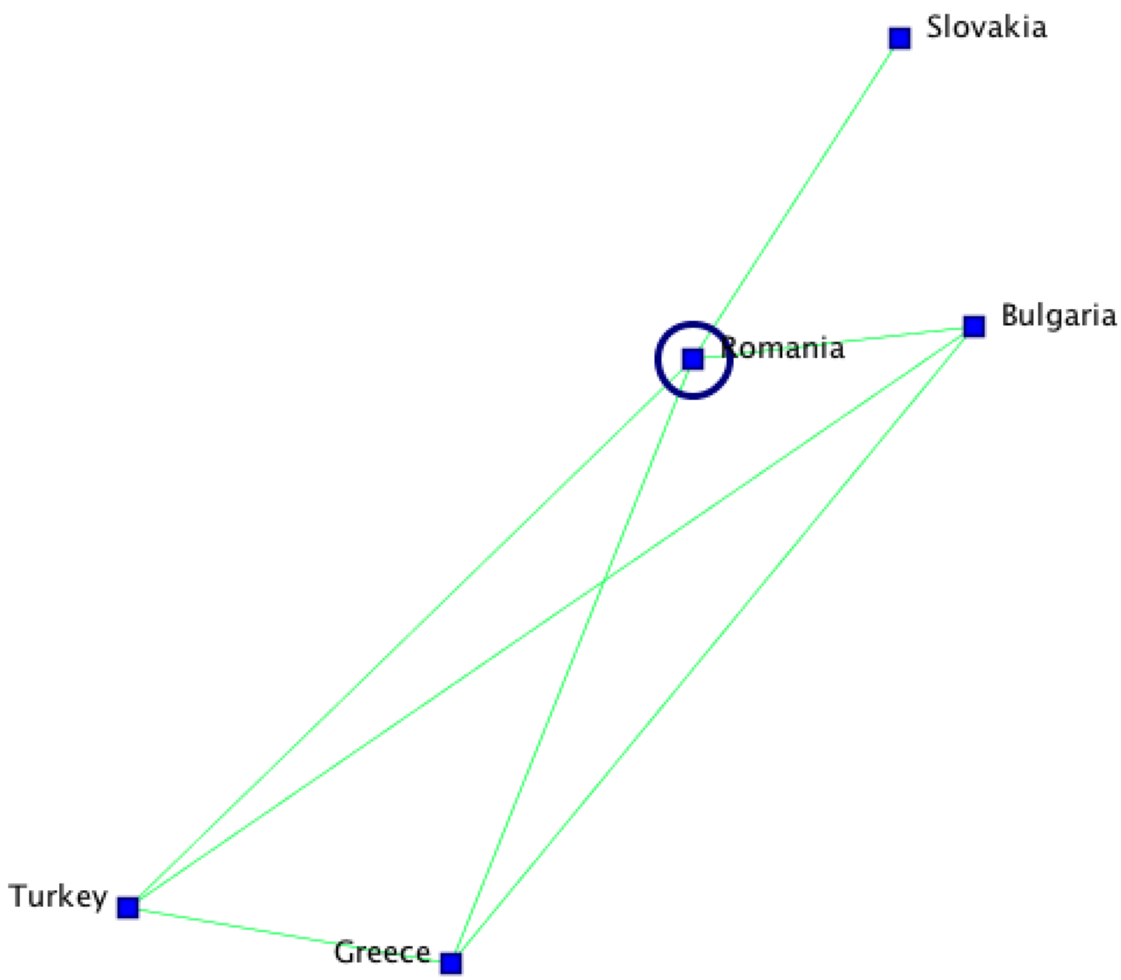


Figure 140. Egonet of Romania

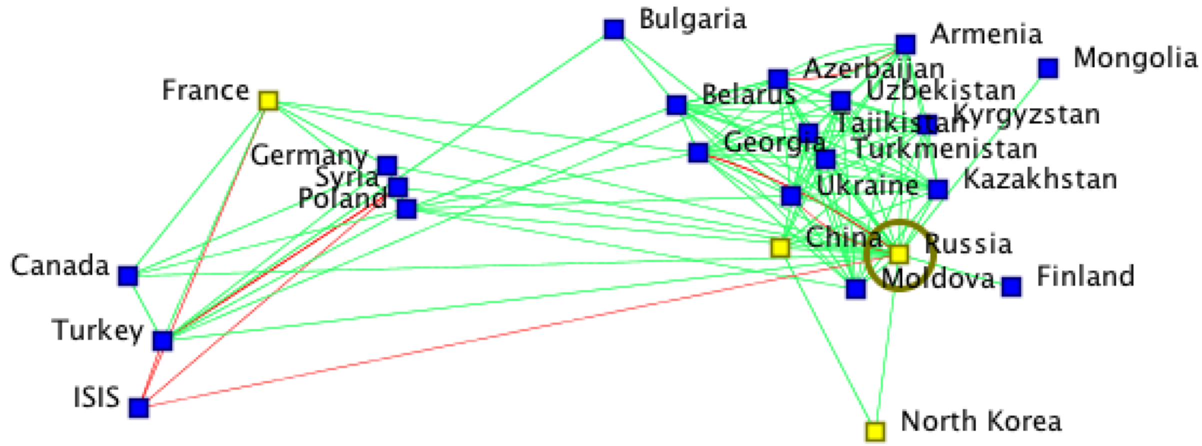


Figure 141. Egonet of Russia

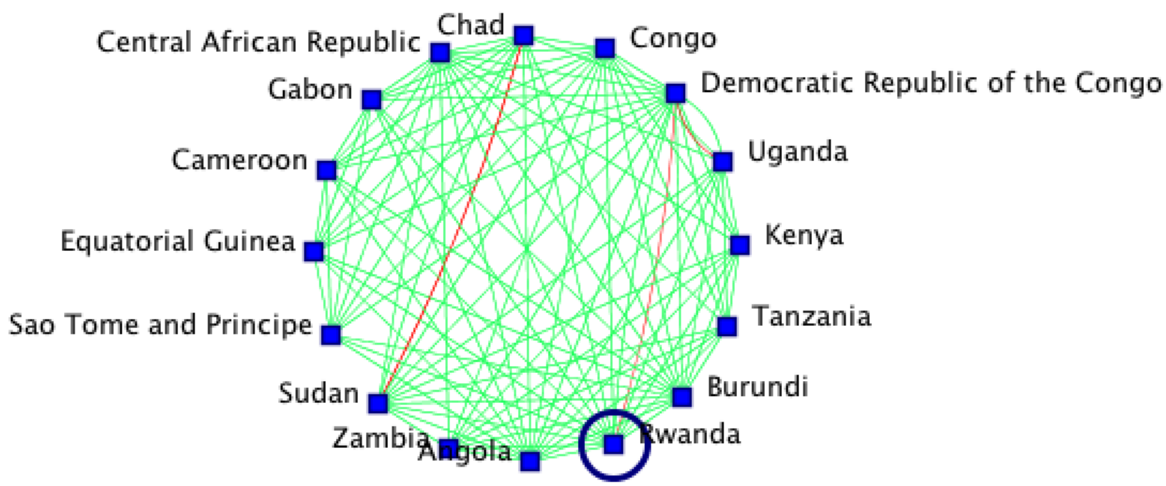


Figure 142. Egonet of Rwanda

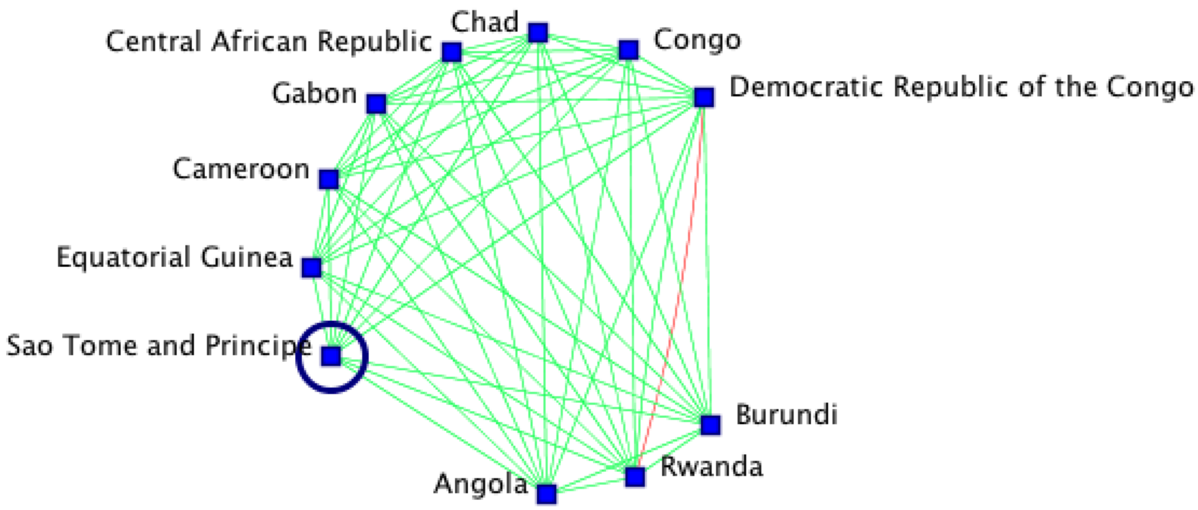


Figure 143. Egonet of Sao Tome

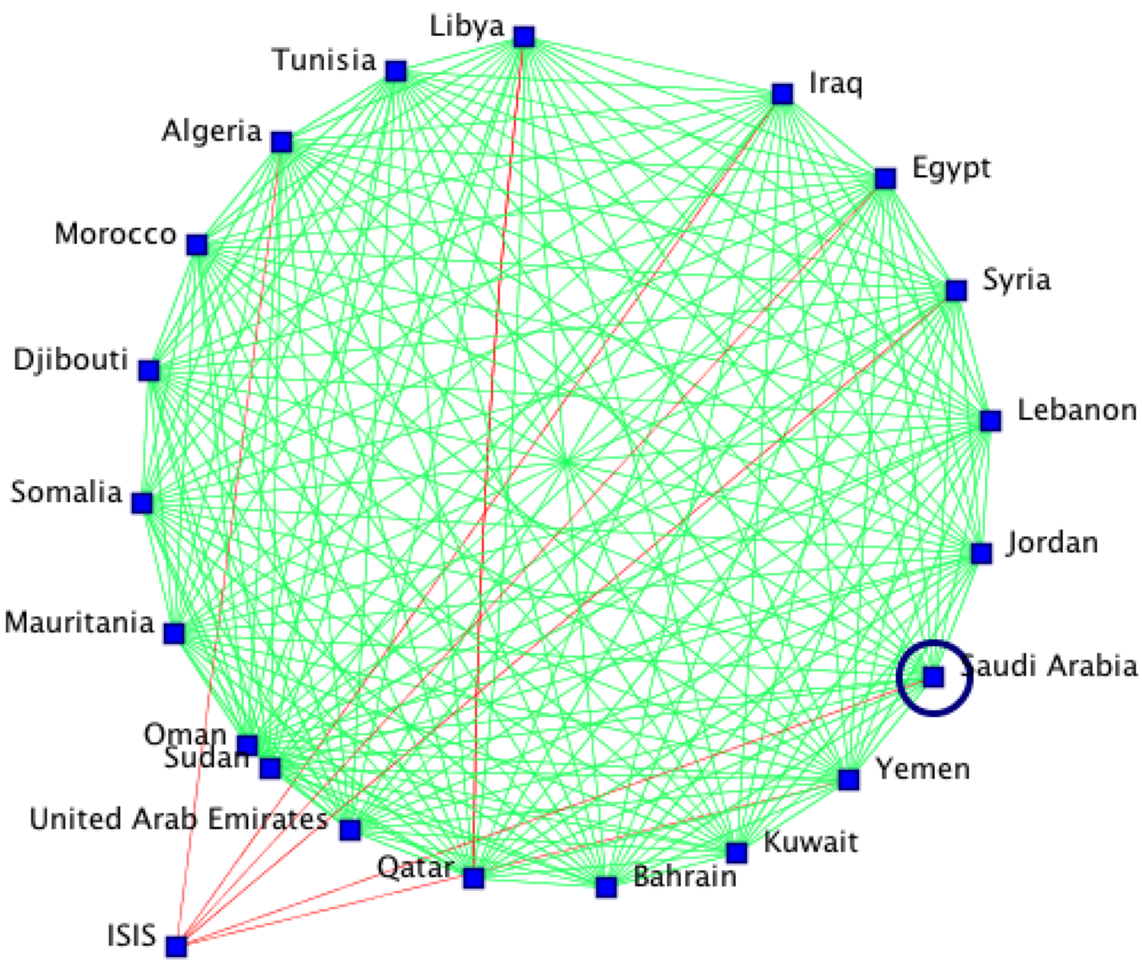


Figure 144. Egonet of Saudi Arabia

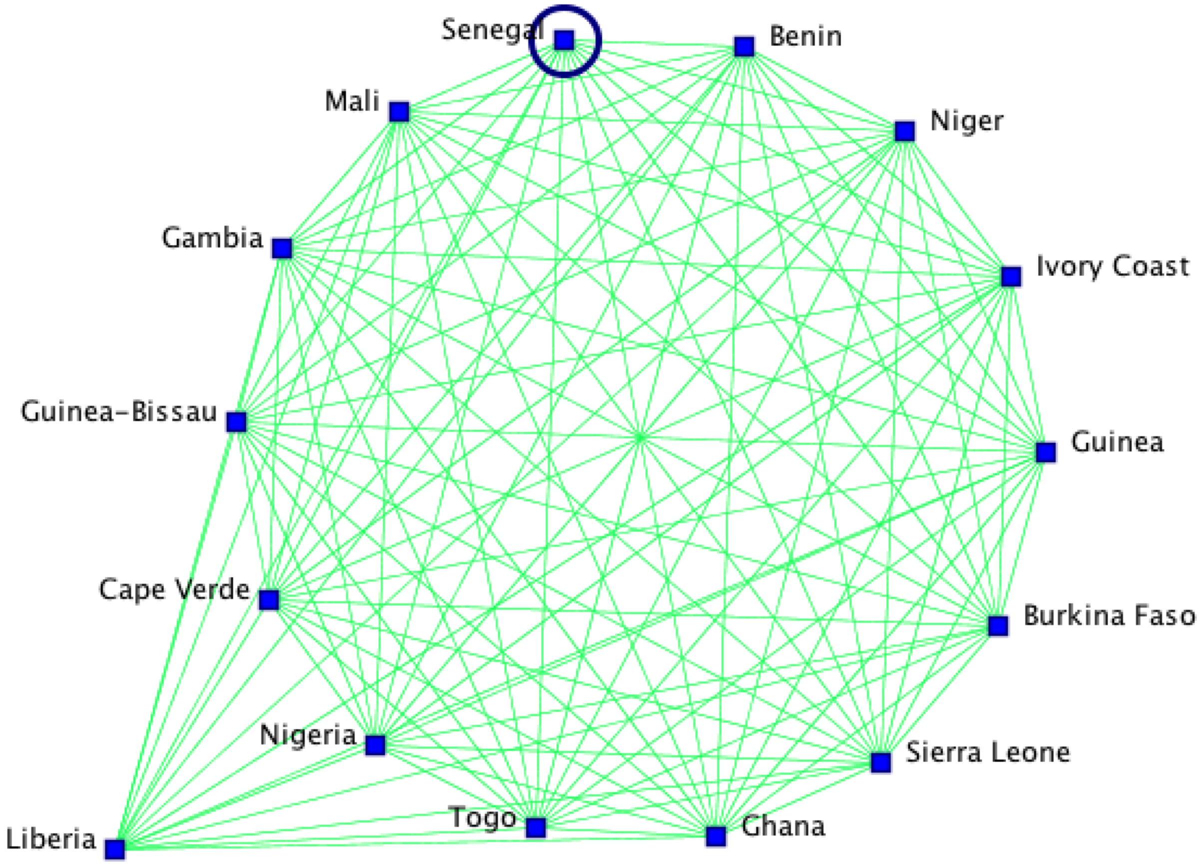


Figure 145. Egonet of Senegal

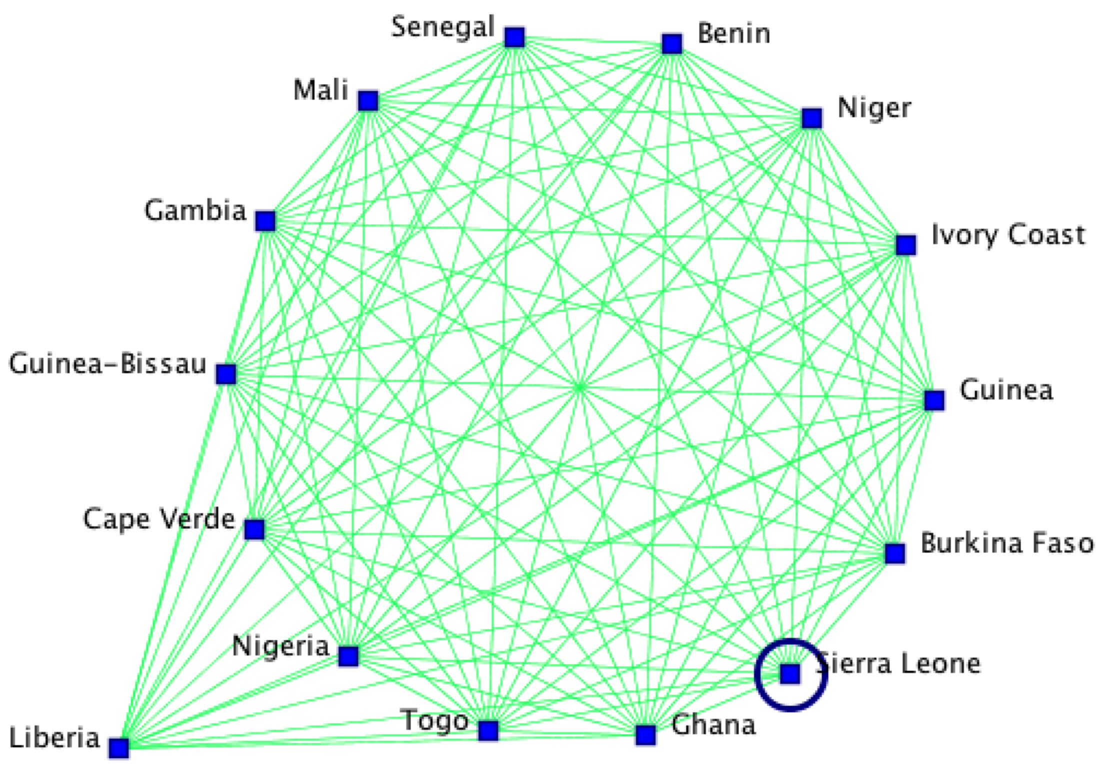


Figure 146. Egonet of Sierra Leone

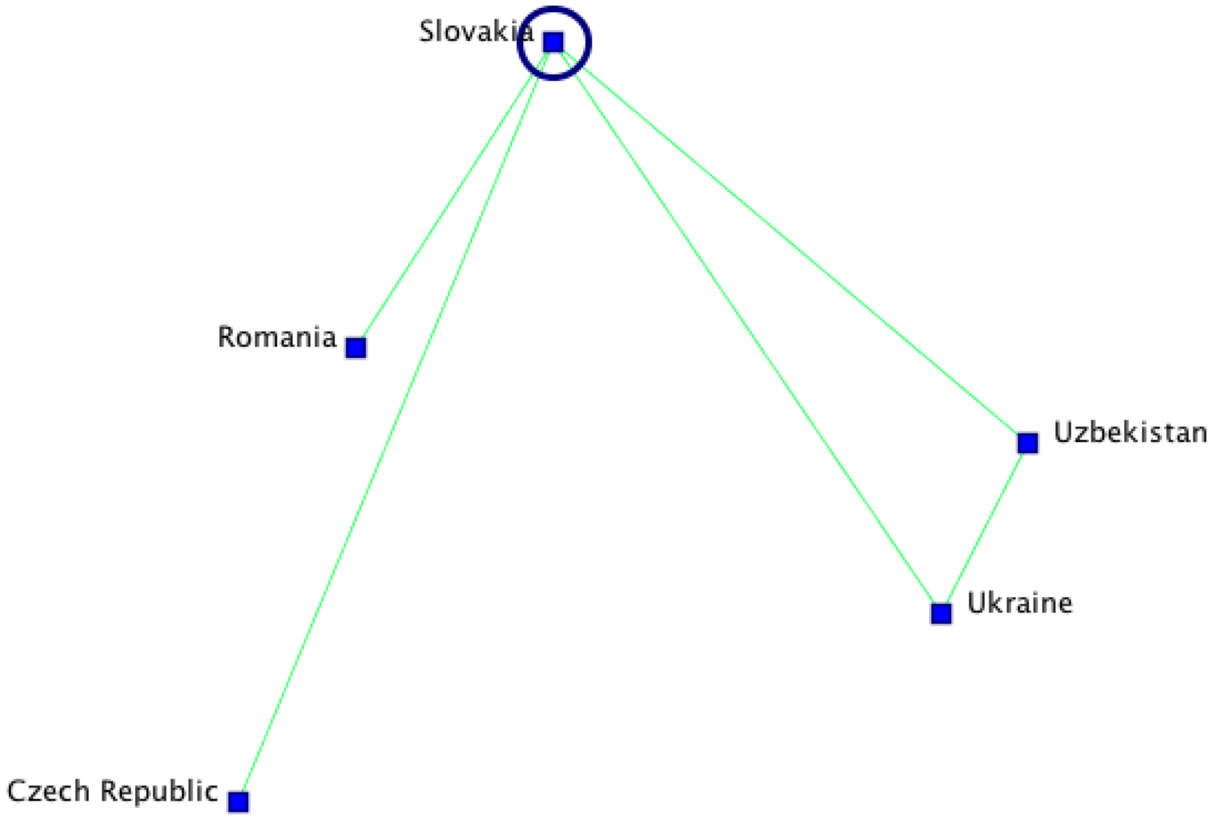


Figure 147. Egonet of Slovakia

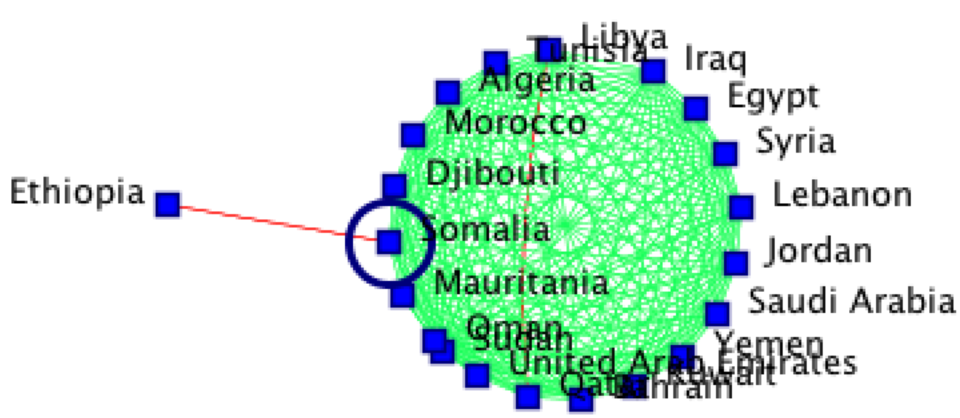


Figure 148. Egonet of Somalia

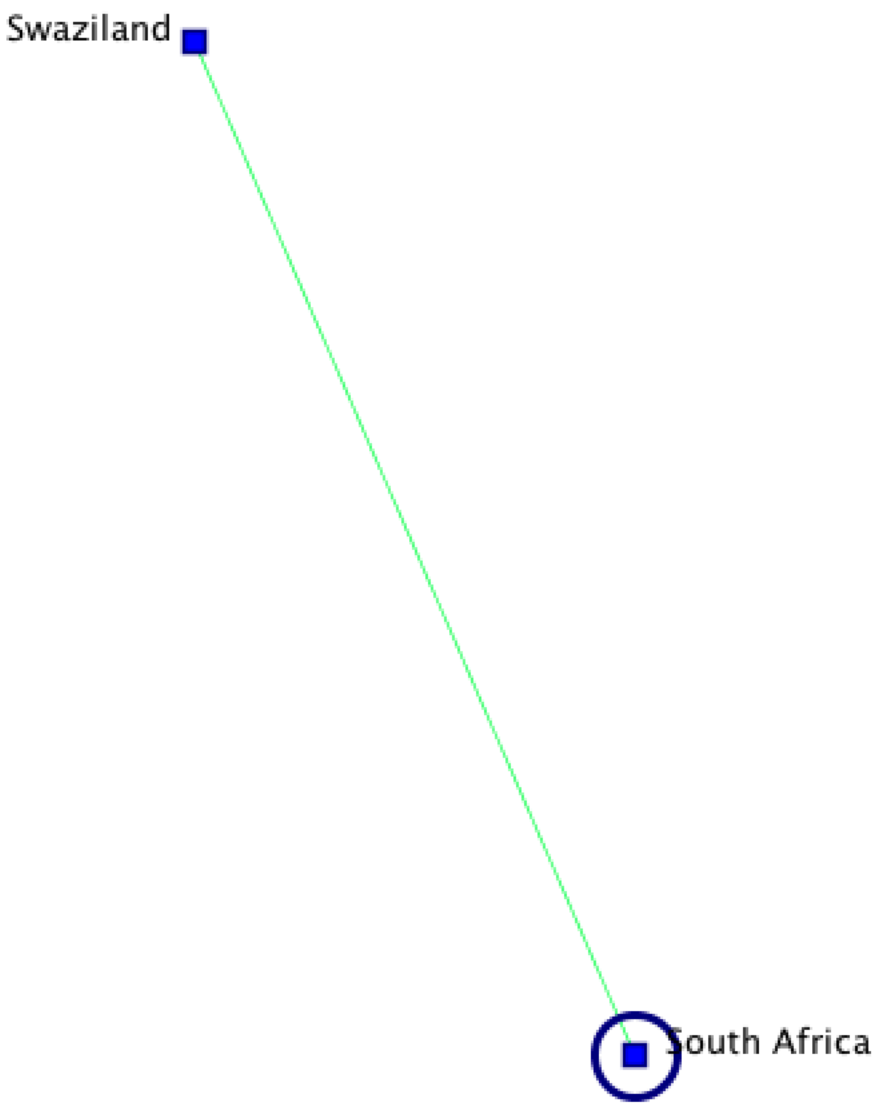


Figure 149. Egonet of South Africa

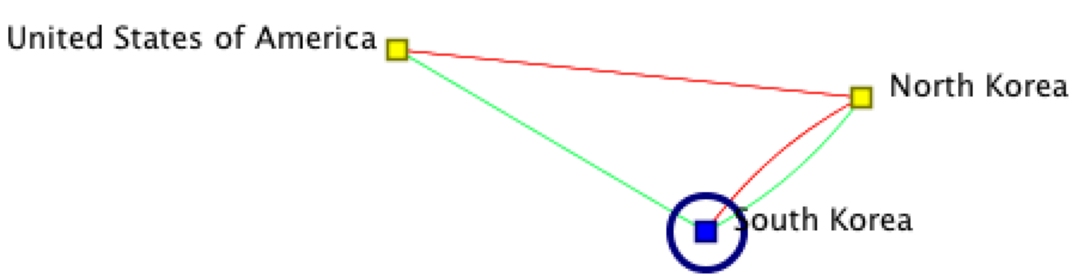


Figure 150. Egonet of South Korea

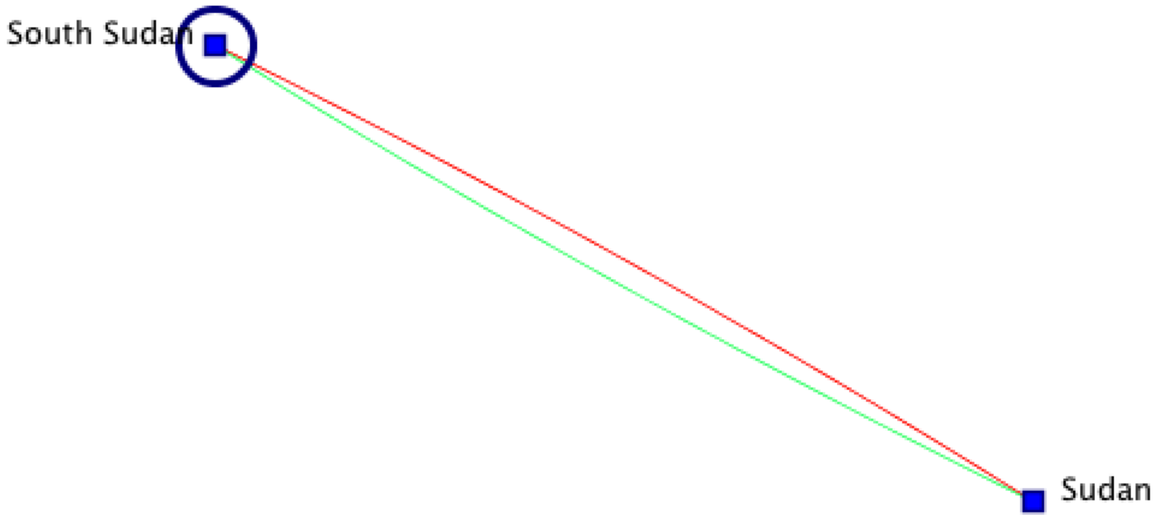


Figure 151. Egonet of South Sudan

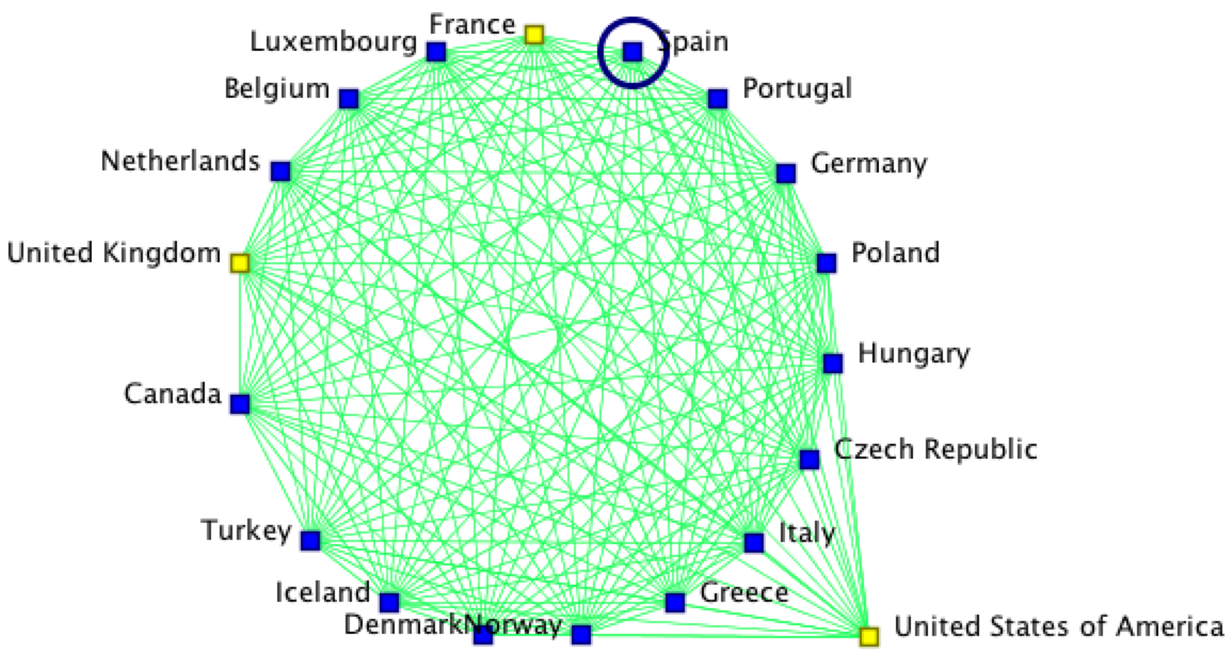


Figure 152. Egonet of Spain

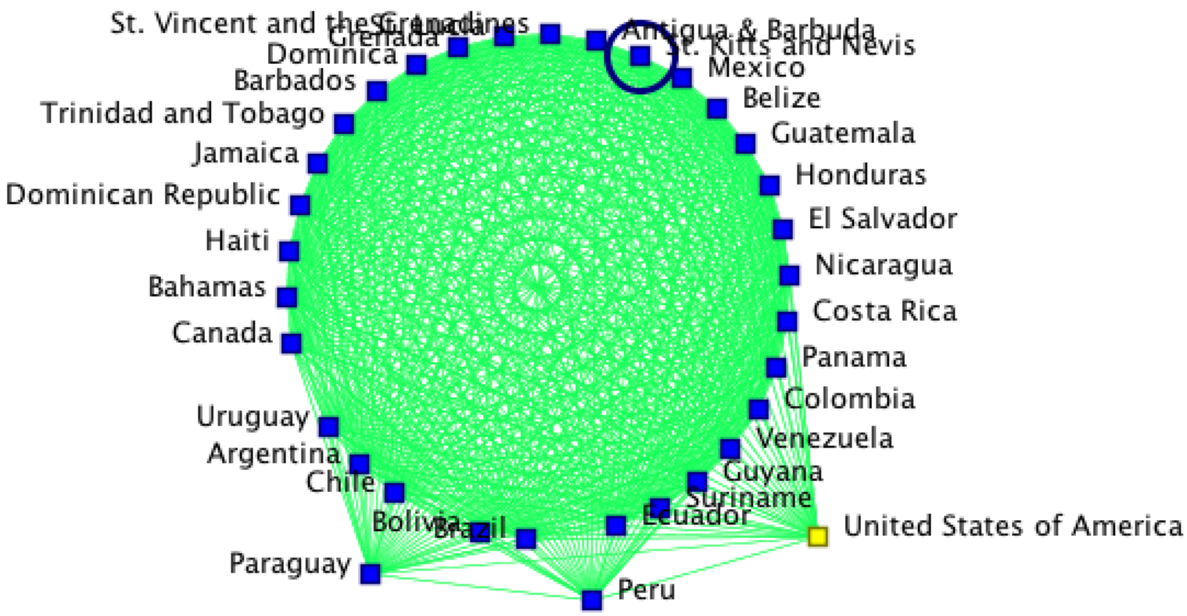


Figure 153. Egonet of St. Kitts

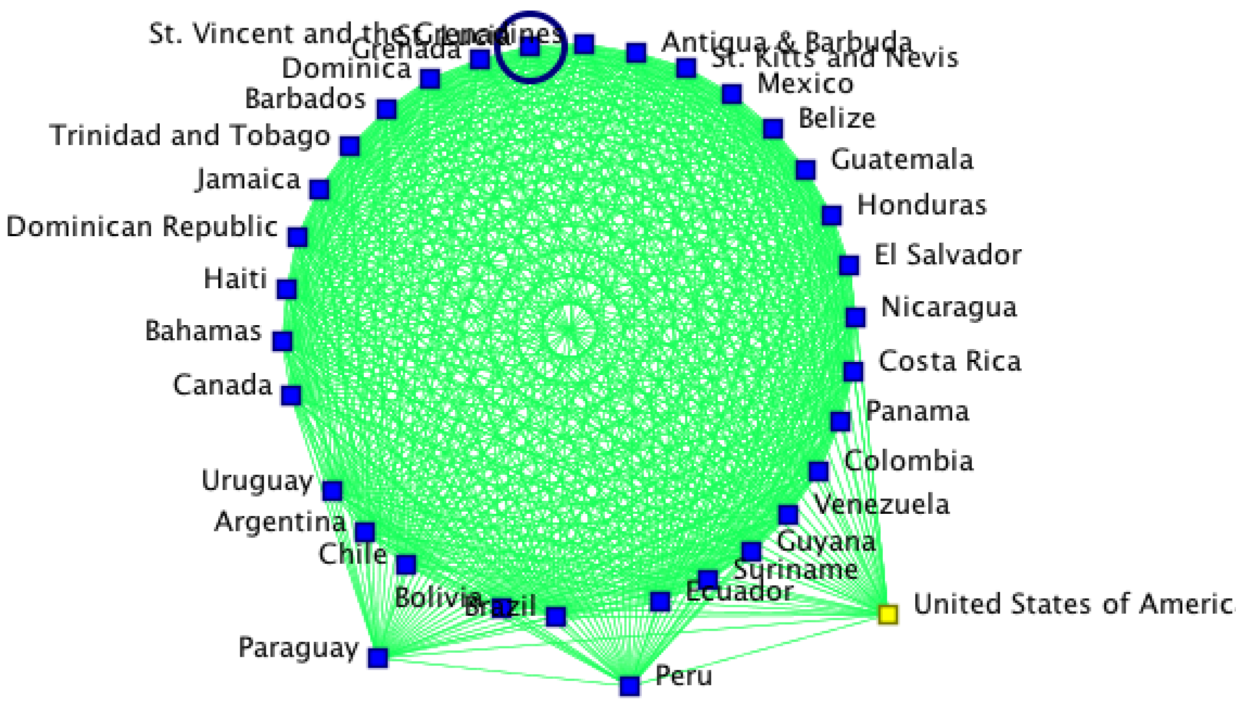


Figure 154. Egonet of St. Lucia

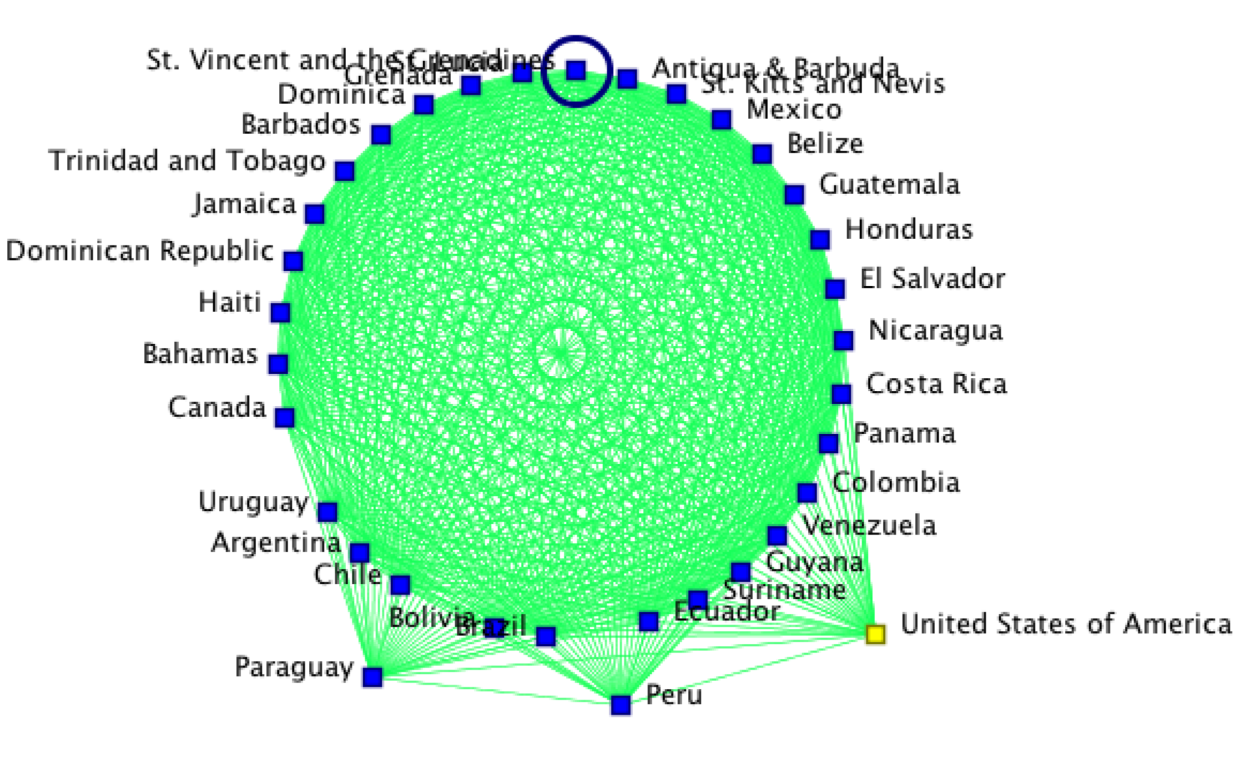


Figure 155. Egonet of St. Vincent

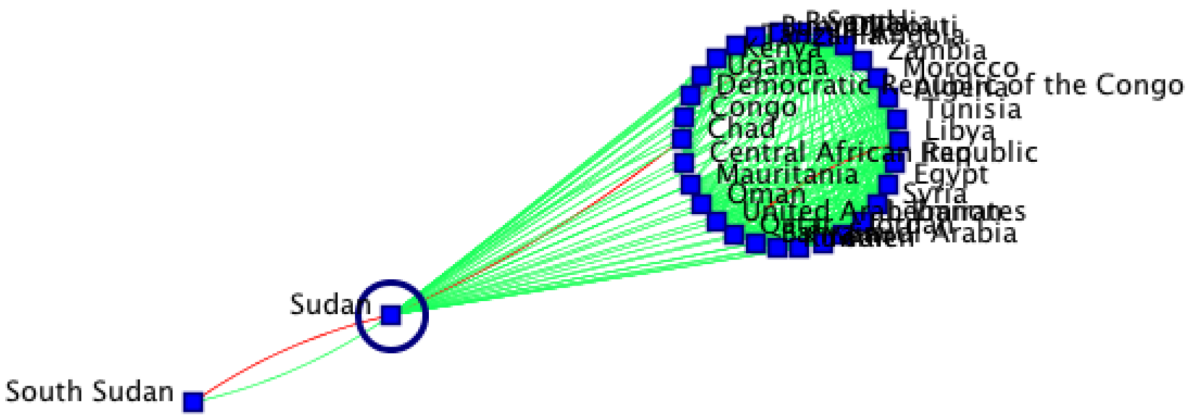


Figure 156. Egonet of Sudan

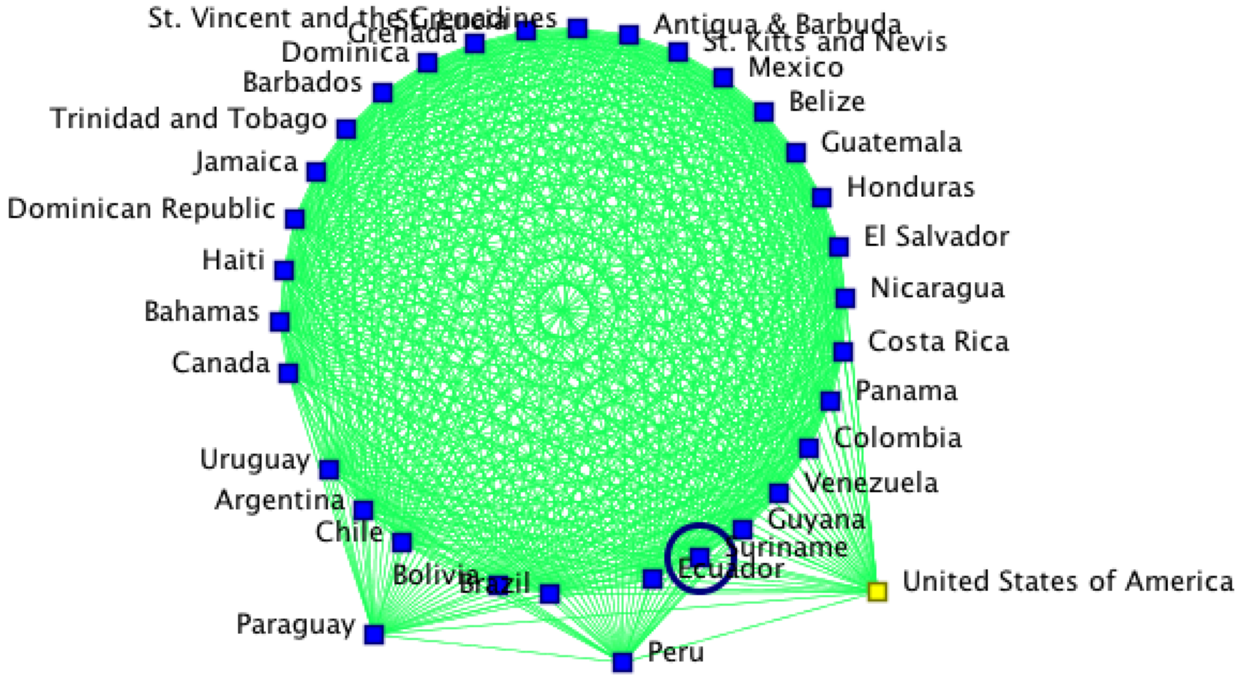


Figure 157. Egonet of Suriname

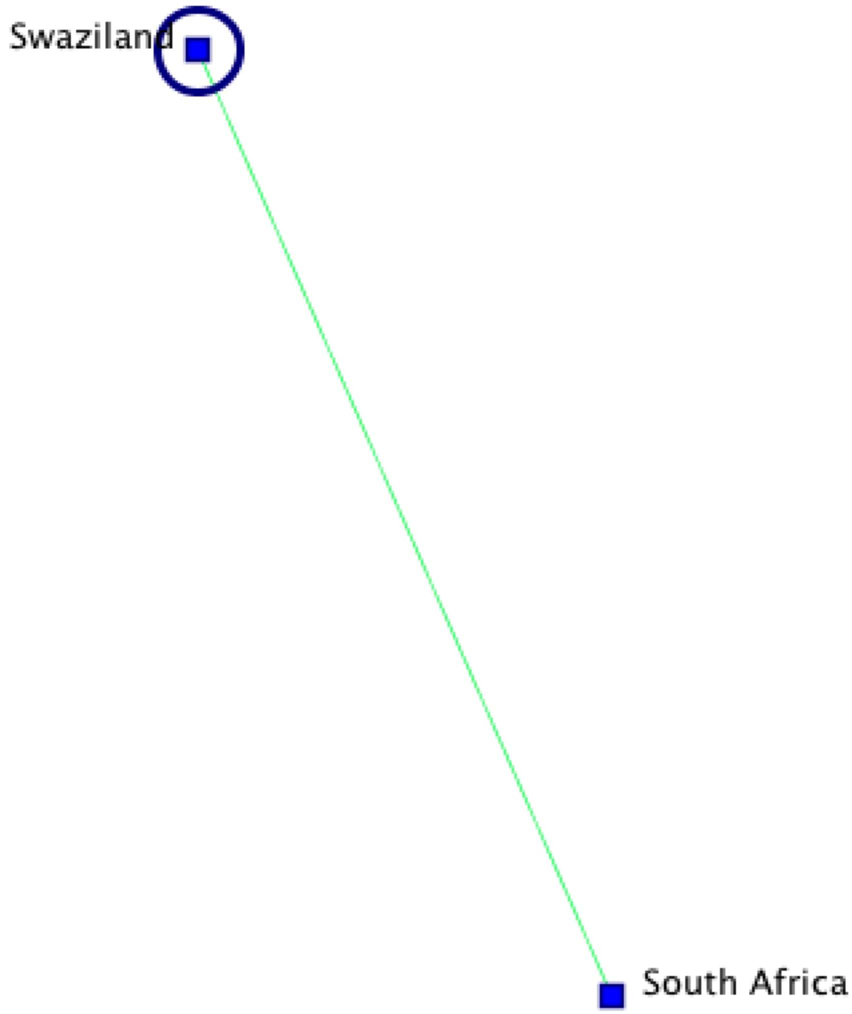


Figure 158. Egonet of Swaziland

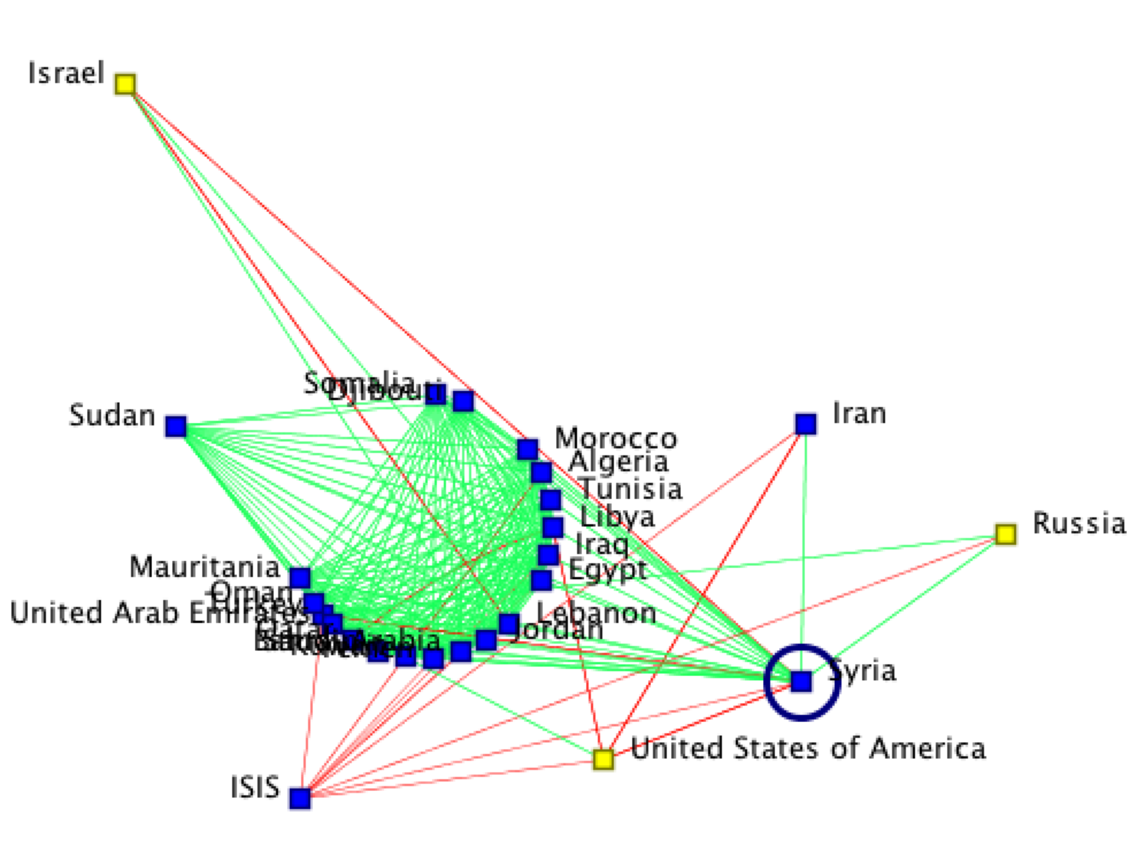


Figure 159. Egonet of Syria

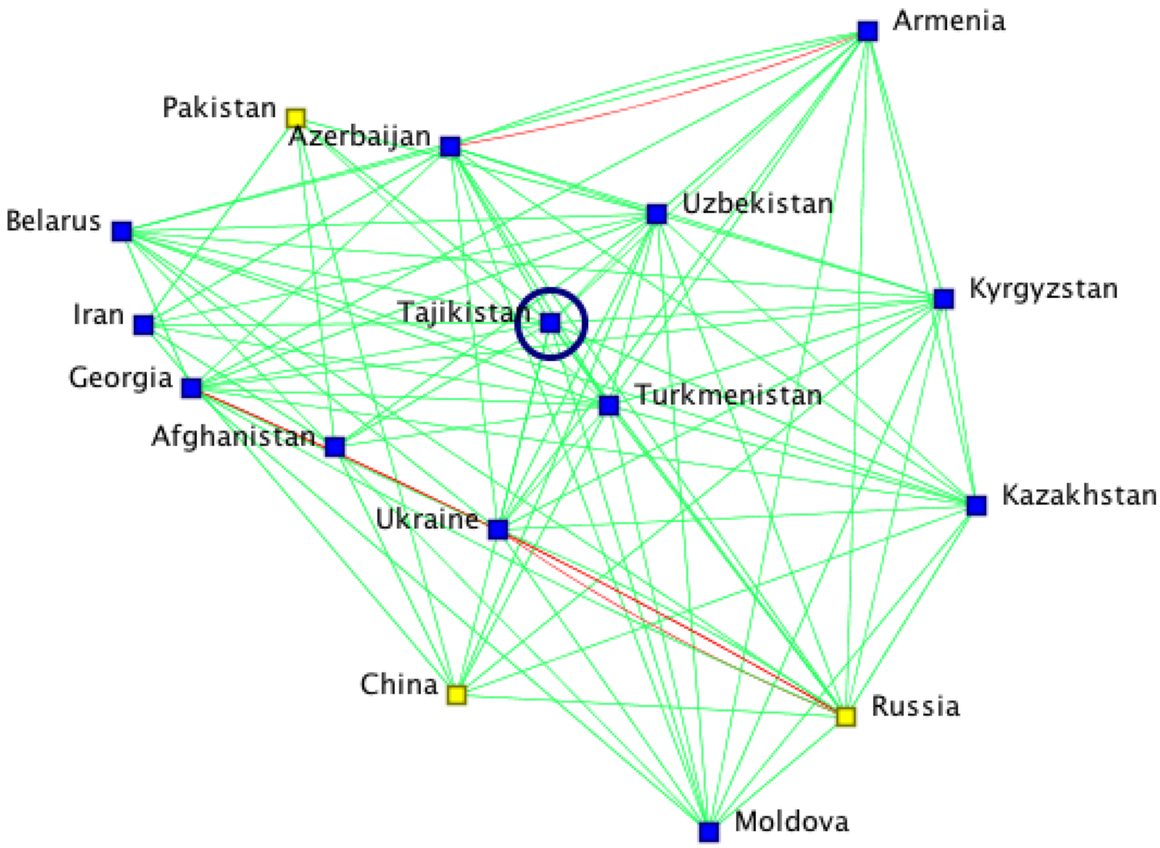


Figure 160. Egonet of Tajikistan

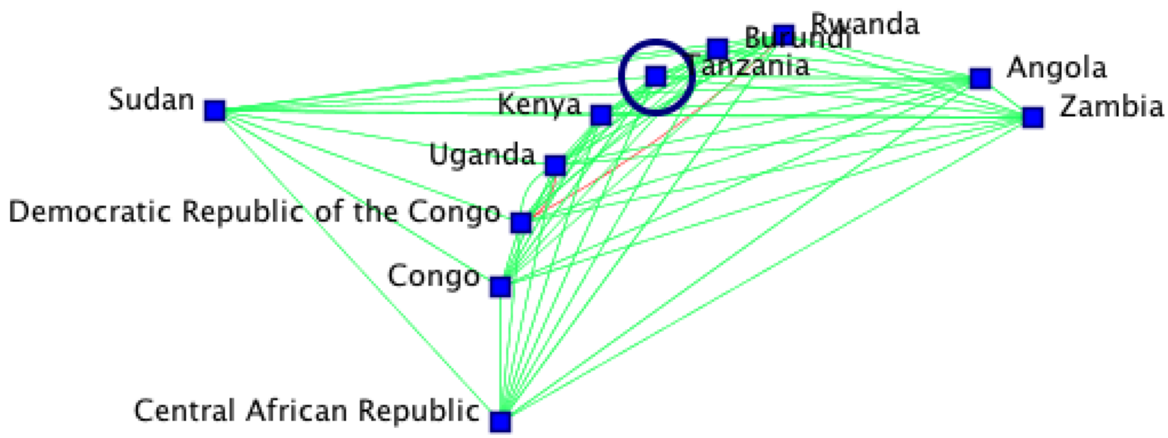


Figure 161. Egonet of Tanazania

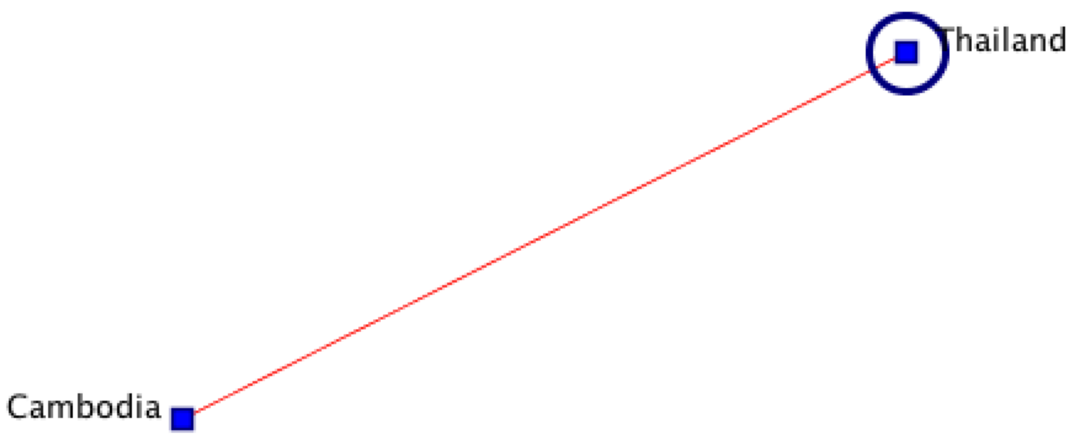


Figure 162. Egonet of Thailand

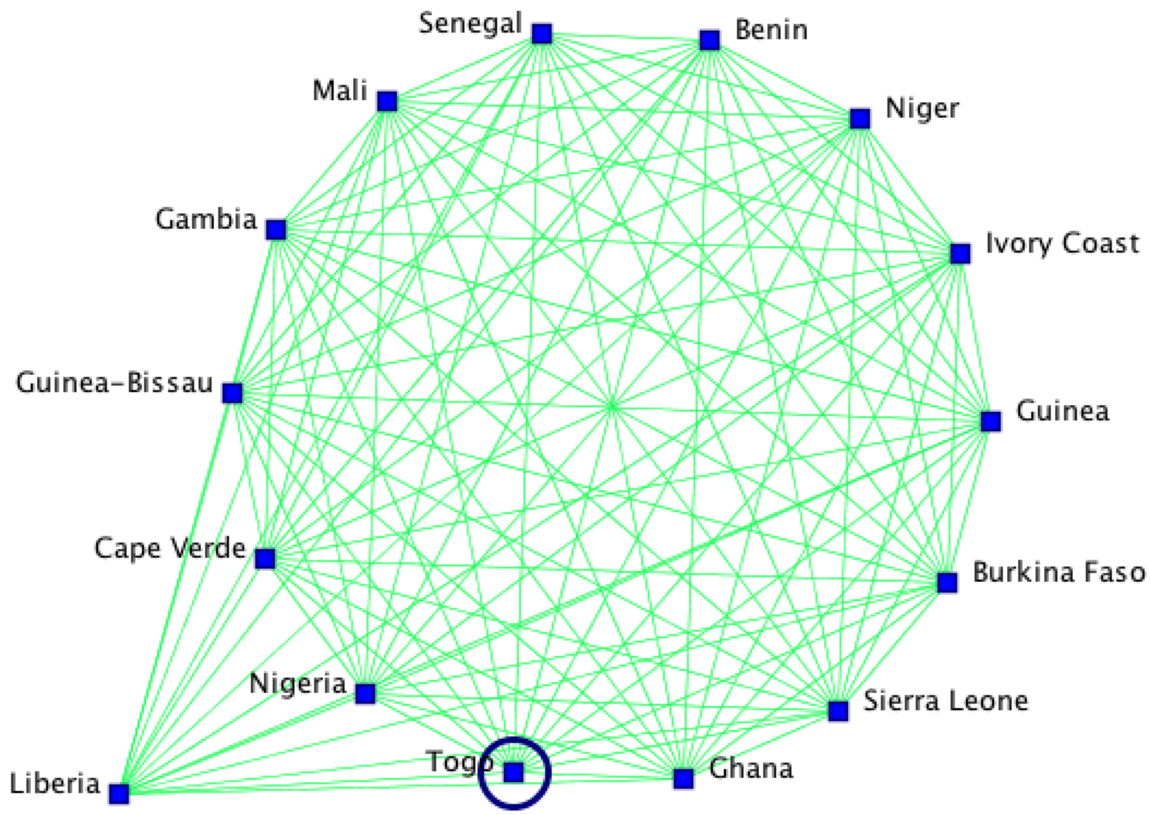


Figure 163. Egonet of Togo

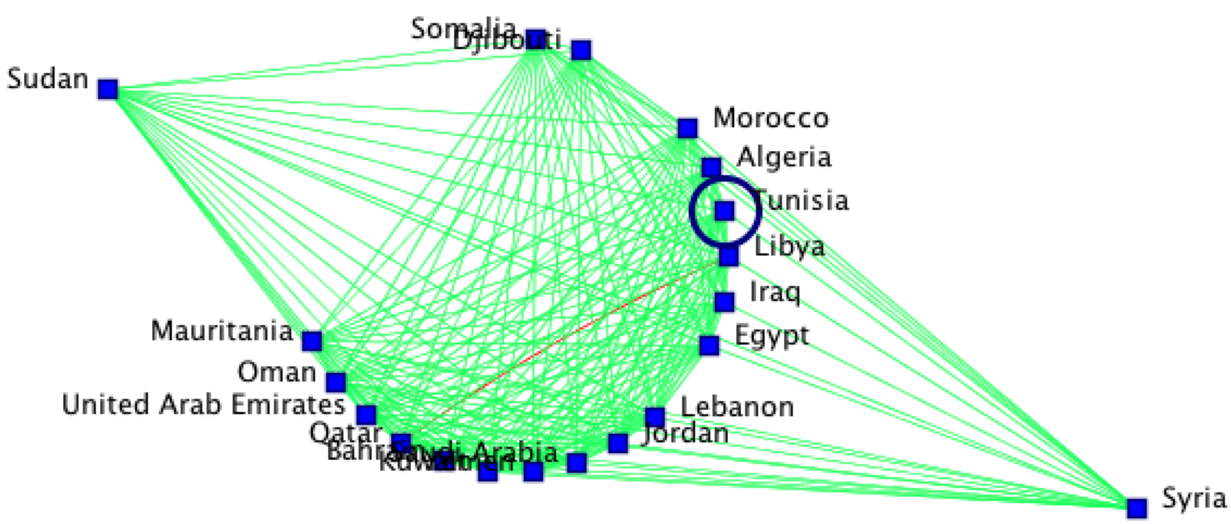


Figure 164. Egonet of Tunisia

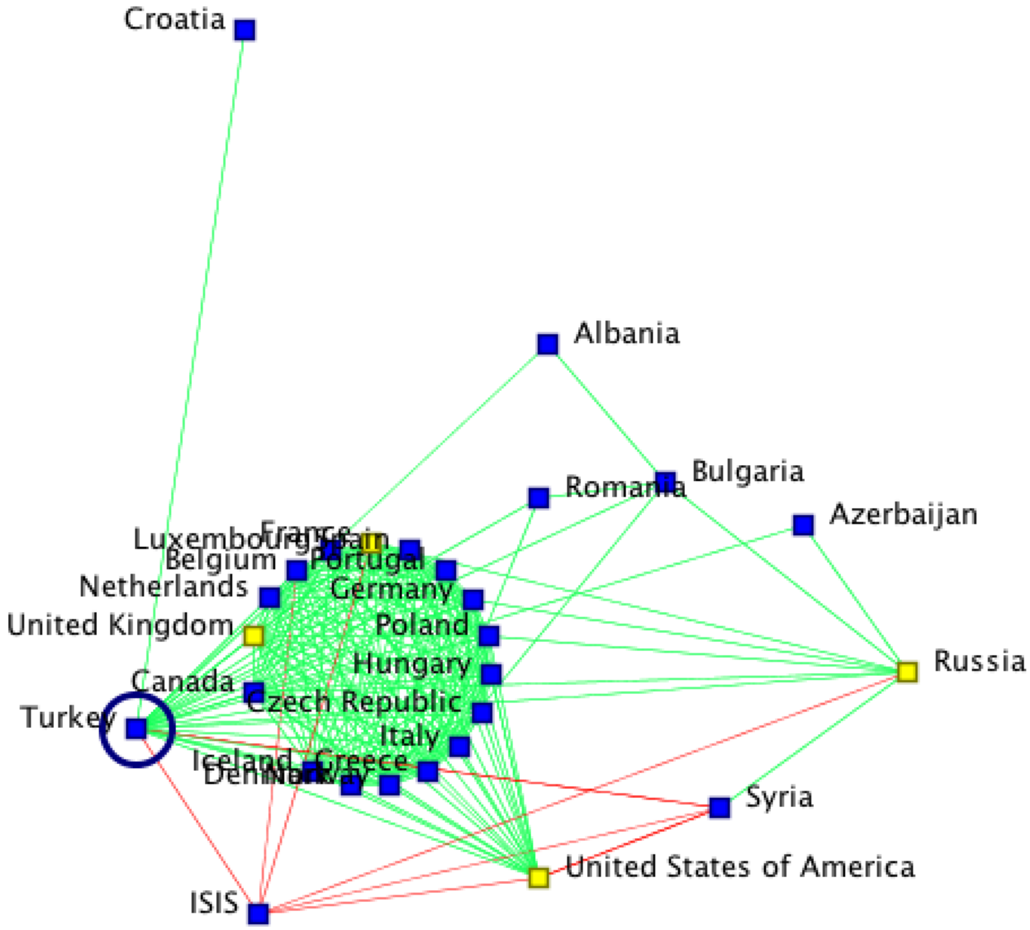


Figure 165. Egonet of Turkey

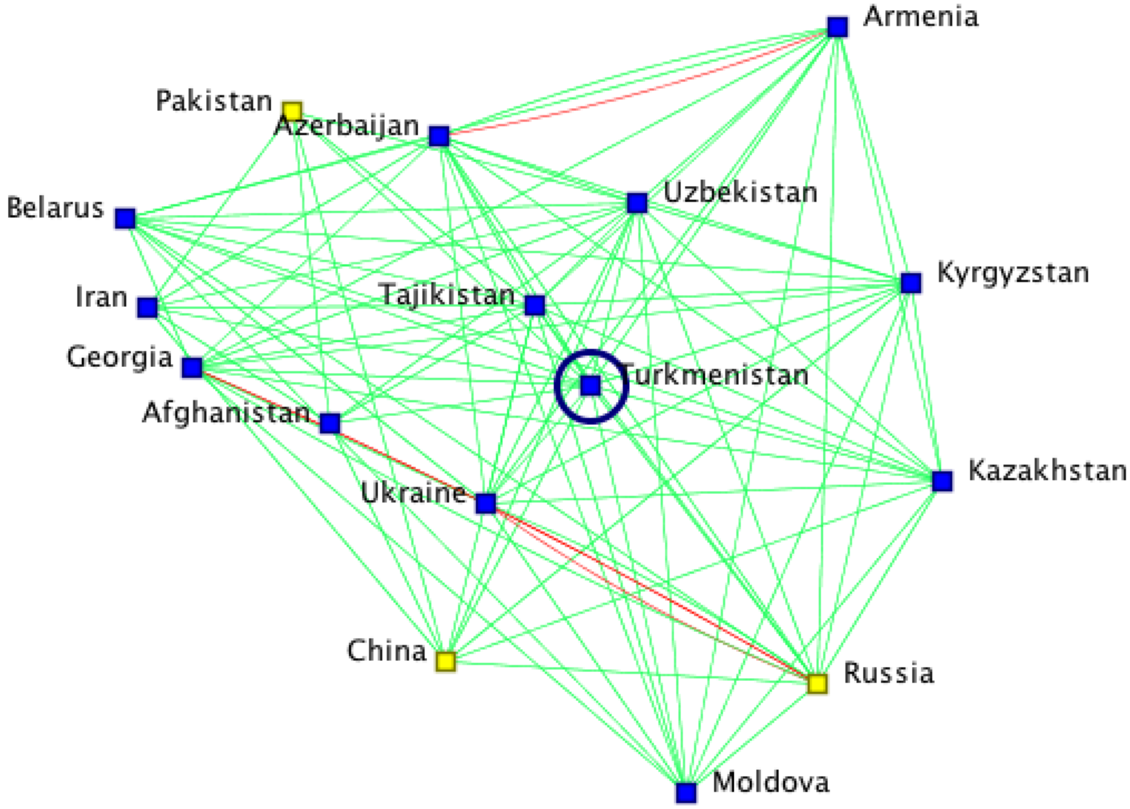


Figure 166. Egonet of Turkmenistan

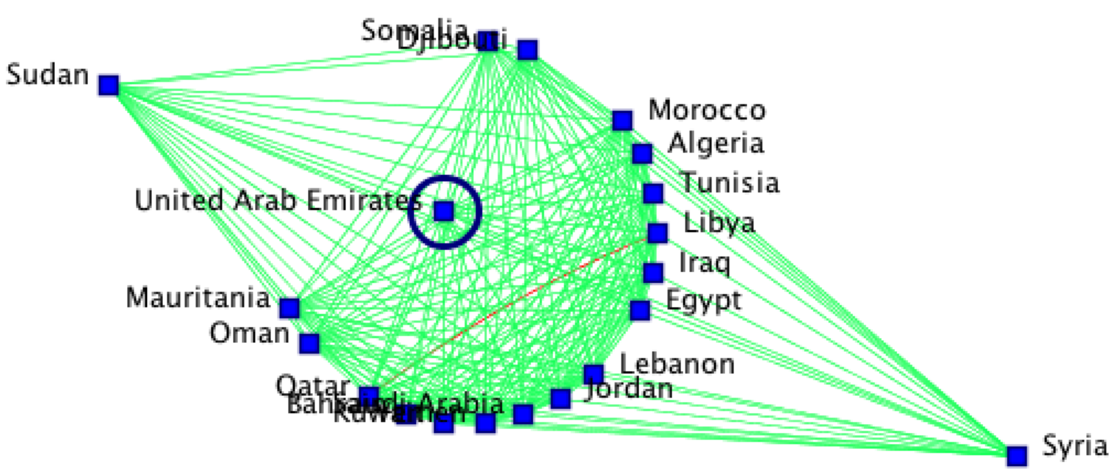


Figure 167. Egonet of UAE

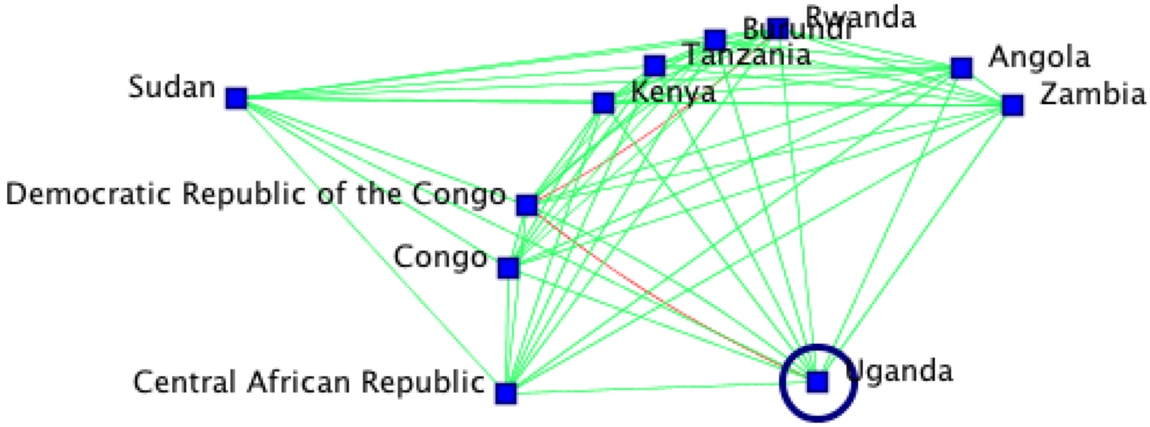


Figure 168. Egonet of Uganda

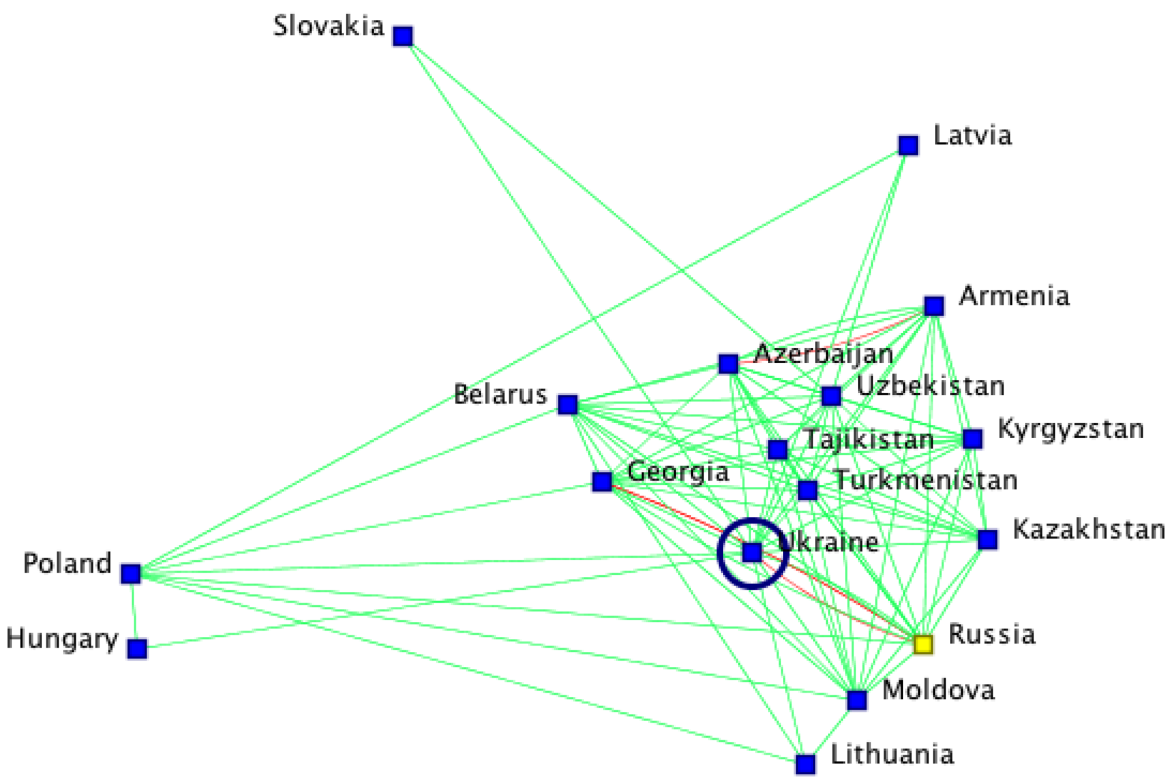


Figure 169. Egonet of Ukraine

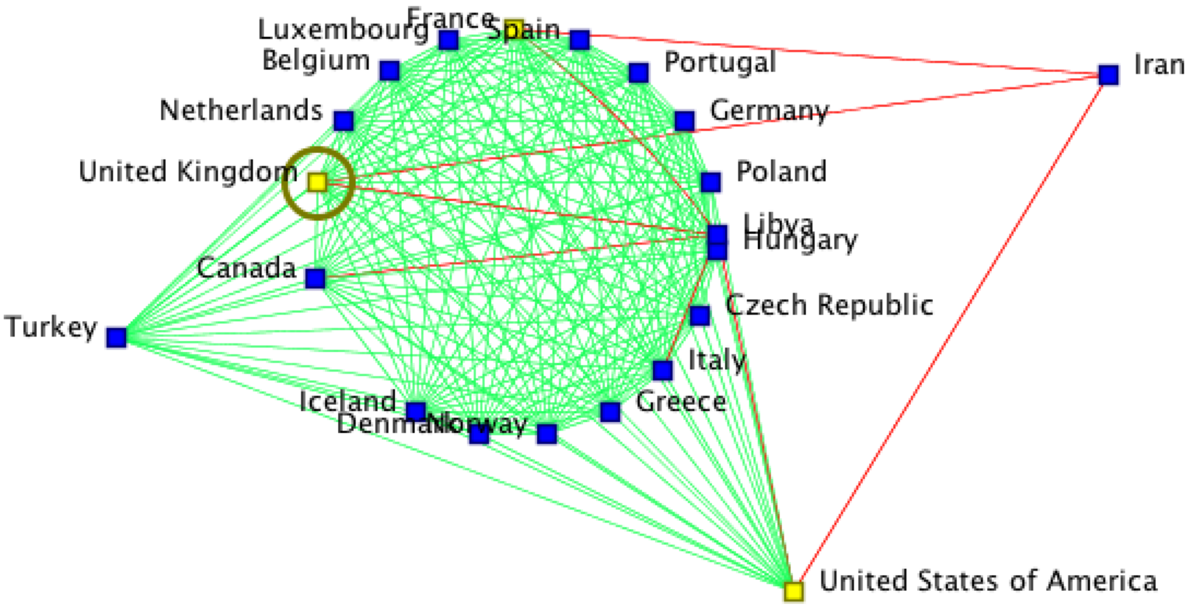


Figure 170. Egonet of United Kingdom

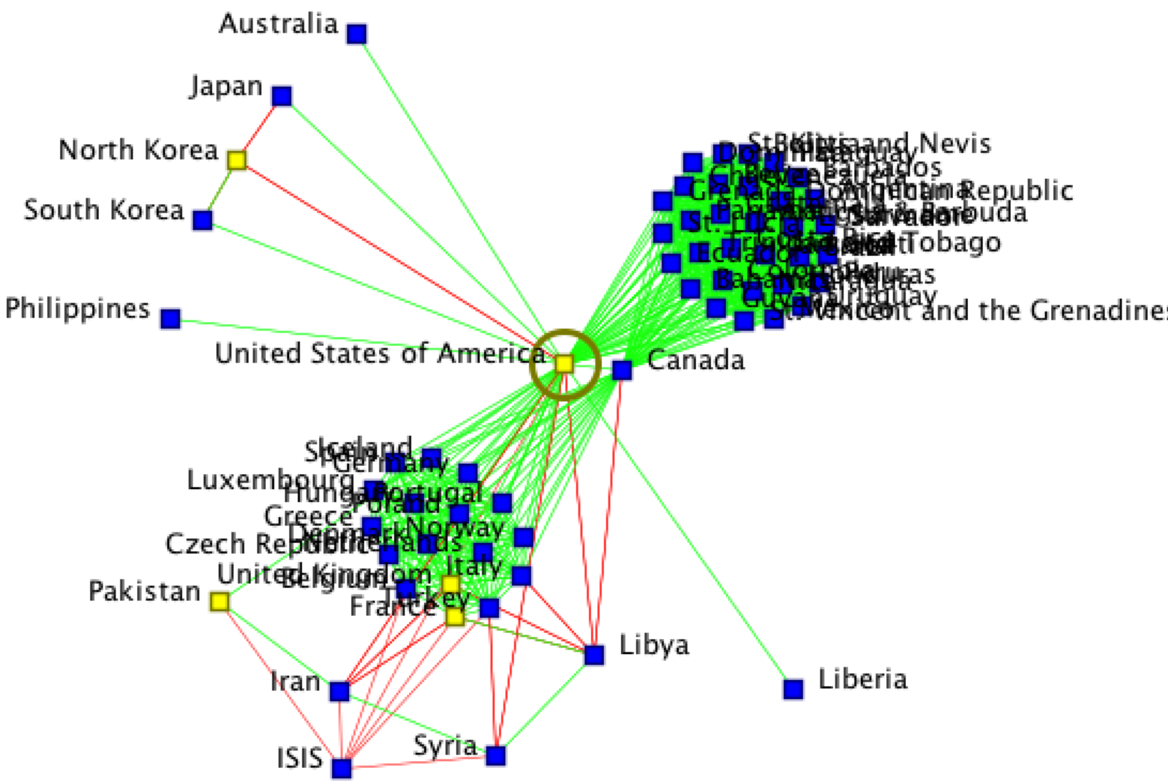


Figure 171. Egonet of United States

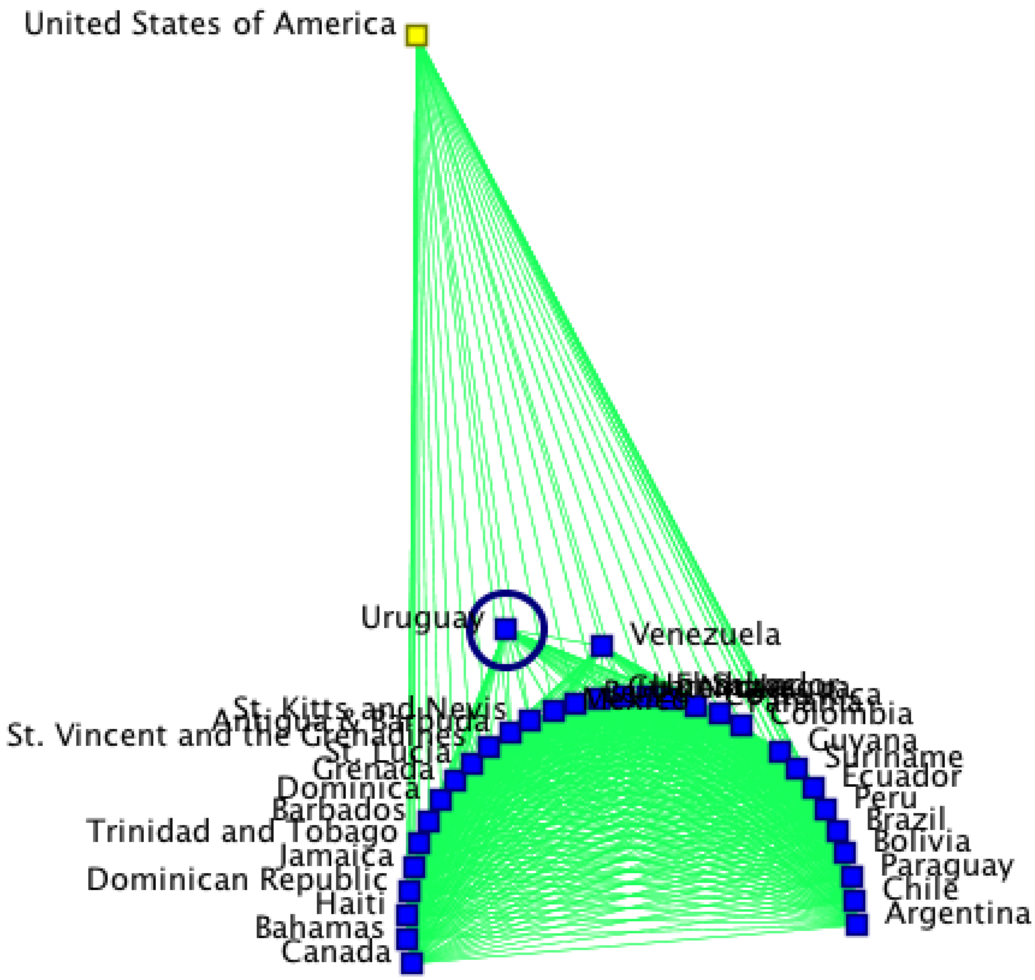


Figure 172. Egonet of Uruguay

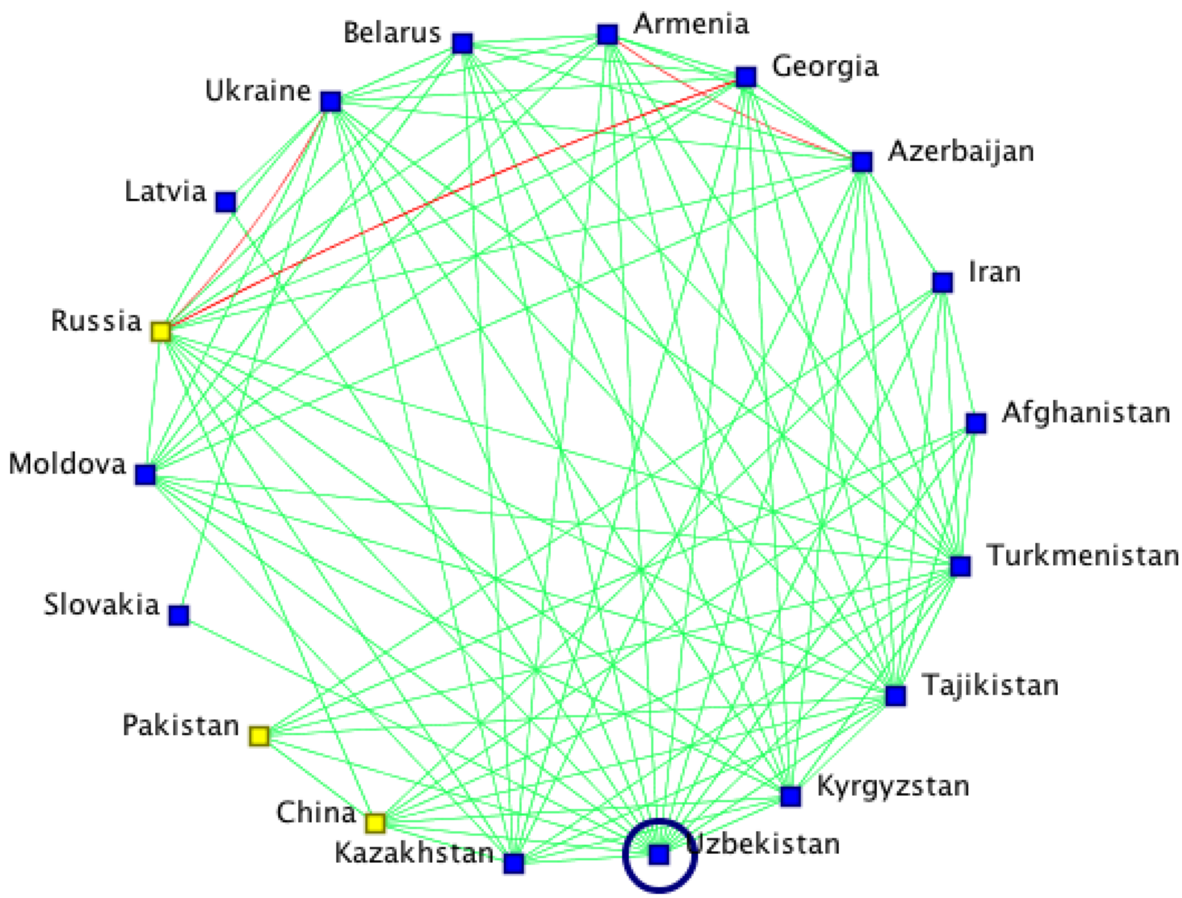


Figure 173. Egonet of Uzbekistan

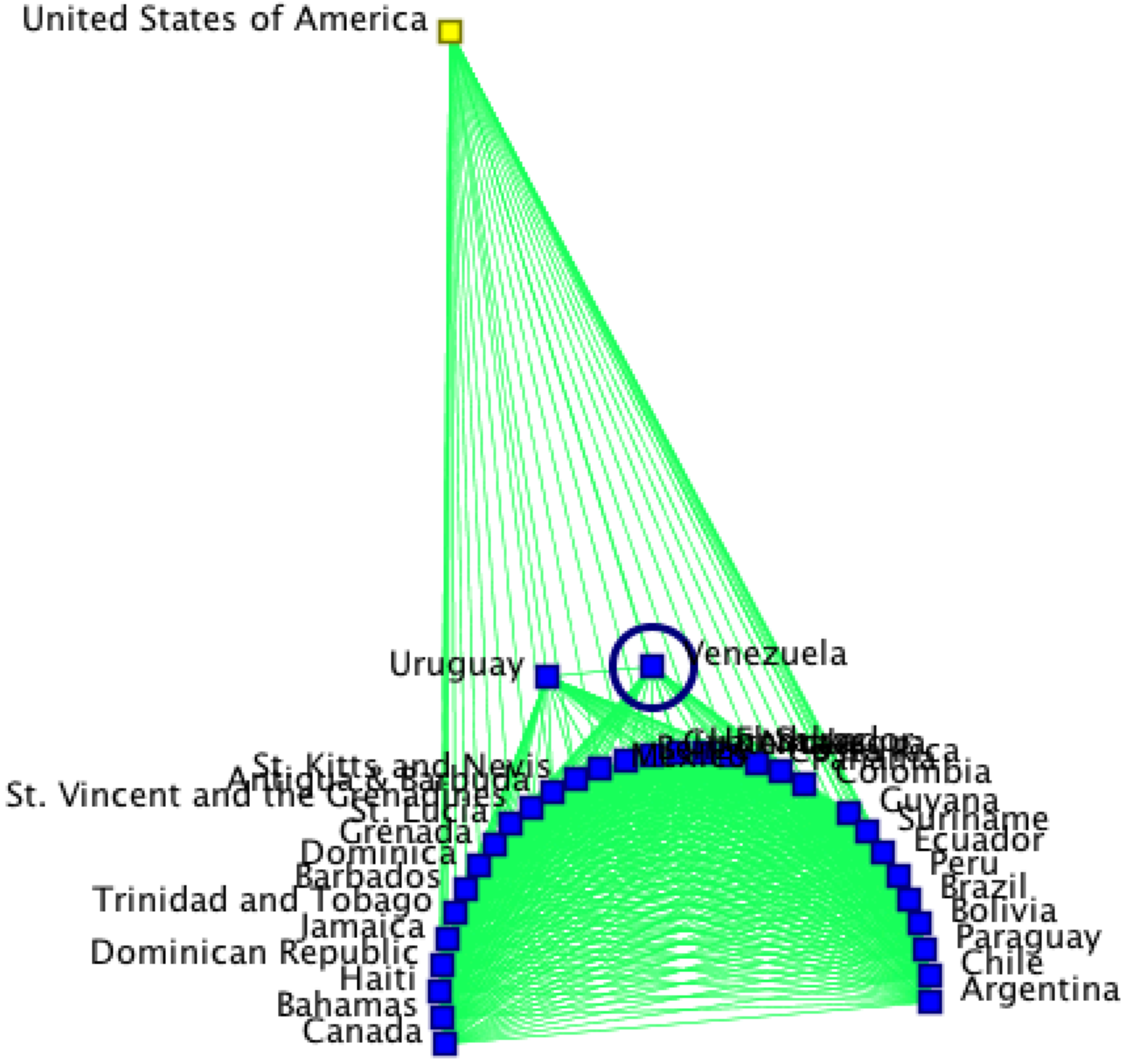


Figure 174. Egonet of Venezuela

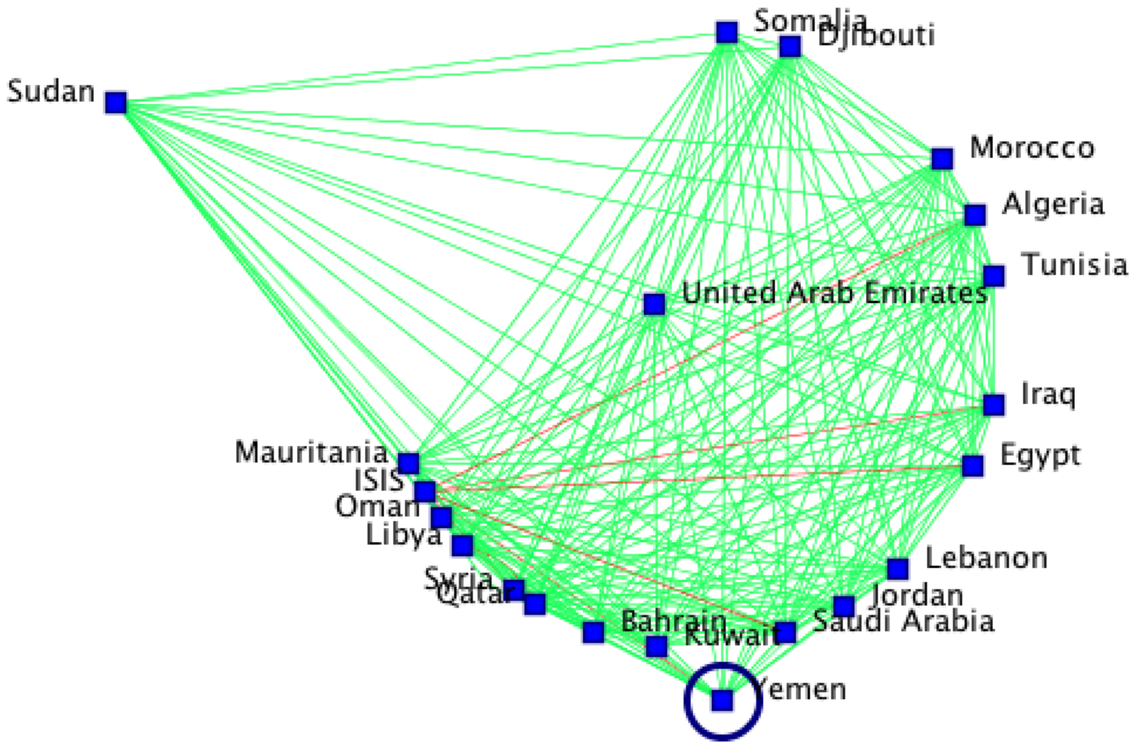


Figure 175. Egonet of Yemen

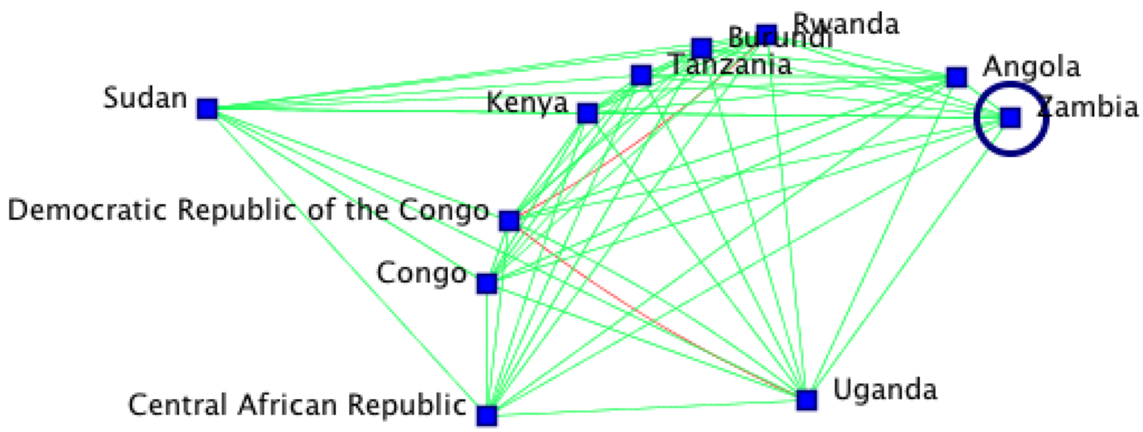


Figure 176. Egonet of Zambia

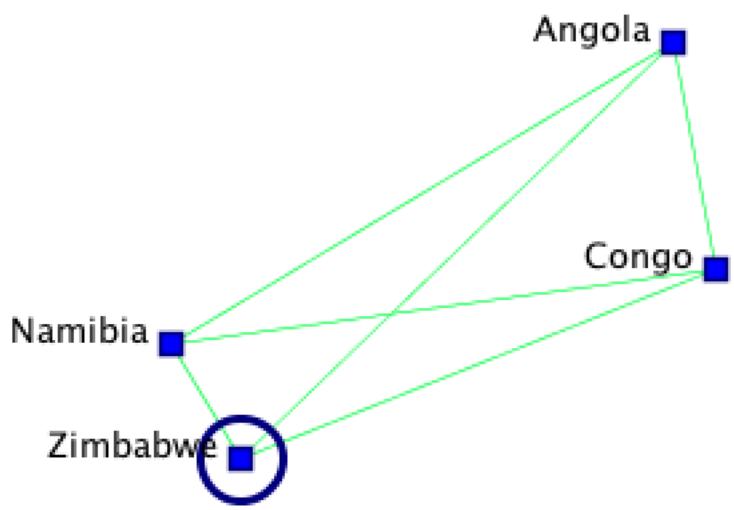


Figure 177. Egonet of Zimbabwe