Collaboration Networks in
Disaster Relief and Humanitarian Aid:
A Case Study of Peru

Mathematical Methods in the Social Sciences
Senior Thesis
6 June 2011

Aditya Kumar
Advised by Professor Irina Dolinskaya
Northwestern University
# Table of Contents

Table of Figures .......................................................................................................................... ii
Acknowledgements ....................................................................................................................... iii

1. Introduction ................................................................................................................................
   1.1 Related Work ......................................................................................................................... 2
   1.2 Paper Overview ...................................................................................................................... 3

2. Networks Background and Terminology ..................................................................................
   2.1 Individual Node Statistics ........................................................................................................ 5
   2.2 Network Statistics ................................................................................................................... 6

3. Data and Methodology ...............................................................................................................
   3.1 Data ...................................................................................................................................... 7
   3.2 Building the Network ............................................................................................................. 9
   3.3 Organizing the Network ........................................................................................................ 10
   3.4 Testing for Small World and Scale-Free Networks ................................................................. 11

4. Analysis & Discussion ................................................................................................................
   4.1 Challenges ............................................................................................................................ 18

5. Conclusion .................................................................................................................................
   5.1 Further Research .................................................................................................................. 19

Bibliography ..................................................................................................................................... 20

Appendices ......................................................................................................................................
   Appendix A – Code ....................................................................................................................... 22
   Appendix B – Centrality for Sub-networks ................................................................................. 36
   Appendix C – Graphs of Sub-networks ....................................................................................... 40
   Appendix D – Power Law Tests ................................................................................................. 45
Table of Figures

Figure 1 – Map of the Affected Region (US Geological Survey) ......................................................... 1
Figure 2 – Example Network ............................................................................................................... 10
Figure 3 – Number of Projects Organizations are Involved With ..................................................... 13
Figure 4 - Illustration of Overall Network. Blue nodes fall in the 90th percentile by degree ...... 14
Figure 5 – Overall Network Sized by Eigenvector Centrality ........................................................... 16
Figure 6 - Degree Distribution Histogram ....................................................................................... 17
Acknowledgements

Many individuals were involved in the process whose eventual output was this paper. Professor Irina Dolinskaya provided immeasurable guidance over the past year and greatly helped define the scope of this project. When it seemed like the data for the type of analysis I wanted to do did not exist, Brooke Jarrett was instrumental in turning me on to Peru and the UNDP. Professor Noshir Contractor first piqued my interest in network analysis. My friends humored me as I pondered various thesis topics ranging from pirates and sea law to predicting scalping markets for concert sales, and read through drafts once I started writing. Finally, I will always be indebted to my family for nurturing my curiosity and encouraging me to find ways to satisfy it. I thank you all.
1. Introduction

On August 15, 2007, an earthquake struck Peru with a magnitude of 8.0 on the Richter scale. Less than one hundred miles from Lima, the earthquake killed at least five hundred people and injured over one thousand. More than thirty-five thousand buildings were destroyed, causing widespread damage to the nation’s infrastructure. Transportation routes and sanitation were significantly affected.

More than three hundred organizations collaborated in the affected area to restore the infrastructure and make the region habitable again. These organizations assisted in all facets of the rebuilding process and filled gaps in the support system created by the earthquake. While some organizations focused on rebuilding damaged infrastructure like latrines or schools, others worked to address the needs of the victims of earthquake by providing children with counseling and support or rehabilitation to those injured.

After a disaster event, the relief effort often becomes crowded with newcomers eager to help. Intuitively, it would seem that the more the merrier when it comes to humanitarian aid. However, the lack of central coordination may lead to redundancies, inefficient allocation of resources and labor, and in some cases may be detrimental to the overall relief effort.

In this paper we examine the reconstruction effort from a network perspective in order to understand interorganizational collaboration efforts and whether a structure of communication emerges from collaboration. Traditional organizational studies have focused on specific organizations or occasionally dyads of organizations. While useful, neither of these lenses sheds light on the entire system of organizations the way a network analysis can. Finding patterns in communication may help identify organizations that should play a leadership role in the ecosystem of humanitarian relief.
1.1 Related Work
Large scale disaster relief efforts have often had to draw upon resources from many different organizations. Kapucu (2007) investigated the emergence of networks of collaboration among disaster relief worker following the September 11, 2001 attacks. Kapucu found that the majority of interviewees believed their organizations benefitted from collaborating with others. Fedorowicz (2007) examined government collaboration using the CapWIN system during Hurricane Katrina and found that interorganizational collaboration is often very difficult. Kapucu (2006) recommended the creation of incentives for communication between organizations as part of preparation for a disaster event.

The unexpected nature of disasters often nullifies plans in place before the event. In a disaster’s wake, many new players emerge to help out with the relief effort (Waugh and Streib 2006). In the United States, hundreds of charities were started after Hurricane Katrina and 9/11 for the purpose of supporting the relief effort either financially or with labor.

With so many players, interorganizational collaboration can be challenging because of complex objectives, scarce resources, and a general lack of trust and accountability between organizations (Vangen and Huxham 2003). Things become even more difficult in international efforts, as cultural and language barriers can hinder collaboration. In 1991, the Department of Humanitarian Affairs was created by the UN in part to address these issues. The UN Office for the Coordination of Humanitarian Affairs (UNOCHA) exists to “mobilize and coordinate effective and principled humanitarian action in partnership with national and international actors in order to alleviate human suffering in disasters and emergencies” (United Nations 2011).

However, an organization like UNOCHA has limited to no authority over the operations of other non-governmental organizations (NGOs). After surveying NGO leaders in Kosovo, Stephenson and Schnitzer (2006) concluded that “humanitarian aid efforts are fraught with competition and confusion. These conditions are part and parcel of the overall environment of relief delivery and cannot simply be fixed by means of a more thoroughgoing top-down coordination. Humanitarian aid implementation is better conceived as a network of actors enmeshed, in part, within a set of preexisting relationships, brought together by an emergency, but with no natural lines of authority existing among them. This suggests that coordination takes place within a relational network of more or less independent organizations (at least from one another).”

Though much has been written about the dynamics of the space, humanitarian relief efforts have not been studied extensively using quantitative methods. However, a few researchers have looked at the nature of collaboration in this humanitarian aid and disaster relief.

In 2000, floods in Mozambique killed hundreds and displaced hundreds of thousands. Over forty-nine countries and thirty international NGOs were involved in the relief effort. Moore et
al. (2003) examined the role of network centrality in humanitarian operations following the floods. Using field interviews and data collected by a local non-profit, they tried to analyze whether organizations with the potential to connect with a number of other organizations were involved in higher impact projects. Moore et al. found that international NGOs were highly central to the network of collaborators. They were able to test the effect of different measures of centrality on the number of people an organization benefited from its work. They found a significant positive correlation between how collaborative an organization was and the number of individuals it benefitted in Mozambique.

Though her work did not focus on collaboration in disaster relief efforts, Word (2005) analyzed the network of non-profit organizations in and around Jacksonville, Florida. Using a survey of 59 local organizations, Word constructed advice, political, and referral networks. Measures of centrality were used to predict the extent of collaboration between organizations. Word found that closeness centrality was a significant factor in predicting collaboration. Her findings indicate that organizational collaboration is driven by two things: the demand for the goods and services an organization provides, and the social structure of the organization’s everyday relationships. Word also found evidence for the emergence of facilitator organizations, who play the role of coordinator in the overall network. This is a similar role as those of international NGOs in Moore’s study. However the statistical methods used in the paper did not control for correlation between ties in a network.

1.2 Paper Overview
In this paper, we adopt and extend the analysis techniques from Moore et al. to the 2007 earthquake in Peru to generate similar insights into and verify the value of network analysis in analyzing humanitarian efforts. We use data collected by the United Nations Development Programme (UNDP) throughout the reconstruction effort. In our analysis of disaster relief organizations in Peru, we find that there is indeed a pattern in the collaboration behavior of organizations involved in the earthquake relief effort. Large international organizations play a more central role in the overall relief effort, interacting with a larger number of organizations and occupying spaces that enable them to facilitate the flow of information. This effect is also present in the subsectors of the relief effort that we investigated. In addition, we test to see if the networks are scale-free or small-world networks. We do not find evidence to support the former claim, but the networks may fall into the category of small-world networks.
2. Networks Background and Terminology

In this section we provide a technical background of network analysis. We begin with an overview of the construction of a network and a review of some of the existing literature in the space. The second part of this section provides descriptions of the measures used to analyze the network in this study.

Sets of relationships can be described as networks. Networks are commonly described as graphs with nodes and edges. Social networks are used to try to describe and simplify human behavior and relationships. In the study of social networks, nodes often represent people or organizations and edges the different relationships that tie them together. Moreno (1934) was one of the first to use graphs to depict social networks when he studied friendships within a group. Borgatti (2009) classifies the relationships depicted by social networks four categories: similarities, social relations, interactions, and flows.

The study of networks can be difficult because of the large amount of data required to do meaningful analysis. For a network of N nodes, information about N^2-N edges needs to be collected to have complete information about the network. Marsden (1990) identifies a number of issues that arise in network analysis. Most social network data is constructed using surveys and other methods requiring human response. For this reason, it is difficult to collect complete information about the nodes in a network and sometimes even to define the boundaries of a network.

However, studying a network often reveals important nodes that otherwise may not have been apparent. Much has been written about nodes with high centrality (Leavitt 1950), with many different methods used to try to identify which nodes are important to the network. The most common centrality measures are degree, closeness, and betweenness. Degree centrality refers simply to the number of nodes that a given node connects to. Closeness uses the average distance of a node to any other node in the network. Betweenness centrality attempts to identify nodes that transmit significant information through the network by virtue of their connections and existing relationships.

Transitivity, or clustering, in networks is often studied to examine grouping behavior. Essentially, the study of clustering tried to answer the question, if Node A and Node B are related, and Node B is tied to Node C, how likely is it that a tie exists between Node A and Node C? In many social networks, the likelihood of a tie between Node A and Node C is much higher than if the network had been created at random. Watts and Strogatz (1998) examined small world networks, or networks with a high degree of clustering. In such networks, “short-cut” relationships exist that greatly cut down the distance between groups of nodes in the network. Small world networks are important because information can quickly spread through them even if the network is sparsely connected or divided into a number of cliques.
The structure of a given network relies heavily on the way that it was constructed. Barabasi (1999) has written about assortativity in the creation of network relationships. Barabasi posited that the distribution of the number of connections a node has in the network can be explained by the fact that new nodes are more likely to connect to nodes that already have a large number of connections. This results in a power-law distribution of node degree. In these networks, dubbed scale-free networks, nodes with high degree are referred to as hubs, and often act as brokers of information. In such networks, there often exist key nodes whose removal would dramatically affect the connectedness of the network.

2.1 Individual Node Statistics

There are a number of measures of centrality that estimate a node’s role in the overall network. Centrality measures how much power a given node has relative to the rest of the network. Those in privileged social positions are often able to exert influence over other members in a social network.

One simple measure of power in a network is the degree, or number of connections, of an organization. In this network, an organization’s degree is a count of the number of other organizations it has collaborated with in Peru.

**Degree centrality** is a normalized measure of degree. A node’s degree centrality is defined as the degree of the node divided by the maximum possible degree in the network. Since the simple networks constructed in this paper do not contain self-links or multiple links between nodes, the maximum possible degree for an organization is simply N-1, where N is the number of organizations in the network. This measure emphasizes organizations that have collaborated with the largest number of other organizations.

**Closeness centrality** \( CC_i \) is a measure of a node’s proximity to the rest of the network. For a given node \( i \), it is defined as:

\[
CC_i = \frac{1}{\text{average distance to all other nodes}_i}
\]

**Betweenness Centrality** \( BC_i \) is a measure of how often a node lies on the shortest path between two nodes. Brandes (2001) defined betweenness centrality for a node \( i \) as:

\[
BC_i = \sum_{j,k \in V} \frac{\sigma(j,k|i)}{\sigma(j,k)}
\]

Where \( \sigma(j,k) \) and \( \sigma(j,k|i) \) are the number of shortest paths between \( j \) and \( k \) and the number of shortest paths between nodes \( j \) and \( k \) that go through node \( i \), respectively.

**Eigenvector centrality** uses the largest eigenvalue of the network’s adjacency matrix to calculate an eigenvector for the network. The idea behind eigenvector centrality is that while the number of connections a node has in the network is important, all connections are not necessarily of equal importance. A connection to an influential organization should have more
weight than a connection to an organization that is not as well connected. Eigenvector centrality was the basis for Google’s original PageRank system for sorting search results (Newman 2010). The networkX algorithm to calculate eigenvector centrality uses the Power Method to calculate the eigenvector of the network matrix.

2.2 Network Statistics
Summary statistics for the network enable comparison across different types of networks. They give a sense of the “connectedness” of a given network. In this study, we examine three summary statistics for networks: Average degree, Density, and Mean Geodesic.

**Average degree** is the average number of other organizations an average organization has collaborated with in Peru. This value establishes a baseline level of interactions for the network.

**Density** is calculated as: \( \frac{\text{Number of Edges}}{\text{Number of possible Edges}} \). For a network that does not allow multiple parallel links or self links, this value simplifies to \( \frac{\text{Number of Edges}}{N(N - 1)} \), where \( N \) is the number of nodes in the network.

**Mean geodesic** is the average shortest path between two nodes in the network. If we consider the collaboration network to be a suitable proxy for a communication network among the organizations in Peru, this statistic could be interpreted as the average number of collaborators an organization would have to communicate through to get in touch with another organization.

To measure how tightly connected groups within the network are, we use the **average clustering coefficient** (\( C \)):

\[
C = \frac{1}{N} \sum_{i \in G} c_i
\]

Where \( N \) is the number of nodes in the network, \( i \) is a node in the network, and \( c_i \) is the clustering coefficient of the node \( i \). The **clustering coefficient** (\( c_i \)) is a measure of how connected node \( i \)'s

\[
c_i = \frac{2T(i)}{\deg(i) [\deg(i) - 1]}
\]

Where \( T(i) \) is the number of triangles through node \( i \). A triangle is a set of three nodes that each have connections to the others and \( \deg(i) \) is the degree of node \( i \).
3. Data and Methodology

In this section we provide an overview of the data used for this study and describe the methodology used to build and analyze the network.

3.1 Data

When the earthquake struck Peru in 2007, a number of organizations were on the scene to assist in the recovery process. One of the international organizations on the ground was the UNDP. The UNDP works with local and national governments, NGOs, and the private sector to help coordinate development efforts and measure the effectiveness of different programs. Because of its position in the development ecosystem, the UNDP was able to collect information about the actions of players involved in the disaster relief effort in Peru.

The data for this project was obtained from a UNDP Group entitled “Coordinación y Transición Perú,” or Coordination and Transition Peru. The group was set up by the UNDP and NGOs operating in the region affected by the 2007 earthquake to share and store project information regarding reconstruction efforts.

This document repository contains project databases for the provinces of Castrovirreyna and Huaytará in the Department of Huancavelica, Chincha, Pisco, and Ica in the Department of Ica, and Cañete in the Department of Lima. Projects are separated into databases grouped by project state, funding, sector, and location. The typical project record contains the following fields:

- Agency Type
- Department
- Province
- Sector - The general sector in which the project is classified (ex. Water Sanitation)
- Organization – Organization(s) responsible for the project
- District - The smallest geographic attribute in the dataset, this captures the district(s) in which the project was active
- Description – A brief description of the project
- Progress
- Financier/Executor – Organization(s) involved in funding or executing the project
- Beneficiaries - Number of families or individuals that will benefit from the successful completion of this project.
In all, 1398 projects were captured in this project database involving 330 organizations. In some cases, the same organization reported projects under multiple, slightly different names. In these instances, the organization name was standardized where possible.

<table>
<thead>
<tr>
<th>Agency Type</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agencias NU</td>
<td>342</td>
</tr>
<tr>
<td>Gobierno</td>
<td>68</td>
</tr>
<tr>
<td>Gobierno Local</td>
<td>97</td>
</tr>
<tr>
<td>ONG</td>
<td>600</td>
</tr>
<tr>
<td>Other</td>
<td>32</td>
</tr>
<tr>
<td>Otras</td>
<td>96</td>
</tr>
<tr>
<td>(blank)</td>
<td>163</td>
</tr>
</tbody>
</table>

Table 1 - Projects by Agency Type

<table>
<thead>
<tr>
<th>Department</th>
<th>Province</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huancavelica</td>
<td>Castrovirreyna</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Huaytara</td>
<td>4</td>
</tr>
<tr>
<td>Huancavelica Total</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>Ica</td>
<td>Chincha</td>
<td>486</td>
</tr>
<tr>
<td></td>
<td>Ica</td>
<td>330</td>
</tr>
<tr>
<td></td>
<td>Pisco</td>
<td>503</td>
</tr>
<tr>
<td>Ica Total</td>
<td></td>
<td>1319</td>
</tr>
<tr>
<td>Lima</td>
<td>Cañete</td>
<td>69</td>
</tr>
<tr>
<td>Lima Total</td>
<td></td>
<td>69</td>
</tr>
</tbody>
</table>

Table 2 - Projects by Geographic Region

Table 1 summarizes the projects in Peru by agency type. Table 2 summarizes projects by geographic location. In this dataset, the majority of projects took place in the Department of Ica, which was most affected by the earthquake. The province with the most activity was Pisco, the epicenter of the earthquake. Table 3 summarizes all projects by the sectors they were classified in by the UNDP. Health, Housing, and Water & Sanitation were the most popular sectors for relief efforts.
### Table 3 - Projects by Sector

<table>
<thead>
<tr>
<th>Sector</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agua</td>
<td>1</td>
</tr>
<tr>
<td>Ayuda Alimentaria-BMZ</td>
<td>1</td>
</tr>
<tr>
<td>Ayuda humanitarian</td>
<td>4</td>
</tr>
<tr>
<td>Capacity Building</td>
<td>4</td>
</tr>
<tr>
<td>Civil Society</td>
<td>1</td>
</tr>
<tr>
<td>Communications Reinforcement</td>
<td>1</td>
</tr>
<tr>
<td>Comunicaciones</td>
<td>1</td>
</tr>
<tr>
<td>Desarrollo de capacidades</td>
<td>2</td>
</tr>
<tr>
<td>Economic Recovery</td>
<td>1</td>
</tr>
<tr>
<td>Education</td>
<td>228</td>
</tr>
<tr>
<td>Emergency</td>
<td>3</td>
</tr>
<tr>
<td>Food Safety</td>
<td>98</td>
</tr>
<tr>
<td>Fortalecimiento institucional</td>
<td>1</td>
</tr>
<tr>
<td>Health</td>
<td>302</td>
</tr>
<tr>
<td>Housing</td>
<td>251</td>
</tr>
<tr>
<td>Livelihood</td>
<td>111</td>
</tr>
<tr>
<td>NGO Coordination</td>
<td>1</td>
</tr>
<tr>
<td>Protection</td>
<td>125</td>
</tr>
<tr>
<td>Red Agua Potable</td>
<td>1</td>
</tr>
<tr>
<td>SNU Coordination</td>
<td>1</td>
</tr>
<tr>
<td>Techos Temporales</td>
<td>2</td>
</tr>
<tr>
<td>Various</td>
<td>6</td>
</tr>
<tr>
<td>Water &amp; Sanitation</td>
<td>250</td>
</tr>
<tr>
<td>Vivienda y gestion Territorio</td>
<td>1</td>
</tr>
<tr>
<td>Casa Hogar Ninos Huérfanos Senor de Luren</td>
<td>1</td>
</tr>
</tbody>
</table>

#### 3.2 Building the Network

One of the fundamental issues in performing network analysis is defining what constitutes a tie between two nodes in a network. Sometimes, a network tie is well defined. For example, one can define a relationship between two Facebook users if they are “friends” with each other. However, it can often be difficult to operationalize a relationship measure. For example, when examining disaster relief efforts in Peru, it is hard to say what constitutes collaboration between entities. Two organizations may be in constant communication and sharing information with each other, while others may have people in the field working together on projects, and it is hard to distinguish among these different types of collaboration. For the purposes of this analysis, we define a tie between two organizations as existing if both organizations appear in the Organization or Financier/Executor field of a given project.
To demonstrate, Table 4 describes the network ties in a hypothetical collaboration of four organizations. Organizations A and B are both listed in the record for Project 1, thus the record indicates that a connection exists between them. In Project 2, Organizations B and C are listed and consequently connected. Figure 2 illustrates the network that results from aggregating over these network ties.

<table>
<thead>
<tr>
<th>Project</th>
<th>Organization</th>
<th>Financier/Executor</th>
<th>Does the project mean a tie exists between</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>A,B</td>
</tr>
<tr>
<td>Project 1</td>
<td>Org A</td>
<td>Org B</td>
<td></td>
</tr>
<tr>
<td>Project 2</td>
<td>Org B</td>
<td>Org C</td>
<td></td>
</tr>
<tr>
<td>Project 3</td>
<td>Org A</td>
<td>Org C,B</td>
<td>YES</td>
</tr>
<tr>
<td>Project 4</td>
<td>Org A</td>
<td>Org D</td>
<td></td>
</tr>
</tbody>
</table>

Table 4 - Definition of Network Tie

3.3 Organizing the Network
In addition to examining the overall network of collaboration in Peru, we examine different sectors of humanitarian relief to see if collaboration structures vary. To do this, we look at the sub-networks created by each sector in the database. To ensure that there is enough data to build a network, we only look at sectors with 50 or more projects in the database. Sectors that
meet this criterion are Education, Food Safety, Health, Housing, Livelihood, Protection and Water & Sanitation. In addition to examining whether collaboration varies across project type, we also look at whether collaboration varies geographically by individually analyzing the three provinces of Ica, the department in which most of the projects were executed.

3.4 Testing for Small World and Scale-Free Networks
We also examine the network and its sub-networks to see if relief networks are small-world networks. In this study, we use the framework laid out by Watts and Strogatz to see if the relief network is a small-world network. Watts and Strogatz defined a small world network as one with short average path length between nodes and a high clustering coefficient.

We test for the existence of a scale-free structure as described by Barabasi (1999) using methods outlined by Clauset et al. (2009). The existence of a scale-free structure is useful to identify because it can mean that a “targeted attack,” or the removal of key organizations could destroy the overall connectedness of the network. In this network, this property would underscore the importance of organizations with a large number of partners because their absence could imply a much lower degree of collaboration in the network.
4. Analysis & Discussion

The average humanitarian relief project involved 2.7 organizations. This was the average project size across different project sectors and geographic regions, although it appears that projects were slightly larger on average in the province of Pisco. It is possible that the nature of the damage in Pisco required larger collaboration teams. Pisco was the epicenter of the earthquake, and its seaside location contributed to a great deal of water damage compared to regions further inland.

<table>
<thead>
<tr>
<th></th>
<th>Projects</th>
<th>Organizations/Project</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>1398</td>
<td>2.775</td>
</tr>
<tr>
<td><strong>Project Sectors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>228</td>
<td>2.627</td>
</tr>
<tr>
<td>Food Safety</td>
<td>98</td>
<td>2.949</td>
</tr>
<tr>
<td>Health</td>
<td>302</td>
<td>2.772</td>
</tr>
<tr>
<td>Housing</td>
<td>251</td>
<td>2.892</td>
</tr>
<tr>
<td>Livelihood</td>
<td>111</td>
<td>2.684</td>
</tr>
<tr>
<td>Protection</td>
<td>125</td>
<td>2.888</td>
</tr>
<tr>
<td>Water &amp; Sanitation</td>
<td>250</td>
<td>2.688</td>
</tr>
<tr>
<td><strong>Geographic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chincha</td>
<td>486</td>
<td>2.84</td>
</tr>
<tr>
<td>Ica</td>
<td>330</td>
<td>2.682</td>
</tr>
<tr>
<td>Pisco</td>
<td>503</td>
<td>2.943</td>
</tr>
</tbody>
</table>

*Table 5 - Average Organizations per Project*

Figure 3 illustrates the number of projects organizations were involved with in the relief effort. Of the 330 organizations in this study, CARE PERU was the most active over the time period, with involvement in 160 projects. ECHO was the only other organization involved with more than 100 projects. This is not surprising since CARE and ECHO are both international organizations with experience dealing with natural disasters.
Table 6 - Network Summary Statistics

Table 6 provides network statistics for each of the sub-networks in this study. Each of the subsets of the overall network has a higher density than the overall network. This implies that the level of collaboration across sectors and across geographic regions is lower than the rate of collaboration within. Organizations likely have more reasons to collaborate with organizations in their field or with groups that operate nearby. Of all the sectors, Food Safety has the densest and tightest network. This can be explained by the fact that, of the reported Food Safety projects, many were repeated collaborations among a number of organizations.
Figure 4 illustrates the entire network of organizations across all sectors and geographic regions. Organizations in the 90th percentile for number of partners are highlighted in blue. Appendix C contains visualizations of all of the sub-networks in this study. These are organizations that, at least by degree, were the most collaborative in the relief effort. Though the network has a giant component, it is clear that the network has many isolated components of a few organizations that work only with each other.
Table 7 lists the top five organizations by each network centrality measure for the overall relief effort. The organizations most central to the overall collaboration network are large international NGOs like UNICEF and ECHO. The government of Pisco, Municipalidad Pisco, also had a high degree in the overall network. The high degree of organizations likely gave them access to a lot of information about the relief effort in Peru, putting them in a prime position to coordinate the overall effort. Further investigation would be required to see if these organizations took on leadership roles in the effort. Appendix B contains tables of the top five organizations for each of the sub-networks in this study.

The organizations with high betweenness centrality are those that consistently fall on the shortest path between the two organizations. If information travels through the network, these organizations are positioned to have an understanding about what is going on across the relief effort. It seems that these organizations like the Humanitarian Aid department of the European Commission (ECHO) and Cooperazione Internazionale (COOPI) are ones that are on the ground and involved with direct disaster relief projects.

Organizations with high eigenvector centrality have collaborated with key members of the network. Eigenvector centrality is often considered as a revealed prestige ranking. Thus it is possible that there is some prestige involved with working with organizations like UNICEF or United Nations Population Fund (UNFPA). These organizations may be “preferred partners” or partners that are especially valued by other organizations in the relief effort. There are a number of reasons for this to be the case. It is possible that UNICEF and UNFPA are organizations that are easy to work with or are able to provide supplies to humanitarian workers. MINSA, the Peruvian Ministry of Health, may be a valued partner in relief efforts because of its knowledge of post-disaster health risks in Peru. Figure 5 illustrates the overall network. Each node’s size it determined by its eigenvector centrality.

The network of organizations in Peru does not have a scale-free structure. This means that the distribution of the number of partners an organization has does not vary according to a power law. Table 8 summarizes the results of power-law tests from Clauset et al. for the overall network as well as the sub-networks described earlier.
Figure 5 – Overall Network Sized by Eigenvector Centrality
<table>
<thead>
<tr>
<th>NETWORK</th>
<th>ALPHA</th>
<th>ALPHA_ERR</th>
<th>KS</th>
<th>PVAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>2.607</td>
<td>0.099</td>
<td>0.771</td>
<td>0.00</td>
</tr>
<tr>
<td>Education</td>
<td>2.011</td>
<td>0.092</td>
<td>0.808</td>
<td>0.00</td>
</tr>
<tr>
<td>Food Safety</td>
<td>2.217</td>
<td>0.197</td>
<td>0.684</td>
<td>0.00</td>
</tr>
<tr>
<td>Health</td>
<td>2.573</td>
<td>0.172</td>
<td>0.774</td>
<td>0.00</td>
</tr>
<tr>
<td>Housing</td>
<td>2.011</td>
<td>0.085</td>
<td>0.773</td>
<td>0.00</td>
</tr>
<tr>
<td>Livelihood</td>
<td>2.247</td>
<td>0.150</td>
<td>0.768</td>
<td>0.00</td>
</tr>
<tr>
<td>Protection</td>
<td>2.117</td>
<td>0.142</td>
<td>0.774</td>
<td>0.00</td>
</tr>
<tr>
<td>Water &amp; Sanitation</td>
<td>2.072</td>
<td>0.112</td>
<td>0.804</td>
<td>0.00</td>
</tr>
<tr>
<td>Chincha</td>
<td>2.011</td>
<td>0.083</td>
<td>0.803</td>
<td>0.00</td>
</tr>
<tr>
<td>Ica</td>
<td>1.983</td>
<td>0.090</td>
<td>0.817</td>
<td>0.00</td>
</tr>
<tr>
<td>Pisco</td>
<td>2.664</td>
<td>0.170</td>
<td>0.802</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 8 - Power Law Tests for Scale-Free Properties

Figure 6 - Degree Distribution Histogram

Figure 6 shows a loglog plot of the degree distribution for the overall network, with an overlay of the bestfit power law distribution from our analysis (α=2.607). This power-law model does not fit the data well. Our test uses the Kolmogorov-Smirnov (KS) statistic to compare the empirical distribution to the expected distribution and finds that the model fits with a p value of 0.00 (in this analysis, higher p values indicate significance). As the p values for the sub-
networks and the additional plots in Appendix D make clear, none of the networks qualify as scale-free networks.

Though it does not qualify as a scale-free network, the relief effort in Peru may follow the form of a small-world network. Watts and Strogatz described small world networks using network metrics: the number of nodes $N$, the average degree $k$, the average path length $L$, and the clustering coefficient $C$. If $L \sim N/2k$ and $C \sim 3/4$, then the network may be classified as a small world network. For the overall network in Peru, $L = 3.722 < N/2k = 330/4.648 = 35.5$ and $C = 0.634$. While the average path length is much shorter than the threshold discussed by Watts and Strogatz, the clustering coefficient for the network is also lower. This may be because the network in question is not fully connected; Watts and Strogatz required $n \gg k \gg \ln(n) \gg 1$ to ensure connectedness. Thus, we cannot conclusively say that the network follows a small-world pattern.

4.1 Challenges
There were a number of roadblocks to doing this analysis. The most pressing is the integrity of the dataset. The data was pulled from an online archive of a UNDP working group, but there is no way to verify whether the project database has been updated. It is also possible that the UNDP record of projects in Peru was not comprehensive, and that many partnerships between organizations went unobserved in the UNDP dataset. Second, though every effort was made to standardize organization names, lack of fluency in Spanish was a significant challenge in this analysis. It is possible that translation errors occurred over the course of this study or that the same organization was reported under different names and not aggregated properly.
5. Conclusion

With such a rough dataset, it is hard to find conclusive insights into the patterns of collaboration present in a disaster relief operation such as the one in Peru. In spite of this obstacle, the analysis in this paper does shed light on what could be gained by applying network analysis techniques to disaster relief and humanitarian aid.

We find empirical evidence for the important roles played by international NGOs and local governments in local relief efforts, which may be important in when it comes time to select lead institutions for future disaster relief efforts.

The fact that the relief network is not scale-free may be useful to the understanding of continuity in relief efforts. Since it is not scale-free, the overall connectedness of the network is not susceptible to a “targeted attack,” where one systematically removes highly central nodes from the network. In the context of this network, this could mean that if some of the highly central international NGOs move on to the next international crisis, the overall relief effort may not suffer much because the connectedness of the overall network will remain intact.

The possibility that the network is a small world network and the fact that it is not a scale-free network may also mean that it is robust to random perturbations. If organizations involved in the relief effort randomly drop out of or come into the network, the overall structure of the network is unlikely to change significantly. This may be useful for long term planning. While it may not be possible to depend on the involvement of a specific organization in the long term, the robustness of the network may ensure that other organizations with similar capabilities will be around when the time comes to call on them.

5.1 Further Research

The results of this study are promising for further research. Most important in any further research is establishing a system to collect accurate and comprehensive data. Since network analysis is data intensive, it is crucial that the data collected fully describe all players in the collaboration network in order to confidently conclude anything about the overall network.

Combining the network analysis from this study with contextual knowledge about the specific disaster will allow researchers to develop deeper insights into the dynamics of a humanitarian aid operation. Researchers working in their native language will have an easier time describing implications of the results of network analysis.

It would be interesting to extend this type of analysis to other humanitarian aid operations around the world to see how they differ. The network of collaboration in the Gulf Coast of the United States or in Japan after the earthquake in 2011 will probably look different than the one we described in Peru. The differences in the network structure could be an artifact of the differences in capabilities between the first world and the developing world.
Bibliography


Monge, P. R. and N. Contractor (2003). Theories of communication networks, Oxford University Press US.


Appendices

Appendix A – Code

```python
Created on May 9, 2011
@author: Aditya Kumar

import csv
import sqlite3
import 22omes22s as nx
from 22omes22s.exception import NetworkXException, NetworkXError
import matplotlib.pyplot as plt

# Define Graph Analysis Function
def analyzeGraph(projTypeGraph, ss):
    plt.figure(figsize=(10,10))
    plt.title(str(projTypeGraph) + “ Network”)  # Network
    gPos = nx.spring_layout(projTypeGraph, iterations=20)
    nx.draw(projTypeGraph, pos=gPos, with_labels=False, node_size=50, alpha=0.5)

    # topNode = nx.Graph(name=“HighDegree”)
    degreeCutoff = sorted(projTypeGraph.degree().viewvalues(), reverse=True)[9]
    topList = []
    topDict = {}
    for n in projTypeGraph:
        if (projTypeGraph.degree()[n] >= degreeCutoff):
            topList.append(n)
            topDict[n] = n

    nx.draw_networkx_nodes(projTypeGraph, pos=gPos, nodelist=topList, node_size=50, alpha=0.5, node_color=’b’)
    nx.draw_networkx_labels(projTypeGraph, pos=gPos, labels=topDict, font_size=9)
    plt.savefig(“output\networkPlot-“ + str(projTypeGraph.name) + “.png”)
    plt.clf()

# Analysis Algorithms
# Export dot file for visualization
nx.write_dot(projTypeGraph, “output\export-“ + str(projTypeGraph.name) + “.dot”)

# POWER LAW CHECKS
import plfit
x = (projTypeGraph.degree().values())
powerFit = plfit.plfit(sorted(x))
[p, ksv] = powerFit.test_pl()
powerFit.plotcdf()
plt.title(str(projTypeGraph) + “ LogLog Degree Distribution”)```
plt.ylabel("log(Number of Organizations)")
plt.xlabel("log(Degree)")
plt.savefig("output\power-"+str(projTypeGraph)+”.png")
plt.clf()
plt.close()
print "PLFIT!"

# SUMMARY INFO
meanGeodesic = 0
for g in nx.connected_component_subgraphs(projTypeGraph):
    meanGeodesic += nx.average_shortest_path_length(g) *
    nx.classes.function.number_of_nodes(g) /
    nx.classes.function.number_of_nodes(projTypeGraph)
avgDegree =
float(sum(nx.classes.function.degree(projTypeGraph).values()))/len(nx.classes
.function.degree(projTypeGraph))
# PRINT SUMMARY DATA
ss.writerow([str(projTypeGraph),
    nx.classes.function.number_of_nodes(projTypeGraph),nx.classes.function.number
_of_edges(projTypeGraph), avgDegree,
    nx.classes.function.density(projTypeGraph),
    meanGeodesic,powerFit._xmin,powerFit._alpha,powerFit._alphaerr,powerFit._like
lihood,powerFit._ks,powerFit._ngtx,p])

# Degree Histogram
# plot(nx.classes.function.degree_histogram(projTypeGraph)
plt.plot(nx.classes.function.degree_histogram(projTypeGraph),’r’,marker
=’o’)  
plt.title(str(projTypeGraph)+” Degree Plot”)
plt.ylabel("Number of Organizations")
plt.xlabel("Degree")
plt.savefig("output\degreePlot-"+str(projTypeGraph)+”.png")
plt.clf()
plt.close()

# Open Centrality File
centralityFile = open(’output\centrality-
+str(projTypeGraph)+’.txt’,’wb’)
f = csv.writer(centralityFile, delimiter=’\t’)
# Degree Centrality Analysis
degreeList = nx.classes.function.degree(projTypeGraph)
# Degree Centrality Analysis
degreeCList = nx.algorithms.centrality.degree_centrality(projTypeGraph)
# Closeness Centrality
closenessList = nx.algorithms.centrality.closeness_centrality(projTypeGraph)
# Betweenness
betweennessList = nx.algorithms.centrality.betweenness_centrality(projTypeGraph)
# Eigenvector
eigenTrue = True
try:
eigenList =
nx.algorithms.centrality.eigenvector_centrality(projTypeGraph)
    except NetworkXError:
        eigenTrue = False
# Print to File
print >> centralityFile,
"NETWORK\ORGANIZATION\tDEGREE\tDEGREE_CENTRALITY\tCLOSENESS\tBETWEENNESS\tEI
genvalue"
for org in degreeList:
    if(eigenTrue):
        f.writerow( [str(projTypeGraph), org, degreeList[org],
                    degreeCList[org], closenessList[org], betweennessList[org], eigenList[org]] )
    else:
        f.writerow( [str(projTypeGraph), org, degreeList[org],
                     degreeCList[org], closenessList[org], betweennessList[org], 'NA'] )
# Close Centrality File
centralityFile.close()
# END OF ANALYZE GRAPH FUNCTION
#Open data file
dataReader =
csv.reader(open('Data_Book_04.txt','rb'),dialect='excel',delimiter='\t')
#Connect to project database in SQLite3
conn = sqlite3.connect('peruDB.db')
c = conn.cursor()
#Create table and store project lists
c.execute('drop table projectTypeDB')
c.execute('''create table projectTypeDB(project integer, agency_type text, department text, province text, proctype text, orgs text, numOrgs integer)''')
print dataReader.next()
for row in dataReader:
    c.execute("insert into projectTypeDB
            values(?,?,?,?,?,?,?)",(str(row[0]),str(row[1]),str(row[2]),str(row[3]),str(row[4]),str(row[14]),str(row[16])))
#Open Project Summary File
summaryFile=open('output\Summary.txt','wb')
ss = csv.writer(summaryFile,delimiter='\t')
ss.writerow(["Network","Nodes","Edges","avgDegree","Density","meanGeodesic","xmin","alpha","alphaErr","Log-Likelihood","KS","N>xmin"])#Build and Analyze Large Network
projectOrgList = c.execute("select orgs from projectTypeDB")
projTypeGraph = nx.Graph(name="Overall")
print "Overall Graph Started"
for orgList in projectOrgList:
    for org1 in orgList[0].split(’,’):
        if org1 == "":
            break
    for org2 in orgList[0].split(’,’):
        if org2 == "":
            break
    if org1.strip().lower()!=org2.strip().lower():
        if not projTypeGraph.has_edge(org1.strip().lower(),org2.strip().lower()):
            projTypeGraph
```python
.add_edge(org1.strip().lower(),org2.strip().lower(),weight=1)
else:
    projTypeGraph[org1.strip().lower()][org2.strip().lower()]['weight'] += 1

print "Large Network Graph Completed"
analyzeGraph(projTypeGraph,ss)

#Iterate Over Project Types and Create Networks
projTypeListQuery = c.execute("select projtype from projectTypeDB group by
projtype having(count(projtype)>50)")
for projType in projTypeListQuery.fetchall():
    projTypeGraph = nx.Graph(name=str(projType[0]))
    print projType[0]+" Started"
    # Query Based on Project Type
    projectOrgList = c.execute("select orgs from projectTypeDB where
projtype=?",projType)
    for orgList in projectOrgList:
        for org1 in orgList[0].split(','):  
            if org1 == "":
                break
        for org2 in orgList[0].split(','):  
            if org2 == "":
                break
            if org1.strip().lower()!=org2.strip().lower():
                if not projTypeGraph.has_edge(org1.strip().lower(),org2.strip().lower()):
                    projTypeGraph
                    .add_edge(org1.strip().lower(),org2.strip().lower(),weight=1)
                else:
                    projTypeGraph[org1.strip().lower()][org2.strip().lower()]['weight'] += 1
        print str(projType[0])+" Graph Completed"
analyzeGraph(projTypeGraph,ss)

#Iterate over Geographic locations for Ica
provinceListQuery = c.execute("select province from projectTypeDB where
department='Ica' group by province ")
for province in provinceListQuery.fetchall():
    provinceGraph = nx.Graph(name=str(province[0]))
    print province[0]+" Started"
    # Query Based on Project Type
    projectOrgList = c.execute("select orgs from projectTypeDB where
province=?",province)
    for orgList in projectOrgList:
        for org1 in orgList[0].split(','):  
            if org1 == "":
                break
        for org2 in orgList[0].split(','):  
            if org2=="":
                break
            if org1.strip().lower()!=org2.strip().lower():
                if not provinceGraph.has_edge(org1.strip().lower(),org2.strip().lower()):
                    provinceGraph
```
```python
.add_edge(org1.strip().lower(),org2.strip().lower(),weight=1)
else:
    provinceGraph[org1.strip().lower()][org2.strip().lower()]['weight'] += 1
    print str(province[0])+" Graph Completed"
analyzeGraph(provinceGraph,ss)

#Close Summary File
summaryFile.close()

#Close Database Connection
conn.commit()
c.close()
```

Modified plfit.py from Clauset et al. 2009

```python
+ intended to implement a power-law fitting routine as specified in.....
+ http://www.santafe.edu/~aaronc/powerlaws/
+
+ The MLE for the power-law alpha is very easy to derive given knowledge
+ of the lowest value at which a power law holds, but that point is
+ difficult to derive and must be acquired iteratively.

"""
plfit.py – a python power-law fitter based on code by Aaron Clauset
http://www.santafe.edu/~aaronc/powerlaws/
Requires pylab (matplotlib), which requires numpy

example use:
from plfit import plfit

MyPL = plfit(mydata)
MyPL.plotpdf(log=True)

"""

import numpy
import time
import pylab
usefortran=False
cyok=False

import numpy.random as npr
from numpy import log,log10,sum,armin,aramax,exp,min,max

class plfit:
    """
    A Python implementation of the Matlab code
    http://www.santafe.edu/~aaronc/powerlaws/plfit.m
    from http://www.santafe.edu/~aaronc/powerlaws/

```
The output "alpha" is defined such that \( p(x) \sim (x/x_{\text{min}})^{-\alpha} \)

```python
def __init__(self, x, **kwargs):
    """
    Initializes and fits the power law. Can pass "quiet" to turn off output (except for warnings; "silent" turns off warnings)
    """
    self.data = x
    self.plfit(**kwargs)

def alpha_(self, x):
    """
    given a sorted data set and a minimum, returns power law MLE fit
    data is passed as a keyword parameter so that it can be
    vectorized
    """
    x = x[x>=xmin]
    n = sum(x>=xmin)
    a = float(n) / sum(log(x/xmin))
    return a
    return alpha

def kstest_(self, x):
    """
    given a sorted data set and a minimum, returns power law MLE ks-test w/data
    data is passed as a keyword parameter so that it can be
    vectorized
    """
    x = x[x>=xmin]
    n = float(len(x))
    a = float(n) / sum(log(x/xmin))
    cx = numpy.arange(n, dtype='float')/float(n)
    cf = 1-(xmin/x)**a
    ks = max(abs(cf-cx))
    return ks
    return kstest

def plfit(self, nosmall=True, finite=False, quiet=False, silent=False, usefortran=usefortran, usecy=False,
         xmin=None, verbose=False):
    """
    A Python implementation of the Matlab code
    http://www.santafe.edu/~aaronc/powerlaws/plfit.m
    from http://www.santafe.edu/~aaronc/powerlaws/
    """
```
nosmall is on by default; it rejects low s/n points

The fortran version is fastest, the C (cython) version is ~10% slower, and the python version is ~3x slower than the fortran version.

Also, the cython code suffers ~2% numerical error relative to the fortran and python for unknown reasons.

```python
x = self.data
z = numpy.sort(x)
t = time.time()
xmins, argxmins = numpy.unique(z, return_index=True)[:-1]
t = time.time()
if xmin is None:
    if usefortran:
        dat, av = fplfit.plfit(z, int(nosmall))
        goodvals = dat > 0
        sigma = (av - 1) / numpy.sqrt(len(z) - argxmins)
        dat = dat[goodvals]
        av = av[goodvals]
        if not quiet: print "FORTRAN plfit executed in %f seconds" % (time.time() - t)
    elif usecy and cyok:
        dat, av = cplfit.plfit_loop(z, nosmall=nosmall, zunique=xmins, argunique=argxmins)
        goodvals = dat > 0
        sigma = (av - 1) / numpy.sqrt(len(z) - argxmins)
        dat = dat[goodvals]
        av = av[goodvals]
        if not quiet: print "CYTHON plfit executed in %f seconds" % (time.time() - t)
    else:
        av = numpy.asarray(list(map(self.alpha_(z), xmins)), dtype='float')
        dat = numpy.asarray(list(map(self.kstest_(z), xmins)), dtype='float')
        if nosmall:
            # test to make sure the number of data points is high enough
            # to provide a reasonable s/n on the computed alpha
            sigma = (av - 1) / numpy.sqrt(len(z) - argxmins + 1)
            goodvals = sigma < 0.1
            nmax = argmin(goodvals)
            if nmax > 0:
                dat = dat[:nmax]
                av = av[:nmax]
            else:
                print "Not enough data left after flagging – using all data."
        If not quiet: print "PYTHON plfit executed in %f seconds" % (time.time() - t)
```
(time.time() - t)
    self._av = av
    self._xmin_kstest = dat
    self._sigma = sigma
    xmin = xmins[argmin(dat)]
    z = z[z>=xmin]
    n = len(z)
    alpha = 1 + n / sum(log(z/xmin))
    if finite:
        alpha = alpha*(n-1.)/n+1./n
    if n < 50 and not finite and not silent:
        print '(PLFIT) Warning: finite-size bias may be present. N=%i' % n
    ks = max(abs(numpy.arange(n)/float(n) - (1-(xmin/z)**(alpha-1))))
    L = n*log((alpha-1)/xmin) - alpha*sum(log(z/xmin))
    # requires another map... Larr = 29omes29s(len(unique(x))) * log((av-1)/unique(x)) - av*sum
    self._likelihood = L
    self._xmin = xmin
    self._xmins = xmins
    self._alpha = alpha
    self._alphaerr = (alpha-1)/numpy.sqrt(n)
    self._ks = ks  # this ks statistic may not have the same value as
                    # min(dat) because of unique()
    self._ngtx = n

    if not quiet:
        if verbose: print "The lowest value included in the power-law fit, ",
        print "xmin: %g" % xmin,
        if verbose: print "The number of values above xmin, ",
        print "n(>xmin): %i" % n,
        if verbose: print "The derived power-law alpha (p(x)~x^-alpha) with MLE-derived error, ",
        print "alpha: %g +/- %g " % (alpha, self._alphaerr),
        if verbose: print "The log of the Likelihood (the minimized parameter), ",
        print "Log-Likelihood: %g " % L,
        if verbose: print "The KS-test statistic between the best-fit power-law and the data, ",
        print "ks: %g" % (ks)

    return xmin, alpha

def alphavsks(self, autozoom=True, **kwargs):
    """
    Plot alpha versus the ks value for derived alpha. This plot can be used
    as a diagnostic of whether you have derived the ‘best’ fit: if there are
    multiple local minima, your data set may be well suited to a broken
    powerlaw or a different function.
    """
    pylab.plot(self._xmin_kstest, 1 + self._av, '.')
    pylab.errorbar([self._ks], self._alpha, yerr=self._alphaerr, fmt='+')
ax = pylab.gca()
if autozoom:
    ax.set_xlim(0.8*(self._ks), 3*(self._ks))
    ax.set_ylim((self._alpha) - 5*self._alphaerr, (self._alpha) + 5*self._alphaerr)
    ax.set_xlabel(“KS statistic”)  
    ax.set_ylabel(r’$\alpha$’)
pylab.draw()

return ax

def plotcdf(self, x=None, xmin=None, alpha=None, **kwargs):
    """
    Plots CDF and powerlaw
    ""
    if not(x): x = self.data
    if not(xmin): xmin = self._xmin
    if not(alpha): alpha = self._alpha

    x = numpy.sort(x)
    n = len(x)
    xcdf = numpy.arange(n, 0, -1, dtype=’float’)/float(n)

    q = x[x >= xmin]
    fcdf = (q/xmin)**(1-alpha)
    nc = xcdf[argmax(x >= xmin)]
    fcdf_norm = nc*fcdf

    plotx = pylab.linspace(q.min(), q.max(), 1000)
    ploty = (plotx/xmin)**(1-alpha) * nc

    pylab.loglog(x, xcdf, marker=’+’, color=’k’, **kwargs)
    pylab.loglog(plotx, ploty, ’r’, **kwargs)
    #pylab.loglog(q, fcdf_norm, color=’r’, **kwargs)

def plotpdf(self, x=None, xmin=None, alpha=None, nbins=50, dolog=True, dnds=False, 
drawstyle=’steps-post’, **kwargs):
    """
    Plots PDF and powerlaw.
    ""
    if not(x): x = self.data
    if not(xmin): xmin = self._xmin
    if not(alpha): alpha = self._alpha

    x = numpy.sort(x)
    n = len(x)

    pylab.gca().set_xscale(’log’)
    pylab.gca().set_yscale(’log’)

    if dnds:
        hb = pylab.histogram(x, bins=numpy.logspace(log10(min(x)), log10(max(x)), nbins))
        h = hb[0]
        b = hb[1]
        db = hb[1][1:] - hb[1][:-1]
h = h/db
pylab.plot(b[:-1],h,drawstyle=drawstyle,color='k',**kwargs)
    # alpha -= 1
elif dolog:
    hb =
      pylab.hist(x,bins=numpy.logspace(log10(min(x)),log10(max(x)),nbins),log=True,fill=False,edgecolor='k',**kwargs)
        alpha -= 1
    h,b=hb[0],hb[1]
else:
    hb =
      pylab.hist(x,bins=numpy.linspace((min(x)),(max(x)),nbins),fill=False,edgecolor='k',**kwargs)
        h,b=hb[0],hb[1]
    # plotting points are at the center of each bin
    b = (b[1:]+b[:-1])/2.0
    q = x[x>=xmin]
    px = (alpha-1)/xmin * (q/xmin)**(-alpha)
    # Normalize by the median ratio between the histogram and the power-law
    # The normalization is semi-arbitrary; an average is probably just as valid
    plotloc = (b>xmin)*(h>0)
    norm = numpy.median( h[plotloc] / ((alpha-1)/xmin * (b[plotloc]/xmin)**(-alpha))  )
    px = px*norm
    plotx = pylab.linspace(q.min(),q.max(),1000)
    ploty = (alpha-1)/xmin * (plotx/xmin)**(-alpha) * norm
    #pylab.loglog(q,px,'r',**kwargs)
    pylab.loglog(plotx,ploty,'r',**kwargs)
    axlims = pylab.axis()
    pylab.vlines(xmin,axlims[2],max(px),colors='r',linestyle='dashed')
    pylab.gca().set_xlim(min(x),max(x))

def plotppf(self,x=None,xmin=None,alpha=None,dolog=True,**kwargs):
    """
    Plots the power-law-predicted value on the Y-axis against the real values along the X-axis. Can be used as a diagnostic of the fit quality.
    """
    if not(xmin): xmin=self._xmin
    if not(alpha): alpha=self._alpha
    if not(x): x=numpy.sort(self.data[self.data>xmin])
    else: x=numpy.sort(x[x>xmin])
    # N = M^(-alpha+1)
    # M = N^(1/(-alpha+1))
    m0 = min(x)
    N = (1.0+numpy.arange(len(x)))[::-1]
    xmodel = m0 * N**(1/(1-alpha)) / max(N)**(1/(1-alpha))
if dolog:
    pylab.loglog(x,xmodel,'.',**kwargs)
    pylab.gca().set_xlim(min(x),max(x))
    pylab.gca().set_ylim(min(x),max(x))
else:
    pylab.plot(x,xmodel,'.',**kwargs)
    pylab.plot([min(x),max(x)],[min(x),max(x)],'k–')
    pylab.xlabel("Real Value")
    pylab.ylabel("Power-Law Model Value")

def test_pl(self,niter=1e3,**kwargs):
    ""
    Monte-Carlo test to determine whether distribution is consistent with a power law
    Runs through niter iterations of a sample size identical to the input sample size.
    Will randomly select values from the data < xmin. The number of values selected will
    be chosen from a uniform random distribution with p(<xmin) = n(<xmin)/n.
    
    Once the sample is created, it is fit using above methods, then the best fit is used to
    compute a Kolmogorov-Smirnov statistic. The KS stat distribution is compared to the
    KS value for the fit to the actual data, and p = fraction of random
    ks values greater than the data ks value is computed. If p<.1, the data may be
    inconsistent with a powerlaw. A data set of n(>xmin)>100 is required to distinguish a PL
    from an exponential, and n(>xmin)>~300 is required to distinguish a log-normal
    distribution from a PL.
    For more details, see figure 4.1 and section
    ""
    **WARNING** This can take a very long time to run! Execution time scales as
    niter * setsize
    ""
    xmin = self._xmin
    alpha = self._alpha
    self.data = numpy.array(self.data)
    niter = int(niter)
    ntail = sum(self.data >= xmin)
    ntot = len(self.data)
    nnot = ntot-ntail # n(<xmin)
    pnot = nnot/float(ntot) # p(<xmin)
    nonpldata = self.data[self.data<xmin]
    nrandnot = sum( numpy.rand(ntot) < pnot ) # randomly choose how many to
    sample from <xmin
    nrandtail = ntot - nrandnot # and the rest will be sampled
from the powerlaw

ksv = []
for I in xrange(niter):
    # first, randomly sample from power law
    # with caveat!
    Nonplind = numpy.floor(npr.rand(nrandnot)*nnot).astype('int')
    fakenonpl = nonpldata[nonplind]
    randarr = npr.rand(nrandtail)
    fakepl = randarr**(1/(1-alpha)) * xmin
    fakedata = numpy.concatenate([fakenonpl,fakepl])
    # second, fit to powerlaw
    TEST = plfit(fakedata,quiet=True,silent=True,nosmall=True,**kwargs)
    ksv.append(TEST._ks)
ksv = numpy.array(ksv)
    p = (ksv>self._ks).sum() / float(niter)
    self._pval = p
    self._ks_rand = ksv
    print "p(%i) = %0.3f" % (niter,p)
return p,ksv

def plfit_lsq(x,y):
    """
    Returns A and B in y=Ax^B
    http://mathworld.wolfram.com/LeastSquaresFittingPowerLaw.html
    """
    n = len(x)
    btop = n * (log(x)*log(y)).sum() - (log(x)).sum()*(log(y)).sum()
    bbottom = n*(log(x)**2).sum() - (log(x).sum())**2
    b = btop / bbottom
    a = ( log(y).sum() - b * log(x).sum() ) / n
    A = exp(a)
    return A, b

def plexp(x,xm=1,a=2.5):
    """
    CDF(x) for the piecewise distribution exponential x<xmin, powerlaw
    x>=xmin
    This is the CDF version of the distributions drawn in fig 3.4a of Clauset
    et al.
    """
    C = 1/(-xm/(1-a) - xm/a + exp(a)*xm/a)
    Fpl = lambda(X): 1+C*(xm/(1-a))*(X/xm)**(1-a))
    Pexp = lambda(X): C*xm/a*exp(a)-C*(xm/a)*exp(-a*(X/xm-1))
    d=Fpl(x)
    d[x<x]=Pexp(x)
    return d

def plexp_inv(p,xm,a):
    """
    Inverse CDF for a piecewise PDF as defined in eqn. 3.10
of Clauset et al.

```python
C = 1/(-xm/(1 - a) - xm/a + exp(a)*xm/a)
Pxm = 1+C*(xm/(1-a))
x = P*0
x[P>=Pxmx] = xm*( (P[P>=Pxmx]-1) * (1-a)/(C*xm) )**(1/(1-a))  # powerlaw
x[P<Pxm] = (log( (C*xm/a*exp(a)-P[P<Pxm])/(C*xm/a) ) - a) * (-xm/a)  # exp

return x
```
def pl_inv(P,xm,a):
    """
    Inverse CDF for a pure power-law
    """
    x = (1-P)**(1/(1-a)) * xm
    return x

def test_fitter(xmin=1.0,alpha=2.5,niter=500,npts=1000,invcdf=plexp_inv):
    """
    Tests the power-law fitter
    """
    Example (fig 3.4b in Clauset et al.):
    xminin=[0.25,0.5,0.75,1,1.5,2,5,10,50,100]
xmarr,af,ksv,nxarr = plfit.test_fitter(xmin=xminin,niter=1,npts=50000)
    loglog(xminin,xmarr.squeeze(),'x')

    Example 2:
    xminin=[0.25,0.5,0.75,1,1.5,2,5,10,50,100]
xmarr,af,ksv,nxarr = plfit.test_fitter(xmin=xminin,niter=10,npts=1000)
    loglog(xminin,xmarr.mean(axis=0),'x')

    Example 3:
    xminin=[0.25,0.5,0.75,1,1.5,2,5,10,50,100]
xmarr,af,ksv,nxarr = plfit.test_fitter(xmin=1.0,niter=1000,npts=1000)
hist(xmarr.squeeze());
    # Test results:
    # mean(xmarr) = 0.70, median(xmarr)=0.65 std(xmarr)=0.20
    # mean(af) = 2.51 median(af) = 2.49 std(af)=0.14
    # biased distribution; far from correct value of xmin but close to correct alpha

    Example 4:
    xminin=[0.25,0.5,0.75,1,1.5,2,5,10,50,100]
xmarr,af,ksv,nxarr = plfit.test_fitter(xmin=1.0,niter=1000,npts=1000,invcdf=pl_inv)
    print("mean(xmarr): %0.2f median(xmarr): %0.2f std(xmarr): %0.2f" %
    (mean(xmarr),median(xmarr),std(xmarr)))
    print("mean(af): %0.2f median(af): %0.2f std(af): %0.2f" %
    (mean(af),median(af),std(af)))
    # mean(xmarr): 1.19 median(xmarr): 1.03 std(xmarr): 0.35
    # mean(af): 2.51 median(af): 2.50 std(af): 0.07

    """
    xmin = numpy.array(xmin)
    if xmin.shape == ():
        xmin.shape = 1
    lx = len(xmin)
sz = [niter, lx]
xmarr, alphaf_v, ksv, nxarr =
numpy.zeros(sz), numpy.zeros(sz), numpy.zeros(sz), numpy.zeros(sz)
for j in xrange(lx):
    for I in xrange(niter):
        randarr = npr.rand(npts)
        fakedata = invcdf(randarr, xmin[j], alpha)
        TEST = plfit(fakedata, quiet=True, silent=True, nosmall=True)
        alphaf_v[I, j] = TEST._alpha
        ksv[I, j] = TEST._ks
        nxarr[I, j] = TEST._ngtx
        xmarr[I, j] = TEST._xmin

return xmarr, alphaf_v, ksv, nxarr
Appendix B – Centrality for Sub-networks

### Education

<table>
<thead>
<tr>
<th>ORG</th>
<th>DEG</th>
<th>ORG</th>
<th>CLOSENESS</th>
<th>ORG</th>
<th>BETWEENNESS</th>
<th>ORG</th>
<th>EIGENVALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>cruz roja espanola</td>
<td>19</td>
<td>cruz roja espanola</td>
<td>0.159</td>
<td>cruz roja espanola</td>
<td>0.027</td>
<td>cruz roja espanola</td>
<td>0.493</td>
</tr>
<tr>
<td>cruz roja peruana</td>
<td>14</td>
<td>cruz roja peruana</td>
<td>0.122</td>
<td>echo</td>
<td>0.015</td>
<td>cruz roja peruana</td>
<td>0.445</td>
</tr>
<tr>
<td>caritas del peru</td>
<td>9</td>
<td>echo</td>
<td>0.119</td>
<td>coopi</td>
<td>0.008</td>
<td>organismos publicos</td>
<td>0.396</td>
</tr>
<tr>
<td>organismos publicos</td>
<td>8</td>
<td>ocha</td>
<td>0.107</td>
<td>itdg</td>
<td>0.007</td>
<td>construccion s parcco</td>
<td>0.317</td>
</tr>
<tr>
<td>ayne peru trinitarias</td>
<td>7</td>
<td>cooperacion sueca</td>
<td>0.107</td>
<td>unicef</td>
<td>0.006</td>
<td>becon</td>
<td>0.317</td>
</tr>
</tbody>
</table>

### Food Safety

<table>
<thead>
<tr>
<th>ORG</th>
<th>DEG</th>
<th>ORG</th>
<th>CLOSENESS</th>
<th>BETWEENNESS</th>
<th>EIGENVALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ayne peru trinitarias</td>
<td>7</td>
<td>ayne peru trinitarias</td>
<td>0.189</td>
<td>echo</td>
<td>0.017</td>
</tr>
<tr>
<td>solidaridad y desarrollo</td>
<td>7</td>
<td>solidaridad y desarrollo</td>
<td>0.189</td>
<td>ayne peru trinitarias</td>
<td>0.007</td>
</tr>
<tr>
<td>echo</td>
<td>6</td>
<td>echo</td>
<td>0.162</td>
<td>solidaridad y desarrollo</td>
<td>0.007</td>
</tr>
<tr>
<td>personal de la institucion</td>
<td>5</td>
<td>personal de la institucion</td>
<td>0.147</td>
<td>prodeiso</td>
<td>0.005</td>
</tr>
<tr>
<td>equipo de emergencia ayne peru</td>
<td>5</td>
<td>equipo de emergencia ayne peru</td>
<td>0.147</td>
<td>personal de la institucion</td>
<td>0.000</td>
</tr>
</tbody>
</table>
### Health

<table>
<thead>
<tr>
<th>ORG</th>
<th>DEGREE</th>
<th>ORG</th>
<th>CLOSENESS</th>
<th>ORG</th>
<th>BETWEEN</th>
<th>ORG</th>
<th>EIGEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>fundacion terre des 37omes (tdh)</td>
<td>19</td>
<td>unicef</td>
<td>0.264</td>
<td>unicef</td>
<td>0.069</td>
<td>unicef</td>
<td>0.2</td>
</tr>
<tr>
<td>unicef</td>
<td>19</td>
<td>fundacion terre des 37omes (tdh)</td>
<td>0.260</td>
<td>diresa</td>
<td>0.051</td>
<td>tierra de hombres</td>
<td>0.2</td>
</tr>
<tr>
<td>tierra de hombres</td>
<td>18</td>
<td>tierra de hombres</td>
<td>0.242</td>
<td>fundacion terre des 37omes (tdh)</td>
<td>0.037</td>
<td>unfpa</td>
<td>0.2</td>
</tr>
<tr>
<td>unfpa</td>
<td>15</td>
<td>unfpa</td>
<td>0.237</td>
<td>coopi</td>
<td>0.037</td>
<td>cedappfun</td>
<td>0.2</td>
</tr>
<tr>
<td>fundacion richmond</td>
<td>15</td>
<td>richmond</td>
<td>0.229</td>
<td>oxfam</td>
<td>0.035</td>
<td>acjfund</td>
<td>0.2</td>
</tr>
</tbody>
</table>

### Housing

<table>
<thead>
<tr>
<th>ORG</th>
<th>DEGREE</th>
<th>ORG</th>
<th>CLOSENESS</th>
<th>ORG</th>
<th>BETWEEN</th>
<th>ORG</th>
<th>EIGEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>municipalidad pisco</td>
<td>19</td>
<td>pnud</td>
<td>0.177</td>
<td>echo</td>
<td>0.201</td>
<td>empresas privadas</td>
<td>0.359</td>
</tr>
<tr>
<td>predes</td>
<td>15</td>
<td>echo</td>
<td>0.177</td>
<td>pnud</td>
<td>0.164</td>
<td>municipalidad pisco</td>
<td>0.280</td>
</tr>
<tr>
<td>empresas privadas</td>
<td>15</td>
<td>indeci</td>
<td>0.171</td>
<td>indeci</td>
<td>0.155</td>
<td>forsur</td>
<td>0.273</td>
</tr>
<tr>
<td>forsur</td>
<td>13</td>
<td>itdg</td>
<td>0.154</td>
<td>itdg</td>
<td>0.122</td>
<td>banmat</td>
<td>0.270</td>
</tr>
<tr>
<td>indeci</td>
<td>12</td>
<td>ocha</td>
<td>0.154</td>
<td>predes</td>
<td>0.112</td>
<td>indeci</td>
<td>0.266</td>
</tr>
</tbody>
</table>

### Livelihood

<table>
<thead>
<tr>
<th>ORG</th>
<th>DEGREE</th>
<th>ORG</th>
<th>CLOSENESS</th>
<th>ORG</th>
<th>BETWEEN</th>
<th>ORG</th>
<th>EIGEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ayne peru trinitarias</td>
<td>10</td>
<td>ayne peru trinitarias</td>
<td>0.147</td>
<td>echo</td>
<td>0.01</td>
<td>ayne peru trinitarias</td>
<td>0.60</td>
</tr>
<tr>
<td>solidaridad y desarrollo</td>
<td>10</td>
<td>solidaridad y desarrollo</td>
<td>0.147</td>
<td>fao</td>
<td>0.01</td>
<td>solidaridad y desarrollo</td>
<td>0.60</td>
</tr>
<tr>
<td>echo</td>
<td>7</td>
<td>echo</td>
<td>0.119</td>
<td>ayne peru trinitarias</td>
<td>0.00</td>
<td>equipo de oficina ayne sur gobierno de las islas baleares</td>
<td>0.34</td>
</tr>
<tr>
<td>personal de la institucion</td>
<td>5</td>
<td>fao</td>
<td>0.099</td>
<td>solidaridad y desarrollo</td>
<td>0.00</td>
<td>equipo de emergencia fons mallorqui de solidaritat</td>
<td>0.22</td>
</tr>
<tr>
<td>equipo de emergencia</td>
<td>4</td>
<td>personal de la institucion</td>
<td>0.098</td>
<td>coopi</td>
<td>0.00</td>
<td>consell de mallorca</td>
<td>0.22</td>
</tr>
<tr>
<td>ayne peru</td>
<td>5</td>
<td>institucion</td>
<td>0.098</td>
<td>coopi</td>
<td>0.00</td>
<td>consell de mallorca</td>
<td>0.22</td>
</tr>
</tbody>
</table>
### Protection

<table>
<thead>
<tr>
<th>Org</th>
<th>Degree</th>
<th>Org</th>
<th>Closeness</th>
<th>Org</th>
<th>Between</th>
<th>Org</th>
<th>Eigen</th>
</tr>
</thead>
<tbody>
<tr>
<td>echo</td>
<td>11</td>
<td>echo</td>
<td>0.247</td>
<td>echo</td>
<td>0.108</td>
<td>unfp3</td>
<td>0.60</td>
</tr>
<tr>
<td>unicef</td>
<td>11</td>
<td>cruz roja peruana</td>
<td>0.206</td>
<td>cruz roja peruana</td>
<td>0.051</td>
<td>38omes38s ramos</td>
<td>0.60</td>
</tr>
<tr>
<td>cruz roja peruana</td>
<td>9</td>
<td>cruz roja espanola</td>
<td>0.192</td>
<td>unfp3</td>
<td>0.042</td>
<td>union europea</td>
<td>0.40</td>
</tr>
<tr>
<td>care peru</td>
<td>6</td>
<td>unicef</td>
<td>0.181</td>
<td>unicef</td>
<td>0.028</td>
<td>naciones unidas</td>
<td>0.24</td>
</tr>
<tr>
<td>cruz roja espanola</td>
<td>6</td>
<td>unfp3</td>
<td>0.176</td>
<td>cruz roja espanola</td>
<td>0.017</td>
<td>union europea</td>
<td>1</td>
</tr>
</tbody>
</table>

### Water & Sanitation

<table>
<thead>
<tr>
<th>Org</th>
<th>Degree</th>
<th>Org</th>
<th>Closeness</th>
<th>Org</th>
<th>Between</th>
<th>Org</th>
<th>Eigen</th>
</tr>
</thead>
<tbody>
<tr>
<td>unicef</td>
<td>13</td>
<td>unicef</td>
<td>0.224</td>
<td>coopi</td>
<td>0.177</td>
<td>echo</td>
<td>0.556</td>
</tr>
<tr>
<td>echo</td>
<td>11</td>
<td>coopi</td>
<td>0.224</td>
<td>emapisco</td>
<td>0.121</td>
<td>acf</td>
<td>0.551</td>
</tr>
<tr>
<td>coopi</td>
<td>10</td>
<td>oxfam</td>
<td>0.217</td>
<td>echo</td>
<td>0.118</td>
<td>donantes espanoles</td>
<td>0.551</td>
</tr>
<tr>
<td>emapisco</td>
<td>10</td>
<td>emapisco</td>
<td>0.200</td>
<td>unicef</td>
<td>0.101</td>
<td>accion contra el hambre</td>
<td>0.279</td>
</tr>
<tr>
<td>oxfam</td>
<td>9</td>
<td>semapach</td>
<td>0.189</td>
<td>oxfam</td>
<td>0.059</td>
<td>itdg</td>
<td>0.028</td>
</tr>
</tbody>
</table>

### Chincha

<table>
<thead>
<tr>
<th>Org</th>
<th>Degree</th>
<th>Org</th>
<th>Closeness</th>
<th>Org</th>
<th>Between</th>
<th>Org</th>
<th>Eigen</th>
</tr>
</thead>
<tbody>
<tr>
<td>ayne peru trinitarias</td>
<td>17</td>
<td>coopi</td>
<td>0.179</td>
<td>coopi</td>
<td>0.118</td>
<td></td>
<td></td>
</tr>
<tr>
<td>solidaridad y desarrollo</td>
<td>17</td>
<td>unicef</td>
<td>0.158</td>
<td>unicef</td>
<td>0.082</td>
<td></td>
<td></td>
</tr>
<tr>
<td>unicef</td>
<td>17</td>
<td>echo</td>
<td>0.151</td>
<td>echo</td>
<td>0.071</td>
<td></td>
<td></td>
</tr>
<tr>
<td>coopi</td>
<td>15</td>
<td>semapach</td>
<td>0.141</td>
<td>cooperacion italiana</td>
<td>0.044</td>
<td></td>
<td></td>
</tr>
<tr>
<td>echo</td>
<td>9</td>
<td>drei</td>
<td>0.140</td>
<td>itdg</td>
<td>0.044</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Ica

<table>
<thead>
<tr>
<th>Org</th>
<th>Degree</th>
<th>Org</th>
<th>Closeness</th>
<th>Org</th>
<th>Between</th>
<th>Org</th>
<th>Eigen</th>
</tr>
</thead>
<tbody>
<tr>
<td>cruz roja peruana</td>
<td>23</td>
<td>cruz roja peruana</td>
<td>0.156</td>
<td>cruz roja peruana</td>
<td>0.132</td>
<td>echo</td>
<td>0.5</td>
</tr>
<tr>
<td>cruz roja espanola</td>
<td>20</td>
<td>federacion internacional</td>
<td>0.155</td>
<td>federacion internacional</td>
<td>0.122</td>
<td>pma</td>
<td>0.3</td>
</tr>
<tr>
<td>unicef</td>
<td>17</td>
<td>cidap</td>
<td>0.146</td>
<td>cidap</td>
<td>0.113</td>
<td>cerf</td>
<td>0.3</td>
</tr>
<tr>
<td>federacion internacional</td>
<td>10</td>
<td>unicef</td>
<td>0.139</td>
<td>fepromu ica</td>
<td>0.107</td>
<td>acf</td>
<td>0.3</td>
</tr>
<tr>
<td>predes</td>
<td>9</td>
<td>ficr</td>
<td>0.138</td>
<td>trocaire</td>
<td>0.095</td>
<td>accion contra el hambre</td>
<td>0.3</td>
</tr>
</tbody>
</table>
## Pisco

<table>
<thead>
<tr>
<th>Org</th>
<th>Degree</th>
<th>Closedness</th>
<th>Betweenness</th>
<th>Eigenvector</th>
</tr>
</thead>
<tbody>
<tr>
<td>unicef</td>
<td>36</td>
<td>0.325</td>
<td>0.284</td>
<td>0.2</td>
</tr>
<tr>
<td>municipalidad pisco</td>
<td>26</td>
<td>0.294</td>
<td>0.258</td>
<td>0.2</td>
</tr>
<tr>
<td>fundacion terre des 39omes (tdh)</td>
<td>19</td>
<td>0.293</td>
<td>0.224</td>
<td>0.2</td>
</tr>
<tr>
<td>tierra de hombres</td>
<td>18</td>
<td>0.279</td>
<td>0.192</td>
<td>0.2</td>
</tr>
<tr>
<td>minsa</td>
<td>17</td>
<td>0.278</td>
<td>0.147</td>
<td>0.2</td>
</tr>
</tbody>
</table>
Appendix C – Graphs of Sub-networks

Education Network

Food Safety Network
Appendix D – Power Law Tests

[Graphs showing log-log degree distributions for Education and Food Safety]