Supplementary Information

Mapping (mis)alignment within a collaborative network using homophily metrics

Kimberly Pugel[a], Amy Javernick-Will[a]*, Cliff Nyaga[b][¶], Muhammed Mussa[c][¶], Desta Dimtse[c][¶], Lucia Henry[c][¶], Karl Linden[a]

[a] Department of Civil, Environmental, and Architectural Engineering, University of Colorado Boulder, USA

[b] FundiFix Ltd, Kenya

[c] Tetra Tech, Ethiopia

[d] Tetra Tech, USA

2022

*Corresponding author: amy.javernick@colorado.edu
[¶]These authors contributed equally to this work.

Pages S1 – S7

Total Number of Tables: 1, Total Number of Figures: 0

Table of Contents

1. Descriptive Statistics

The following descriptive statistics (Borgatti et al., 2009; Scott, 2017) quantify the network in a variety of ways and are useful for interpreting the homophily analysis (Table 1). Descriptive statistics that are generally used to describe the size of the network include number of actors (*nodes*) and number of relationships (*ties*). The spread of the network is described by *diameter,* which is the maximum length of the shortest path between two nodes, where a path is the number of steps it takes to get from one node to another. How 'crowded' the network is can be described using *density* and *centralization. Density* is the number of ties that exist divided by the total possible number of ties. *Centralization* is the overall cohesion of the graph, or how centered nodes are around a particular set of nodes. The *transitivity* represents the concept of "I tend to be friends with my friends' friends" in that if three actors have two connections between them, the transitivity represents the likelihood of the third connection being made. Finally, fragmentation of the network is reflected by the number of *connected components*, in which components are a set of nodes connected to one another and where more than one connected component means the graph is segmented or siloed. These calculations are standard practice for network studies (Wasserman and Faust, 1994; Scott, 2017) and were calculated through R's igraph package (Csardi and Nepusz, 2006). Detailed R code for all network manipulations, calculations, and visualizations is in the Supplementary Information.

The network statistics show that the Kitui and Debre Birhan networks are relatively small for network studies (9 and 22 actors, respectively), that are relatively well-connected with low average path lengths of less than 2. Information sharing occurs more frequently than support in both cases, as all networks have a higher density, or number of ties that exist divided by the total number of possible ties in the network. The networks remain somewhat dispersed, with low centralization

scores <0.5 for information-sharing and support, indicating that the ties are not highly clustered around a core group of individuals. Transitivity represents the likelihood of a third connection being made if three actors have two connections between them, representing the notion that "I tend to be friends with my friends' friends". Transitivity ranges quite widely across cases, relationships, and instances in time, with the highest rate represented in the Debre Birhan 2020 networks (0.76) and the Kitui 2018 information-sharing network (0.80). Because these networks had low average path lengths and high transitivity, these networks may exhibit some "small-world" properties where ideas may flow relatively quickly through the networks. This theory of small world networks was first developed by Watts and Strogatz (1999) to explain disease spreading in highly-connected networks where actors are only a few steps away from each other.

Table 1. Descriptive statistics of networks. Definitions of various parameters are provided in the preceding paragraph.

Network		Tie type	# Nodes	# Ties	Dia-meter	Avg. path length	Density	Centrali-zation	Transitivity	# Connected components
Debre Birhan	2018	information	9	15	3	1.8	0.21	0.29	0.72	2
		support	9	16	3	1.4	0.22	0.28	0.39	1
	2019	information	9	25	3	1.7	0.35	0.44	0.62	1
		support	9	24	3	1.7	0.33	0.38	0.47	1
	2020	information	9	42	2	1.4	0.58	0.38	0.76	1
		support	9	18	3	1.7	0.25	0.40	0.76	1
Kitui	2018	information	22	319	2	1.3	0.69	0.26	0.82	1
		support	22	105	5	2.1	0.23	0.35	0.48	1
	2020	information	22	284	3	1.5	0.61	0.48	0.70	1
		support	22	148	4	1.8	0.32	0.41	0.53	1

2. R Code

```
install.packages("tables")
install.packages("igraph")
library(tables)
library(igraph)
```

2.1. Debre Birhan

```
nodes<-read.csv("DB-nodes.csv",header=TRUE, as.is=TRUE)
info.b<-read.csv("DB-info-b.csv", header=T, as.is=T)
info.m<-read.csv("DB-info-m.csv", header=T, as.is=T)
info.e<-read.csv("DB-info-e.csv", header=T, as.is=T)
support.b<-read.csv("DB-support-b.csv", header=T, as.is=T)
support.m<-read.csv("DB-support-m.csv", header=T, as.is=T)
support.e<-read.csv("DB-support-e.csv", header=T, as.is=T)

info.b.net<-graph_from_data_frame(d=info.b,vertices=nodes,directed=T)
info.m.net<-graph_from_data_frame(d=info.m,vertices=nodes,directed=T)
info.e.net<-graph_from_data_frame(d=info.e,vertices=nodes,directed=T)
support.b.net<-graph_from_data_frame(d=support.b,vertices=nodes,directed=T)
support.m.net<-graph_from_data_frame(d=support.m,vertices=nodes,directed=T)
support.e.net<-graph_from_data_frame(d=support.e,vertices=nodes,directed=T)

pib<-assortativity_nominal(info.b.net, V(info.b.net)$Prioritynum_b)
pim<-assortativity_nominal(info.m.net, V(info.m.net)$Prioritynum_m)
pie<-assortativity_nominal(info.e.net, V(info.e.net)$Prioritynum_e)
psb<-assortativity_nominal(support.b.net, V(support.b.net)$Prioritynum_b)
psm<-assortativity_nominal(support.m.net, V(support.m.net)$Prioritynum_m)
pse<-assortativity_nominal(support.e.net, V(support.e.net)$Prioritynum_e)

pscores<-c(pib,pim,pie,psb,psm,pse)
ptab<-matrix(data=pscores, nrow=2, ncol=3, byrow=TRUE)
rownames(ptab)<- c('info', 'support')
colnames(ptab)<-c('baseline', 'midline', 'endline')
ptab<-as.table(ptab)

##export as table
write.table(ptab,file="priority_assortativity.csv",append=FALSE,
quote=FALSE,sep=",",eol="\r",na="NA", dec=".", row.names=TRUE, col.names=TRUE,
qmethod="escape")


##network statistics

###baseline

##number of nodes
gorder(info.b.net)

##number of ties by network
gsize(info.b.net)
gsize(support.b.net)
```

```
##longest geodesic
diameter(info.b.net, directed = TRUE, unconnected = TRUE, weights = NULL)
diameter(support.b.net, directed = TRUE, unconnected = TRUE, weights = NULL)

##average path length for small network calcs
average.path.length(info.b.net)
average.path.length(support.b.net)

##density of network
edge_density(info.b.net, loops = FALSE)
edge_density(support.b.net, loops = FALSE)

##centralization by degree

centr_degree(info.b.net, mode="all", loops = FALSE, normalized= TRUE)
centr_degree(support.b.net, mode="all", loops = FALSE, normalized= TRUE)

##average clustering coefficient
transitivity(info.b.net, type = "average")
transitivity(support.b.net, type = "average")

##connceted components
components(info.b.net, mode = "weak")
components(support.b.net, mode = "weak")


###midline

##number of nodes
gorder(info.m.net)

##number of ties by network
gsize(info.m.net)
gsize(support.m.net)

##longest geodesic
diameter(info.m.net, directed = TRUE, unconnected = TRUE, weights = NULL)
diameter(support.m.net, directed = TRUE, unconnected = TRUE, weights = NULL)

##average path length for small network calcs
average.path.length(info.m.net)
average.path.length(support.m.net)

##density of network
edge_density(info.m.net, loops = FALSE)
edge_density(support.m.net, loops = FALSE)

##centralization by degree

centr_degree(info.m.net, mode="all", loops = FALSE, normalized= TRUE)
centr_degree(support.m.net, mode="all", loops = FALSE, normalized= TRUE)
```

```
##average clustering coefficient
transitivity(info.m.net, type = "average")
transitivity(support.m.net, type = "average")


##connceted components
components(info.m.net, mode = "weak")
components(support.m.net, mode = "weak")

###endline

##number of nodes
gorder(info.e.net)

##number of ties by network
gsize(info.e.net)
gsize(support.e.net)

##longest geodesic
diameter(info.e.net, directed = TRUE, unconnected = TRUE, weights = NULL)
diameter(support.e.net, directed = TRUE, unconnected = TRUE, weights = NULL)

##average path length for small network calcs
average.path.length(info.e.net)
average.path.length(support.e.net)

##density of network
edge_density(info.e.net, loops = FALSE)
edge_density(support.e.net, loops = FALSE)

##centralization by degree

centr_degree(info.e.net, mode="all", loops = FALSE, normalized= TRUE)
centr_degree(support.e.net, mode="all", loops = FALSE, normalized= TRUE)

##average clustering coefficient
transitivity(info.e.net, type = "average")
transitivity(support.e.net, type = "average")

##connceted components
components(info.e.net, mode = "weak")
components(support.e.net, mode = "weak")
```

2.2. Kitui

```
nodes<-read.csv("KT-nodes.csv",header=TRUE, as.is=TRUE)
info.b<-read.csv("KT-info-b.csv", header=T, as.is=T)
info.e<-read.csv("KT-info-e.csv", header=T, as.is=T)
support.b<-read.csv("KT-support-b.csv", header=T, as.is=T)
support.e<-read.csv("KT-support-e.csv", header=T, as.is=T)

info.b.net<-graph_from_data_frame(d=info.b,vertices=nodes,directed=T)
info.e.net<-graph_from_data_frame(d=info.e,vertices=nodes,directed=T)
```

```r
support.b.net<-graph_from_data_frame(d=support.b,vertices=nodes,directed=T)
support.e.net<-graph_from_data_frame(d=support.e,vertices=nodes,directed=T)

pib<-assortativity_nominal(info.b.net, V(info.b.net)$bprioritynum)
pie<-assortativity_nominal(info.e.net, V(info.e.net)$eprioritynum)
psb<-assortativity_nominal(support.b.net, V(support.b.net)$bprioritynum)
pse<-assortativity_nominal(support.e.net, V(support.e.net)$eprioritynum)

pscores<-c(pib,pie,psb,pse)
ktptab<-matrix(data=pscores, nrow=2, ncol=2, byrow=TRUE)
rownames(ktptab)<- c('info', 'support')
colnames(ktptab)<-c('baseline', 'endline')
ktptab<-as.table(ktptab)

ktptab

##network statistics

###baseline

##number of nodes
gorder(info.b.net)

##number of ties by network
gsize(info.b.net)
gsize(support.b.net)

##longest geodesic
diameter(info.b.net, directed = TRUE, unconnected = TRUE, weights = NULL)
diameter(support.b.net, directed = TRUE, unconnected = TRUE, weights = NULL)

##average path length for small network calcs
average.path.length(info.b.net)
average.path.length(support.b.net)

##density of network
edge_density(info.b.net, loops = FALSE)
edge_density(support.b.net, loops = FALSE)

##centralization by degree

centr_degree(info.b.net, mode="all", loops = FALSE, normalized= TRUE)
centr_degree(support.b.net, mode="all", loops = FALSE, normalized= TRUE)

##average clustering coefficient
transitivity(info.b.net, type = "average")
transitivity(support.b.net, type = "average")

##connceted components
components(info.b.net, mode = "weak")
components(support.b.net, mode = "weak")
```

```
###endline

##number of nodes
gorder(info.e.net)

##number of ties by network
gsize(info.e.net)
gsize(support.e.net)

##longest geodesic
diameter(info.e.net, directed = TRUE, unconnected = TRUE, weights = NULL)
diameter(support.e.net, directed = TRUE, unconnected = TRUE, weights = NULL)

##average path length for small network calcs
average.path.length(info.e.net)
average.path.length(support.e.net)

##density of network
edge_density(info.e.net, loops = FALSE)
edge_density(support.e.net, loops = FALSE)

##centralization by degree

centr_degree(info.e.net, mode="all", loops = FALSE, normalized= TRUE)
centr_degree(support.e.net, mode="all", loops = FALSE, normalized= TRUE)

##average clustering coefficient
transitivity(info.e.net, type = "average")
transitivity(support.e.net, type = "average")

##connceted components
components(info.e.net, mode = "weak")
components(support.e.net, mode = "weak")
```

3. Description of attached data files

Files attached:

File name	Description
S1_DB-nodes.csv	Nodelist for Debre Birhan Learning Alliance
S2_DB-info-2018.csv	Information-sharing network for Debre Birhan in 2018
S3_DB-info-2019.csv	Information-sharing network for Debre Birhan in 2019
S4_DB-info-2020.csv	Information-sharing network for Debre Birhan in 2020
S5_DB-support-2018.csv	Support network for Debre Birhan in 2018
S6_DB-support-2019.csv	Support network for Debre Birhan in 2019
S7_DB-support-2020.csv	Support network for Debre Birhan in 2020
S8_KT-nodes.csv	Nodelist for Kitui WASH Forum
S9_KT-info-2018.csv	Information-sharing network for Kitui in 2018
S10_KT-info-2020.csv	Information-sharing network for Kitui in 2020
S11_KT-support-2018.csv	Support network for Kitui in 2018
S12_KT-support-2020.csv	Support network for Kitui in 2020