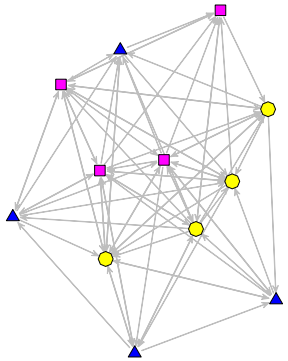
SURF:
Using Social Network Analysis to Reducing Risk

Justin Castagna



Agenda

- Introduction
- My journey with Analytics & R…so far
- Definition
- How to interpret
- Data code and basic visualisation
- Data Structure
- Measurement
- Scenarios – Applied examples
  - UDT
  - Risk Mapping
- Questions

My journey with Analytics & R…so far

- My Name is Justin Castagna
- http://au.linkedin.com/in/justincastagna

- Accountant – Not a programmer or statistician.
- Background with Superannuation, Funds management.

- Largely self taught with ‘R’.
- Began this journey by being involved in quality program, Lean-Six Sigma at AMP. Discovered the value of applied stats with Minitab.
- Changed jobs - no statistics tools.
- Discovered open-source “R” environment in 2006 – right price
- Slow learning curve with command line interface. Often frustrated by my lack of progress it getting ‘R’ to do what I could imagine and what I had been able to do with Minitab. Perseverance and lots of small steps.
- “R” is like a Swiss army knife for your work toolkit


The SIX SIGMA WAY
How GE, Motorola, and Other Top Companies Are Honing Their Performance
Peter S. Pande, Robert P. Neuman, Roland R. Cavanagh
HARD FACTS
DANGEROUS HALF-TRUTHS & TOTAL NONSENSE
PROFITING FROM EVIDENCE-BASED MANAGEMENT
Jeffrey Pfeffer
Robert I. Sutton
DRIVING RESULTS THROUGH SOCIAL NETWORKS
HOW TOP ORGANIZATIONS LEVERAGE NETWORKS FOR PERFORMANCE AND GROWTH
ROB CROSS
ROBERT J. THOMAS
The Hidden Power of Social Networks
Understanding How Work Really Gets Done in Organizations
Rob Cross • Andrew Parker
Statistics and Computing
Peter Dalgaard
Introductory Statistics with R
Springer
Cambridge Series in Statistical and Probabilistic Mathematics
Data Analysis and Graphics Using R
An Example-Based Approach
Second Edition
John Maindonald and John Braun
A FIRST COURSE IN STATISTICAL PROGRAMMING WITH R
R-bloggers
R news and tutorials contributed by (320) R bloggers
Plotting forecast() objects in ggplot part 2: Visualize Observations, Fits, and Forecasts
Revolutions
R Chart featured in Facebook IPO
DOUGLAS W. HUBBARD
THE FAILURE OF RISK MANAGEMENT
Why It's Broken and How to Fix It
The Visual Display of Quantitative Information
EDWARD R. TUFTE
Visual Business Intelligence
DATA DRIVEN
THOMAS C. REDMAN
Thomas H. Davenport • Jeanne G. Harris
Competing on Analytics
The New Science of Winning
Harvard Business School Press
TRANSFORMING PERFORMANCE MEASUREMENT
Rethinking the Way We Measure and Drive Organizational Success
DEAN R. SPITZER, PH.D.
HOW TO MEASURE ANYTHING
FINDING THE VALUE OF INTANGIBLES IN BUSINESS
Douglas W. Hubbard

Definition:
What is Social Network Analysis?

- Social network analysis is based on an assumption of the importance of relationships among interacting units. The social network perspective encompasses theories, models, and applications that are expressed in terms of relational concepts or processes. Along with growing interest and increased use of network analysis has come a consensus about the central principles underlying the network perspective. In addition to the use of relational concepts, we note the following as being important:
- Actors and their actions are viewed as interdependent rather than independent, autonomous units
- Relational ties (linkages) between actors are channels for transfer or "flow" of resources (either material or nonmaterial)
- Network models focusing on individuals view the network structural environment as providing opportunities for or constraints on individual action
- Network models conceptualize structure (social, economic, political, and so forth) as lasting patterns of relations among actors
- The unit of analysis in network analysis is not the individual, but an entity consisting of a collection of individuals and the linkages among them. Network methods focus on dyads (two actors and their ties), triads (three actors and their ties), or larger systems (subgroups of individuals, or entire networks.
- Wasserman, S. and K. Faust, 1994, *Social Network Analysis.* Cambridge: Cambridge University Press.

How to read

Vertex
aka
Actor
Node
Point

vertex.col
vertex.sides
vertex.rot
vertex.cex
vertex.border
vertex.lty

Edge
aka
Tie
Link
Line

Relationship indicated by arrow direction

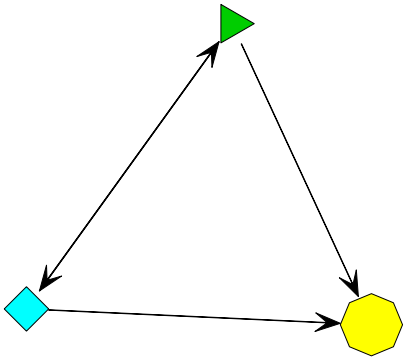
Introduction: Data and Code

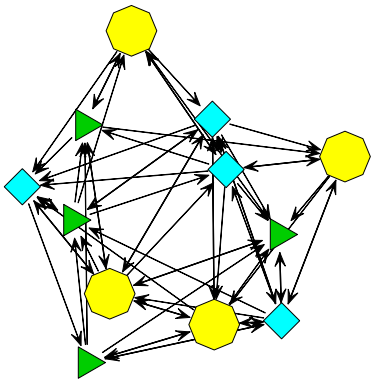
library(sna) # There are other packages on R available

gplot() # Key plotting command

gplot(rgraph(3)) # creates randomised graph, 3 indicates number of nodes

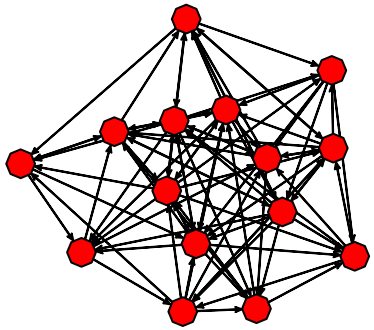
gplot(rgraph(3),vertex.col=c(3,5,7), vertex.cex=c(1,1,2), vertex.sides=c(3,4,8))



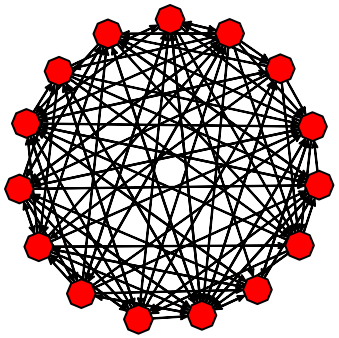


Alternate Plot Visualisations

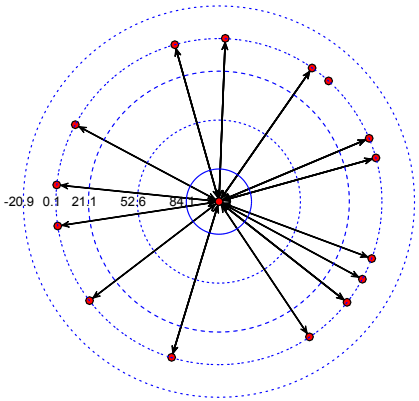
Random



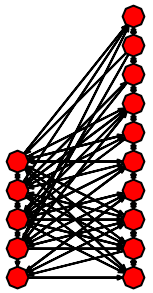
Circle



Target



Co-ordinates



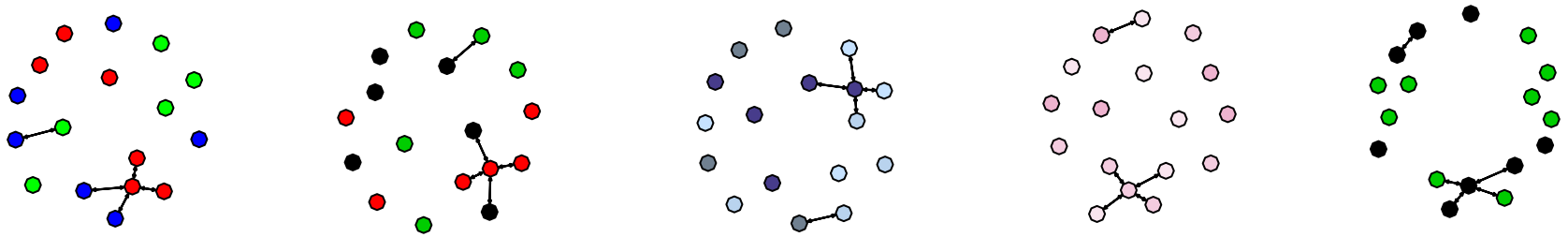
1. gplot(rgraph(15))

2. gplot(rgraph(15),mode="circle")

3. mydata=gplot(rgraph(15))
   gplot.target(mydata,betweenness(mydata)) #Note metric used

4. locx=c(1,1,1,1,1,5,5,5,5,5,5,5,5,5,5) locy=c(1,2,3,4,5,1,2,3,4,5,6,7,8,9,10)
   Loc=matrix(cbind(locx,locy),nrow=15,ncol=2)
   gplot(rgraph(15),coord=Loc,jitter=FALSE)

Vertex Colour



mydata=gplot(rgraph(15),vertex.cex=4)

1. gplot(mydata, vertex.col=c("blue","red","green"))
2. gplot(mydata, vertex.col=c(1,2,3))
3. gplot(mydata, vertex.col=colours()[c(598,599,600,601)])
4. gplot(mydata,vertex.col=c(rgb(226,126,174,c(50,100,150),maxColorValue=255)))
5. gplot(mydata, vertex.cex=4, vertex.col=ifelse(mydata>-2,1,3))

Data Structure

- gplot(dat, g = 1, gmode = "digraph", diag = FALSE,
- label = NULL, coord = NULL, jitter = TRUE, thresh = 0,
- usearrows = TRUE, mode = "fruchtermanreingold",
- displayisolates = TRUE, interactive = FALSE, interact.bycomp = FALSE,
- xlab = NULL, ylab = NULL, xlim = NULL, ylim = NULL, pad = 0.2,
- label.pad = 0.5, displaylabels = !is.null(label), boxed.labels = FALSE,
- label.pos = 0, label.bg = "white", vertex.sides = NULL, vertex.rot = 0,
- arrowhead.cex = 1, label.cex = 1, loop.cex = 1, vertex.cex = 1,
- edge.col = 1, label.col = 1, vertex.col = NULL, label.border = 1,
- vertex.border = 1, edge.lty = 1, label.lty = NULL, vertex.lty = 1,
- edge.lwd = 0, label.lwd = par("lwd"), edge.len = 0.5,
- edge.curve = 0.1, edge.steps = 50, loop.steps = 20,
- object.scale = 0.01, uselen = FALSE, usecurve = FALSE,
- suppress.axes = TRUE, vertices.last = TRUE, new = TRUE,
- layout.par = NULL, ...)

Matrix of Relationships

Actor_1	Actor_2	Actor_3	Actor_4	Actor_5	Actor_6	Actor_7
0	1	0	0	0	0	0
1	0	9	1	0	0	0
0	0	0	0	0	0	0
0	5	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	9	0

List of Actor's Attributes

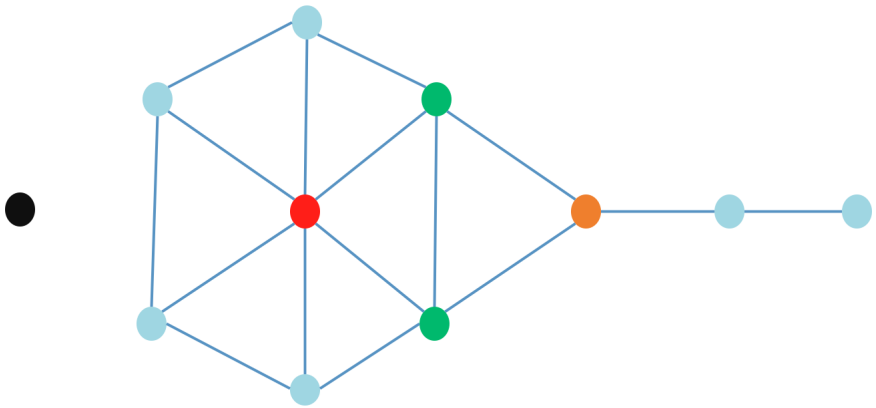
ID	Name	Office	My_Colour	MyShape
1	Actor_1	Australia	1	15
2	Actor_2	Australia	1	
3	Actor_3	Australia	1	
4	Actor_4	Australia	1	
5	Actor_5	Asia	2	
6	Actor_6	Asia	2	
7	Actor_7	Europe	3	

Measurement

- Behind the visualisations is a lot of measurement. Measure the centrality of nodes to help understand the various roles of groups in network.
  - Roles include clusters, brokers, boundary spanners, peripheral players and isolates

- Degree – The number of direct connections a node has. Shows how active the node is in the entire network.
  - Direct connections can show the clique a node may belowng to.
  - More not necessarily better, its important how the node is connected to other nodes.
- Betweenness – measures influence of what does/does not flow in network.
  - Indicates broker role connecting between different groups
- Closeness – the shortest path to other nodes.
- Network Centralization – provides insight into the network structure of individual nodes and entire network.
  - Centralised structure has very few nodes binding network together. Single points of failure om key nodes (indicated by high betweenness). If removed you get network fragmentation as split into smaller groups.
  - Less centralised networks are more resilient to failure.

Interpretation



- Hub
- Broker
- Boundary Spanner
- Isolate

- Location of node in network is important.
- Highest scores - Degree (red), Betweenness (amber), Closeness (green).
- Is possible to be not connected (black).
- Other networks may not be measured in this instance thus other players (blue) may be far more important than shown.

Other commands to explore

- gplot
  - thresh = 0,
  - displayisolates = FALSE,
  - interactive = TRUE,
- gplot3d
- sociomatrixplot
- Brokerage
  - brokerage(Matrix1,A_Type) #UDT Scenario
    - MyM=brokerage(Matrix1,A_Type)
    - MyM$n
    - MyM$raw.nli[,2] #$raw.nli w_O

Scenarios

- Two Scenarios
- UDT
- Risk Mapping

- Q: How do you use SNA to reduce risk?
- A: Use SNA to;
  - Visualise and measure nodes to help facilitate a dialogue on risk, and
  - Use as tool to focus scarce resources on remediation of key nodes.

Scenario: UDT

- Organisation has many User Defined Tool's.
  - Excel spreadsheets or Access Databases that are used in daily work.
- With UDT ease of use (especially Excel), they enable easy workarounds to system limitations to become entrenched.
- With quantity of UDT's its easy to loose visibility of their pervasive use across an organisation.
- Inherent risk to organisation. Often Risk team will try to mitigate with a centralised list to control spread and provide some visibility to senior management. May measure risk with a traffic light, "RAG", rating.

- A spreadsheet list with a thousand rows does not do justice to the issue. Difficult to visualise and few will attempt to measure .
- Use SNA to put organisation on page. Visualise the problem.

Scenario: UDT

UDTID	Name	Owner	Team	Type	SRC Systems	EXP Systems
1	Reconiliation manager	James C	Corp Actions	Excel	Coral, UDT98	UDT269
2	Import tool	John S	Corp Actions	Excel	Coral, Pacific	N/A
3	Rec tool - big	Emma H	Accounting	Access	Oracle	N/A
4	Rec tool small	Emma H	Accounting	Excel	Oracle, Client Website	N/A
5	Format	Sofia K	Accounting	Excel	UDT6, Oracle, Pacific	N/A
6	Share databae	Nicholas Z	Static Data	Access	Arctic	N/A
7	Client reporter	Nicholas Z	Static Data	Excel	Arctic, Pacific	N/A
8	Div Tool	Janelle N	Inc Team	Excel	UDT16	Oracle
9	RE Client	James M	Ops	Excel	Client, ABS	NA
10	RE Client Rec	James M	Ops	Excel	Client, Website	UDT90

- 80-20 Rule: Only 20% is the fun stuff.
- Organisations likely to already have a data collated in a central list.
  - Not usable in current state.
- Need to transform into usable relationship matrix and adjacent list of nodes and characteristics.
- Manual – takes time. Need to think about your data.

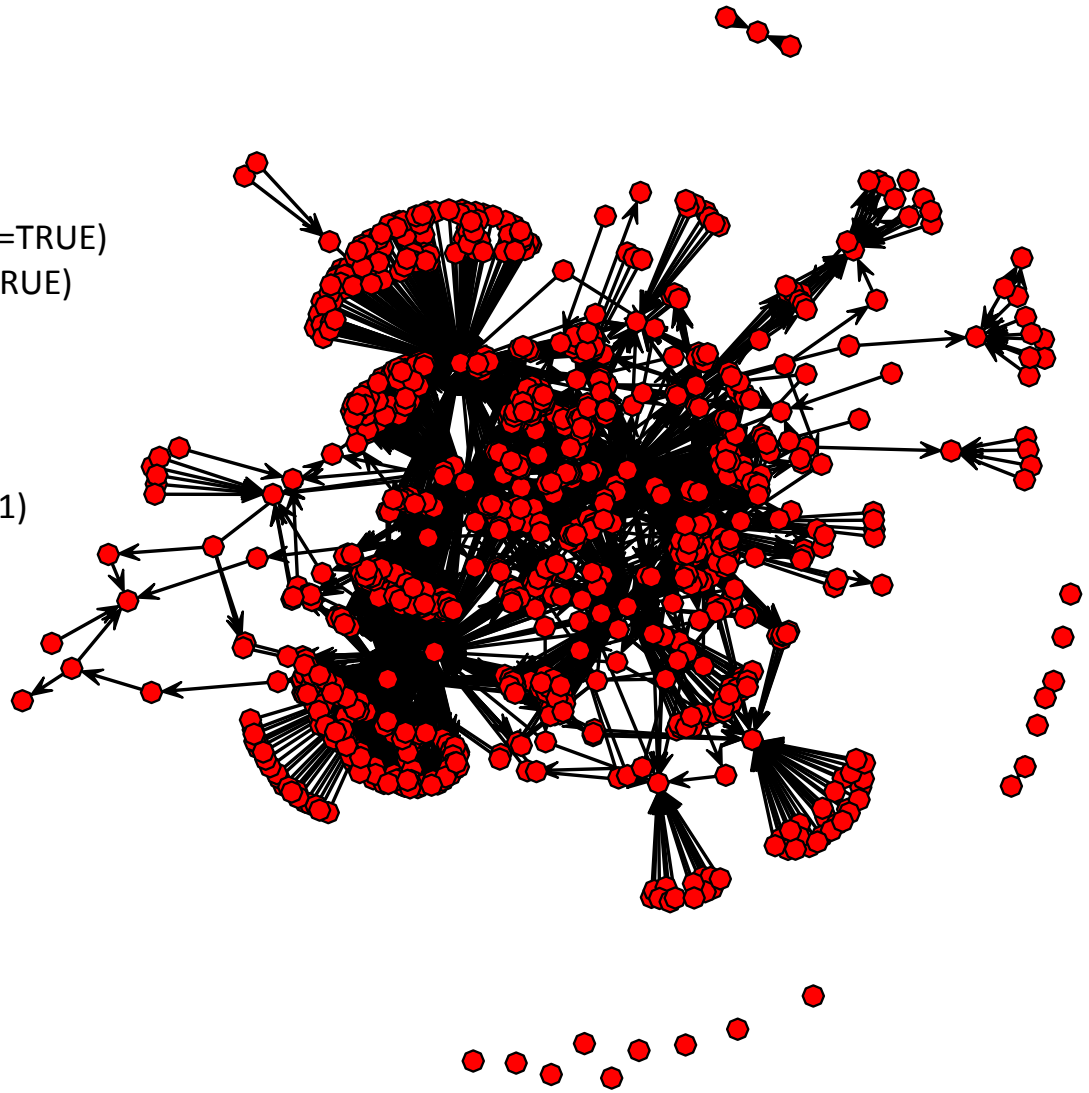
Scenario: UDT

UDTID	Name	Owner	Team	Type	SRC Systems	EXP Systems
1	Reconiliation manager	James C	Corp Actions	Excel	Coral, UDT98	UDT269
2	Import tool	John S	Corp Actions	Excel	Coral, Pacific	N/A
3	Rec tool - big	Emma H	Accounting	Access	Oracle	N/A
4	Rec tool small	Emma H	Accounting	Excel	Oracle, Client Website	N/A
5	Format	Sofia K	Accounting	Excel	UDT6, Oracle, Pacific	N/A
6	Share databae	Nicholas Z	Static Data	Access	Arctic	N/A
7	Client reporter	Nicholas Z	Static Data	Excel	Arctic, Pacific	N/A
8	Div Tool	Janelle N	Inc Team	Excel	UDT16	Oracle
9	RE Client	James M	Ops	Excel	Client, ABS	NA
10	RE Client Rec	James M	Ops	Excel	Client, Website	UDT90

- Potential Nodes
  - UDTID, Owner, Team, Type, Source & Export Systems.
  - Think how how to describe node attributes – colour, shape, size etc
- Relationships
  - Transform into a relationship matrix.
  - UDT > TEAM, UDT > SOURCE SYSTEM & EXPORT SYSTEM, UDT>UDT
  - Relationship = 1 (or more)
  - No relationship = 0
- This is the 80% of the hack work.

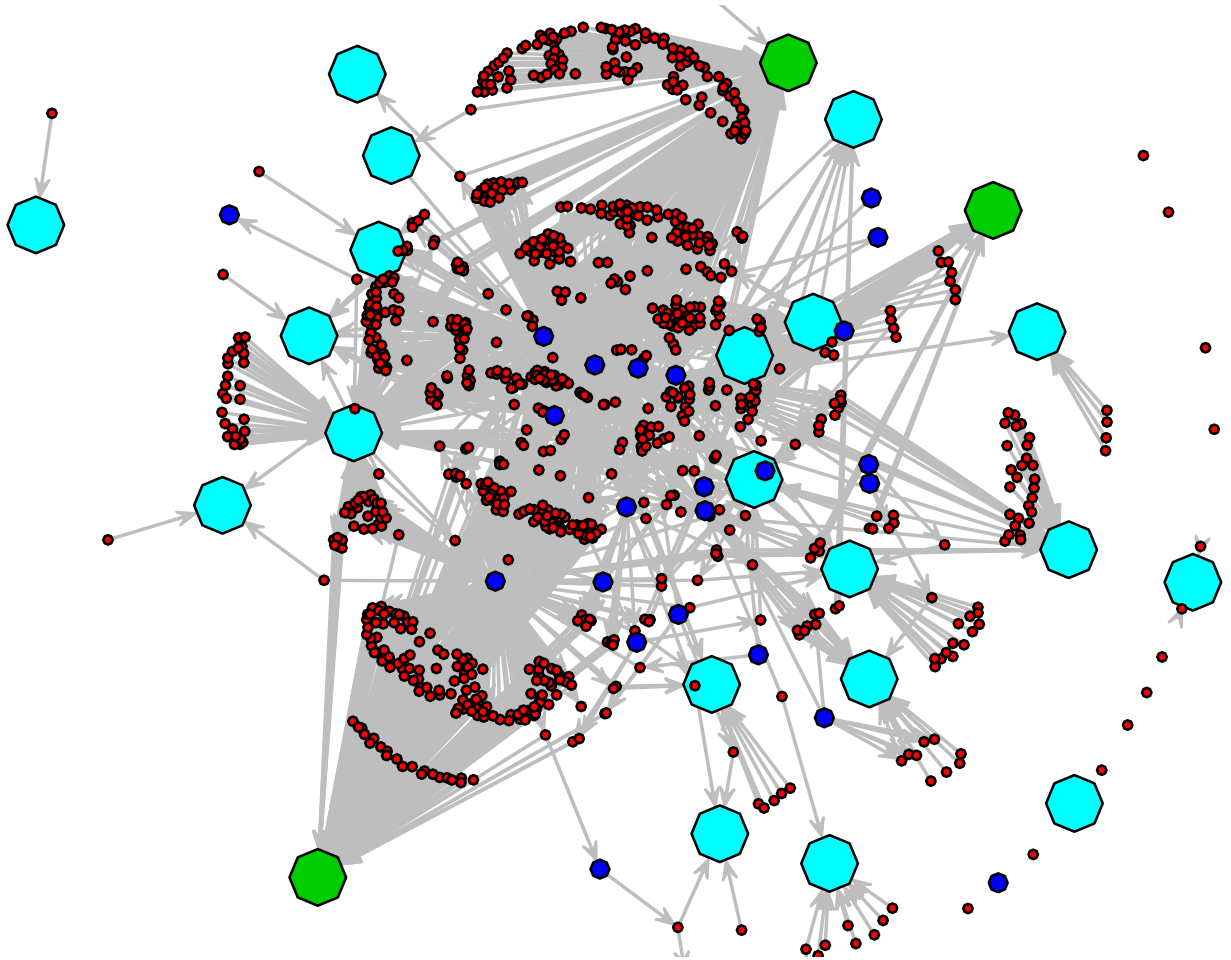
Scenario:UDT

```
par(ask=TRUE)

#load UDT script
MyMatrix1=read.csv("UdtMatrix.csv",header=TRUE)
MyActor1=read.csv("UdtActor.csv",header=TRUE)

attach(MyMatrix1)
attach(MyActor1)

row.names(MyMatrix1)=colnames(MyMatrix1)
Matrix1=as.matrix(MyMatrix1)

gplot(Matrix1)
```



Scenario: UDT

```
gplot(Matrix1,
+       vertex.col=Col1,
+       vertex.pch=Pch1,
+       vertex.cex = Cex1,
+       interactive = TRUE,
+       displaylabels=FALSE,
+       boxed.labels=FALSE,
+       label=Lab1,
+       edge.col=8,
+       object.scale=0.008,
+       label.cex=.
4,display.isolates=FALSE)
```



Colour & Symbol Codes

- Key Team
- Team
- UDT
- System

Scenario: UDT

- Degree(Matrix1)
  [1] 6 10 34 307 23 2 9 5 35 3 10 239 21 2 42 0 1 1
- Betweenness(Matrix1)
  [1] 2.061600e+03 4.407000e+03 0.000000e+00 1.853461e+05 0.000000e+00
  [6] 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
- rev(sort(betweenness(Matrix1)))
  [1] 1.883373e+05 1.853461e+05 1.480844e+05 1.265350e+05 9.857516e+04
  [6] 7.895839e+04 5.749865e+04 3.764665e+04 3.177727e+04 2.685657e+04
- rev(order(betweenness(Matrix1)))
  [1] 1000 4 992 478 998 26 305 411 379 919 976 290 277 918 909
  [16] 484 980 20 915 106 2 123 981 847 279 1 298 882 297 214
  Atlantic, Accounting, Pacific, U452, Indian, Zulu….

- As could be expected, key systems and teams get ranked highest.
- Note U452 – What makes this UDT special?
- Use this to help prioritise & target UDT's for more robust solutions.

Scenario: UDT

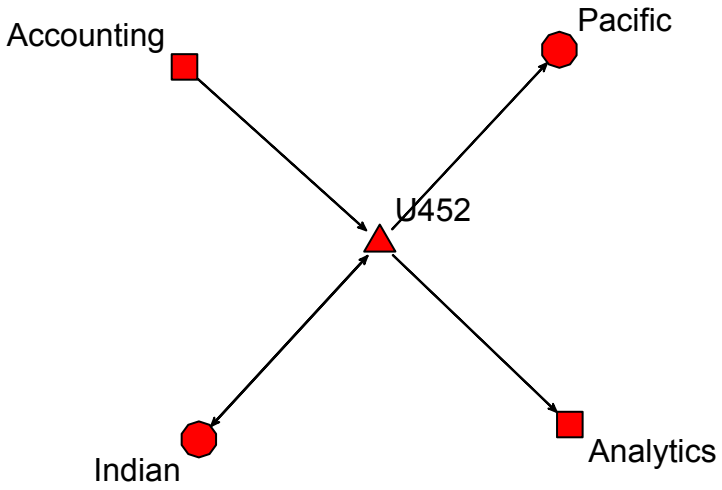
- Investigate why U452 is special.

- MyEgo=ego.extract(MyMatrix1)
- print(MyEgo$U452)
- gplot(MyEgo$U452,displaylabels=TRUE, vertex.sides=c(3,4,4,10,10), vertex.rot=c(90,45,45,90,90),vertex.cex=4)

> print(MyEgo$U452)

	U452	Accounting	Analytics	Pacific	Indian
U452	0	0	1	1	1
Accounting	1	0	0	0	0
Analytics	0	0	0	0	0
Pacific	0	0	0	0	0
Indian	1	0	0	0	0

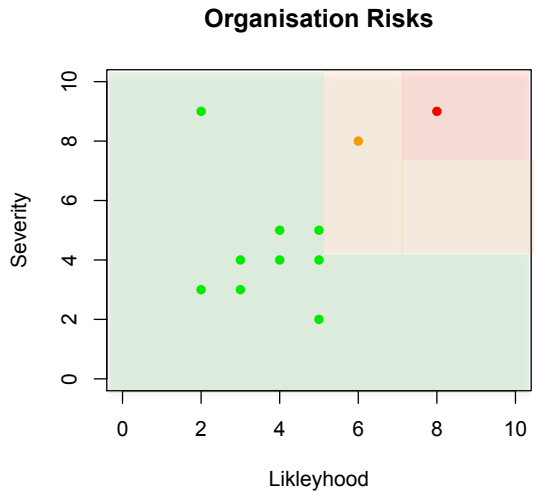
>



Scenario: Risk Mapping

Risk Log:

Risk No.	Name	Risk detail	Category	Severity	Likleyhood	RATING	Mitigation Plan	Owner
1	Key Person risk in IT	Mary is only one who knows system	People	8	6	48	Train backup	John S
2	Team Poaching	Analytics team may be poached	People	4	3	12	Review Package	HR
3	Oracle system	Rollover of ye data	System	3	2	6	Tech to test	James B
4	Oracle system - import module	UDT not parseing information for importing correctly	System	9	2	18	TBA	TBA
5	Custom Client Report	Custom results monthly report is dependent on excel reformating prior to emailing	Client	5	5	25	Try and automate report	Tech



- Many Risk teams will have comprehensive logs of an organisation's risks, risk categorisation, assessment of risk and action plans.
- Potential problem - each risk is independent and viewed in isolation.
- Use SNA to visualise and measure relationships.

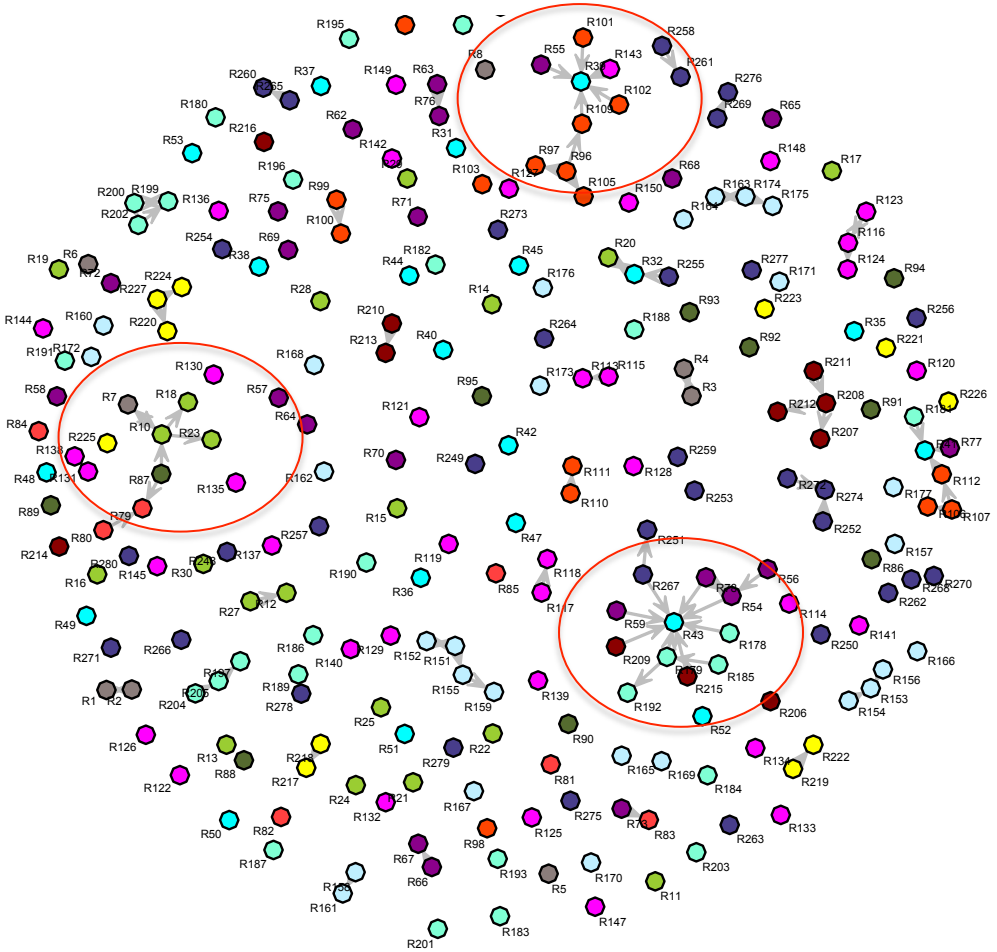
Scenario: Risk Mapping

- 80-20 Rule: Again only 20% is the fun stuff.
- Based on published Corporate Executive Board scenario.
  - Network Risk Maps: Unearthing Risk Interdependencies Nov 2009
- Need to utilise group based decision making to determine relationships.
- Path to reducing overall risk is targeting key nodes as priority. Those which have higher potential to 'trigger' another risk (indicated by relationship).
  - Increase fragmentation of network, i.e. less centralised
  - Want to avoid cascading failures. Ie "things happen in threes"
- Need to utilise group based decision making to determine relationships.
  - Does this node have a relationship with another node. What direction is relationship?
- Group should determine "rules" up front.
  - majority vs unanimous, scoring of relationship strength
- Recommend a smaller warm up exercise/trial first.
- Group must review each risk in relation to all other risks logged.
  - May want to simplify by doing a category at a time; then looking at relationships between categories
- Use this to building relationship matrix.

Scenario: Risk Mapping

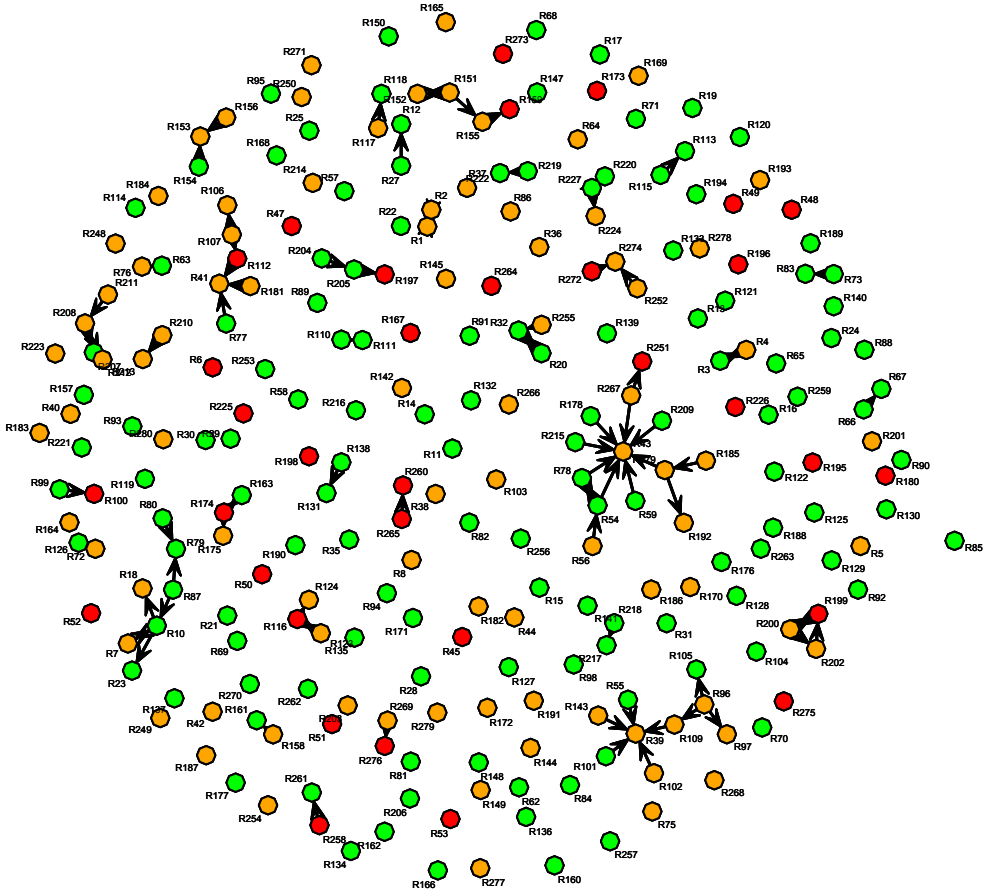
Colour & Symbol Codes

- Accounting
- Alpha
- Analytics
- Audit
- Bravo
- Charlie
- Echo
- HR
- Reporting
- Tax
- Technology
- Management
- Zulu



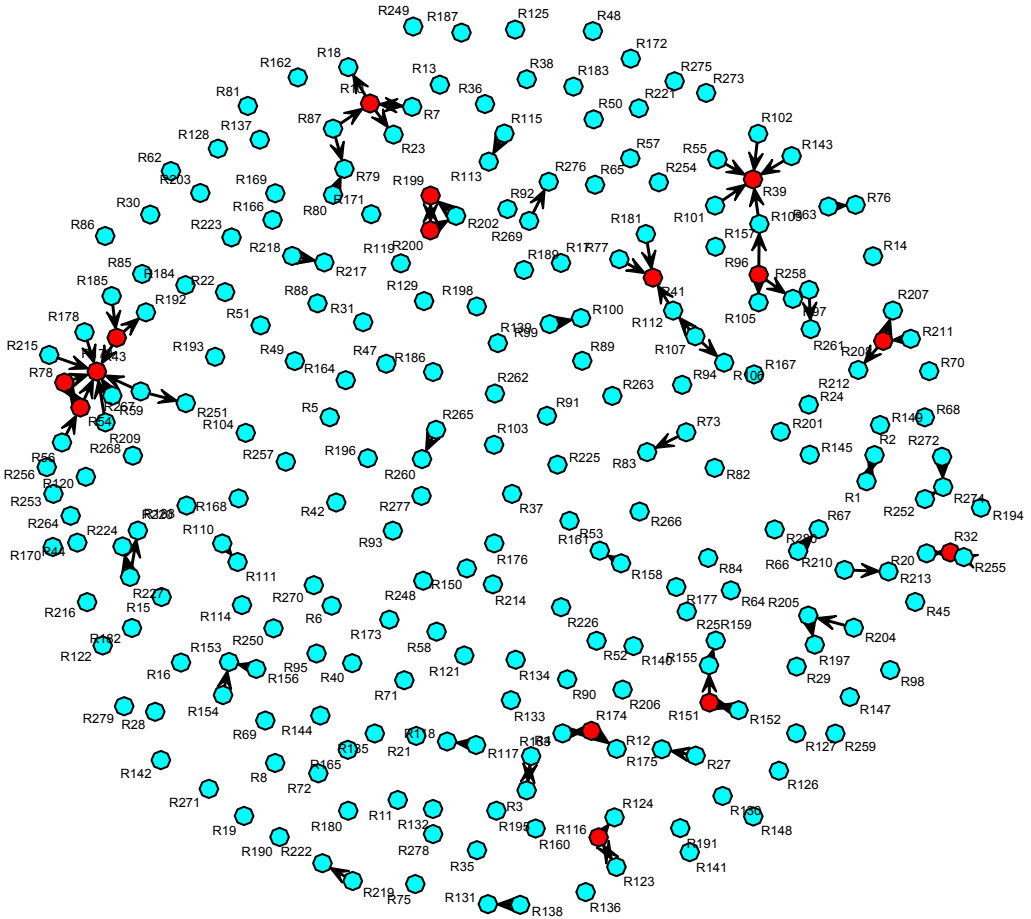
Path to reducing overall risk is targeting key nodes as priority. Those which have higher potential to 'trigger' another risk (indicated by relationship).

Scenario: Risk Mapping



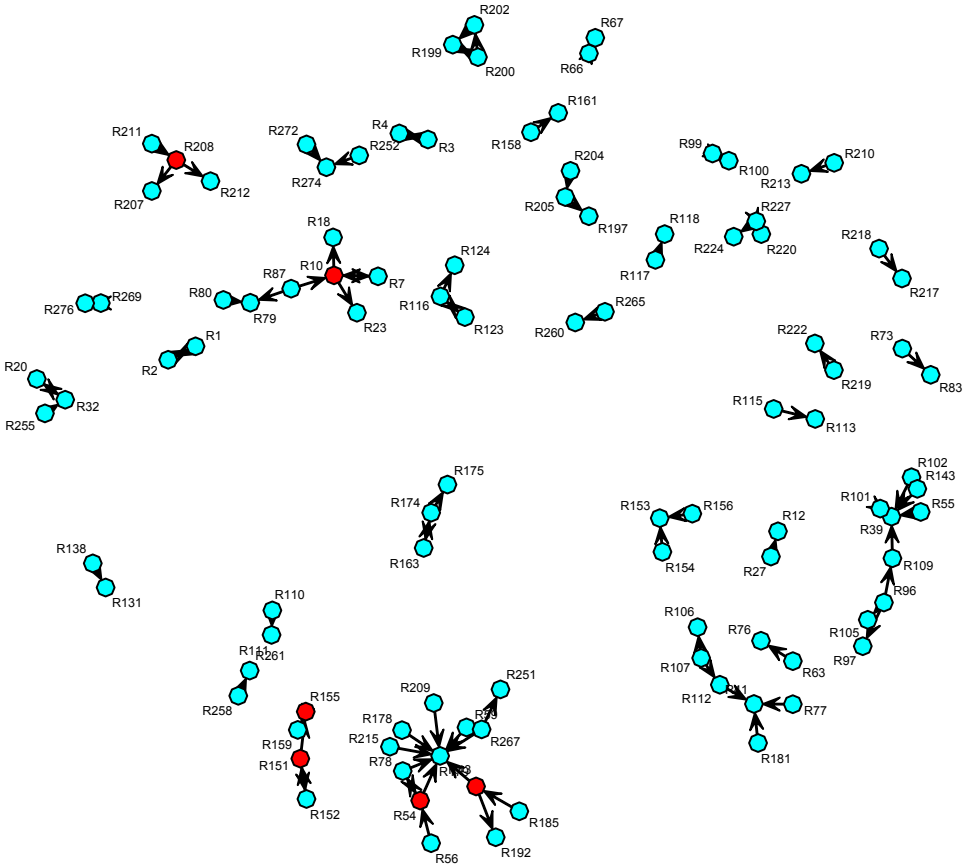
- Using a Traffic Light concept
- RPN=Severity * Likleyhood
- gplot(Matrix1,vertex.col=ifelse(RPN>7,"red",ifelse(RPN>3,"orange","green")))

Scenario: Risk Mapping



Mydegree=degree(Matrix1)
gplot(Matrix1,vertex.col=ifelse(Mydegree>2,2,5))

Scenario: Risk Mapping



- Mybetweenness=(betweenness(Matrix1))
- Mybetweenness
- gplot(Matrix1,vertex.col=ifelse(Mybetweenness>1,2,5), displaylabels=TRUE, label.cex=.4,displayisolates=FALSE)

Further Readings

Six Myths About Informal Networks — and How To Overcome Them
MIT SLOAN MANAGEMENT REVIEW Spring 2002
Rob Cross, Nitin Nohria & Andrew Parker

“Mapping Value of Employee Collaboration”
Mckinsey Quarterly 2006 No 3
Robert Cross, Roger Martin Leigh Weiss

“Harnessing the power of informal employee networks”
Mckinsey Quarterly 2007 No 4
Lowell Bryan, Eric Matson, Leigh Weiss

The role of networks in organisational change
Robert Cross, Salvtore Parise, Leigh Weiss
Mckinsey Quarterly April 2007

Informal Networks: The company behind the chart
Harvard Business Review: July Aug 1993
David Krackhardt & Jeffery Hanson

Network Risk Maps: Unearthing Risk Interdependencies
Corporate Executive Board: OPERATIONS COUNCIL November 2009

Managing Core Competencies of the Organization:
Organizational Network Mapping (Tactic #17)
The Advisory Board Company. 1996

Questions