Game theoretic centrality analysis of terrorist networks

Herbert Hamers

Presentation is based on this joint work with:
Peter Borm, Twan Burg, Tjeerd Campen, Bart Husslage, Roy Lindelauf

Lunteren, January 19, 2017

Literature

Presentation is based on the following three papers:

Roy Lindelauf, Herbert Hamers, Bart Husslage (2013). Cooperative game theoretic centrality analysis of terrorist networks: The cases of Jemaah Islamiyah and Al Qaeda. European Journal of Operational Research, 229(1), 230-238.

Bart Husslage, Peter Borm, Twan Burg, Herbert Hamers, Roy Lindelauf (2015). Ranking terrorists in networks: a sensitivity analysis of Al Qaeda's 9/11 attack. Social Networks, 42, 1-7.

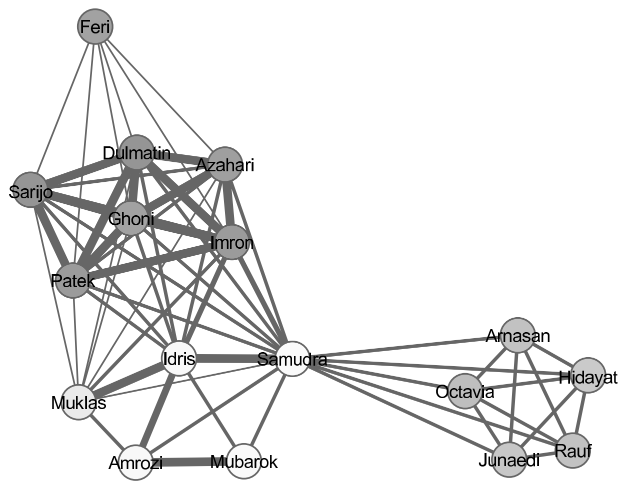
Herbert Hamers, Bart Husslage, Roy Lindelauf, Tjeerd Campen(2016). A New Approximation Method for the Shapley Value Applied to the WTC 9/11 Terrorist Attack. CentER Discussion Paper, 2016-042.

Outline:

- Networks
- Centrality measures
- Games and centrality measures
- Case: Jemaah Islamiyah, Bali attack
- Sensitivity analysis ranking
- Case: 9/11 attack Al Qaeda
- Approximation Shapley value
- Case: 9/11 attack Al Qaeda (revisited)

Networks

Interaction between terrorists can be described by a network.
Each terrorist is represented by one node in the network.
An edge between two nodes indicates that there is interaction between these two terrorists.
Interaction can be communication (e.g., phone, internet), exchanging goods (e.g., bomb devices)



The identification of key players in a terrorist network can lead to prevention of attacks, due to efficient allocation of surveillance means or isolation of key players in order to destabilize the network.

Centrality measures

Standard centrality measures from graph theory use only network structure (i.e. communication).

Game theoretical measures takes both network structure and non-network features, usually individual parameters (i.e. financial means, bomb building skills) into account.

The application of all these centrality measures results in rankings of the terrorists in the network.

Graph theoretical centrality measures

The normalized degree centrality of person i is expressed as the fraction of the network to which person i is directly related:

$$C_{\textsf{degree}}(i) = \frac{d(i)}{|N| - 1},$$

where $d(i)$ represents the number of direct relations of person i and $|N|$ is the total number of persons in the network.

Graph theoretical centrality measures

Let s_{kj} denote the total number of shortest paths between person k and j and let s_{kij} denote the number of shortest paths between k and j that pass through person i. The normalized betweenness centrality of person i is defined by

$$C_{\text{between}}(i) = \frac{2}{(|N|-1)(|N|-2)} \cdot \sum_{\substack{k,j \in N \setminus \{i\} \\ k<j}} \frac{s_{kij}}{s_{kj}},$$

The idea of betweenness centrality is that a person is important when he enables the flow of information between other persons in the network.

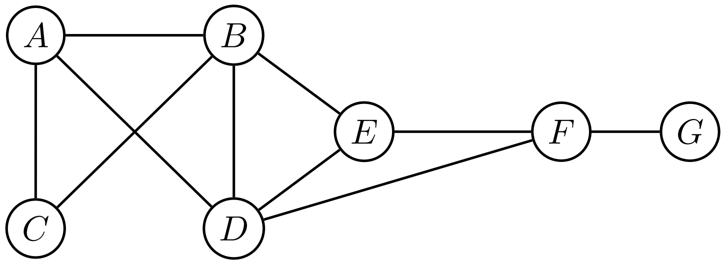
Graph theoretical centrality measures

The normalized closeness centrality of person i is defined by

$$C_{\mathsf{close}}(i) = \frac{|N| - 1}{\sum_{j \in N} l_{ij}},$$

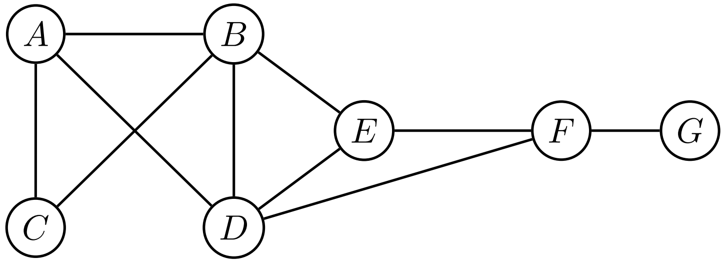
where l_{ij} denotes the shortest distance between person i and j.
The normalized closeness centrality quantifies the distance from a certain person to all other persons in the network.

Example:



Person	Degree	Betweenness	Closeness
A	0.5000	0.0778	0.6000
B	0.6667	0.2222	0.6667
C	0.3333	0	0.4615
D	0.6667	0.3222	0.7500
E	0.5000	0.1111	0.6667
F	0.5000	0.3333	0.6000
G	0.1667	0	0.4000

Graph theoretical centrality measures



Degree	Betweenness	Closeness
B^*	F	D
D^*	D	B^*
$A^\bullet$	B	E^*
$E^\bullet$	E	$A^\bullet$
$F^\bullet$	A	$F^\bullet$
C	C^*	C
G	G^*	G

Limitations of graphs centrality measures in (terroristic) network:

1. Takes only structure of network into account
2. Additional (individual) data is not included
3. Players in rankings are not distinguished enough

Shapley value as centrality measure

A cooperative game is a tuple (N, v) where

- $N = \{1, 2, ..., n\}$ is the set of players
- $v : 2^N \to \mathbb{R}$ is its characteristic function

By convention, $v(\emptyset) = 0$.
A set $S \in 2^N$ is called a coalition and N is called the grand coalition.

For example, the value of the grand coalition can express:
1. money (profit)
2. power (voting)
3. importance (terrorism)

Objective is finding an allocation (to all players) of value of the grand coalition.

The Shapley value of a game (N, v) is defined as

$$\varphi(v) = \frac{1}{n!} \sum_{\sigma \in \Pi(N)} m^{\sigma}(v),$$

where

- n is cardinallity of N,
- $\Pi(N)$the set of all permutations of N,
- $m_i^{\sigma}(v) = v(\{j \mid \sigma(j) \leq \sigma(i)\}) - v(\{j \mid \sigma(j) < \sigma(i)\})$ for all $i \in N$.

An undirected graph G is a pair $G = (N, E)$ where

- N: Vertex set of G
- E: Edge set of G

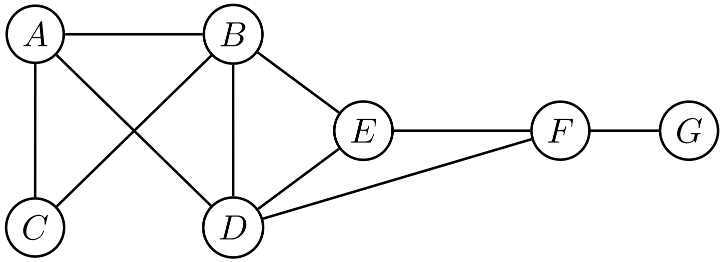
For $S \subseteq N$,

- $G[S]$: the subgraph of G induced by $S \subseteq V$

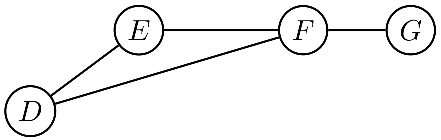
The connectivity game on a graph $G = (N, E)$ is defined as

$$v^{\mathsf{conn}}(S) = \begin{cases} 1 & \text{if } G[S] \text{ is connected and } |S| > 1, \\ 0 & \text{otherwise.} \end{cases}$$

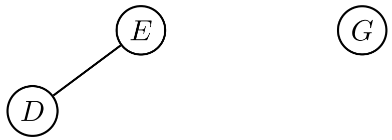
Consider the connectivity game corresponding to:



Then, for example, coalition $\{D, E, F, G\}$ is connected and coalition $\{D, E, G\}$ is not.



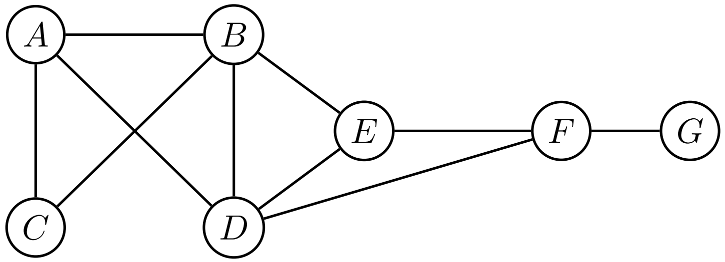
Subgraph for coalition $\{D, E, F, G\}$.



Subgraph for coalition $\{D, E, G\}$.

Hence, $v^{\text{conn}}(\{D, E, F, G\}) = 1$ and $v^{\text{conn}}(\{D, E, G\}) = 0$.

In example of a weighted connectivity game (additional individual information is included)



Additional information:

- Person E participated in previous attack
- Person C and E have sufficient financial means

Based on this information the following weights are assigned:
Person C: 4, Person E: 11, All others: 1.

$$v^{\mathsf{wconn}}(S) = \begin{cases} \sum_{i \in S} w_i & \text{if } G[S] \text{ is connected,} \\ 0 & \text{otherwise,} \end{cases}$$

Rankings based on graph theoretical centralities and Shapley value of weighted connectivity game

Degree	Betweenness	Closeness	Shapley
B^*	F	D	E
D^*	D	B^*	F
$A^\bullet$	B	E^*	B
$E^\bullet$	E	$A^\bullet$	D
$F^\bullet$	A	$F^\bullet$	C
C	C^*	C	A
G	G^*	G	G

Observe:

- B and F in top 3 of all rankings
- Shapley value better able to distinguish individuals than standard centrality
- the use of additional information ranks E and C higher.

Application of game theoretical centrality

The application of game theoretic centrality to a terrorist network consists of three steps:

1. **Construct the network (input)**
2. **Define a game theoretic model (modeling)**
3. **Analyze the rankings of players (output)**

Application of game theoretical centrality

1. **Construct the network (input)**

- data collection with respect to target group
- identify the relationships
- assign weights to individuals and their relationships

Result: a weighted graph

Application of game theoretical centrality

2. **Define a game theoretic model (modeling)**

- define a cooperative game based on the information in step 1.
  (game depends on information at hand!)

Result: (a set of) cooperative games

Application of game theoretical centrality

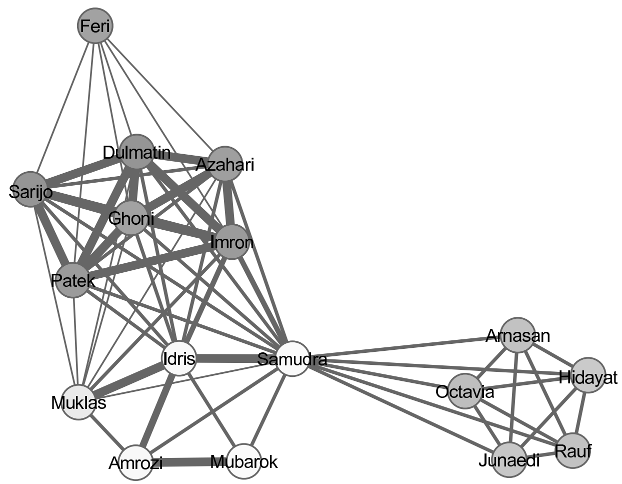
3. **Analyze the rankings of players (output)**

- Use a game theoretic centrality measure (Shapley value)
- analyse the ranking(s)

Result: identification of key players in the network

Case: Jemaah Islamiyah, Bali attack

The network of attack Bali, 2002, by Jemaah Islamiya:

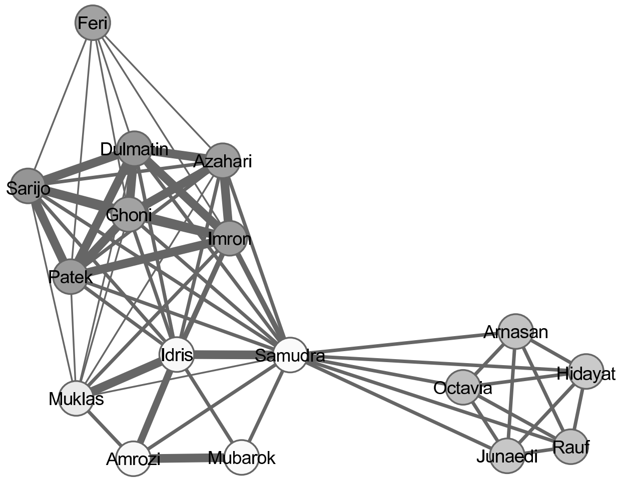


Weighted connectivity game is based on the following:

- Data from publication of Koschade (2005)
- Frequency and duration of interaction in a coalition
- The number of connections in a coalition

Case: Jemaah Islamiyah, Bali attack

The network of attack Bali, 2002, by Jemaah Islamiya:



Formally, we have

$$v^{\mathsf{wconn1}}(S) = \begin{cases} \max\limits_{\substack{i,j \in S \\ i \neq j}} f_{ij} & \text{if } S_G \text{ is connected,} \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

Case: Jemaah Islamiyah, Bali attack

Rankings for the Jemaah Islamiyah network

Degree	Betweenness	Closeness	Shapley
Samudra	Samudra	Samudra	Samudra
Idris	Idris	Idris	Muklas
Muklas*	Muklas	Muklas*	Feri
Ali Imron*	Ali Imron*	Ali Imron*	Azahari
Dulmatin*	Dulmatin*	Dulmatin*	Sarijo
Azahari*	Azahari*	Azahari*	Patek
Patek*	Patek*	Patek*	Dulmatin
Ghoni*	Ghoni*	Ghoni*	Idris
Sarijo*	Sarijo*	Sarijo*	Ghoni
Feri	Amrozi	Arnasan•	Octavia*
Arnasan•	Feri•	Junaedi•	Abdul Rauf*
Junaedi•	Arnasan•	Abdul Rauf•	Hidayat*
Abdul Rauf•	Junaedi•	Octavia•	Arnasan*
Octavia•	Abdul Rauf•	Hidayat•	Junaedi*
Hidayat•	Octavia•	Amrozi	Amrozi
Amrozi	Hidayat•	Mubarok	Mubarok
Mubarok	Mubarok•	Feri	Ali Imron

Observe:

- Samudra was the key player in this operation
- the rankings in standard centrality of the 5 most important persons are ambiguous
- Shapley value creates a real top 5
- Shapley introduces 3 new top 5 persons: Feri, Azahari and Sarijo.
- Feri was first suicide bomber
- Azahari bomb expert and "brain" behind attack

Sensitivity analysis rankings

How robust are rankings with respect to:

- network structure (adding or removal egdes)
- individual strength (weight individual)
- relational strength (weight edge)

We focus on Al Qaeda 9/11 attack.

Sensitivity analysis rankings Al Qaeda

The individuals and their relations of the 19 crew members of the four planes

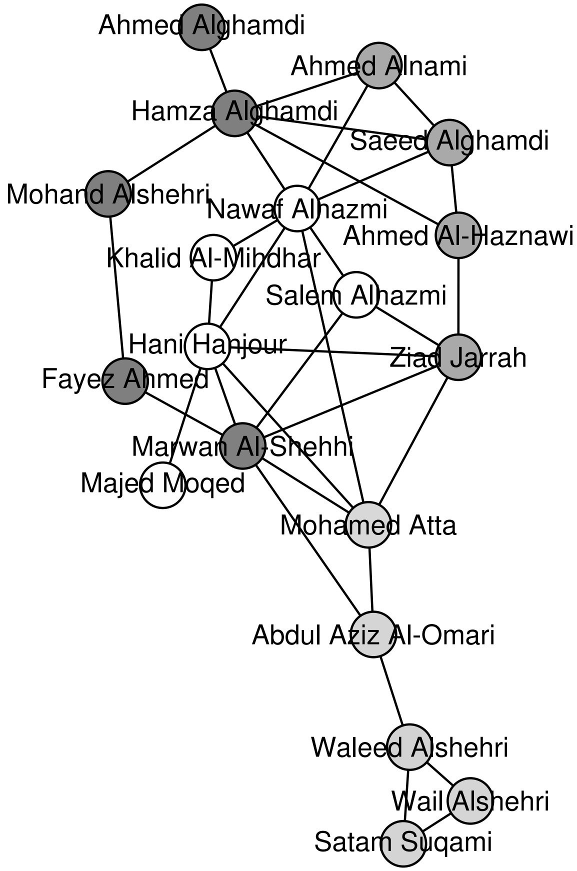


Figure 1: Operational network of hijackers of Al Qaeda's 9/11 attack. AA-77 (white), AA-11 (lightgray), UA-93 (gray) and UA-175 (darkgray).

Sensitivity analysis rankings Al Qaeda

We have only some additional information with respect to the individual strength.

Hijacker	Weight	Hijacker	Weight
Ahmed Alghamdi	1	Nawaf Alhazmi	2
Hamza Alghamdi	1	Khalid Al-Mihdhar	3
Mohand Alshehri	1	Hani Hanjour	1
Fayez Ahmed	1	Majed Moqed	1
Marwan Al-Shehhi	3	Mohamed Atta	4
Ahmed Alnami	1	Abdul Aziz Al-Omari	1
Saeed Alghamdi	1	Waleed Alshehri	1
Ahmed Al-Haznawi	1	Satam Suqami	1
Ziad Jarrah	4	Wail Alshehri	1
Salem Alhazmi	1		

Table 1: Weight assigned to each hijacker of Al Qaeda's $9/11$ attack.

Sensitivity analysis rankings Al Qaeda

We use the following game:

For a connected coalition we define

$$v(S) = \left(\sum_{i \in S} w_i\right) \cdot \max_{ij \in E_S} k_{ij}.$$

and for a not connected coalition we define

$$v^{\mathsf{mwconn}}(S) = \max_{T \subset S,\ T \text{connected}} v^{\mathsf{mwconn}}(T).$$

Sensitivity analysis rankings Al Qaeda

The ranking using game theoretic centrality measure (Shapley value)

Ranking R^{m}
Mohamed Atta
Ziad Jarrah
Marwan Al-Shehhi
Nawaf Alhazmi
Hani Hanjour
Khalid Al-Midhar
Abdul Aziz Al-Omari
Hamza Alghamdi
Waleed Alshehri
Ahmed Al-Haznawi
Salem Alhazmi
Fayez Ahmed
Saeed Alghamdi
Mohand Alshehri
Ahmed Alnami
Majed Moqed
Ahmed Alghamdi
Satam Suqami
Wail Alshehri

Change in network: four edges removed

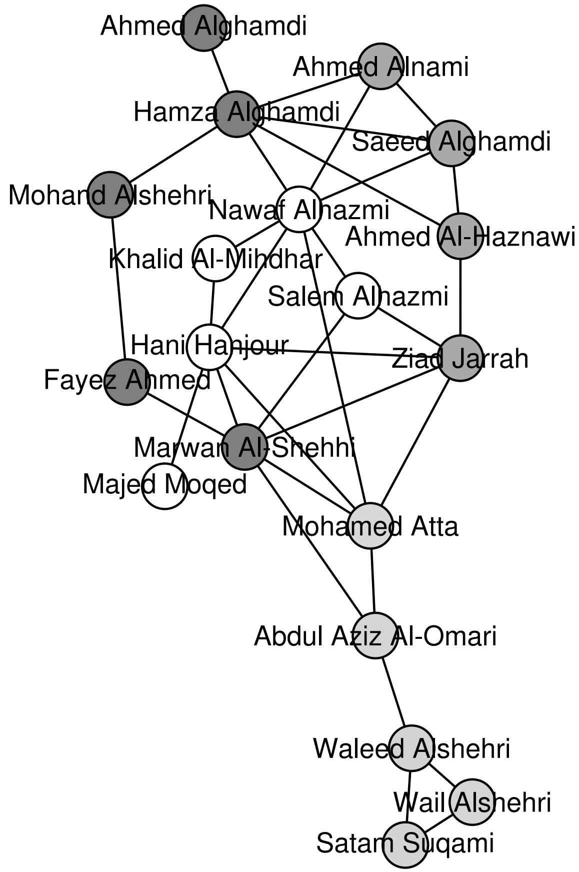


Figure 2: Operational network of hijackers of Al Qaeda's $9/11$ attack. AA-77 (white), AA-11 (lightgray), UA-93 (gray) and UA-175 (darkgray).

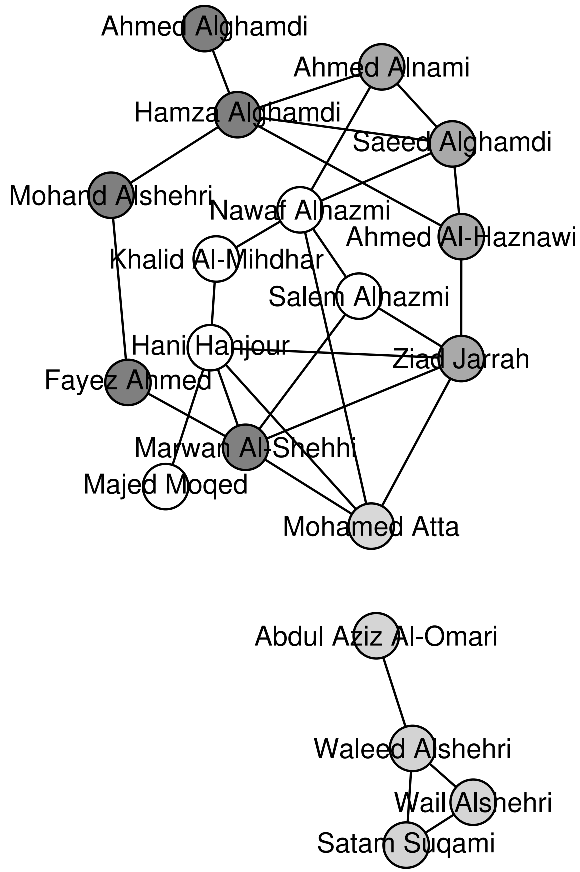


Figure 3: Operational network of hijackers of Al Qaeda's $9/11$ attack with four (random) links removed. AA-77 (white), AA-11 (lightgray), UA-93 (gray) and UA-175 (darkgray).

Effect on ranking

Ranking R^{m}
Mohamed Atta
Ziad Jarrah
Marwan Al-Shehhi
Nawaf Alhazmi
Hani Hanjour
Khalid Al-Midhar
Abdul Aziz Al-Omari
Hamza Alghamdi
Waleed Alshehri
Ahmed Al-Haznawi
Salem Alhazmi
Fayez Ahmed
Saeed Alghamdi
Mohand Alshehri
Ahmed Alnami
Majed Moqed
Ahmed Alghamdi
Satam Suqami
Wail Alshehri

Table 2: Ranking for the original network (Figure 2).

Ranking R_1
Ziad Jarrah
Mohamed Atta
Marwan Al-Shehhi
Nawaf Alhazmi
Khalid Al-Midhar
Hani Hanjour
Hamza Alghamdi
Ahmed Al-Haznawi
Salem Alhazmi
Fayez Ahmed
Saeed Alghamdi
Mohand Alshehri
Ahmed Alnami
Majed Moqed
Ahmed Alghamdi
Waleed Alshehri
Satam Suqami
Wail Alshehri
Abdul Aziz Al-Omari

Table 3: Ranking for the changed network (Figure 3).

Comparing rankings

Value assigned to each position in ranking R^m

Position	1	2	3	4	5	6	7	8	9	10
Value	1	4/5	3/5	2/5	1/5	1/14	2/14	3/14	4/14	5/14
Position	11	12	13	14	15	16	17	18	19	
Value	6/14	7/14	8/14	9/14	10/14	11/14	12/14	13/14	1	

Table 4: Value assigned to each position in ranking R^{m}.

The difference between ranking R^m and new ranking R_1 is expressed by ρ.

ρ is defined as:
the sum of
the values of all hijackers that leave the top-5 in R^{m}
and
enter the top-5 in R_1 is taken.

Effect on ranking

Ranking R^{m}
Mohamed Atta
Ziad Jarrah
Marwan Al-Shehhi
Nawaf Alhazmi
Hani Hanjour^{out}
Khalid Al-Midhar^{in}
Abdul Aziz Al-Omari
Hamza Alghamdi
Waleed Alshehri
Ahmed Al-Haznawi
Salem Alhazmi
Fayez Ahmed
Saeed Alghamdi
Mohand Alshehri
Ahmed Alnami
Majed Moqed
Ahmed Alghamdi
Satam Suqami
Wail Alshehri

Ranking R_1
Ziad Jarrah
Mohamed Atta
Marwan Al-Shehhi
Nawaf Alhazmi
Khalid Al-Midhar^{in}
Hani Hanjour^{out}
Hamza Alghamdi
Ahmed Al-Haznawi
Salem Alhazmi
Fayez Ahmed
Saeed Alghamdi
Mohand Alshehri
Ahmed Alnami
Majed Moqed
Ahmed Alghamdi
Waleed Alshehri
Satam Suqami
Wail Alshehri
Abdul Aziz Al-Omari

Difference between these rankings:
$\rho(R^{\mathsf{m}}, R_1) = \frac{1}{5} + \frac{1}{14} = \frac{19}{70} \approx 0.2714$. Note maximum value $\rho \approx 7.29$.

Three types of simulations

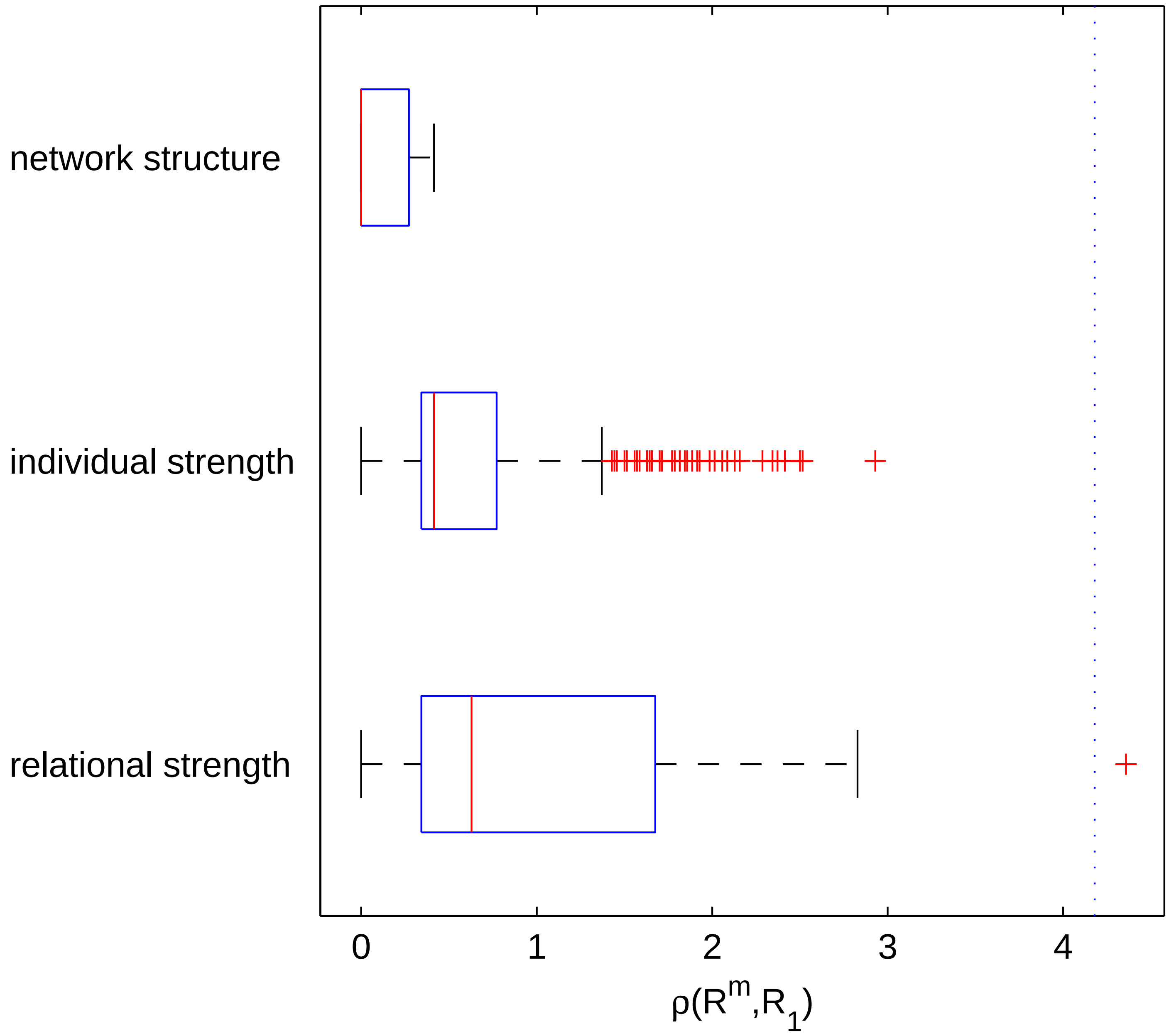
Network structure: adding or removing up to four edges (1000 simulations).

Individual strength: the weight for each is randomly equal to 1,2,3,4 (1000 simulations).

Relational strength: the weight of a single link is randomly increased to 4 (33 computation).

Furthermore, 1000 simulation of random rankings to generate expected ρ (so, a ranking obtained using no additional information about network structure or weights). For these simulations $\rho = 4.18$.

Results of three types of simulations



Approximation Shapley value

Calculation is important in practice, e.g.,

- covert networks
- social networks
- voting problems
- .....

Time efficient calculation Shapley value in general not possible.

Presence of structure in game or an underlying network may lead to time efficiency calculations of Shapley value.
But even if structure is present, a time efficient calculation may not be possible.

We need approximations for Shapley value!

Approximation Shapley value

Recall

$$\varphi(v) = \frac{1}{n!} \sum_{\sigma \in \Pi(N)} m^{\sigma}(v),$$

Procedure random sampling (Castro, Gómez,Tejeda (2009)):

Input: n-person cooperative game (N, v).

1. Select a subset Π_r of r orderings from all $n!$ possible orderings, i.e., $\Pi_r \subset \Pi$.
2. Compute the marginal contributions $m_v^{\sigma}(i)$ for all players $i \in N$ and for all orderings $\sigma \in \Pi_r$.
3. Approximate the Shapley value for each player i by averaging the marginal contributions obtained at step 2, i.e., $\hat{\varphi}_i(v) = \frac{1}{r} \sum_{\sigma \in \Pi_r} m_v^{\sigma}(i)$.

Approximation Shapley value

Example

S	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$v(S)$	0	1	3	0	5	7	4	10

$\varphi(v) = (3\frac{5}{6}, 3\frac{1}{3}, 2\frac{5}{6})$

Using random sampling procedure:

σ	$m_v^\sigma(1)$	$m_v^\sigma(2)$	$m_v^\sigma(3)$
$(1,2,3)$	1	4	5
$(1,3,2)$	1	3	6
$(3,1,2)$	7	3	0

$\hat{\varphi}(v) = (3, 3\frac{1}{3}, 3\frac{2}{3})$.

Approximation Shapley value

Procedure structured random sampling:

Input: n-person cooperative game (N, v).

1. Select a subset Π_r of r orderings from all $n!$ possible orderings, i.e., $\Pi_r \subset \Pi$, with $r = t \cdot n$ and $t \in \mathbb{N}$.
2. Divide the subset Π_r in n groups of size t.
3. For each player i:
   (a) Swap player i with the player at position j for each of the t orderings in group j, where $j \in \{1, \ldots, n\}$, resulting in a set Π'_r of r new orderings.
   (b) Compute the marginal contributions $m_v^\sigma(i)$ of player i for all new orderings $\sigma \in \Pi'_r$.
   (c) Approximate the Shapley value of player i by averaging the marginal contributions obtained at step $3b$, i.e., $\hat{\varphi}_i(v) = \frac{1}{r} \sum_{\sigma \in \Pi'_r} m_v^\sigma(i)$.

Approximation Shapley value

Example

Group	Ordering	Swap 1	$m_v^\sigma(1)$	Swap 2	$m_v^\sigma(2)$	Swap 3	$m_v^\sigma(3)$
1	$(1,2,3)$	$(1,2,3)$	1	$(2,1,3)$	3	$(3,2,1)$	0
2	$(1,3,2)$	$(3,1,2)$	7	$(1,2,3)$	4	$(1,3,2)$	6
3	$(3,1,2)$	$(3,2,1)$	6	$(3,1,2)$	3	$(2,1,3)$	5

$\hat{\varphi}(v) = (4\frac{2}{3}, 3\frac{1}{3}, 3\frac{2}{3})$

Observations:

1. Both procedures use the same number of marginals. But structured procedure also includes a swap.
2. Random procedure is efficient, structured procedure is not.

Nevertheless, structured procedure outerperforms random sampling.

Approximation Shapley value

Two error measures to compare performance of the two procedures.

Average Average Absolute Error (AAAE)

$$\text{AAAE} = \frac{1}{50}\sum_{j=1}^{50}\left(\frac{1}{n}\sum_{i=1}^{n}|\hat{\varphi}_i(v_j) - \varphi_i(v_j)|\right)$$

Average Average Percentage Error (AAPE)

$$\text{AAPE} = \frac{1}{50}\sum_{j=1}^{50}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{\varphi}_i(v_j) - \varphi_i(v_j)|}{|\varphi_i(v_j)|}\right)$$

Approximation Shapley value

Procedure error measures

1. Randomly generate 50 SOUG games and normalize the value of the grand coalition in each game.
2. Compute the exact Shapley values for all players in all 50 games.
3. Use random sampling to approximate the Shapley values for all players in all 50 games and compute the error measures AAAE en AAPE.
4. Use structured random sampling to approximate the Shapley values for all players in all 50 games and compute the error measures AAAE en AAPE.

Approximation Shapley value

Result with respect to number of orderings

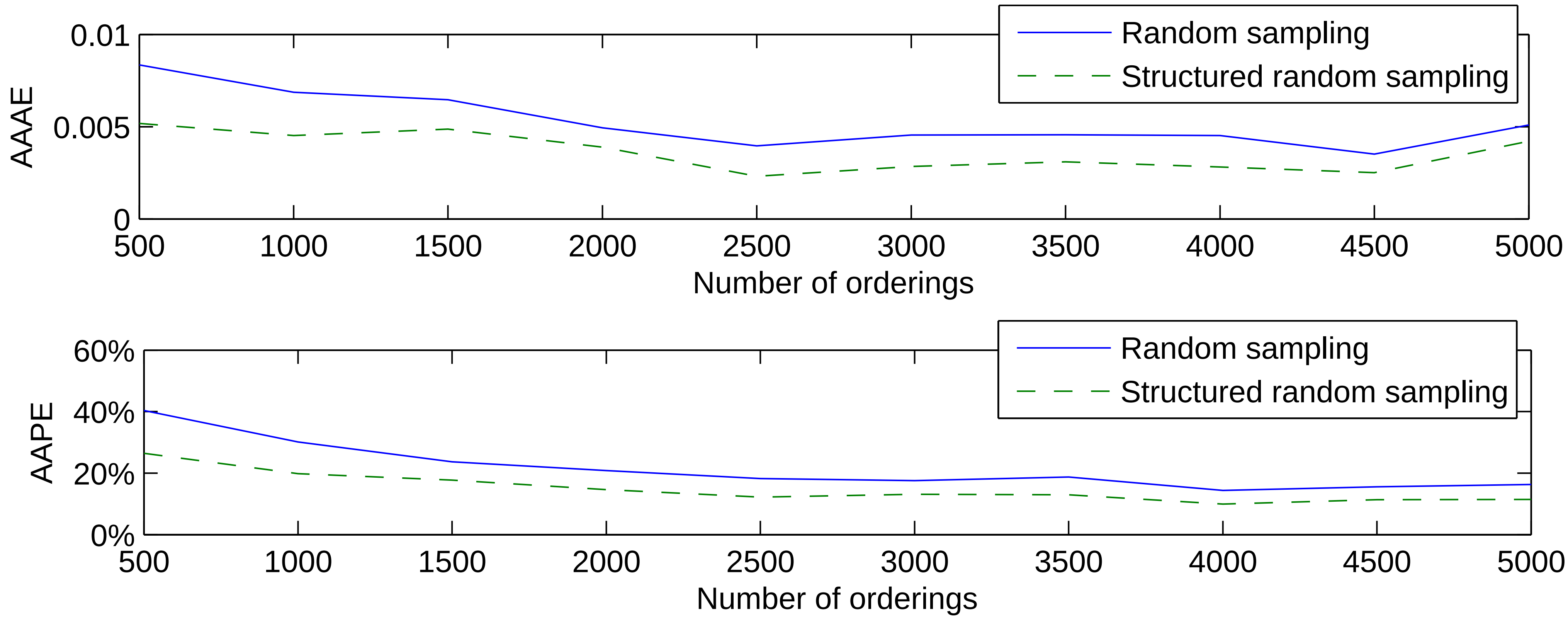


Figure 5: Performance analysis on the number of orderings.

Approximation Shapley value

Result with respect to the number of players

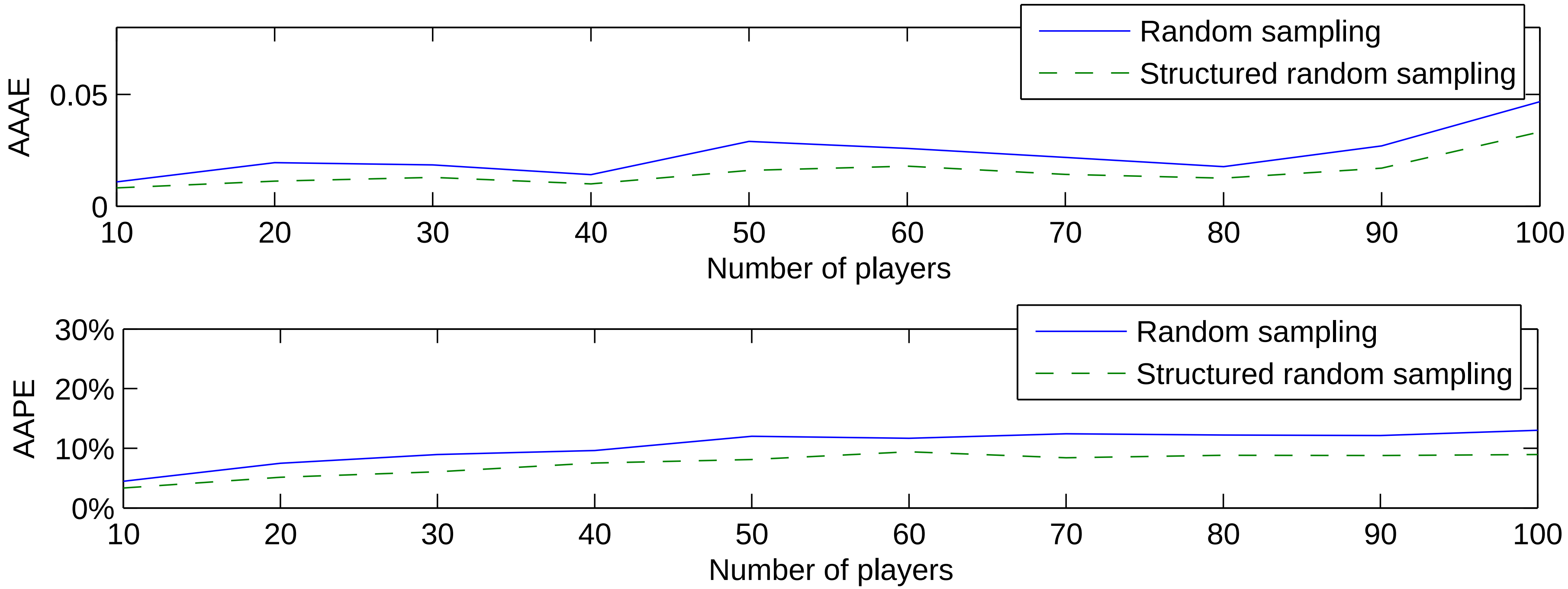
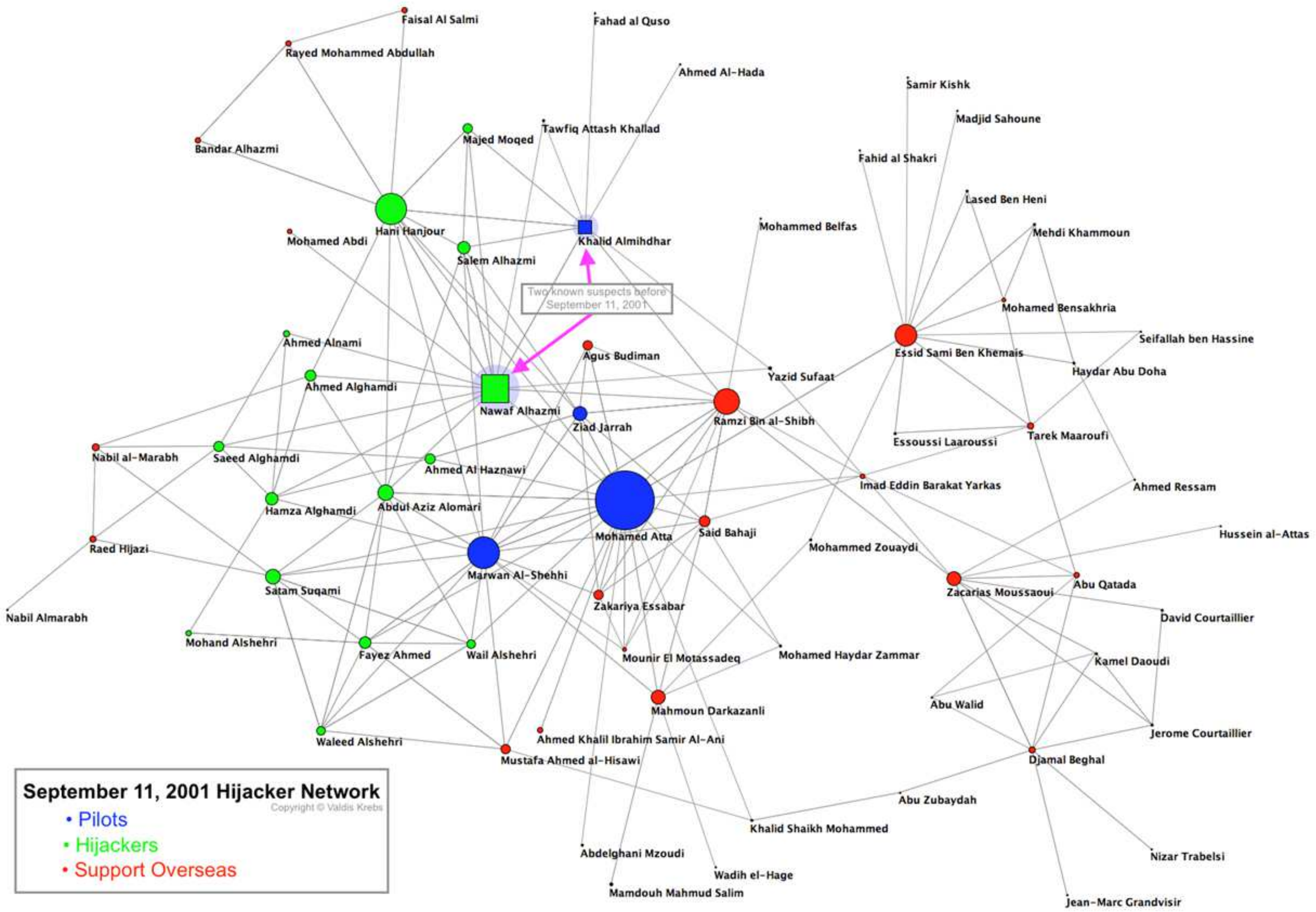


Figure 6: Performance analysis on the number of players.

Case 9/11 attack Al Qaeda (revisited)



Case 9/11 attack Al Qaeda (revisited)

Ranking	Name	Appr. Shapley value
1	Mohamed Atta	0.1137
2	Essid Sami Ben Khemais	0.1111
3	Hani Hanjour	0.1107
4	Djamal Beghal	0.1070
5	Khalid Almihdhar	0.1069
6	Mahmoun Darkazanli	0.1067
7	Zacarias Moussaoui	0.1009
8	Nawaf Alhazmi	0.0995
9	Ramzi Bin al-Shibh	0.0985
10	Raed Hijazi	0.0949
11	Hamza Alghamdi	0.0090
12	Fayez Ahmed	0.0088
13	Marwan Al-Shehhi	0.0046
14	Satam Suqami	0.0038
15	Saeed Alghamdi	0.0037

Table 5: First 15 members in WTC network according to the approximated Shapley value.

Concluding remarks

Game theoretical centrality measure takes into account structure network, individual and relationship features

Rankings are not too sensitive in case of missing edges or weight information about individuals

Approximation methods to Shapley value are important to analyze large networks.

Further research

1. Create better approximation methods Shapley value
2. Include dynamic aspects to incorporate change network
3. Use of real life data to fine tune framework