



Σύνθετα Δίκτυα

com+plex: with+ -fold (having parts)

Διδάσκων –
Δημήτριος Κατσαρός



Μετρικές κεντρικότητας

Centrality measures



Περιεχόμενα

Παρουσιάσαμε

- Degree centrality (DC) και centralization
- Shortest-Path Betweenness Centrality (SPBC)

σε μη-κατευθυνόμενα δίκτυα

Θα παρουσιάσουμε

- Closeness centrality
- Bridging centrality
- επεκτάσεις όλων σε κατευθυνόμενα δίκτυα



Κεντρικότητα εγγύτητας (closeness centrality)

- Εάν δεν μας ενδιαφέρει να έχουμε πολλούς “άμεσους” φίλους;
- Ή να είμαστε “μεταξύ” άλλων;
- Αλλά, εξακολουθεί να μας ενδιαφέρει να είμαστε στο “μέσον” των πραγμάτων, όχι μακριά από το κέντρο;



Κεντρικότητα εγγύτητας

Η κεντρικότητα εγγύτητας βασίζεται στο μέσο μήκος του ελάχιστων μονοπατιών μεταξύ του κόμβου και όλων των υπόλοιπων κόμβων του δικτύου

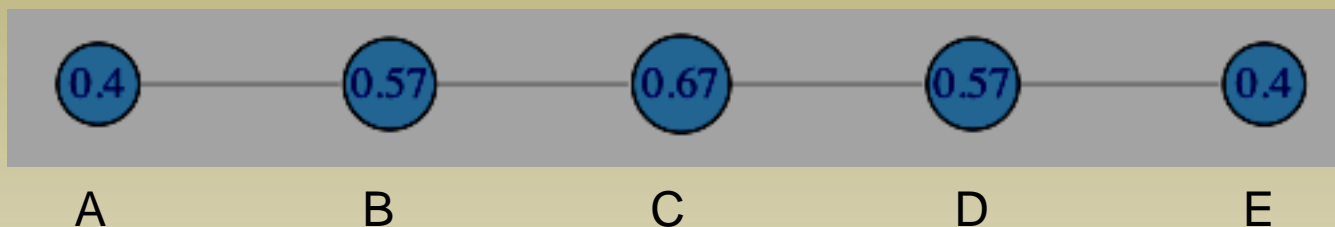
Closeness Centrality (CC):

$$C_c(i) = \left[\sum_{j=1}^N d(i, j) \right]^{-1}$$

Κανονικοποιημένη CC

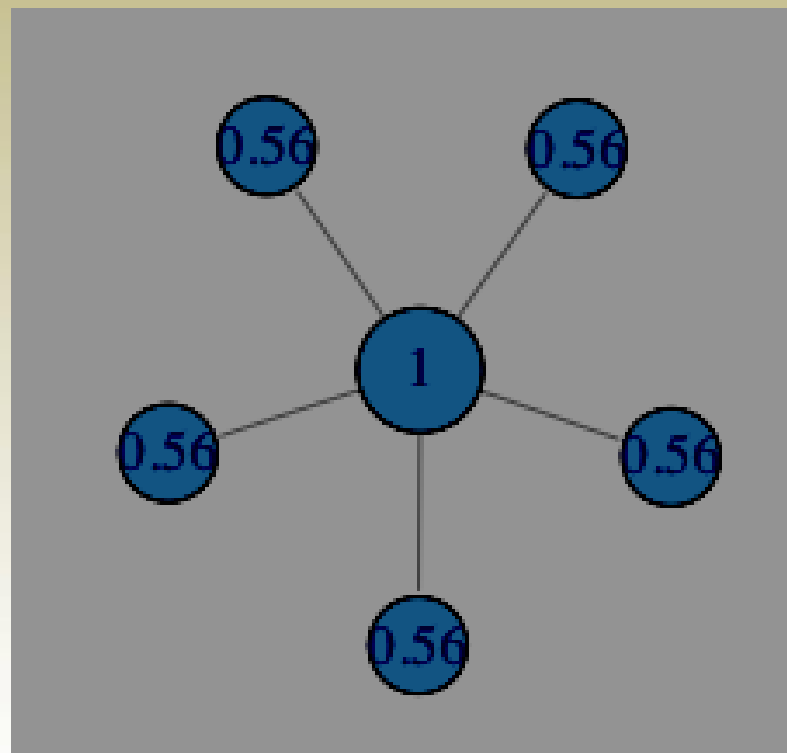
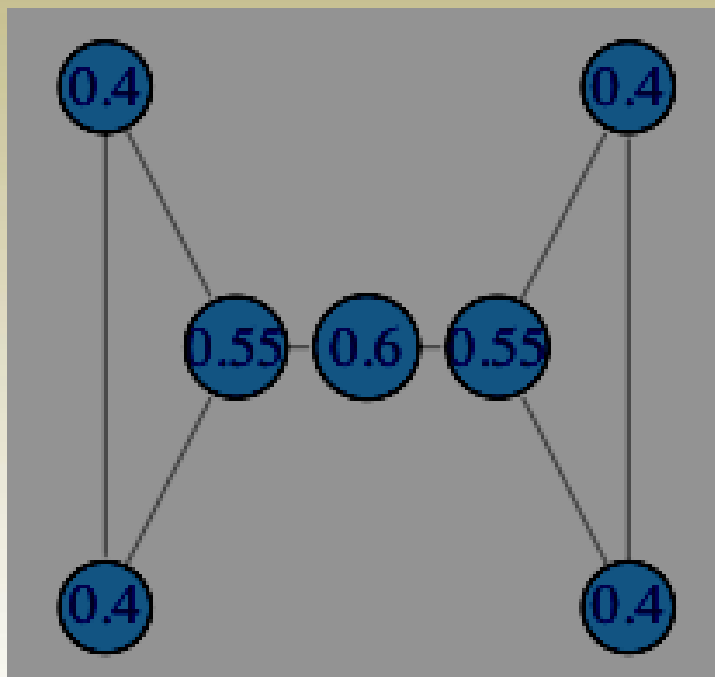
$$C'_c(i) = (C_c(i)) / (N - 1)$$

Κεντρικότητα εγγύτητας: Παράδειγμα



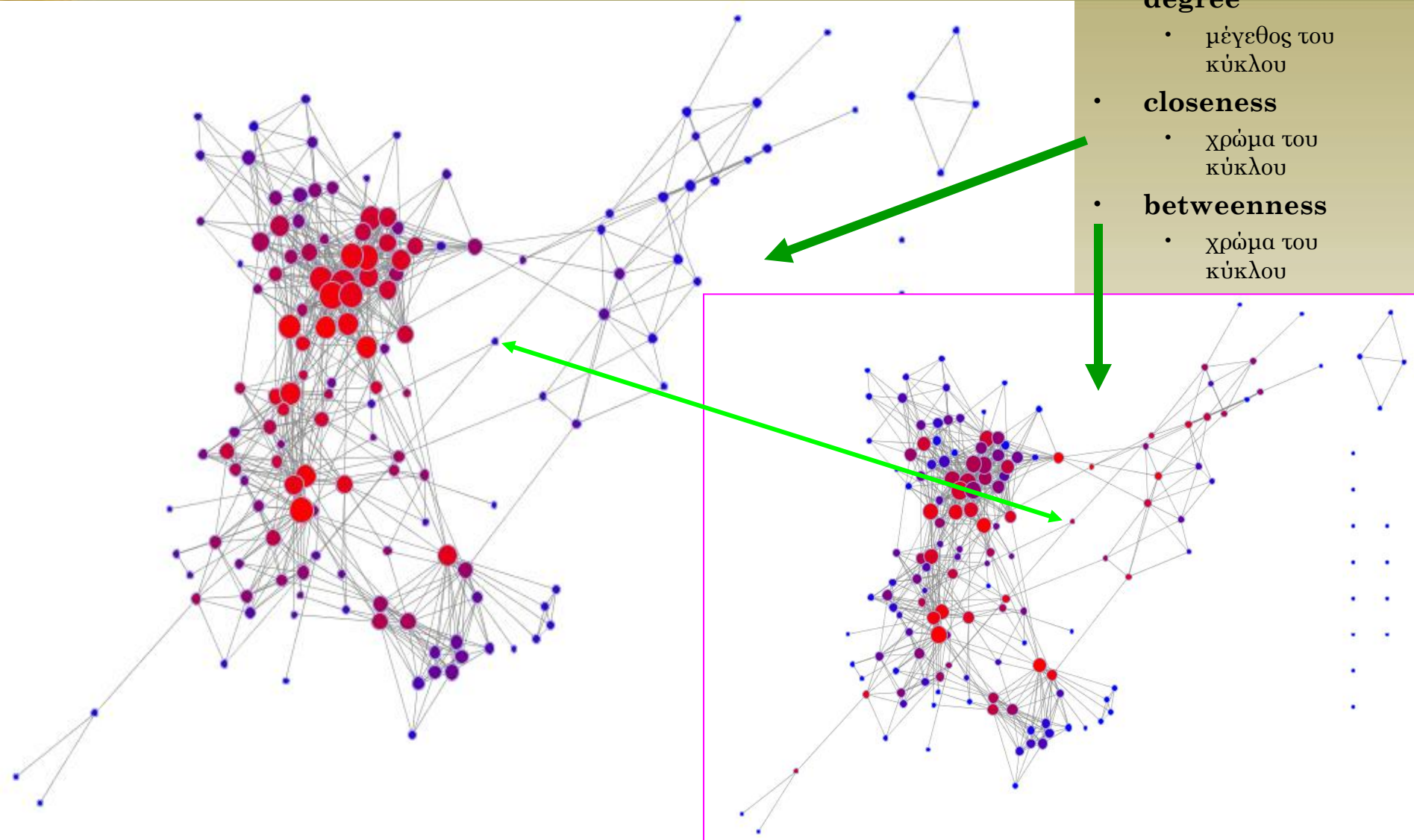
$$C'_c(A) = \left[\frac{\sum_{j=1}^N d(A, j)}{N-1} \right]^{-1} = \left[\frac{1+2+3+4}{4} \right]^{-1} = \left[\frac{10}{4} \right]^{-1} = 0.4$$

Κεντρικότητα εγγύτητας: Παράδειγμα





Πόσο κοντά είναι η DC και η SPBC στην CC;





Περιεχόμενα

Παρουσιάσαμε:

- Degree centrality (DC) και centralization
- Shortest-Path Betweenness Centrality (SPBC)

σε μη-κατευθυνόμενα δίκτυα

Θα παρουσιάσουμε:

- Closeness centrality
- Bridging centrality
- επεκτάσεις σε κατευθυνόμενα δίκτυα

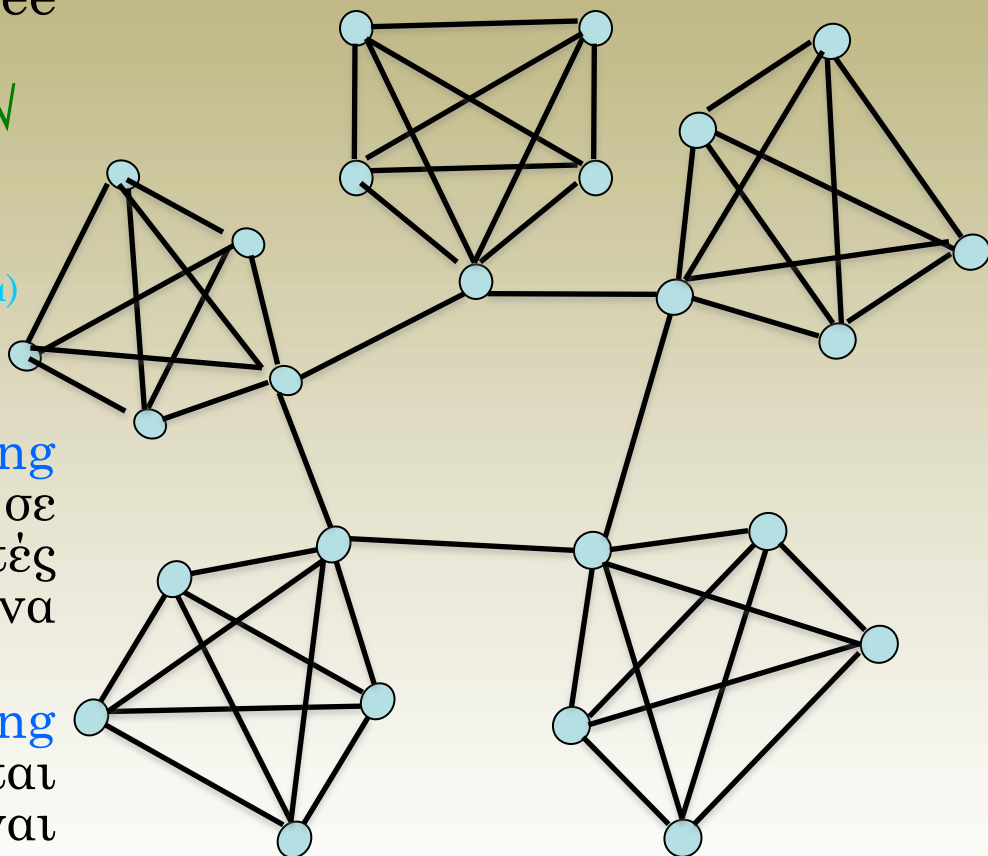
Παρατήρηση

- Βασικές ιδιότητες των scale-free δικτύων

- Power-law degree distribution ✓
- Small world ✓
- Robustness (αργότερα)
- Hierarchical modularity (αργότερα)
- Assortative (αργότερα)

- Θα υπάρχουν μερικές/οί bridging κόμβοι/ακμές μεταξύ modules σε scale-free δίκτυα με βάση αυτές τις ιδιότητες, και μπορούμε να τις/τους αναγνωρίσουμε οπτικά

- Η εύρεση των bridging κόμβων/ακμών, που βρίσκονται μεταξύ των modules, είναι ενδιαφέρουσα και χρήσιμη εργασία (ευρωστία δικτύου, προστασία μονοπατιών, κ.τ.λ.)





Κίνητρο

- Οι υπάρχουσες μετρικές δεν είναι ικανές να αναγνωρίσουν τους bridging nodes/edges: κυρίως κυριαρχούνται από το βαθμό του κόμβου που εξετάζουμε
- Η ενδιαμεσότητα μιας ακμής(!?) έχει επίσης ισχυρή ροπή προς τους κόμβους με μεγάλο βαθμό
- Υψηλή τάση για cluttering στο κέντρο του δικτύου. Έτσι, είναι δύσκολο να διαχωρίσουμε τους bridging nodes/edges από άλλα είδη nodes/edges



Γέφυρα

- Μια γέφυρα θα πρέπει να βρίσκεται πάνω σε κάποιο **σημαντικό μονοπάτι**, π.χ., ελάχιστο μονοπάτι
- Μια γέφυρα θα πρέπει να βρίσκεται **μεταξύ modules**
- Οι γειτονικές περιοχές ενός κόμβου θα πρέπει να έχουν **low range of public domain** μεταξύ τους



Ενδιαμεσότητα και Συντελεστής Γέφυρας (bridging coefficient)

- **Betweenness**: καθολική σημαντικότητα ενός κόμβου/ακμής από την σκοπιά των ελαχίστων μονοπατιών

$$SPBC(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

$$SPBC(e) = \sum_{s \neq t \in V} \frac{\sigma_{st}(e)}{\sigma_{st}}$$

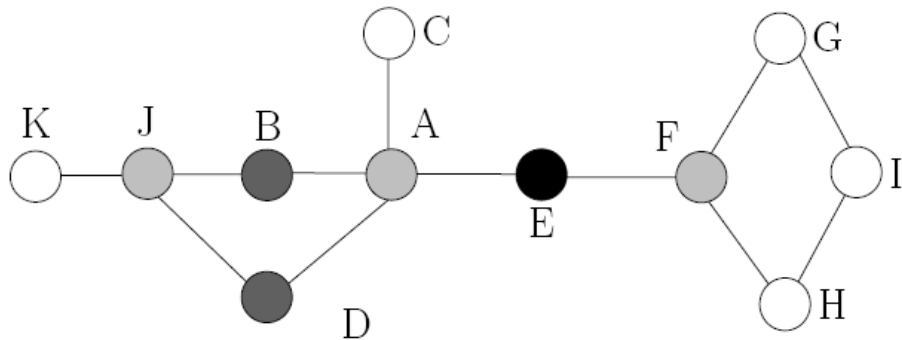
- **Bridging Coefficient**: μετρική που μετρά την έκταση του πόσο καλά ένας κόμβος ή ακμή βρίσκεται μεταξύ καλά διασυνδεδεμένων περιοχών
 - Ο αριθμός των ακμών που εισέρχονται ή εξέρχονται το άμεσο “γειτονικό” υπογράφημα ενός κόμβου v σχετικά με τον αριθμό των ακμών που είναι εντός του άμεσου “γειτονικού” υπογραφήματος του v είναι υψηλός στις γέφυρες
 - Η μέση πιθανότητα του να εγκαταλείψεις το άμεσο υπογράφημα ενός κόμβου v



Bridging coefficient κόμβου

Bridging Coefficient: $\Psi(v) = \frac{1}{d(v)} \sum_{i \in N(v)} \frac{\delta(i)}{d(i)-1}$

όπου $d(v)$ είναι ο βαθμός ενός κόμβου v και $\delta(v)$ είναι ο αριθμός των ακμών που εξέρχονται από το άμεσο γειτονικό υπογράφημα (direct neighbor subgraph) του κόμβου v από εκείνες που προσπίπτουν σε κάθε άμεσο γειτονικό κόμβο i του κόμβου v



C_b : SP Betweenness centrality

BC: Bridging coefficient

C_R : Bridging centrality

Node	Degree	C_B	BC	C_R
E	2	0.53333	0.85714	0.45713
B	2	0.15555	0.85714	0.13333
D	2	0.15555	0.85714	0.13333
F	3	0.47777	0.22222	0.10617
A	4	0.65555	0.10000	0.06555
J	3	0.21111	0.16666	0.03518



Bridging coefficient ακμής

Bridging Coefficient:
$$\Psi(e) = \frac{d(i)\Psi(i) + d(j)\Psi(j)}{(d(i) + d(j))(|C(i, j)| + 1)}$$

όπου οι κόμβοι i και j είναι τα δυο άκρα της ακμής e , $d(i)$ είναι ο βαθμός του κόμβου i , $\Psi(i)$ είναι ο bridging coefficient του κόμβου i , $C(i, j)$ είναι το σύνολο των κοινών άμεσων γειτόνων των κόμβων i και j

Δηλαδή:

Ο bridging coefficient για μια ακμή e ορίζεται ως το γινόμενο του διαβαμισμένου μέσου όρου (weighted average) των bridging coefficient των δυο άκρων i και j της ακμής e και του αντίστροφου του αριθμού των κοινών άμεσων γειτόνων των κόμβων i και j



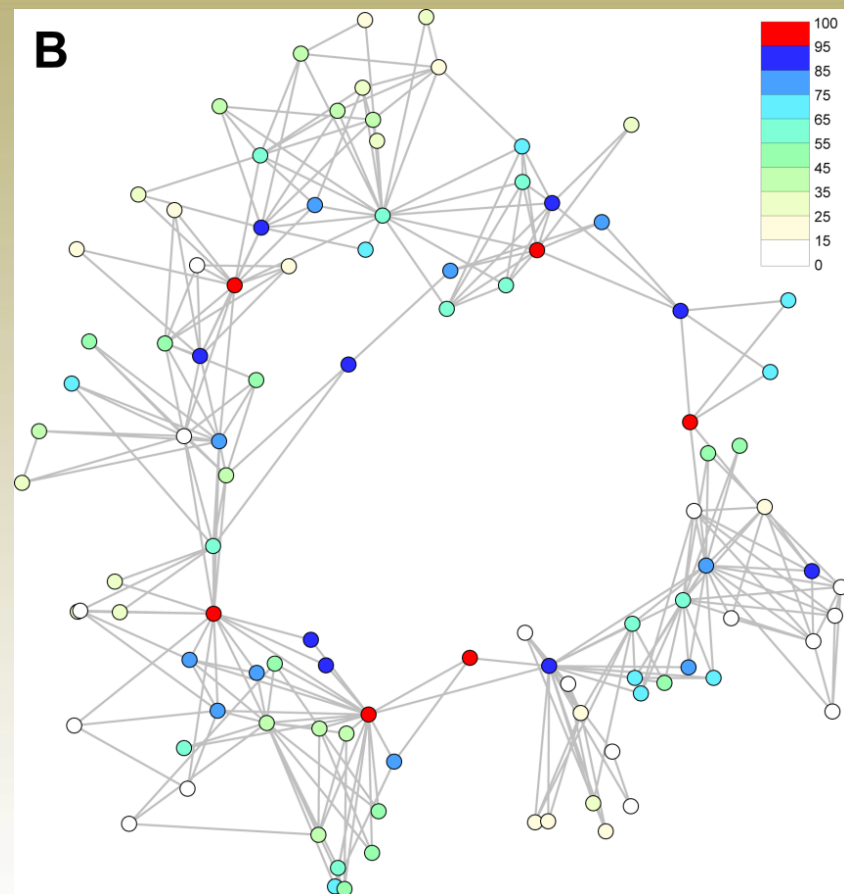
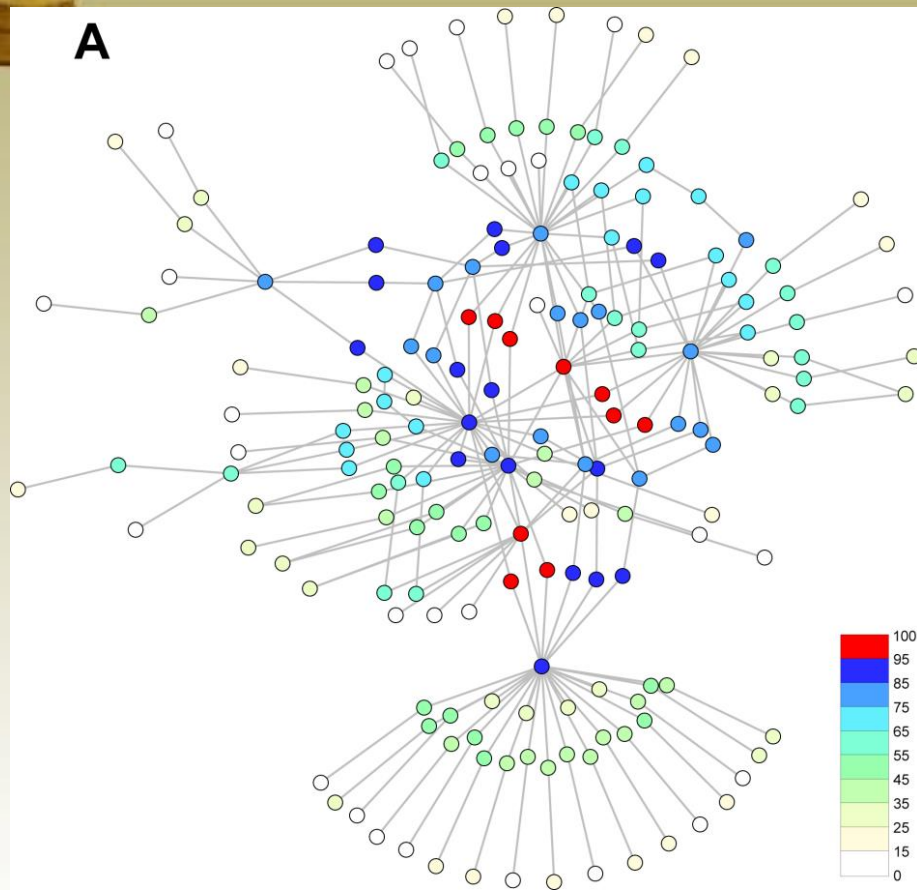
Bridging Centrality

Η bridging centrality ορίζεται ως το γινόμενο του rank της ενδιάμεσότητας επί το rank του the bridging coefficient

$$C_{Br}(v) = R_{SPBC(v)} \bullet R_{\Psi(v)}$$

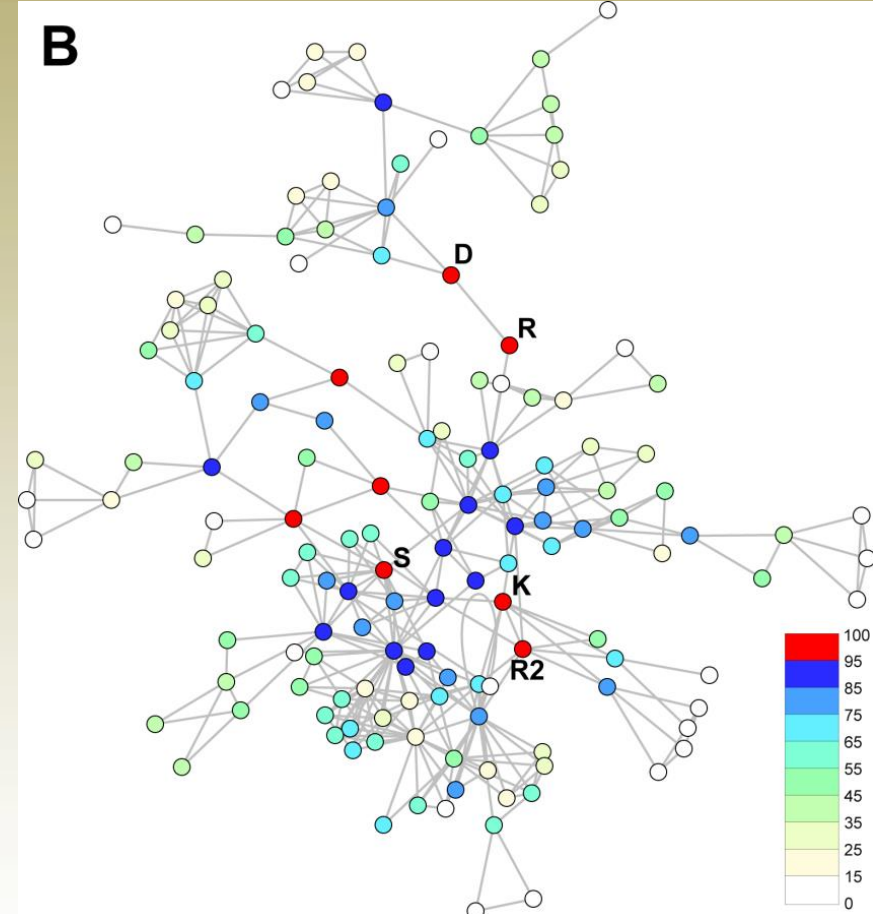
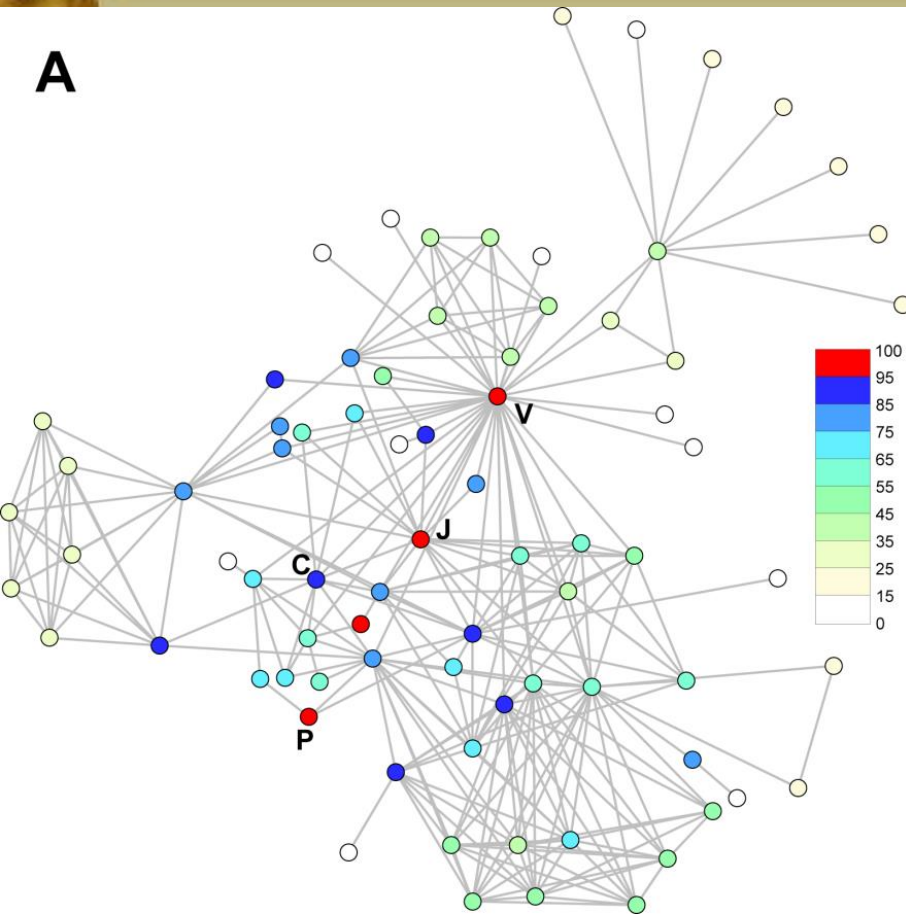
$$C_{Br}(e) = R_{SPBC(e)} \bullet R_{\Psi(e)}$$

ΒC: Παραδείγματα από το Web



Results for Web Networks: Figures A and B show the results for the AT&T Web Network and RPI Web Network, respectively. The nodes with the highest 0-5th percentile of values for the bridging centrality are highlighted in red circles; the nodes with the lowest values of bridging centrality are the 85th-100th percentiles and are highlighted in white circles. The color map for the percentile values is shown in the Figure.

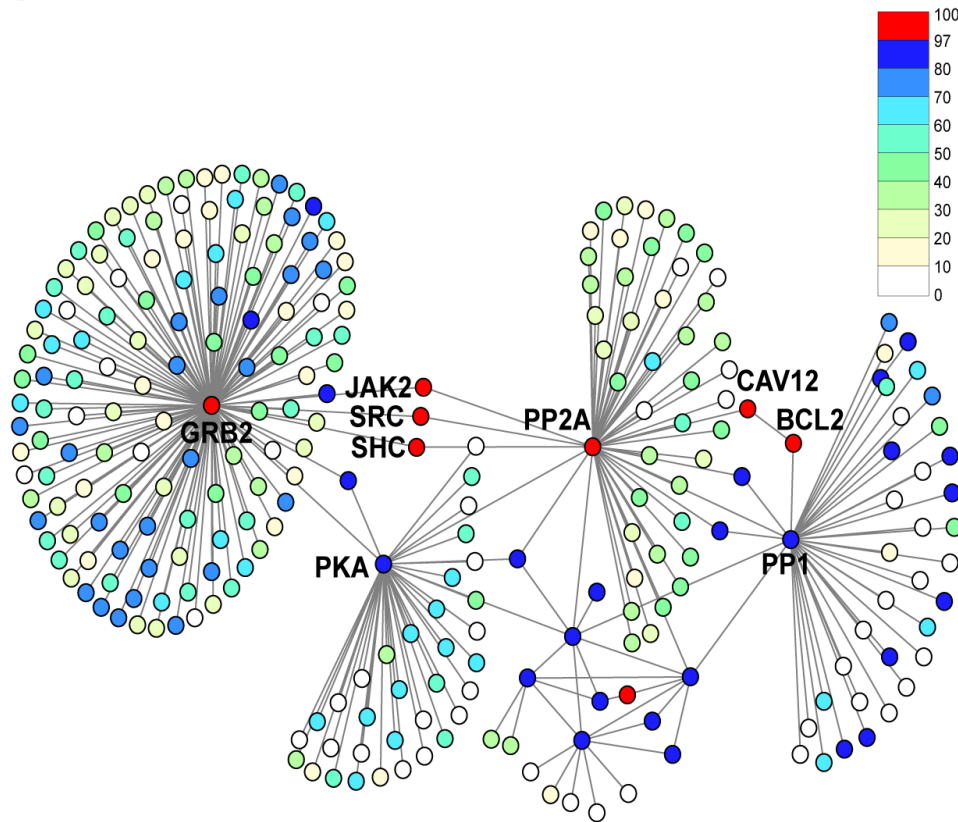
ΒC: Παραδείγματα από κοινωνικά δίκτυα



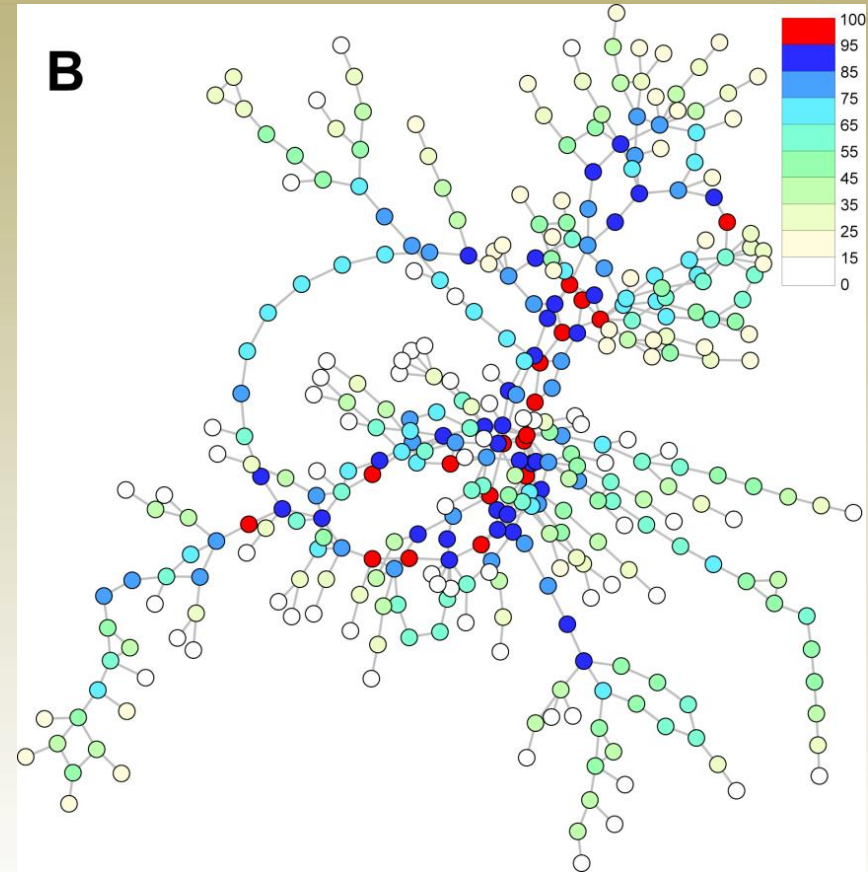
Results for Social Networks: Figures A and B show the results for the Les Miserable Character Network and Physics Collaboration Network, respectively. The nodes with the highest 0-5th percentile of values for the bridging centrality are highlighted in red circles; the nodes with the lowest values of bridging centrality are the 85th-100th percentiles and are highlighted in white circles. The nodes corresponding to Valjean (V), Javert (J), Pontmercy (P) and Cosette (C) are labeled in Figure A. The nodes corresponding to Rothman (R), Redner (R2), Dodds (D), Krapivsky (K) and Stanley (S) are labeled in Figure B. The color map for the percentile values is shown in the Figure.

BC: Παραδείγματα από βιολογικά δίκτυα

A



B



- **Results for Biological Networks:** Figures A and B show the results the Cardiac Arrest Network and Yeast Metabolic Network, respectively. The nodes corresponding to Src, Shc and Jak2 (J2) are labeled in Figure A. The nodes with the highest 0-5th percentile of values for the bridging centrality are highlighted in red circles; the nodes with the lowest values of bridging centrality are the 85th-100th percentiles and are highlighted in white circles. The color map for the percentile values is shown in the Figure.



Περιεχόμενα

Παρουσιάσαμε

- Degree centrality (DC) και centralization
- Shortest-Path Betweenness Centrality (SPBC)

σε μη-κατευθυνόμενα δίκτυα

Θα παρουσιάσουμε

- Closeness centrality
- Bridging centrality
- επεκτάσεις σε κατευθυνόμενα δίκτυα



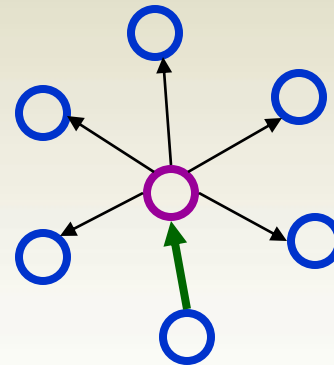
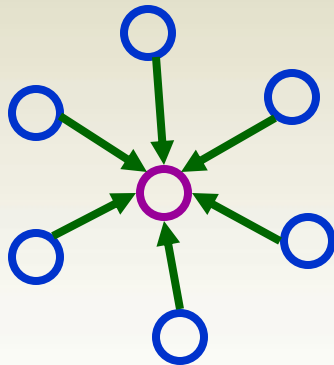
Prestige σε κατευθυνόμενα δίκτυα

- Όταν το ‘prestige’ είναι η σωστή λέξη:
 - Θαυμασμός (admiration)
 - Επιρροή (influence)
 - Χορηγός (gift-giving)
 - Εμπιστοσύνη (trust)
- Η κατεύθυνση ειδικά σημαντική σε περιπτώσεις όπου δεν υπάρχει αμοιβαιότητα (π.χ., δίκτυο επιλογής συνδαιτυμόνων)
- Όταν το ‘prestige’ δεν είναι η σωστή λέξη
 - Σύμβουλος gives advice to (can reverse direction)
 - Εντολέας (gives orders to)
 - Δανειστής (lends money to)
 - Αντιπαθεί (dislikes)
 - Δεν εμπιστεύεται (distrusts)



Επεκτάσεις της DC σε κατευθυνόμενα δίκτυα

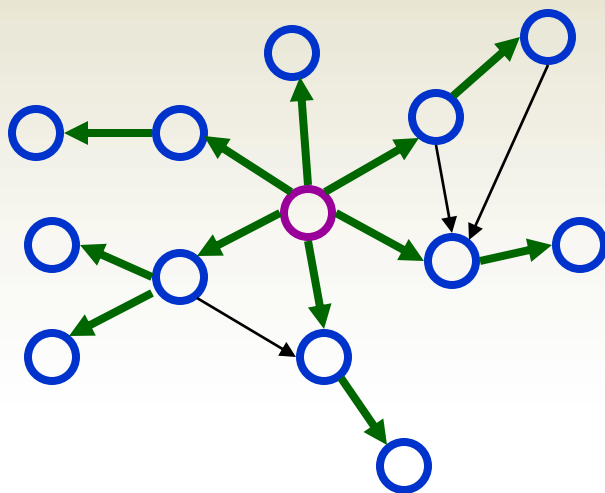
- degree centrality
 - indegree centrality
 - Ένα άρθρο που αναφέρεται (is cited) από πολλά άλλα έχει υψηλό prestige
 - Κάποιο πρόσωπο που υποδεικνύεται από πολλούς άλλους για π.χ., κάποιο έπαινο έχει υψηλό prestige





Επεκτάσεις της CC σε κατευθυνόμενα δίκτυα

- Η CC συνήθως υπονοεί
 - Όλα τα μονοπάτια οδηγούν σε σένα
 - Τα μονοπάτια από σένα θα πρέπει να οδηγούν παντού
- Συνήθως, εξετάζει μόνο κόμβους στους οποίους ο κόμβος i έχει πρόσβαση





Κατευθυνόμενα geodesics

- Κάποιος κόμβος δεν βρίσκεται απαραίτητα επάνω σε ένα ελάχιστο μονοπάτι από το j προς το k , εάν βρίσκεται επάνω σε ένα ελάχιστο μονοπάτι από το k προς το j





Επέκταση της SPBC σε κατευθυνόμενα δίκτυα

- Εξετάζουμε το κλάσμα όλων των κατευθυνόμενων μονοπατιών μεταξύ δυο οποιονδήποτε κόμβων που διέρχονται από κάποιο κόμβο

betweenness του κόμβου i

μονοπάτια μεταξύ j και k που διέρχονται από τον i

$$C_B(i) = \sum_{j,k} g_{jk}(i) / g_{jk}$$

όλα τα μονοπάτια μεταξύ j και k

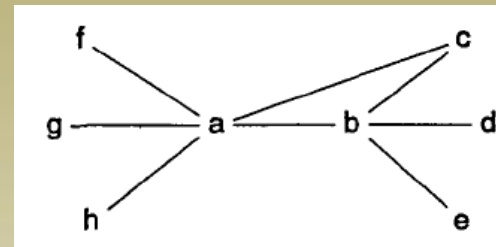
- Μόνη τροποποίηση: όταν κανονικοποιούμε, έχουμε $(N-1)*(N-2)$ αντί για $(N-1)*(N-2)/2$, επειδή υπάρχουν διπλάσια ordered ζεύγη απ' ότι unordered ζεύγη

$$C'_B(i) = C_B(i) / [(N-1)(N-2)]$$

Centralities για groups κόμβων

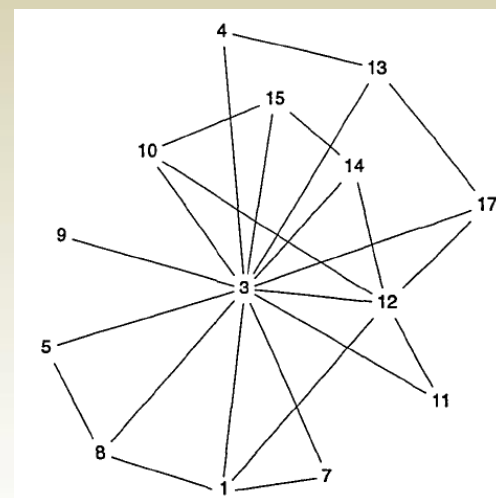
- Group Degree Centrality (GDC)

- Ο αριθμός των non-group κόμβων που συνδέονται με μέλη του group
 - Πολλαπλοί σύνδεσμοι στον ίδιο κόμβο: άπαξ
 - $GDC(\{a,b\}) = 6$



- Group Closeness Centrality (GCC)

- Το άθροισμα των αποστάσεων από το group προς τους non-group κόμβους
Ορίζεται: ως η μέγιστη, ή η ελάχιστη, ή η μέση



- Απόσταση του 12 από {8,1,7}=
1(minimum), 2 (maximum), 1,67 (mean)

- Group Betweenness Centrality (GBC)

- Το ποσοστό των shortest paths μεταξύ ζευγών non-group κόμβων τα οποία διέρχονται από τουλάχιστον έναν κόμβο του group



Centralities για groups κόμβων

Group Betweenness Centrality (GBC)

Ένας αλγόριθμος υπολογισμού

- I. Μέτρηση του αριθμού των geodesies μεταξύ κάθε ζεύγους από nongroup members, και συνακόλουθη κατασκευή ενός node-by-node πίνακα με τους μετρητές αυτούς
- II. Διαγραφή όλων των links που αφορούν σε group members και επανάληψη του προηγούμενου υπολογισμού, και κατόπιν δημιουργία ενός νέου node-by-node πίνακα με τους μετρητές
- III. Διαίρεση κάθε κελιού του νέου πίνακα με το αντίστοιχο κελί του πρώτου πίνακα
- IV. Υπολόγισε τα αθροίσματα όλων αυτών των κλασμάτων



Sink Group Betweenness Centrality

- Sink Betweenness Centrality (SBC)

- Η SBC ενός κόμβου v ισούται με το ποσοστό των shortest paths μεταξύ κάθε κόμβου i του δικτύου και ενός συγκεκριμένου *sink* κόμβου s , τα οποία διέρχονται από τον κόμβο v :

$$SBC(v) = \sum_{\substack{i=1 \\ i \neq s \neq v}}^n \frac{\sigma_{is}(v)}{\sigma_{is}}$$

s is the sink node

- Sink Group Betweenness Centrality

- Groups $\{k=1, \dots, z\}$ C_k από sink κόμβους

$$SGBC(v) = \sum_{i=1}^n \sum_{\substack{j \in \cup_{k=1}^z C_k \\ v \notin \cup_{k=1}^z C_k}} \frac{\sigma_{ij}(v)}{\sigma_{ij}}$$

- Αντί για έναν κόμβο v , μπορούμε να σκεφτούμε ένα σύνολο από κόμβους, όπως στην group version της betweenness centrality

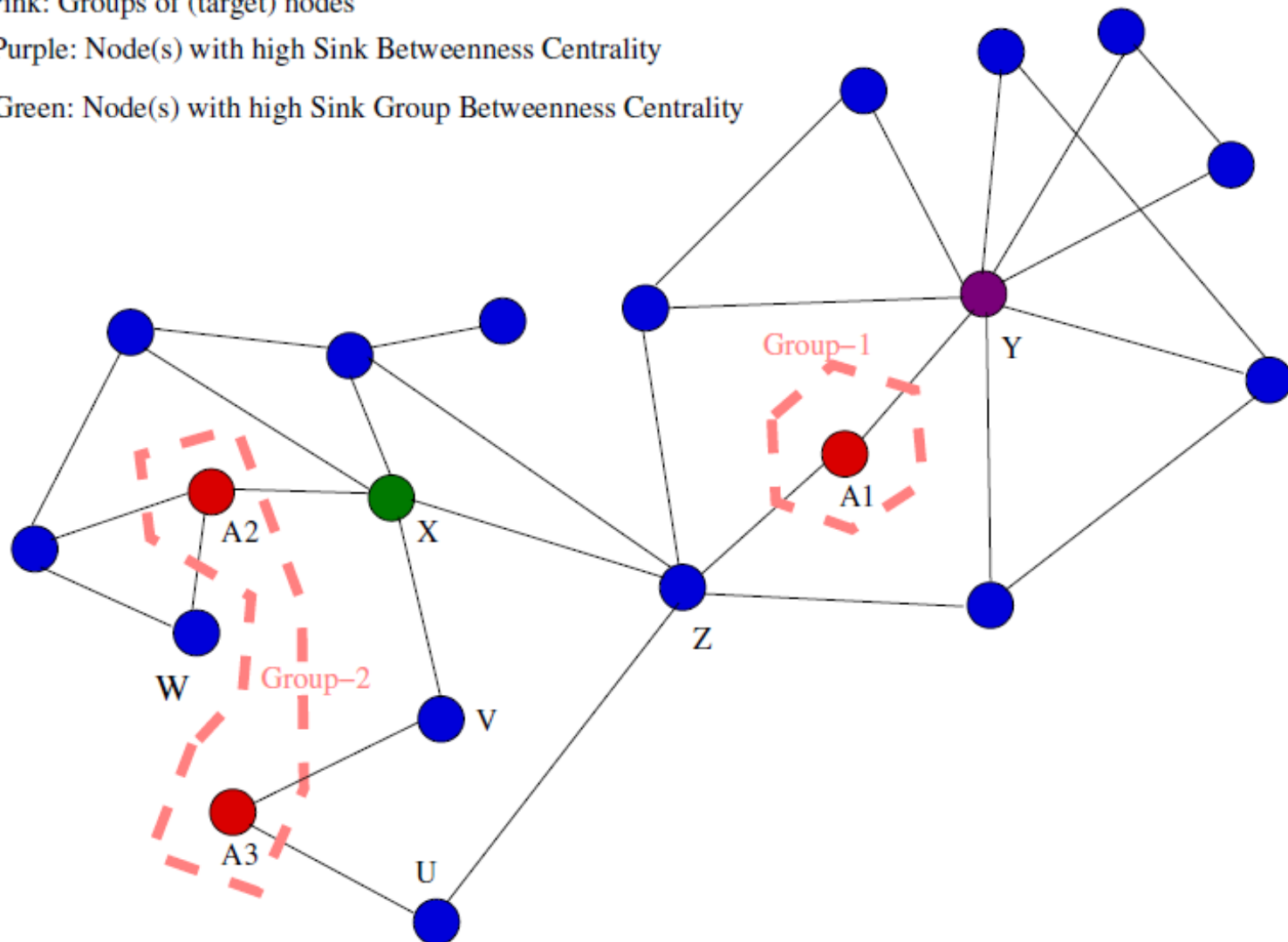
Παράδειγμα SGBC (1/2)

Red: Node(s) which are our target

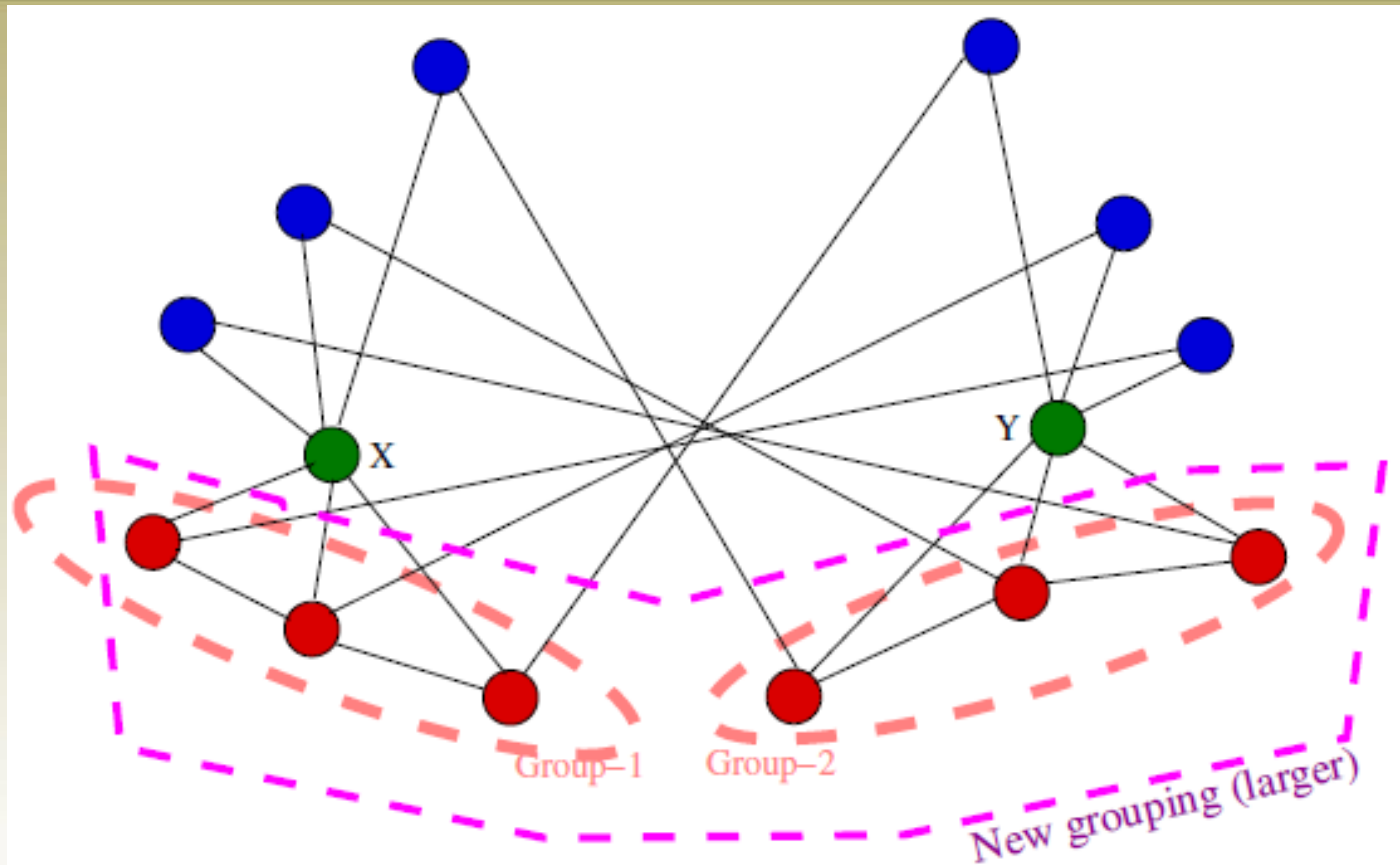
Pink: Groups of (target) nodes

Purple: Node(s) with high Sink Betweenness Centrality

Green: Node(s) with high Sink Group Betweenness Centrality



Παράδειγμα SGBC (2/2)



Red: Node(s) which are our target

Pink: Initial grouping of target nodes

Green: Nodes with high SGBC with respect to the initial grouping

Purple: New grouping (all red nodes into the same group)



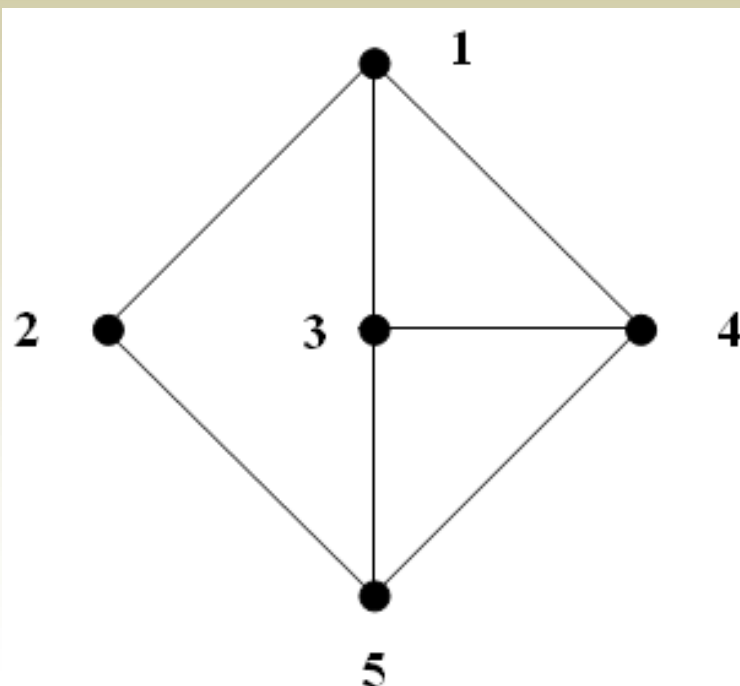
Σύνοψη

- Κεντρικότητα
 - Πολλές μετρικές: degree, betweenness, closeness, bridging και οι **group versions** αυτών, π.χ., Group Betweenness Centrality
 - Θα δούμε και τις *φασματικές κεντρικότητες* (*spectral centralities*), όπως η Katz, PageRank, Bonacich στην επόμενη διάλεξη
 - Μπορεί να είναι ανομοιόμορφα κατανεμημένες
 - Μέτρηση με την έννοια της centralization
 - Επεκτάσεις σε κατευθυνόμενα δίκτυα:
 - Prestige

Ασκήσεις

Άσκηση 1.

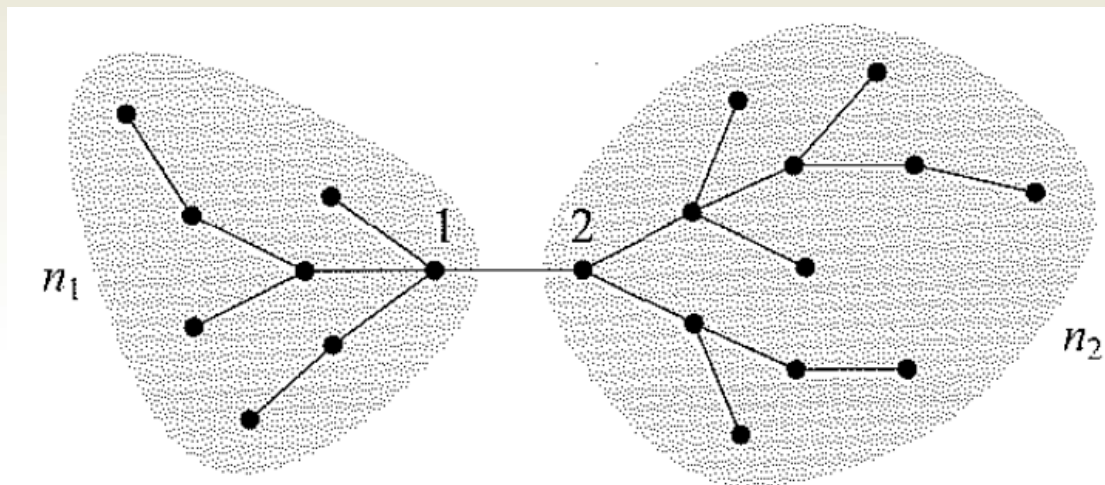
Να βρεθεί η closeness centrality των κόμβων του γραφήματος.



Ασκήσεις

Άσκηση 2.

Να βρεθεί η σχέση που συνδέει την closeness centrality του κόμβου 1 (C_1) με την closeness centrality του κόμβου 2 (C_2) συναρτήσει του συνολικού αριθμού των κόμβων n του δικτύου, και των αριθμών των κόμβων n_1 και n_2 των δυο συνιστωσών του παρακάτω δικτύου, δηλ., μια αναλυτική παράσταση στη μορφή π.χ., $C_1 = f(C_2, n_1, n_2, n)$ για κάποια συνάρτηση f , η οποία ζητάται.





Ασκήσεις

Άσκηση 3.

- I. Να οριστεί η betweenness centrality για edges.
- II. Να οριστεί η degree centrality για edges.

Λύση.

- I. Προφανώς με τον ανάλογο τρόπο.
- II. Διάφορες προσεγγίσεις, αλλά μπορούμε να ορίσουμε την *nearest-neighbor edge centrality* ως ακολούθως:

$$DC(e = (a, b)) = \frac{DC(a) + DC(b) - 2}{|DC(a) - DC(b)| + 1}$$

Δηλαδή, οι κόμβοι της ακμής να συνδέονται με πολλούς άλλους κόμβους (ο αριθμητής), και μεταξύ τους να είναι παρόμοιοι (ο παρονομαστής)

Οικογένειες μετρικών κεντρικότητας

