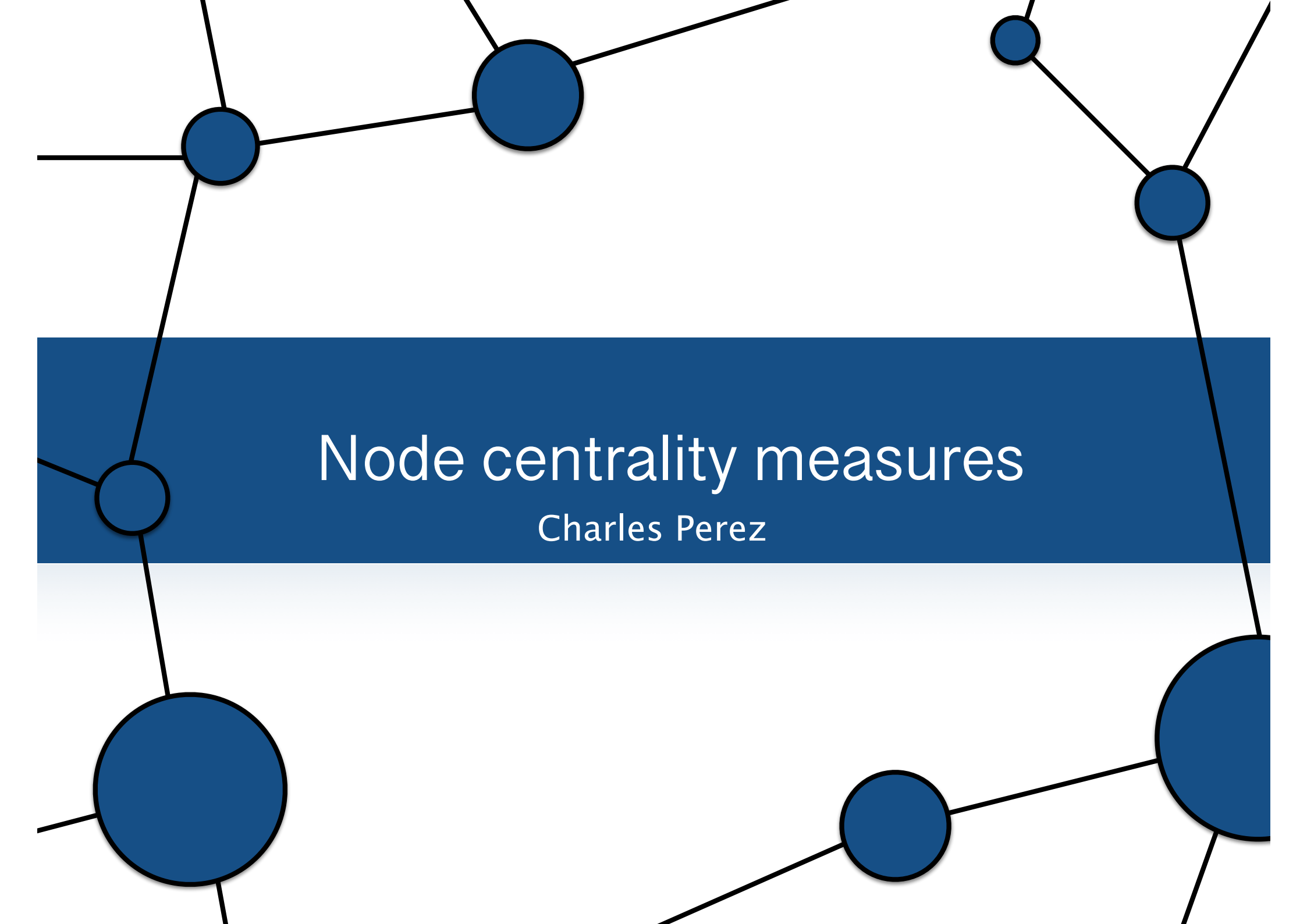
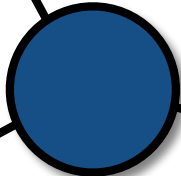




Node centrality measures

Charles Perez



Why studying centrality?

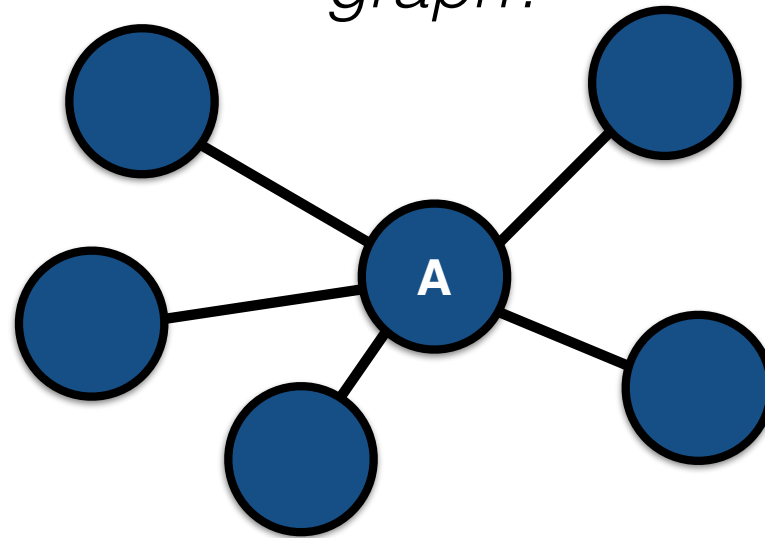
- Identify key nodes of a graph
 - Influent users
 - Celebrities, popular individuals
 - Experts
 - Key webpages (search engines)
- Applications : digital marketing, medecine, recommandation systems.



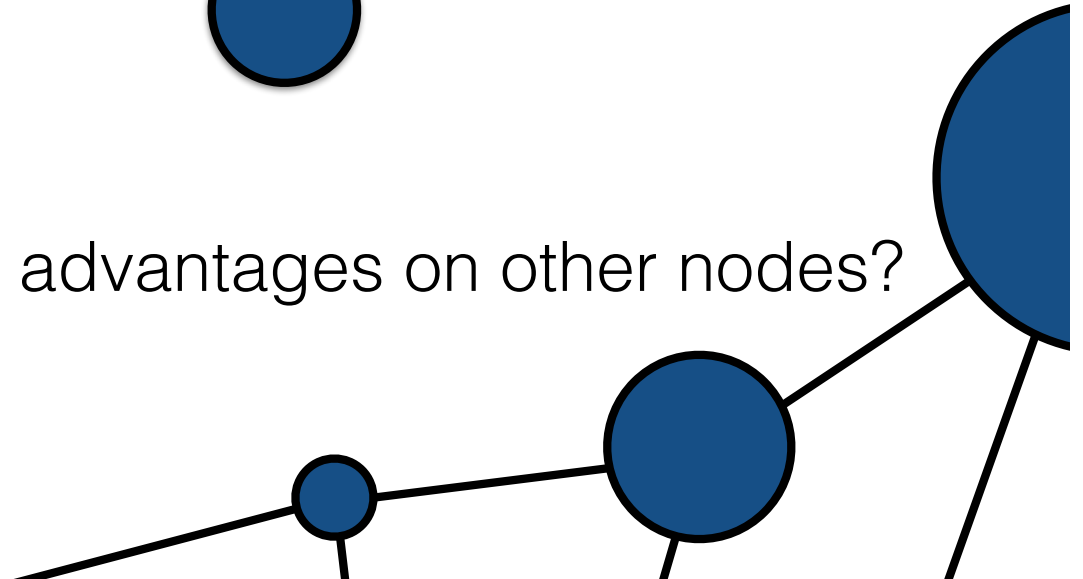


Measuring node importance

How to define important nodes from their position in the graph?



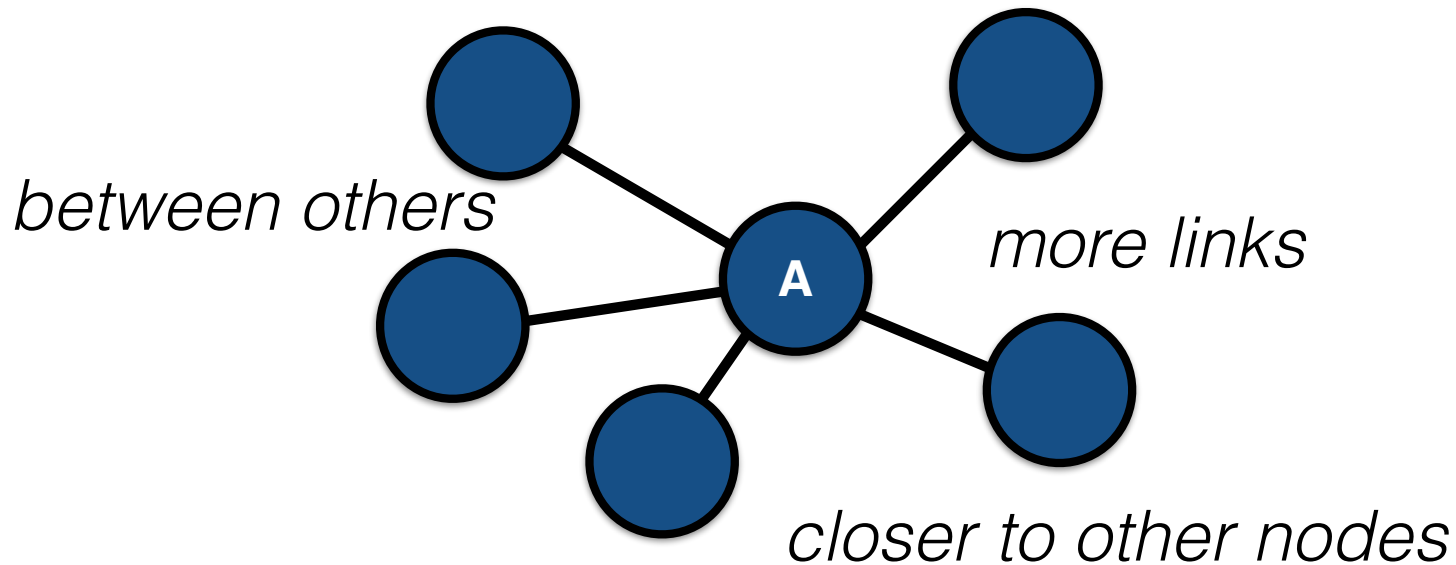
Freeman : Node **A** has three advantages on other nodes?





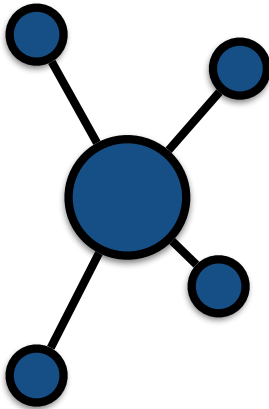
Measuring node importance

How to define important nodes from their position in the graph?

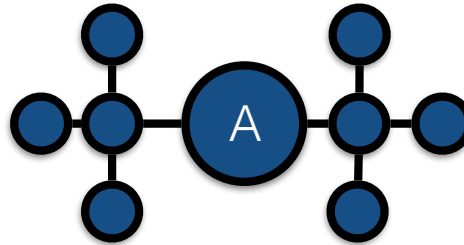


Node centrality

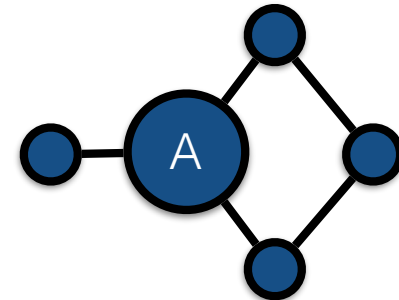
**Degree
centrality**



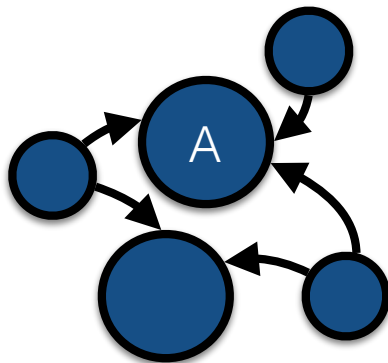
**Betweenness
centrality**



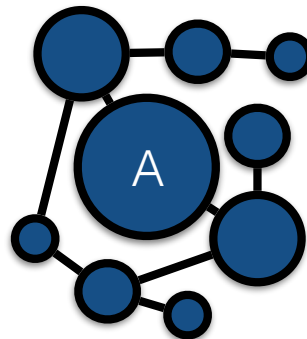
**Closeness
centrality**



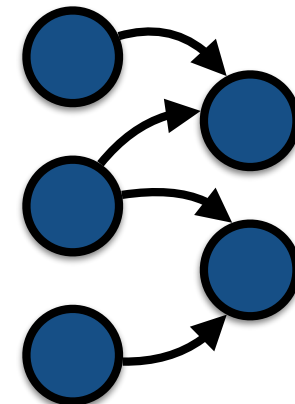
Page Rank

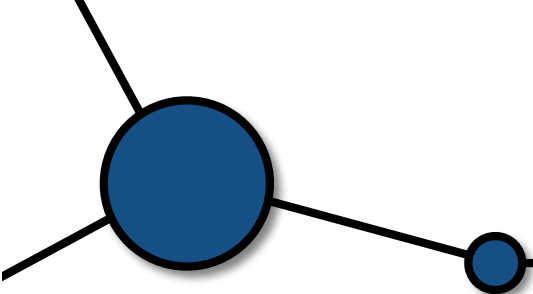


**Eigenvalues
centrality**



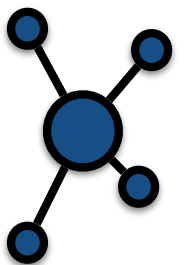
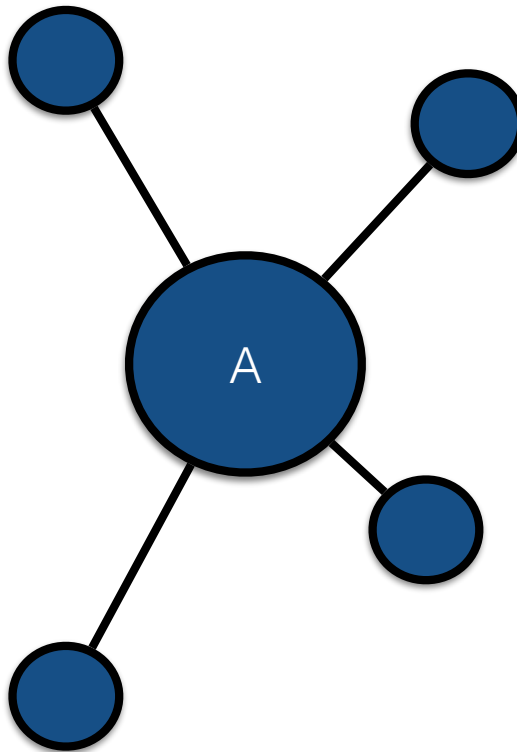
HITS





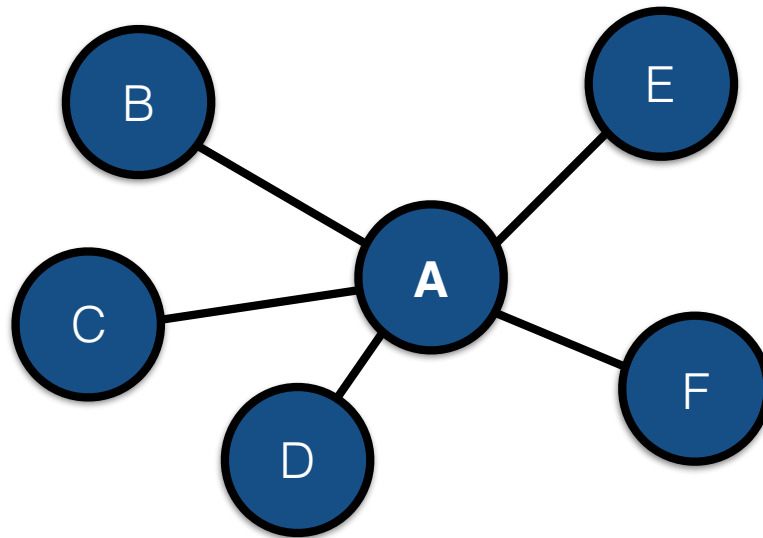
Degree centrality

- Hypothesis : More links a node has more important it is.



Degree centrality

- Definition : Degree equals to the number of links the node has (number of neighbours).



$d_a=5$

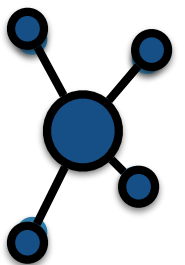
$d_b=1$

$d_c=1$

$d_d=1$

$d_e=1$

$d_f=1$



Profils ayant mentionné PSB + Profils

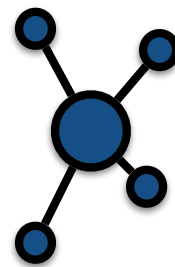
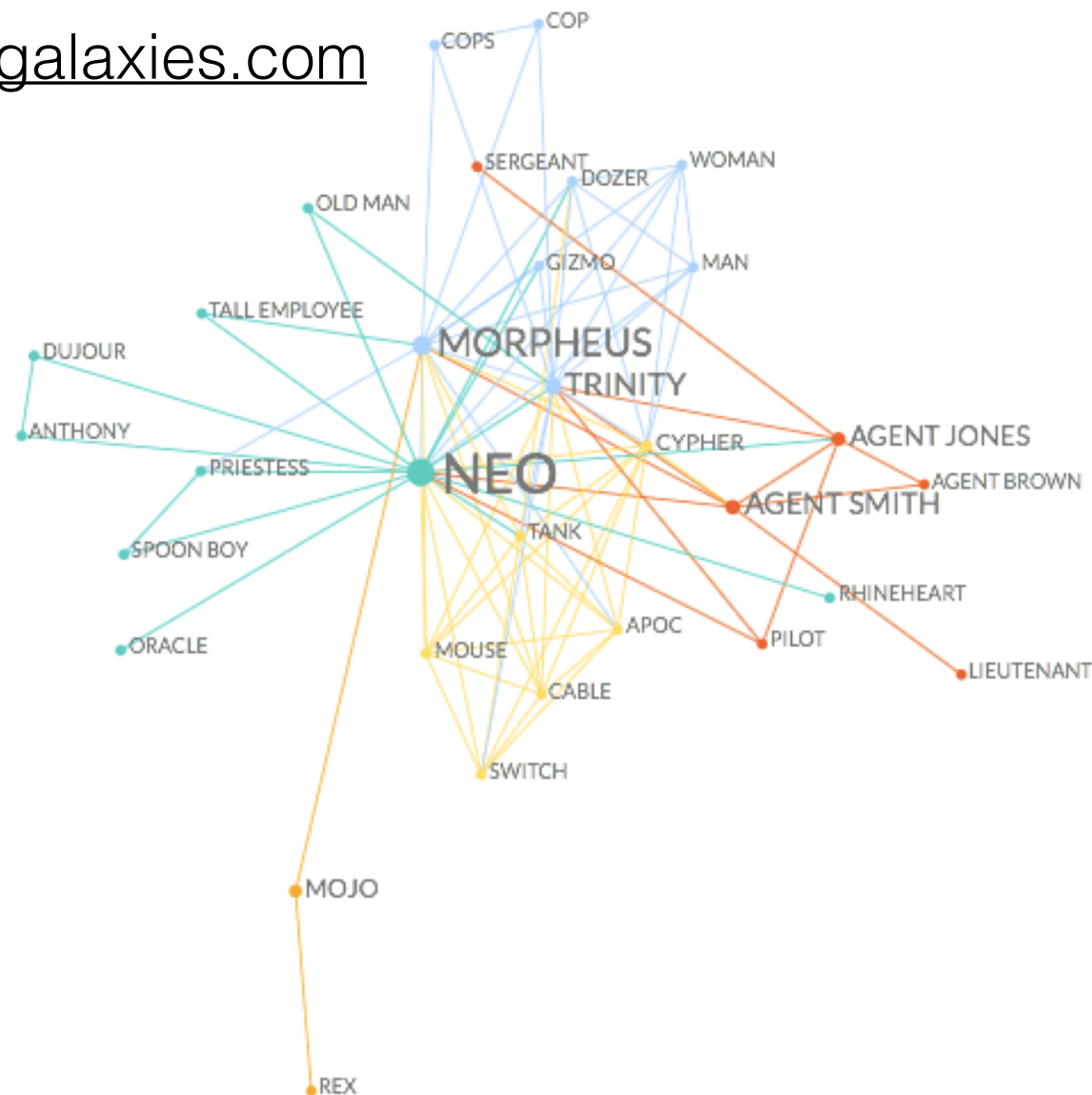
Profils qui ont mentionné PSB

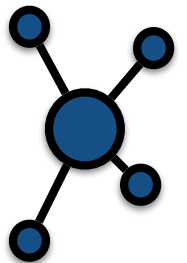
Profils ayant mentionné PSB + Profils

Profils qui ont mentionné PSB

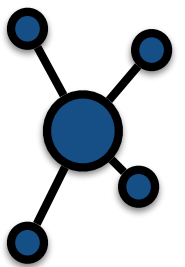
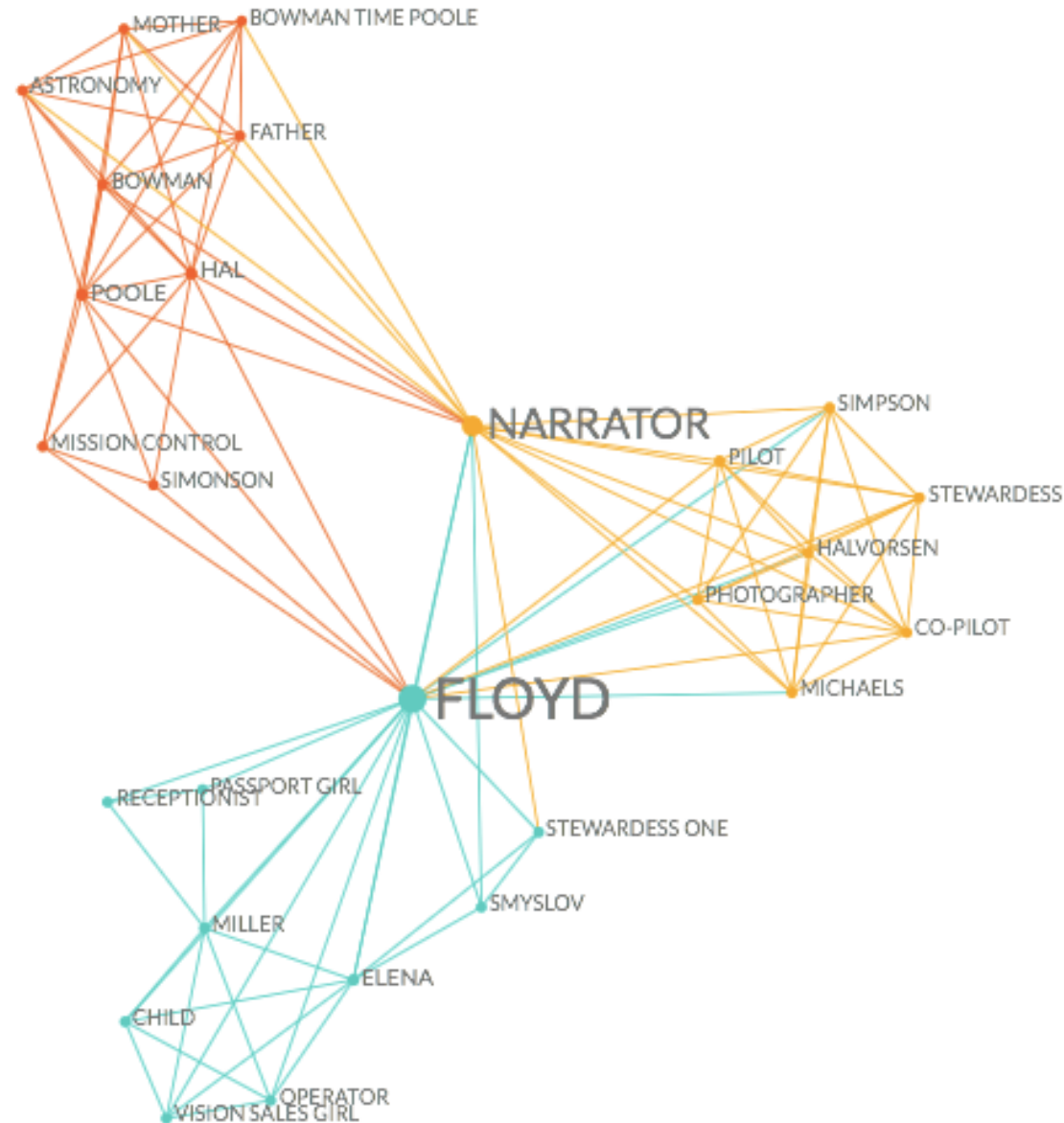
THE MATRIX 1999

<http://www.moviegalaxies.com>

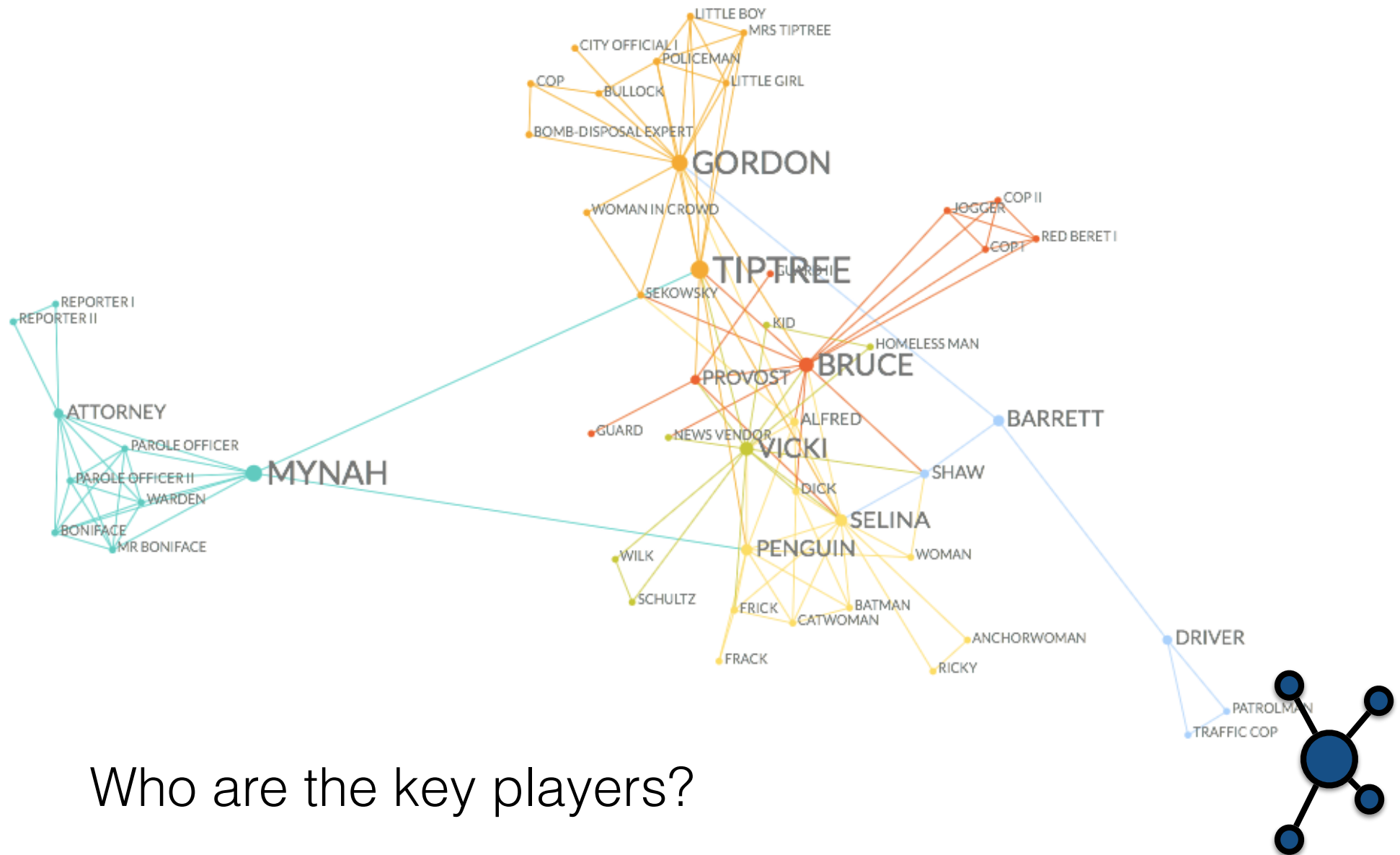




2001: A SPACE ODYSSEY 1968



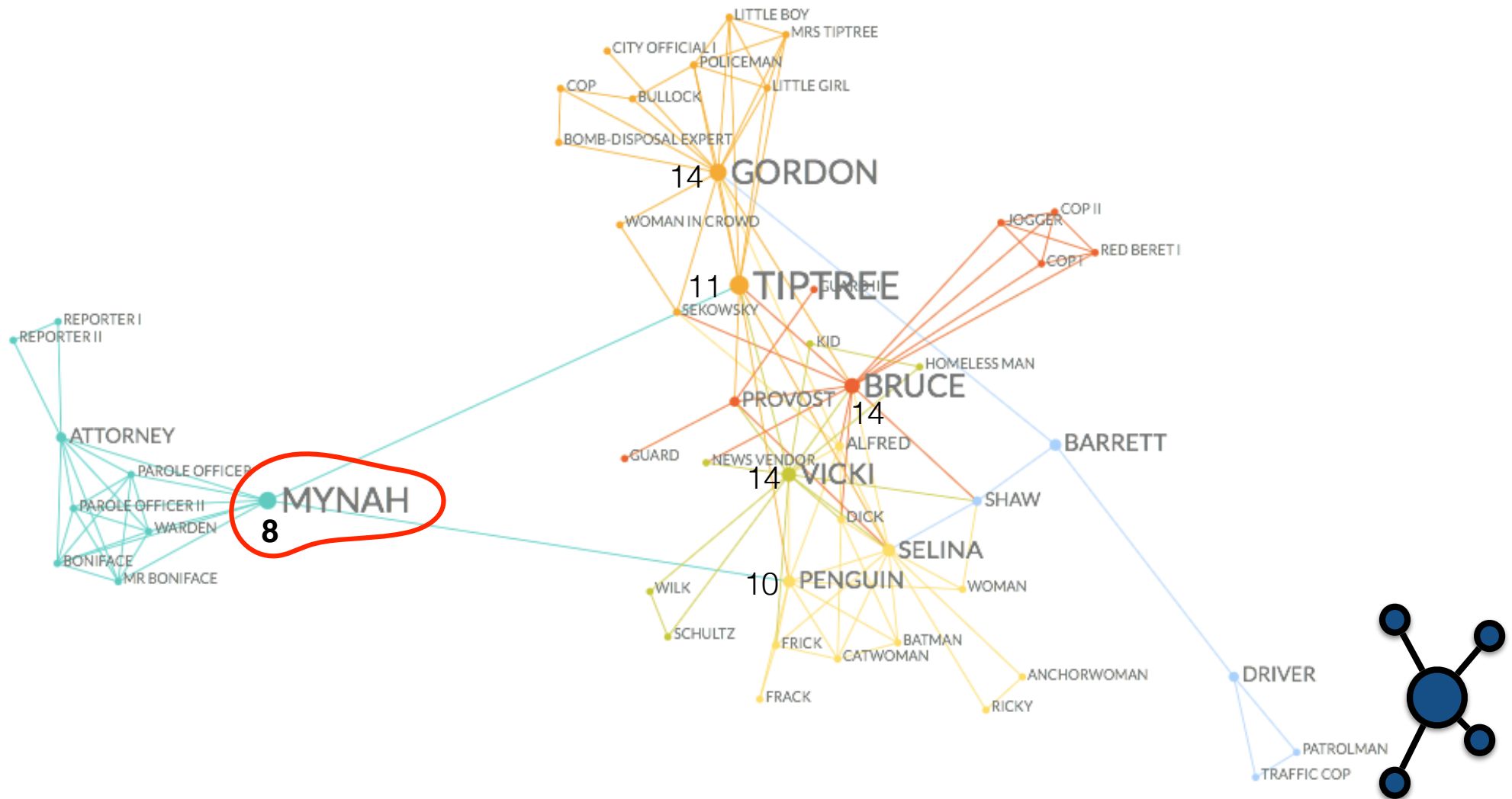
BATMAN RETURNS 1992



Who are the key players?

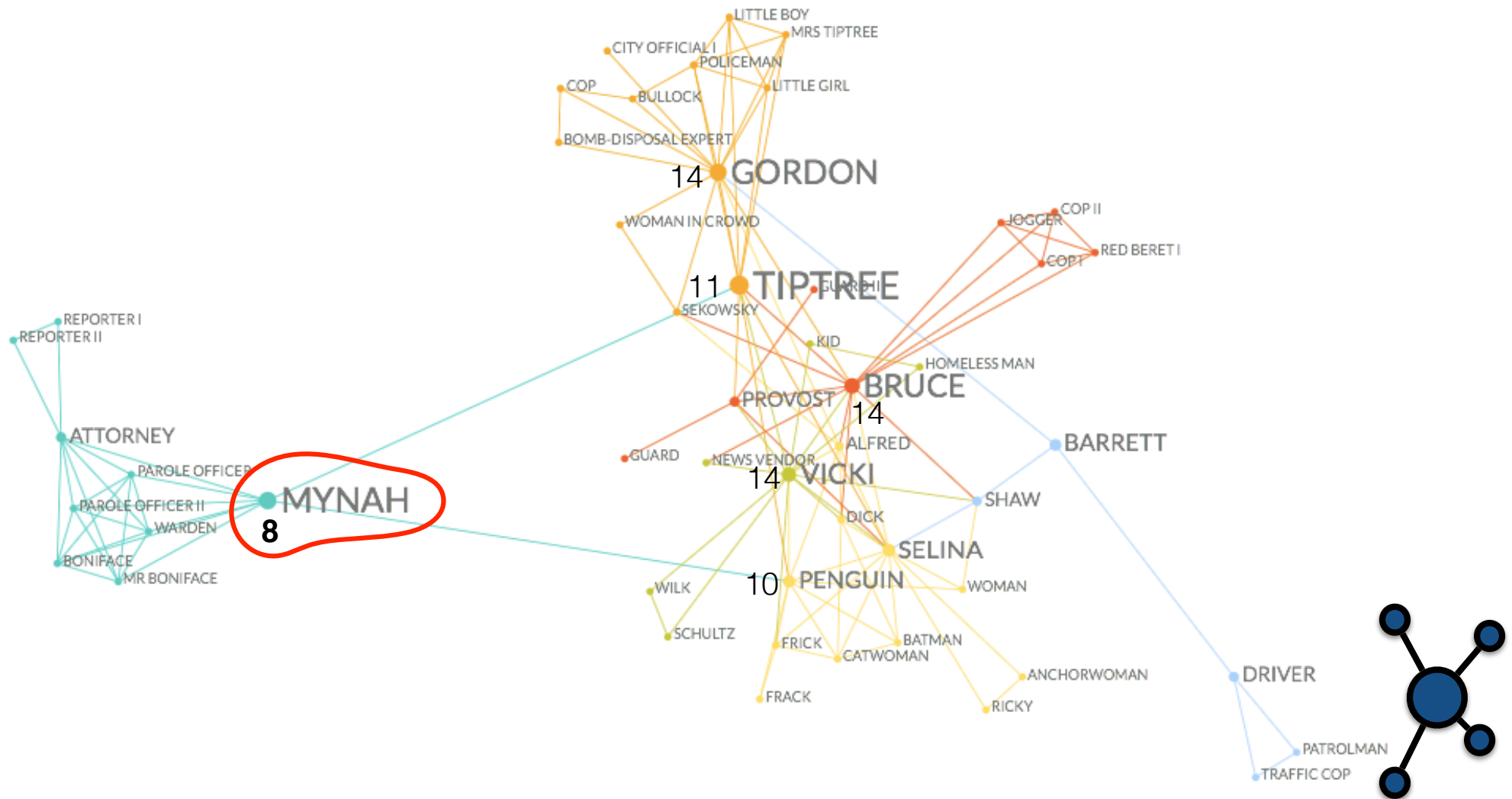
Number of nodes only?

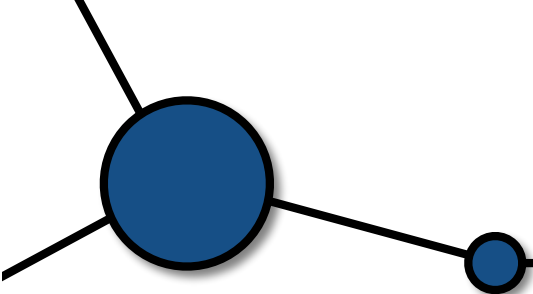
Despite a low degree, Mynah has a key role between groups.



Number of nodes only?

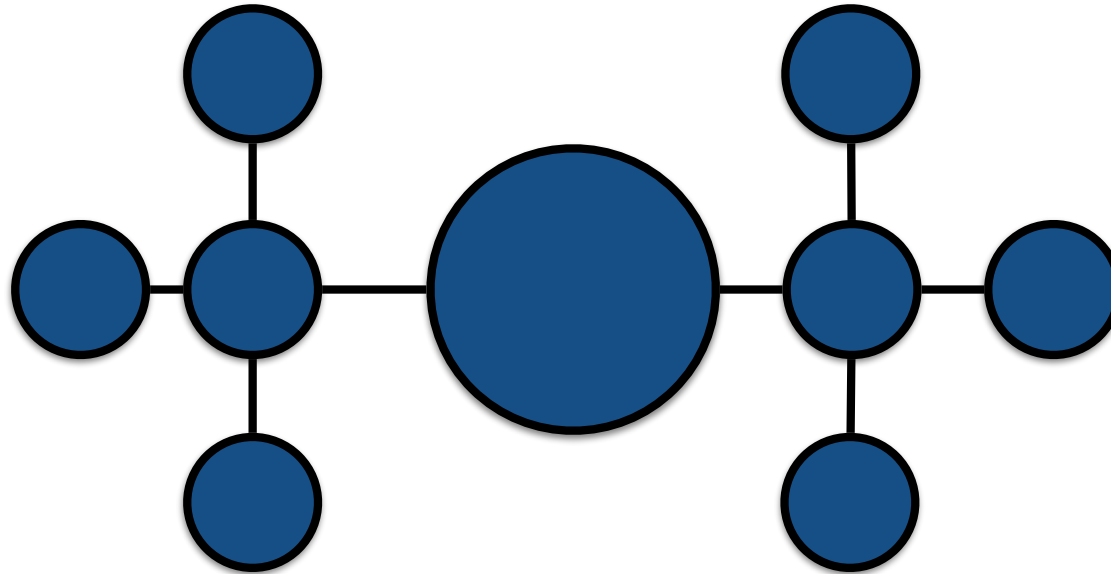
How to measure this betweenness centrality?





Betweenness centrality

- Hypothesis: An important node appears in-between many other nodes (on the path between)



Betweenness centrality

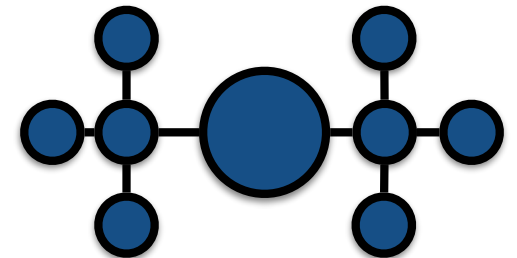
- Definition: Betweenness centrality of a node i equals the the number of shortest path (between any couple of node) that passes through this nodes divided by the number of shortest path existing between these two nodes.

Number of shortest path between j and k passing through x

Number of shortest path between j and k

$$C_B(x) = \sum_{j < k} g_{jk}(x) / g_{jk}$$

Sum over all couples of nodes

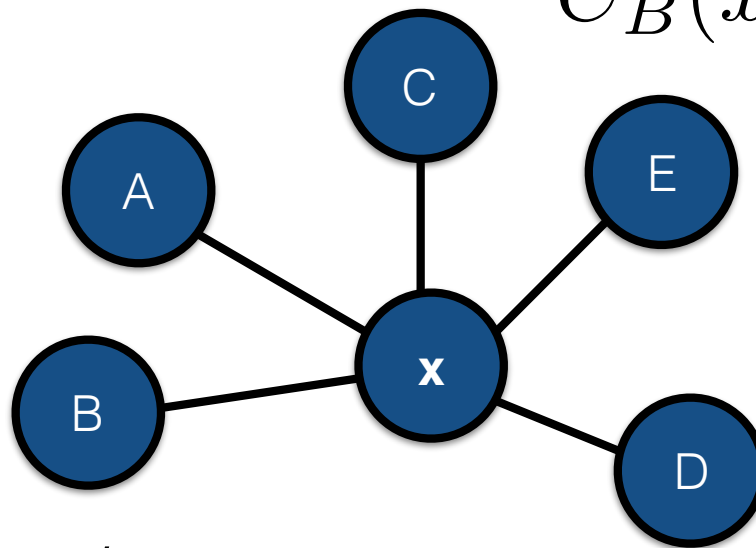


Betweenness centrality: example

Betweenness centrality of a node i equals the the number of shortest path (between any couple of node) that passes through this nodes divided by the number of shortest path existing between these two nodes.

$$C_B(x) = \sum_{j < k} g_{jk}(x) / g_{jk}$$

$$C_B(x) = 10$$

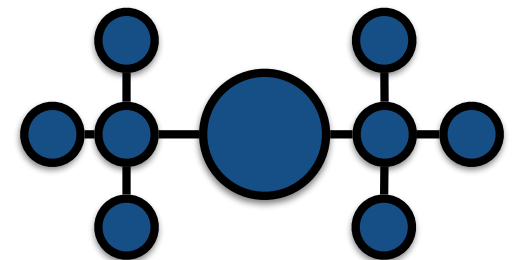


A->B : 1/1

A->C : 1/1 B->C : 1/1

A->D : 1/1 B->D : 1/1 C->D : 1/1

A->E : 1/1 B->E : 1/1 C->E : 1/1 D->E : 1/1

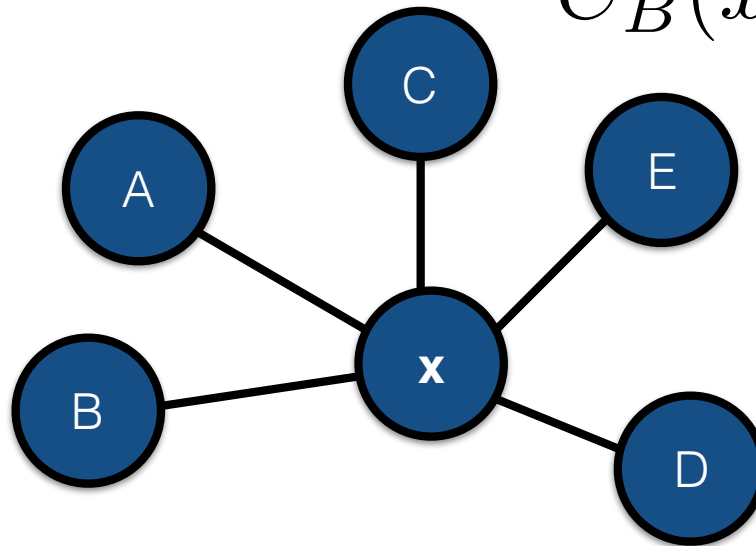


Betweenness centrality: example

Betweenness centrality of a node i equals the the number of shortest path (between any couple of node) that passes through this nodes divided by the number of shortest path existing between these two nodes.

$$C_B(x) = \sum_{j < k} g_{jk}(x) / g_{jk}$$

$$C_B(A)=0$$

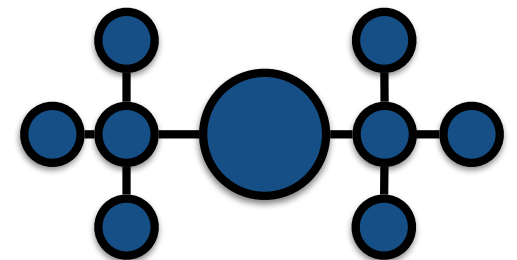


x->B : 0/1

x->C : 0/1 B->C : 0/1

x->D : 0/1 B->D : 0/1 C->D : 0/1

x->E : 0/1 B->E : 0/1 C->E : 0/1 D->E : 0/1

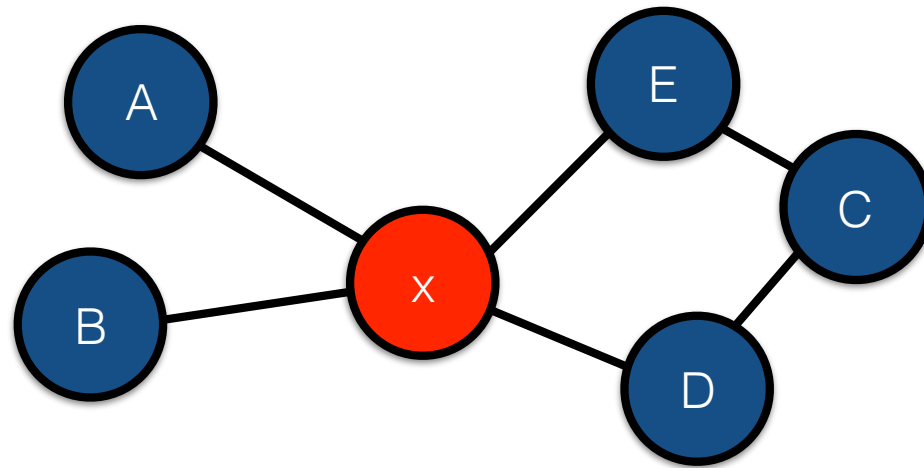


Betweenness centrality: example

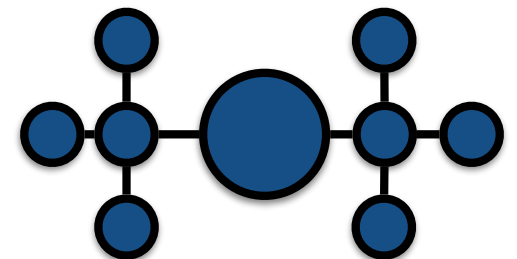
Betweenness centrality of a node i equals the the number of shortest path (between any couple of node) that passes through this nodes divided by the number of shortest path existing between these two nodes.

$$C_B(x) = 7.5$$

$$C_B(x) = \sum_{j < k} g_{jk}(x) / g_{jk}$$



A → B : 1/1 D → E : 1/2
A → C : 2/2 B → C : 2/2
A → D : 1/1 B → D : 1/1 C → D : 0/1
A → E : 1/1 B → E : 1/1 C → E : 0/1

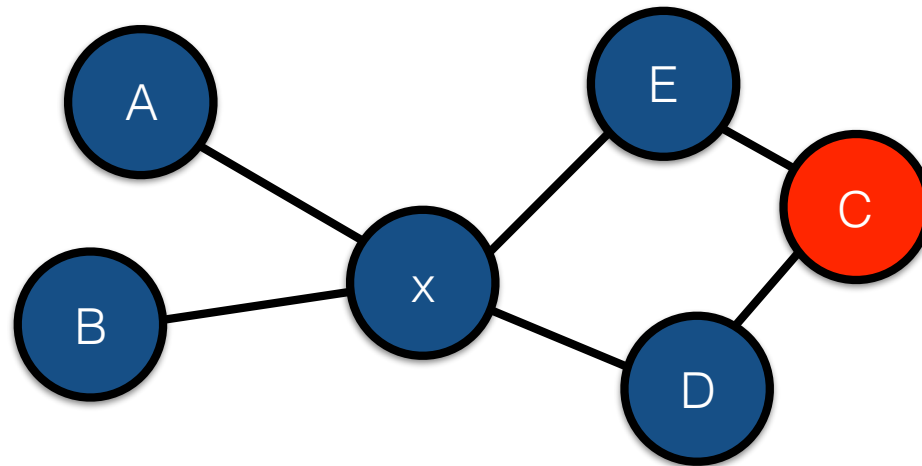


Betweenness centrality: example

Betweenness centrality of a node i equals the the number of shortest path (between any couple of node) that passes through this nodes divided by the number of shortest path existing between these two nodes.

$$C_B(x) = \sum_{j < k} g_{jk}(x) / g_{jk}$$

$$C_B(c) = 0.5$$

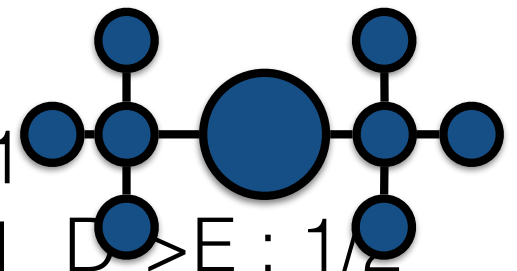


A->B : 0/1

A->X : 0/1 B->X : 0/1

A->D : 0/1 B->D : 0/1 X->D : 0/1

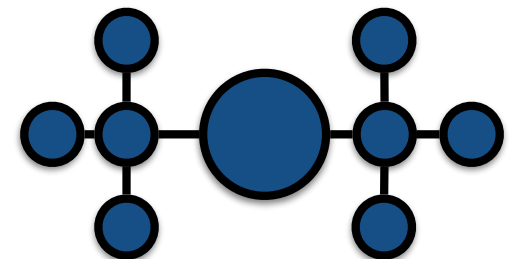
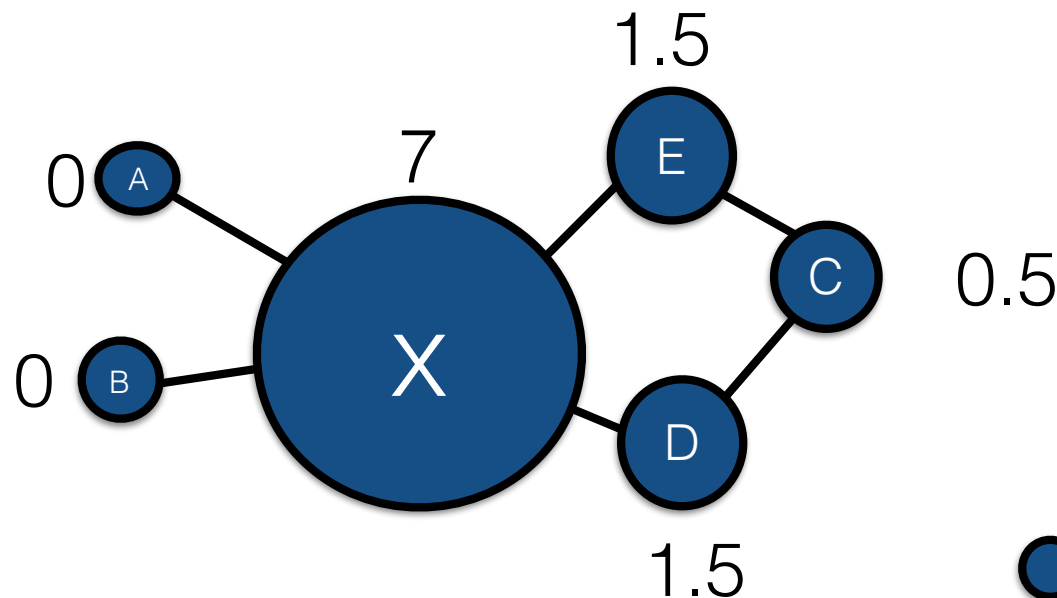
A->E : 0/1 B->E : 0/1 X->E : 0/1 D->E : 1/2



Betweenness centrality: example

Betweenness centrality of a node i equals the the number of shortest path (between any couple of node) that passes through this nodes divided by the number of shortest path existing between these two nodes.

$$C_B(x) = \sum_{j < k} g_{jk}(x) / g_{jk}$$





2nd

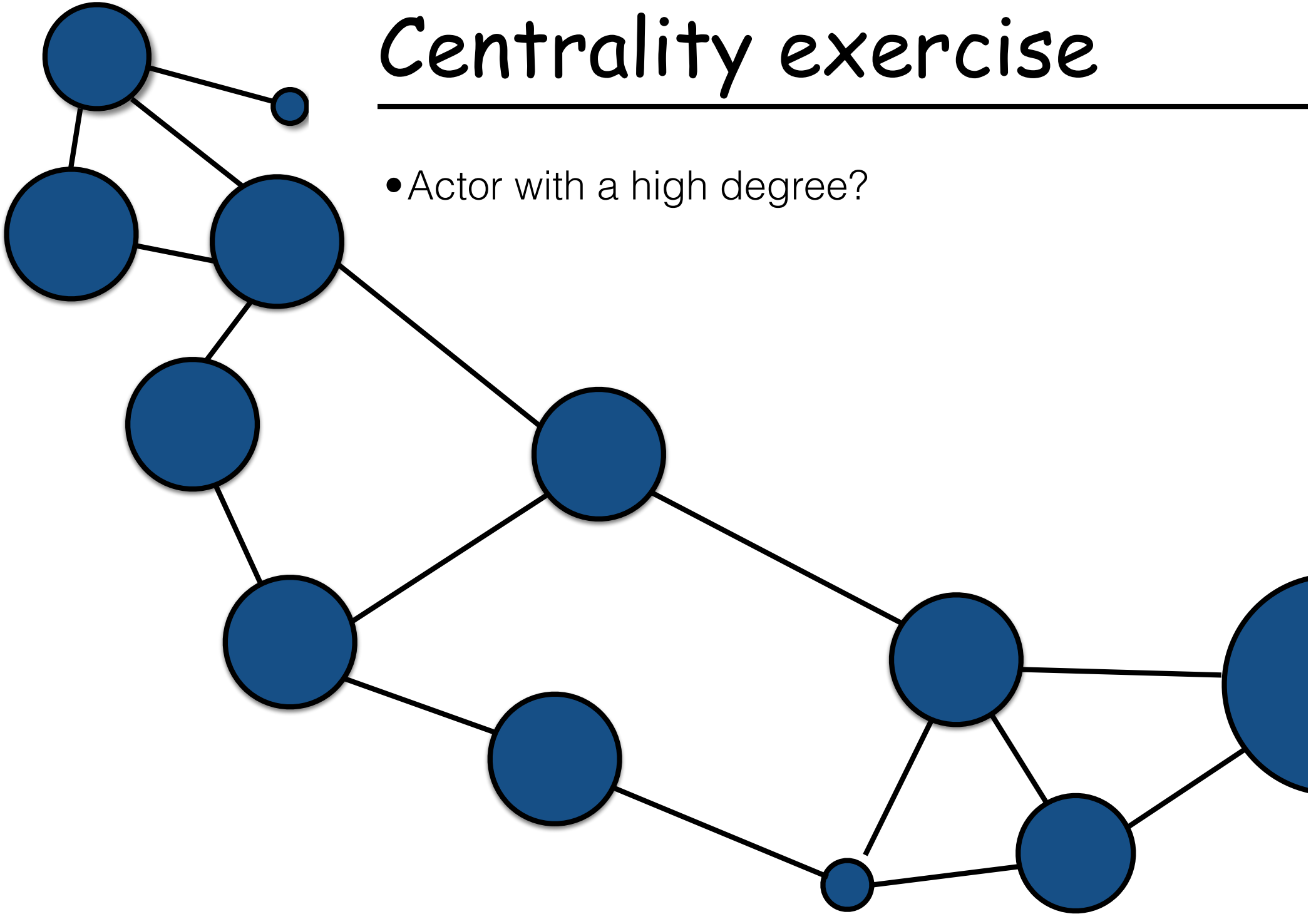
Education Sciences Po Grenoble (Institute of Political Studies)

Send Léa InMail

500+
connections

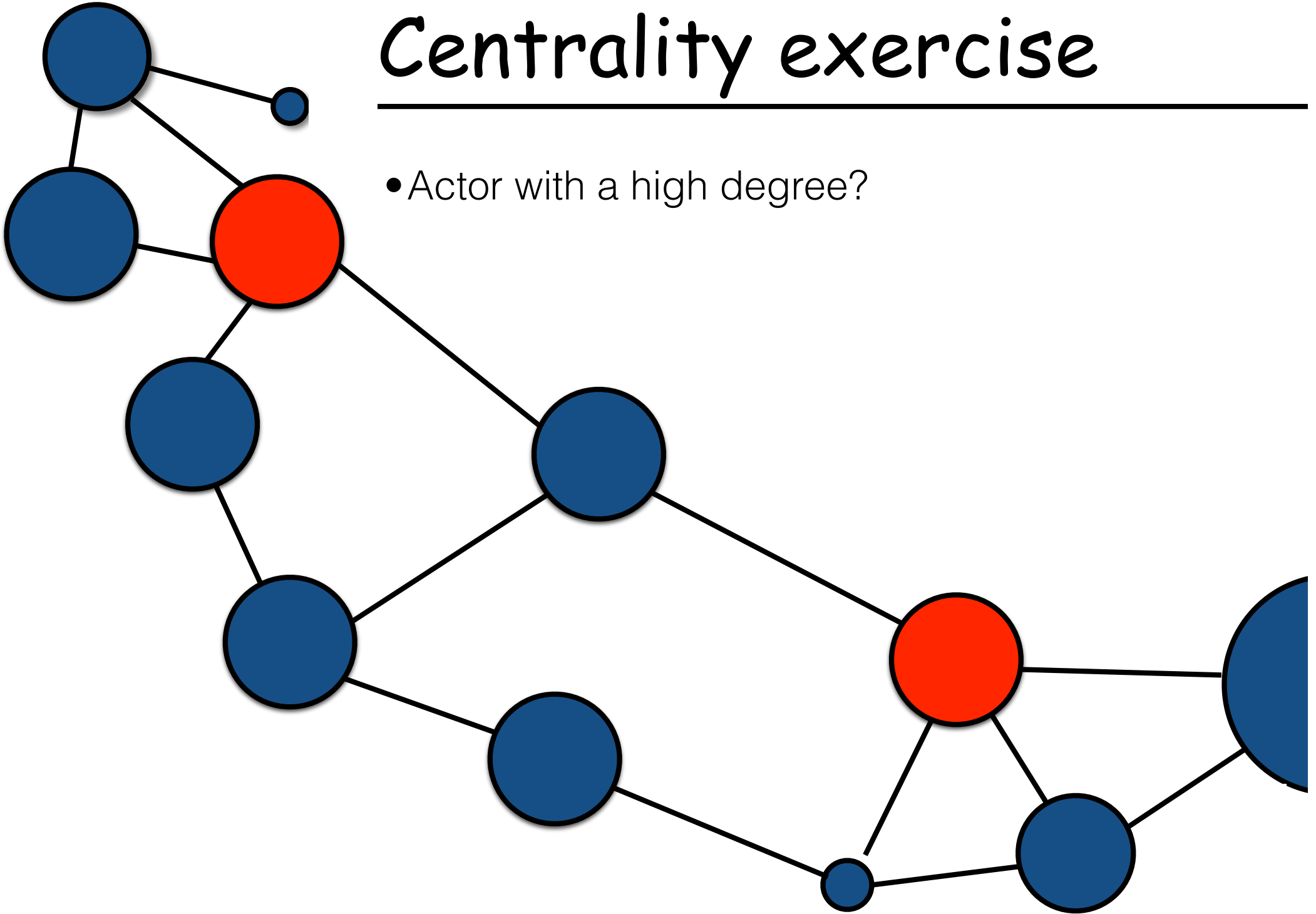
Centrality exercise

- Actor with a high degree?



Centrality exercise

- Actor with a high degree?

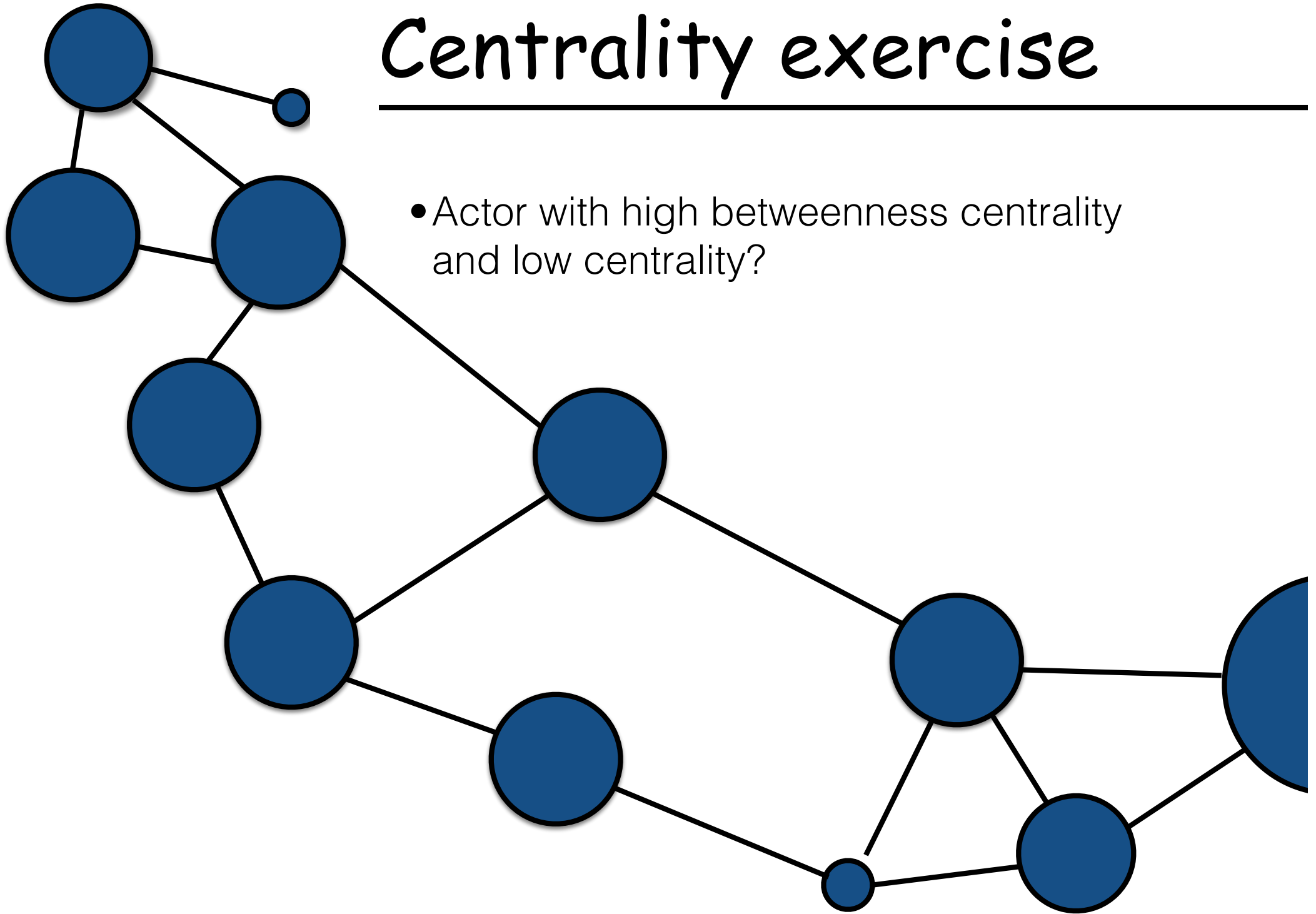


- Actor with high betweenness centrality?

- # Centrality exercise
-
- Actor with high betweenness centrality?
-
- ```
graph TD; N1(()) --- N2(()); N1 --- N3(()); N2 --- N4(()); N3 --- N4; N4 --- N5(()); N4 --- N6(()); N5 --- N7(()); N6 --- N8(()); N7 --- N9(()); N8 --- N9; N9 --- N10((()))
```

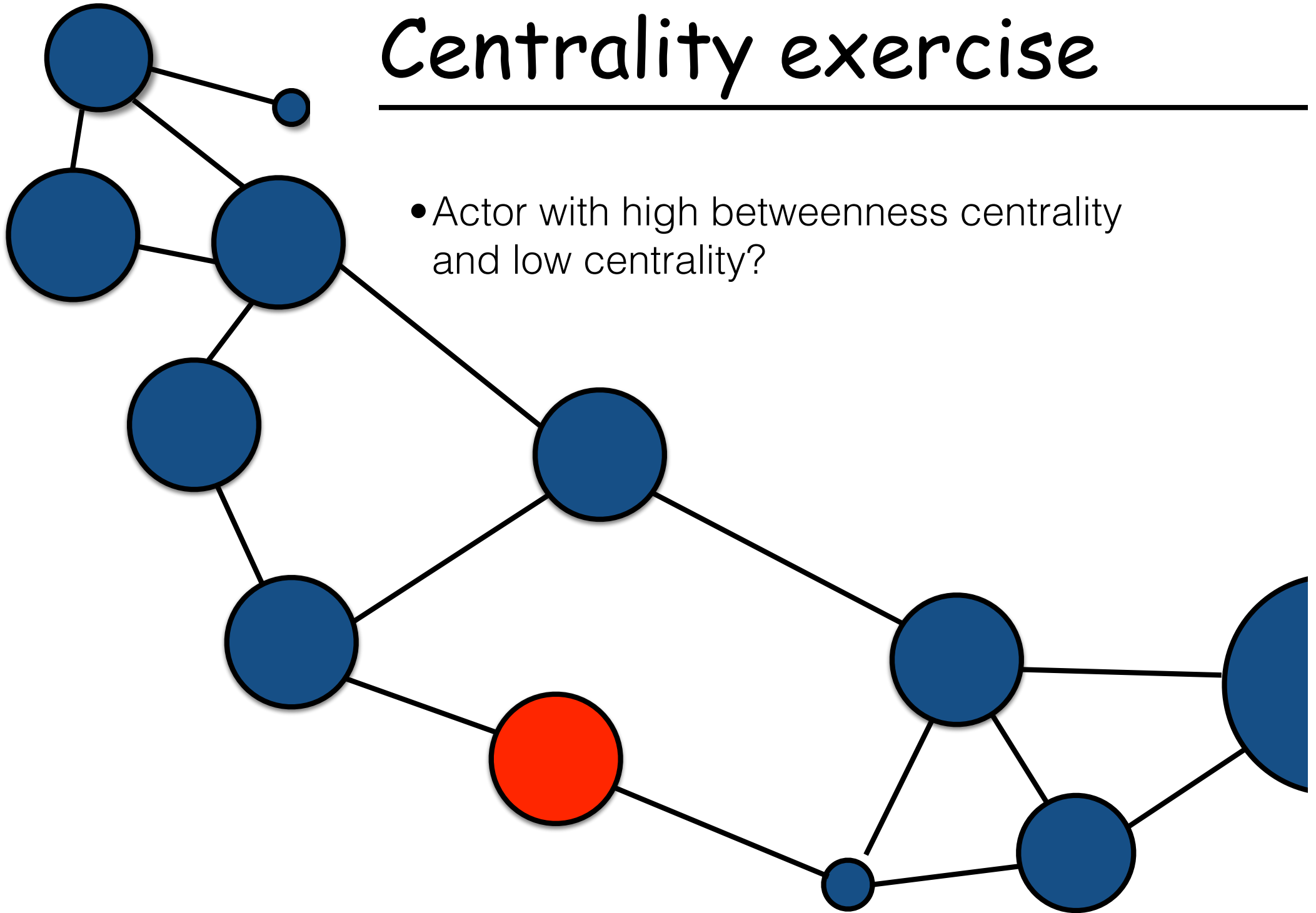
- Actor with high betweenness centrality?

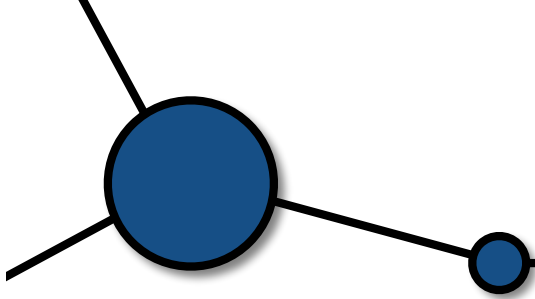
- # Centrality exercise
- Actor with high betweenness centrality?
- 
- The graph consists of 12 nodes and 15 edges. Two nodes are highlighted in red, indicating high betweenness centrality. The graph shows a central path of nodes connecting different clusters.
- ```
graph LR; N1(( )) --- N2(( )); N1 --- N3(( )); N2 --- N4(( )); N3 --- N5(( )); N4 --- N6(( )); N5 --- N7(( )); N6 --- N8(( )); N7 --- N9(( )); N8 --- N10(( )); N9 --- N11(( )); N10 --- N12(( )); N11 --- N13(( )); N12 --- N14(( )); N13 --- N15(( )); N14 --- N16(( )); N15 --- N17(( )); N16 --- N18(( )); N17 --- N19(( )); N18 --- N20(( )); N19 --- N21(( )); N20 --- N22(( )); N21 --- N23(( )); N22 --- N24(( )); N23 --- N25(( )); N24 --- N26(( )); N25 --- N27(( )); N26 --- N28(( )); N27 --- N29(( )); N28 --- N30(( )); N29 --- N31(( )); N30 --- N32(( )); N31 --- N33(( )); N32 --- N34(( )); N33 --- N35(( )); N34 --- N36(( )); N35 --- N37(( )); N36 --- N38(( )); N37 --- N39(( )); N38 --- N40(( )); N39 --- N41(( )); N40 --- N42(( )); N41 --- N43(( )); N42 --- N44(( )); N43 --- N45(( )); N44 --- N46(( )); N45 --- N47(( )); N46 --- N48(( )); N47 --- N49(( )); N48 --- N50(( )); N49 --- N51(( )); N50 --- N52(( )); N51 --- N53(( )); N52 --- N54(( )); N53 --- N55(( )); N54 --- N56(( )); N55 --- N57(( )); N56 --- N58(( )); N57 --- N59(( )); N58 --- N60(( )); N59 --- N61(( )); N60 --- N62(( )); N61 --- N63(( )); N62 --- N64(( )); N63 --- N65(( )); N64 --- N66(( )); N65 --- N67(( )); N66 --- N68(( )); N67 --- N69(( )); N68 --- N70(( )); N69 --- N71(( )); N70 --- N72(( )); N71 --- N73(( )); N72 --- N74(( )); N73 --- N75(( )); N74 --- N76(( )); N75 --- N77(( )); N76 --- N78(( )); N77 --- N79(( )); N78 --- N80(( )); N79 --- N81(( )); N80 --- N82(( )); N81 --- N83(( )); N82 --- N84(( )); N83 --- N85(( )); N84 --- N86(( )); N85 --- N87(( )); N86 --- N88(( )); N87 --- N89(( )); N88 --- N90(( )); N89 --- N91(( )); N90 --- N92(( )); N91 --- N93(( )); N92 --- N94(( )); N93 --- N95(( )); N94 --- N96(( )); N95 --- N97(( )); N96 --- N98(( )); N97 --- N99(( )); N98 --- N100(( )); N99 --- N101(( )); N100 --- N102(( )); N101 --- N103(( )); N102 --- N104(( )); N103 --- N105(( )); N104 --- N106(( )); N105 --- N107(( )); N106 --- N108(( )); N107 --- N109(( )); N108 --- N110(( )); N109 --- N111(( )); N110 --- N112(( )); N111 --- N113(( )); N112 --- N114(( )); N113 --- N115(( )); N114 --- N116(( )); N115 --- N117(( )); N116 --- N118(( )); N117 --- N119(( )); N118 --- N120(( )); N119 --- N121(( )); N120 --- N122(( )); N121 --- N123(( )); N122 --- N124(( )); N123 --- N125(( )); N124 --- N126(( )); N125 --- N127(( )); N126 --- N128(( )); N127 --- N129(( )); N128 --- N130(( )); N129 --- N131(( )); N130 --- N132(( )); N131 --- N133(( )); N132 --- N134(( )); N133 --- N135(( )); N134 --- N136(( )); N135 --- N137(( )); N136 --- N138(( )); N137 --- N139(( )); N138 --- N140(( )); N139 --- N141(( )); N140 --- N142(( )); N141 --- N143(( )); N142 --- N144(( )); N143 --- N145(( )); N144 --- N146(( )); N145 --- N147(( )); N146 --- N148(( )); N147 --- N149(( )); N148 --- N150(( )); N149 --- N151(( )); N150 --- N152(( )); N151 --- N153(( )); N152 --- N154(( )); N153 --- N155(( )); N154 --- N156(( )); N155 --- N157(( )); N156 --- N158(( )); N157 --- N159(( )); N158 --- N160(( )); N159 --- N161(( )); N160 --- N162(( )); N161 --- N163(( )); N162 --- N164(( )); N163 --- N165(( )); N164 --- N166(( )); N165 --- N167(( )); N166 --- N168(( )); N167 --- N169(( )); N168 --- N170(( )); N169 --- N171(( )); N170 --- N172(( )); N171 --- N173(( )); N172 --- N174(( )); N173 --- N175(( )); N174 --- N176(( )); N175 --- N177(( )); N176 --- N178(( )); N177 --- N179(( )); N178 --- N180(( )); N179 --- N181(( )); N180 --- N182(( )); N181 --- N183(( )); N182 --- N184(( )); N183 --- N185(( )); N184 --- N186(( )); N185 --- N187(( )); N186 --- N188(( )); N187 --- N189(( )); N188 --- N190(( )); N189 --- N191(( )); N190 --- N192(( )); N191 --- N193(( )); N192 --- N194(( )); N193 --- N195(( )); N194 --- N196(( )); N195 --- N197(( )); N196 --- N198(( )); N197 --- N199(( )); N198 --- N200(( )); N199 --- N201(( )); N200 --- N202(( )); N201 --- N203(( )); N202 --- N204(( )); N203 --- N205(( )); N204 --- N206(( )); N205 --- N207(( )); N206 --- N208(( )); N207 --- N209(( )); N208 --- N210(( )); N209 --- N211(( )); N210 --- N212(( )); N211 --- N213(( )); N212 --- N214(( )); N213 --- N215(( )); N214 --- N216(( )); N215 --- N217(( )); N216 --- N218(( )); N217 --- N219(( )); N218 --- N220(( )); N219 --- N221(( )); N220 --- N222(( )); N221 --- N223(( )); N222 --- N224(( )); N223 --- N225(( )); N224 --- N226(( )); N225 --- N227(( )); N226 --- N228(( )); N227 --- N229(( )); N228 --- N230(( )); N229 --- N231(( )); N230 --- N232(( )); N231 --- N233(( )); N232 --- N234(( )); N233 --- N235(( )); N234 --- N236(( )); N235 --- N237(( )); N236 --- N238(( )); N237 --- N239(( )); N238 --- N240(( )); N239 --- N241(( )); N240 --- N242(( )); N241 --- N243(( )); N242 --- N244(( )); N243 --- N245(( )); N244 --- N246(( )); N245 --- N247(( )); N246 --- N248(( )); N247 --- N249(( )); N248 --- N250(( )); N249 --- N251(( )); N250 --- N252(( )); N251 --- N253(( )); N252 --- N254(( )); N253 --- N255(( )); N254 --- N256(( )); N255 --- N257(( )); N256 --- N258(( )); N257 --- N259(( )); N258 --- N260(( )); N259 --- N261(( )); N260 --- N262(( )); N261 --- N263(( )); N262 --- N264(( )); N263 --- N265(( )); N264 --- N266(( )); N265 --- N267(( )); N266 --- N268(( )); N267 --- N269(( )); N268 --- N270(( )); N269 --- N271(( )); N270 --- N272(( )); N271 --- N273(( )); N272 --- N274(( )); N273 --- N275(( )); N274 --- N276(( )); N275 --- N277(( )); N276 --- N278(( )); N277 --- N279(( )); N278 --- N280(( )); N279 --- N281(( )); N280 --- N282(( )); N281 --- N283(( )); N282 --- N284(( )); N283 --- N285(( )); N284 --- N286(( )); N285 --- N287(( )); N286 --- N288(( )); N287 --- N289(( )); N288 --- N290(( )); N289 --- N291(( )); N290 --- N292(( )); N291 --- N293(( )); N292 --- N294(( )); N293 --- N295(( )); N294 --- N296(( )); N295 --- N297(( )); N296 --- N298(( )); N297 --- N299(( )); N298 --- N300(( )); N299 --- N301(( )); N300 --- N302(( )); N301 --- N303(( )); N302 --- N304(( )); N303 --- N305(( )); N304 --- N306(( )); N305 --- N307(( )); N306 --- N308(( )); N307 --- N309(( )); N308 --- N310(( )); N309 --- N311(( )); N310 --- N312(( )); N311 --- N313(( )); N312 --- N314(( )); N313 --- N315(( )); N314 --- N316(( )); N315 --- N317(( )); N316 --- N318(( )); N317 --- N319(( )); N318 --- N320(( )); N319 --- N321(( )); N320 --- N322(( )); N321 --- N323(( )); N322 --- N324(( )); N323 --- N325(( )); N324 --- N326(( )); N325 --- N327(( )); N326 --- N328(( )); N327 --- N329(( )); N328 --- N330(( )); N329 --- N331(( )); N330 --- N332(( )); N331 --- N333(( )); N332 --- N334(( )); N333 --- N335(( )); N334 --- N336(( )); N335 --- N337(( )); N336 --- N338(( )); N337 --- N339(( )); N338 --- N340(( )); N339 --- N341(( )); N340 --- N342(( )); N341 --- N343(( )); N342 --- N344(( )); N343 --- N345(( )); N344 --- N346(( )); N345 --- N347(( )); N346 --- N348(( )); N347 --- N349(( )); N348 --- N350(( )); N349 --- N351(( )); N350 --- N352(( )); N351 --- N353(( )); N352 --- N354(( )); N353 --- N355(( )); N354 --- N356(( )); N355 --- N357(( )); N356 --- N358(( )); N357 --- N359(( )); N358 --- N360(( )); N359 --- N361(( )); N360 --- N362(( )); N361 --- N363(( )); N362 --- N364(( )); N363 --- N365(( )); N364 --- N366(( )); N365 --- N367(( )); N366 --- N368(( )); N367 --- N369(( )); N368 --- N370(( )); N369 --- N371(( )); N370 --- N372(( )); N371 --- N373(( )); N372 --- N374(( )); N373 --- N375(( )); N374 --- N376(( )); N375 --- N377(( )); N376 --- N378(( )); N377 --- N379(( )); N378 --- N380(( )); N379 --- N381(( )); N380 --- N382((
```



Centrality exercise

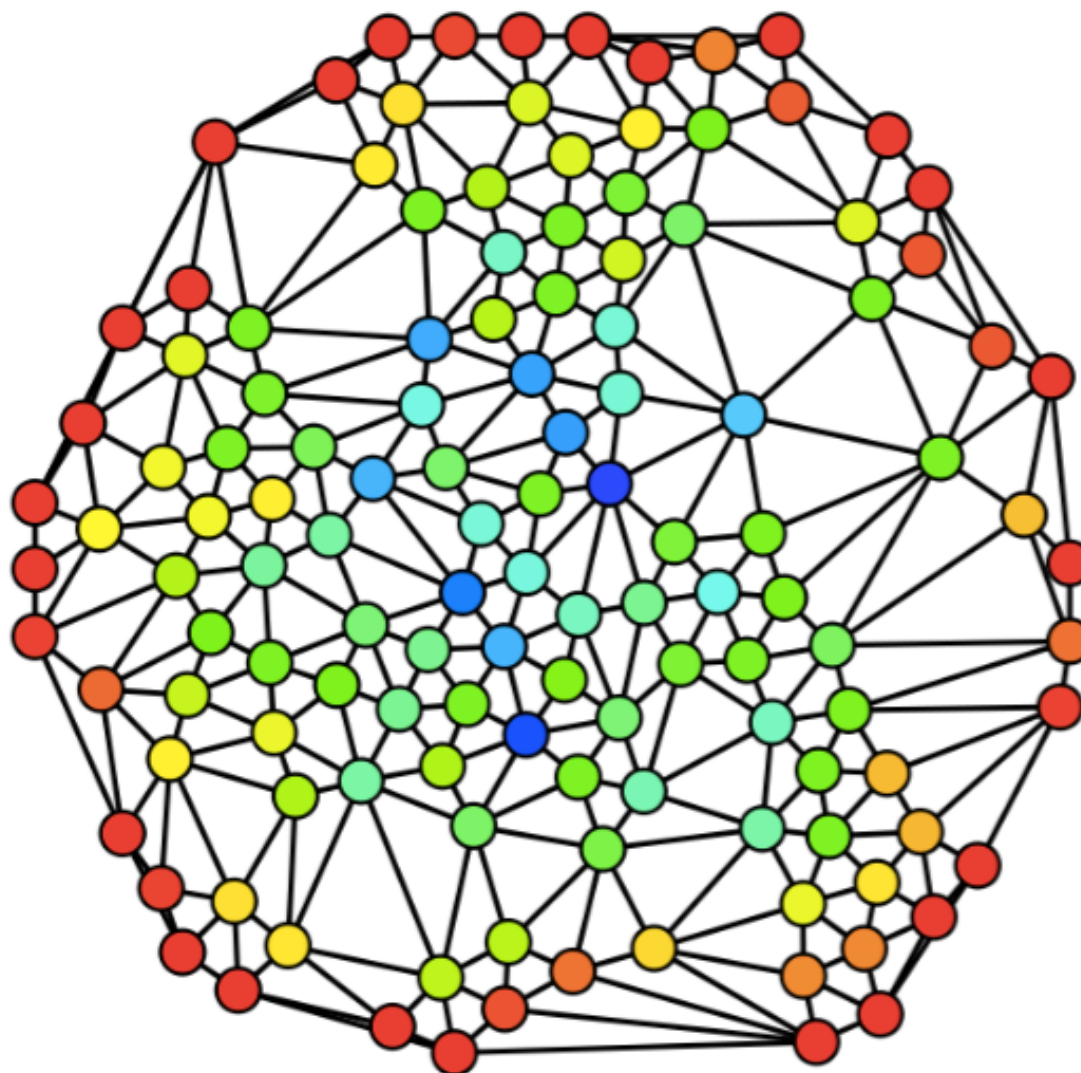
- Actor with high betweenness centrality and low centrality?

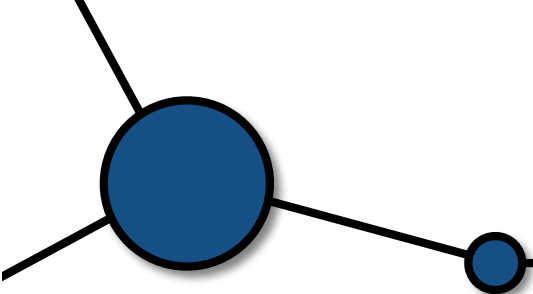




Centrality exercise

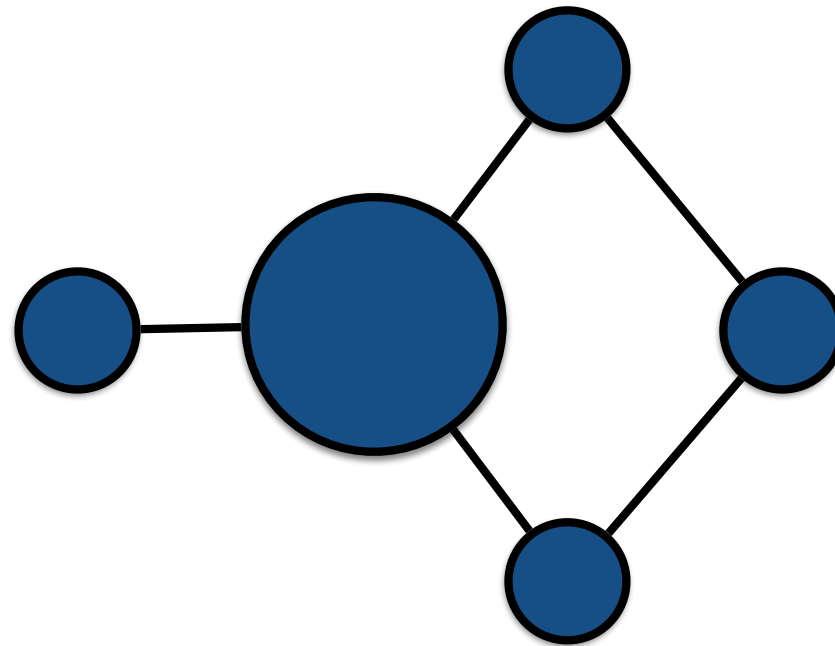
Actor with high betweenness centrality?

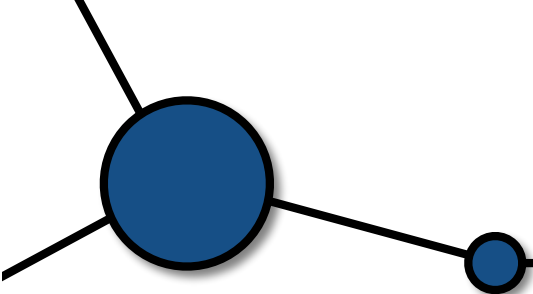




Closeness centrality

- Hypothesis : an important nodes is a node that is close to every other nodes





Closeness centrality

- Definition : closeness centrality of node i measures the degree to which this node is close to other nodes.

minimum total distance

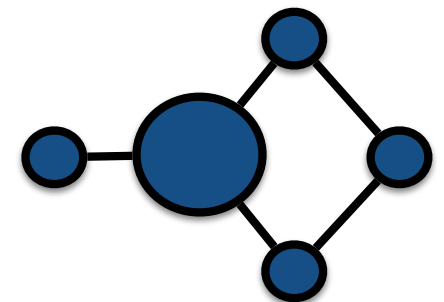
by convention

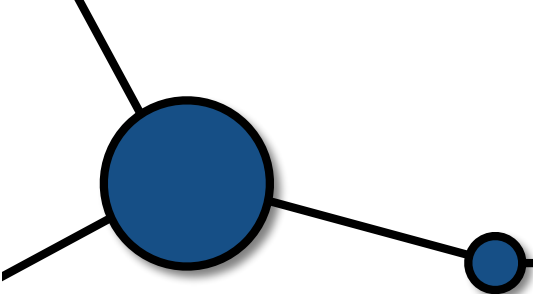
$$\frac{1}{\infty} = 0$$

$$C_i(x) = \sum_{j=1}^N \frac{(N-1)}{d(x, j)}$$

distance to other nodes

for all nodes



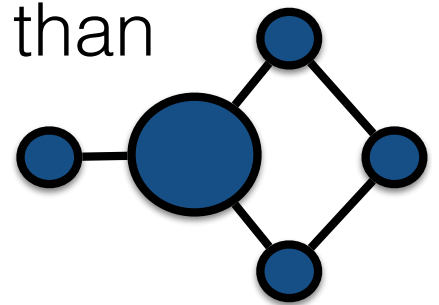


Closeness centrality

- Definition : closeness centrality of node i measures the degree to which this node is close to other nodes.

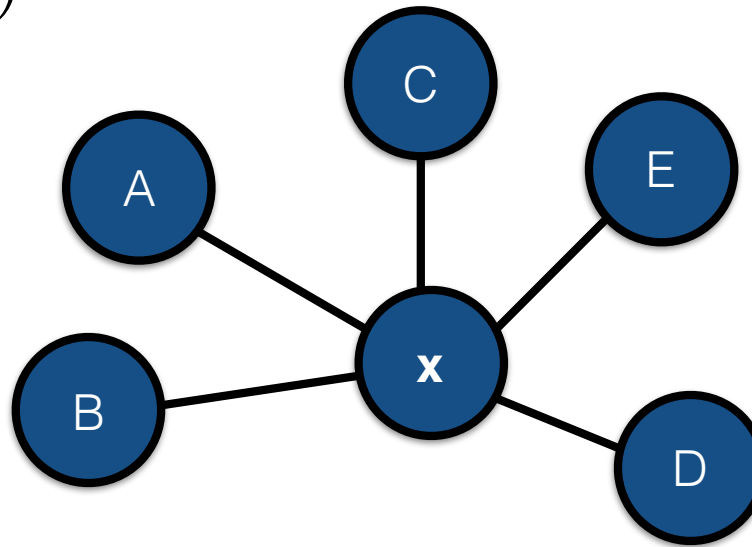
$$C_i(x) = \sum_{j=1}^N \frac{(N - 1)}{d(x, j)}$$

- A node with high closeness centrality is closer to other nodes
- This node can reach any other node quicker than other nodes



Closeness centrality

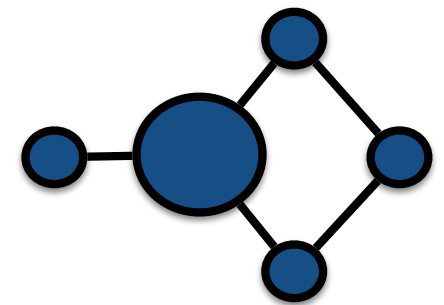
$$C_i(x) = \sum_{j=1}^N \frac{(N-1)}{d(x, j)}$$



$$C_P(x) = (6-1) * (1/1 + 1/1 + 1/1 + 1/1 + 1/1)$$

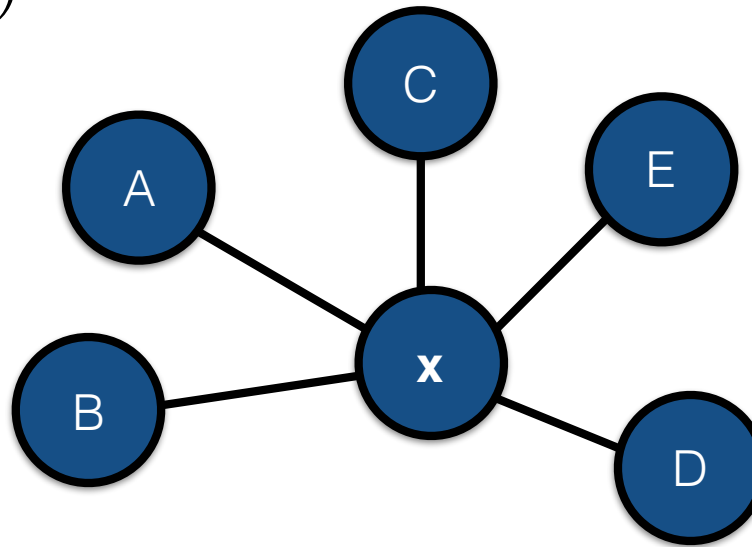
$$C_P(x) = 5/5$$

$$C_P(x) = 1$$



Closeness centrality

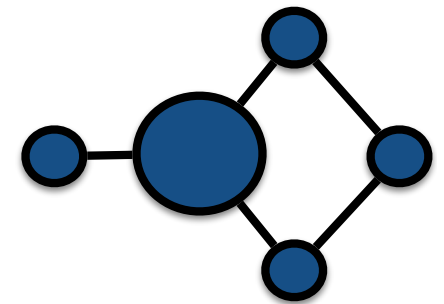
$$C_i(x) = \sum_{j=1}^N \frac{(N-1)}{d(x, j)}$$



$$C_P(A) = (6-1) * (1/1 + 1/2 + 1/2 + 1/2 + 1/2)$$

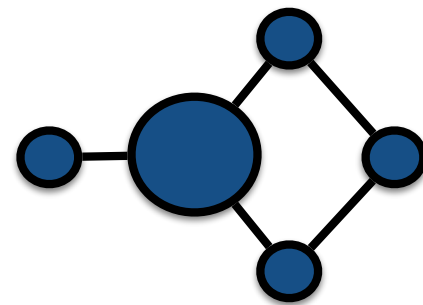
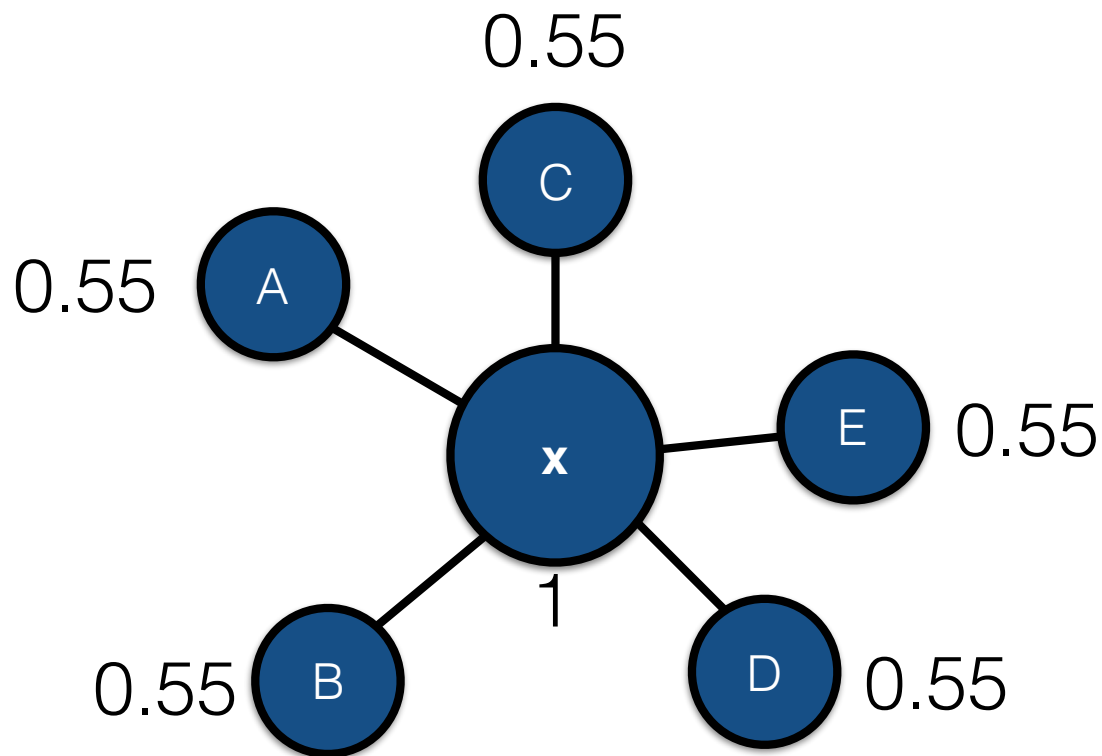
$$C_P(A) = 5/9$$

$$C_P(A) = 0.55$$



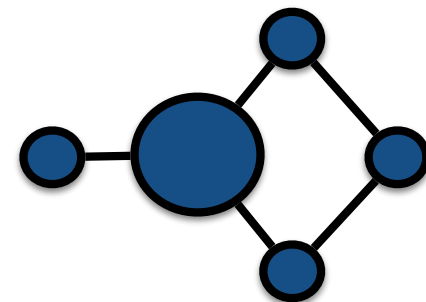
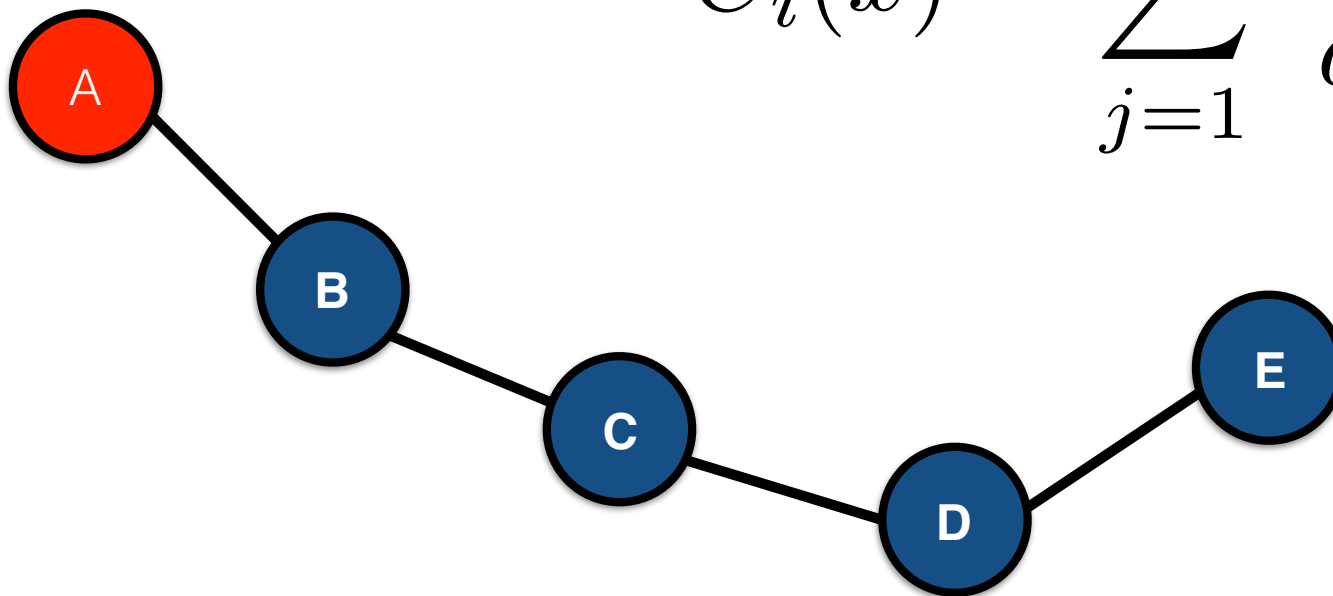
Closeness centrality

$$C_i(x) = \sum_{j=1}^N \frac{(N-1)}{d(x, j)}$$



Closeness centrality

$$C_i(x) = \sum_{j=1}^N \frac{(N-1)}{d(x, j)}$$

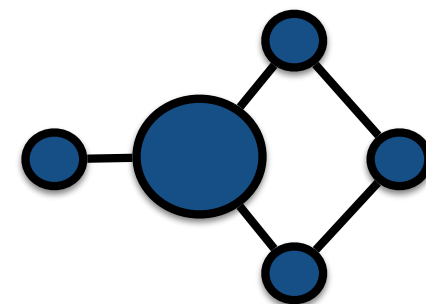
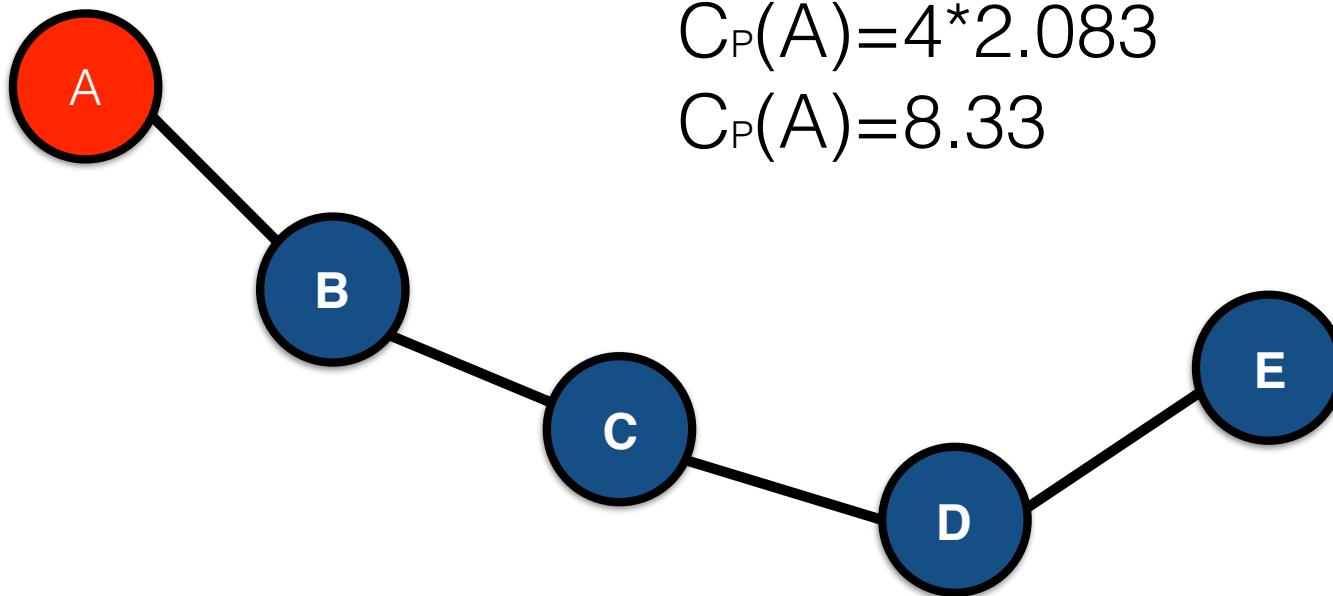


Closeness centrality

$$C_P(A) = (5-1) * (1/1 + 1/2 + 1/3 + 1/4)$$

$$C_P(A) = 4 * 2.083$$

$$C_P(A) = 8.33$$



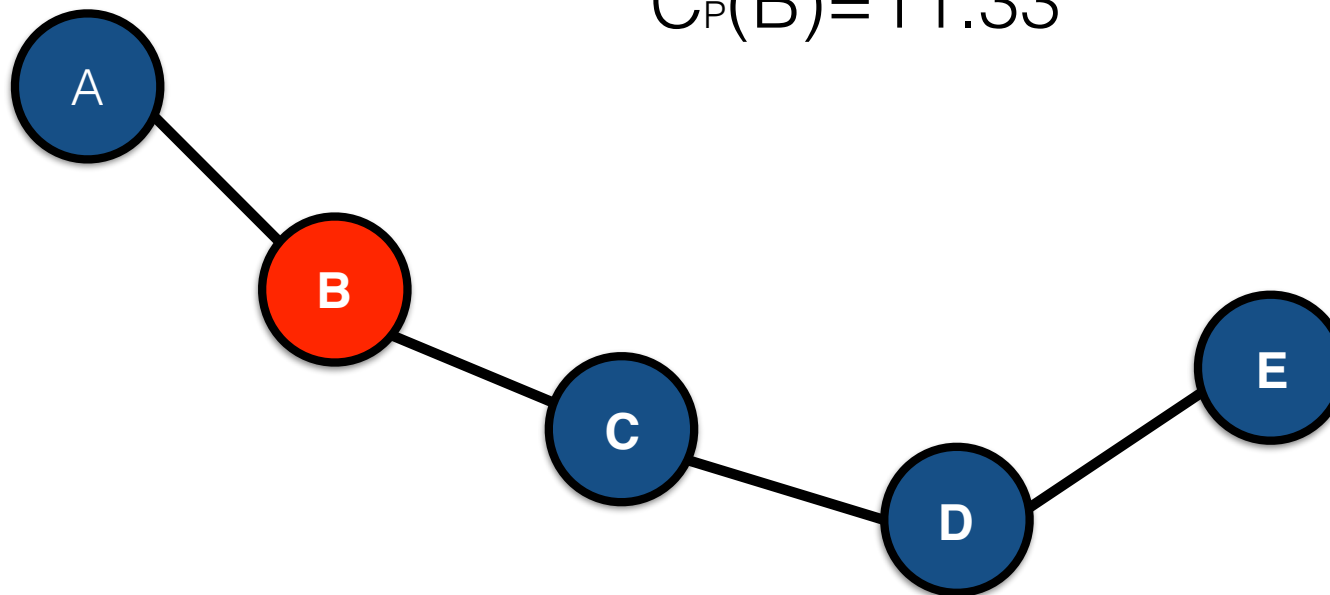
Closeness centrality

$$C_P(B) = (5-1) * (1/1 + 1/1 + 1/2 + 1/3)$$

$$C_P(B) = 4 * 2.83$$

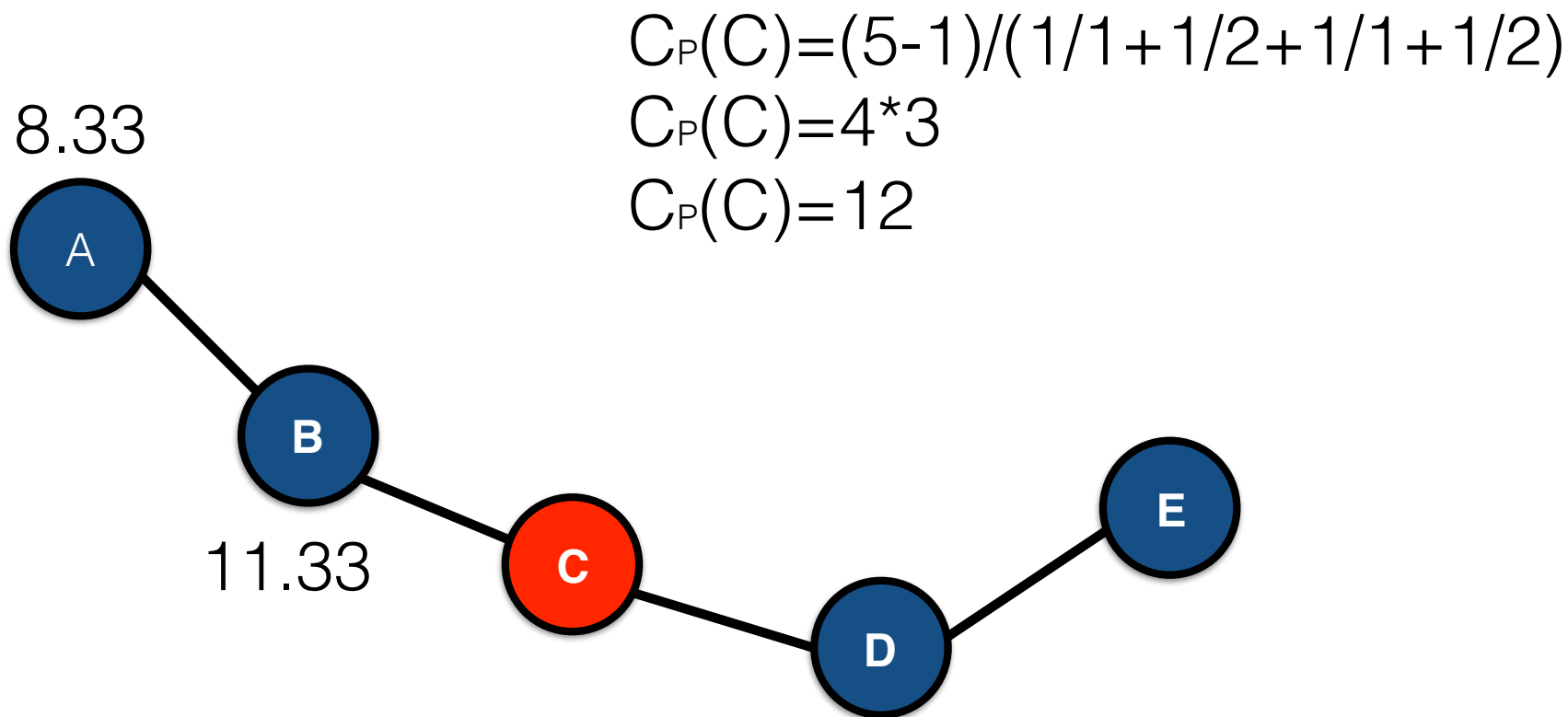
$$C_P(B) = 11.33$$

8.33



$$C_i(x) = \sum_{j=1}^N \frac{(N-1)}{d(x, j)}$$

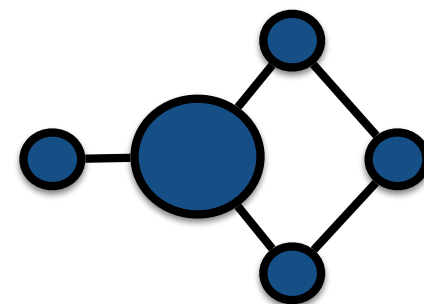
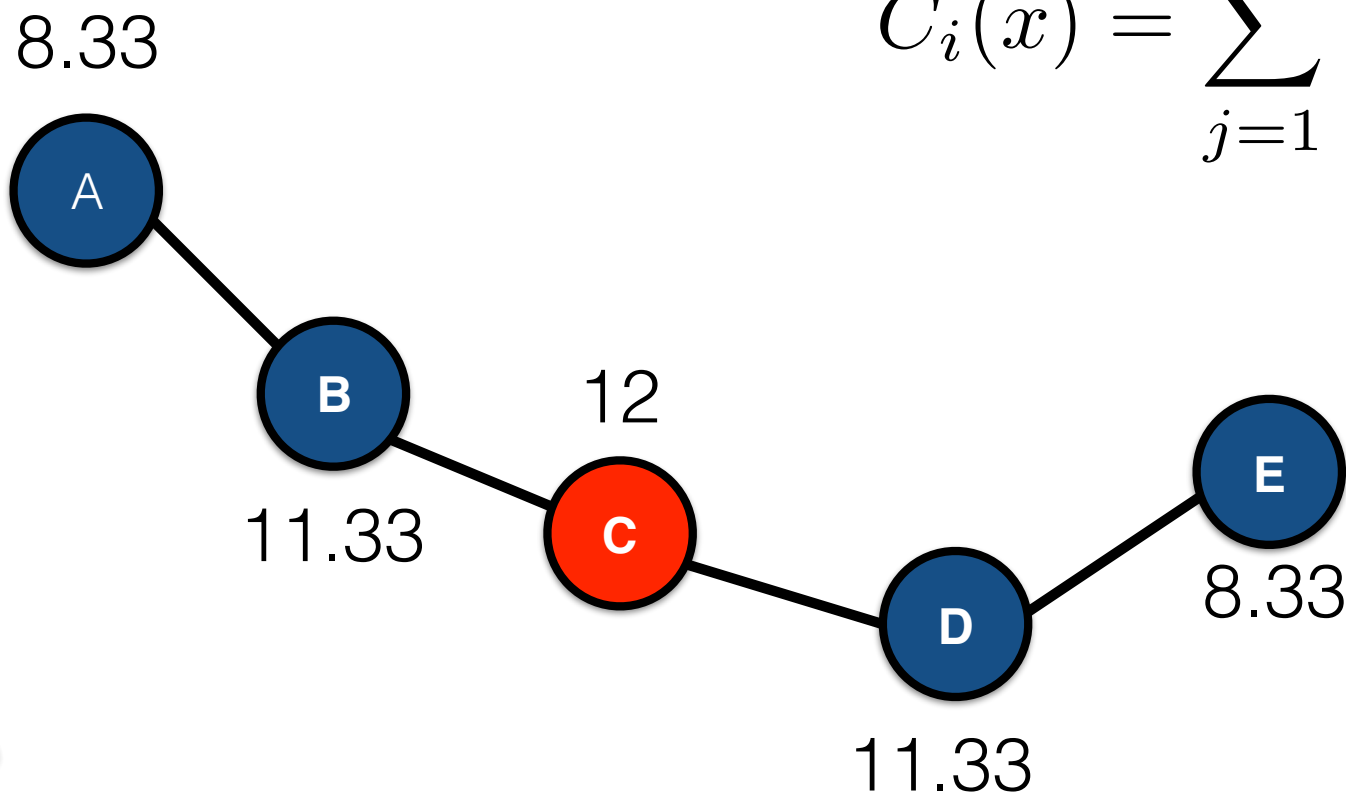
Closeness centrality



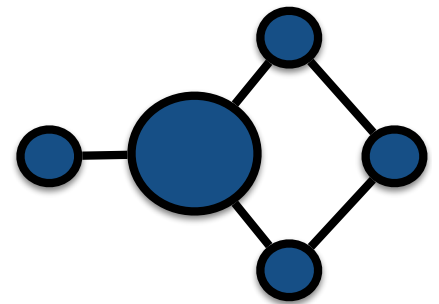
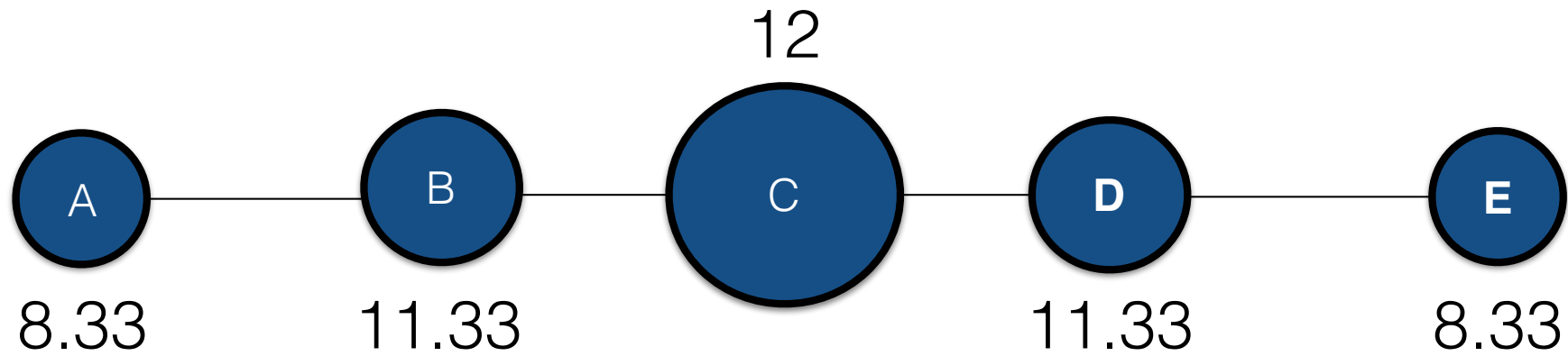
$$C_i(x) = \sum_{j=1}^N \frac{(N-1)}{d(x, j)}$$

Closeness centrality

$$C_i(x) = \sum_{j=1}^N \frac{(N-1)}{d(x, j)}$$



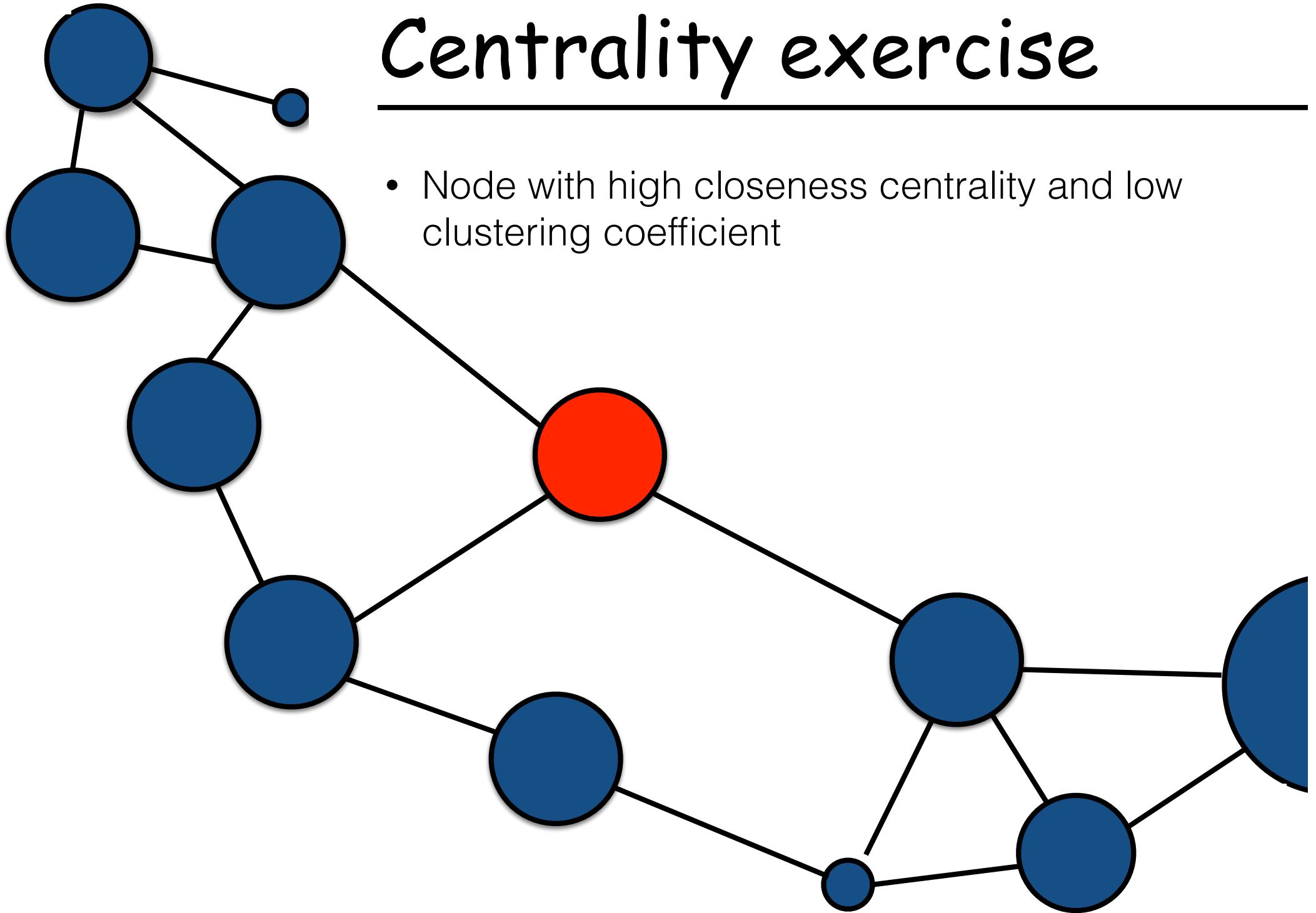
Closeness centrality



- Node with high closeness centrality and low clustering coefficient
-
- The graph consists of 10 nodes and 12 edges. The nodes are arranged in a way that shows a path of nodes on the left, a central node connected to two others, and a cluster of three nodes on the right. The node in the center, which is connected to the most other nodes (4), is highlighted as having high closeness centrality and low clustering coefficient.

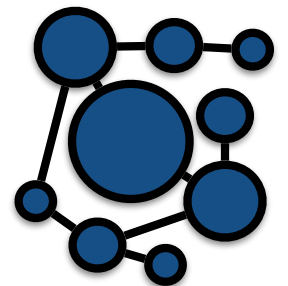
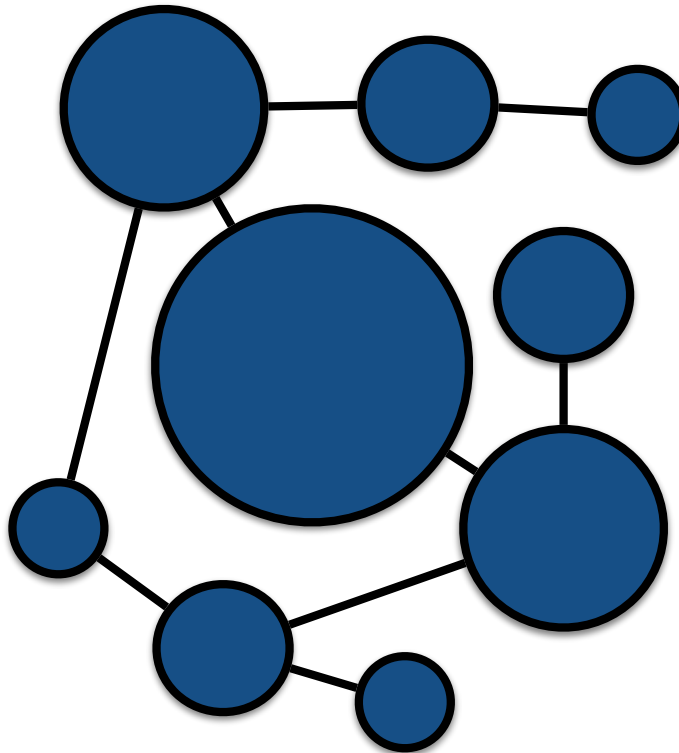
Centrality exercise

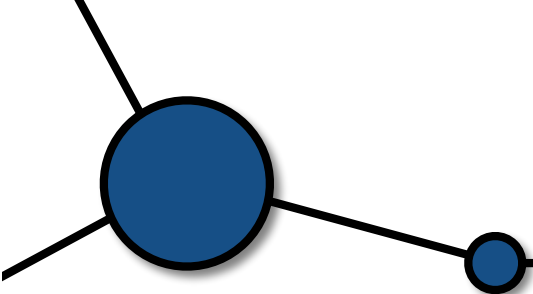
- Node with high closeness centrality and low clustering coefficient



Eigenvalues centrality

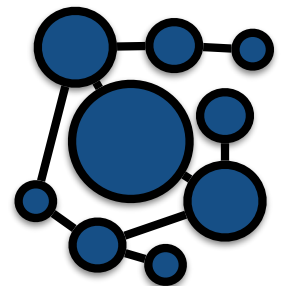
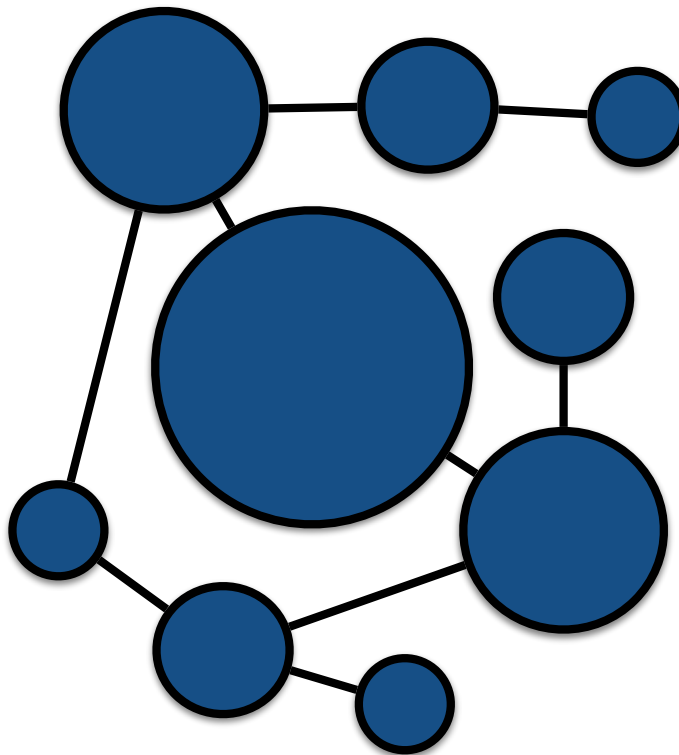
- Hypothesis : A node is central if it is connected to central nodes.



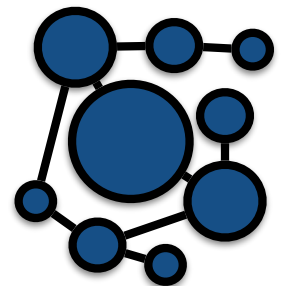
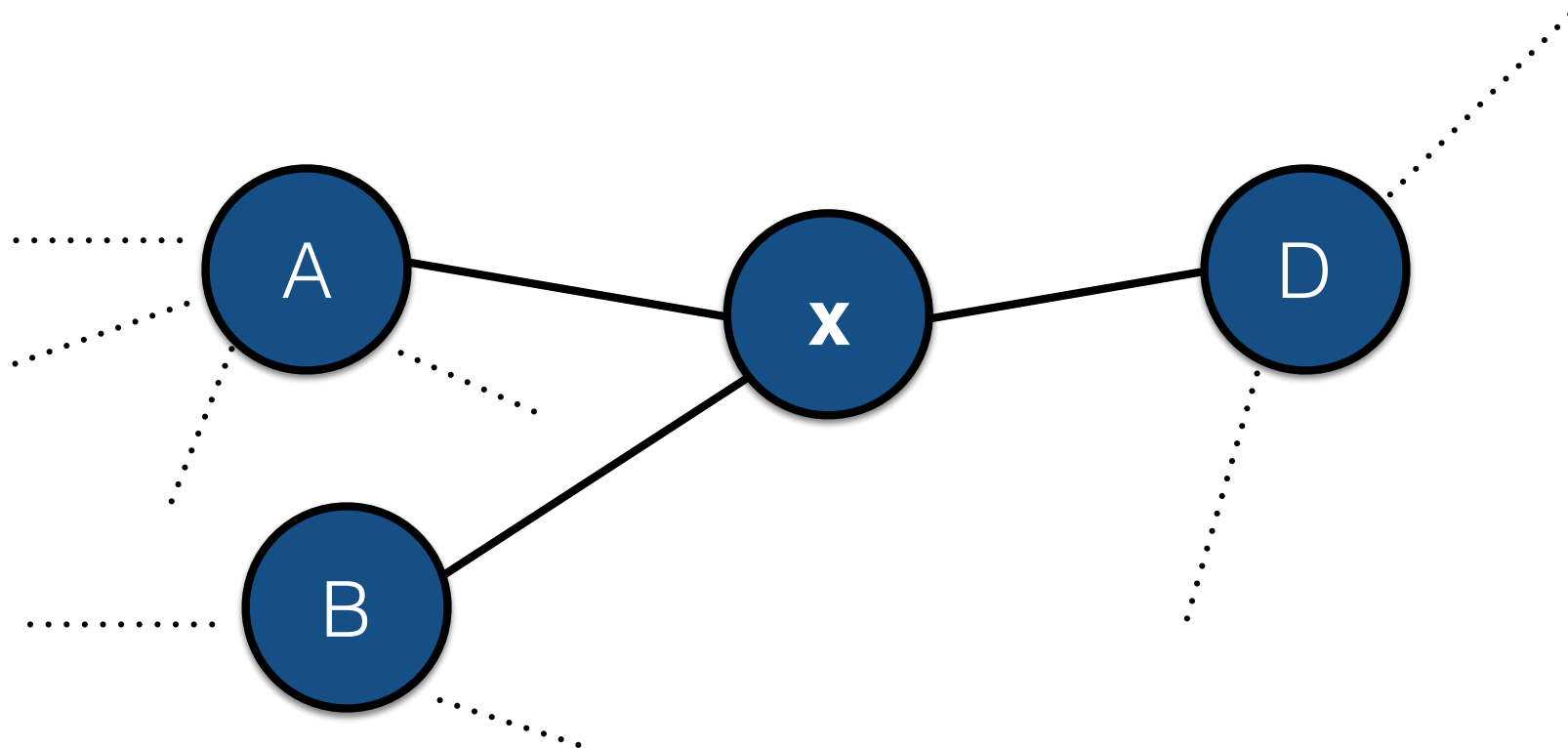


Eigenvalues centrality

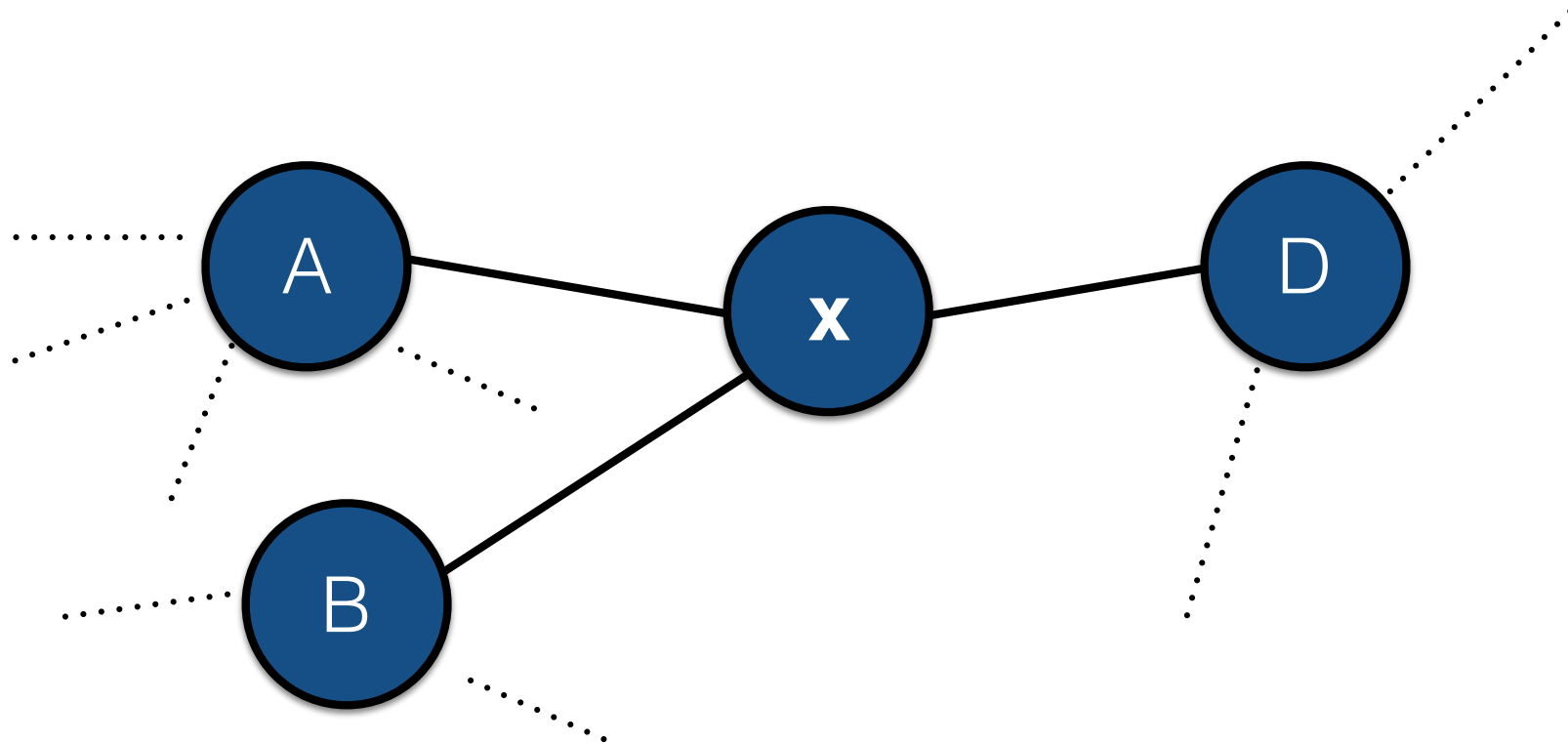
- Similar to a vote mechanism where each page vote for the election of the best page. Each page gives its vote to its neighbours.



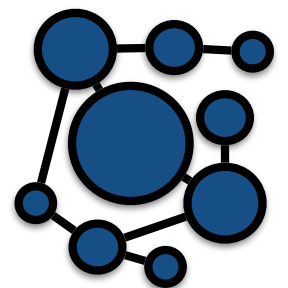
Eigenvalues centrality



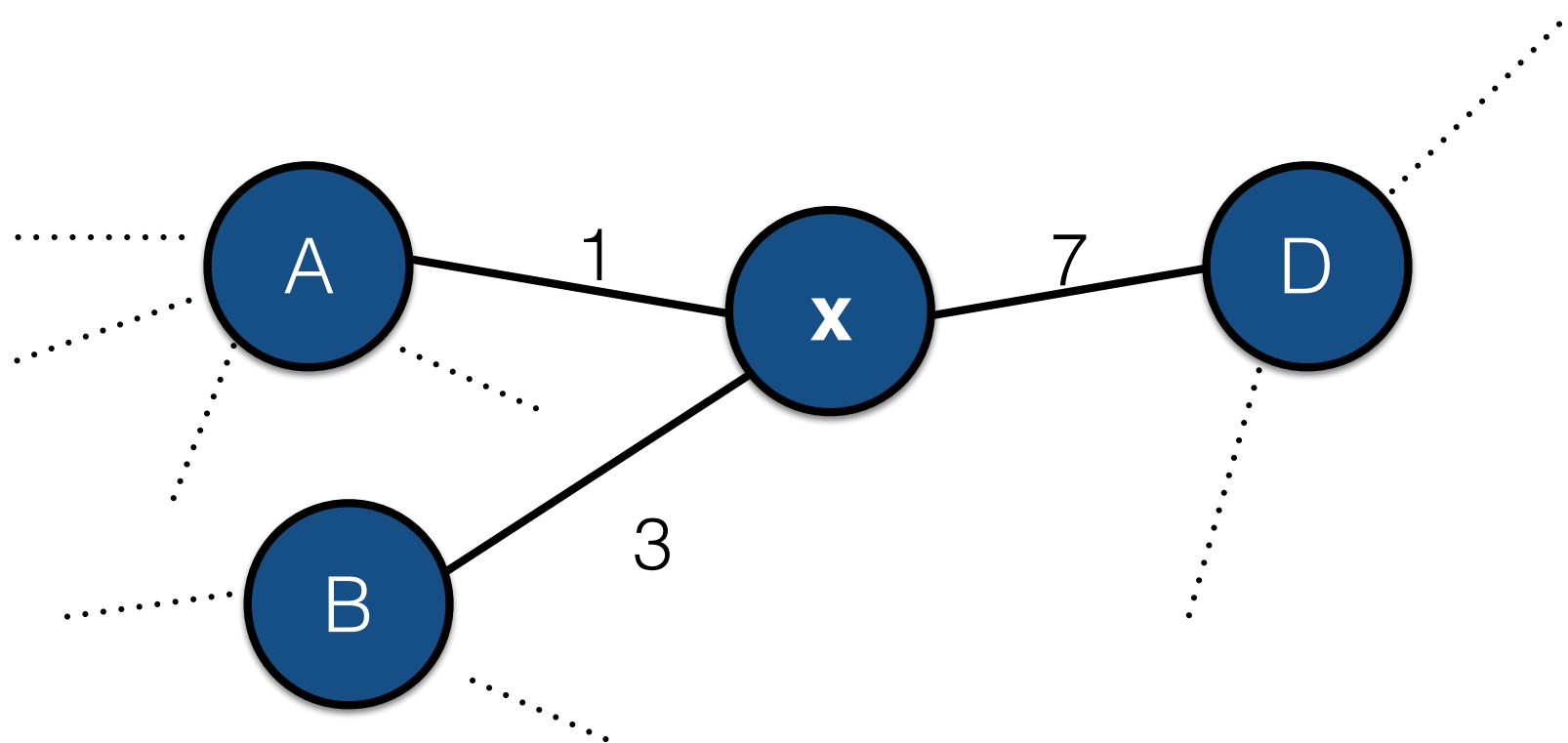
Eigenvalues centrality



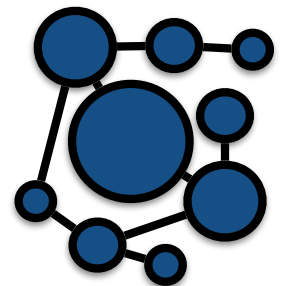
$$C_{EV}(X) = C_{EV}(D) + C_{EV}(A) + C_{EV}(B)$$



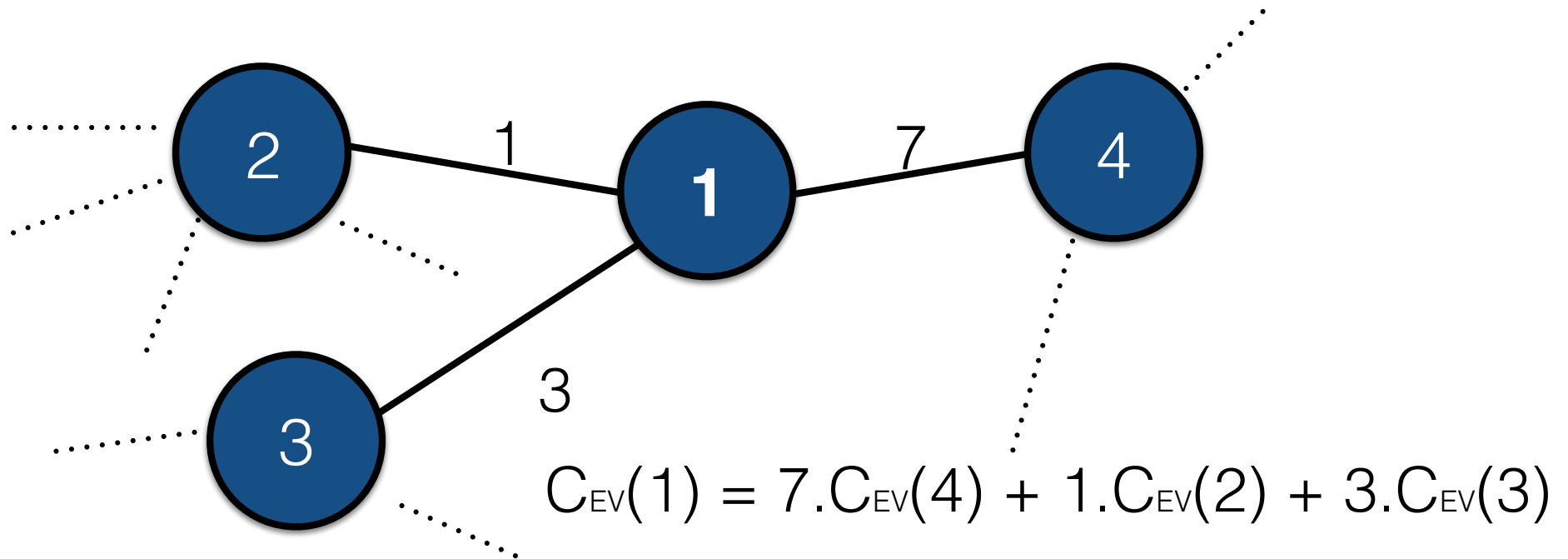
Eigenvalues centrality



$$C_{EV}(X) = 7.C_{EV}(D) + 1.C_{EV}(A) + 3.C_{EV}(B)$$

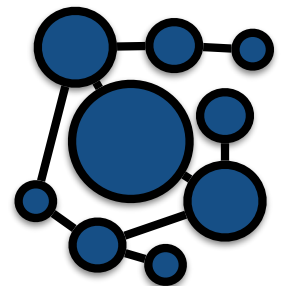


Eigenvalues centrality



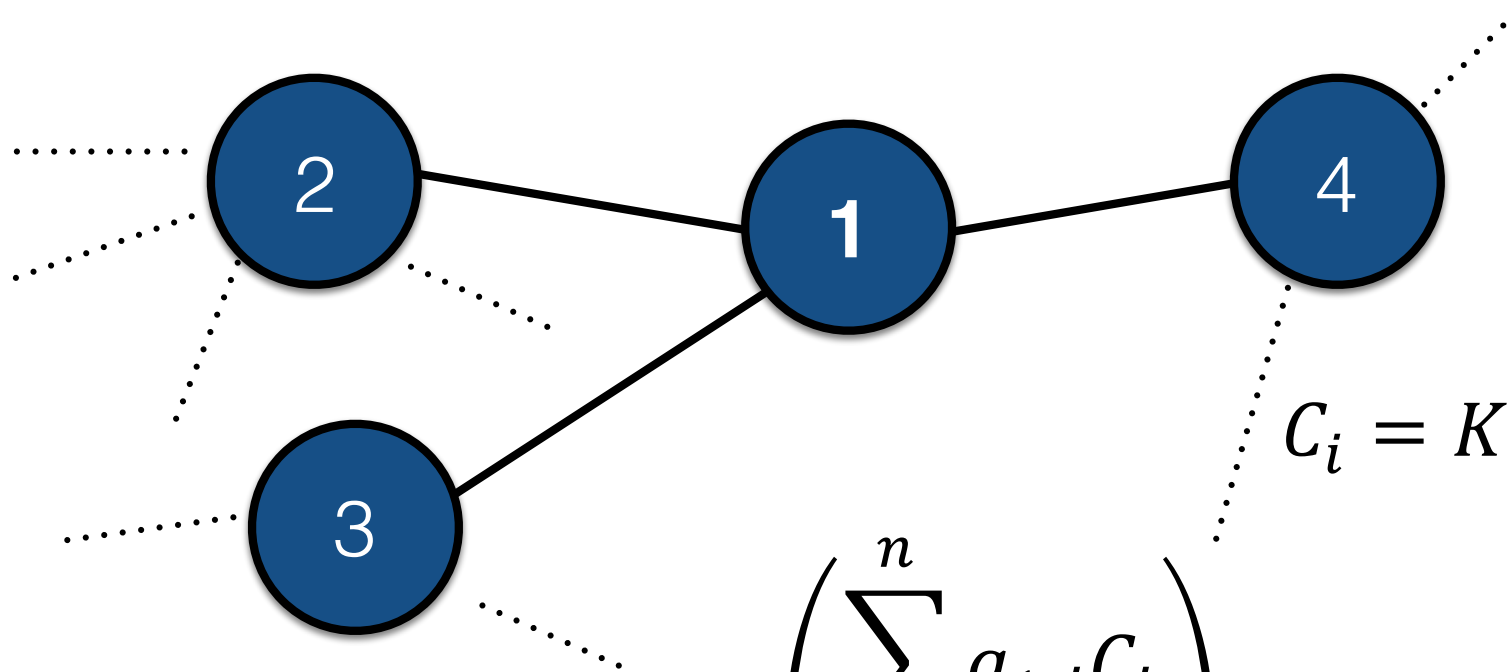
$$C_{EV}(i) = K (a_{i,1}.C_{EV}(1) + a_{i,1}.C_{EV}(2) + \dots + a_{i,n}.C_{EV}(n))$$

$$C_i = K \sum_{j=1}^n a_{i,j} C_j$$



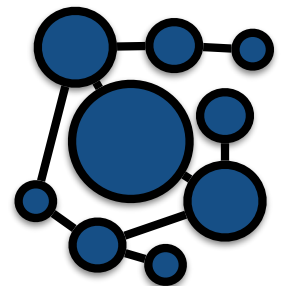
Eigenvalues centrality

$$C_{EV}(1) = 7.C_{EV}(4) + 1.C_{EV}(2) + 3.C_{EV}(3)$$



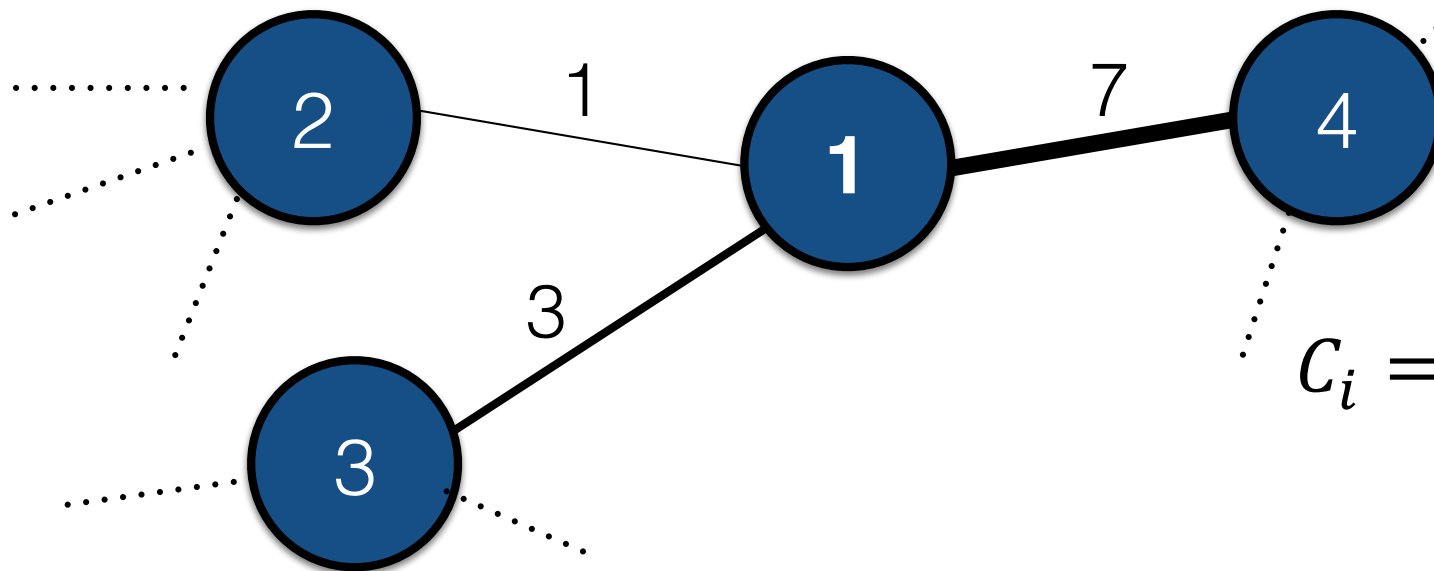
$$C_i = K \sum_{j=1}^n a_{i,j} C_j$$

$$\begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix} = K \begin{pmatrix} \sum_{j=1}^n a_{1,j} C_j \\ \vdots \\ \sum_{j=1}^n a_{n,j} C_j \end{pmatrix}$$



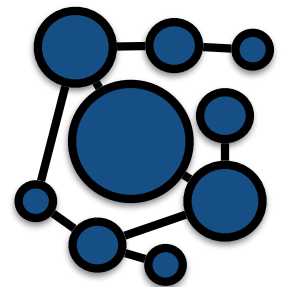
Eigenvalues centrality

$$C_{EV}(1) = 7.C_{EV}(4) + 1.C_{EV}(2) + 3.C_{EV}(3)$$



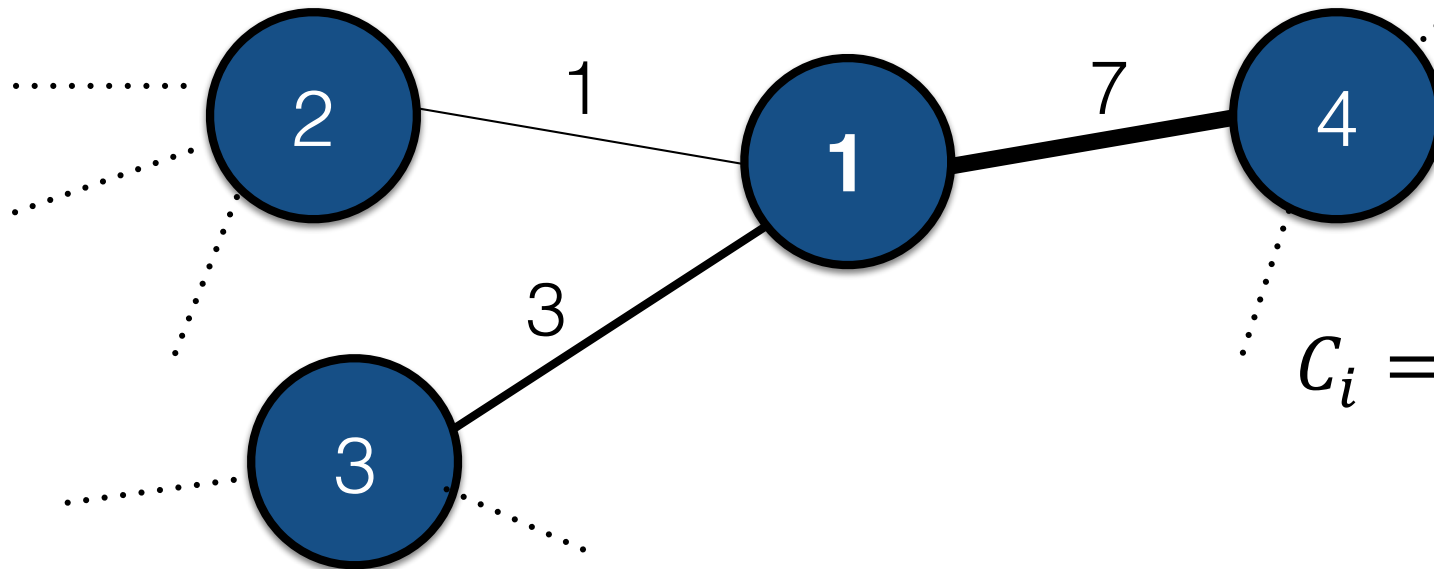
$$C_i = K \sum_{j=1}^n a_{i,j} C_j$$

$$\begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix} = K \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix} \begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix}$$



Eigenvalues centrality

$$C_{EV}(1) = 7.C_{EV}(4) + 1.C_{EV}(2) + 3.C_{EV}(3)$$

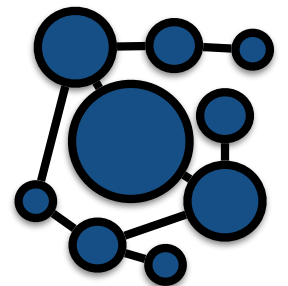


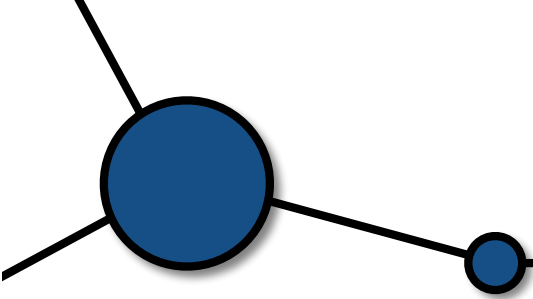
$$C_i = K \sum_{j=1}^n a_{i,j} C_j$$

$$C = K(A * C)$$

$$\frac{1}{K} C = AC$$

$$\lambda C = AC$$



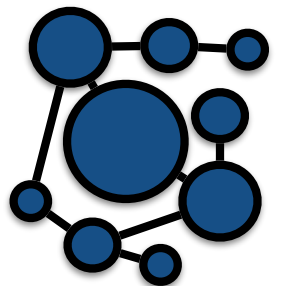


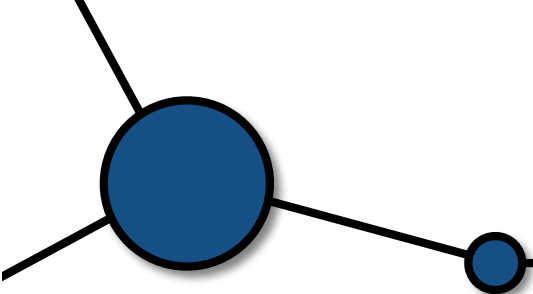
Eigenvalues centrality

- **C** is an Eigenvector of **A** associated with the Eigenvalue **λ**
- We choose the highest Eigenvalue to make sure that every Eigenvector components are positive (Perron-Frobenius theorem)

$$\lambda C = AC$$

$$(A - \lambda I)C = 0$$

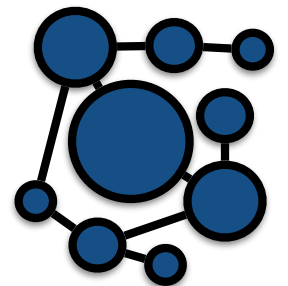




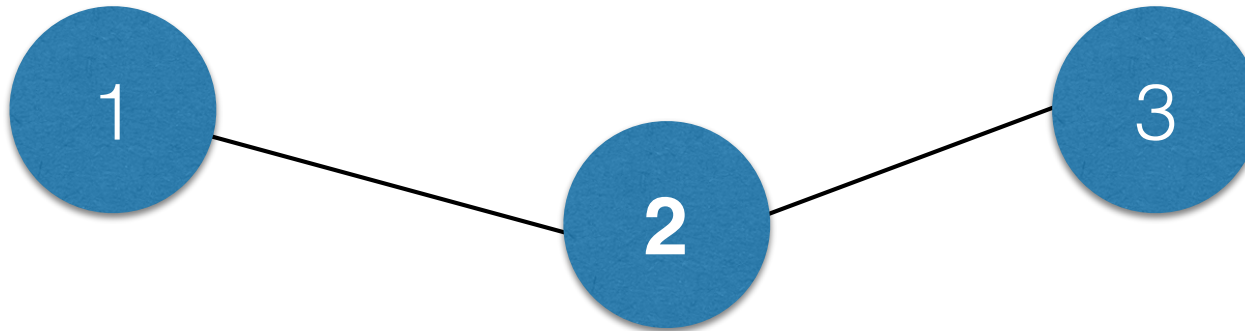
Eigenvalues centrality

- Definition : Eigenvalue centrality of a node i equals to the i^{th} component du Eigenvector associated with the highest Eigenvalues of the adjacency matrix of the graph.

$$C_{EV}(i) = x_i = \frac{1}{\lambda} \sum_{j=1}^n A_{i,j} x_j$$



Eigenvalues centrality

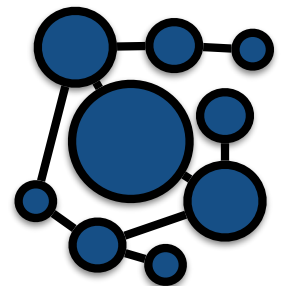


$$(A - \lambda I)C = 0$$

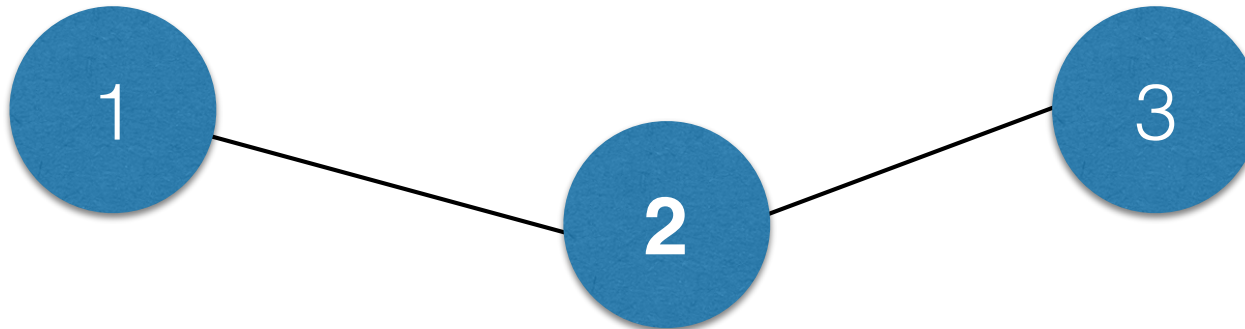
$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$



$$\begin{pmatrix} 0 - \lambda & 1 & 0 \\ 1 & 0 - \lambda & 1 \\ 0 & 1 & 0 - \lambda \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = 0$$



Eigenvalues centrality

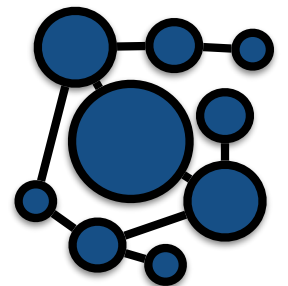


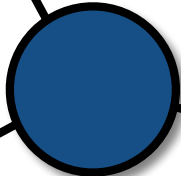
$$\begin{pmatrix} 0 & -\lambda & 1 \\ 1 & 0 & -\lambda \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = 0$$

Eigenvectors are the roots of the characteristic polynomial

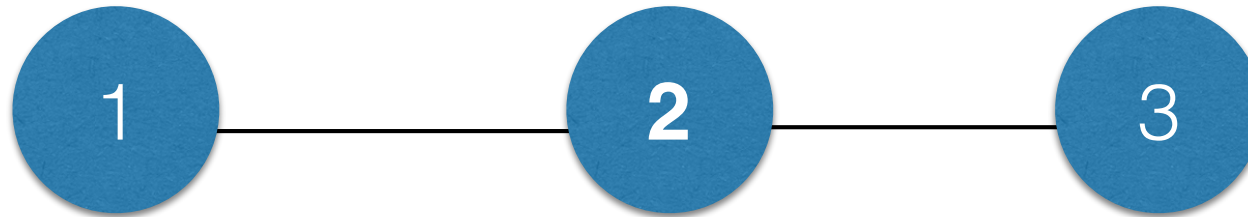


$$\begin{vmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{vmatrix} = 0$$





Eigenvalues centrality



$$\begin{vmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{vmatrix} = 0$$

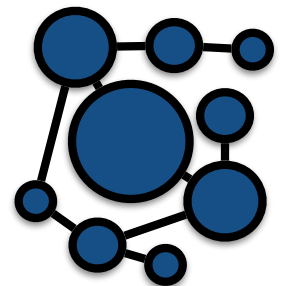
$$-\lambda * (\lambda^2 - 1) - (-\lambda) = 0$$

$$-\lambda^3 + 2\lambda = 0$$

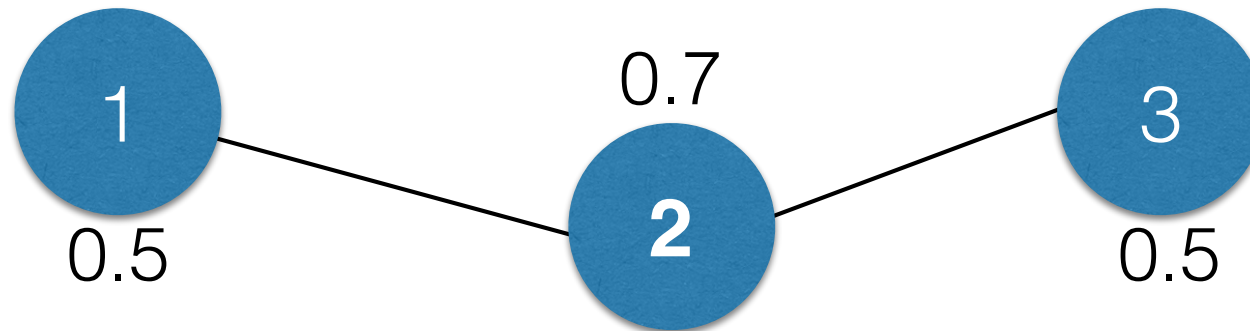
$$\lambda(-\lambda^2 + 2) = 0$$

$$\lambda = 0$$

$$\lambda = \pm\sqrt{2}$$



Eigenvalues centrality



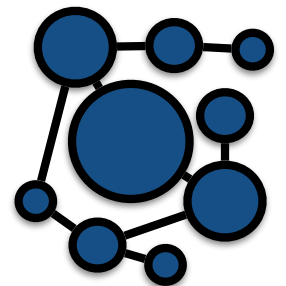
$$\begin{cases} -\sqrt{2}C_1 + C_2 = 0 \\ C_1 - \sqrt{2}C_2 + C_3 = 0 \\ C_2 - \sqrt{2}C_3 = 0 \end{cases}$$

$$\begin{cases} C_3 = C_1 \\ C_2 = \sqrt{2}C_1 \end{cases}$$

$$c = \begin{pmatrix} 0.5 \\ 0.7 \\ 0.5 \end{pmatrix}$$

$$\begin{cases} C_2 = \sqrt{2}C_1 \\ C_3 = \frac{1}{\sqrt{2}}C_2 \\ C_1 - \sqrt{2}C_2 + C_3 = 0 \end{cases}$$

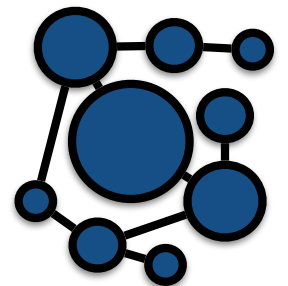
$$c = \begin{pmatrix} C_1 \\ \sqrt{2}C_1 \\ C_1 \end{pmatrix}$$

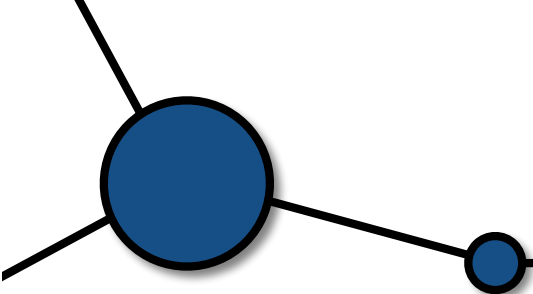




The PageRank algorithm

- Algorithm proposed by Larry Page
- Modification of the Eigenvalue centrality to fix some large scale issues
- Algorithm still partially used by Google and published in 1997



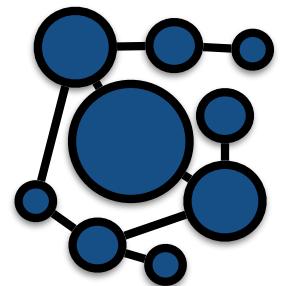


Eigenvalues centrality

- Eigenvalues centrality

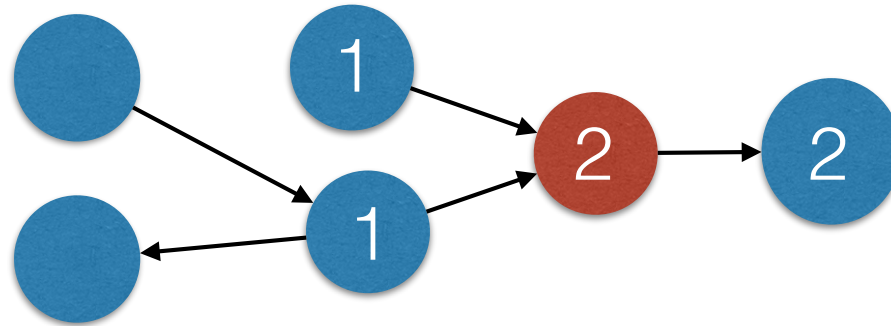
$$C_i = K \sum_{j=1}^n a_{i,j} C_j$$

- Two modifications are made

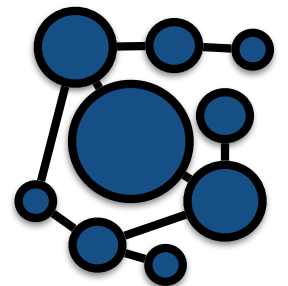
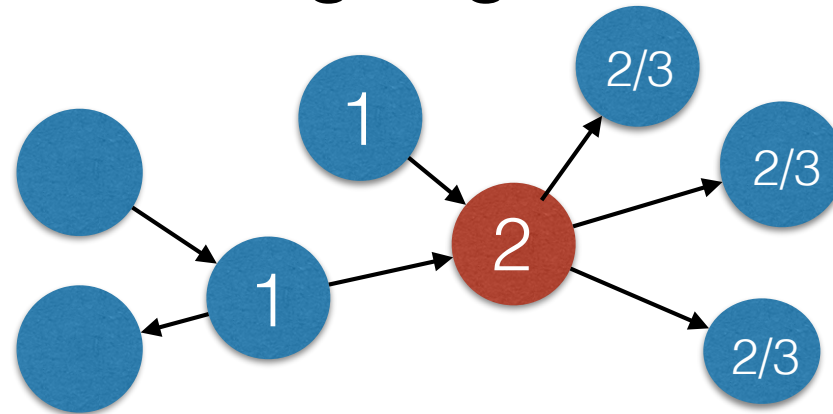


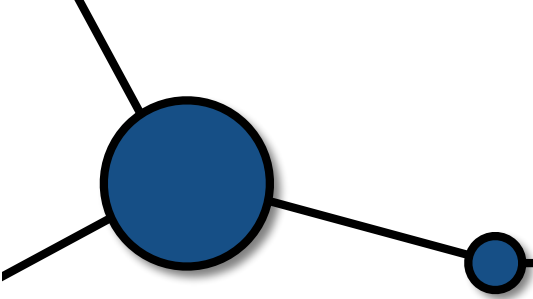
Modification 1

- A page receives **less weight** from a page that has **a high number of outgoing links**



- A page receives **more weight** from a page that has **a low number of outgoing links**





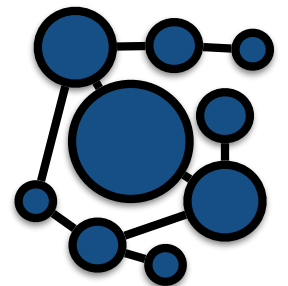
Modification 1

- Eigenvalue centrality

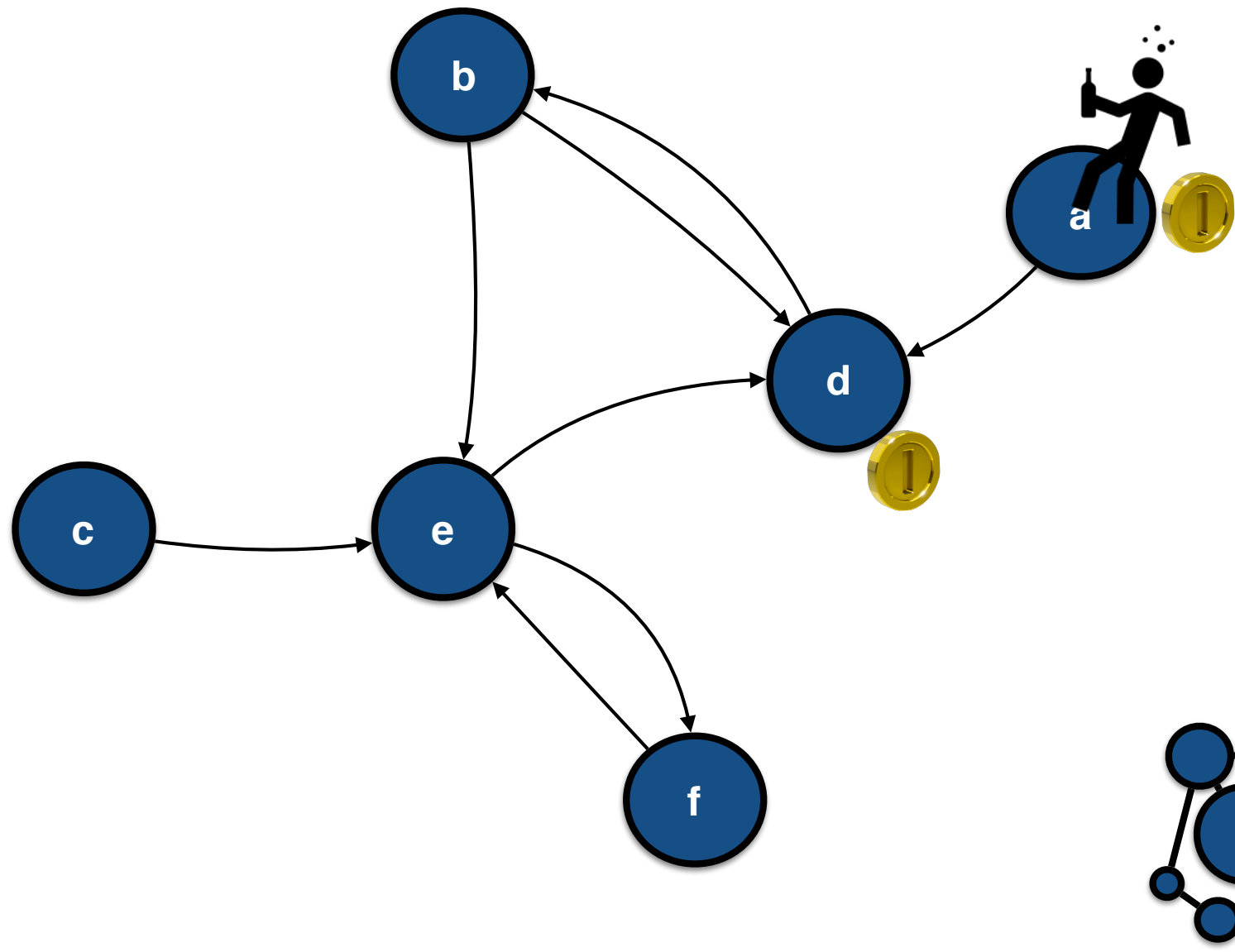
$$C_i = K \sum_{j=1}^n a_{i,j} C_j$$

- PageRank : first modification

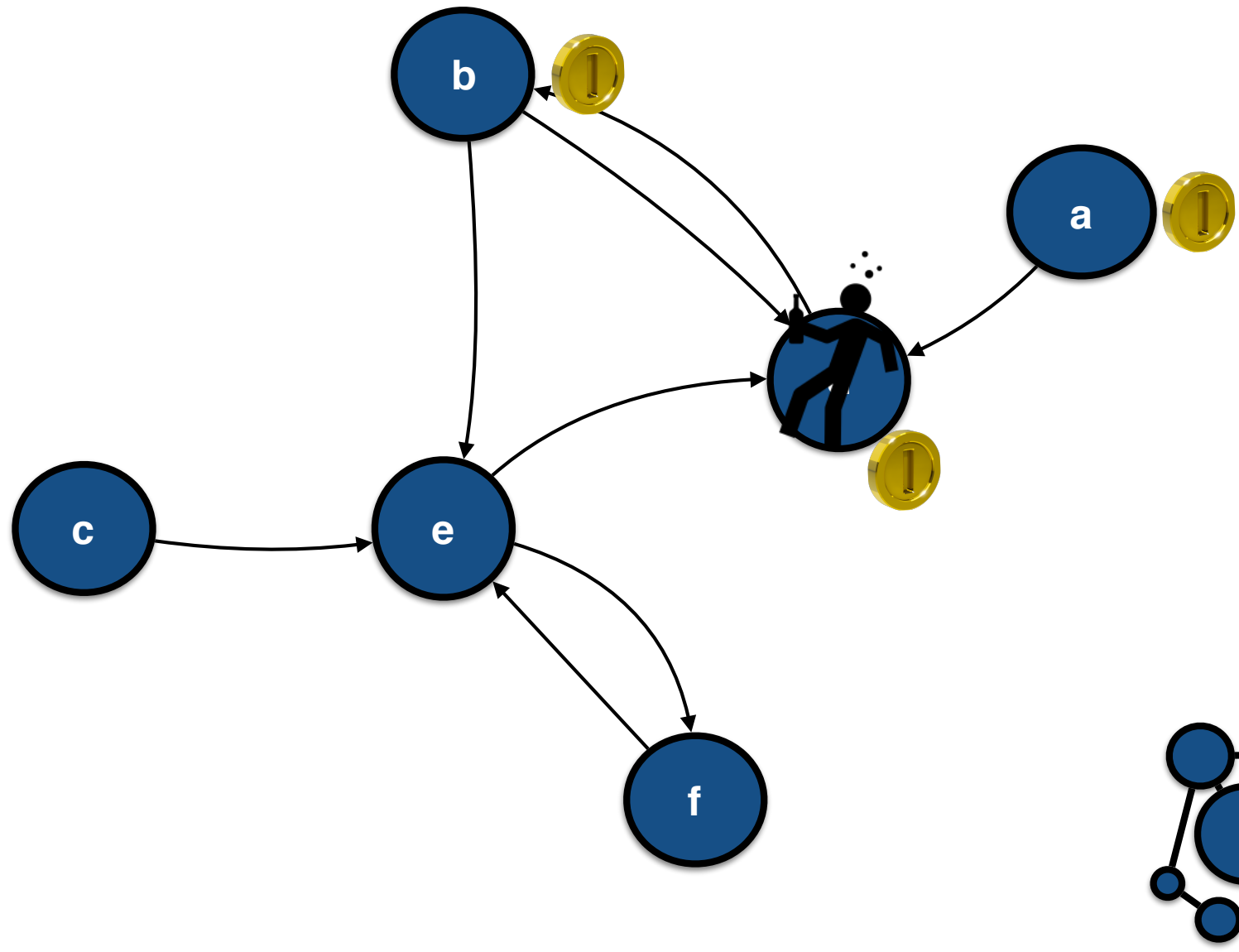
$$C_i = K \sum_{j=1..N} a_{j,i} \frac{C_j}{k_j^{out}}$$

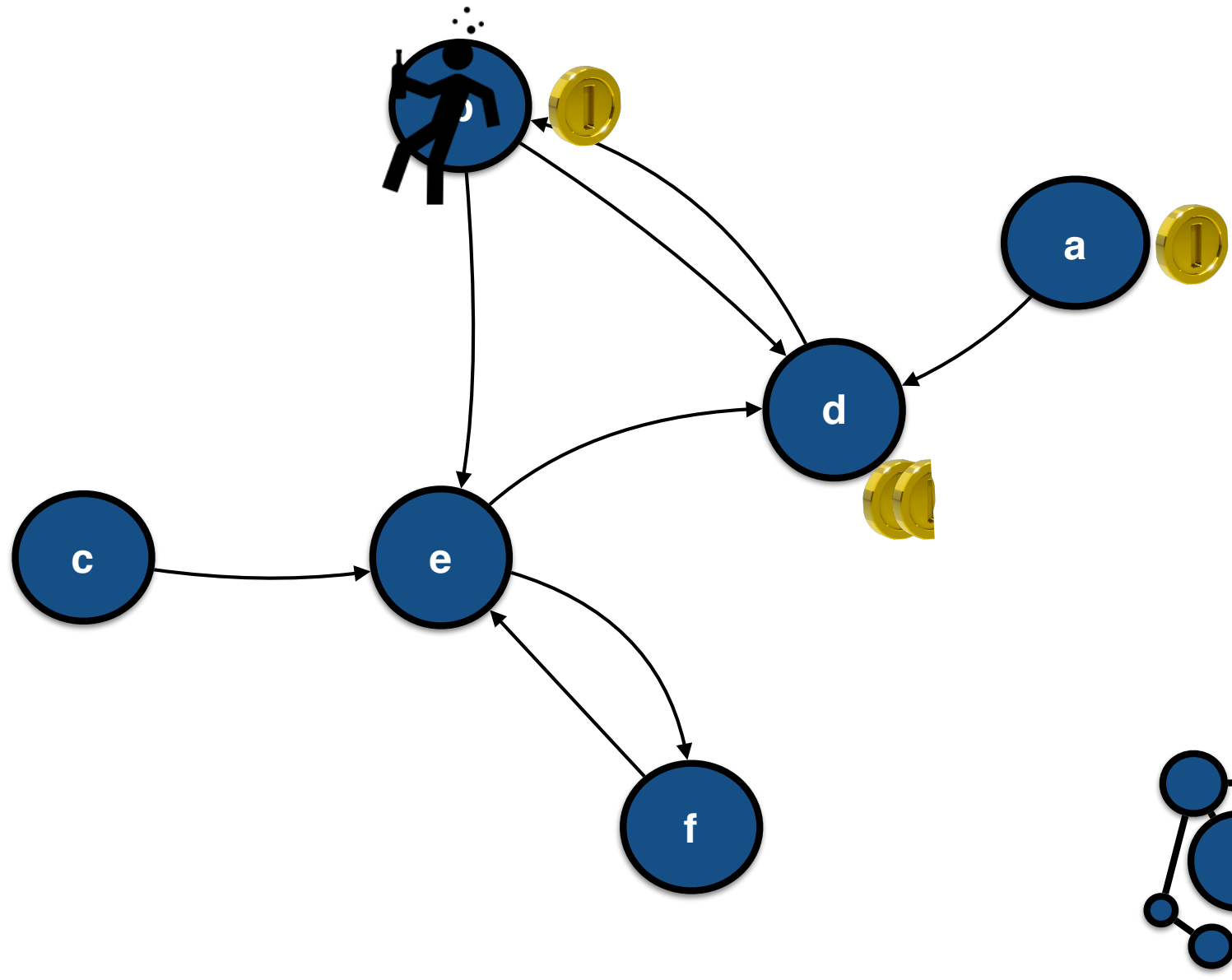
Random walk similarity



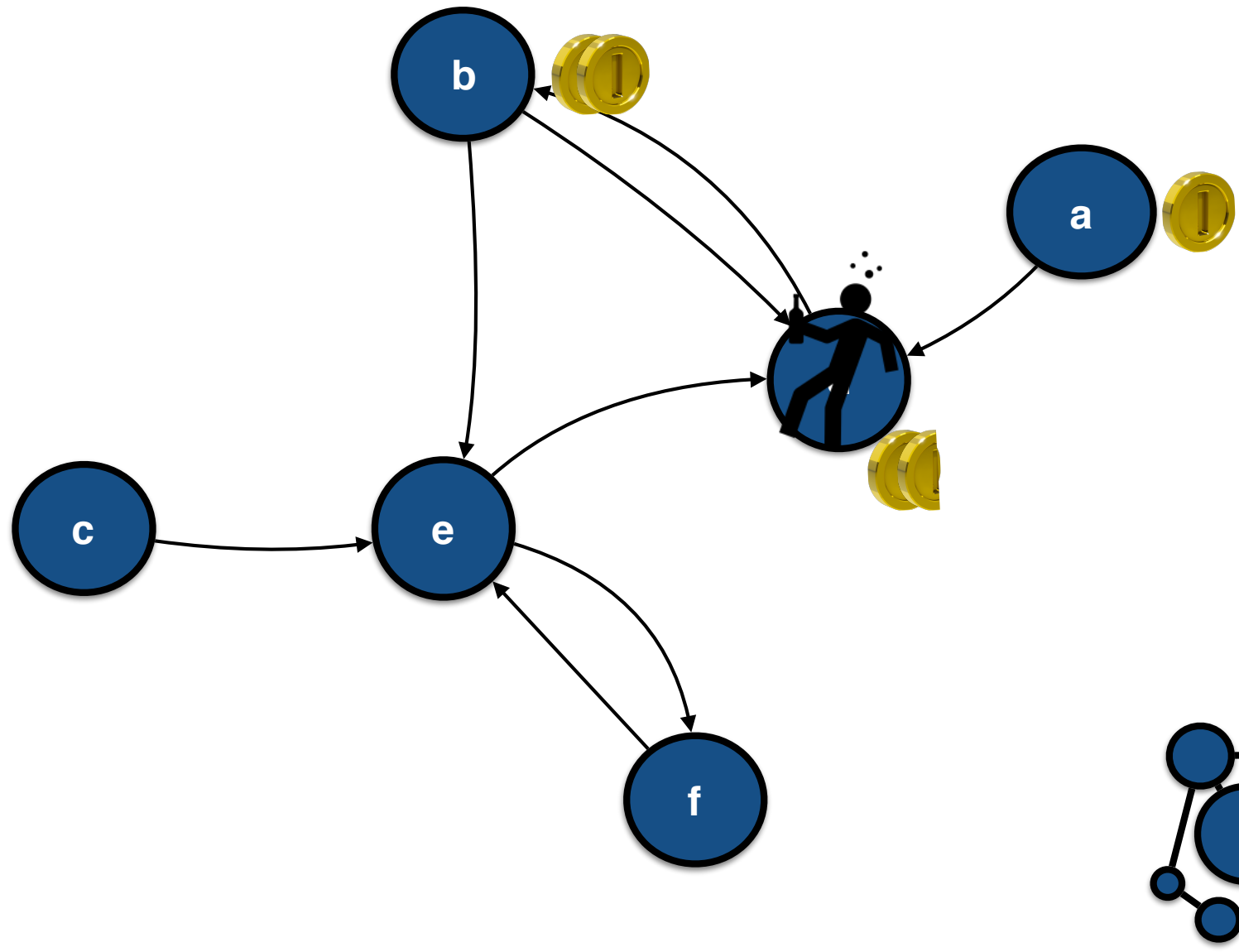
Random walk similarity



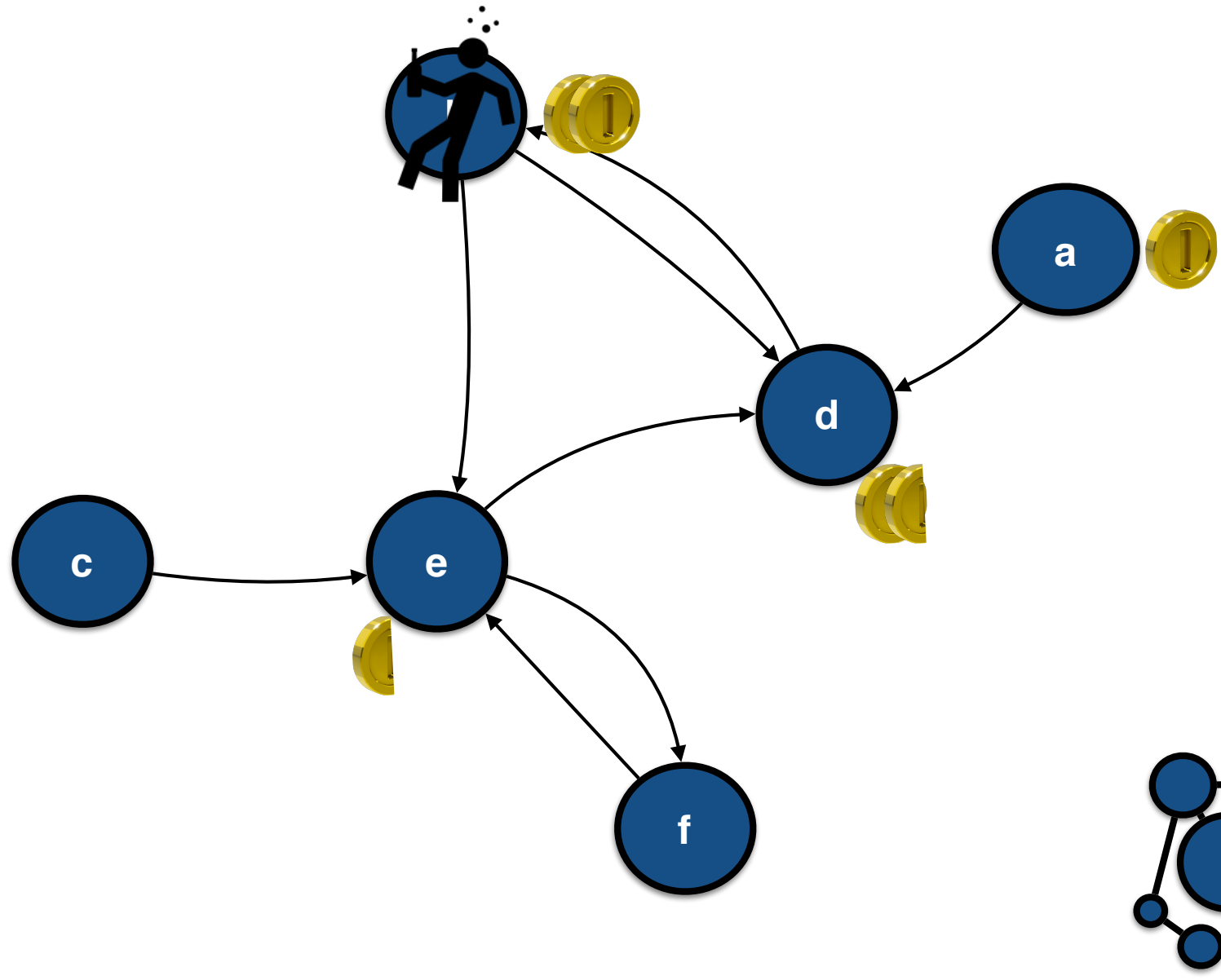
Random walk similarity



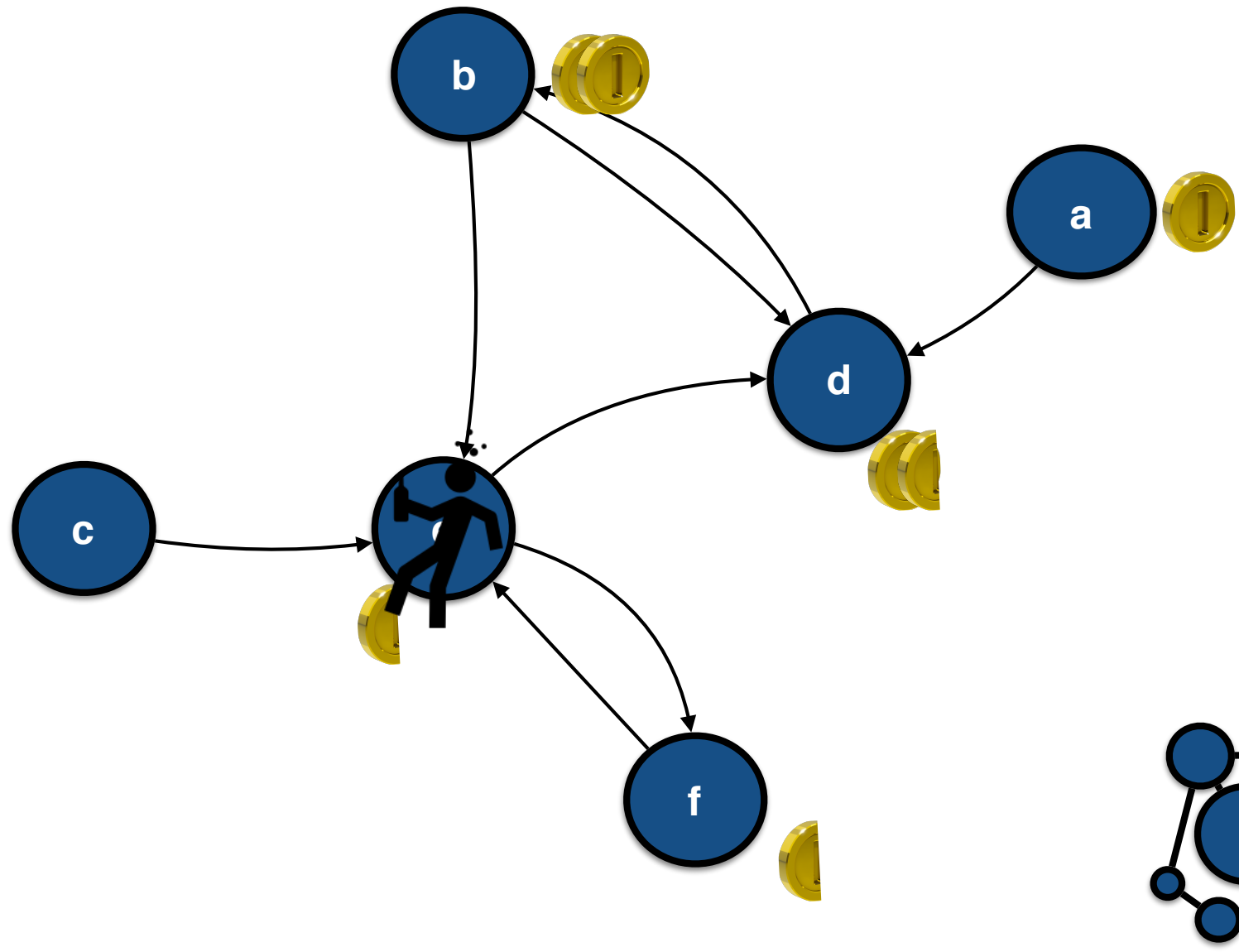
Random walk similarity



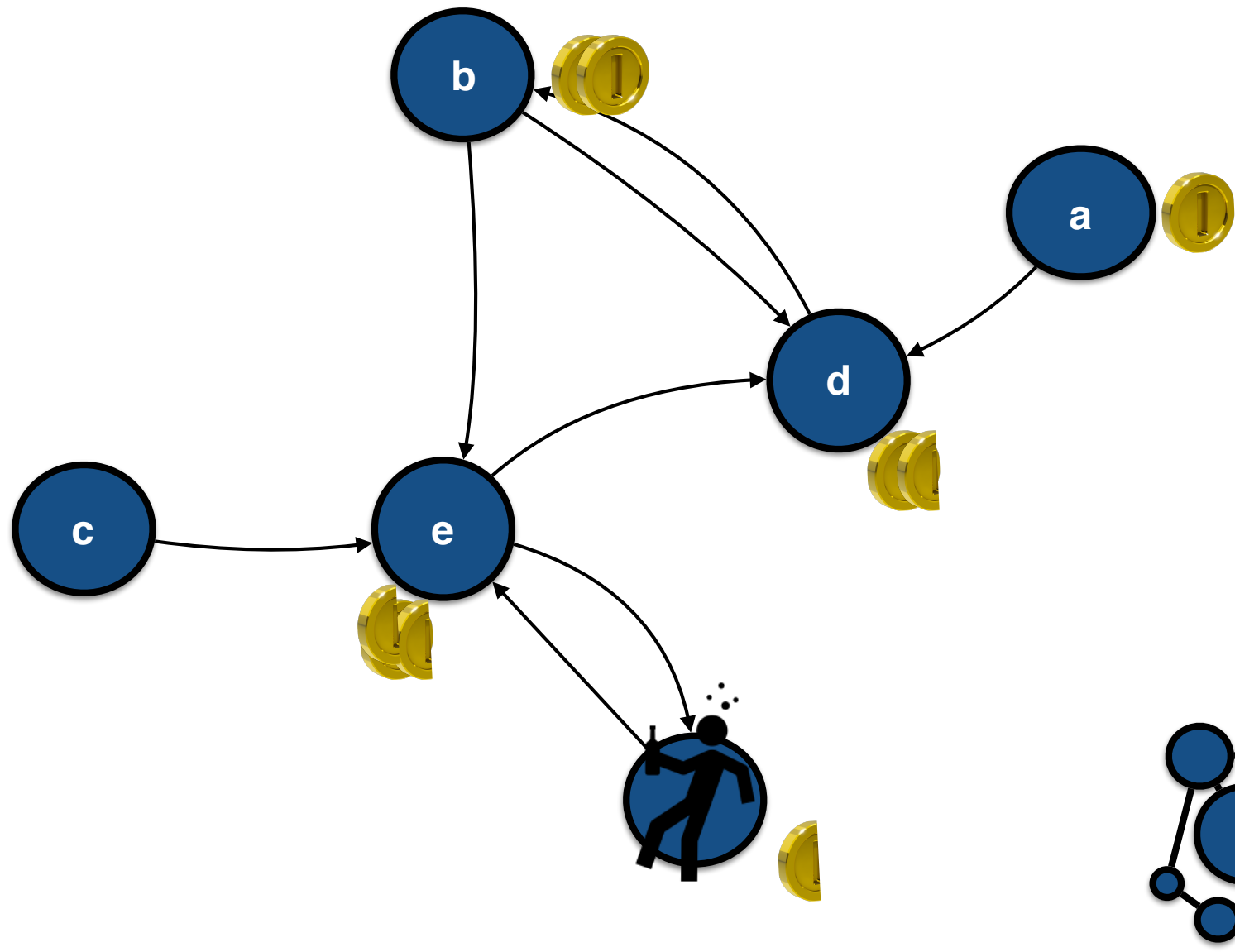
Random walk similarity



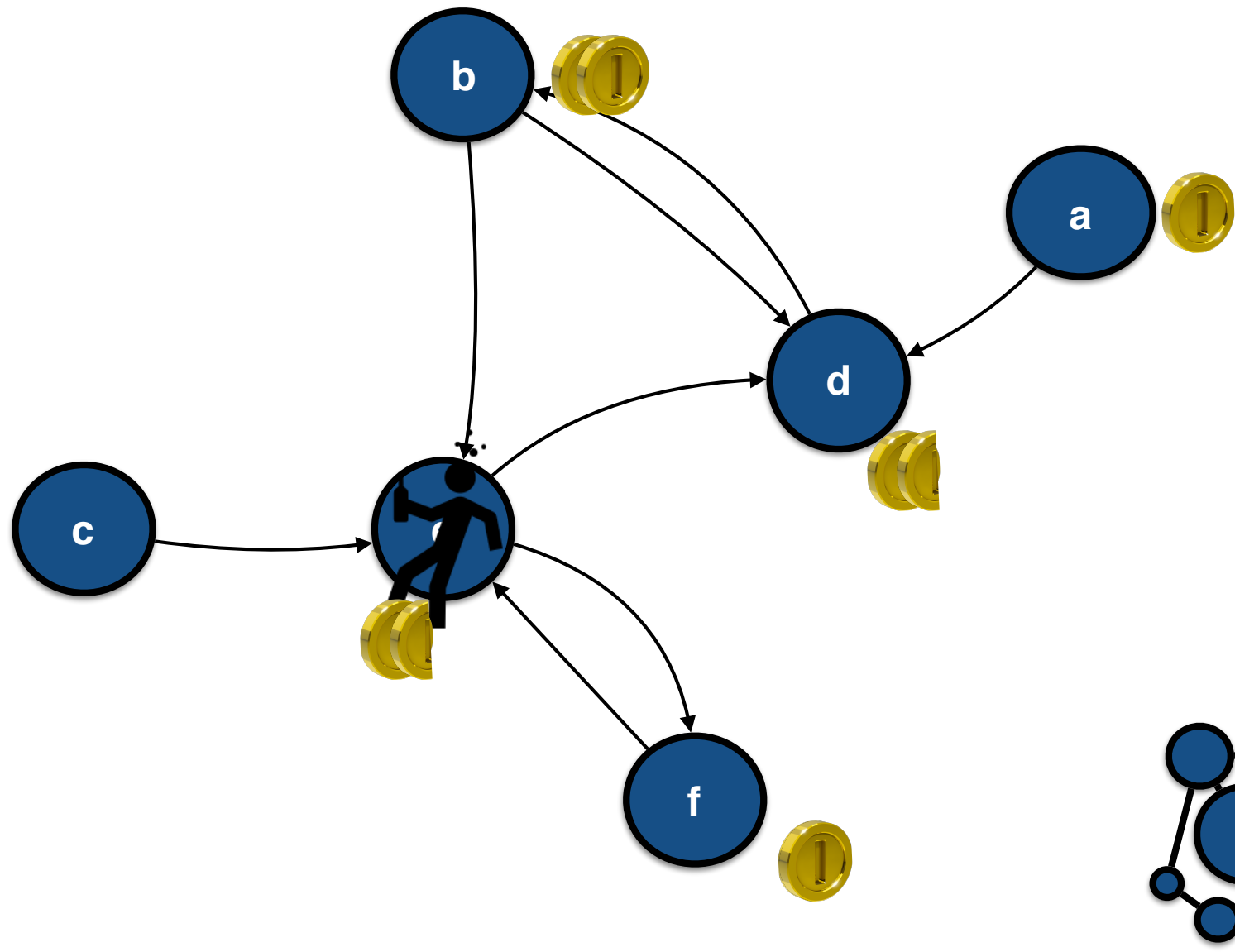
Random walk similarity



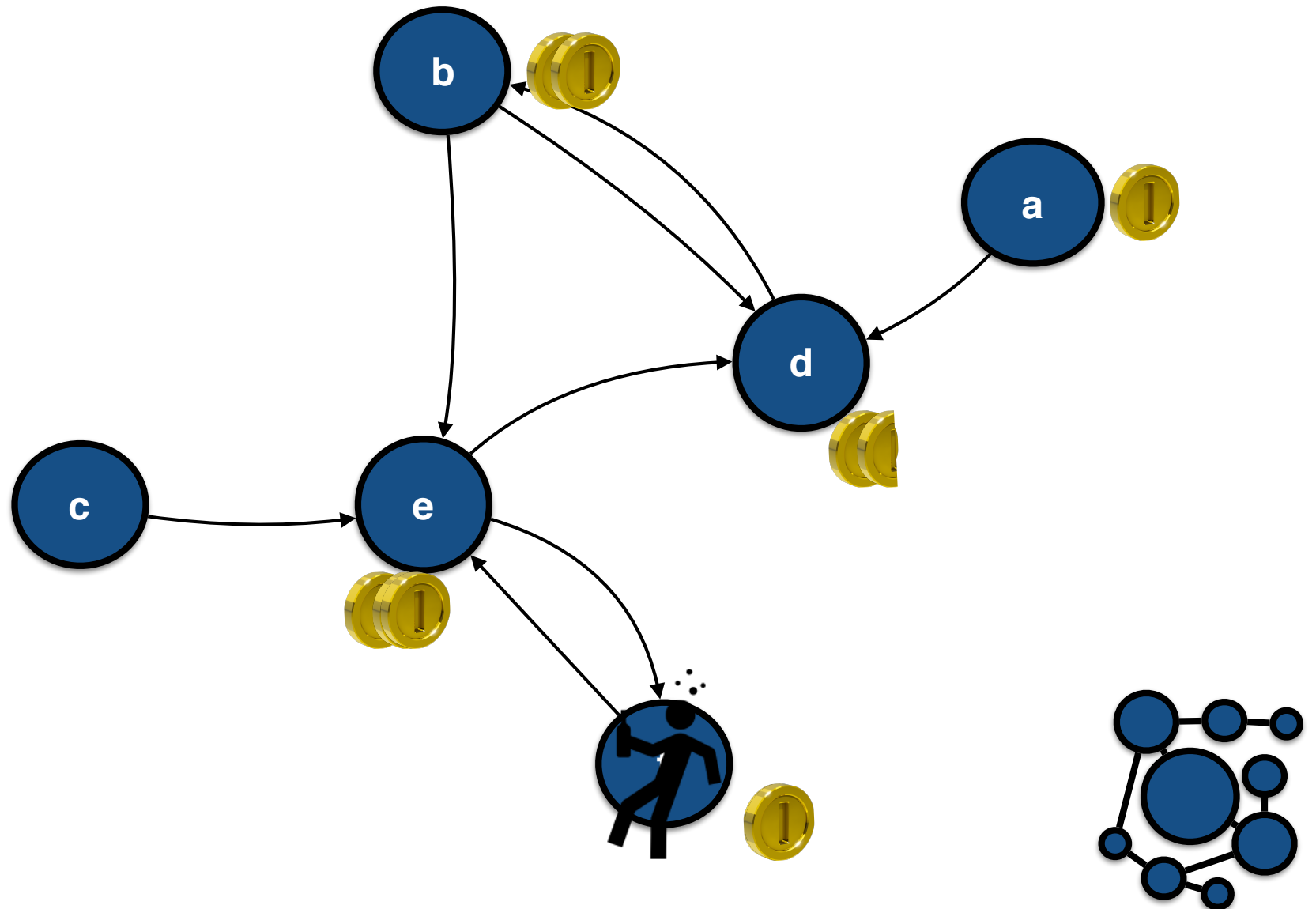
Random walk similarity



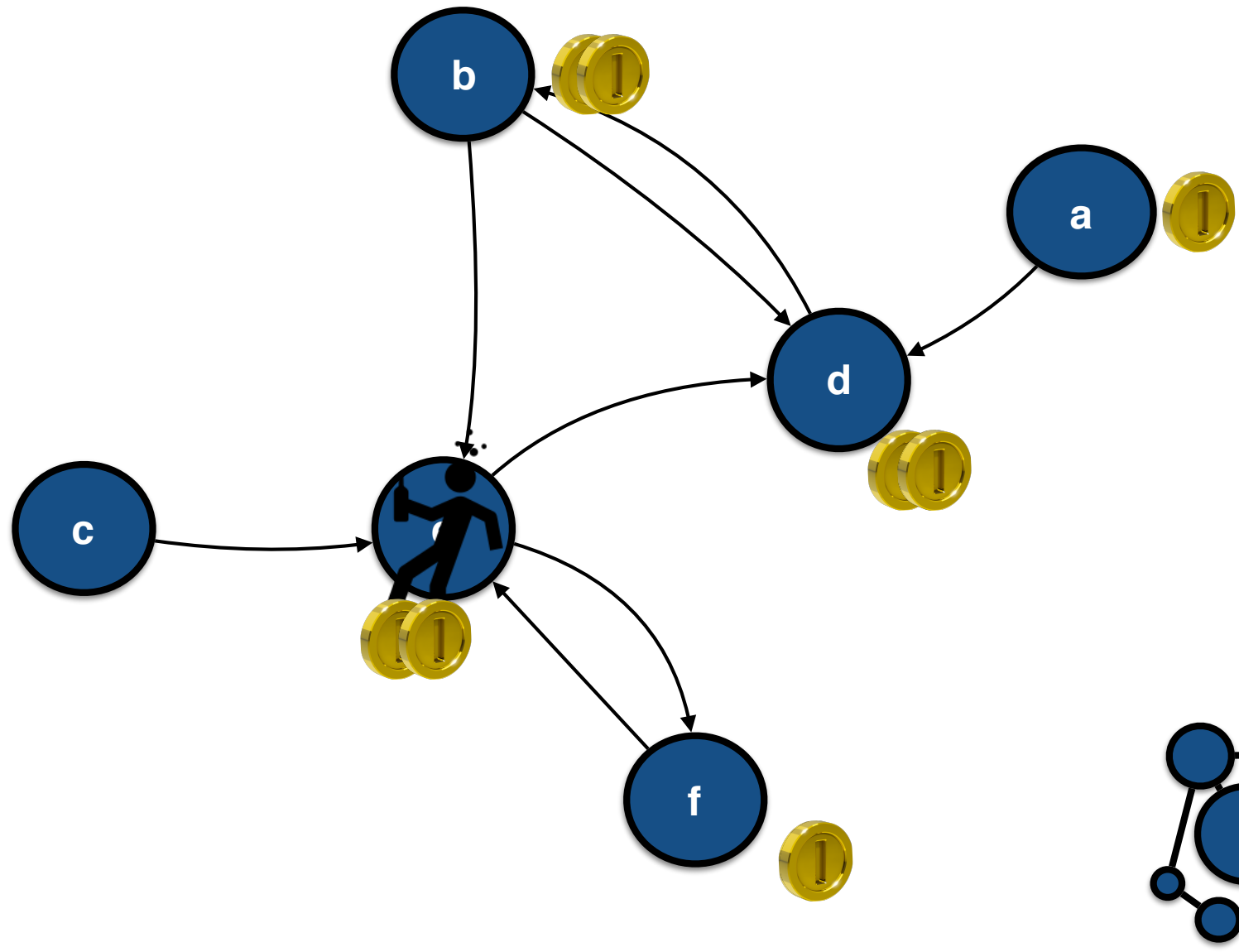
Random walk similarity



Random walk similarity

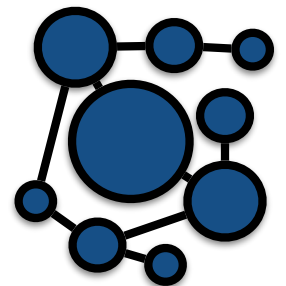
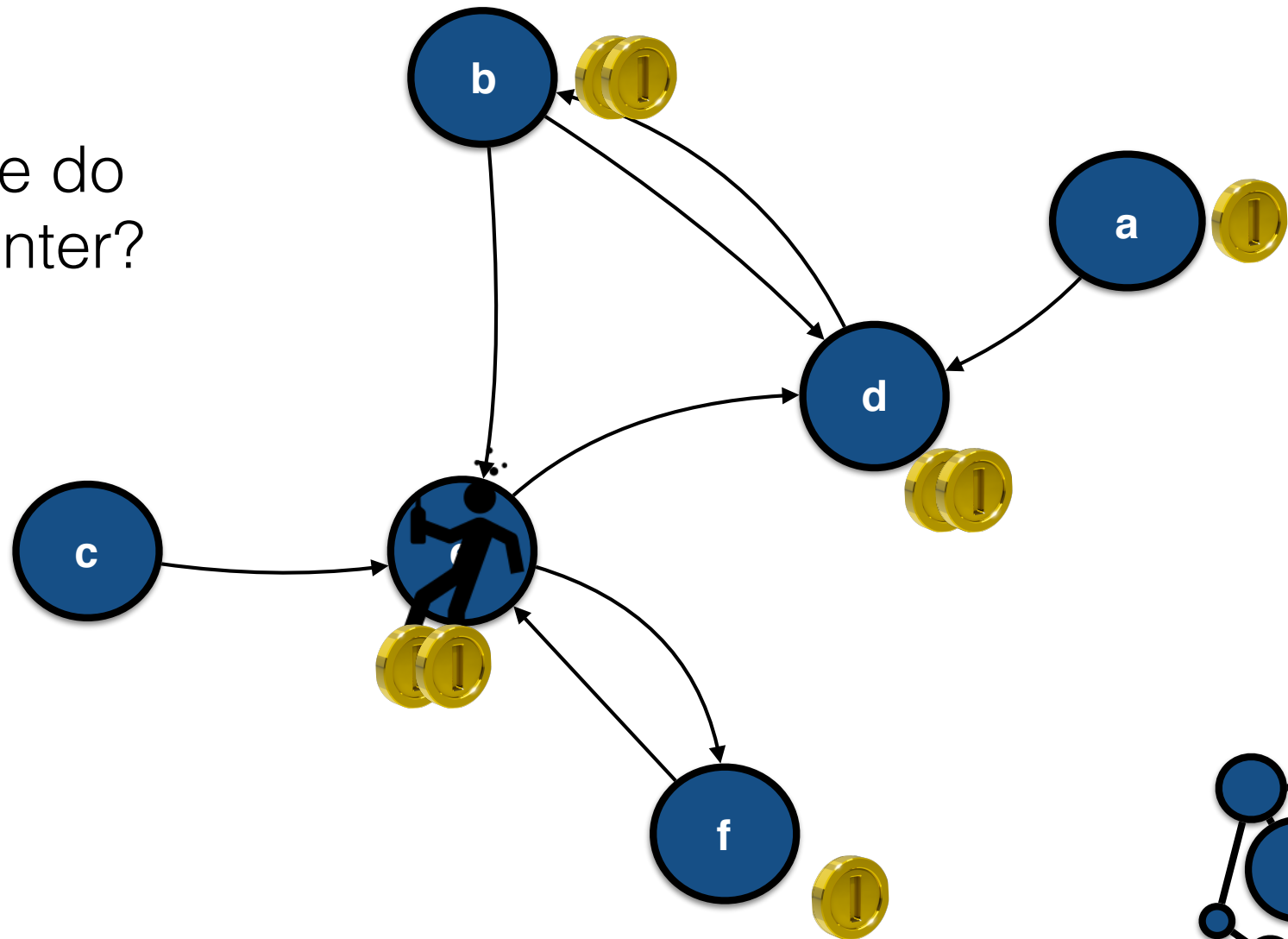


Random walk similarity



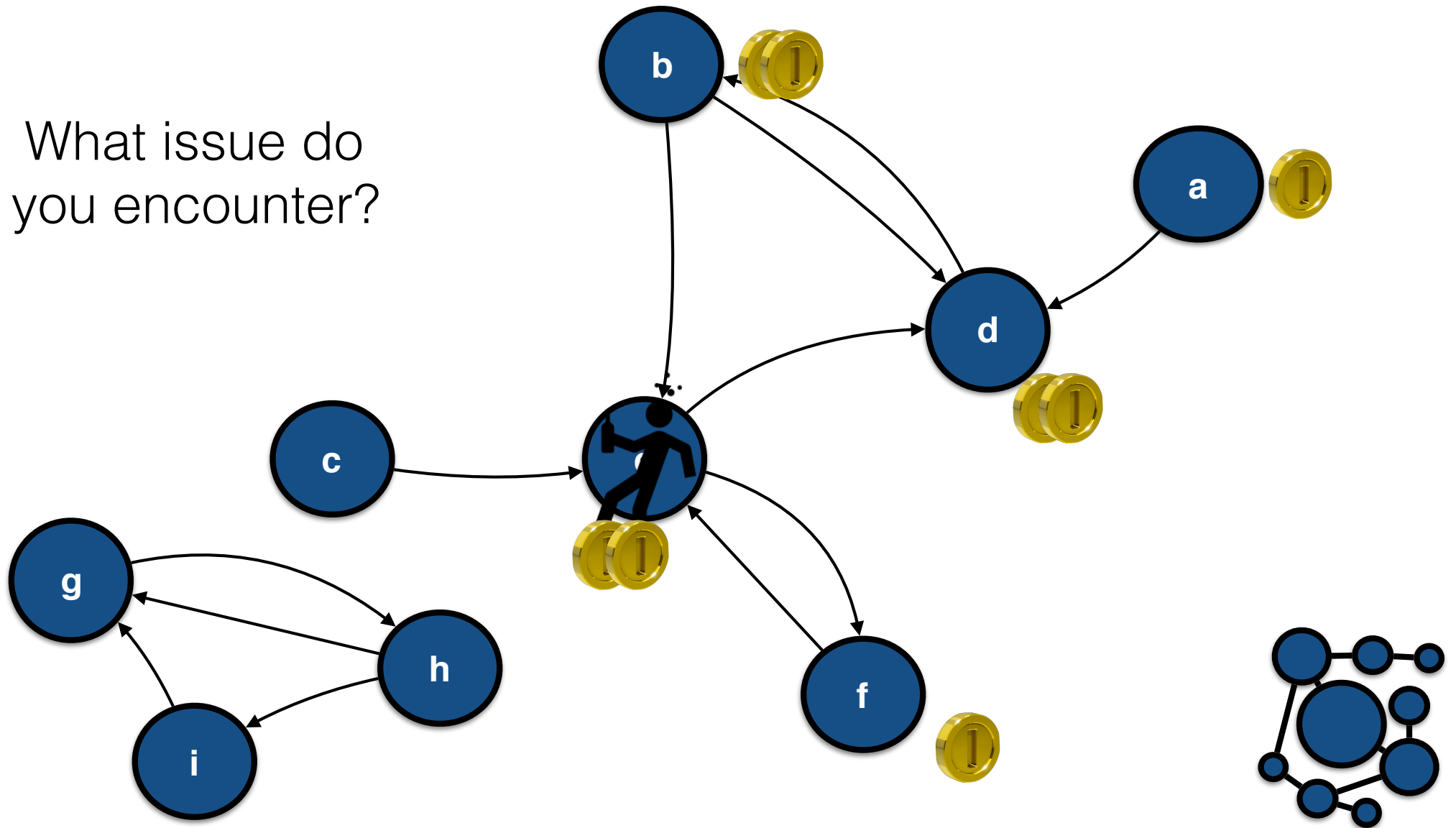
Random walk similarity

What issue do you encounter?



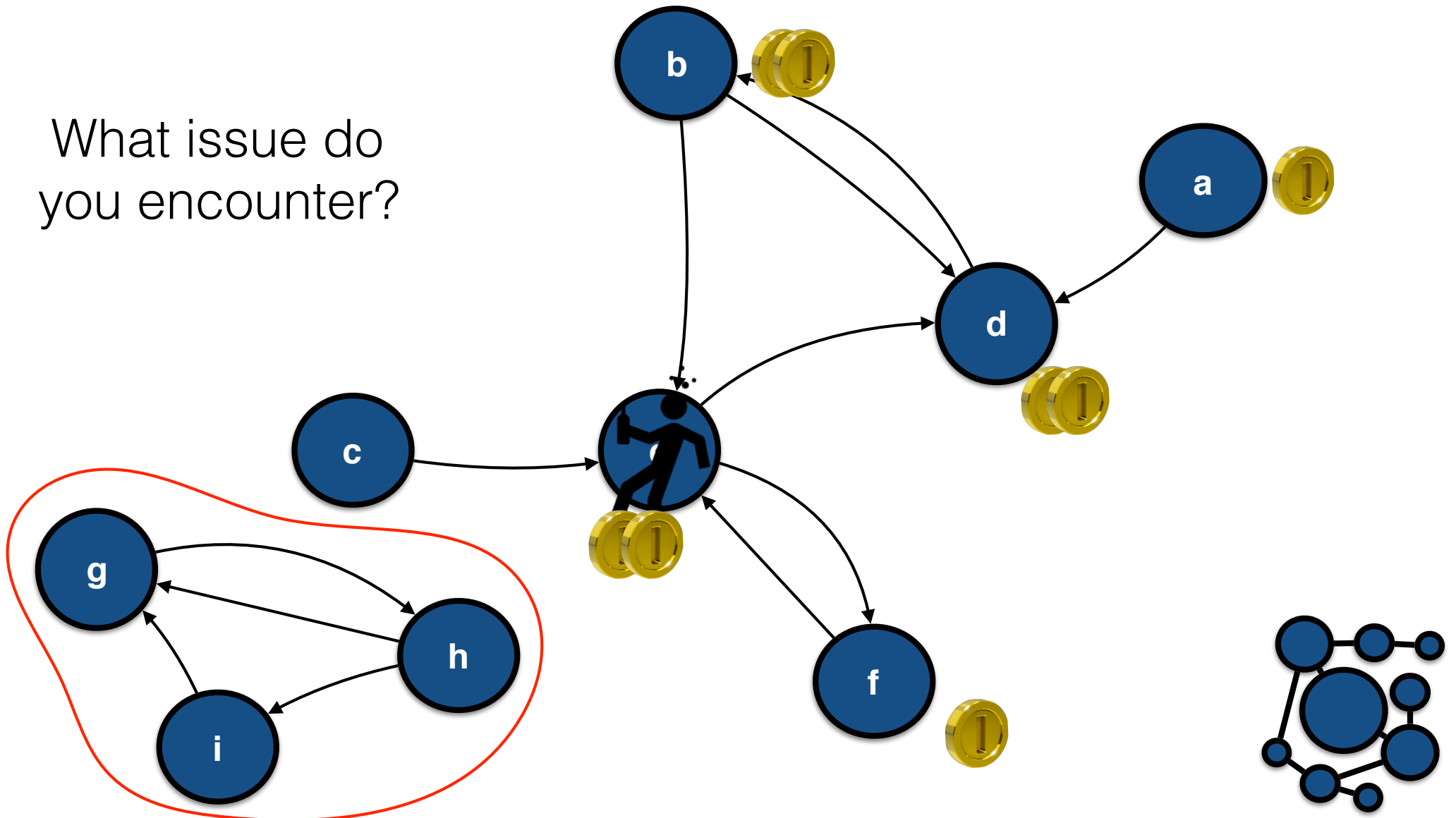
Random walk similarity

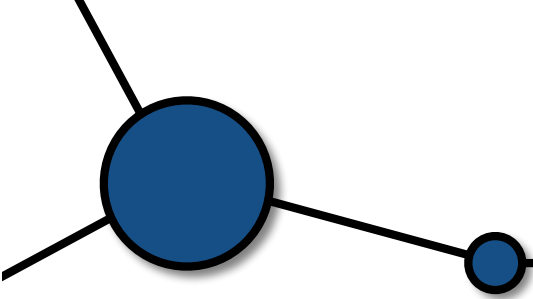
What issue do you encounter?



Random walk similarity

What issue do you encounter?



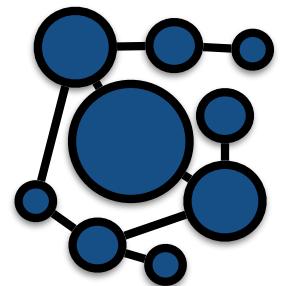


Modification 2

- It is now possible to move at any position of the graph with a certain probability

$$C_i = \frac{(1 - d)}{N} + d \sum_{j=1..N} a_{j,i} \frac{C_j}{k_j^{out}}$$

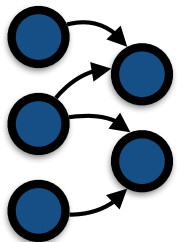
- Usually $d=0.85$
- Teleportation probability equals 15%



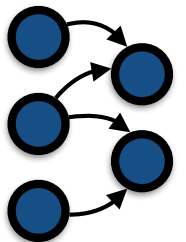


The HITS algorithm

- Algorithm proposed by John Kleinberg
- Used for ranking webpages
- Issues :
 - Too many pages are results of a query
 - Identify only the most important pages



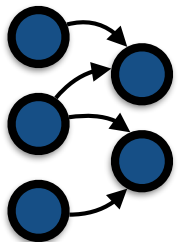
Starting point



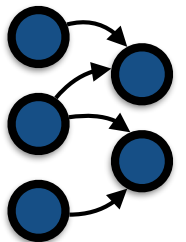
Starting point



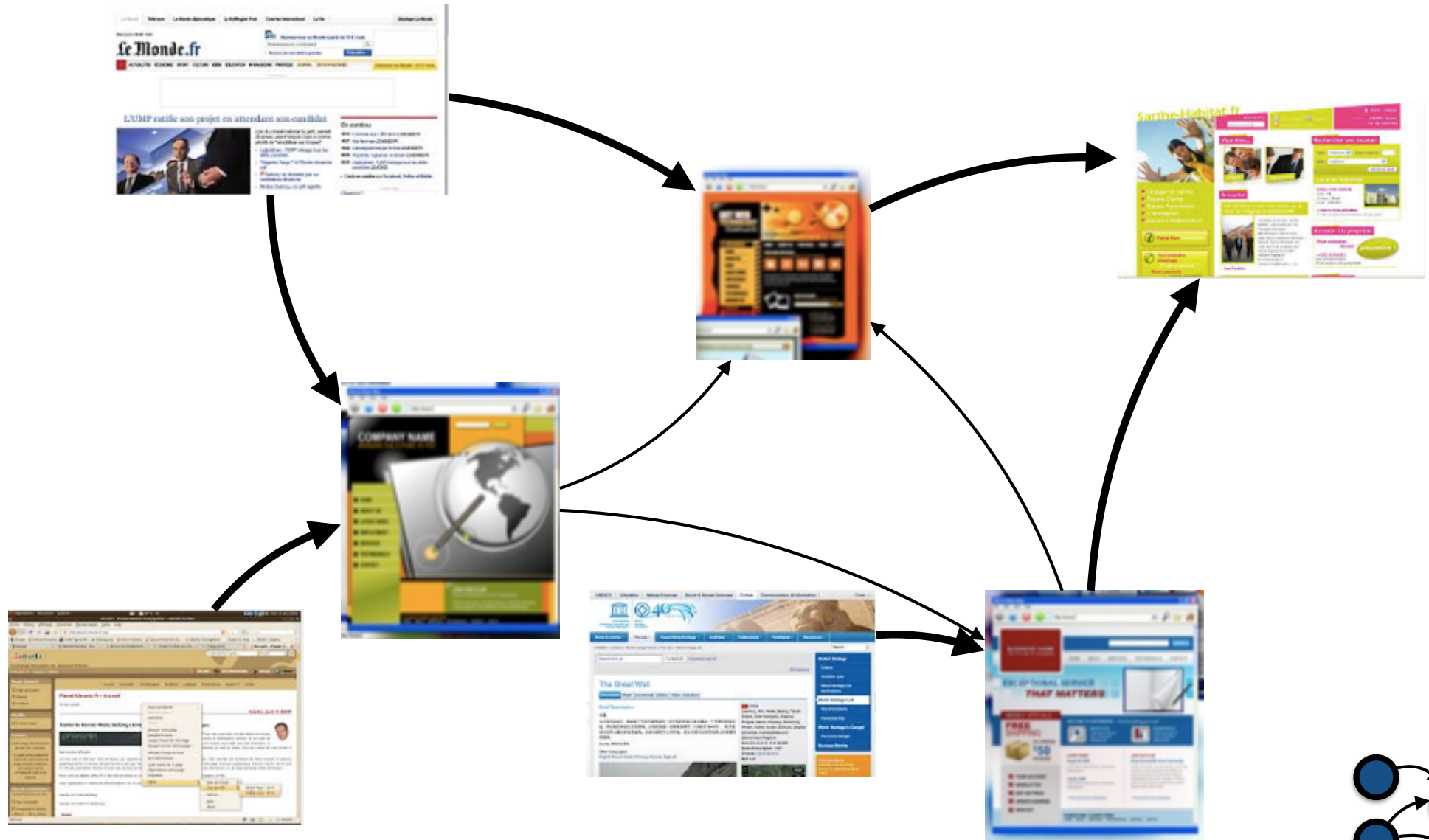
Pages that contain the keywords searched



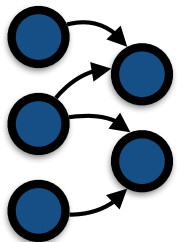
Starting point



Starting point



Enlarging to all pages citing or cited by these pages



Define the analysis graph

Subgraph(s,E,t,d)

s : query string

E : a text-based search engine.

t , d : natural numbers.

Let R denote the top t results of E on s.

Set S := R

For each page p of R

Let $\Gamma^+(p)$ denote the set of all pages p points to.

Let $\Gamma^-(p)$ denote the set of all pages pointing to p .

Add all pages in $\Gamma^+(p)$ to S .

If $|\Gamma^-(p)| < d$ then

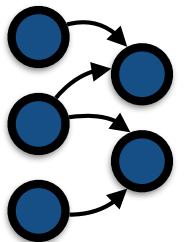
Add all pages in $\Gamma^-(p)$ to S .

Else

Add an arbitrary set of d pages from $\Gamma^-(p)$ to S .

End

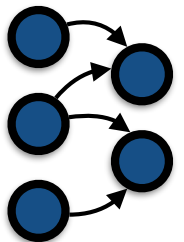
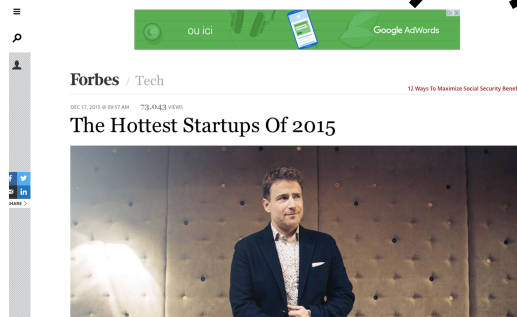
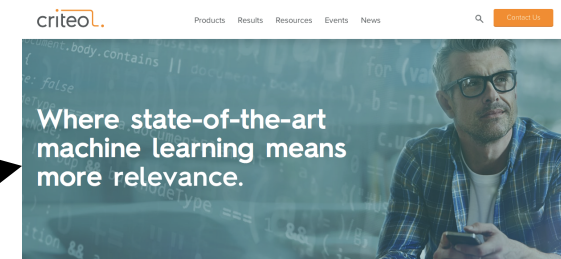
Return S



Principe : Hubs et Authorities

Hubs

Authorities

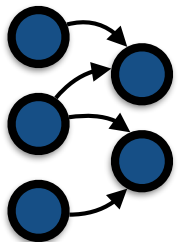
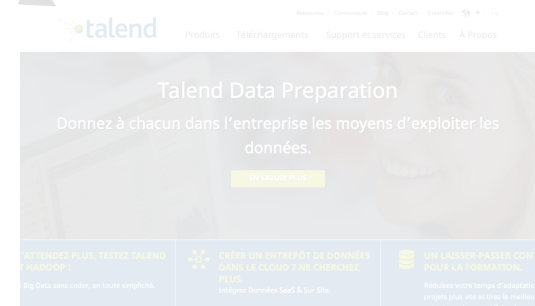
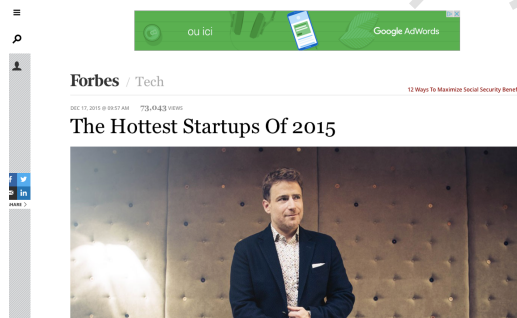
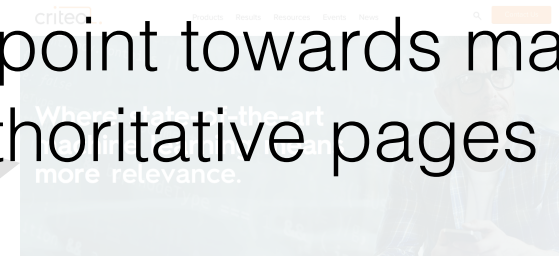
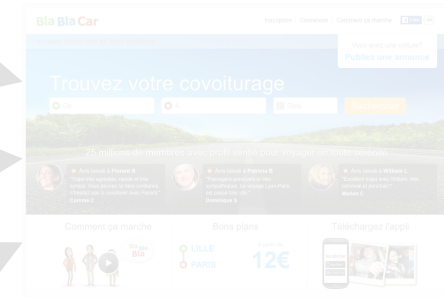


Principe : Hubs ?

Hubs

Authorities

Hubs point towards many authoritative pages

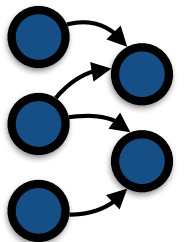
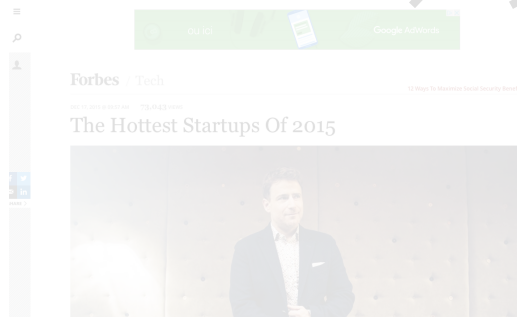
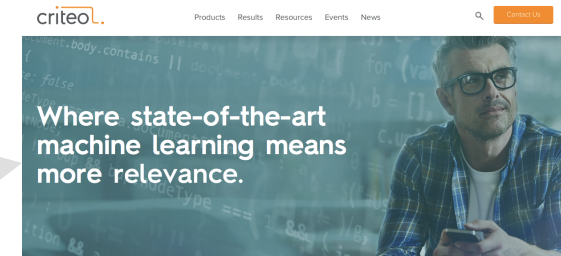


Principe : authorities?

Authorities

Hubs

Authorities are pages that receives many incoming links from hubs

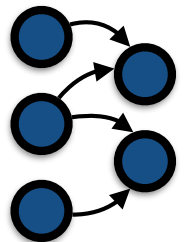
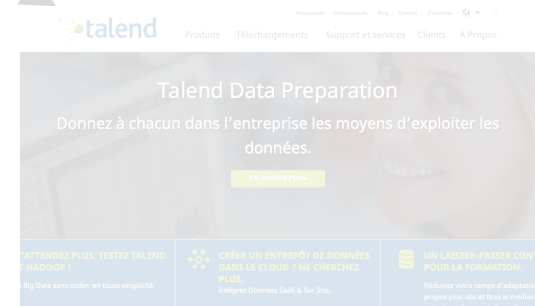
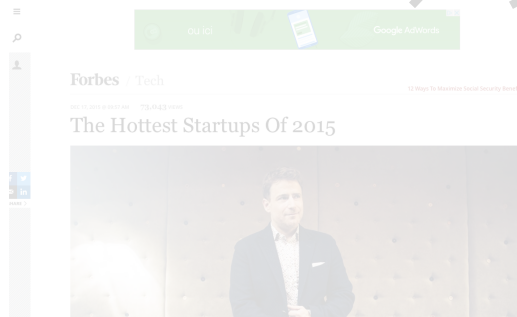
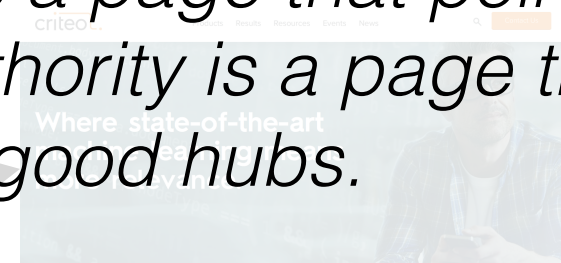
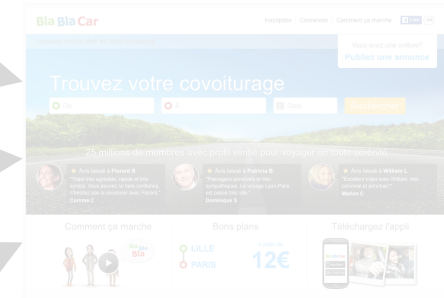


Algorithme paradigm

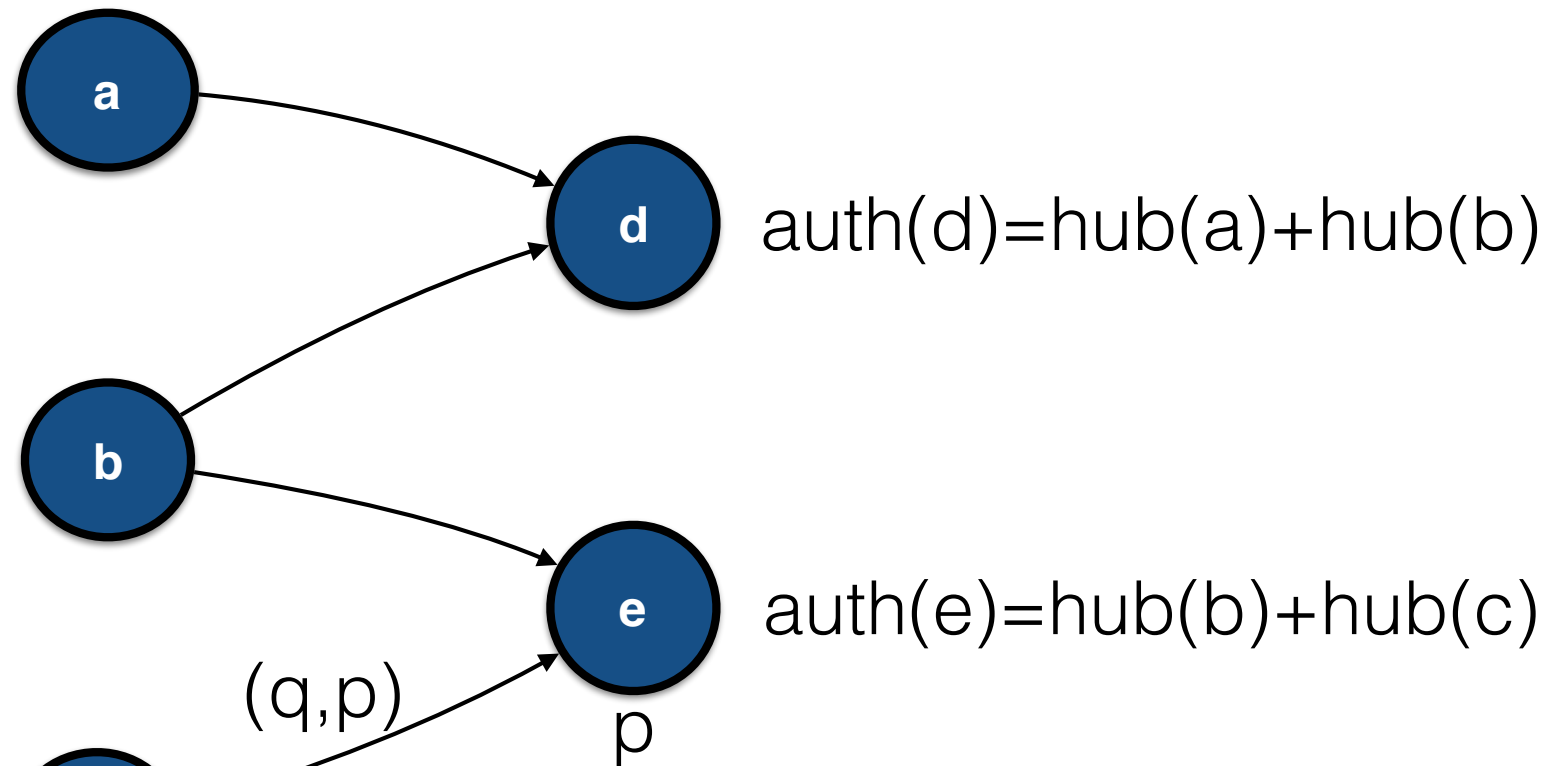
Autorités

Hubs

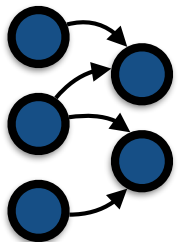
Mutual reinforcement: A good hub is a page that points to many good authorities. A good authority is a page that receives many links from good hubs.



Algorithm : authority score



$$\text{auth}(p) \leftarrow \sum_{q: \{q,p\} \in E} \text{hub}(q)$$



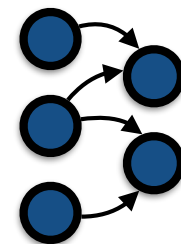
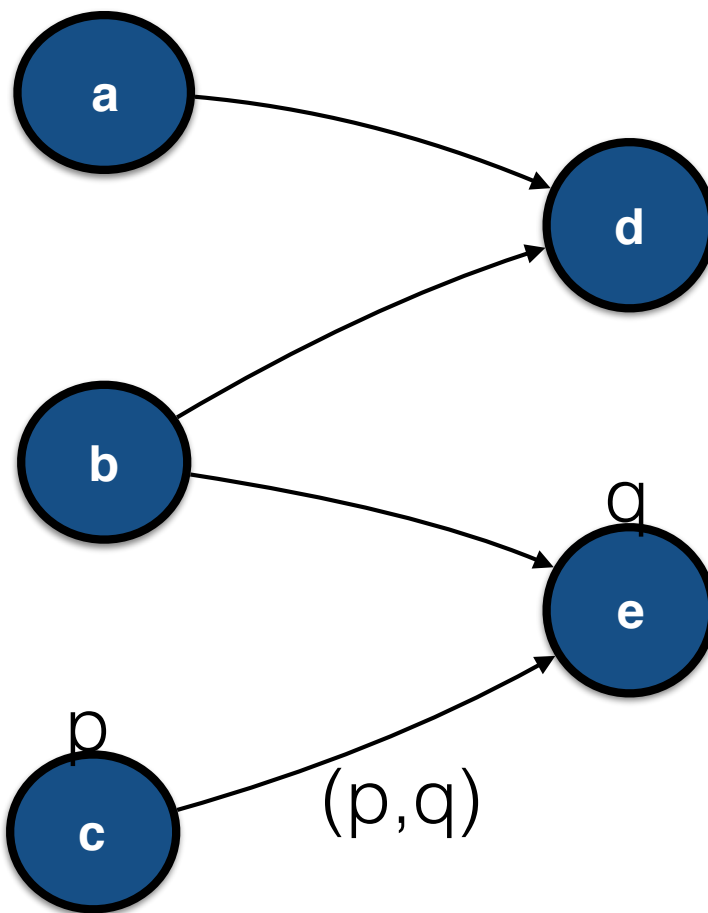
Algorithm : hub score

$$\text{hub}(a) = \text{auth}(d)$$

$$\text{hub}(b) = \text{auth}(d) + \text{auth}(e)$$

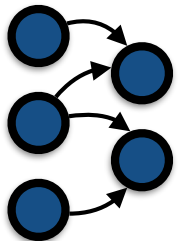
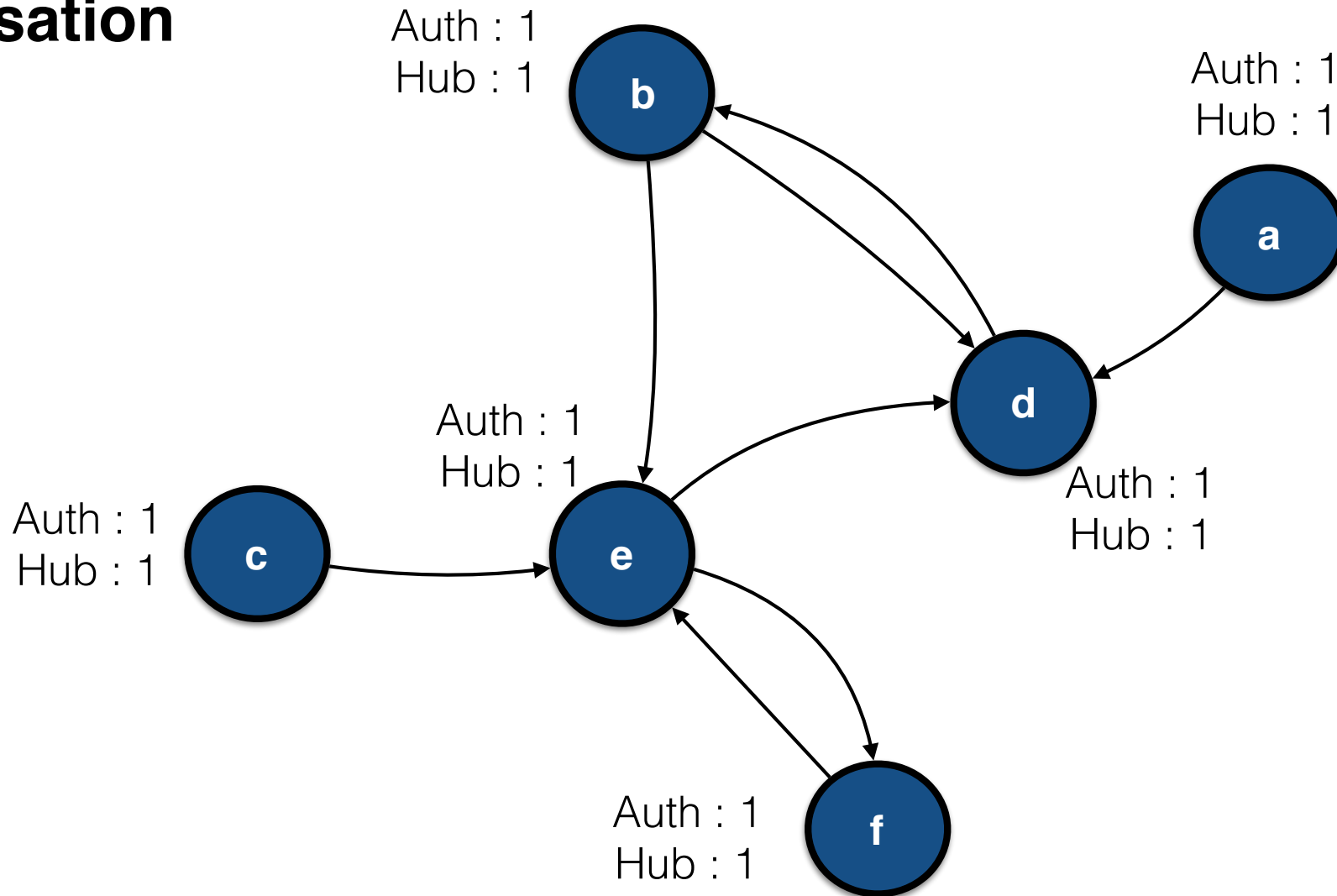
$$\text{hub}(c) = \text{auth}(e)$$

$$\text{hub}(p) \leftarrow \sum_{q: \{p,q\} \in E} \text{auth}(q)$$



Example

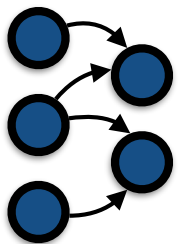
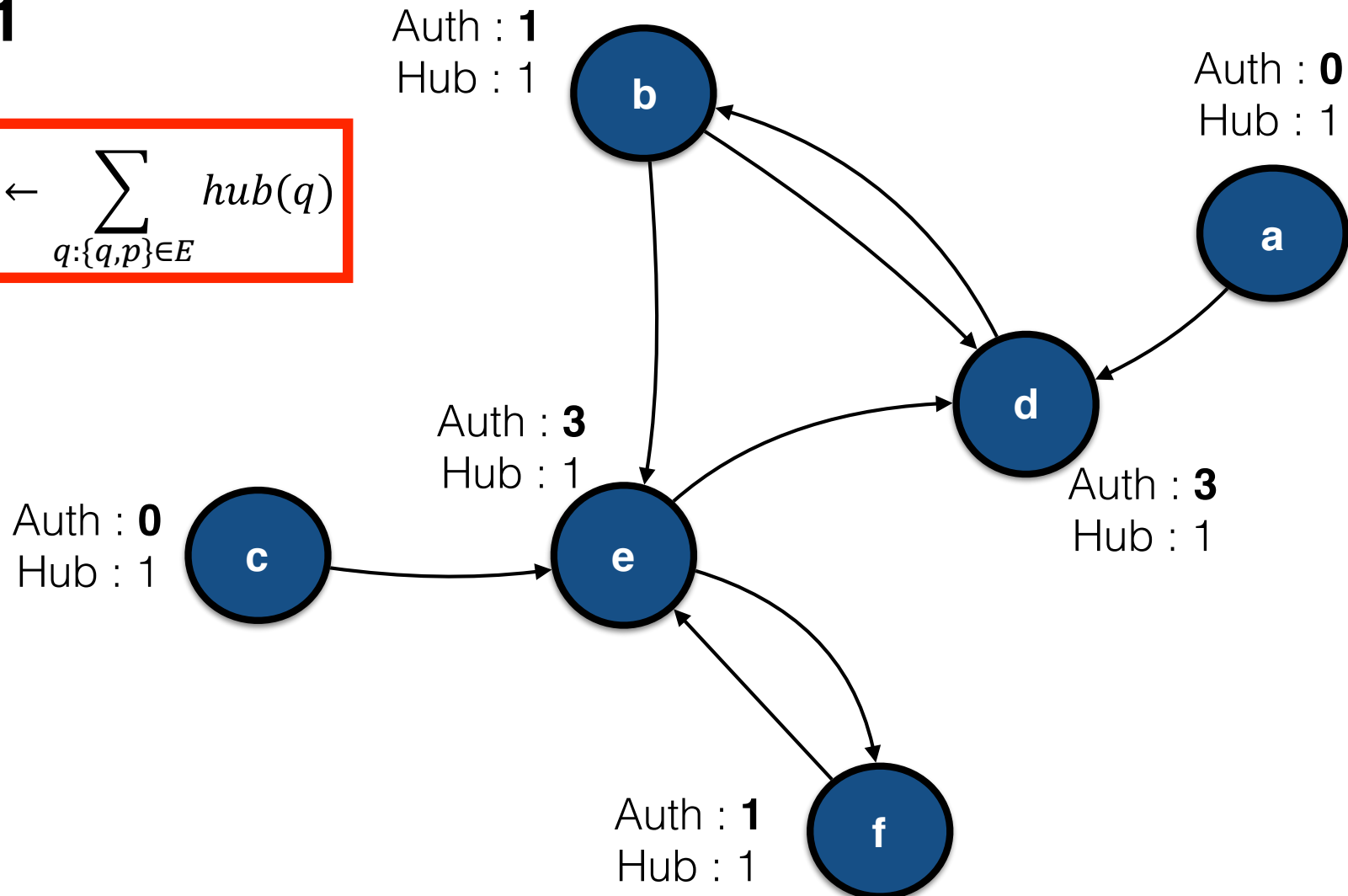
Initialisation



Example

auth_1

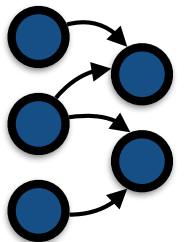
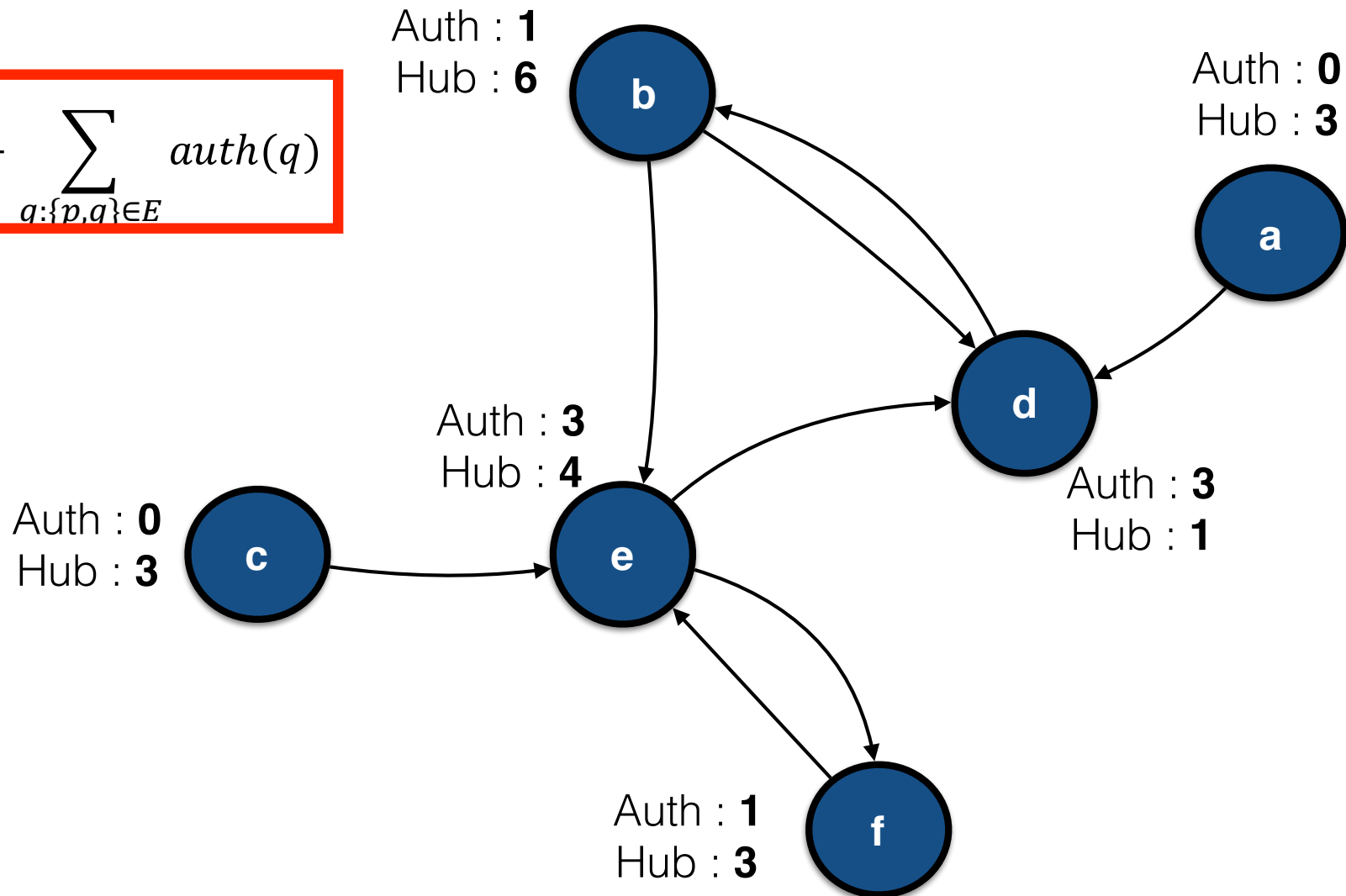
$$auth(p) \leftarrow \sum_{q:\{q,p\} \in E} hub(q)$$



Example

hub_1

$$hub(p) \leftarrow \sum_{q:\{p,q\} \in E} auth(q)$$



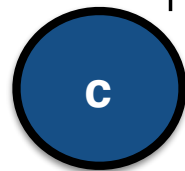
Example

Normalisation

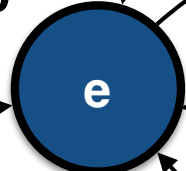
$$\sum_i \text{hub}(i)^2 = 1$$

$$\sum_i \text{auth}(i)^2 = 1$$

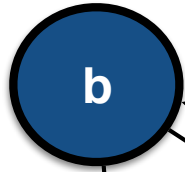
Auth : 0
Hub : $3/\sqrt{80}$



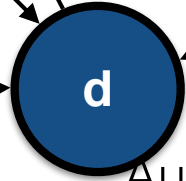
Auth : $3/\sqrt{20}$
Hub : $4/\sqrt{80}$



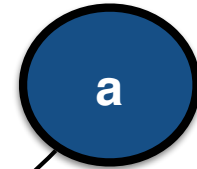
Auth : $1/\sqrt{20}$
Hub : $6/\sqrt{80}$



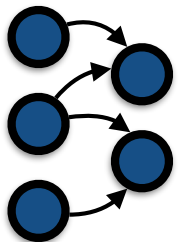
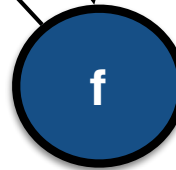
Auth : $3/\sqrt{20}$
Hub : $1/\sqrt{80}$



Auth : 0
Hub : $3/\sqrt{80}$



Auth : $1/\sqrt{20}$
Hub : $3/\sqrt{80}$



Normalisation

$$\sum_i x_i'^2 = 1$$

$$x_i' = \alpha x_i$$

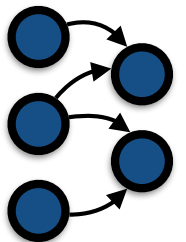
Required for
convergence

$$\sum_i x_i'^2 = \sum_i (\alpha x_i)^2$$

$$\sum_i x_i'^2 = \alpha^2 \sum_i x_i^2$$

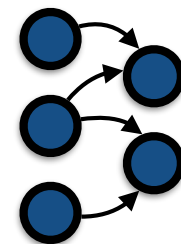
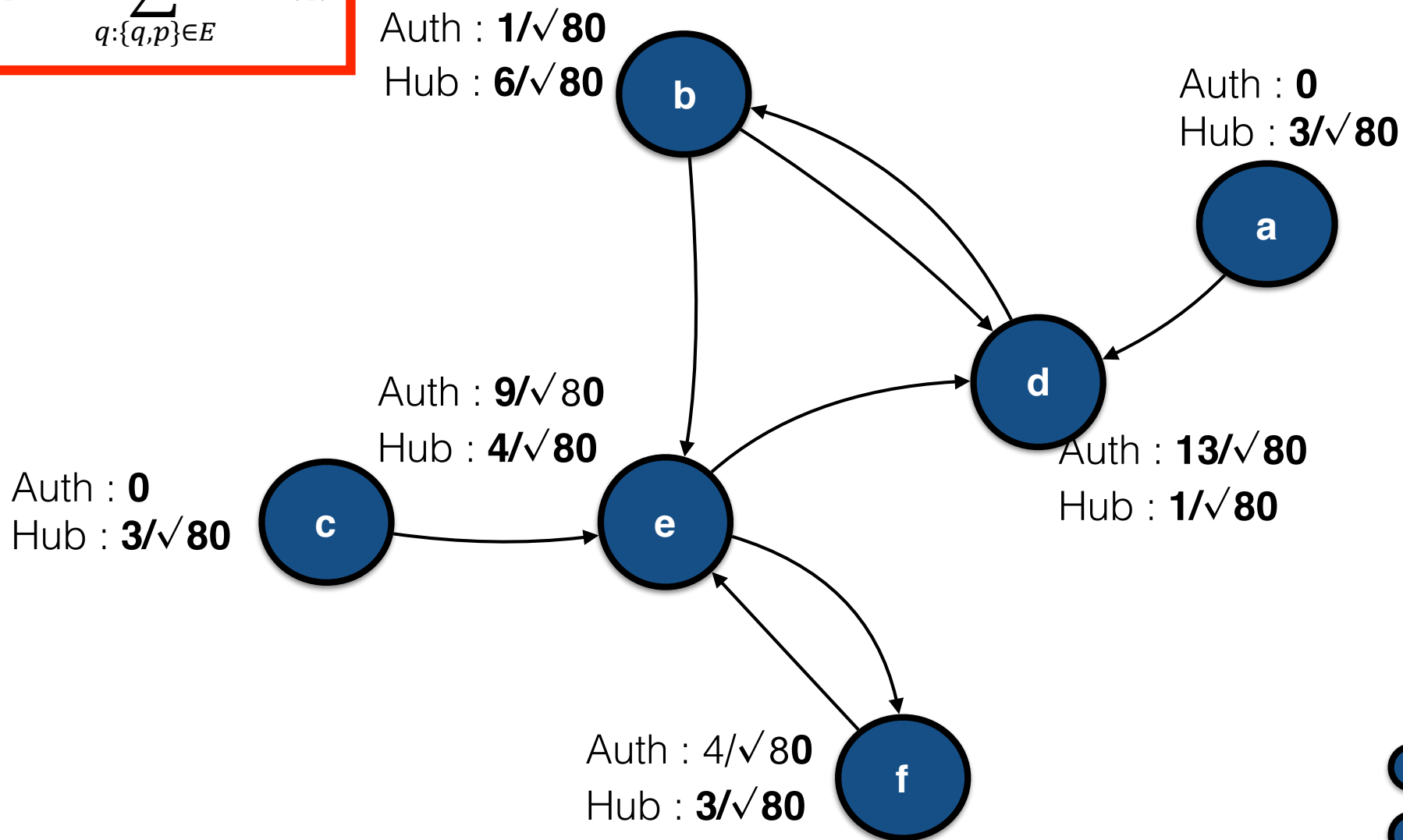
$$\alpha^2 \sum_i x_i^2 = 1 \quad \alpha = \frac{1}{\sqrt{\sum_i x_i^2}}$$

$$x_i' = \frac{x_i}{\sqrt{\sum_i x_i^2}}$$



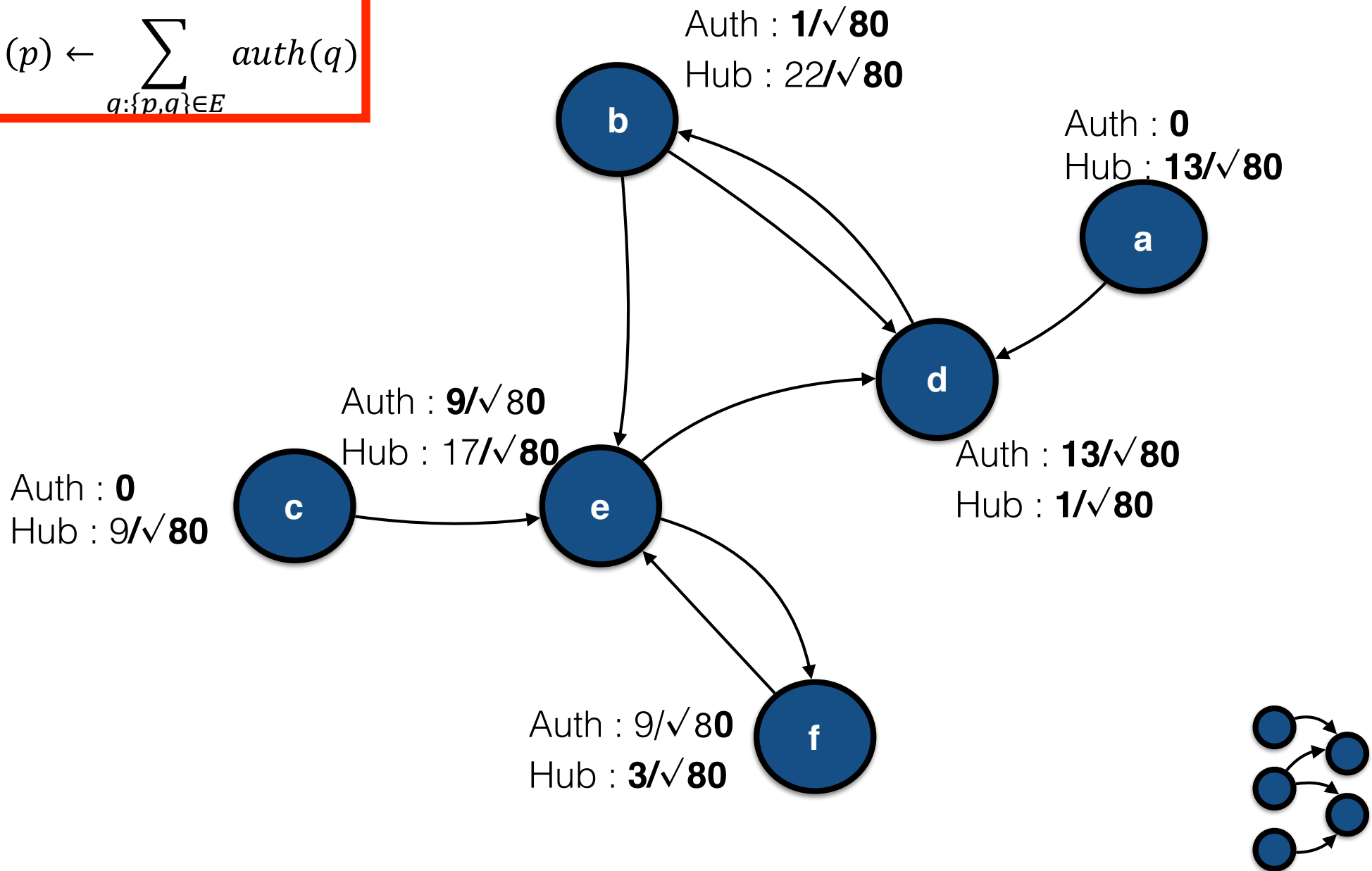
Example

$$auth(p) \leftarrow \sum_{q:\{q,p\} \in E} hub(q)$$



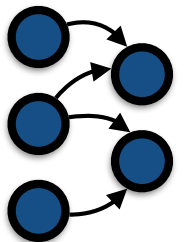
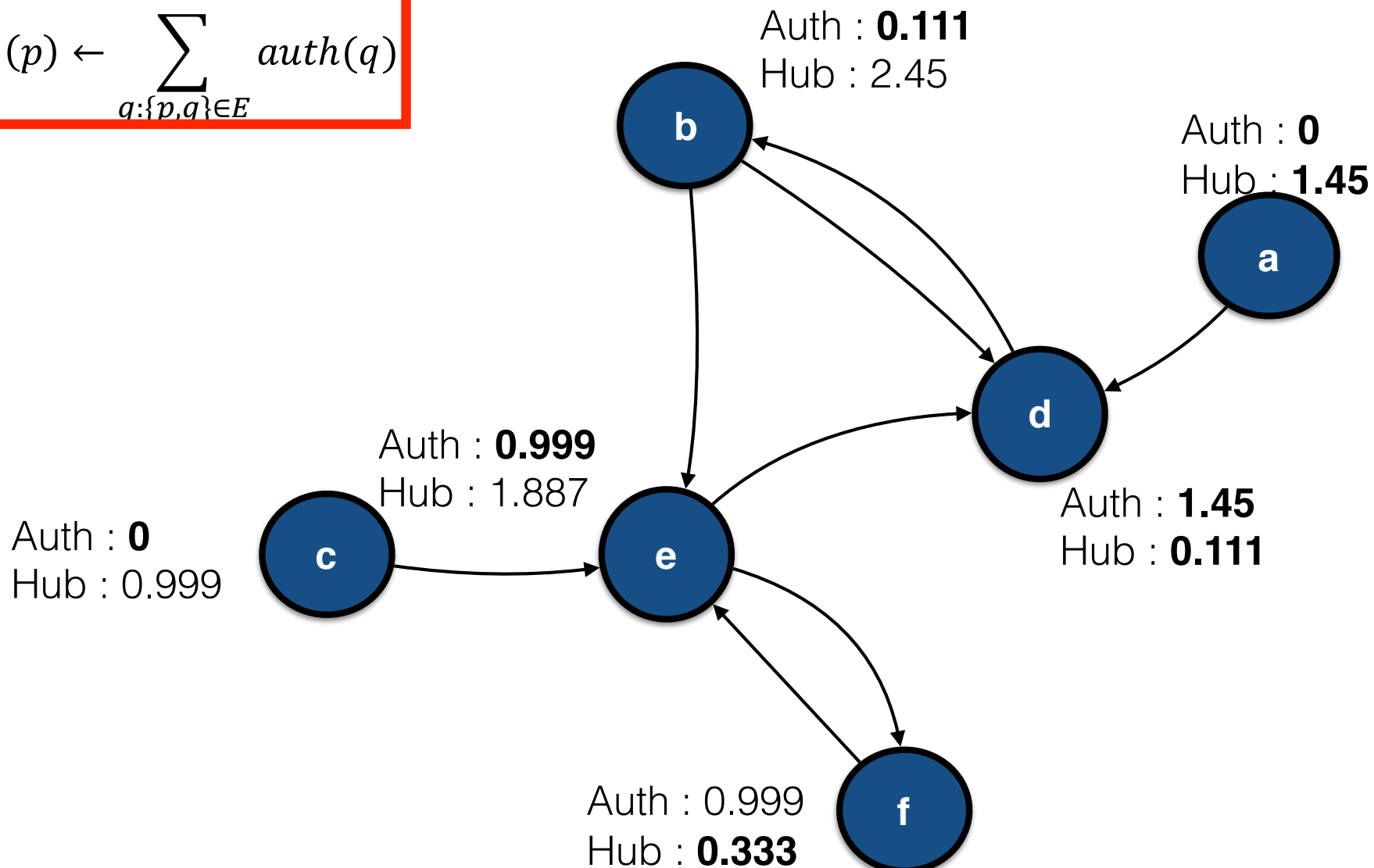
Example

$$hub(p) \leftarrow \sum_{q:\{p,q\} \in E} auth(q)$$



Example

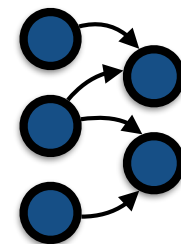
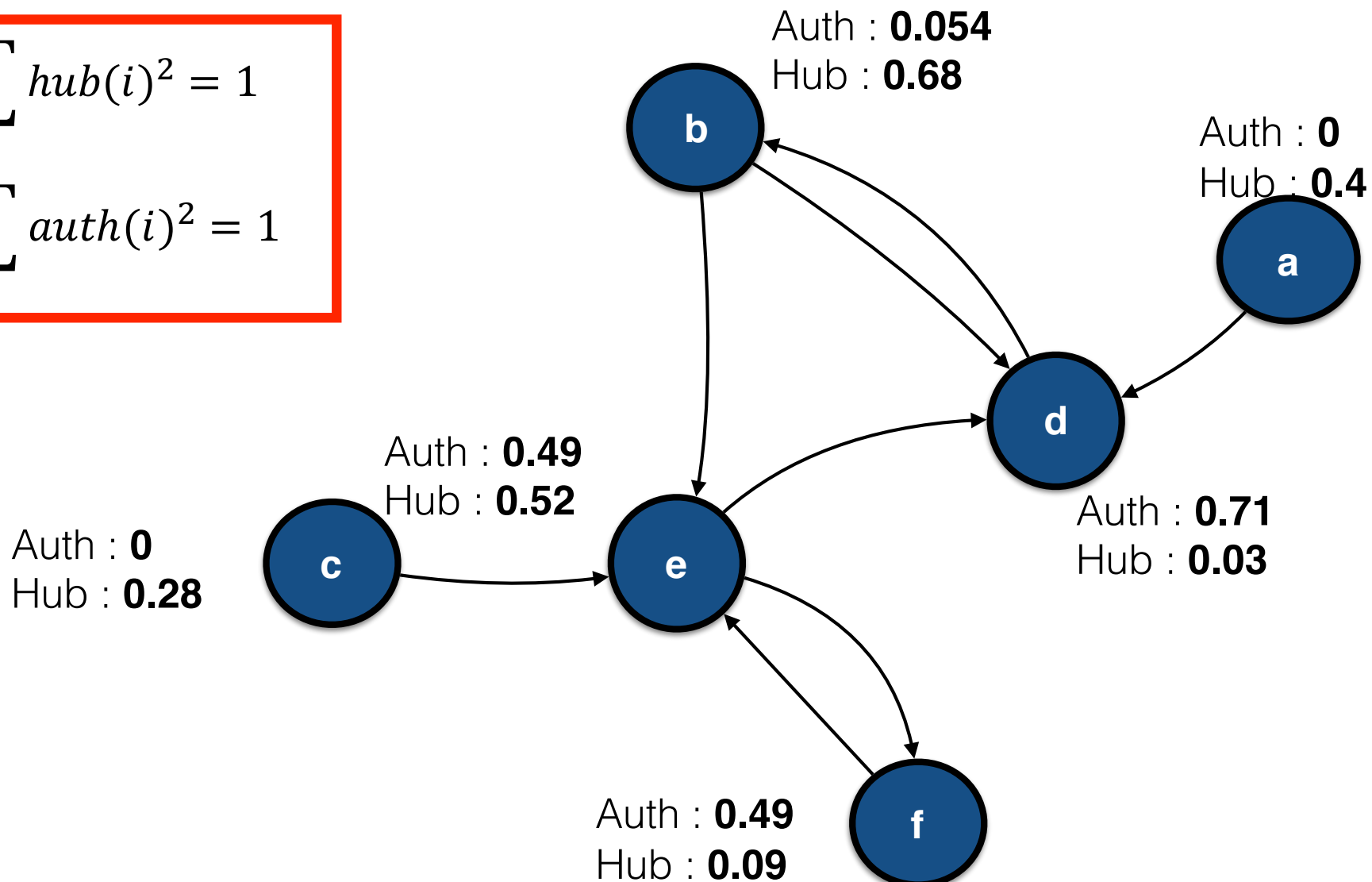
$$hub(p) \leftarrow \sum_{q:\{p,q\} \in E} auth(q)$$



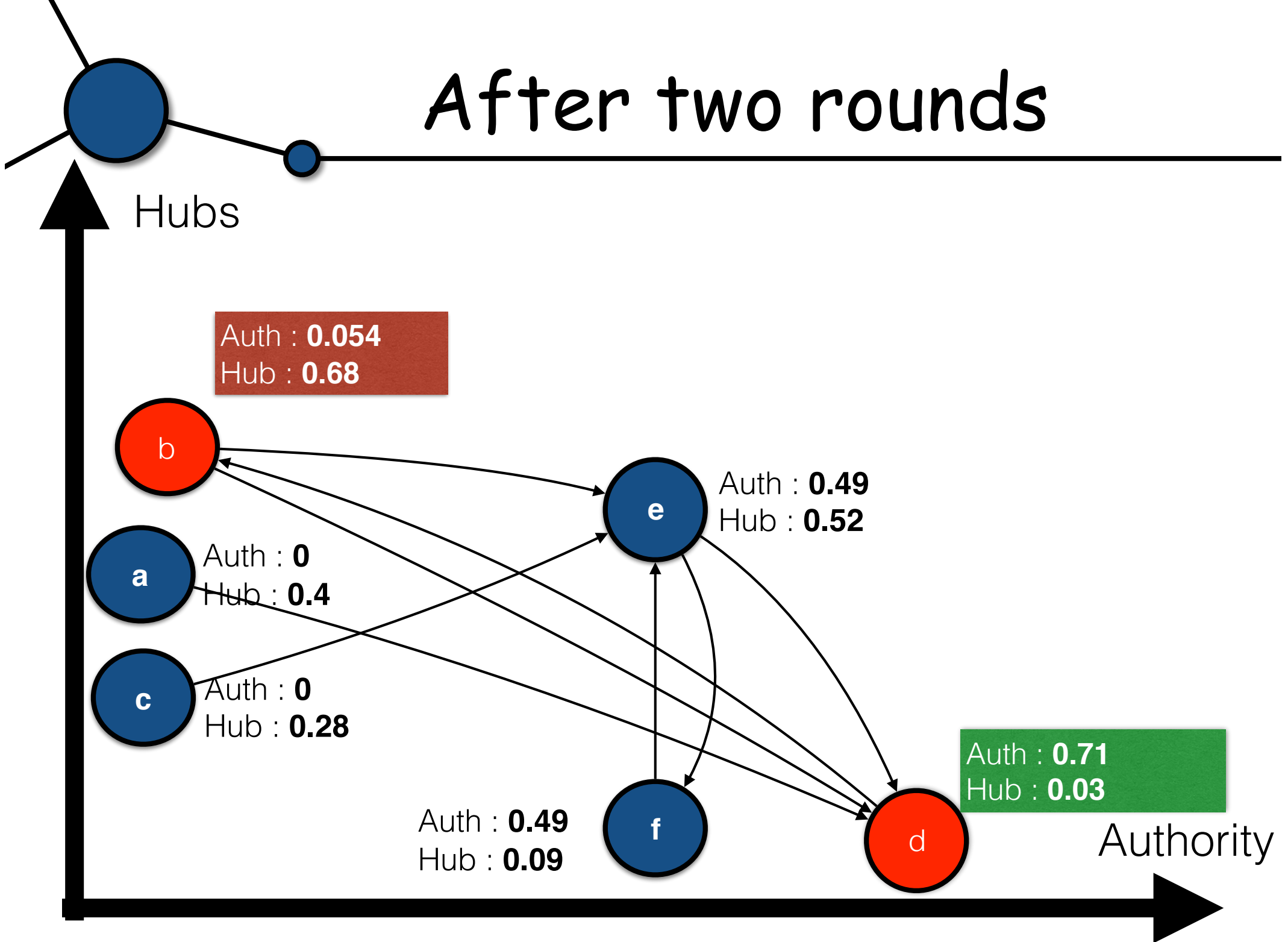
Example

$$\sum_i \text{hub}(i)^2 = 1$$

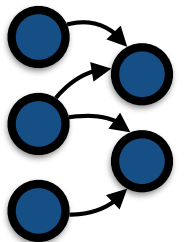
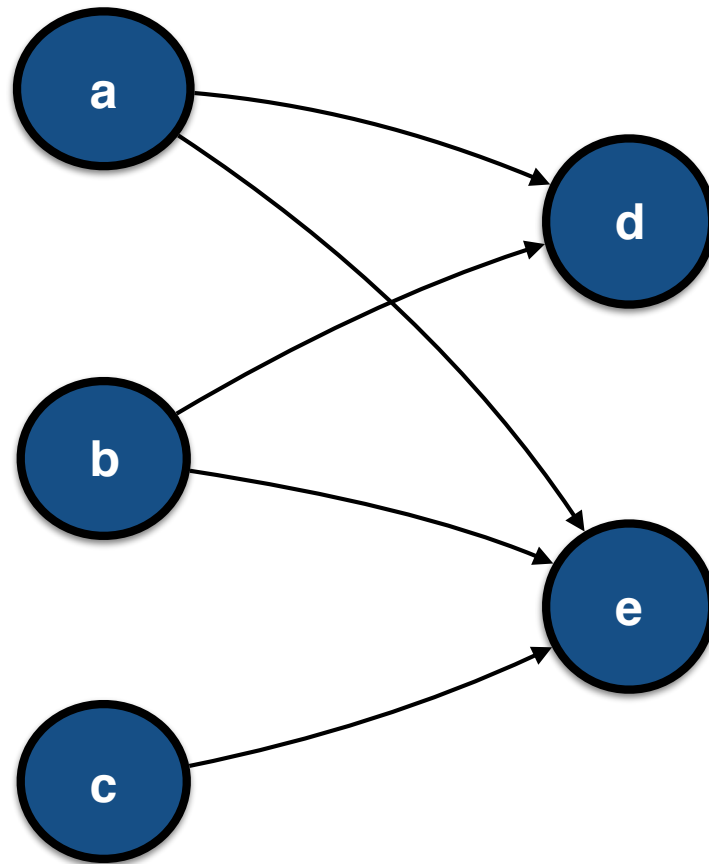
$$\sum_i \text{auth}(i)^2 = 1$$



After two rounds

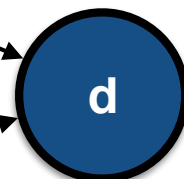
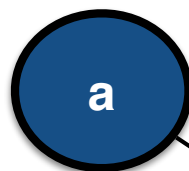


Second example



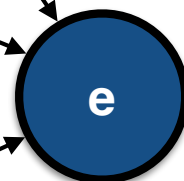
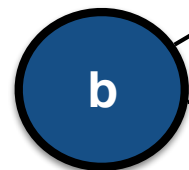
HITS Algorithm - 2 rounds

Auth : **1**
Hub : **1**



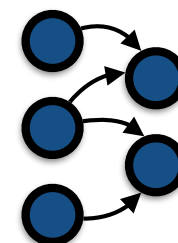
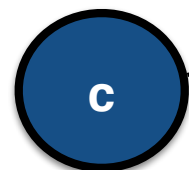
Auth : **1**
Hub : **1**

Auth : **1**
Hub : **1**



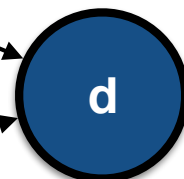
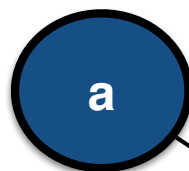
Auth : **1**
Hub : **1**

Auth : **1**
Hub : **1**



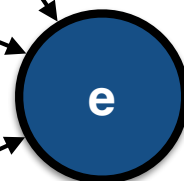
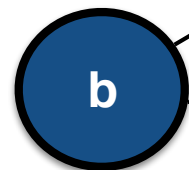
HITS - Auth1

Auth : **0**
Hub : **1**



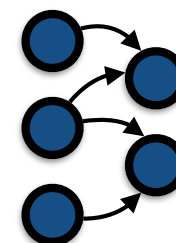
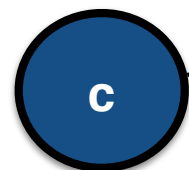
Auth : **2**
Hub : **1**

Auth : **0**
Hub : **1**



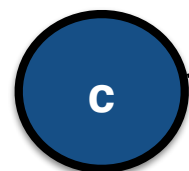
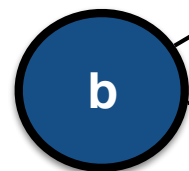
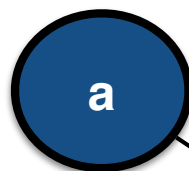
Auth : **3**
Hub : **1**

Auth : **0**
Hub : **1**



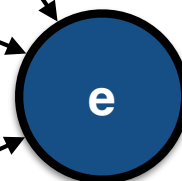
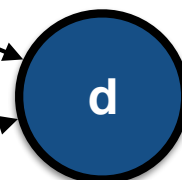
HITS - Hub1

Auth : 0
Hub : 5



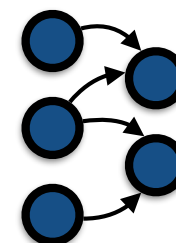
Auth : 0
Hub : 5

Auth : 0
Hub : 3

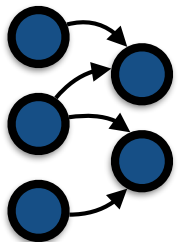
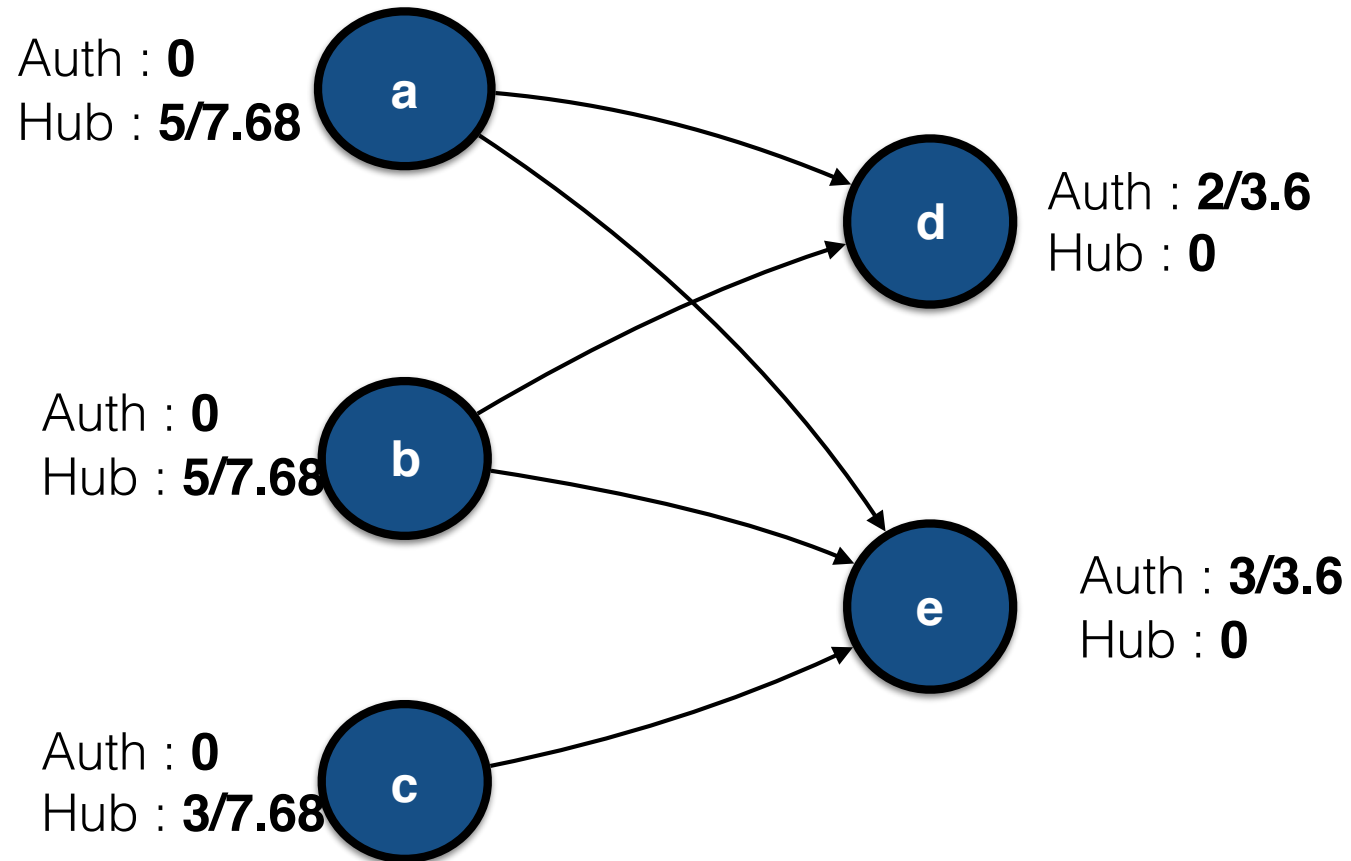


Auth : 2
Hub : 0

Auth : 3
Hub : 0

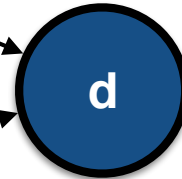
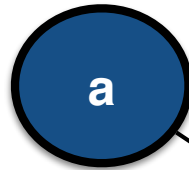


HITS - Normalisation



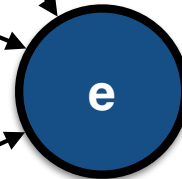
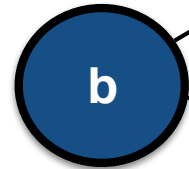
HITS - Normalisation

Auth : **0**
Hub : **0.65**



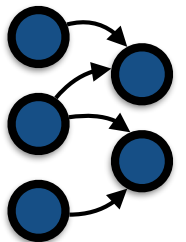
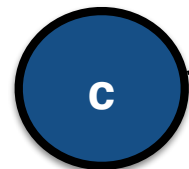
Auth : **0.55**
Hub : **0**

Auth : **0**
Hub : **0.65**



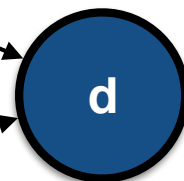
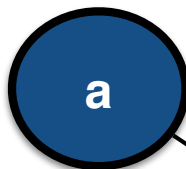
Auth : **0.83**
Hub : **0**

Auth : **0**
Hub : **0.39**



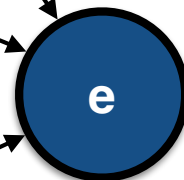
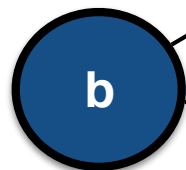
HITS - Auth2

Auth : **0**
Hub : **0.65**



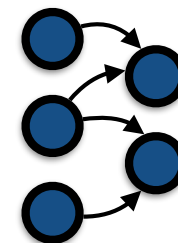
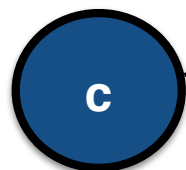
Auth : **1.3**
Hub : **0**

Auth : **0**
Hub : **0.65**



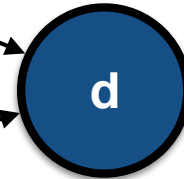
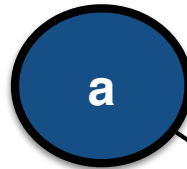
Auth : **1.69**
Hub : **0**

Auth : **0**
Hub : **0.39**



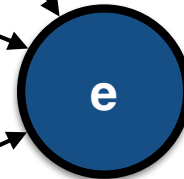
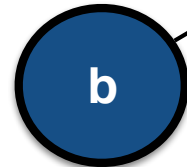
HITS - Hub2

Auth : **0**
Hub : **2.99**



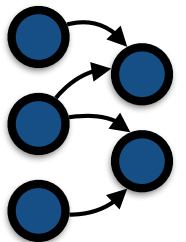
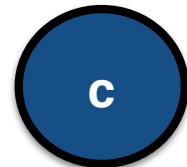
Auth : **1.3**
Hub : **0**

Auth : **0**
Hub : **2.99**



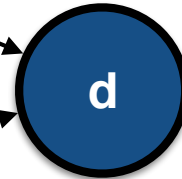
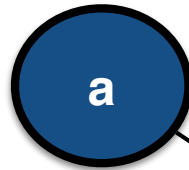
Auth : **1.69**
Hub : **0**

Auth : **0**
Hub : **1.69**



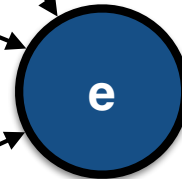
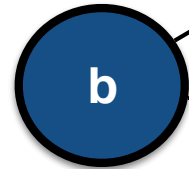
HITS - Normalisation

Auth : **0**
Hub : **0.65**



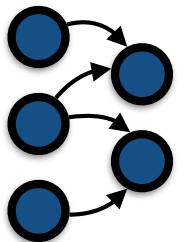
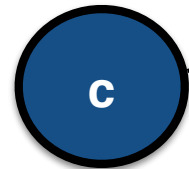
Auth : **0.61**
Hub : **0**

Auth : **0**
Hub : **0.65**



Auth : **0.79**
Hub : **0**

Auth : **0**
Hub : **0.37**



HITS Pseudo-algorithm

Iterate(G, k)

G : a collection of n linked pages

k : a natural number

Let z denote the vector $(1, 1, 1, \dots, 1) \in \mathbf{R}^n$.

Set $x_0 := z$.

Set $y_0 := z$.

For $i = 1, 2, \dots, k$

Apply the $\mathcal{I}$ operation to (x_{i-1}, y_{i-1}) , obtaining new x -weights x'_i .

Apply the $\mathcal{O}$ operation to (x'_i, y_{i-1}) , obtaining new y -weights y'_i .

Normalize x'_i , obtaining x_i .

Normalize y'_i , obtaining y_i .

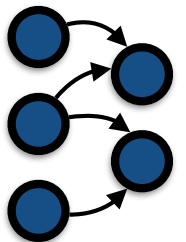
End

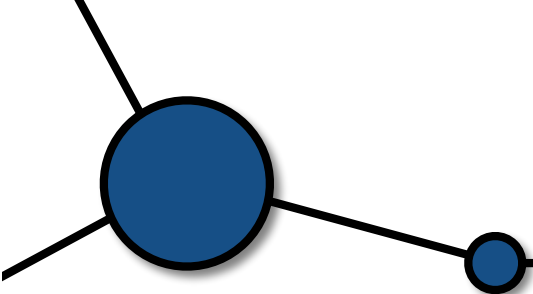
Return (x_k, y_k) .

$$\mathcal{I} \text{ auth}(p) \leftarrow \sum_{q: \{q,p\} \in E} \text{hub}(q)$$

$$\mathcal{O} \text{ hub}(p) \leftarrow \sum_{q: \{p,q\} \in E} \text{auth}(q)$$

x = auth
y = hub





Conclusion

- Many different approaches for measuring centrality
 - Degree : number of links
 - Betweenness : intermediary
 - Closeness : proximity to others
 - HITS : identify hubs & authorities
 - PageRank : measure relative importance
- Choice of a metric depends on each situation

