CSEDU 2018
10TH INTERNATIONAL CONFERENCE ON COMPUTER SUPPORTED EDUCATION
FUNCHAL, MADEIRA - PORTUGAL / 15 - 17 MARCH, 2018

Evaluating the Complementarity of Communication Tools for Learning Platforms

Leonardo Carvalho[1], Laura Assis[1], Leonardo Lima[1], Eduardo Bezerra[1], Gustavo Guedes[1], Artur Ziviani[2], Fabio Porto[2], Rafael Barbastefano[1], Eduardo Ogasawara[1]

[1] CEFET/RJ – Federal Center for Technological Education of Rio de Janeiro
[2] LNCC – National Laboratory for Scientific Computing

CEFET/RJ

Laboratório Nacional de Computação Científica

Agenda

- Introduction
- Related work
- Background
- MGF: Mixed Graph Framework
- Experimental evaluation
- Final remarks

Learning Platforms (LP)

- Many Learning Platforms (LP) (ex.: Moodle)
  - Specialized features: Instant messaging, wikis, social applications
- Communications tools are constantly evolving
  - Emerge new features
    - Comments, private messages, blogging, media file sharing
    - Support for mobile devices
  - Reduce barriers among students and between student-teacher
- Commonly LP tools have their own social features
  - Due to security, pedagogical decisions
  - Features of web 2.0/3.0 similar to Facebook, LinkedIn

Choices for LP

- Choice for a particular LP can be time-consuming and expensive
- Measure the effectiveness of New Communication Tool (NCT)
  - Check if NCT brings benefits to LP
  - How NCT is providing a complementary communication flow with respect to the Current Communication Tool (CCT)

Problem statement / approach

- Problem
  - Measure the complementarity of a NCT when CCT is already established in a LP
- Proposal
  - Mixed Graph Framework (MGF) to evaluate the complementariness of CCT with respect to NCT
  - CCT and NCT are modeled as graphs, respectively G_c and G_n
  - Create a Mixed Graph G_m
  - Measure G_m with respect to G_c

Related work

- Analysis of social networks
  - Widely studied for many years
  - Analyze the structure and dynamics of networks
- In educational environments
  - Research is expanding
  - Usage of social networks in LP
  - Learning and teaching achievements

Related work

- Being social?
  - Data mining on network metrics extracted from information flow modeled as graphs
  - Identification of groups (clusters and cliques)
  - Metrics such as cohesion and average distance used in Network Science to gain insights
- Usage on distance learning education
  - Analysis of communication flow of students to draw conclusions and improve the e-learning courses
- Frameworks for understanding social media
  - User contributions behavior and interrelationship
- Mixed Graph Framework is novel

Background: Graph Representation

- Graph G(V, E)
- Vertices: $i \in V$
- Edges: $(i,j) \in E$
  - Adjacent matrix (A)
- $w_{i,j}$ is the communication flow between i, j
  - Weighed adjacent matrix
- Directed graph – $w_{i,j} \neq w_{j,i}$

Background: Graph Centrality Measures

- Closeness: how close a member is to the others
  - $C_c(v) = \frac{1}{\sum_{x \in V \backslash \mathrm{v}} d(v,x)}$
  - $d(v, x)$ is the distance between nodes
- Betweenness: summarize if a vertex is between other pair of vertices
  - $C_b(v) = \sum_{s \neq t \neq v \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}$
  - $\sigma(s, t)$ is the number of minimum paths connecting s, t
- Kleinberg centrality: identify important members
  - Hubs and Authorities
  - Eigenvectors of AA^T and A^TA

Background: Statistical Analysis

- Distribution
  - Parametric (normal distribution)
  - Non-parametric (scale-free with power-law)
- Comparison of samples
  - Distribution
    - Wilcoxon Rank Sum Test
  - Correlation
    - Spearman Rank Correlation Test

MGF: Mixed Graph Framework

Algorithm 1 Main *MGF* Algorithm

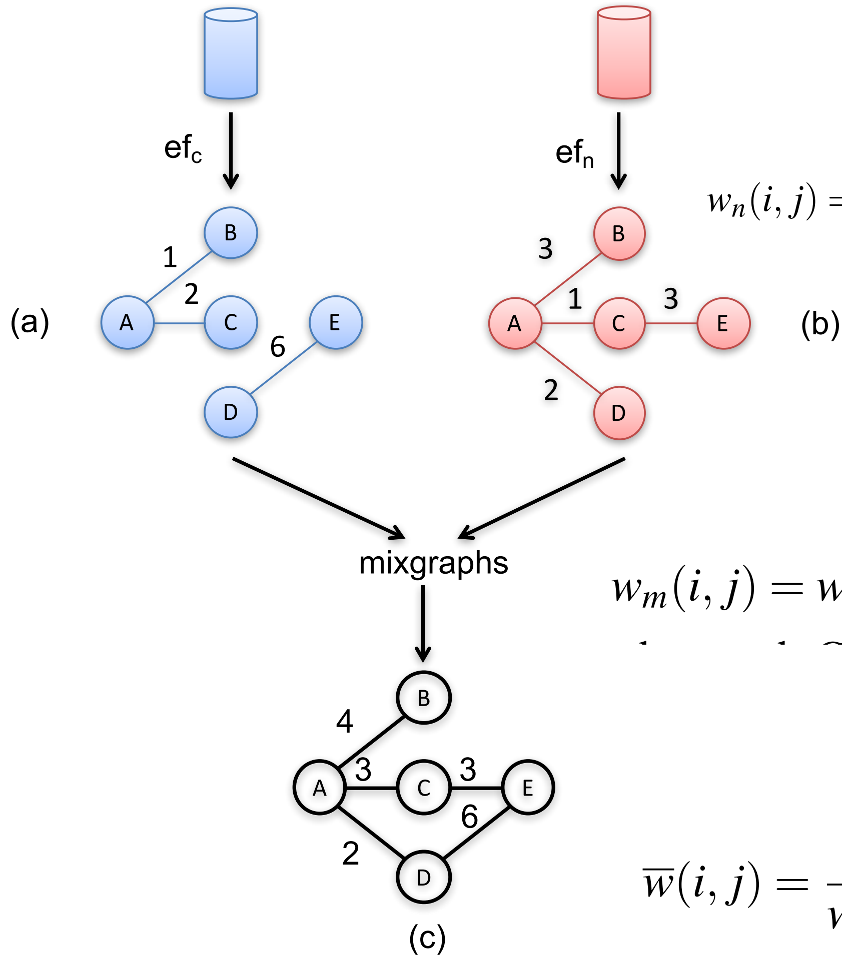
function $MGF(D\ d_c, D\ d_n, ef_c, ef_n)$
    $G_c \leftarrow fExtract_c(d_c)$
    $G_n \leftarrow fExtract_n(d_n)$
    $G_m \leftarrow fMix(G_c,\ G_n)$
    return $fAnalyze(G_c,\ G_m)$
end function

function $fAnalyze(G_c,\ G_m)$
    $r_1 \leftarrow analyzeClosenessDist(G_c,\ G_m)$
    $r_2 \leftarrow analyzeClosenessCorr(G_c,\ G_m)$
    $r_3 \leftarrow analyzeBetweennessCorr(G_c,\ G_m)$
    $r_4 \leftarrow analyzeEigenTopK(G_c,\ G_m)$
    return $\{r_1, r_2, r_3, r_4\}$
end function

MGF: Toy example

$$w_c(i,j) = |posts(i,j)|$$

$$w_n(i,j) = \frac{\beta|comments(i,j)| + \gamma|likes(i,j)|}{\beta + \gamma}$$



$$w_m(i,j) = w_c(i,j) + w_n(i,j)$$

$$\overline{w}(i,j) = \frac{1}{w(i,j)}$$

Figure 1: Communication flow: (a) G_c extracted from the CCT dataset; (b) G_n extracted from NCT dataset; (c) G_m produced by mixing G_c with G_n

MGF – Complementarity analysis

Algorithm 3 Analysis of Centrality

function $analyzeClosenessDist(G_c,\ G_m)$
  $vc_c \leftarrow closeness(convertDist(G_c))$
  $vc_m \leftarrow closeness(convertDist(G_m))$
  return $wilcox.test(vc_m, vc_m)$
end function

function $analyzeClosenessCorr(G_c,\ G_m)$
  $vc_c \leftarrow closeness(convertDist(G_c))$
  $vc_m \leftarrow closeness(convertDist(G_m))$
  return $spearman.cor.test(vc_m, vc_m)$
end function

function $analyzeBetweennessCorr(G_c,\ G_m)$
  $vb_c \leftarrow betweenness(convertDist(G_c))$
  $vb_m \leftarrow betweenness(convertDist(G_m))$
  return $spearman.cor.test(vb_c, vb_m)$
end function

- Closeness centrality distribution
  - Intuition is to measure the intensity change of communication when introducing NCT
- Closeness and Betweenness correlation
  - Intuition is to measure if the introduction of NCT changes the way in which people interact concerning CCT
- These sentinels may observe difference signals

Experimental Evaluation
Synthetic data generation

- Gc – hierarchical teacher-students communication
  - Organizational structure
  - Teacher-students or Tutors-students
- Gn – social network communication among students
  - Does not impose an organization structure

Table 1: Parameters used in the experimental evaluation

Parameter	Description
$v_c = v_n$	Number of nodes in both graphs, G_c and G_n
k	Number of groups in G_c
e_c	Number of edges (communication flows) in G_c
e_n	Number of edges (communication flows) in G_n

Experimental Evaluation
Synthetic data generation

Algorithm 4 Synthetic dataset production

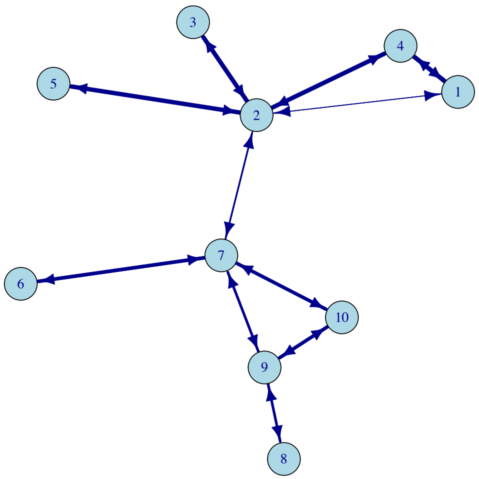
function $SyntheticDatasets(k, v, e, e_n)$
  for all $i \leftarrow 1 \ to \ k$ **do**
    $G_c^i \leftarrow new \ ScaleFreeGraph(v, e)$
    $G_c \leftarrow G_c \cup G_c^i$
  end for
  for $i \leftarrow 1 \ to \ k_E - 1$ **do**
    for $j \leftarrow i + 1 \ to \ k_E$ **do**
      $e_l \leftarrow connect(G_c^i, G_c^j)$
      $E_c \leftarrow E_c \cup e$
    end for
  end for
  $v_c \leftarrow v \cdot k$
  $v_n \leftarrow v_c$
  $G_n \leftarrow new \ ScaleFreeGraph(v_n, e_n)$
  return $(\{G_c, G_n\})$
end function

Network growth

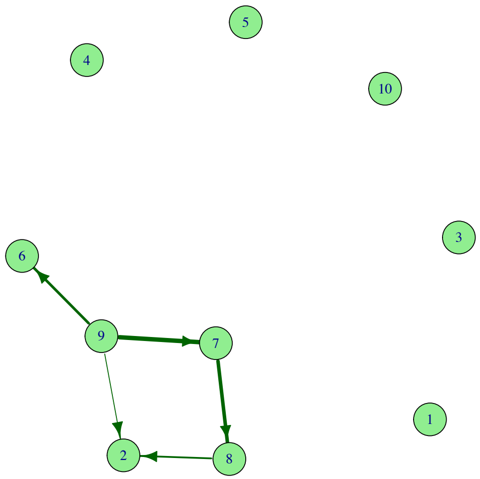
Algorithm 5 Network Growth

```
function NetGrowth(w_c, w_n, r)
    RS ← {}
    for all δ ← 0 to 100 step r do
        w_{n,δ} ← Filter(δ, δ/100 · w_n)
        w_{m,δ} ← fMix(w_c, w_{n,δ})
        RS ← RS ∪ fAnalyze(w_c, w_{m,δ})
    end for
    plotCharts(RS)
end function
```

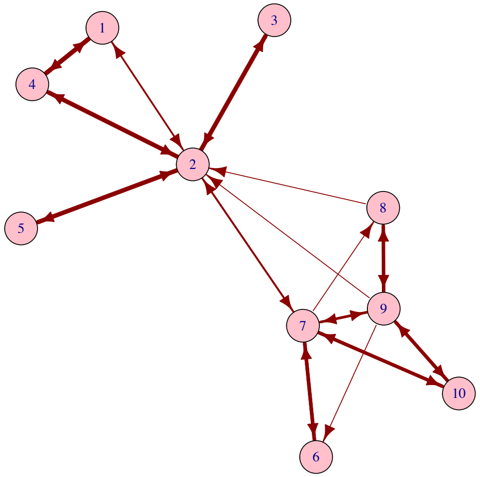
Toy example



$(a)\ G_c(v_c = 10, k_c = 2, e_c = 10)$

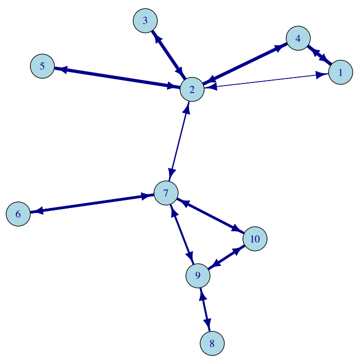


$(b)\ G_n^{25\%}(v_n = 10, e_n = 10)$

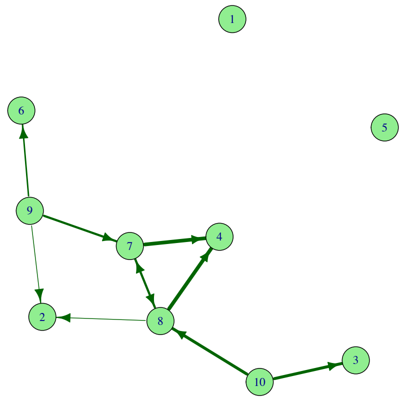


$(c)\ G_m^{25\%} = G_c + G_n^{25\%}$

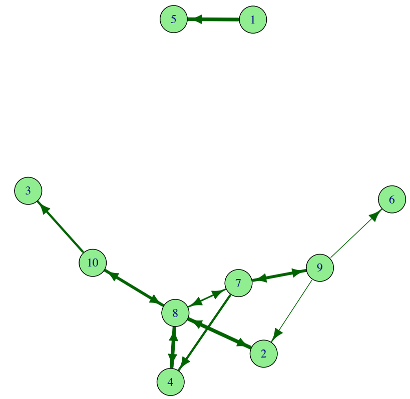
Toy example – NCT growth



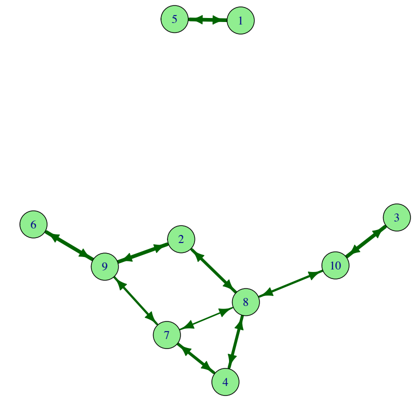
$(a)\ G_c(v_c = 10, k_c = 2, e_c = 10)$



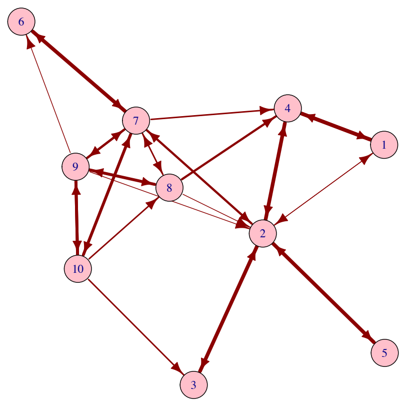
$(d)\ G_n^{50\%}(v_n = 10, e_n = 10)$



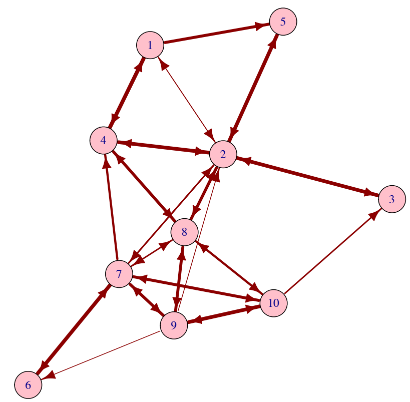
$(e)\ G_n^{75\%}(v_n = 10, e_n = 10)$



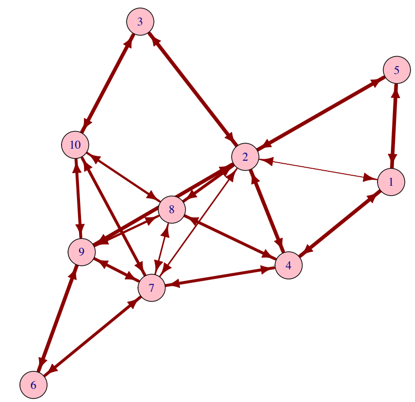
$(f)\ G_n^{100\%}(v_n = 10, e_n = 10)$



$(g)\ G_m^{50\%} = G_c + G_n^{50\%}$

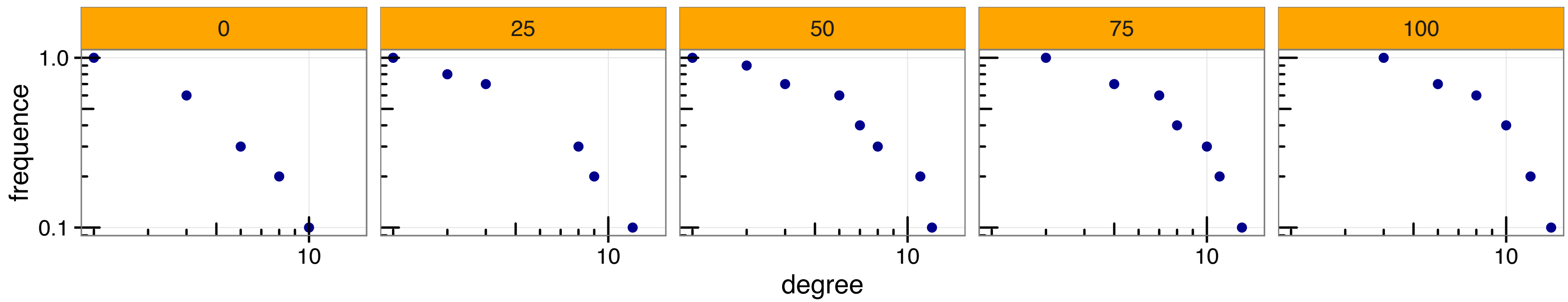


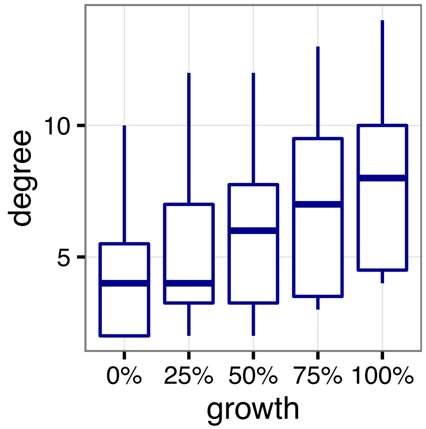
$(h)\ G_m^{75\%} = G_c + G_n^{75\%}$

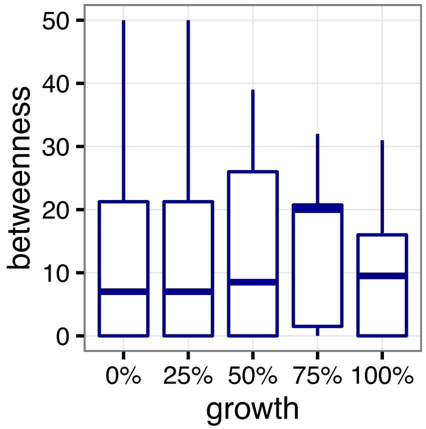


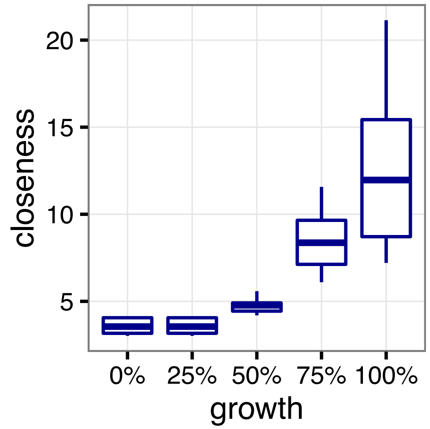
$(i)\ G_m^{100\%} = G_c + G_n^{100\%}$

Analysis of degree, betweenness, closeness

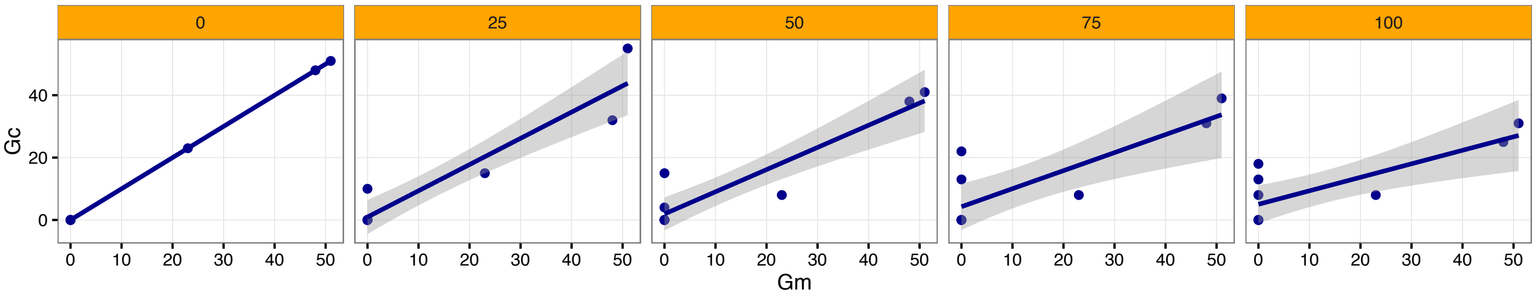




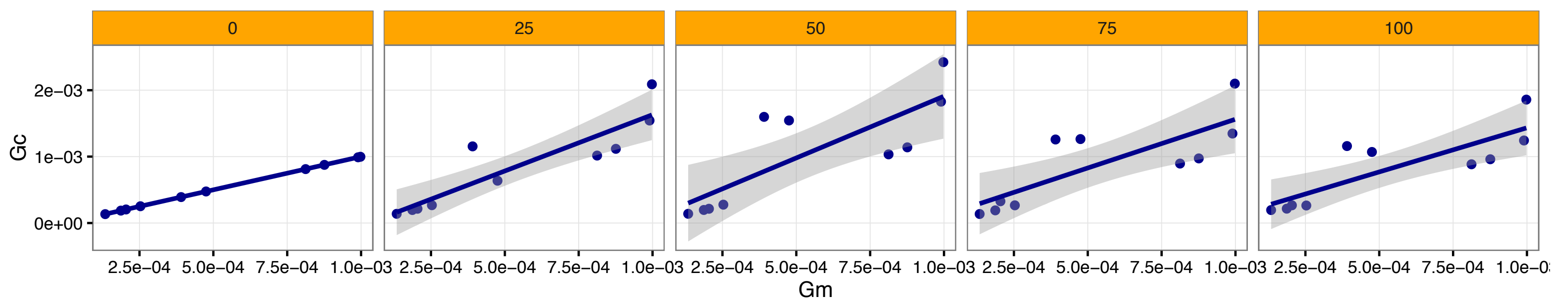




Correlation analysis



Betweenness



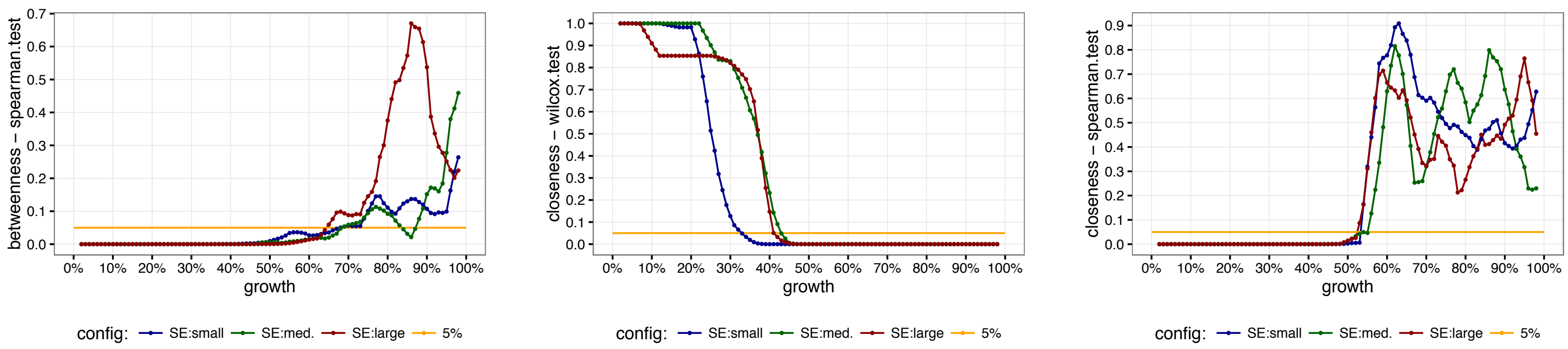
Closeness

Sensitive Analysis

Table 2: LP Scenarios

Scenario	Description
SC (G_n scale)	$v_c = 30,\ k_c = 3,\ e_c = 60$ $small:\ e_n = 25$ $medium:\ e_n = 45$ $large:\ e_n = 55$
SC (G_c groups)	$v_c = 30,\ e_c = 60,\ e_n = 45$ $low:\ k_c = 2$ $moderated:\ k_c = 3$ $high:\ k_c = 5$
MC (G_c groups)	$v_c = 150,\ e_c = 60$ $low:\ k_c = 10,\ e_n = 120$ $moderated:\ k_c = 15,\ e_n = 180$ $high:\ k_c = 25,\ e_n = 300$

Sensitive analysis: Small course
Varying the number of edges in G_n



(a) (b) (c)

Figure 4: Scenario of Small Course - varying number of edges in G_n: betweenness correlation analysis (a), closeness median analysis (b), closeness correlation analysis (c)

Conclusions

- Proposed MGF to analyze if NCT is complementary to CCT
- Evaluated MGF using synthetic data
- Future work
  - Analyze a real-world scenario
  - Analyze the timely evolution of a CCT
  - Analyze the network increase

Contact

CEFET/RJ

Eduardo Ogasawara
eogasawara@ieee.org
http://eic.cefet-rj.br/~eogasawara

Acknowledgements

CNPq
Conselho Nacional de Desenvolvimento Científico e Tecnológico

CAPES

FAPERJ
Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro