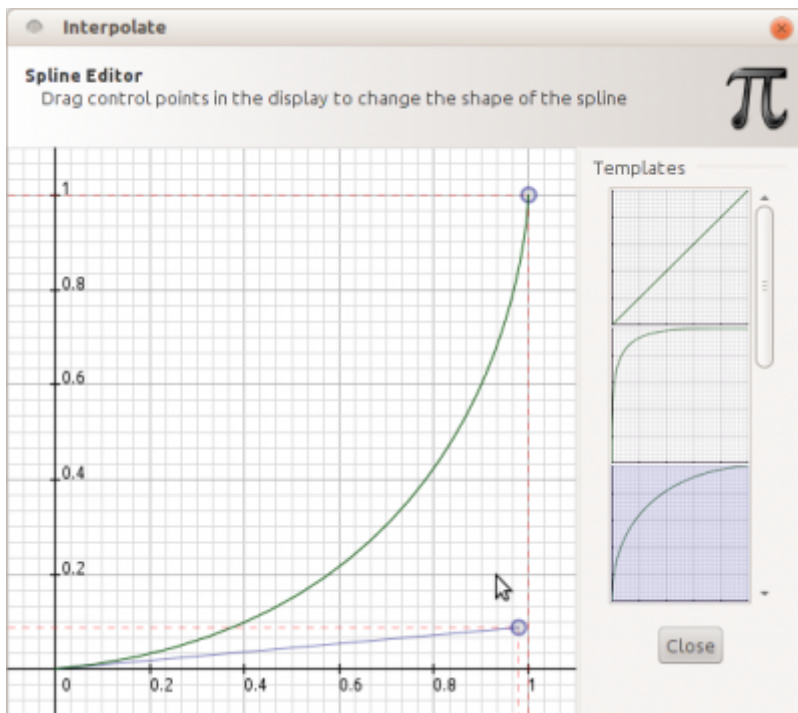


Analiza sieci społecznościowych w narzędziu Gephi (część I)

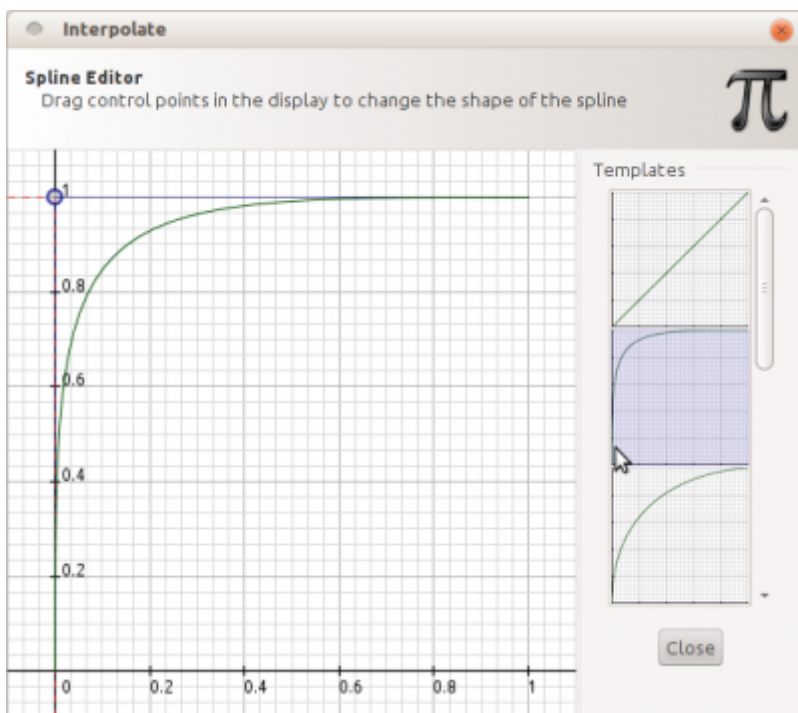
Celem pierwszej części ćwiczenia jest zapoznanie się z narzędziem Gephi i wykonanie prostych analiz oraz wizualizacji sieci. W ramach ćwiczenia studenci zapoznają się z algorytmami rozkładu grafów, poznają ogólny interfejs narzędzia i wykonują samodzielnie ćwiczenie dotyczące wizualizacji grafu powiązań między postaciami z powieści W.Hugo "Nędznicy"

podstawowa obsługa Gephi

1. Załaduj zbiór danych dining.gephi i otwórz go w narzędziu Gephi. Upewnij się, że znajdujesz się na zakładce „Overview”. W oknie *Layout* rozwiń listę i wybierz rozkład o nazwie ForceAtlas 2. W opcjach rozkładu zaznacz opcję *Dissuade Hubs* i ustaw parametr *Gravity* na 5.0. Uruchom rozkład i zaobserwuj wyniki, po krótkiej chwili zatrzymując działanie algorytmu.
2. Włącz wyświetlanie etykiet i dostosuj ich rozmiar do wykresu, możesz także zmienić kolor czcionki do wyświetlania etykiet.
3. W oknie *Partition* kliknij w przycisk *Edges* i odśwież listę atrybutów, które można wykorzystać do partycjonowania zbioru krawędzi. Wybierz z listy wartość *choice* i kliknij przycisk *Apply*. Zauważ zmianę, jaka się dokonała w wizualizacji grafu.
4. Wykorzystaj dodatkowe narzędzia z belki bocznej do zmiany wyglądu grafu. Pokoloruj wierzchołek Jane na czerwono i zwiększ nieco jego rozmiar. Włącz narzędzie *Dragging*, zwiększ promień selekcji i przetestuj jego działanie. Wyznacz najkrótszą ścieżkę między Hildą i Evą.
5. Dodaj nowy wierzchołek do grafu, etykietuj go swoim imieniem i utwórz dwie krawędzie prowadzące z „Twojego” wierzchołka do dwóch dowolnych innych wierzchołków. Następnie, przejdź na zakładkę „Data Laboratory” i sprawdź identyfikator „Twojego” wierzchołka. Kliknij na przycisku *Edges* i zaktualizuj krawędzie prowadzące od „Twojego” wierzchołka, wskazując na jedną z krawędzi jako krawędź pierwszego wyboru, a drugą krawędź jako krawędź drugiego wyboru. Powróć do wizualizacji sieci i sprawdź, czy zmiany zostały uwzględnione. Odśwież wizualizację.
6. Przejdź do okna *Ranking* i zaznacz przycisk *Nodes*. Zauważ, że stopnie wyjściowe i wejściowe wierzchołków zostały już policzone. Zaznacz chęć zróżnicowania wielkości węzłów w zależności od stopnia wejściowego, ustawiając minimalny rozmiar wierzchołka na 2 a maksymalny rozmiar wierzchołka na 10. Kliknij na link [Spline](#) i ustaw funkcję wielkości wierzchołka jak na rysunku poniżej.



7. Kliknij na przycisk **Apply** i zaobserwuj wynik. Ponownie otwórz link [Spline](#) i zmień funkcję na pokazaną poniżej. Jeszcze raz zastosuj różnicowanie rozmiaru wierzchołka względem stopnia wejściowego.

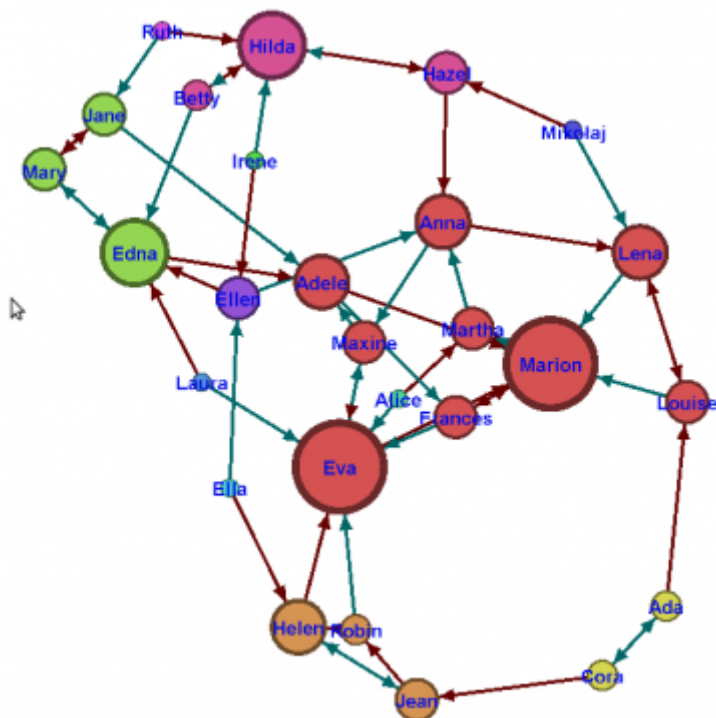


8. Powróć do linowego związku między rozmiarem wierzchołka i jego stopniem wejściowym.

9. Wyświetl rozkład stopni wejściowych i wyjściowych wierzchołków klikając we właściwym miejscu w oknie *Statistics*. Zwróć uwagę, że wyliczone wartości zostały automatycznie dodane do tabeli w zakładce „Data Laboratory”.

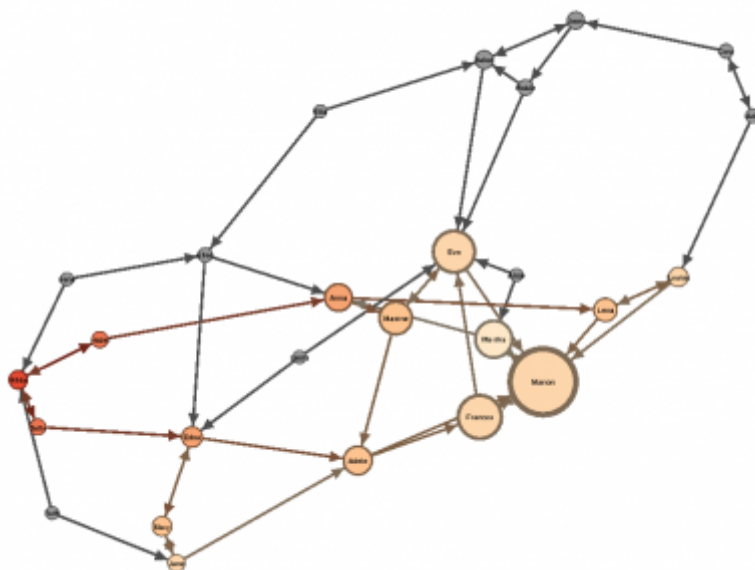
10. Wyznacz silnie spójne komponenty występujące w sieci (posłuż się przyciskiem **Connected Components** w oknie *Statistics*). Wykorzystaj znalezione komponenty do pokolorowania wierzchołków. W tym celu przejdź do okna *Partition*, zaznacz przycisk **Nodes**, odśwież listę

dostępnych atrybutów i wybierz Strongly-Connected ID. Jeśli zaproponowane kolory są zbyt do siebie podobne, kliknij prawym klawiszem myszy i z menu kontekstowego wybierz opcję Randomize. Kliknij przycisk Apply i obejrzyj uzyskaną wizualizację. W chwili obecnej Twoja sieć powinna wyglądać mniej więcej w ten sposób:



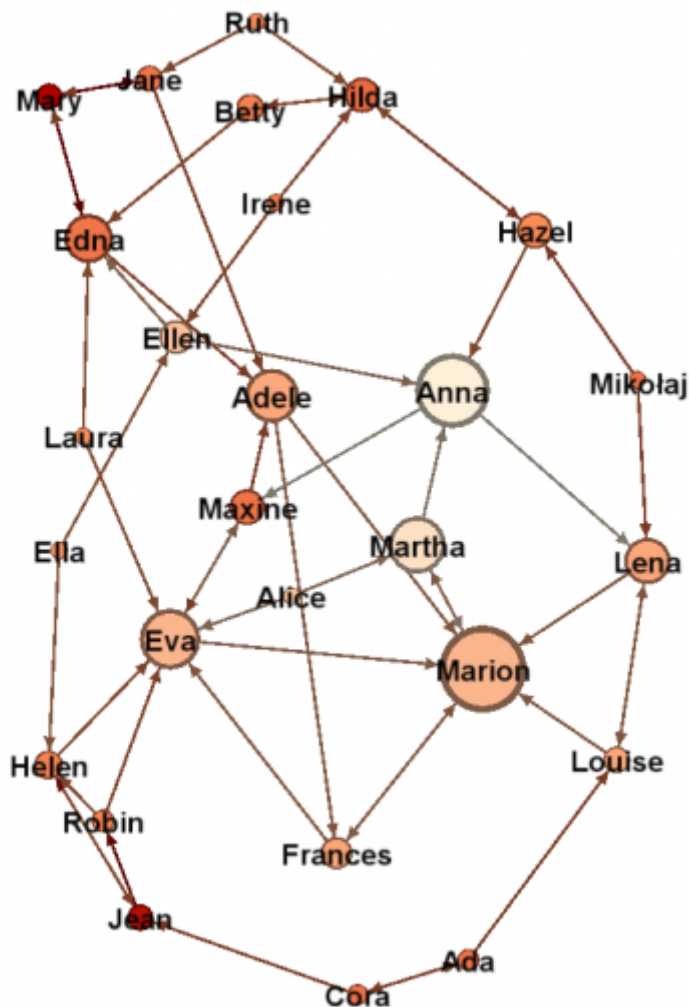
11. Zresetuj wielkości i kolory węzłów i etykiet. Wyznacz odległości między wszystkimi parami wierzchołków klikając na przycisk Avg. Path Length w oknie Statistics. Upewnij się, że traktujesz krawędzie jako łuki posiadające kierunek. Nie włączaj normalizacji wyznaczanych miar centralności. Obejrzyj uzyskane rozkłady.

12. Wykorzystaj narzędzie do rysowania odległości od zadanego wierzchołka. Wyświetl mapę odległości wierzchołków od wierzchołka Hilda.



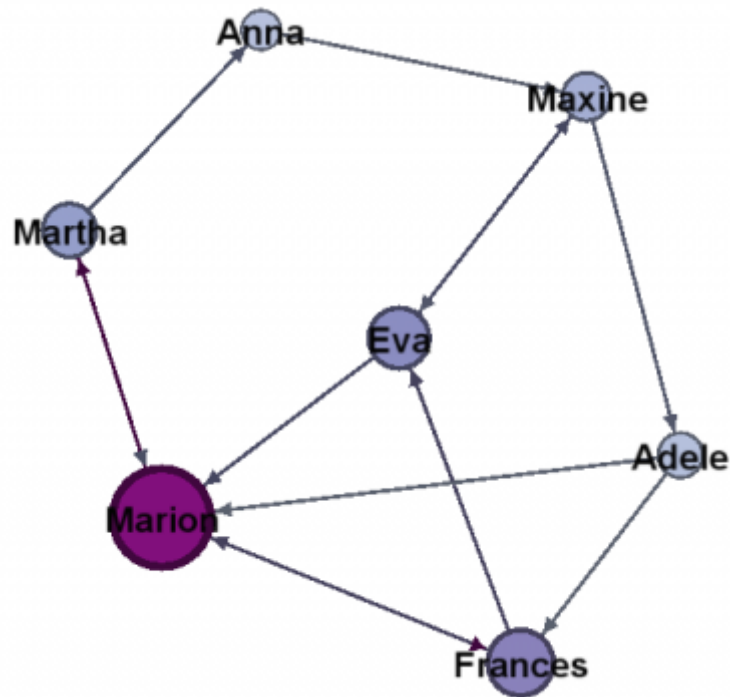
13. Przejdź do okna *Ranking* i zaznacz przycisk *Nodes*. Zaznacz chęć zmiany koloru wierzchołka w zależności od wskazanego kryterium, a z listy jako kryterium wybierz centralność wg. miary bliskości. Zmień używany schemat kolorów i zastosuj kryterium. Obejrzyj wynik w wizualizacji, wyświetl także ranking wierzchołków w postaci tabelarycznej.

14. Jako kolejne kryterium wybierz centralność wg. pośrednictwa i wskaż rozmiar wierzchołka jako modyfikowaną cechę. Zastosuj kryterium do sieci i zaobserwuj wynik. Twoja sieć powinna wyglądać mniej więcej tak:



15. Krawędzi w analizowanej sieci można interpretować jako jednoznaczne głosowanie wierzchołka A na wierzchołek B. W takiej sytuacji dobre przybliżenie globalnego rankingu ważności wierzchołków produkuje algorytm PageRank. Uruchom ten algorytm z prawdopodobieństwem losowego przeskoczenia do innego węzła równym 15% i progiem zbieżności w wysokości 0.001. Użyj wyznaczonej miary PageRank aby narysować sieć w taki sposób, aby wielkość i kolor wierzchołka odpowiadały wyznaczonej centralności wg. PageRank. Zidentyfikuj 5 najważniejszych dziewczyn w sieci.

16. Wybierz tylko te wierzchołki w grafie, które posiadają miarę centralności wg. PageRank powyżej 0.05 i utwórz z nich nową sieć. W tym celu przejdź na zakładkę „Filters” i rozwiń gałąź *Attributes-Range*, wybierz węzeł PageRank i przenieś metodą *Drag & Drop* do okna *Queries*. Wykorzystaj suwak aby ograniczyć zbiór wierzchołków do pożądanego progu miary PageRank. Twoja sieć powinna wyglądać tak:



17. Otwórz ponownie pełną sieć i zresetuj kolory i wielkości wierzchołków. Następnie wybierz dowolne kryterium centralności i użyj go do zróżnicowania wyglądu wierzchołków (przykładowo: rozmiar wierzchołka reprezentuje pośrednictwo, kolor wierzchołka odpowiada komponentowi). Przejdź na zakładkę „Preview”. Z listy rozwijanej wybierz szablon „Default Curved”, a następnie dokonaj niewielkich zmian. Zmień czcionkę na Arial Bold 8, wyłącz etykiety krawędzi, zmień promień na 1.0, wyłącz rysowanie granicy węzłów. Wygeneruj i obejrzyj plik *.png. Następnie wyłącz krzywiznę krawędzi, ustaw rozmiar strzałek na 3.0 i ponownie wygeneruj plik.

zadanie samodzielne

Pobierz plik [les.miserables.gephi](#) potrzebny do wykonania ćwiczenia. Plik zawiera informacje o współwystępowaniu poszczególnych postaci w kolejnych scenach w powieści Wiktora Hugo pt. Nędznicy. Zapoznaj się z fabułą powieści.

- <http://pl.wikipedia.org/wiki/N%C4%99dznicy>
- D. E. Knuth, The Stanford GraphBase: A Platform for Combinatorial Computing, Addison-Wesley, Reading, MA (1993).

1. Wczytaj sieć do narzędzia Gephi. Eksperymentalnie dobierz taki rozkład, który zwiększa czytelność. Zauważ, że wierzchołki są już przydzielone do klas modularności.
2. Wyświetl sieć w taki sposób, aby uwypuklić centralność wierzchołków według ich pośrednictwa.
3. Wyświetl sieć z wyznaczonymi odległościami od wierzchołka Gavroche.
4. Zbuduj podsieć zawierającą 10% najważniejszych wierzchołków, gdzie miarą ważności wierzchołka jest jego stopień.
5. Przygotuj najbardziej atrakcyjną wizualizację sieci. Jeśli chcesz, możesz ograniczyć sieć do pewnego podzbioru wierzchołków. Eksperymentuj z doбором kolorów, kształtów, czcionek, tła.

From:

<https://semantic.cs.put.poznan.pl/wiki/TSiSS/> - **Technologie semantyczne i sieci społecznościowe**

Permanent link:

<https://semantic.cs.put.poznan.pl/wiki/TSiSS/doku.php?id=laboratorium-gephi-1>

Last update: **2017/03/29 18:15**

