NOTE: Obliczenia dla „Lizaka" były dość kosztowne (dla k=5, program działał 20godzin), można to argumentować bardzo długim czasem naszego random walk'a (i ew. moimi niskimi umiejętnościami optymalizacji kodu w Matlabie).

Zacznę od tego, że najszybsze przejście grafu następowało kolejno dla:

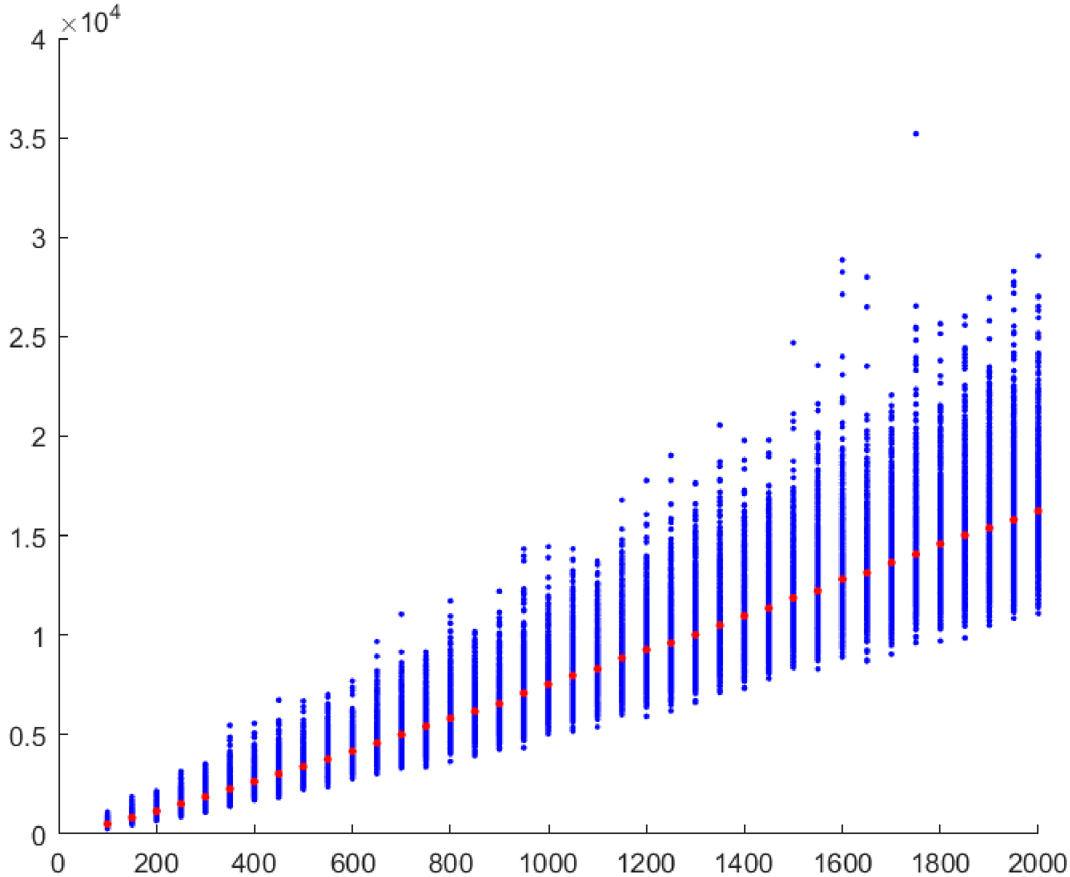
1. Kliki +/-10^4 kroków
2. Drzewa +/-10^5 kroków
3. Ścieżki (z dowolnego punktu) +/-10^6 kroków (pojedyncze wartości w 10^7)
4. Lizaka +/-10^9 kroków

Koncentracja wyników wokół wartości średniej dla wszystkich grafów malała wraz ze wzrostem n. Jednak zdecydowanie najszybciej następowało to dla ścieżek (i lizaka, aczkolwiek niskie k sprawia, że nie mogę tego powiedzieć ze 100% przekonaniem).

CLIQUE

Patrząc na wykres na pierwszy rzut oka, ciężko ocenić asymptotykę wykresu (wątpliwości między O(n) a O(n ln n)). Po dokładniejszych oględzinach (i przyłożeniu linijki do ekranu), byłbym skłonny uznać, że czas pokrycia kliki jest O(n ln n).

W literaturze znalazłem czas pokrycia dla kliki:
n (ln n). Co potwierdzałoby zaobserwowaną (i zmierzoną linijką) aproksymacje.



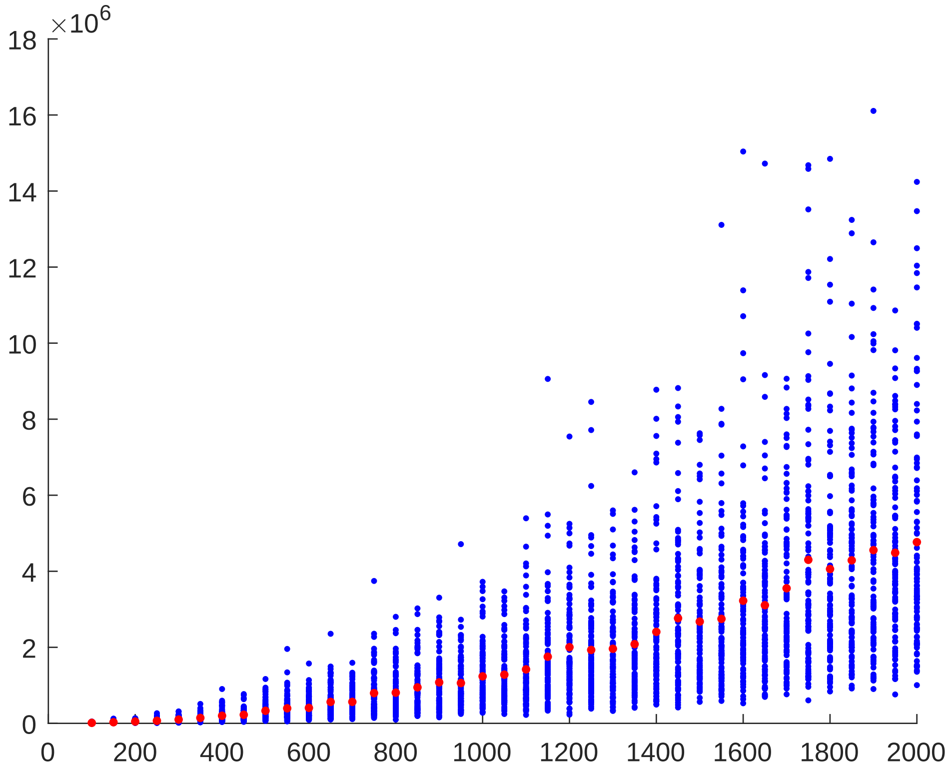
×10^4
4
3.5
3
2.5
2
1.5
1
0.5
0
0
200
400
600
800
1000
1200
1400
1600
1800
2000

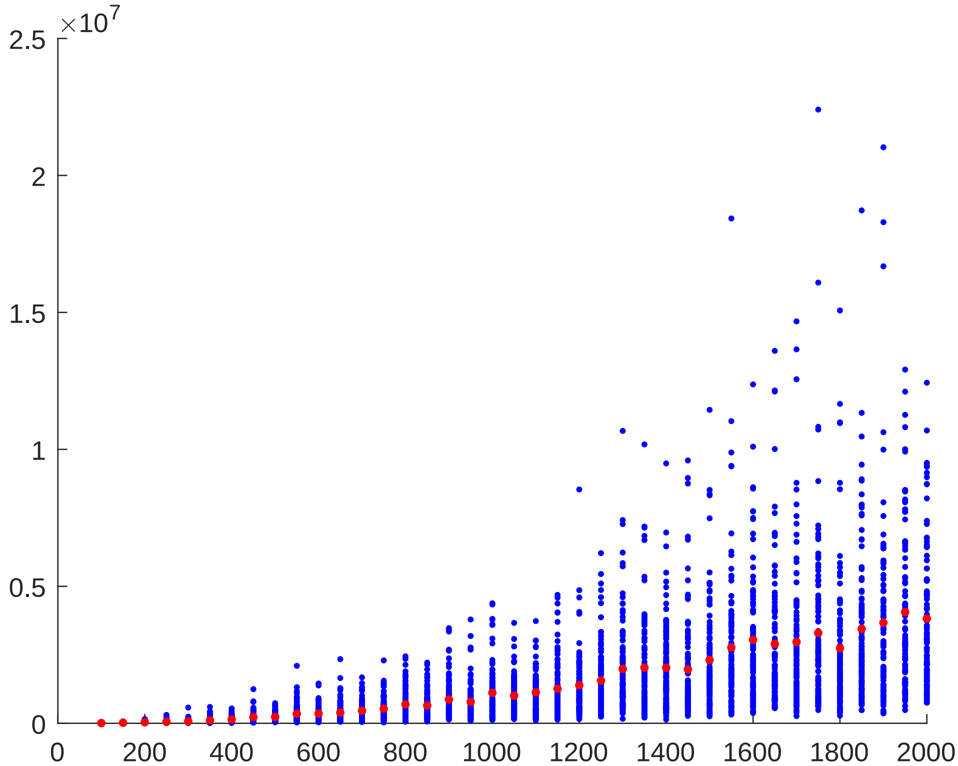
PATH

Dla ścieżek możemy zauważyć, że końcowe wykresy są takie same dla obu wybranych punktów. Co więcej, będą takie same dla dowolnego punktu ze ścieżki, ma to związek z faktem, że nasz rozkład prawdopodobieństwa będzie dążył do pewnego rozkładu stacjonarnego <=> niezależnie gdzie zaczynamy, koniec końców szanse na wylądowanie w punkcie X są takie same (dla odpowiednio dużej liczby kroków) dla każdego punktu ze ścieżki. Jednakowy wygląd naszych wykresów możemy traktować jako swego rodzaju potwierdzenie powyższego faktu.

Patrząc na wykresy widzimy że sama wartość średnia nie rośnie nad wyraz szybko. Jednak pojedyncze wartości wydają się rosnąć w tempie kwadratowym. O(n^2) wydaje się być dobrą aproksymacją (myślę, że można zauważyć kształt paraboli idąc po najwyższych wartościach dla każdego n, dla obu ścieżek).

W literaturze znalazłem czas pokrycia dla ścieżki:
5/4*(n-1)^2 - ¼. Co potwierdzałoby zaobserwowaną aproksymacje.


×10^6
18
16
14
12
10
8
6
4
2
0
0
200
400
600
800
1000
1200
1400
1600
1800
2000



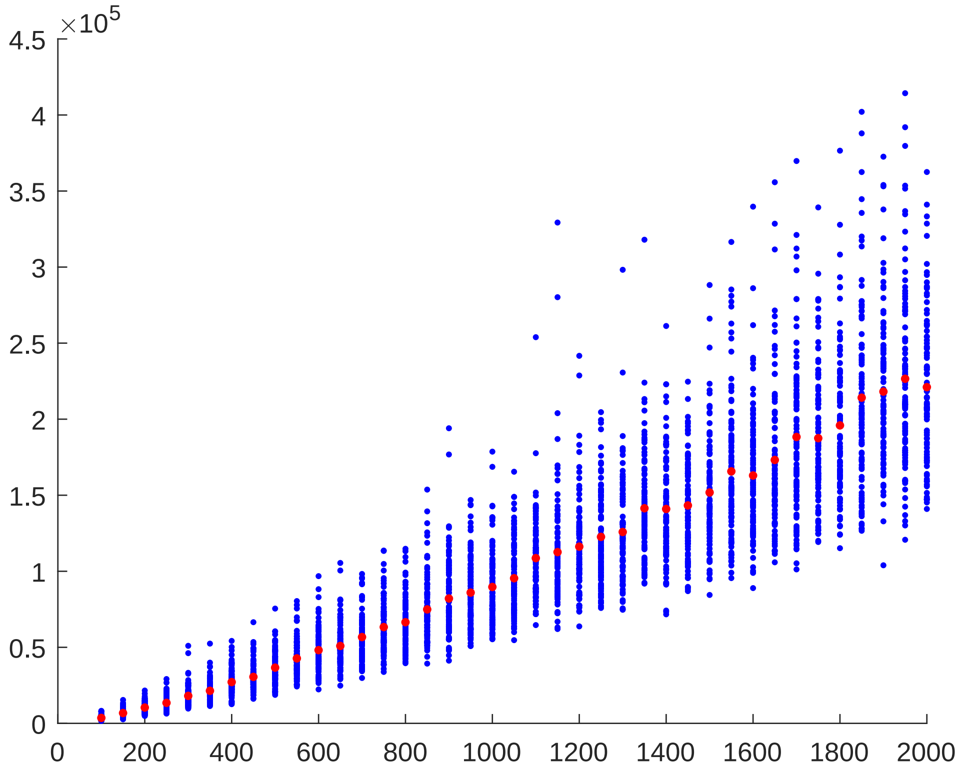
×10^7
2.5
2
1.5
1
0.5
0
0
200
400
600
800
1000
1200
1400
1600
1800
2000

TREE

Wykres wygląda podobnie do kliki, główna różnica to wartości o rząd wielkości większe i nieco mniejsze zagęszczenie wokół średniej (to ostatnie może być spowodowane niższym k).

Tu nieco wyraźniej widać, że wartości nie są liniowe, jednak nie są również n^2. Biorąc to pod uwagę jak i podobieństwo wykresu do wykresu kliki sugerowałbym aproksymację O(n ln n).

W literaturze znalazłem czas pokrycia dla drzewa binarnego:
n (ln n). Co potwierdzałoby zaobserwowaną aproksymacje.



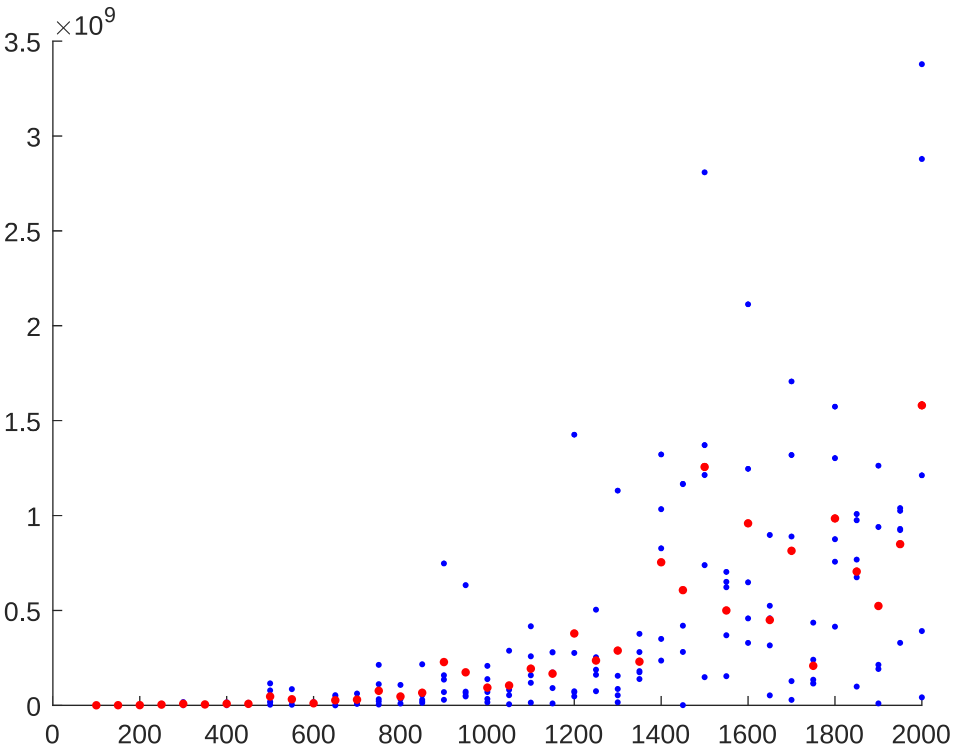
×10^5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
0
200
400
600
800
1000
1200
1400
1600
1800
2000

LOLLIPOP

Na wykresie widać bardzo szybki wzrost wartości (szybszy niż n^2, lecz wolniejszy niż n! czy 2^n) oraz szybkie pojawienie się bardzo dużych (w porównaniu do innych grafów) wartości. Oba te fakty powodują, że skłaniałbym się do aproksymacji O(n^3).

W literaturze znalazłem czas pokrycia dla lizaka:
n^3. Co potwierdzałoby zaobserwowaną aproksymacje.

.


×10^9
3.5
3
2.5
2
1.5
1
0.5
0
0
200
400
600
800
1000
1200
1400
1600
1800
2000

Skąd takie a nie inne wyniki?

Szybkość Kliki można argumentować faktem, że każdy node jest jednakowo prawdopodobny do bycia następnym nodem gdzie „zawitamy" w naszym random walku.

Ścieżki były stosunkowo szybkie jednak dla zdecydowanie większych n byłby drugim najwolniejszym z analizowanych wykresów. Jest tak gdyż średnio szansa na pójście w kierunku nie odwiedzonego node'a jest taka sama jak powrót do miejsca gdzie byliśmy ruch wcześniej, także możemy spodziewać się dużo ruchów które nas nie zbliżają i/lub oddalają od pełnego pokrycia.

W przypadku drzewa, sytuacja była dość ciekawa. (QUICK NOTE: dzieci == wierzchołki z rzędu niżej, rodzic == wierzchołek rząd niżej, rodzic i dzieci danego wierzchołka są z nim bezpośrednio połączone tj. nie ma innych node'ów po drodze). Dla wierzchołków które mają 2 dzieci i jednego rodzica szansa na zejście w dół jest 2x większa niż wrócenie na górę. Początkowo ma to niewątpliwą zaletę tj. szansa na powrót do pkt. poprzednie/startowego jest mniejsza niż do nowych niezbadanych node'ów. Jednak później gra to na naszą niekorzyść tj. „ciężko się przepiąć" do nowej gałęzi. Jest jednak w pewien sposób „niwelowane" przez fakt, że odległość z pkt. początkowego do dowolnego innego jest zdecydowanie mniejsza niż w przypadku np. ścieżki. Oznacza to, że jeśli ugrzęźniemy w jakiejś gałęzi to nie jesteśmy aż tak daleko od pkt. początkowego a tym samym innych nowych gałęzi.

Dla lizaka mamy dramat jeśli chodzi o czas pokrycia. O ile pokrycie samej kliki jest łatwe i szybkie to tu fakt, że każdy punkt (z kliki) ma jednakową szansę na bycie wybranym sprawia, że bardzo ciężko trafić do punktu połączonego ze ścieżką i równie ciężko wejść na samą ścieżkę. Dodajmy do tego fakt, że ze ścieżki bardzo łatwo wypaść i wrócić do kliki i widzimy, że mamy bardzo dobrą recepturę na długi czas pokrycia.