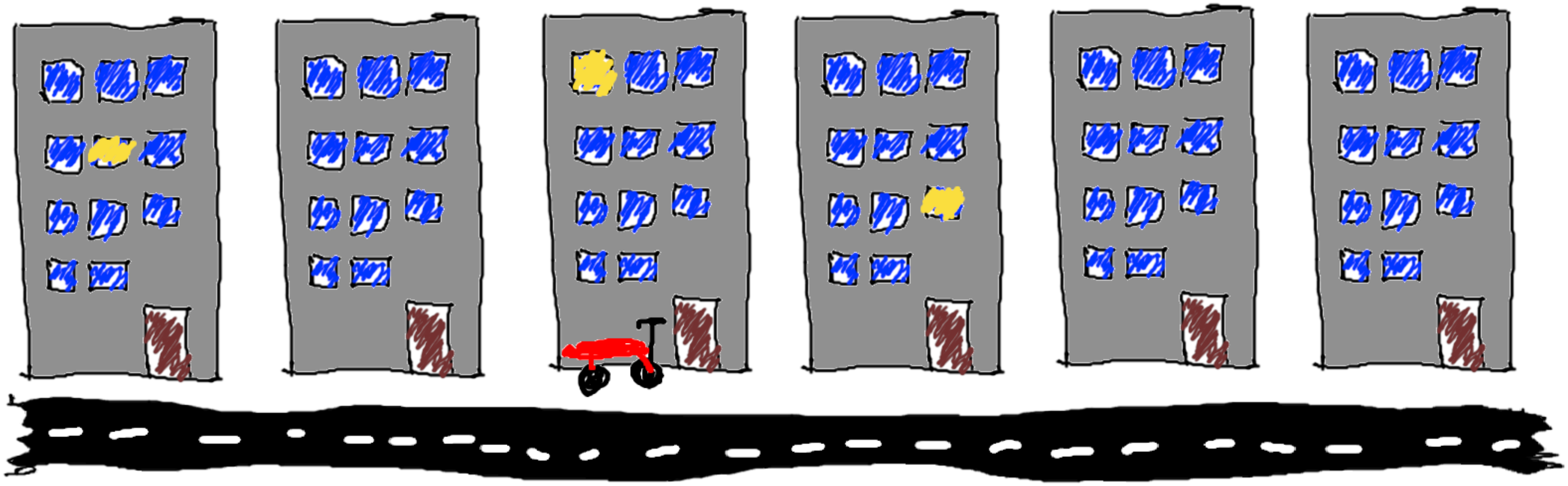
Ograniczenie górne współczynnika aproksymacji odpornego na manipulację mechanizmu lokalizacyjnego na grafach z cyklem

Problem lokalizacyjny na przykładzie

Grupa znajomych (A)dam, (B)artek, (C)zesiek, ... **(agenci)** mieszkający przy jednej ulicy **(graf)** kupili razem motor. Zastanawiają się gdzie powinni go parkować tak, aby sumarycznie razem mieli do niego jak najbliżej **(problem lokalizacyjny)**.

Najprostszy przypadek

Chłopcy mieszkają przy jednej prostej drodze.



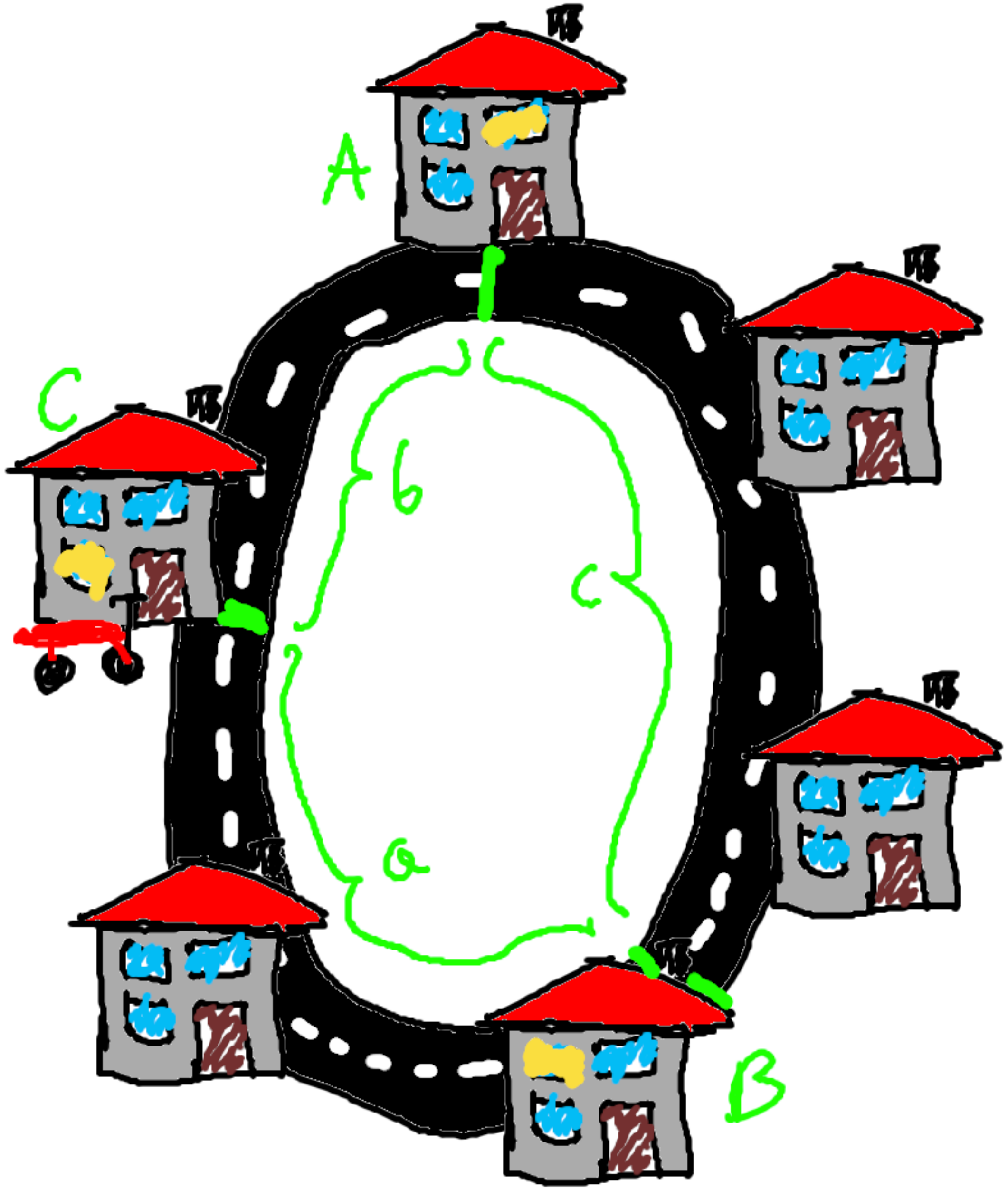
Rozwiązanie: Parkowanie motoru przy mieszkaniu którego lokalizacja jest medianą pozostałych (proste zadanie, że to minimalizuje sumę odległości).

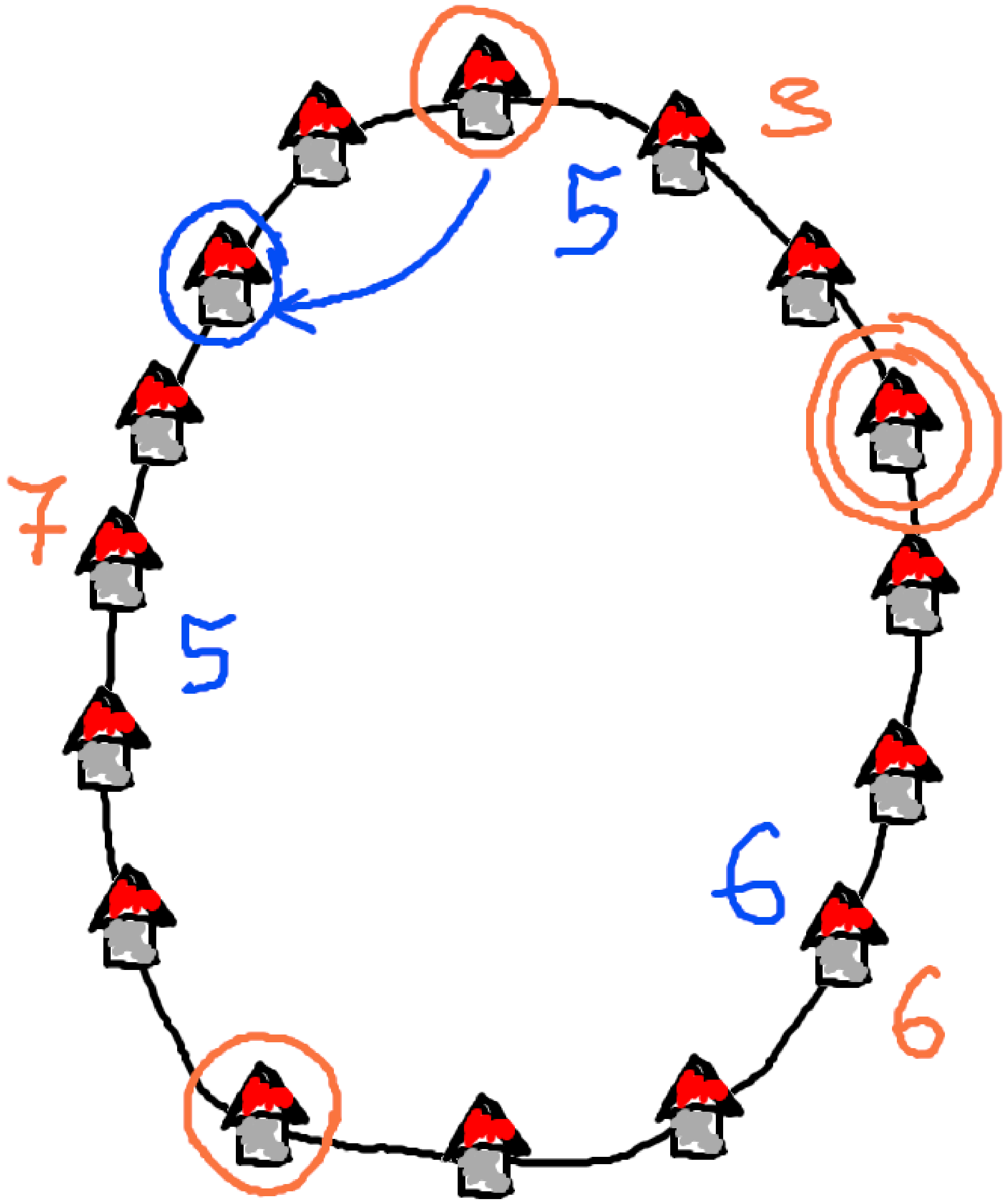
Chytrość nie popłaca (odporność na manipulację)

- założenie niewiedzy o wzajemnych lokalizacjach jest trochę naciągane
- agent, którego miejsce jest medianą, ma już najlepszy możliwy dla siebie rezultat
- pozostali agenci mogą co najwyżej pogorszyć swój wynik, przesuwając medianę dalej od swojej prawdziwej lokalizacji

Osiedle rondowe

- brak mediany
- koszt powiązany z każdym miejscem
- punkt na przeciwko najdłuższego łuku generuje najmniejszy koszt





Chytrość popłaca :(

Przykład na obrazku

- tradeoff między optymalnością a odpornością na manipulację
- definicja współczynnika aproksymacji - czyli jak dobrze sobie radzimy

Dyktatura na ratunek

- ustalamy jednego agenta i zawsze wybieramy jego propozycję
- trywialnie odporny na manipulację
- może skutkować wyborem skrajnie nieefektywnego miejsca
- *nieuczciwy*?
- w aktualnym modelu to najlepsze co możemy mieć (literatura)

Losowa dyktatura (RD) na ratunek

- losowo ustalamy agenta po głosowaniu i wybieramy zaproponowane przez niego miejsce
- trywialnie odporny na manipulację (loteria)
- działa lepiej (skupiska mają większą szansę wyboru)

Garść oznaczeń

- G zbiór wierzchołków grafu
- N zbiór agentów o mocy n
- G^n zbiór możliwych stanów świata (preferencji agentów odnośnie lokalizacji)
- $x \in G^n$ dowolny stan świata, $y \in G$ wierzchołek generujący najmniejszy koszt
- $sc_x(l)$ funkcja obliczająca koszt społeczny (sumę odległości powiązana z loterią l)
- $OPT = sc_x(y)$ - najmniejszy koszt
- rd funkcja obliczająca loterię wierzchołków zgodnie z mechanizmem RD

Zły przypadek dla RD

- współczynnik aproksymacji oszacowany od dołu przez $2 - 2/n$
- dla dowolnego grafu z dwoma różnymi lokacjami a, b odległymi o d ustalmy stan świata x, w którym 1 agent preferuje a, a pozostałych $n - 1$ preferuje b
- $sc_x(a) = (n-1)d$, $sc_x(b) = d$
- minimalny koszt generuje wybranie b ($OPT = d$)
- koszt rd:

$$sc_x\left(rd(x)\right) = sc_x\left(\frac{1}{n}a + \frac{n-1}{n}b\right) = 2\frac{n-1}{n}d = \left(2 - \frac{2}{n}\right)OPT$$

Ograniczenie górne dla RD

Na podstawie artykułu:

$$sc_x(rd(x)) = \sum_{i\in N} \frac{1}{n} \sum_{j\in N} d(x_i, x_j) \overset{(1)}{\leq} \frac{1}{n} \sum_{i\in N, j\in N\setminus\{i\}} (d(x_i, y) + d(y, x_j))$$

$$= \frac{1}{n} \sum_{i\in N} ((n-1)d(x_i, y) + (OPT - d(y, x_i))) = OPT + \frac{n-2}{n} \sum d(x_i, y)$$

$$= \left(2 - \frac{2}{n}\right) OPT$$

(1) - z nierówności trójkąta

Czy można lepiej?

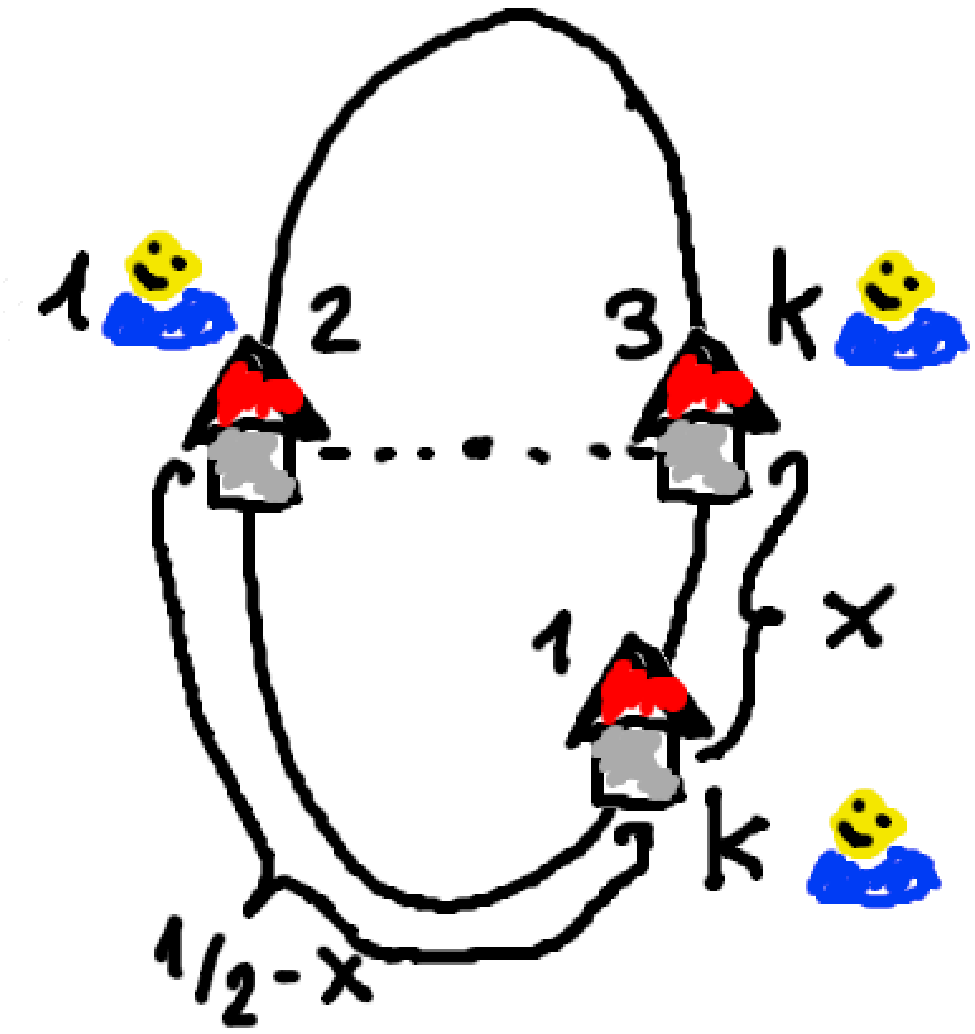
- znane ograniczenia dolne na współczynnik aproksymacji w klasie mechanizmów odpornych na manipulację: 1 (w niektórych przypadkach wiemy więcej - artykuł)
- ograniczenie górne wynikające z pokazanego wcześniej mechanizmu: $2 - \frac{n}{2}$ (dla $n \to \infty$ to 2)
- co pomiędzy? (moja praca magisterska)

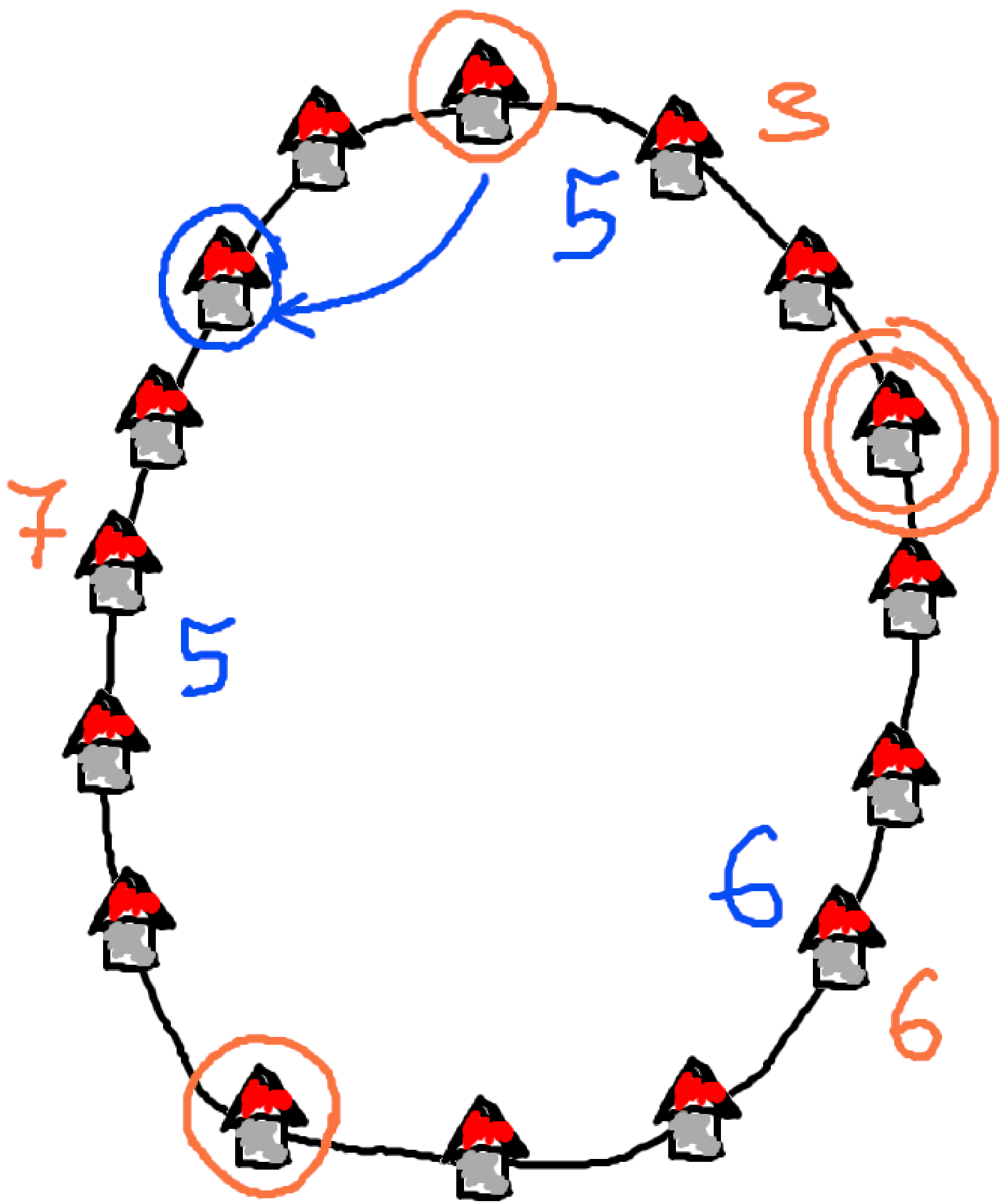
Mechanizm PCD

- definicja dla nieparzystej liczby agentów
- rozcinamy cykl po środku losowej krawędzi (z równym prawdopodobieństwem) i wybieramy medianę zgłoszeń (wedle naturalnego porządku na powstałym odcinku)
- "trywialnie" odporny na manipulację

Zły przypadek PCD

- oznaczenia jak na obrazku
- $x = \frac{1}{4\sqrt{k}}$
- $c_1 = kx + \frac{1}{2} - x \leq \frac{\sqrt{k}}{4} + \frac{1}{2}$
- $c_2 = k(\frac{1}{2} - x) + \frac{k}{2}$
- $c_3 = kx + \frac{1}{2}$
- koszt: $\frac{1}{2}c_1 + xc_2 + (\frac{1}{2} - x)c_3$
  $= -2kx^2 + (2k-1)x + \frac{1}{2}$
  $= \frac{\sqrt{k}}{2} - \frac{1}{4\sqrt{k}} + \frac{3}{8} \geq \frac{\sqrt{k}}{2}$





Porównanie wyników

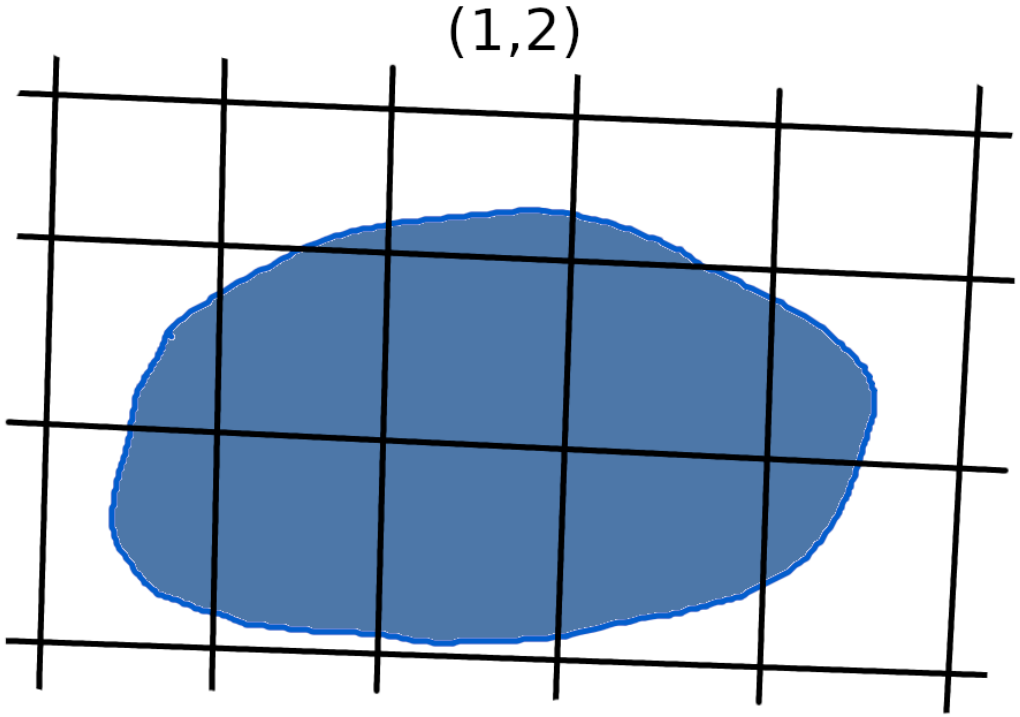
Intuicja lepszości PCD od RD przez obrazek.

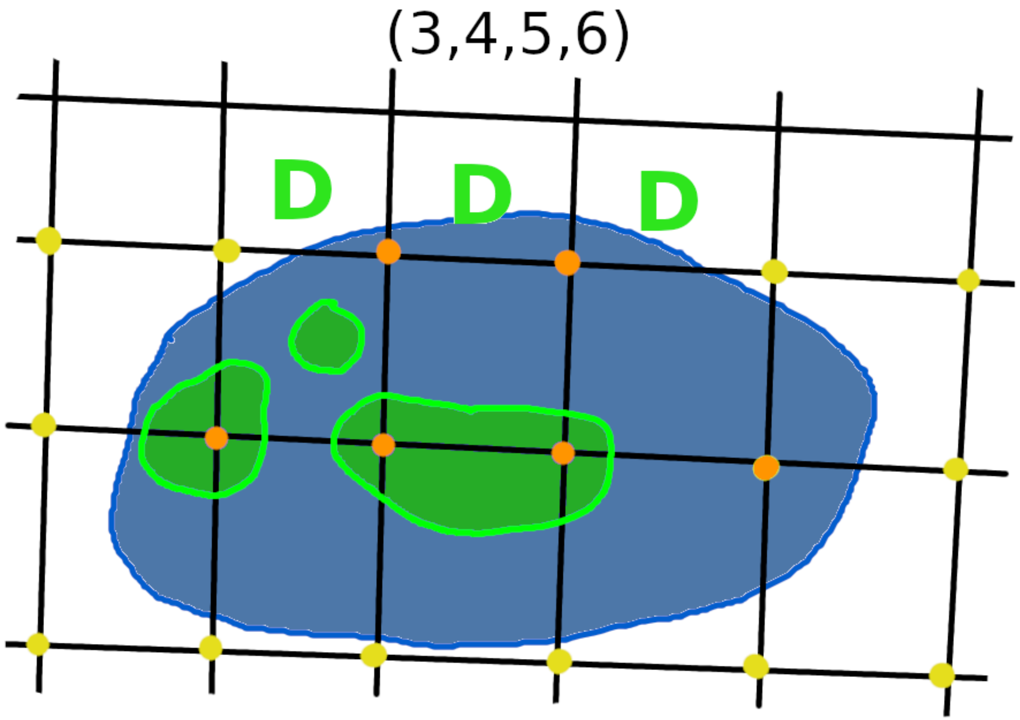
Metoda	Współczynnik aproksymacji	SP
OPT	1	Nie
RD	≈ 1.19	Tak
PCD	1.125	Tak
BEST	< 1.04	Tak

Ograniczenie górne PCD (plan dowodu)

1. podział stanów świata na przypadki rozróżniające PCD
2. klasyfikacja przypadków realizowalnych
3. rozpatrzenie dominujących PCD przypadków
4. pozbycie się wierzchołków nieistotnych (lemat o nieoptymalności i niepopularności wierzchołków nieistotnych)
5. oszacowanie górne współczynnika aproksymacji w PCD przypadku z użyciem programowania liniowego
6. globalne szacowanie

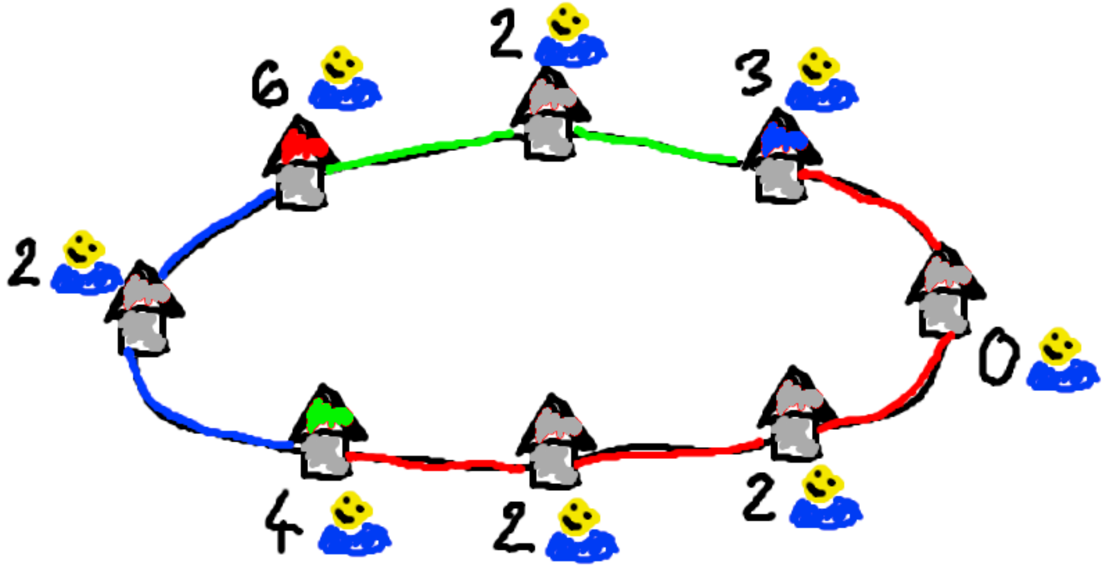
Ograniczenie górne PCD (przez obrazki)





Anatomia PCD przypadku

- grupowanie agentów ze względu na wybrany wierzchołek (ze względu na anonimowość stan świata można reprezentować funkcją popularności)
- przypadek PCD - przyporządkowanie krawędzi grafu jego wierzchołkom



Anatomia PCD przypadku kontynuacja

- wierzchołki przypisane krawędziom nie mogą się przeplatać
- grupując po przypisanych wierzchołkach PCD przypadek jest funkcją różnowartościową z parami rozłącznych spójnych podzbiorów krawędzi (łuków) w wierzchołki
- przeciwdziedzinę PCD przypadku nazwiemy wierzchołkami istotnymi

Charakteryzacja realizowalnych PCD przypadków

- przypisywane łuki (dziedzina PCD przypadku) kończą się w wierzchołkach istotnych
- wierzchołki istotne mają przypisane łuki na przeciwko (ze względu na liczbę wierzchołków istotnych)
- inne spojrzenie - ustalamy nieparzyście wiele wierzchołków istotnych i każdemu odpowiada łuk między przeciwnymi wierzchołkami istotnymi

PCD przypadek dominujący

- wszystkie krawędzie są przypisane jednemu wierzchołkowi v
- v ma przypisaną sąsiadującą krawędź $\Rightarrow$ v ma popularność > 0.5
- v jest wierzchołkiem generującym najmniejszy koszt
- PCD ma współczynnik aproksymacji 1 (zawsze wybiera optymalny wierzchołek)

Lemat o nieoptymalności (punktów nieistotnych)

- dla każdego nieistotnego wierzchołka v, któryś z jego istotnych sąsiadów (v_b, v_e) generuje nie większy koszt
- oznaczmy przez v_o wierzchołek, któremu w danym PCD przypadku przypisany jest łuk (v_b, v_e)
- "przesuwanie się" na kole od v w stronę v_o (po krótszym łuku) zbliża nas do ściśle więcej niż połowy agentów, więc zmniejsza koszt

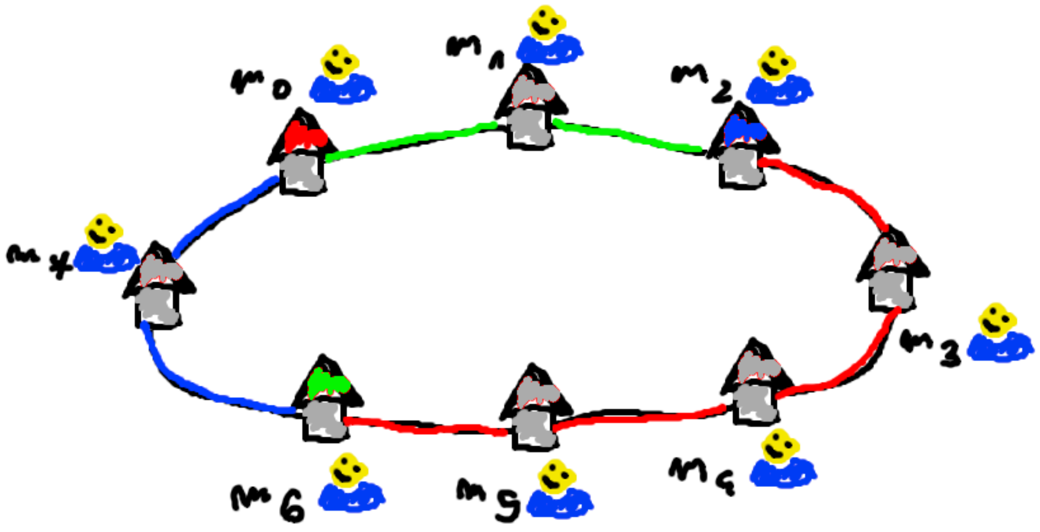
Lemat o niepopularności (punktów nieistotnych)

- istnieją "złe" stany świata, w których punkty nieistotne mają zerową popularność
- popularność punktów nieistotnych możemy przenieść na istotnych sąsiadów nie zmniejszając współczynnika aproksymacji

Szacowanie współczynnika aproksymacji realizowalnego PCD przypadku z użyciem LFP

- dla dowolnego wyboru wierzchołka generującego optymalny koszt
- funkcja celu jest postaci $\frac{\text{funkcja liniowa popularności}}{\text{funkcja liniowa popularności}}$
- ograniczenia są nierównościami liniowymi:
  - koszt powiązany z wierzchołkiem jest minimalny
  - sumaryczne popularności łuków są ograniczone (tak by realizować zadany przypadek)
  - popularności sumą się do liczby agentów
- niektóre nierówności trzeba "stępić"

Przykładowy model LFP



- zmienne: $m_0, m_1, \ldots, m_7$
- cel: minimalizacja $\frac{sc_x(pcd(x))}{sc_x(opt(x))}$
- normalizacja popularności:
  $m_0 + m_1 + \ldots + m_7 = 1$
- minimalność kosztu powiązanego z wierzchołkiem 0: $\forall_i c_0 \leq c_i$
  (każdy koszt c_i jest wielkością liniową zmiennych $m_0, \ldots, m_7$)
- ograniczenia dolne popularności łuków np.: $m_0 + m_1 + m_2 \geq 0.5$
- ograniczenia górne popularności łuków np.: $m_1 + m_2 + m_3 + m_4 + m_5 \leq 0.5$

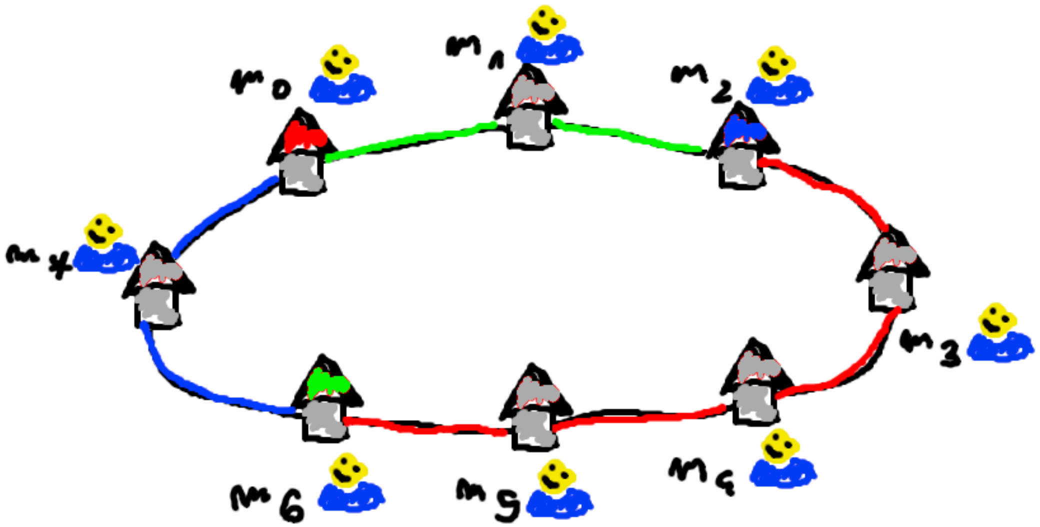
Redukcja do LP

- transformacja Charnes'a Cooper'a
- nowa funkcja celu
- nowa zmienna
- nowe ograniczenia
- dodatkowo rezygnujemy z części ograniczeń

Przykładowy model po uproszczeniach

- zmienne: $m_0, m_1, \ldots, m_7, t$
- cel: minimalizacja $sc_x(pcd(x)) = \frac{2m_0+3m_1+4m_2+5m_3+6m_4+5m_5+4m_6+3m_7}{16}$
- normalizacja popularności:

  $m_0 + m_1 + \ldots + m_7 = t$
- ograniczenia dolne popularności łuków:
  - $m_0 + m_1 + m_2 \geq \frac{t}{2}$
  - $m_2 + m_3 + m_4 + m_5 + m_6 \geq \frac{t}{2}$
  - $m_0 + m_6 + m_7 \geq \frac{t}{2}$
- ograniczenie mianownika: $sc_x(opt(x)) = \frac{m_1+2m_2+3m_3+4m_4+3m_5+2m_6+m_7}{8} = 1$



Maksimum modelu

- maksimum jest realizowane w konfiguracjach spełniających możliwie dużo ograniczeń w sposób równościowy
- nie istnieje konfiguracja spełniająca równościowo wszystkie ograniczenia
- istnieją konfiguracje spełniające równościowo wszystkie ograniczenia bez dowolnego jednego
- maksimum funkcji celu jest realizowane w konfiguracjach z popularnością rozłożoną po równo między dwa wierzchołki istotne

Globalne szacowanie

- "złe" stany świata mają dwa wierzchołki istotne z popularnością 0.5
- niech l oznacza krótszy łuk między owymi wierzchołkami, a d jego długość
- wierzchołki należące do l generują koszt $\frac{d}{2}$ (bo $d \leq 0.5$)
- wierzchołki należące do l mają przypisane prawdopodobieństwo ze wszystkich krawędzi nie należących do l (przynajmniej $1-d$)
- wierzchołki spoza l mają koszt nie większy niż $\frac{1-d}{2} \geq \frac{d}{2}$
- koszt powiązany z PCD jest nie większy niż:

$$\frac{(1-d)\frac{d}{2} + \frac{1-d}{2}d}{\frac{d}{2}} = 2 - 2d$$

Wnioski

- współczynnik aproksymacji PCD na kole jest ograniczony z góry przez $2 - \frac{2}{k}$, gdzie k to liczba wierzchołków, dla dowolnej nieparzystej liczby agentów (najmniejsza możliwa długość łuku to $\frac{1}{k}$)
- przeglądając uważnie powyższy dowód możemy wywnioskować jakie stany świata (PCD przypadki) są złe dla PCD. Odpowiadają one mniej więcej temu przedstawionemu wcześniej jako zły przypadek

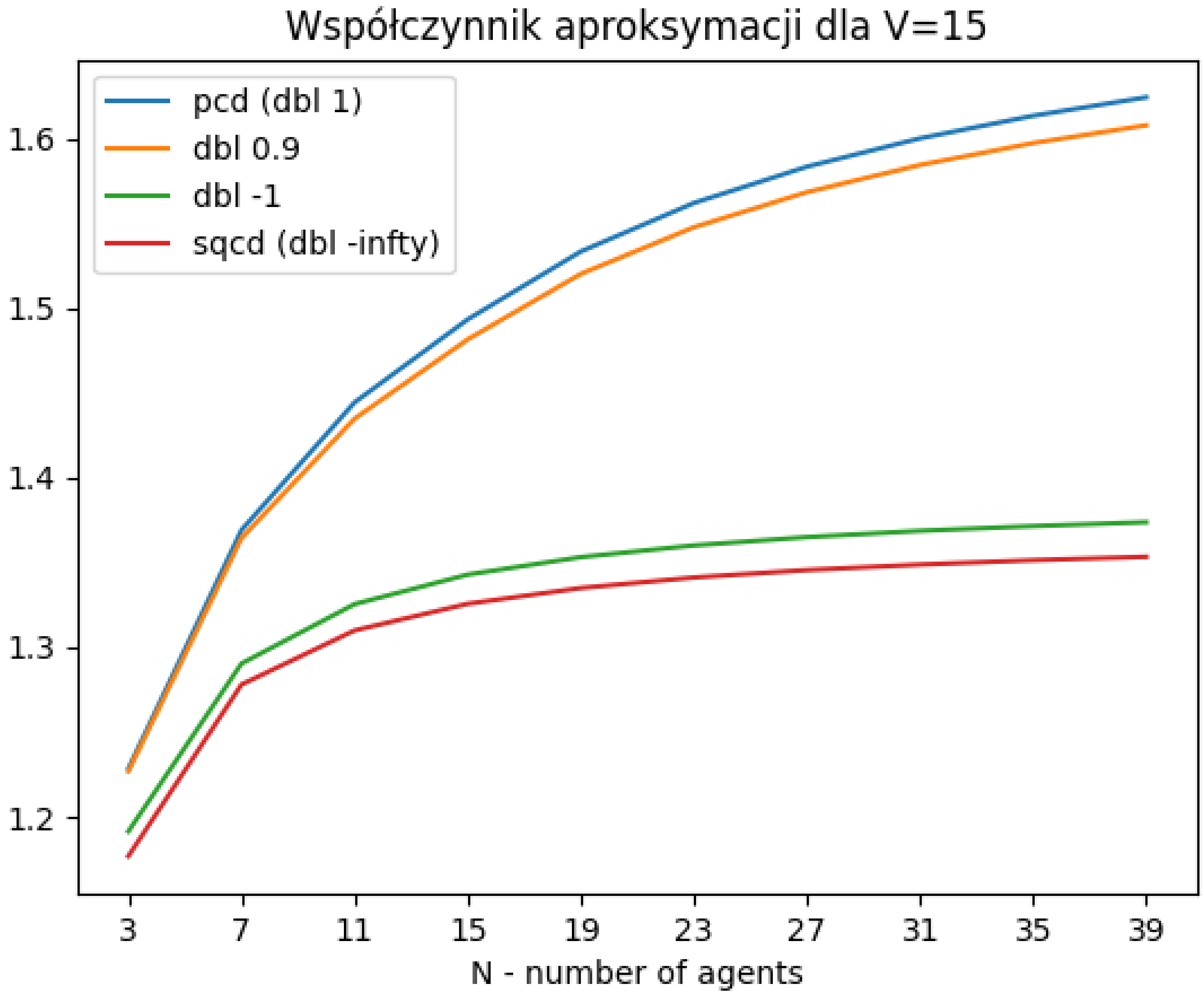
Co dalej?

- PCD dla parzystej liczby agentów
- twierdzenie o mieszaniu PCD z RD
- jeszcze bardziej obiecujące wyniki eksperymentalne

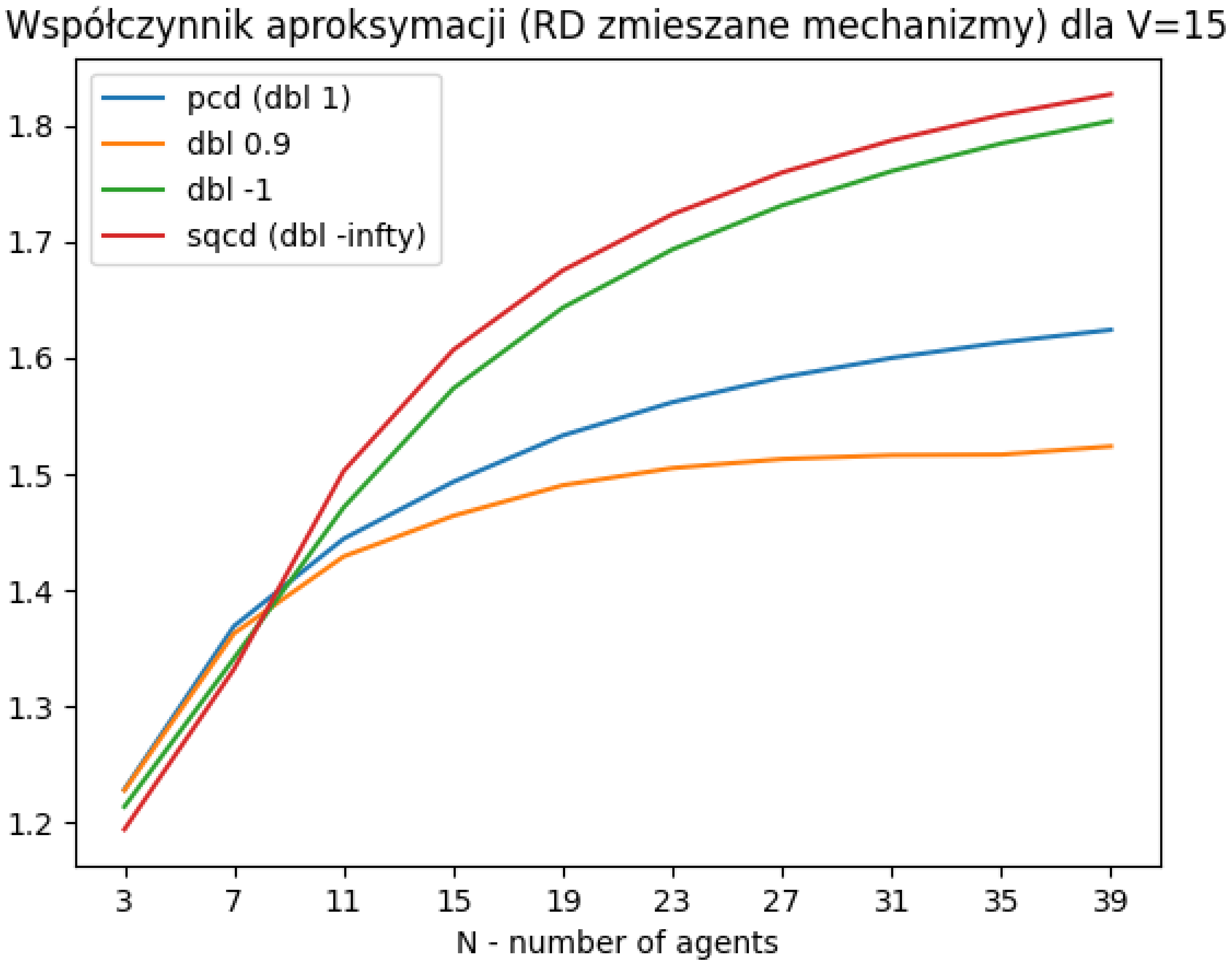
Koniec

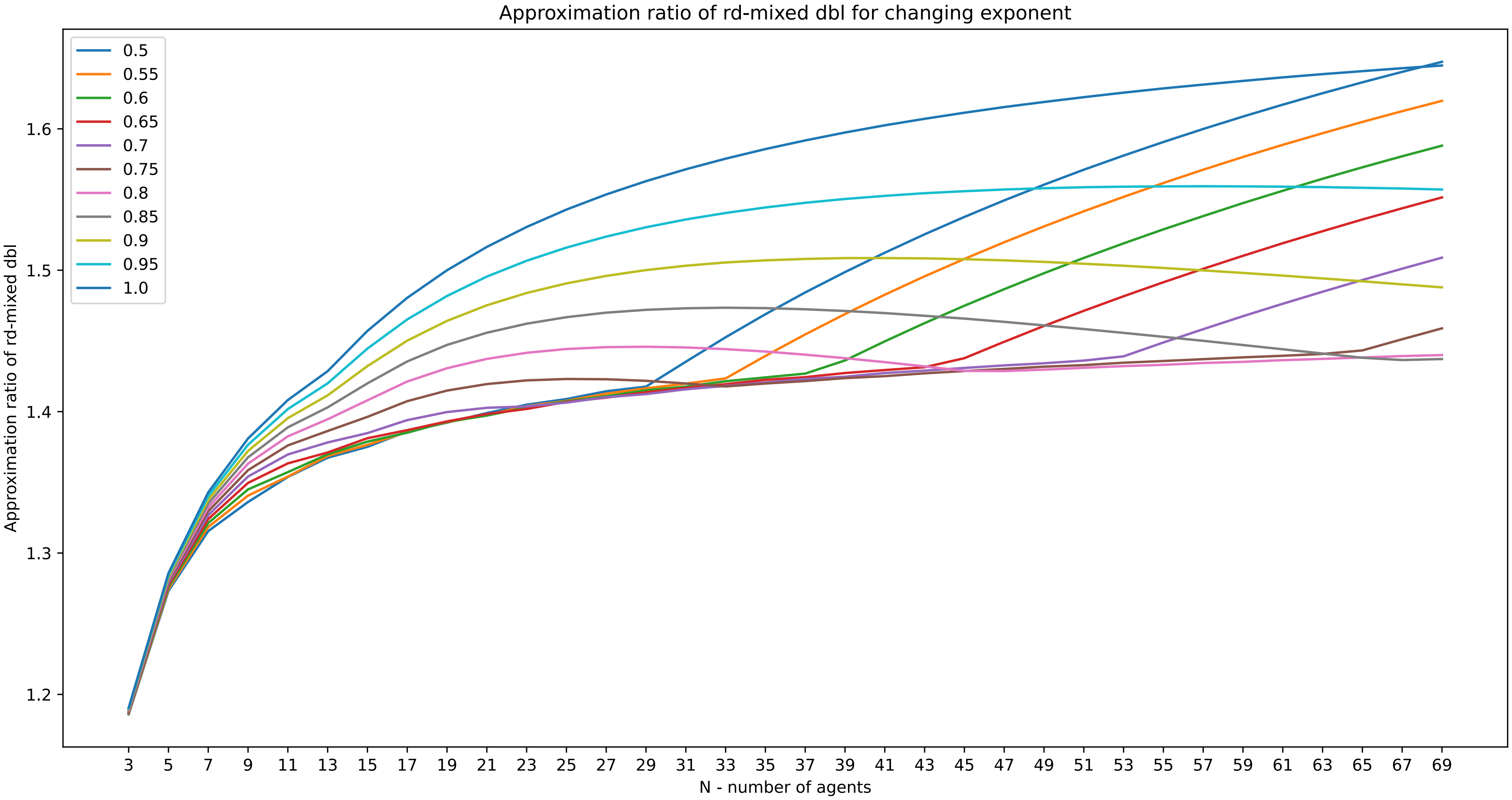
Posłowie

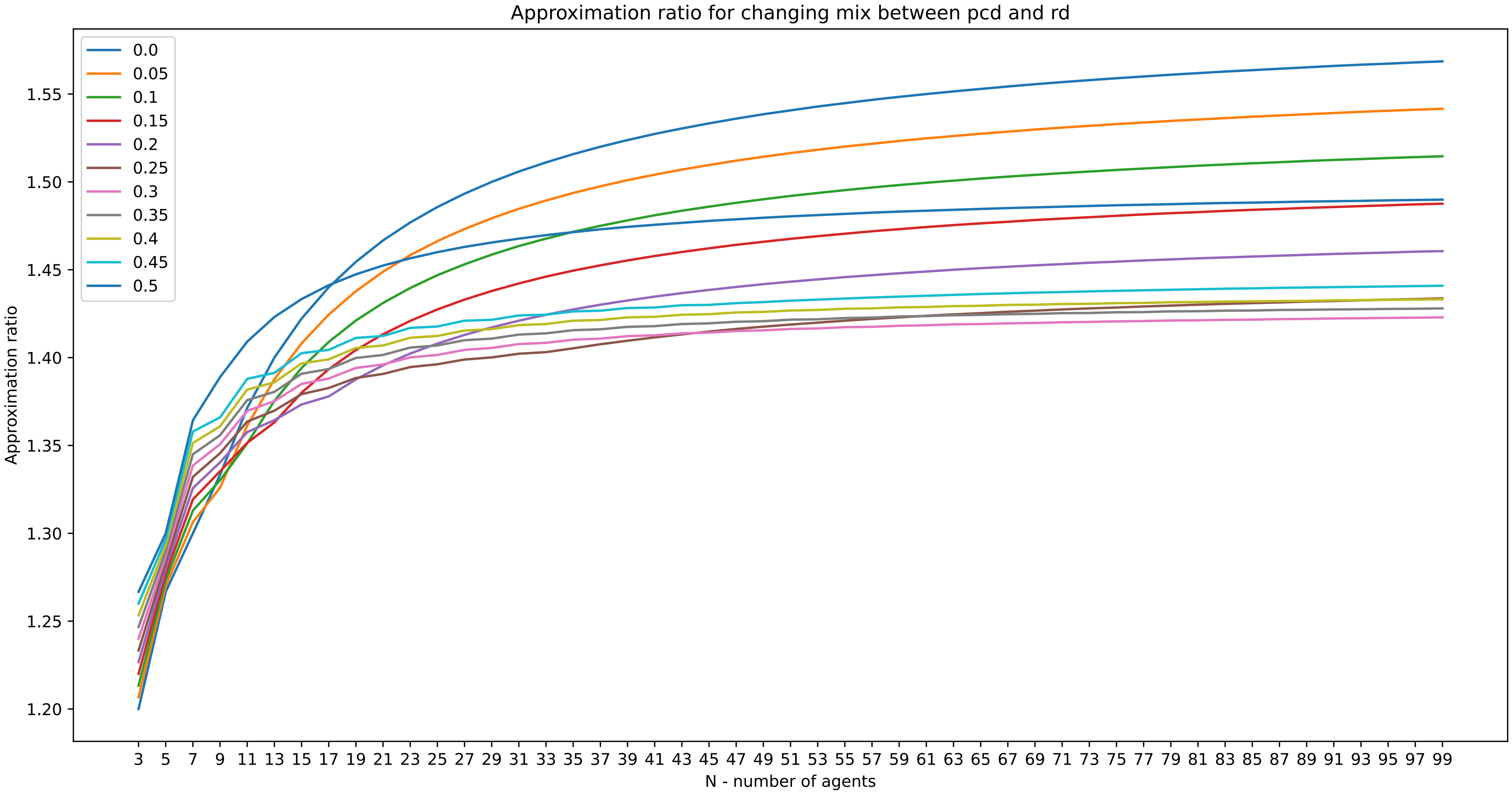
- przemyślenia o szukaniu tematu pracy
- od zaplecza o pracy nad magisterką

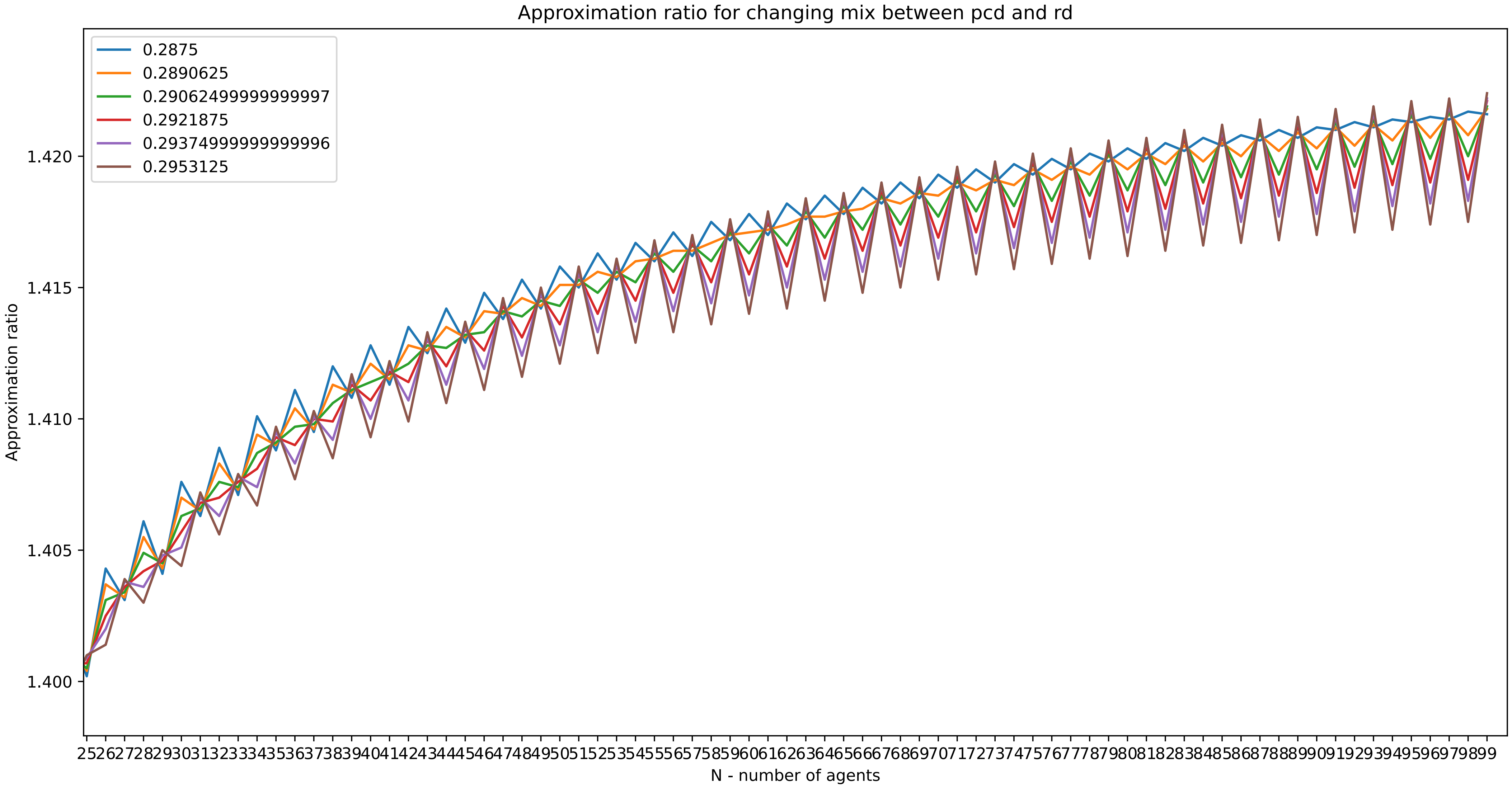


Współczynnik aproksymacji dla V=15
pcd (dbl 1)
dbl 0.9
dbl -1
sqcd (dbl -infty)
1.2
1.3
1.4
1.5
1.6
3
7
11
15
19
23
27
31
35
39
N - number of agents


Współczynnik aproksymacji (RD zmieszane mechanizmy) dla V=15
pcd (dbl 1)
dbl 0.9
dbl -1
sqcd (dbl -infty)
1.2
1.3
1.4
1.5
1.6
1.7
1.8
3
7
11
15
19
23
27
31
35
39
N - number of agents


Approximation ratio of rd-mixed dbl for changing exponent
0.5
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9
0.95
1.0
1.2
1.3
1.4
1.5
1.6
Approximation ratio of rd-mixed dbl
3
5
7
9
11
13
15
17
19
21
23
25
27
29
31
33
35
37
39
41
43
45
47
49
51
53
55
57
59
61
63
65
67
69
N - number of agents


Approximation ratio for changing mix between pcd and rd
0.0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
Approximation ratio
1.20
1.25
1.30
1.35
1.40
1.45
1.50
1.55
N - number of agents
3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59 61 63 65 67 69 71 73 75 77 79 81 83 85 87 89 91 93 95 97 99


Approximation ratio for changing mix between pcd and rd
0.2875
0.2890625
0.29062499999999997
0.2921875
0.29374999999999996
0.2953125
1.420
1.415
1.410
1.405
1.400
Approximation ratio
25262728293031323334353637383940414243444546474849505152535455565758596061626364656667686970717273747576777879808182838485868788899091929394959697989
N - number of agents