



Przejście fazowe w sieciach złożonych w modelu Axelroda

Korzeń W., Maćkowski M., Rozwadowski P., Szczepilewska P., Sznajder W.¹

Opiekun: Tomasz Raducha

¹Uniwersytet Warszawski, Wydział Fizyki

Streszczenie

W poniższej pracy przeanalizowano zachowanie sieci złożonych w modelu Axelroda pomiędzy stanem uporządkowanym a nieuporządkowanym (przejście fazowe). Do badania użyto sieci złożonej typu Small-world, Barabási-Albert oraz Klemm-Eguiluz.

1. Wprowadzenie

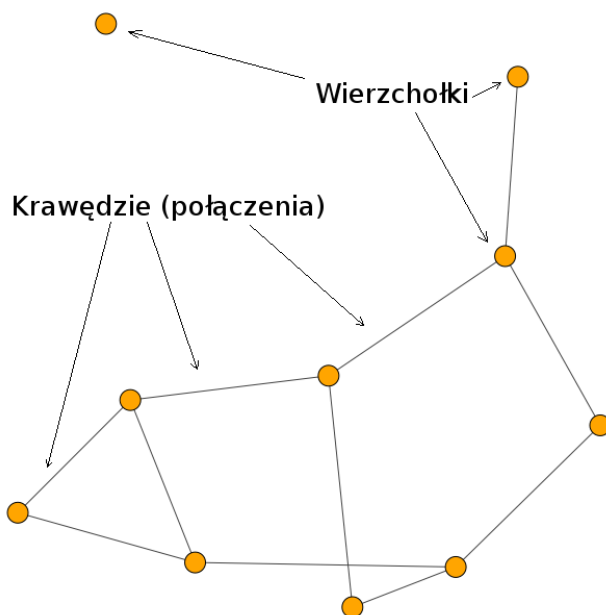
Zadziwiająca ilość układów w otaczającym nas świecie posiada strukturę sieci: sieci społeczne, sieci komputerowe (w tym Internet), sieci komunikacyjne w miastach czy sieci krystaliczne. Pierwsze historyczne próby uporządkowania tej wiedzy zaczęły się już w XVIII wieku za sprawą Eulera, który stworzył teorię grafów. Graf (G) jest strukturą matematyczną, która w uproszczeniu jest zbiorem wierzchołków (V) wraz ze zbiorem krawędzi (E), które je łączą (rysunek 1). Przy pomocy grafów możliwe jest więc przedstawianie i badanie relacji między obiektami w sieci.

Już latach 30. XX wieku teoria grafów znalazła swoje zastosowanie w socjologii. W 1967 roku Stanley Milgram przeprowadził eksperyment ("świat jest mały"), który miał przetestować hipotezę, że członkowie jakiegokolwiek dużej społeczności mogą być połączeni poprzez krótkie sieci pośrednich znajomych. Milgram wysyłał listy do kilkuset losowo wybranych osób z poleceniem by zostały one przekazane do jego przyjaciela (w przypadku, gdyby go nie znali to list miał być wysłany komuś, kto mógłby być bliższy tej osobie). Po średnio sześciokrotnym przekazaniu większości listów trafiły one do przyjaciela Milgrama.

Powstanie ogromnych baz danych pod koniec XX wieku pozwoliło badać sieci rzeczywiste. Najciekawszym elementem badań wydaje się sieć internetowa. Pomimo ogromnych rozmiarów, wykazano, że ma ona cechy przypominający "mały świat" Milgrama.

Obecnie takie sieci zwane sieciami złożonymi (ang. Complex Network) są przedmiotem badań naukowców na całym świecie.

Do symulacji użyto języka programowania *python* wraz z bibliotekami *igraph*, *random*, *math* i *numpy*. Program udostępniono na licencji GNU GPL na stronie internetowej <https://github.com/rozwadowski/AxelrodModel>. W rozdziale drugim opisano wszystkie algorytmy oraz badane wartości. W rozdziale trzecim przedstawiono wyniki symulacji, a w czwartym znajduje się podsumowanie pracy.



Rys. 1.— Przykład grafu.

2. Model teoretyczny

2.1. Model Axelroda

Algorytm:

- Każdy wierzchołek sieci i opisany jest wektorem $s_i = (s_{i,1}, s_{i,2}, \dots, s_{i,F})$, gdzie F to liczba cech (ang. cultural features). Każda cecha może przyjmować jedną z q wartości (ang. cultural traits): $s_{i,f} \in 1, 2, 3, \dots, q$.
- Wartości $s_{i,f}$ są wybrane w sposób losowy, każde z prawdopodobieństwem $p = \frac{1}{q}$.
- Powtarzanie kroku czasowego dopóki nie zostanie osiągnięty tzw. frozen state (w tym przypadku, gdy po 100000 krokach czasowych nie zostanie podjęta żadna czynność)
 - Losowanie wierzchołka i .
 - Losowanie połączenia wychodzącego z wierzchołka i : (i, j) .
 - Obliczenie ilości cech wspólnych l między wierzchołkami i oraz j

$$l(i, j) = \sum_{f=1}^F \delta_{s_{if}s_{jf}}, \quad (1)$$

gdzie δ_{ab} to tak zwana Kroneckera równa 1 (gdy $a = b$) lub 0 (gdy $a \neq b$).

- Jeśli $0 < l(i, j) < F$ to wierzchołki i oraz j wchodzi ze sobą w interakcję z prawdopodobieństwem $l(i, j)/F$.
- W przypadku interakcji wartość g zostaje wylosowana tak by $s_{ig} \neq s_{jg}$, a następnie wartości s_{ig} zostaje przypisana wartość s_{jg} .

2.2. Badane parametry w modelu Axelroda

- *Domena* - grupa wierzchołków w grafie, która posiada taki sam wektor s_i oraz istnieje ścieżka łącząca dowolne dwa wierzchołki.

- *Rozmiar domeny* - ilość wierzchołków w ramach pojedynczej domeny.
- *Rozmiar największej domeny* S_{max} - największa ilość wierzchołków w ramach jednej domeny sieci złożonej.
- *Liczba domen*: N_{domain} - ilość domen w sieci złożonej.
- *Odchylenie standardowe rozkładu* S_{max} : σ - wartość odchylenia standardowego dla rozkładu rozmiaru największej domeny dla zadanego q pochodzących z wielu realizacji symulacji.
- *Moment przejścia fazowego*: q_c - wartość q dla którego σ jest największa.
- *Stan uporządkowany* - stan sieci złożonej w której wierzchołki grupują się w znaczącej wielkości domeny. Umownie, dla $q < q_c$.
- *Stan nieuporządkowany* - stan sieci złożonej w której wierzchołki nie grupują się w znaczącej wielkości domeny bądź nawet występują licznie jako odseparowane. Umownie, dla $q > q_c$.

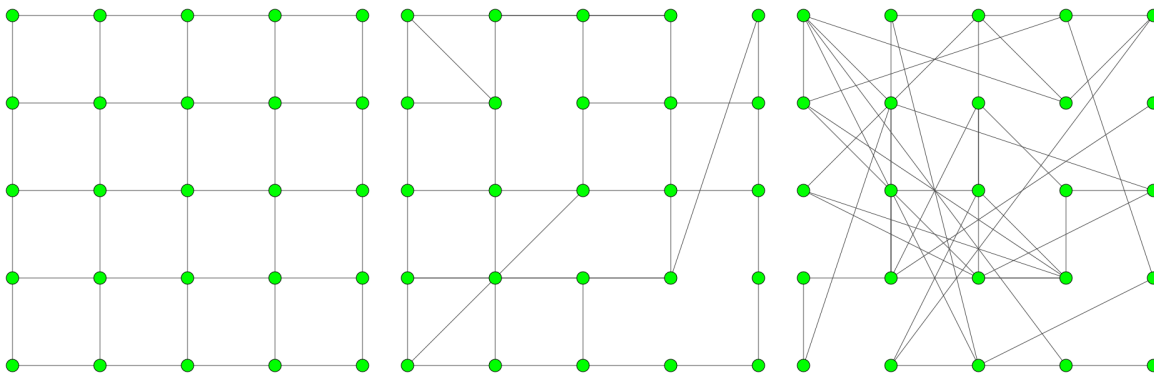
2.3. Model sieci Small-world

Sieci społeczne nie są sieciami w pełni losowymi, lecz mają z nimi pewne cechy wspólne. Najciekawszą własnością jest to, że dowolna para wierzchołków połączona jest ścieżką (d) o wiele mniejszą niż rozmiar sieci (N).

Siecią spełniającą tę cechę jest Sieć Małego Świata zaproponowana w 1998 roku przez Duncana Wattsa i Stevena Strogatza. Model ten powstaje z sieci regularnej w której w sposób losowy zamieniane są wierzchołki w połączeniach. Dzięki takiemu zabiegowi połączenia między dowolnymi wierzchołkami stają się krótsze.

Przykładowe sieci zostały przedstawione na rysunku 2. Dla sieci $p = 0$ mamy do czynienia z siecią regularną, którą charakteryzuje duża ścieżka d . W przypadku $p = 1$ sieć jest bliska losowej, ale wartość d jest mała. Dla wartości pośrednich p sieć wykazuje się małym d . Algorytm tworzenia sieci typu Small-world zastosowany w niniejszej pracy przedstawia się następująco:

- Tworzymy sieć regularną o postaci dwuwymiarowej kraty $L \times L$, gdzie $N = L^2$.
- Z prawdopodobieństwem p krawędź (i, j) zostaje usunięta.
- Na miejsce usuniętej krawędzi losowana jest nowa (i, k) przy założeniu, że połączenia nie mogą się duplikować oraz nie mogą powstawać połączenia typu (i, i) .



Rys. 2.— Przykładowe sieci typu Small-world dla p równego 0.0 (dwuwymiarowa krata), 0.1 i 0.9.

2.4. Model sieci Barabási–Albert

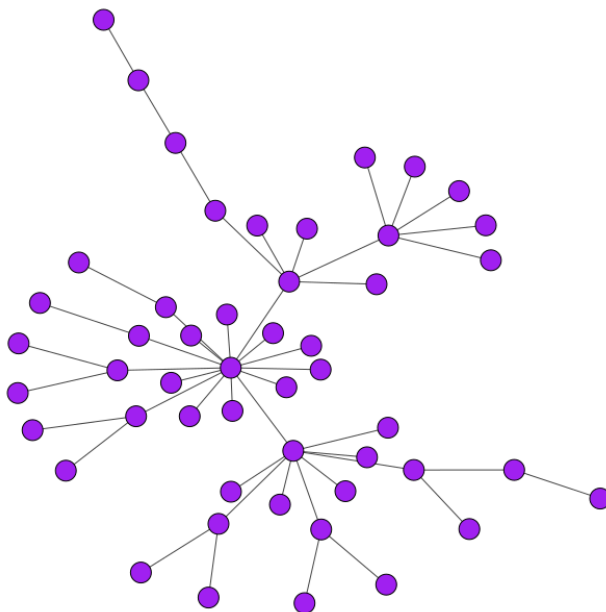
Znaczna ilość sieci wykazuje własności fraktalne. Nazywane są one bezskalowymi. Są to sieci w których rozkład liczby połączeń jest potęgowy (o wykładniku między 2 i 3). Do takich sieci należą sieci społeczne i sieć Internet. W przyrodzie również można znaleźć sieci tego typu jak na przykład sieć interakcji pomiędzy białkami w organizmach żywych. W 1999 roku Albert-László Barabási i Réka Albert zaproponowali algorytm tworzenia sieci bezskalowej.

Sieć taka rośnie w trakcie jej tworzenia na zasadzie preferencyjnego dołączania nowych wierzchołków (jeśli wierzchołek ma duży stopień to zwiększa to prawdopodobieństwa dołączenia do niego kolejnego). W wyniku rozrostu sieci niewielka liczba wierzchołków posiada dużą liczbę połączeń, a duża liczba wierzchołków posiada niewielką liczbę połączeń. Przykładowy graf Barabási–Albert pokazano na rysunku 3. W niniejszej pracy do wygenerowania grafu Barabási–Albert użyto funkcji *Graph.Barabasi()* z biblioteki *igraph* (język programowania *python*). Algorytm tworzenia sieci Barabási–Albert:

- Tworzymy graf pełny o wielkości $m_0 \geq 1$.
- Do sieci dodawany jest węzeł, który tworzy $m \leq m_0$ połączeń do istniejących wierzchołków sieci. Prawdopodobieństwo, że nowy wierzchołek utworzy połączenie do istniejącego już węzła jest proporcjonalne do stopnia tego węzła k_i :

$$P(k_i) = \frac{k_i}{\sum k_i} \quad (2)$$

- Sieć rośnie do momentu, gdy ilość wierzchołków osiągnie liczbę N



Rys. 3.— Przykładowa sieć Barabási–Albert.

2.5. Model sieci Klemm-Eguiluz

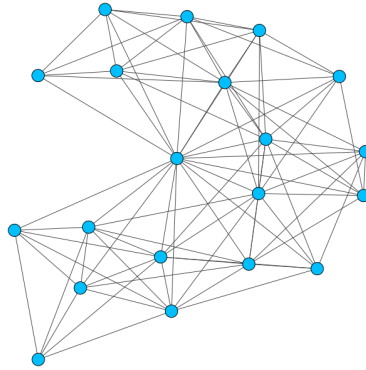
Główną wadą sieci bezskalowych (jak na przykład Barabási–Albert) jest to, że posiadają niski w porównaniu z rzeczywistymi sieciami współczynnik klasteryzacji C , który określa stosunek liczby połączeń między sąsiadami wierzchołka do liczby wszystkich możliwych połączeń.

Konstantin Klemm i Victor Eguiluz zaproponowali algorytm, który tworzy sieci bezskalowe wraz ze współczynnikiem klasteryzacji na poziomie około $5/6$. Przykładową sieć Klemm-Eguiluz przedstawiono na rysunku 4. Algorytm:

- Każdy nowy wierzchołek znajduje się w stanie aktywnym (jest zdolny do tworzenia nowych połączeń) lub dezaktywowanym (nie jest zdolny do tworzenia nowych połączeń).
- Tworzymy graf pełny o wielkości m ze wszystkimi aktywnymi wierzchołkami.

- Dodanie nowego wierzchołka oraz m połączeń do aktywnych wierzchołków.
- Dezaktywacja jednego z wierzchołków, wybór z prawdopodobieństwem odwrotnie proporcjonalnym do stopnia tego węzła:

$$P(k_i) = \frac{\frac{1}{k_i}}{\sum k_i} \quad (3)$$

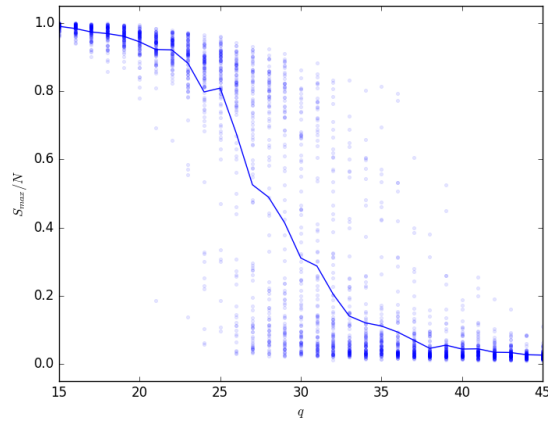


Rys. 4.— Przykładowa sieć Klemm-Eguluz.

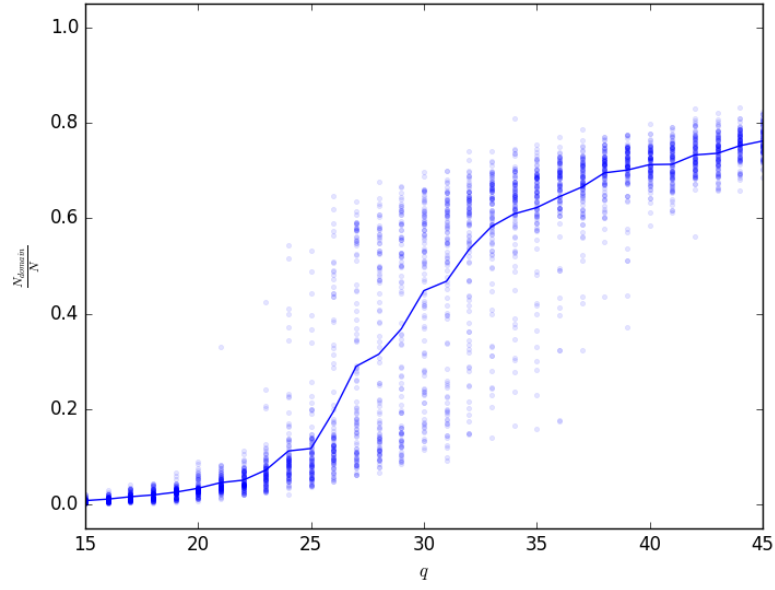
3. Wyniki

3.1. Stan uporządkowany, nieuporządkowany, przejście fazowe.

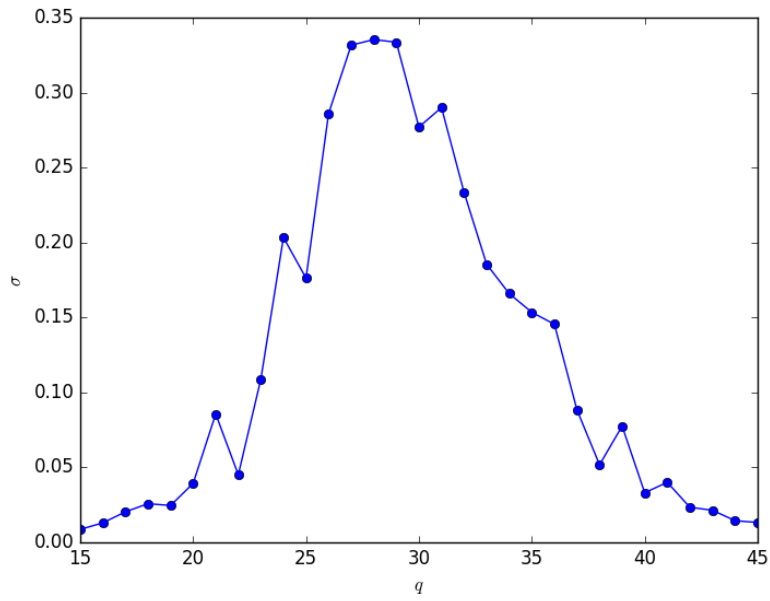
Na rysunkach 5 i 6 przedstawiono wyniki symulacji algorytmu Axelroda na sieci typu Small-world o rozmiarze 20×20 wierzchołków z parametrami $F = 5$ oraz $p = 0.1$. Dla odpowiednio niskiego q (czyli ilości możliwych wartości dla każdego elementu wektora s_{if}) rezultatem symulacji jest niewielka liczba dużych domen. Taki stan nazywamy uporządkowanym. Analogicznie dla odpowiednio dużego q powstaje duża liczba małych domen - stan nieuporządkowany. Głównym celem analizy jest zbadanie przejścia fazowego q_c , które charakteryzuje moment przejścia ze stanu uporządkowanego w nieuporządkowany. Na rysunku 5 widać wyraźnie dwie grupy wokół których zbierają się wyniki - skupione wokół $S_{max} = 1$ dla małych q oraz skupione wokół $S_{max} = 0$ dla dużych q . W okolicy szukanego q_c znajdują się punkty z obydwu grup. W związku z tym najlepszą metodą na ustalenie q_c jest analiza odchylenia standardowego. Tam gdzie jest ona największa, tam znajduje się q_c . Dla podanego przykładu $q_c = 28$.



Rys. 5.— Wykres S_{max}/N - q dla 96 realizacji algorytmu Axelroda (linia ciągła oznacza średnią).



Rys. 6.— Wykres $N_{\text{domain}}/N-q$ dla 96 realizacji algorytmu Axelroda (linia ciągła oznacza średnią).



Rys. 7.— Wykres $\sigma-q$ dla 96 realizacji algorytmu Axelroda (linia ciągła oznacza średnią).

3.2. Model sieci Small-world

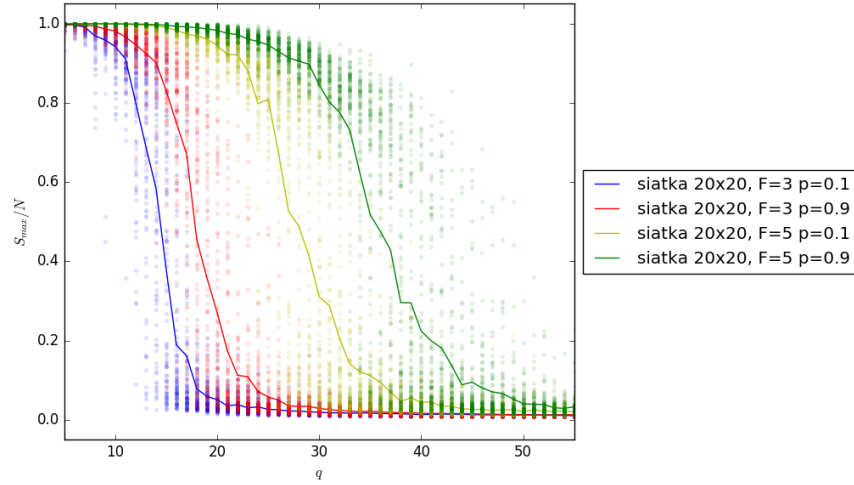
Przeprowadzono analizę miejsca przejścia fazowego w zależności od parametrów F i p (rysunki 8-10) oraz N (rysunki 11-13). Na rysunkach 8 i 9 przy pomocy kropek pokazano wynik każdej realizacji symulacji, a linią ciągłą wartość średnią. Ze względu na czytelność wykresów, w dalszej części pracy znajdują się jedynie linie ciągłe reprezentujące średnie wartości. W tabeli 1 przedstawiono otrzymane wartości q_c .

Wzrost liczby wierzchołków N powoduje przesunięcie przejścia fazowego w kierunku większych wartości. Można to zinterpretować w ten sposób, że większa populacja ma większe szanse na utworzenie się dużych grup (domen) i stanu uporządkowanego, a populacja niewielka szybko kończy proces uzgadniania cech s_{if} . W przypadku siatki 10x10 nie zaobserwowano zmian q_c dla różnych wartości p . Dopiero przy zwiększeniu liczby wierzchołków (do 400) wyraźnie występuje podobna zależność - q_c przesuwa się w kierunku większych wartości. Sieć o większych cechach sieci losowej ($p \rightarrow 1$) ma lepszą zdolność do utworzenia jednej dużej domeny. Sieci bliskie kraty regularnej mają tę własność, że krótka ścieżka d nie jest w stanie objąć dużej liczby wierzchołków. Wzrost parametru F również przesuwa przejście fazowe w prawo - wiąże się to z tym, że im więcej cech można uzgadniać tym łatwiej wierzchołki mogą wchodzić w interakcję.

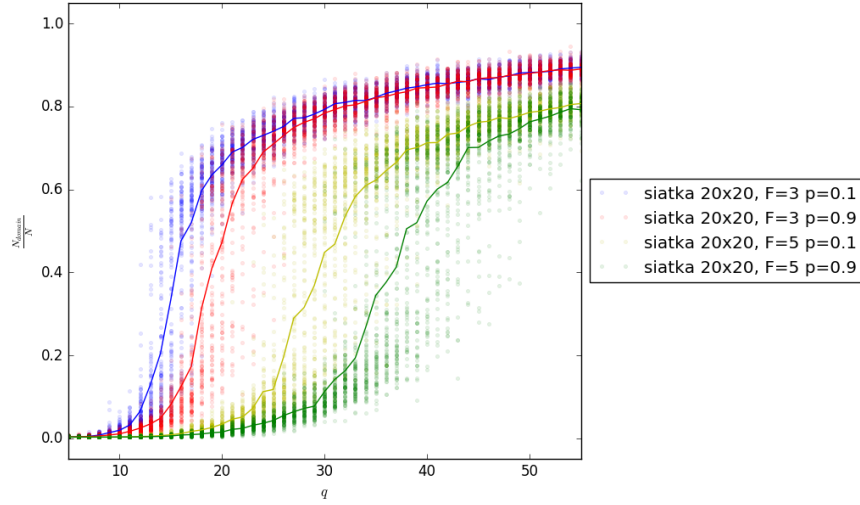
Dla wartości $q \ll q_c$ obserwujemy stan uporządkowany, w którym cała sieć znajduje się w jednej domenie. Dla wartości $q \gg q_c$ następuje bardzo szybki skok do stanu nieuporządkowanego, w którym cała sieć składa się z domen o wielkości jednego wierzchołka.

Tabela 1: Wartości przejścia fazowego q_c w modelu Axelroda w zależności od parametrów N , p oraz F sieci typu Small-world.

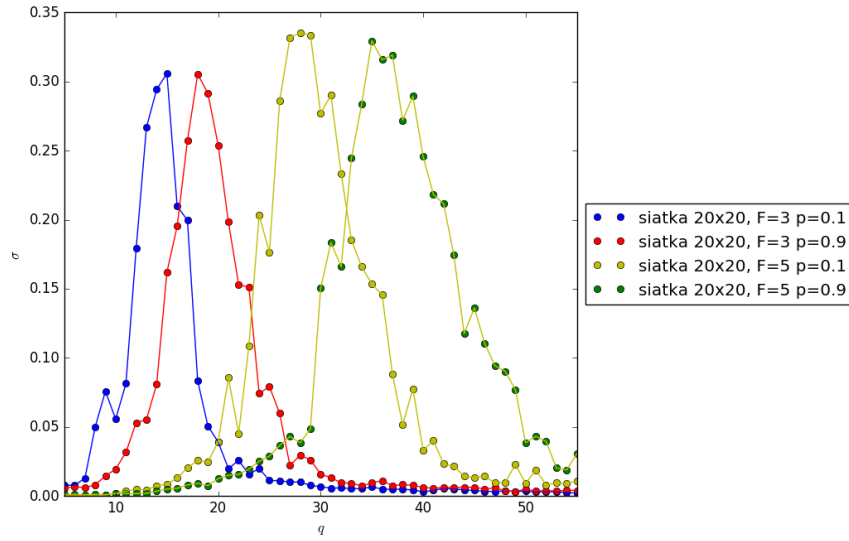
F	p	$N = L^2$	q_c
3	0.1	400	15
3	0.9	400	18
5	0.1	400	28
5	0.9	400	36
3	0.1	100	9
3	0.9	100	9



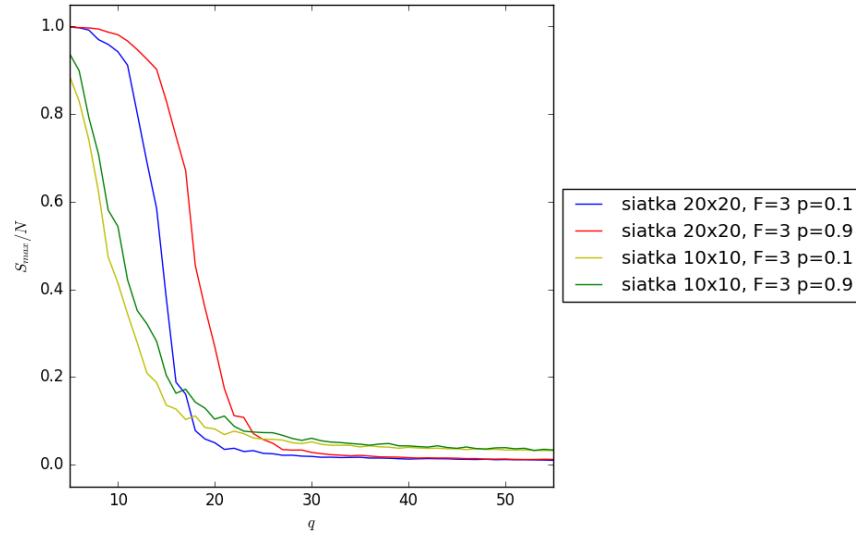
Rys. 8.— Wykres S_{max}/N - q dla różnych F i p w modelu Small-world. Dla każdego q wynik uśredniono z 96 realizacji.



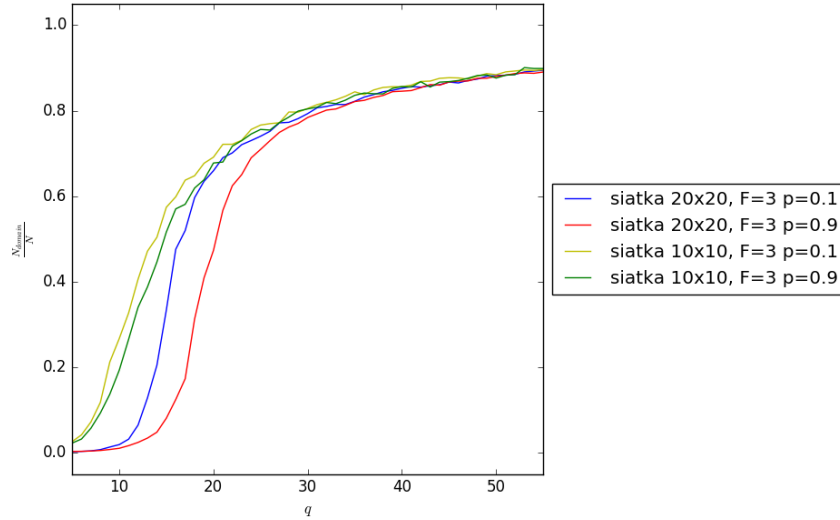
Rys. 9.— Wykres $N_{domain}/N-q$ dla różnych F i p w modelu Small-world. Dla każdego q wynik uśredniono z 96 realizacji.



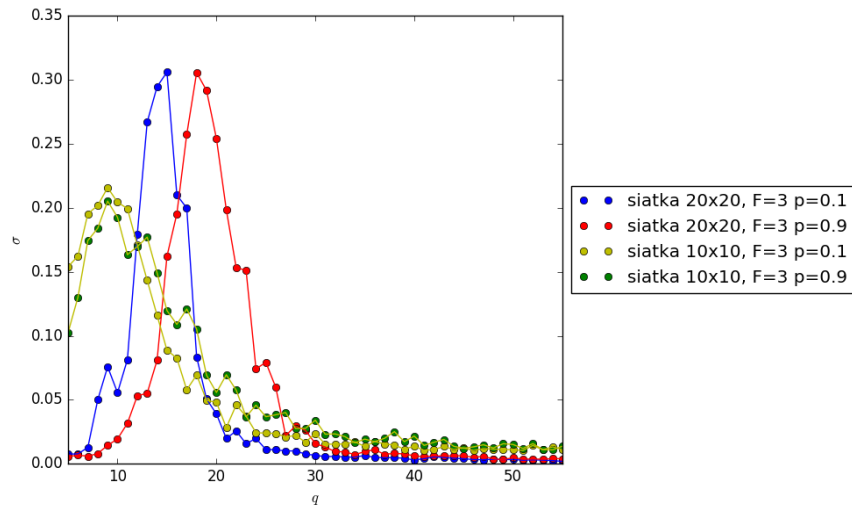
Rys. 10.— Wykres $\sigma-q$ dla różnych F i p w modelu Small-world. Dla każdego q wynik uśredniono z 96 realizacji.



Rys. 11.— Wykres $S_{max}/N-q$ dla różnych N w modelu Small-world. Dla każdego q wynik uśredniono z 96 realizacji.



Rys. 12.— Wykres $N_{domain}/N-q$ dla różnych N w modelu Small-world. Dla każdego q wynik uśredniono z 96 realizacji.



Rys. 13.— Wykres σ - q dla różnych N w modelu Small-world. Dla każdego q wynik uśredniono z 96 realizacji.

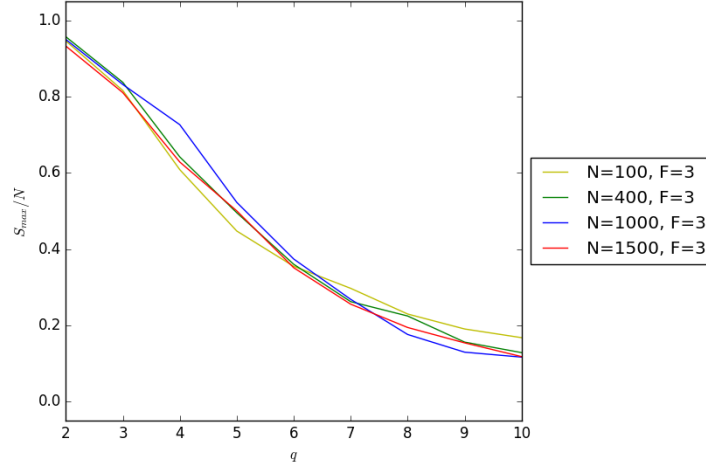
3.3. Model sieci Barabási–Albert

Przeprowadzono analizę miejsca przejścia fazowego w zależności od parametrów F (rysunki 17-19) oraz N (rysunki 14-16). W tabeli 2 przedstawiono otrzymane wartości q_c .

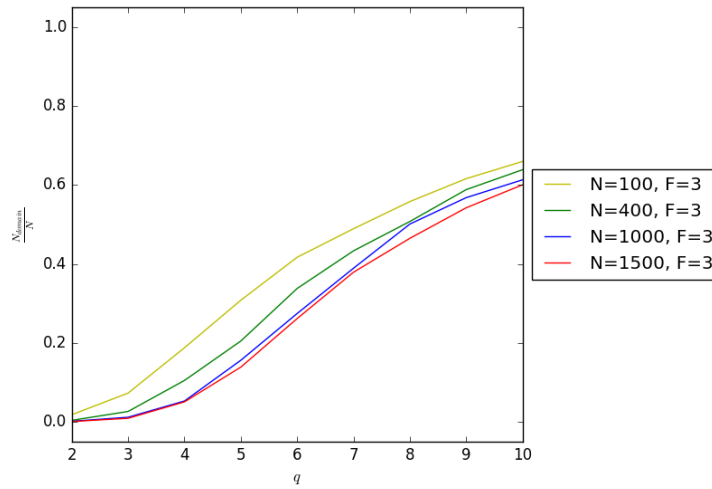
Symulacja dla $N = 100$ i $N = 200$ nie wykazała zmiany przejścia fazowego, dlatego zasymulowano dodatkowo $N = 1000$ i $N = 1500$. Dopiero wtedy zaobserwowano przesunięcie q_c , jest one jednak dużo wolniejsze niż w przypadku modelu Small-world. Zwiększenie parametru F również zwiększa q_c . Zastosowanie sieci bezskalowej z niskim współczynnikiem klastrowania (czyli Barabási–Albert) powoduje, że sieć jest bardziej odporna na zmiany liczby wierzchołków i liczby cech.

Tabela 2: Wartości przejścia fazowego q_c w modelu Axelroda w zależności od parametrów N oraz F sieci Barabási–Albert.

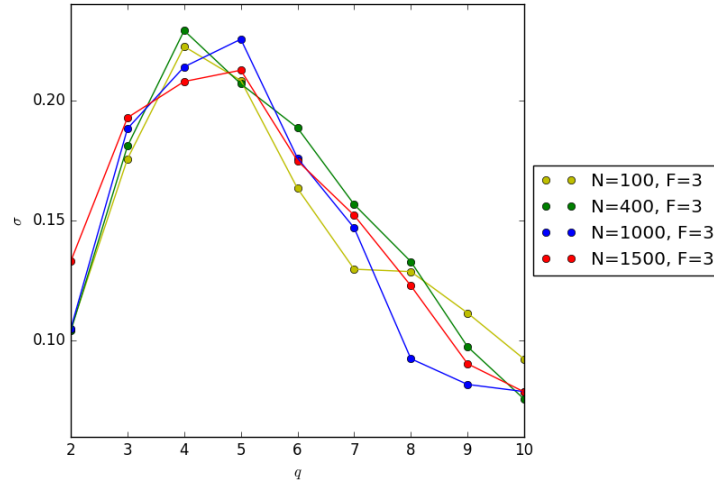
F	N	q_c
3	400	4
3	1000	5
3	1500	5
3	100	4
4	100	6
5	100	7



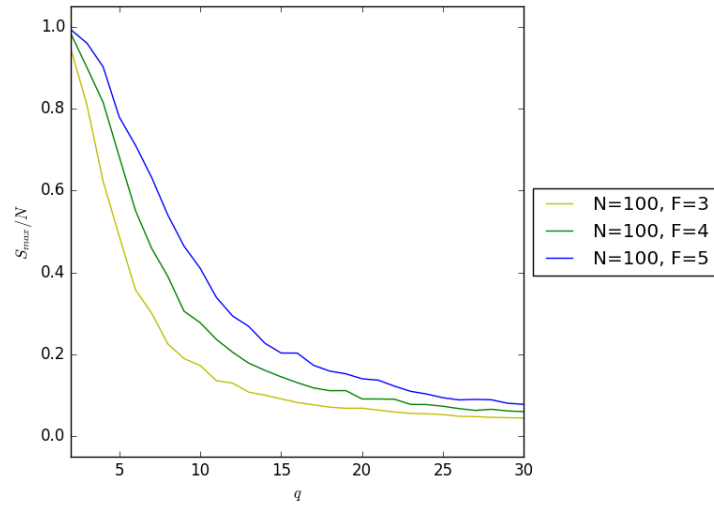
Rys. 14.— Wykres $S_{max}/N-q$ dla różnych N w modelu Barabási–Albert. Dla każdego q wynik uśredniono z 96 realizacji.



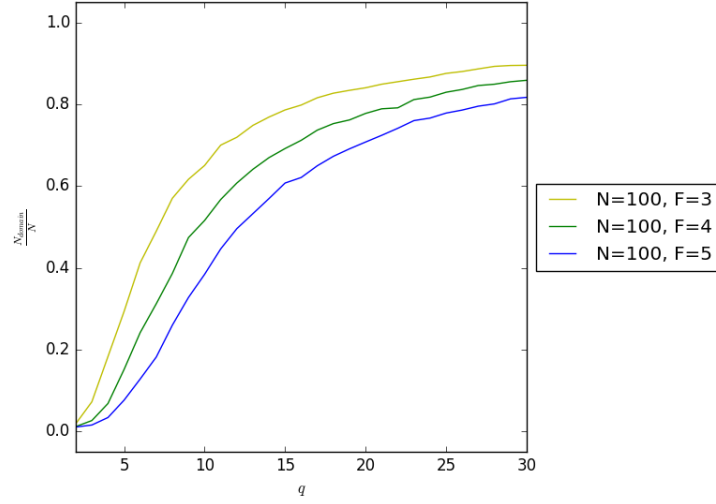
Rys. 15.— Wykres $N_{domain}/N-q$ dla różnych N w modelu Barabási–Albert. Dla każdego q wynik uśredniono z 96 realizacji.



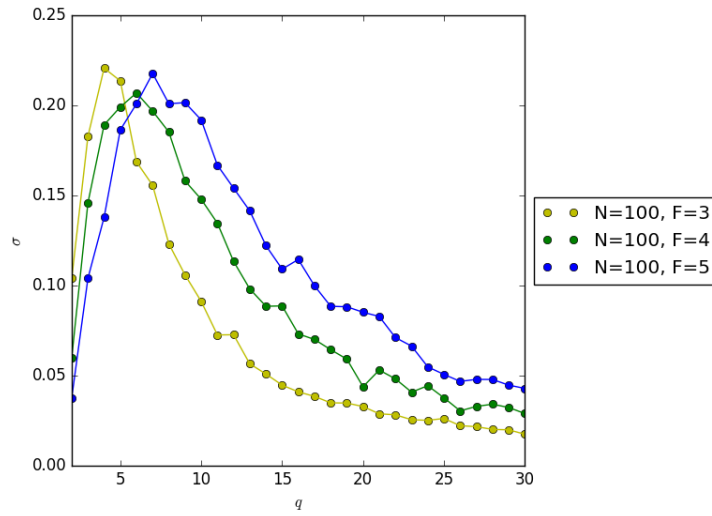
Rys. 16.— Wykres σ - q dla różnych N w modelu Barabási–Albert. Dla każdego q wynik uśredniono z 96 realizacji.



Rys. 17.— Wykres S_{max}/N - q dla różnych F w modelu Barabási–Albert. Dla każdego q wynik uśredniono z 256 realizacji.



Rys. 18.— Wykres $N_{domain}/N-q$ dla różnych F w modelu Barabási–Albert. Dla każdego q wynik uśredniono z 256 realizacji.



Rys. 19.— Wykres $\sigma-q$ dla różnych F w modelu Barabási–Albert. Dla każdego q wynik uśredniono z 256 realizacji.

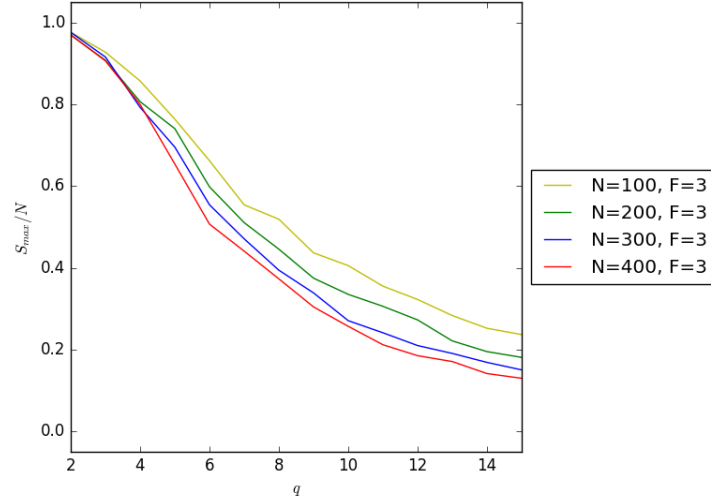
3.4. Model sieci Klemm-Eguiluz

Przeprowadzono analizę miejsca przejścia fazowego w zależności od parametrów F (rysunki 23-25) oraz N (rysunki 14-22). W tabeli 3 przedstawiono otrzymane wartości q_c .

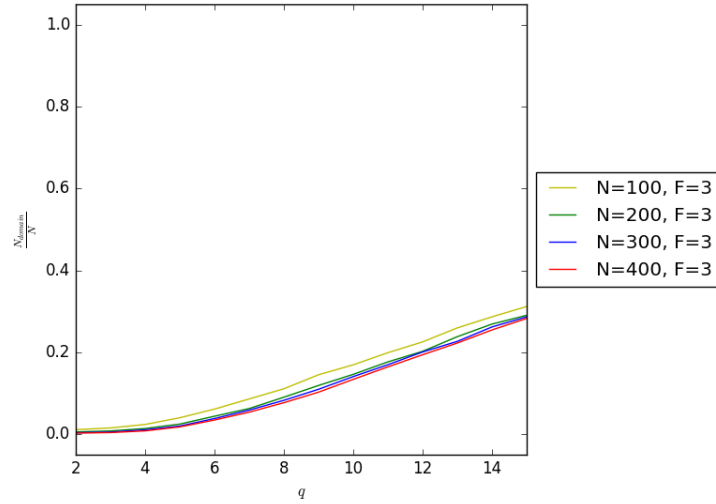
Symulacja dla różnych N nie wykazała zmiany przejścia fazowego. Ze względu na złożoność obliczeniową potrzebną do stworzenia sieci i realizacji symulacji ograniczono się do $N = 400$. Zwiększenie parametru F zwiększa q_c , jednakże "skok" pomiędzy $F = 4$ i $F = 5$ jest znacznie mniejszy niż pomiędzy $F = 3$ i $F = 4$. Warto zauważyć, że dla $q > q_c$ stosunek liczby domen do wszystkich wierzchołków dąży do pewnej wartości równej około 40%. S_{max}/N wykazuje podobną zależność. Wskazuje to na pewne uporządkowanie pomimo rosnącego q .

Tabela 3: Wartości przejścia fazowego q_c w modelu Axelroda w zależności od parametrów N oraz F sieci Klemm-Eguiluz.

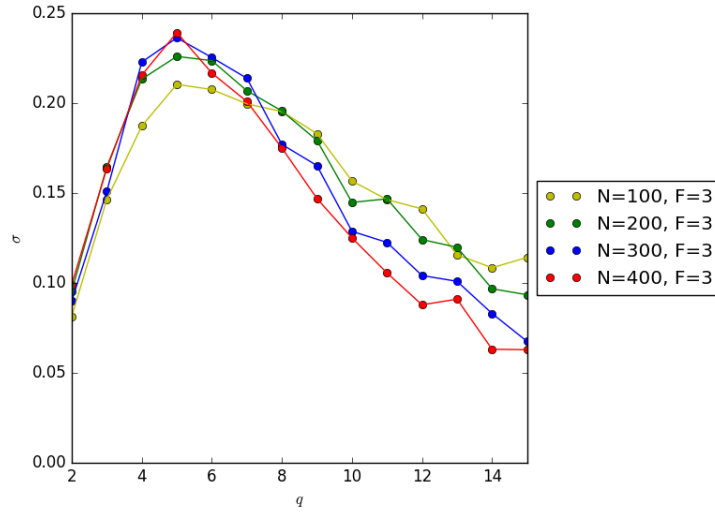
F	N	q_c
3	200	5
3	300	5
3	400	5
3	100	5
4	100	11
5	100	13



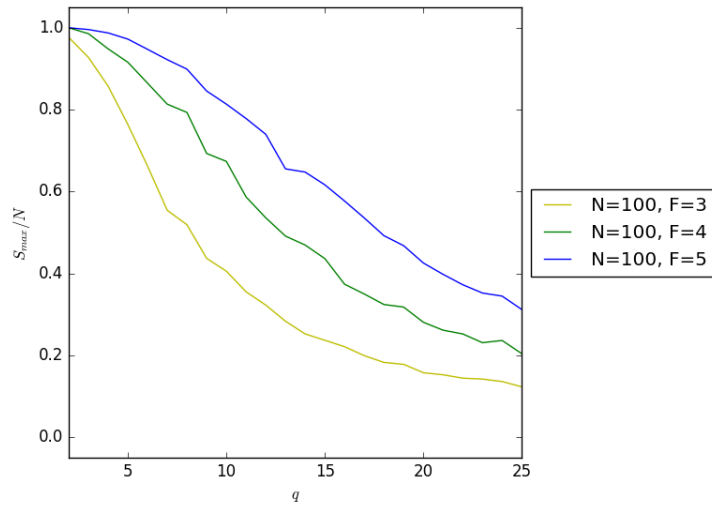
Rys. 20.— Wykres $S_{max}/N-q$ dla różnych N w modelu Klemm-Eguiluz. Dla każdego q wynik uśredniono z 256 realizacji.



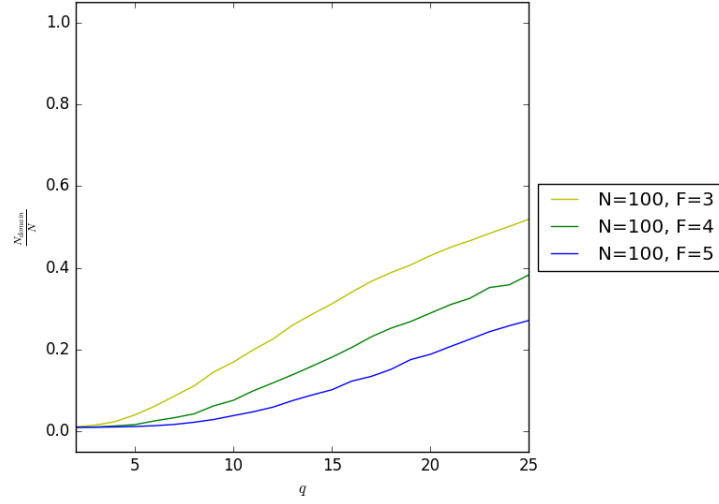
Rys. 21.— Wykres $N_{domain}/N-q$ dla różnych N w modelu Klemm-Eguiluz. Dla każdego q wynik uśredniono z 256 realizacji.



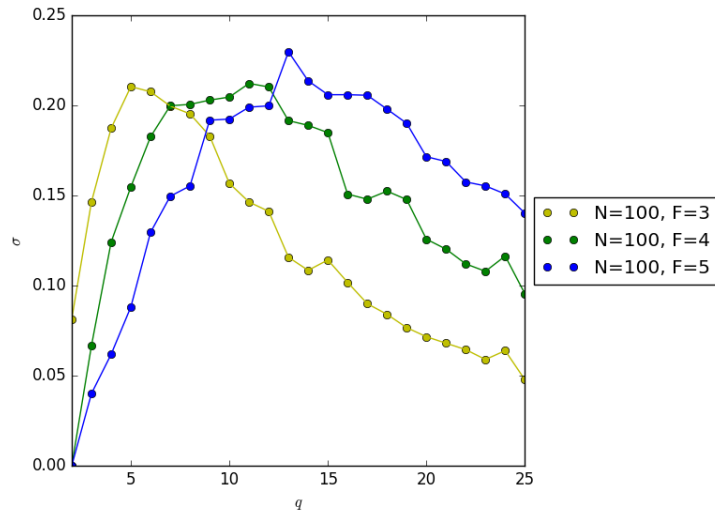
Rys. 22.— Wykres σ - q dla różnych N w modelu Klemm-Eguiluz. Dla każdego q wynik uśredniono z 256 realizacji.



Rys. 23.— Wykres S_{max}/N - q dla różnych F w modelu Klemm-Eguiluz. Dla każdego q wynik uśredniono z 256 realizacji.



Rys. 24.— Wykres $N_{domain}/N-q$ dla różnych F w modelu Klemm-Eguiluz. Dla każdego q wynik uśredniono z 256 realizacji.



Rys. 25.— Wykres $\sigma-q$ dla różnych F w modelu Klemm-Eguiluz. Dla każdego q wynik uśredniono z 256 realizacji.

4. Podsumowanie i wnioski

W powyższej pracy znaleźliśmy punkty przejścia fazowego w modelu Axelroda dla trzech typów sieci złożonej: sieci typu Small-world utworzonej z kraty regularnej, sieci bezskalowej Barabási–Albert oraz sieci bezskalowej z dużym klastrowaniem Klemm-Eguiluz.

Wykazaliśmy, że dla sieci Small-world funkcje $q_c(N)$, $q_c(p)$ i $q_c(F)$ są rosnące co jest zgodne z intuicją: w dużych sieciach (np. społecznościowych) łatwiej o stworzenie uporządkowania poprzez interakcje. Zmniejszenie odległości między wierzchołkami sieci poprzez losowe przepisywanie połączeń (a także zwiększenie liczby cech, które mogłyby być uzgadniane) również zwiększa szansę na stworzenie uporządkowania w sieci. Siecią rzeczywistą, która przypomina taką sieć jest sieć krystaliczna, a losowe przepisane połączone są analogiczne do defektów w kryształach. Przy użyciu takiej analogii q_c jest miejscem, w którym następuje na przykład przejście ze stanu stałego (stan uporządkowany) do ciekłego (stan nieuporządkowany).

Dla sieci rosnącej Barabási–Albert zaobserwowaliśmy mniejsze tempo wzrastania funkcji $q_c(N)$, które jest konsekwencją zbliżonej do fraktalnej natury sieci - niezależnie od wielkości cechy sieci są zbliżone (choć nie identyczne). Sieci rzeczywiste, wykazują podobne własności, ale są one sieciami z wysokim współczynnikiem klastrowania. Do takich sieci należy sieć Klemm-Eguiluz. Wykazuje ona stałą funkcję $q_c(N)$, a ponadto tworzą się w niej uporządkowane strefy nawet dla $q > q_c$. Analogicznym przykładem z rzeczywistości mogą być sieci społecznościowe w Internecie, w których w dużych skalach (miliony użytkowników) sieć wydaje się nieuporządkowana, podczas gdy, w małej skali tworzą się uporządkowane grupy znajomych (związane z np. wspólnym miejscem zamieszkania, pracą, szkołą).

Różne modele sieci złożonych są świetnym narzędziem do testowania i badania sieci rzeczywistych. Podczas, gdy sieć rzeczywistą często mamy w jednym egzemplarzu (np. Internet) to sieci złożone można analizować przy użyciu metod Monte Carlo, które

polega na losowaniu wielu zbliżonych do siebie egzemplarzy danego modelu. Umożliwia to statystyczne badanie rzeczywistych sieci przy użyciu komputerów. Jedynym ograniczeniem jest moc obliczeniowa komputerów, która była główną przeszkodą w skutecznym badaniu przejścia fazowego w modelu Axelroda. Przy użyciu komputerów do użytku domowego możliwe jest uzyskanie wyników na sieciach składających się do około 1000 wierzchołków. Sieci rzeczywiste, takie jak Internet, zawierają miliony wierzchołków, a do efektywnego badania zachowań takiej sieci potrzebne są superkomputery.

Bibliografia

- Castellano C., Marsili M., Vespignani A., "Nonequilibrium Phase Transition in a Model for Social Influence", Physical Review Letters, 85, 3536-3539
- Fronczak A., Fronczak P., "Świat sieci złożonych. Od fizyki do Internetu.", Wydawnictwo Naukowe PWN, 2009
- Klemm K., Eguiluz V. M. , "Nonequilibrium transitions in complex networks: A model of social interaction", Physical Review E 67, 026120
- Klemm K., Eguiluz V. M. , "Growing scale-free networks with small-world behavior", Physical Review , Volume 65, 057102
- Vazquez, F. et al., "Time-scale competition leading to fragmentation and recombination transitions in the coevolution of network and states", Physical Review E 76, 046120