



**Andrzej Jarynowski**

(UJ, CIOP-PIB, NUM, US, WIHiE)

**Andrzej Buda**

(IFJ PAN)

**Piotr Nyczka**

(KU UW<sub>r</sub>)

# **OBLICZENIOWE NAUKI SPOŁECZNE W PRAKTYCE**

## **WSPÓŁCZESNA SOCJOFIZYKA I EKSTRAKCJA WIEDZY**

Socjofizyka, społeczne układy złożone, modelowanie matematyczne i symulacje komputerowe w naukach społecznych, cyfrowa humanistyka, obliczeniowe nauki społeczne (ang. *computational social science*)

Wydanie 1 - pilotażowe

ISBN 978-83-63089-92-4

Wydawnictwo Niezależne, Głogów/ Wrocław

Wydrukowane w Szczecinie przez booksfactory.pl

Projekt okładki: Anna Jankowska

© 2014 Wykorzystywanie treści i fragmentów tej książki w komercyjnych publikacjach (z wyjątkiem recenzji) bez zezwolenia autorów – zabronione. Kopiowanie i rozpowszechnianie oraz wykorzystanie w innych dziełach z podaniem źródła – dozwolone.

## **OBLICZENIOWE NAUKI SPOŁECZNE W PRAKTYCE - WSPÓŁCZESNA SOCJOFIZYKA I EKSTRAKCJA WIEDZY**

### Współpraca przy rozdziałach książki:

Wprowadzenie do metodologii, filozofii oraz historii nauk obliczeniowych w kontekście społecznym – Andrzej Buda (IFJ), Piotr Nyczka (KU UWr)  
Analiza sieci społecznych i sieci złożonych – Andrzej Buda (IFJ)  
Analiza sieciowa w zastosowaniach – Anita Zbieg (IP UWr), Andrzej Buda (IFJ)  
Modelowanie epidemiologiczne – Andrzej Grabowski (CIOP-PIB)  
Modele rynku fonograficznego – Andrzej Buda (IFJ)  
Formowanie się opinii – Piotr Nyczka (KU UWr)  
Polaryzacja partyjna po katastrofie smoleńskiej - hierarchiczna struktura partii politycznych w Polsce – Andrzej Buda (IFJ)

### Rozwinięcie afiliacji twórców książki:

KU UWr- Katedra Unesco Studiów Interdyscyplinarnych, Uniwersytet Wrocławski  
IP UWr- Instytut Psychologii, Uniwersytet Wrocławski  
IFJ- Zakład Układów Złożonych, Instytut Fizyki Jądrowej, Polska Akademia Nauk  
US- Instytut Socjologii, Uniwersytet w Sztokholmie  
UJ- Zakład Teorii Układów Złożonych, Uniwersytet Jagielloński  
WIHiE- Wojskowy Instytut Higieny i Epidemiologii  
CIOP-PIB- Pracownia Techniki Wirtualnej Rzeczywistości, CIOP - Państwowy Instytut Badawczy  
PUM- Państwowy Uniwersytet Mołdawski

W pracy wykorzystano wtórnie przetworzone fragmenty opublikowanych przez autorów tekstów:

Wirtualne aspekty nauki i techniki (A. Jarynowski, racjonalista.pl), Socjofizyka, czyli wkład nauk przyrodniczych w analizę społeczeństw (A. Jarynowski, Nauka Prowadzi w Przyszłość, UJ), Modelowanie epidemiologiczne przy wykorzystaniu analizy tymczasowych sieci społecznych (A. Jarynowski, Postępy inżynierii biomedycznej, UR), Dynamika i struktura na rynkach finansowych i fonograficznych (A. Buda, doktorat IFJ PAN), Modelowanie q-p wyborcy (P. Nyczka, doktorat KU UWr)

### Podziękowania:

Książka powstała dzięki zaangażowaniu: środowiska Polskiego Towarzystwa Fizycznego – Sekcja Fizyki w Naukach Społecznych w Ekonomii, Katedry UNESCO Studiów Interdyscyplinarnych UWr, Centrum Zastosowań Matematyki PG, serii seminariów SNA@PWr oraz Enigmas of Change – UJ, współpracy z Instytutami Socjologii na UWr, UJ oraz Uniwersytecie Sztokholmskim oraz współpracy z grupą NLP przy PWr. Imienne podziękowania należą się Fredrikowi Liljerosowi, Luisowi Roche, Lisie Brouwers, Florentinowi Paladinemu, Jarosławowi Kwapieniowi, Wojciechowi Okrasińskiemu, Januszowi Miśkiewiczowi, Barbarze Pabjan, Zdzisławowi Burdzie, Ewie Rudnickiej, Maciejowi Piaseckiemu, Agnieszce Czaplickiej, Marcie Kliś, Hernanowi Mondanemu, Xinowi Lu, Krzysztofowi Kułakowskiemu, Katarzyna Sznajd-Weron, Danowi Frankowi. Autor mógł poświęcić czas na przygotowanie materiałów dzięki wsparciu Erasmus Mundus (IANUS II).

## SŁOWEM WSTĘPU OD AUTORA

Co łączy zjawiska występujące w przyrodzie i procesy społeczne? To, że ich charakter próbuje się ująć w sztywne ramy modelu. Dlaczego reprezentanci nauk ścisłych zajmują (czy mogą zajmować) się naukami społecznymi? Czym są sieci społeczne a czym układy złożone? Socjofizyka – czy szerzej obliczeniowa socjologia, bo o niej jest mowa w tej książce i zajmuje się opisanymi wyżej problemami, łączy metodologię i narzędzia matematyczne przyrodników ze sferą ludzkich zachowań społecznych. Socjologowie od dawna zastanawiali się, jak mierzyć nastroje społeczne, czy inne charakterystyki społeczeństwa, a także jak przewidzieć i zrozumieć przemiany społeczne. Przyrodnicy i matematycy wykorzystali wyniki tych pomiarów i postanowili wykorzystać je w modelach bazujących zazwyczaj na analogii ze znanymi zjawiskami przyrodniczymi. Bogactwo modeli i koncepcji rozwijanych przez cały wiek XX, zaowocowały powstaniem samodzielnych dziedzin nauki: cybernetyki społecznej, czy młodszej - socjologii analitycznej. Do tej pory ukazało się kilka pozycji literaturowych prezentujących część zagadnień poruszonych w tej książce jak napisane przez fizyków: „Modelowanie Rzeczywistości”, „Świat sieci złożonych”, czy przez socjologów: „Wprowadzenie do dynamiki społecznej”, „Układy złożone w naukach społecznych” „Modelowanie matematyczne i symulacje komputerowe w naukach społecznych”, jednak postanowiliśmy przedstawić książkę w szerszej perspektywie z naciskiem na praktyczne zastosowania. Modele układów społecznych są nazywane różnie przez reprezentantów nauk ścisłych: badaniami systemowymi (przez informatyków), układami dynamicznymi (przez matematyków), czy układami złożonymi (przez fizyków) a dotyczą interdyscyplinarnych zagadnień na miarę XXI wieku jak cyfrowa humanistyka, obliczeniowe nauki społeczne (część I książki). Dodatkowo poprzez dostępność i ogrom cyfrowych danych dotyczących działań społecznych, np. w Internecie, powstała przestrzeń do zastosowań innowacyjnych dla nauk społecznych technik. Badania obliczeniowe (*in silico*) pozwalają na przeprowadzanie eksperymentów niemożliwych (i prawdopodobnie nieetycznych) do przeprowadzenia na rzeczywistych społecznościach. Efekty niektórych takich deskrypcyjnych i eksplanacyjnych analiz (praktycznie bez żadnych wzorów) takich jak własne badania: dynamika wiązania się w pary w kontekście zmiany społecznej sieci kontaktów a rozprzestrzenianie się patogenów szpitalnych/przenoszonych drogą płciową; rozprzestrzenianie się opinii (virali); sieci kryminalne/ muzyczne/ literackie/ polityczne / elit; zostały tu przedstawione.

Modele obliczeniowe (socjofizyczne) odtwarzają już to, co zostało zbadane i opisane przez socjologię (część II książki). Metody obliczeniowe są jednak tanie i chyba to jest ich największą zaletą. Wspomniane w tej książce stworzone narzędzia są na tyle uniwersalne, że mogą zostać użyte po dokonaniu niewielkich modyfikacji, w analizie innych różnych zagadnień społecznych. Modelowe metody ekstrakcji informacji (część III) jak analiza hierarchiczna, analiza sieci społecznych, czy przetwarzanie języka naturalnego wydają się być technologiami przyszłości, ze względu na powszechny dostęp do rejestrowanych danych (Big Data), choćby z Internetu i podstawowym problemem staje się umiejętne przetwarzanie w sposób inteligentny i automatyczny tym zasobów. Na koniec (część IV), odnosząc się do jednej z dyrektyw nauk społecznych: „Socjologia jako służba społeczna”, również modele obliczeniowe można traktować jako służbę społeczną, gdyż docelowo wyniki symulacji służą dobru tegoż społeczeństwa, pomimo krytyki nauki w rodzaju filara metodologii polskiej socjologii „O osobliwościach nauk społecznych” S. Ossowskiego.

Książka jest skierowana do osób zainteresowanych tematyką modelowania w socjologii wywodzących się czy to z nauk społecznych, medycznych, inżynierskich czy ścisłych lecz nie wymaga uniwersyteckiego przygotowania matematycznego.

Kiszyniów, 12. 2014

CZĘŚĆ I PROBLEMATYKA I METODYKA OBLICZENIOWYCH NAUK SPOŁECZNYCH

CZĘŚĆ II MODELOWANIE PROCESÓW SPOŁECZNYCH W PRAKTYCE (SOCJOFIZYKA)

CZĘŚĆ III ANALIZA DANYCH SPOŁECZNYCH W UJĘCIU MODELOWO - OBLICZENIOWYM

CZĘŚĆ IV DYSKUSJA NAD SKUTECZNOŚCIĄ I ROZWOJEM METOD OBLICZENIOWYCH

**Zawartość**

1	Co dają metody matematyczno- obliczeniowe socjologom? .....	2-7
2	Skrócona lista najważniejszych zastosowań obliczeniowych w socjologii .....	2-10
2.1	Formacja opinii – model Isinga magnetyzmu .....	2-11
2.2	Podjęcie decyzji w kontekście społecznym – ewolucyjna teoria gier .....	2-11
2.3	Rozprzestrzenianie się idei – dyfuzja .....	2-12
2.4	Sprzedaż oraz czas życia produktów – cykl przyrodniczy .....	2-12
2.5	Dynamika tłumów – zespół oddziałujących cząstek w gazie (dynamika molekularna) .....	2-12
2.6	Ewakuacja obiektów – dynamika przepływów .....	2-13
2.7	Aktorzy społeczni - modele agentowe .....	2-13
2.8	Korupcja – zjawiska anomalne .....	2-13
2.9	Masa krytyczna – masa krytyczna .....	2-13
2.10	Krach i załamania społeczne – trzęsienia ziemi .....	2-14
2.11	Język i komunikacja społeczna – rozkłady energetyczne .....	2-15
2.12	Stratyfikacja społeczna – zasada alokacji energii .....	2-15
2.13	Zmiana norm – przejścia fazowe .....	2-15
2.14	Analiza sieci społecznych – sieci złożone .....	2-15
2.15	Analiza hierarchiczna – klasteryzacja .....	2-16
3	Wprowadzenie do metodologii, filozofii oraz historii nauk obliczeniowych w kontekście społecznym	3-17
3.1	Jednostka a społeczeństwo .....	3-18
3.2	Fizyka społeczna i fizyka ekonomiczna .....	3-19
3.3	Metody obliczeniowe - statystyczne zdobywają naukę .....	3-21
3.4	Wspomagane komputerowo modelowanie i myślenie systemowe .....	3-23
3.5	Układy złożone jako inspiracja do modelowania procesów społecznych .....	3-27
3.6	Redukcjonizm operacyjny a rzetelność badań .....	3-30
3.7	Model a rzeczywistość .....	3-31
3.8	Hierarchiczność .....	3-33
3.9	Pomiar złożoności .....	3-33
4	Nauka, technika i zmiana społeczna w kontekście informatyzacji .....	4-35
4.1	Postęp technologiczny naszej epoki .....	4-35
4.2	Rewolucja internetowa .....	4-36
4.3	Rzeczywistość Wirtualna (VR) .....	4-38
4.4	Społeczne sieci Internetu a nauka .....	4-38
4.5	Społeczeństwo informacyjne i e-styl życia .....	4-39
4.6	Podsumowanie zmian w nauce i technologiach wirtualnych .....	4-40
5	Rodzaje modelowania wykorzystywanego w naukach społecznych obliczeniowych .....	5-41

5.1	Modelowanie deterministyczne .....	5-42
5.2	Modelowanie stochastyczne .....	5-43
5.3	Modelowaniu przy wykorzystaniu sieci społecznych (rzeczywistych i przypadkowych) .....	5-44
5.4	Modelowanie agentowe .....	5-46
6	Analiza sieci społecznych i sieci złożonych .....	6-49
6.1	Podstawowe pojęcia w analizie sieci społecznych .....	6-49
6.2	Sieci społeczne jako sieci złożone .....	6-52
7	Modelowanie epidemiologiczne .....	7-56
7.1	Model Bernoulli'ego – pierwszy opis epidemii (Bernoulli, 1766) .....	7-56
7.2	Projekt System Informatyczny Redukcji Szpitalnych Zakażeń (SIRS-Z) .....	7-56
7.3	Badanie sieci kontaktów: Polymod .....	7-58
7.4	Modele zagrożeń bio-terrorystycznych .....	7-59
7.5	Grypa pandemiczna – zwłaszcza przykład H1N1, tzw. Świńskiej grypy w 2009 .....	7-60
7.6	Rozprzestrzenianie się chorób przenoszonych drogą płciową – wirus HPV .....	7-61
8	Ewolucyjna Teoria gier .....	8-63
8.1	Próba zamodelowania korupcji w piłce nożnej – teoria gier .....	8-63
8.2	Gry ewolucyjne i mniejszościowe .....	8-64
8.3	Teoria gier dla graczy postępujących wedle norm społecznych .....	8-64
9	Modelowanie ruchu .....	9-70
9.1	Modele ewakuacji .....	9-70
9.2	Modele ruchu drogowego .....	9-72
10	Formowanie się opinii .....	10-73
10.1	Model oligopolu .....	10-77
11	Modelowanie wejścia i wyjścia z małżeństwa .....	11-78
11.1	Zachowania kolektywne i wpływ zmiany norm na przykładzie rozwodów .....	11-78
11.2	Dynamiczny model wiązania się w pary .....	11-80
12	Modele rynku fonograficznego .....	12-83
12.1	Szczegółowa analiza zmian na obecnym rynku fonograficznym .....	12-83
12.2	Sformułowanie modelu rynku fonograficznego .....	12-87
12.3	Analiza innych układów społeczno-ekonomicznych, rynków towarowych i pieniężnych .....	12-89
13	13-89	
14	Analiza procesów ewolucyjnych, dyfuzyjnych, korupcji i hierarchiczności .....	14-90
14.1	Modele ewolucji społecznej .....	14-90
14.2	Modele pseudo-epidemiologiczne .....	14-91
14.3	Analiza mediów społecznościowych w czasie krytycznych wydarzeń .....	14-93
14.4	Dynamika emocji .....	14-94
14.5	Korupcja w piłce nożnej .....	14-94

14.6	Zastosowanie technik wizualizacji sieciowych zależności na przykładzie oceny jakości życia osób ze zdiagnozowaną dysfunkcją tarczycy.....	14-96
15	Analiza sieciowa w zastosowaniach .....	15-97
15.1	Sieci seksualne .....	15-97
15.2	Analiza motywów sieciowych – sieci rosnące i malejące .....	15-99
15.3	Sieci geograficzne w układach człowiek - przyroda.....	15-99
15.4	Sieci w literaturze .....	15-101
15.5	Analiza sieciowa organizacji .....	15-103
16	Dynamika wyborcza.....	16-107
16.1	Czy można sfałszować wybory w Polsce?.....	16-107
16.2	Polaryzacja partyjna po katastrofie smoleńskiej - hierarchiczna struktura partii politycznych w Polsce	16-109
17	Techniki wykorzystania języka naturalnego .....	17-115
18	Krytyka metod obliczeniowych.....	18-120
18.1	Brak zdolności predykcyjnej .....	18-120
18.2	Problem pomiaru, przetwarzanie danych .....	18-121
18.3	Nauki ścisłe a socjologia: kontrowersje.....	18-121
18.4	Socjocybernetyka i formułowanie modelu.....	18-122
	Brak głębi poznania danego tematu.....	18-122
18.5	Konsekwencje wprowadzenia modeli w życie.....	18-123
18.6	Metody matematyczne w naukach społecznych.....	18-124
18.7	Metody obliczeniowe i co dalej.....	18-125
19	Spis ilustracji i tabel .....	19-127
20	Bibliografia.....	20-131
21	Nota o autorach.....	21-135

# CZĘŚĆ I

## PROBLEMATYKA I METODYKA OBLICZENIOWYCH NAUK SPOŁECZNYCH

### 1 Co dają metody matematyczno-obliczeniowe socjologom?

W tej książce proponujemy wsparcie nauk społecznych poprzez modele matematyczne (socjofizyka) oraz metody statystyczno-obliczeniowe. W pierwszej kategorii wykorzystywane są środki matematyczne: równania różniczkowe, rachunek prawdopodobieństwa itp. W drugiej metody statystyczne jak regresja, analiza czynnikowa itp., czy informatyczne jak klasyfikacja, predykcja, podejmowanie decyzji itp. Wszystko po to by opisać zjawiska społeczne, będące jednymi z najbardziej skomplikowanych w przyrodzie.

Modele matematyczne są w ostatnich latach coraz częściej i efektywniej wykorzystywane do modelowania zjawisk obserwowanych w układach złożonych. Są to układy powszechne w przyrodzie (np. żywe komórki, sieci neuronów, organizmy i ich populacje). Należą do nich też społeczeństwa (Green & Bossomaier, 2000) i ich dzieła (np. rynki finansowe, Internet). Układy te ze względu na swój skomplikowany opis matematyczny (w istotnym stopniu związany z ich nieliniowością), który uniemożliwia uzyskiwanie dokładnych wyników obliczeń analitycznych, można badać głównie metodami przybliżonymi lub numerycznymi, obserwując ewolucję czasową ich stanów. W efekcie tych badań rozwijających się intensywnie dzięki rozwojowi technik komputerowych, powstało wiele prac opisujących układy, których struktura i dynamika opierały się dotąd na analizie ilościowej. Badania te oparte były na wykorzystaniu różnych modeli jak automaty komórkowe, modele oddziałujących agentów, układy perkolacyjne i sztuczne sieci neuronowe (Kwapień, 2012). Ostatnio pojawiły się prace modelujące układy społeczne, opisujące m.in. rozprzestrzenianie się opinii i formowania gustów, np. muzycznych (Koen, 2001). Opis matematyczny umożliwia lepsze zrozumienie dynamiki procesów w sieciach złożonych (Holme, 2012). Istotnym elementem takich badań jest aspekt poznawczy, gdyż sieci złożone są interesującymi układami, powszechnie występującymi w społeczeństwie. Zjawiska w nich zachodzące mają z jednej strony charakter specyficzny dla danego układu, z drugiej zaś w wielu przypadkach mają charakter uniwersalny, wykazując podobieństwo do zjawisk występujących w innych układach przyrodniczych (np. do przejść fazowych w kryształach ferromagnetycznych).

Rzeczywistość społeczna od dawna jest fascynującym obszarem do opisu matematycznego. Do tej pory jednak większość prób budowy modeli matematycznych dotyczących zagadnień społecznych kończyło się raczej mało praktycznymi uproszczeniami. Od kiedy do aparatu matematycznego dołączyły metody, jak i sposoby rozwiązywania problemów wypracowane przez przyrodników, modele stały się bardziej realistyczne i powoli zaczyna się je wykorzystywać w predykcji, socjotechnice czy cybernetyce społecznej. Najbardziej uderzającą właściwością obiektów badań nauk społecznych – ludzi, organizacji, społeczeństw – jest ich złożoność. Połączenie systemowego ujęcia struktury społecznej (np. u Durkheima) z uwzględnieniem interakcji na poziomie indywidualnych i zbiorowym (np. u Goffmana) wywołało potrzebę opisu modelu przy użyciu pojęć, jak sprzężenie zwrotne, zdefiniowanych na gruncie nauk przyrodniczych. W gwooli wyjaśnienia, nauki ścisłe nie są bynajmniej monolitem i tak dla przykładu dziedzina wiedzy zajmująca się opisanym wyżej przykładem oddziaływań nieliniowych jest nazywana: badaniami systemowymi (przez informatyków), układami dynamicznymi (przez matematyków), czy układami złożonymi (przez fizyków). Socjofizyka powstała na styku socjologii oraz fizyki wspieranej przez matematykę i informatykę, lecz opiera się na przekonaniu, że istnieje analogia między cząstkami, atomami, molekułami, żywymi organizmami,

ludźmi, a nawet całymi ekosystemami i społeczeństwami. Jest zatem przestrzeń naukowa zawierająca wspólne tezy, postulaty oraz dyrektywy badawcze (Sztompka, 1974) i podobne, ścisłe metody badawcze można stosować nie tylko w naukach przyrodniczych, ale również w naukach społecznych. I tak multidyscyplinarne połączenie głębi dociekań charakteryzujących nauki społeczne ze ścisłością warsztatu fizyka stworzyło warunki rzetelności i autentycznej innowacyjności badań naukowych. Jednakże przełom w praktycznym wykorzystaniu nowych dla nauk społecznych technik nastąpił dopiero w XXI w. kiedy ilość i złożoność zdigitalizowanych informacji o czynnościach społecznych wymusiła zmianę paradygmatów w badaniach społecznych. Właśnie w pierwszych latach XXI w. socjofizyka wykrystalizowała się jako nowa dziedzina wiedzy i tak pierwszy raz określenie SocioPhysics zostało wykorzystane 2002 roku na konferencji w Bielefeld. Warto wspomnieć, że polscy naukowcy byli jedną z najliczniej reprezentowanych nacji w procesie budowy socjofizyki i dalej jest to prężnie rozwijająca się w Polsce dyscyplina, o czym świadczy istnienie Sekcji Fizyki w Ekonomii i Naukach Społecznych przy Polskim Towarzystwie Fizycznym<sup>1</sup>, czy Katedry UNESCO Studiów Interdyscyplinarnych we Wrocławiu<sup>2</sup>. Chęć dokładniejszego zrozumienia zjawisk społecznych wyłoniła naturalną potrzebę ujmowania wyników badań szczegółowych w szerszej, interdyscyplinarnej perspektywie. Socjofizyka umożliwia budowanie bardziej całościowych, multidyscyplinarnych płaszczyzn wiedzy społecznej.



Rys. 1 Mały świat powiązań społecznych. Źródło: obraz Idahliyy Stanley

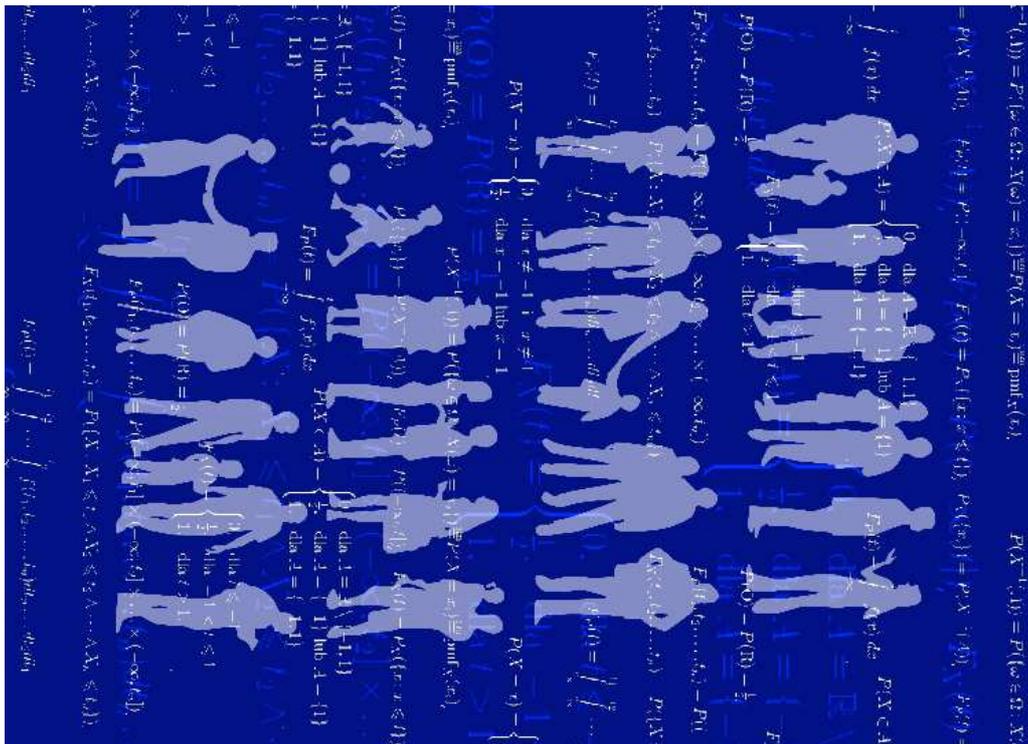
Ta książka ma na celu przedstawienie w pierwszej części kilku ważniejszych aspektów, w których fizycy próbują modelować społeczeństwo z wykorzystaniem metod wypracowanych do analogicznych problemów z fizyki (ze szczególnym uwzględnieniem tych, którymi autor zajmował się osobiście). Na podstawie tych wybranych zagadnień rysuje się główna motywacja stojąca za ingerencją fizyków w socjologię, a mianowicie „wyczuwanie” analogii między tym, czym zajmują się, na co dzień, a tym, co obserwują w życiu społecznym. Poniższe tematy zostały wybrane pod względem wykorzystania aparatu matematycznego opracowanego wcześniej przez fizyków na potrzeby fizyki, nie uwzględniając roli klasycznej matematyki czy

<sup>1</sup> <http://www.ptf.fuw.edu.pl/fens/>

<sup>2</sup> <http://www.kusi.ift.uni.wroc.pl/>

statystyki. Ominięte zostało przy tym całe bogactwo modeli i koncepcji rozwijanych przez cały wiek XX, które zaowocowały powstaniem samodzielnych dziedzin nauki: cybernetyki społecznej (Wiener, 1950), czy młodszej - socjologii analitycznej (Hedström, 2009). Mimo iż w obu dziedzinach wykorzystuje się matematyczny opis zachowań ludzkich, to jednak w większości przypadków brakuje w nich nawet luźnej analogii do praw przyrody, gdyż stawia się w nich na *stricte* inżynierski (cybernetyka), czy socjologiczny (socjologia analityczna) opis procesów. Metody opracowane przez ekonomistów i statystyków jak wszelkiego rodzaju regresje (bardzo chętnie wykorzystywane i rozwijane przez naukowców ze Szkoły Głównej Handlowej w Warszawie w wielu obszarach nauk społecznych), czy informatyczne metody eksploracji danych również nie mieszczą się w zdefiniowanej przez definicji socjofizyki, chociaż fizycy wykorzystują je jako narzędzia pomocnicze. W kolejnej części opisane zostało wykorzystanie analizy sieci społecznych, która wydaje się być najważniejszym wkładem fizyki (dokładnie fizyki układów złożonych) we współczesne badania społeczne. Na koniec zaprezentowane zostało spojrzenie krytyczne na praktyczne wykorzystanie modeli socjofizyki.

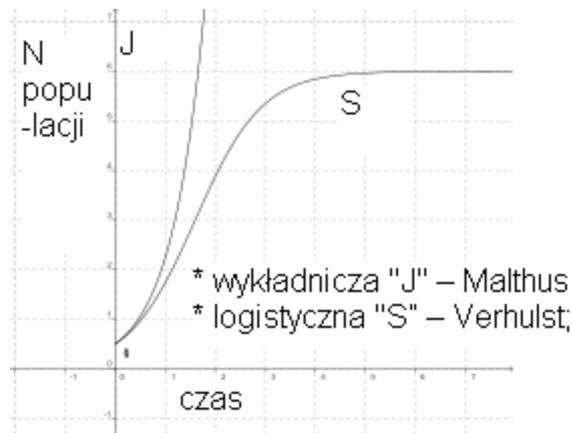
Mimo iż metodologia socjologii jest traktowana jako zasadniczo różna od metodologii fizyki, to obiekt badany – czyli społeczeństwo, jest wspólny. Ten artykuł nie dotyczy sporu metodologicznego (Pabjan, 2004), ani nie porusza kwestii filozofii nauki (Kułakowski, 2012), ale kładzie nacisk na wykorzystaniu metod i analogii fizycznych oraz konsekwencjach takiego wyboru. Zdaniem autora, fizyka nie powinna stawiać problemów nowych w socjologii tylko proponować nowe metody patrzenia na problemy w socjologii znane. Dlatego już na wstępie zostaną wprowadzone główne problemy socjologii w ujęciu socjofizyki oraz zaawansowanych metod eksploracji danych.



Rys. 2 Wzory I ludzie, Grafika A. Jankowskiej

## 2 Skrócona lista najważniejszych zastosowań obliczeniowych w socjologii

Wielość możliwych podejść do modelowania wiąże się cechami tych modeli i np. czy interesuje nas zgodność z rzeczywistością: czy ma być wysoka (precyzyjne i skomplikowane) czy godzimy się niską precyzją (proste modele i łatwe w analizie). Czy w naszym układzie reguły wydają się być deterministyczne lub losowe, jak opiszemy oddziaływania: siły, energie, reguły, a jak zmienne: dyskretne lub ciągłe. Modele można podzielić ponadto na dwa główne, zasadniczo różniące się między sobą rodzaje: makroskopowe i mikroskopowe. W przypadku modeli makroskopowych chcemy odpowiedzieć na pytania „jak i ile”. Nie interesuje nas co się dzieje w mikroskali na poziomie indywidualnych jednostek analizy, tylko jak zachowują się odpowiednie wartości średnie. Mamy tutaj głównie do czynienia z różnego rodzaju równaniami strukturalnymi. Taki opis jest podobny do makroskopowego opisu układów złożonych, tak jak ma to miejsce w termodynamice, gdzie operujemy: temperaturą, ciśnieniem, objętością etc. Dzięki takiemu podejściu można odpowiedzieć na wiele ilościowych pytań, można też generować mniej lub bardziej trafne przewidywania. Przykładem jest model Malthusa przyrostu populacji, gdzie liczebność ludności świata przyrastała do połowy XX wieku właściwie wykładniczo (krzywa „J”), czy model Verhulsta, gdzie pod połowę XX nastąpiło spowolnienie, co nazywane jest krzywą logistyczną („S”).



Rys. 3 Modele makroskopowe liczebności społeczności w matematycznym opisie dynamiki populacji

Zasadniczą wadą takich modeli jest brak odpowiedzi na pytanie o przyczyny występujących zjawisk („dlaczego?”). Modele mikroskopowe tworzy się w celu próby odpowiedzenia na pytanie „dlaczego?”. W przypadku nauk ekonomicznych i społecznych wyróżnia się modele typu:

- mikrosymulacje (ang. Microsimulation), gdzie zmiany stanów obiektów określone są pewnymi regułami deterministycznymi lub stochastycznymi;
- modelowanie agentowe (ang. Agent Based Models), w których badany układ jest zbiorem „agentów” oddziałujących wzajemnie, według pewnych, zależnych od modelu, reguł.

Agent, jako podstawowy element układu, posiada pewne (opisane liczbowo) cechy i zwykle oddziałuje z innymi agentami lub czynnikami zewnętrznymi. Zarówno cechy pojedynczego agenta, jak i reguły oddziaływania są zależne od konkretnego modelu. Można powiedzieć, że obiekt zwany agentem, jest pewnym uogólnieniem koncepcji cząstki, ciała itd. znanymi z fizyki.

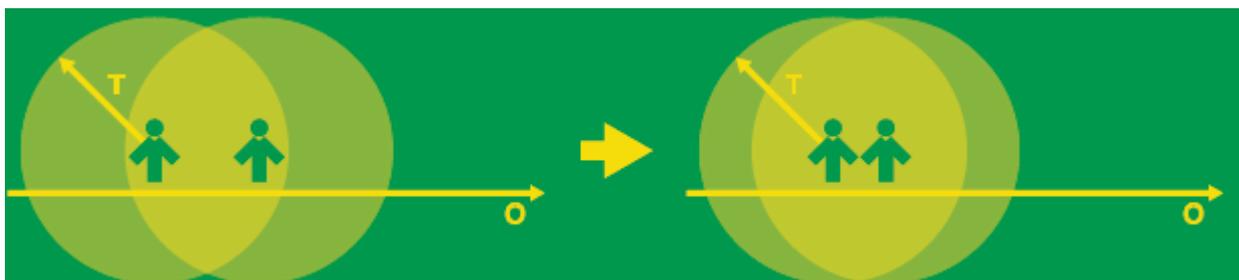
Na konkretnych przykładach modele przedstawiają się następująco. Oto właśnie krótka lista ciekawszych zagadnień społecznych z analogami fizyczno-przyrodniczymi ze szczególnym uwzględnieniem przykładów dotyczących Polski.



Rys. 4 Przykłady zastosowań metod obliczeniowych układów złożonych opracowane przez Agatę Fronczak: Źródło [www.if.pw.edu.pl/~agatka](http://www.if.pw.edu.pl/~agatka)

## 2.1 Formacja opinii – model Isinga magnetyzmu

Socjofizyka choć początkowo traktowana bardzo nieufnie przez fizyków, dziś jest szeroko akceptowaną dziedziną nauki. Jednym z najpopularniejszych zagadnień badawczych rozważanych w ramach tej dziedziny jest modelowanie dynamiki opinii. W modelach formowania się opinii, bazujących na teorii wpływu społecznego, opinia jednostki jest opisana przez spin z modelu Isinga (spin jest cechą cząstki i może być skierowany w górę bądź w dół). Przykładowo analogią spinu w dół może być popieranie partii lewicowej, a w górę prawicowej. Najistotniejszą regułą tego modelu jest podejmowanie decyzji przez większość głosów – jednostka przyjmuje opinię dominującą wśród swoich sąsiadów. Twórcami jednego z pierwszych tego typu modelu, nazwanego na jej cześć (od jej nazwiska) „modelem Sznajdów”, jest rodzina fizyków z Wrocławia (uznawana za prekursorów socjofizyki na równi z francuskimi i niemieckimi naukowcami). Aktorzy tego wirtualnego świata zostali pieszczotliwie nazwani „Spinsonami”. Wykorzystane w modelu fizyczne mechanizmy odpowiadają znanym z socjologii fenomenom typu społeczny dowód słuszności, czy konformizm. Emergencja (czyli efekt zespołowy większy niż suma efektów elementów składowych) pojawia się tu jako wynik oddziaływań pomiędzy osobnikami. W ramach przeprowadzonych symulacji komputerowych uzyskano wiele zaskakujących wyników, między innymi taki, że opinia publiczna może spontanicznie zmienić się drastycznie bez żadnej konkretnej przyczyny (Sznajd-Weron, 2001).



Rys. 5 Wybór opinii pod wpływem interakcji z innymi ludźmi

## 2.2 Podejmowanie decyzji w kontekście społecznym – ewolucyjna teoria gier

Opracowana przez słynnego matematyka, Nasha, teoria gier została uogólniona do tego stopnia, aby opisywać schemat podejmowania decyzji wedle ekonomicznych przesłanek (Axelrod, 1981). Obecnie zamiast opisywać *homo oeconomicus*, próbuje zmierzyć się z *homo sociologicus* w grze zwanej dylematem więźnia. W tym zagadnieniu każdy z dwóch graczy może wybrać, czy będzie współpracował czy zdradzał. W zależności od przyjętej przez obu graczy strategii pojedyncza gra kończy się pewnym ustalonym rozdziałem wypłat/kar. I tak prosty dynamiczny model, w którym gracze - jednostki mogą ze sobą współpracować lub się zdradzać, odtwarza wiele zbiorowych zachowań. Model został rozwinięty przez

uwzględnienie działania normy społecznej (przyzwalającej, czy zabraniającej zdradę) i dynamicznych procesów adaptacyjnych. W zależności od warunków początkowych i parametrów modelu można wygenerować sztuczne społeczeństwa, które będą wyłącznie współpracować lub zdradzać. Dla pewnych obszarów niestabilnych możliwe są jednak przeskoki między dwoma stabilnymi obszarami jednorodnymi (gdzie wszyscy jedynie zdradzają bądź współpracują). Takim przykładem jest zachowanie na drodze, np. Austriaków, którzy ze sobą cały czas współpracują (panuje na tyle powszechne zaufanie, że kierowcy czekający w rzędzie na czerwonym świetle ruszają jednocześnie w momencie zapalenia się światła zielonego). Autor sugeruje, co mogłoby się stać, gdyby zbyt wielu polskich, czy włoskich kierowców przyjechało do Austrii ze swoimi nawykami (Jarynowski, Kułakowski, Gawroński, 2012).

### 2.3 Rozprzestrzenianie się idei – dyfuzja

Mechanizm zaraźliwości społecznej ma miejsce wówczas, gdy osoba zmienia postawę lub zachowanie pod wpływem swojego znajomego. Wspólnie z homofilią (powinowactwo do jednostek podobnych do siebie), zaraźliwość społeczna jest uważana za główny motor dynamiki grup społecznych. Fizyczny proces dyfuzyjny (rozprzestrzenianie się cząsteczek w jakimś ośrodku) można porównać do epidemiologii chorób zakaźnych. Postawy zaraźliwe mogą w ten sposób dynamicznie się rozprzestrzeniać zarażając kolejne osoby. Studium ludzkich zachowań seksualnych i ich wpływ na rozprzestrzenianie się chorób przenoszonych drogą płciową, to społeczno-epidemiologiczne zagadnienie, którym autor obecnie się zajmuje.



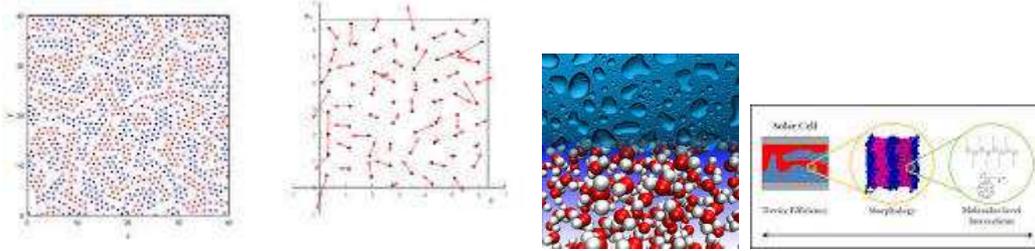
Rys. 6 Model rozprzestrzeniania się opinii ze źródłami rozproszonymi skoncentrowanymi na Ukrainie (stan przed A i po B oraz ścieżki rozchodzenia się opinii C). Źródło: (Dybiec, 2014)

### 2.4 Sprzedaż oraz czas życia produktów – cykl przyrodniczy

Jeżeli między elementami występują zależności, jak w łańcuchu pokarmowym, to zdarzają się sytuacje, kiedy jeden lub więcej gatunków wymiera, przy czym ekosystem jako taki pozostaje. Podobnie dzieje się na rynku, gdzie klienci mają wybór różnych produktów, a innowacje odpowiadają wtedy zmianom genotypów. Na podstawie tych zależności za pomocą modeli opisujących m. in. wyginiecie dinozaurów, można próbować przewidywać, jak długo dany produkt będzie cieszył się aprobatą klientów.

### 2.5 Dynamika tłumów – zespół oddziałujących cząstek w gazie (dynamika molekularna)

Cząsteczki fizyczne w gazie oddziałują ze sobą poprzez przyciąganie, odpychanie, zderzenia itp. Analogiczne oddziaływania zostały nałożone w stosunku do ludzi, przy czym ludziom przypisuje się inne funkcje. W takich modelach minimalizuje się energię układu zazwyczaj za pomocą symulacji komputerowych. Wynikiem czego są różnego rodzaju efekty zbiorowe jak formowanie się korków, które wychodzą wprost z działania podstawowych praw fizyki. Dzięki tym modelom można m.in. opisać falę meksykańską.



Rys. 7 Standardowe techniki symulacji metodami dynamiki molekularnej. Jeżeli zamiast cząstek wstawimy ludzi, to możemy mówić o dynamice społecznej. Źródło: udel.edu

## 2.6 Ewakuacja obiektów – dynamika przepływów

Tak jak ruch cząstek gazowych charakteryzuje względna chaotyczność, tak pewne celowe zachowania (np. chęć dostania się, w możliwie krótkim czasie z punktu A do punktu B) lepiej modeluje się poprzez prawa hydrodynamiczne, czyli jak ciecze. Obecnie każda duża inwestycja budowlana już na poziomie projektu jest konsultowana pod względem szybkości ewakuacji ocenianej właśnie za pomocą modeli (Kosiński, Grabowski, 2010). W Polsce pionierskie analizy z tej dziedziny zostały zrobione na istniejącym już stadionie miejskim w Krakowie oraz budowanym nowym stadionie miejskim w Tychach.

## 2.7 Aktorzy społeczni - modele agentowe

Interakcjom między jednostkami (np. ludźmi) można przypisać pewne reguły i zaimplementować algorytmy odpowiadające dynamice procesu. Oddziaływania mogą mieć charakter deterministyczny (agenci zachowują się za każdym razem tak samo przy tym samym zestawie warunków początkowych) jak również statystyczny.

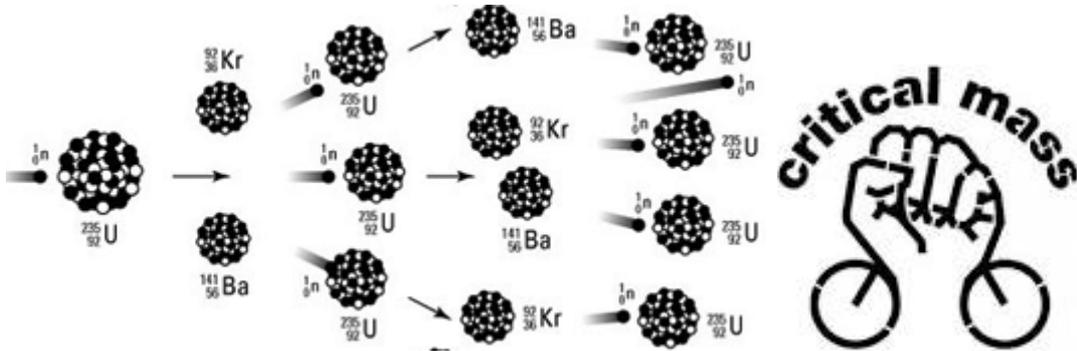
## 2.8 Korupcja – zjawiska anomalne

Poszukiwanie anomalnych zachowań jest bardzo częstym zagadnieniem badawczym w wielu dziedzinach nauki. W medycynie jednym z podstawowych pojęć jest patologia (w ogólniejszym znaczeniu nieprawidłowość), która jako dziedzina zajmuje się rozpoznawaniem przyczyn, mechanizmów i skutków chorób. Podobnie fizycy, opisując otaczającą nas rzeczywistość odkrywają reguły rządzące światem, a także muszą ustosunkować się do pewnych nietypowych zachowań. Dla przykładu ostatnio bardzo popularnym zagadnieniem są austetyki: materiały, które rozciągane powiększają, a ściskane pomniejszą swoje rozmiary, to znaczy, że mają one ujemne współczynniki rozszerzalności.

## 2.9 Masa krytyczna – masa krytyczna<sup>3</sup>

W fizyce często zdarza się, że procesy zmieniają swoją dynamikę w zależności od ilości elementów biorących w nich udział. Co ciekawe, termin *masa krytyczna* został zaadaptowany do socjologii w identycznym brzmieniu jak termin fizyczny. W wielu układach społecznych zaobserwowano takie właściwości. Przykładowo, przeprowadzono badania w Wielkiej Brytanii sprawdzające, jak zależy efektywność ośrodka naukowego (liczona w ilości punktowanych publikacji) od jego wielkości. Okazuje się, że dla małych ośrodków zależność jest liniowa, a od pewnej „masy krytycznej” następuje dramatyczny wzrost efektywności.

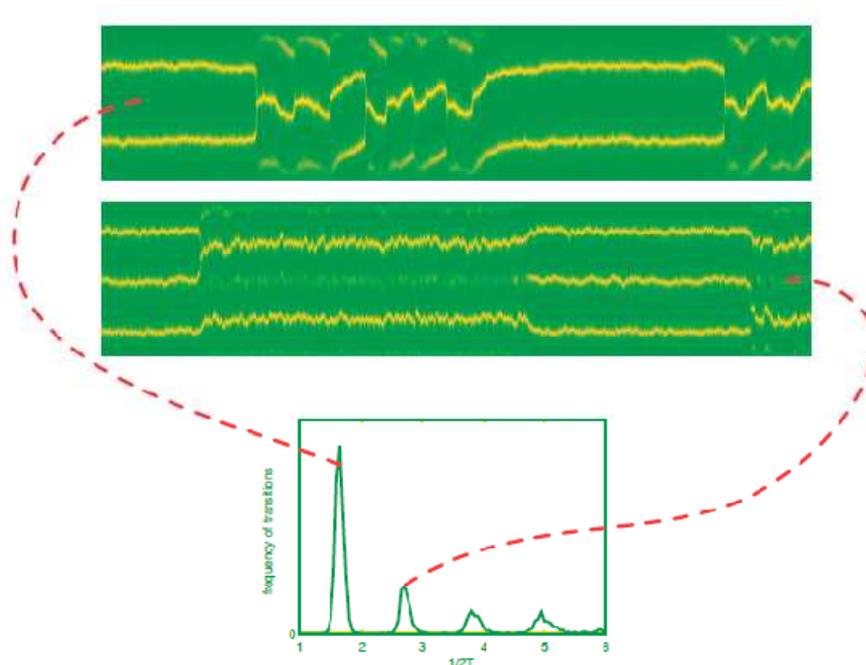
<sup>3</sup> Pojęcie masa krytyczna jest zazwyczaj kojarzone z ruchem rowerowym polegającym na organizowaniu spotkań maksymalnie licznej grupy rowerzystów i manifestacji ich obecności w krajobrazie miast.



Rys. 8 Literalne znaczenie masy krytycznej (lewy). Źródło: <http://longorshortcapital.com/critical-mass-supplier.htm> , ugruntowane społecznie znaczenie (prawy). Źródło: bikeblognyc.com

## 2.10 Krach i załamania społeczne – trzęsienia ziemi

Można zauważyć, że katastrofy, jakie mają miejsca w różnych dyscyplinach, wykazują podobieństwo zarówno w swoim przebiegu, jak i w sposobie rozwijania się. Istnieje ogólna metodologia oparta na pojęciach i technikach statystycznej i nieliniowej fizyki. Takie ujęcie zostało wykorzystane do problemów tak różnych jak: rewolucje społeczne, krachy giełdowe czy trzęsienia ziemi, od których tak naprawdę wszystko się zaczęło. Krytyczność w takich układach pojawia się poniżej/powyżej pewnej krytycznej wartości jakiejś obserwowanej zmiennej.



Rys. 9 Model przejścia układu z jednego stanu do drugiego w sposób gwałtowny

Co wspólnego mogą mieć uskoki sejsmiczne i ludzie grający na giełdzie? Ostatnie badania sugerują, że mogą one zostać opisane w dużym stopniu w tym samym podstawowym fizycznym terminie: jako samoorganizujące się układy złożone. I w obu przypadkach mamy do czynienia z ekstremalnymi zachowaniami: trzęsieniami i krachami (Jarynowski, 2007). Innym przykładem są układy społeczno-ekologiczne, w których ingerencja człowieka w kruchą równowagę naturalną doprowadza do katastrof ekologicznych. W przypadku przetrzebienia łowisk, w wyniku nadmiernych połowów, rybacy często tracą zupełnie swoje miejsce utrzymania. Dlatego wykorzystując narzędzia matematyczne, przy posiadanej wiedzy o warunkach naturalnych, można opracować limity połowów tak aby nie przekroczyć tej bariery, a jednocześnie korzystając z dostępnych złóż naturalnych. Tego typu techniki nazywają się z angielskiego Resilience & Sustainability.

## 2.11 Język i komunikacja społeczna – rozkłady energetyczne

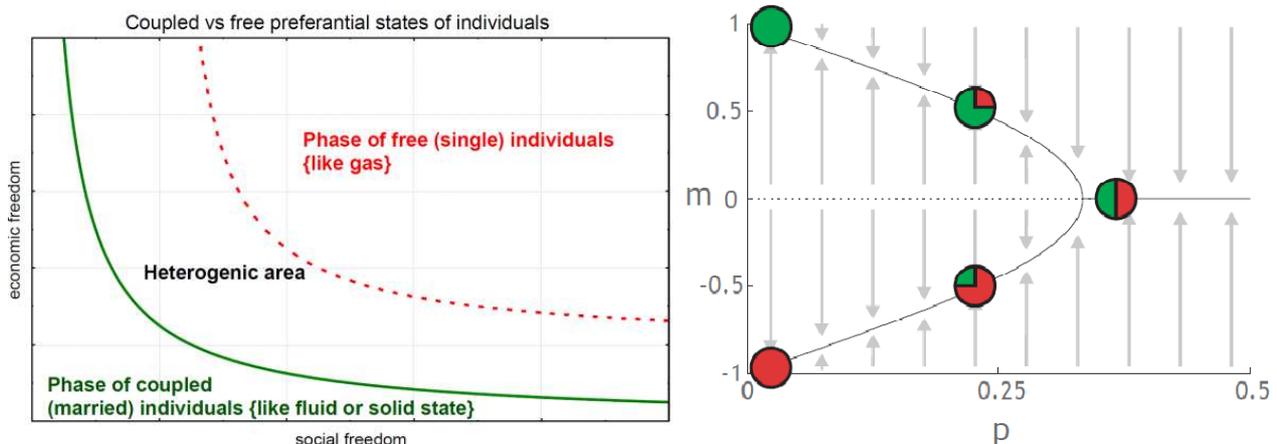
W związku z rozwojem metod komputerowych w przetwarzaniu informacji pojawiła się możliwość wykorzystania procedur automatycznych w analizie literackiej (Jarynowski, Boland, 2013). Proste czynności zliczeniowe, które dało się zautomatyzować, zostały adaptowane do warsztatu badacza literatury. Nie jest to jednak ich jedyne zastosowanie, gdyż badania ilościowe w literaturze zagościły już na dobre w literaturoznawstwie czy językoznawstwie. Najlepszym przykładem takiego podejścia jest analiza Zipfa (Kulig, Kwapien, et al., 2012), opisująca rozkład częstości występowania słów w tekście. Okazuje się, że na podstawie obserwacji częstości występowania słów oraz ich asocjacji można odróżnić pisarzy czy gatunki literackie.

## 2.12 Stratyfikacja społeczna – zasada alokacji energii

W przyrodzie występuje bardzo często zasada Pareto: 20% populacji posiada 80% środków. Wynika to z właściwości bardzo ważnych dla fizyki tak zwanych „rozkładów długoogonowych”. Okazuje się, że społeczności również w podobny sposób alokują swoje dobra; tak też budowane są schematy komunikacyjne. Wiąże się to też z dynamiką wzrostu związaną z efektem św. Mateusza (Merton 1968) „Każdemu bowiem, kto ma, będzie dodane, tak że nadmiar mieć będzie. Temu zaś, kto nie ma, zabiorą nawet to, co ma” [Mt 25:14-30]. Innym ciekawym aspektem stratyfikacji jest analiza powstawania i działania układów hierarchicznych (Czaplicka, et al., 2013) z akcentem na efektywność struktur organizacyjnych. W związku z tym można przewidzieć jakie układy hierarchiczne będą optymalne, np. z punktu widzenia przepływu informacji.

## 2.13 Zmiana norm – przejścia fazowe

Normy obyczajowe podlegają dynamicznym zmianom, czego przykładem może być społeczna akceptacja rozwodów w powojennej Europie. Zakładając, że ludzie w zależności od posiadanej wolności ekonomicznej i społecznej będą preferowali stan związany (małżeństwa), bądź wolny (single). Porównywanym zjawiskiem jest przemiana fazowa ze stanu ciekłego do gazowego pod wpływem zmiany ciśnienia i temperatury, do dynamiki rozwodów i małżeństw pod presją ekonomiczną czy społeczną. Stanowi to teoretyczną bazę analizy. Głównym problemem operacjonalizacji jest mierzalność presji. Presja ekonomiczna lub wolność ekonomiczna mogą zostać wyznaczone na podstawie dochodów gospodarstw domowych (po uwzględnieniu zgromadzonego kapitału oraz zadłużenia). Trudność stanowi presja/wolność społeczna – w tym wypadku, zamiast presji postanowiono zmierzyć zmianę presji.

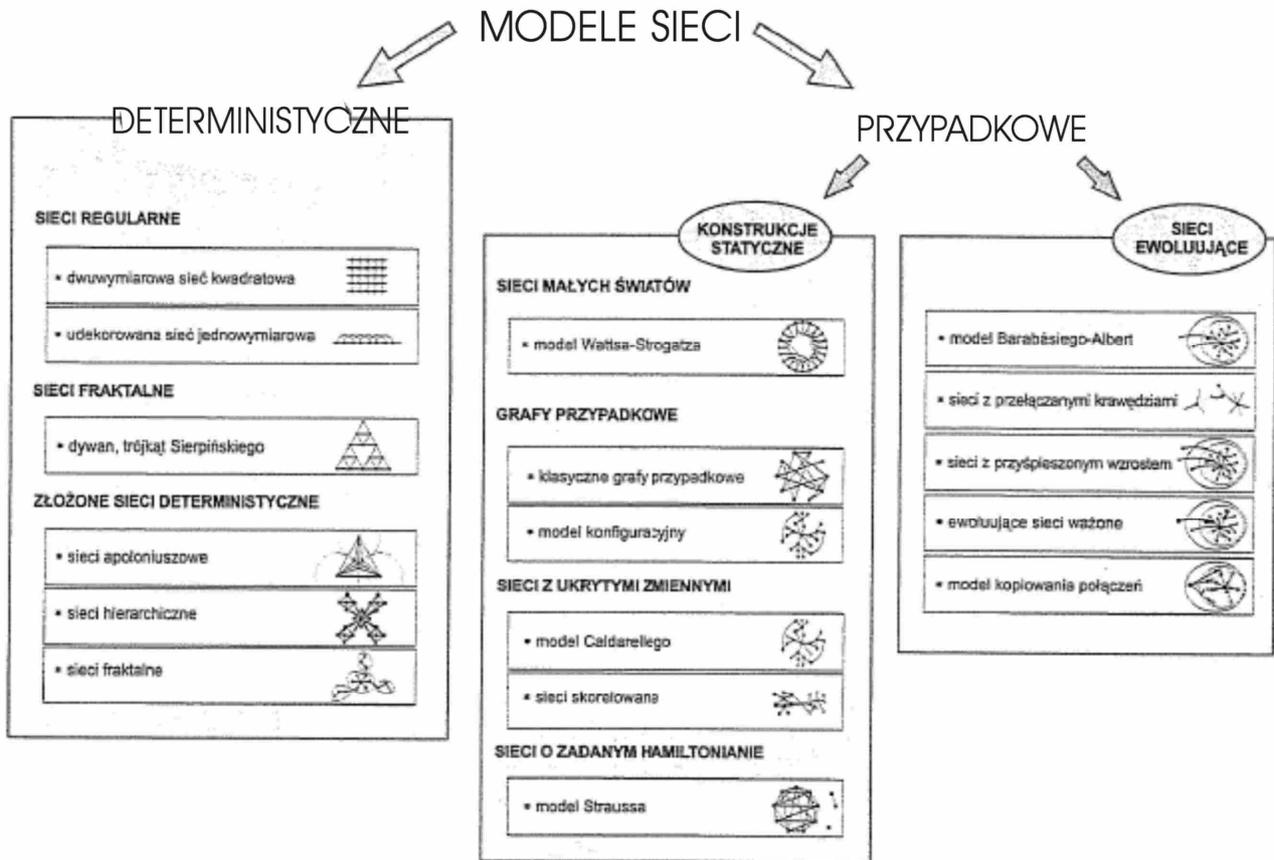


Rys. 10 Schemat analogii przejścia fazowego między społeczeństwem preferującym stan wolny a stanem związanym

## 2.14 Analiza sieci społecznych – sieci złożone

Sieci społeczne daje się w wielu wypadkach odpowiednio zaklasyfikować do sieci złożonych, co pozwala na wykorzystanie wyników badań teoretycznych w praktyce (o czym dokładniej w rozdziałach poświęconym tym zagadnieniom). Charakterystyka kontaktów w postaci czasowej i sieciowej struktury (kiedy i z kim/z

czym) wiadomo że ma wpływ na wiele aspektów życia społecznego. Np. W kontekście chorób przenoszonych drogą płciową analiza sieciowa pozwala stwierdzić, że społeczeństwa zachodnie są bardziej podatne na zarażenia Chlamydia, ze względu na dużą liczbę kontaktów przygodnych, natomiast społeczeństwa z dopuszczalnymi poligamicznymi związkami, jak na przykład społeczeństwo arabskie, są bardziej narażone na zarażenie wirusem HIV, ze względu na intensywność współżycia seksualnego z wieloma partnerami.



Rys. 11 Klasyfikacja różnych modeli sieci. Źródło (Fronczak, Fronczak, 2009)

## 2.15 Analiza hierarchiczna – klasteryzacja

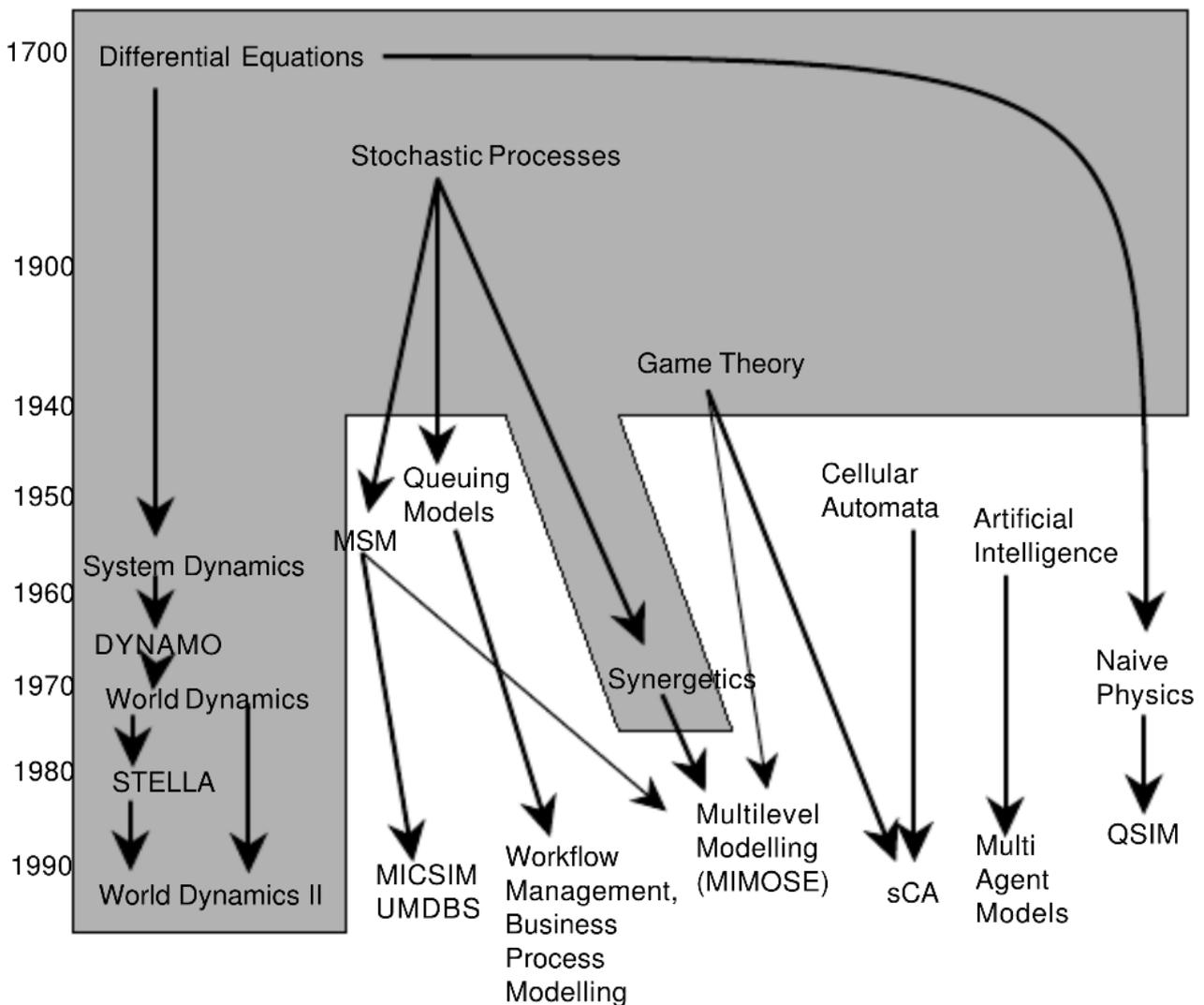
Hierarchiczne przedstawienie zależności między podmiotami jest jedną ze statystycznych metod dotyczących analizy skupień, czyli wyłuskiwania istotnych informacji. Wyróżnia się 2 główne strategii dotyczące grupowania informacji.

- Klasyfikowanie, które polega na wyznaczaniu różnic/podobieństw na podstawie znanych reguł eksperckich.
- Klastrowanie, gdzie aprioryczna wiedza o potencjalnych zależnościach w systemie jest zerowa.

Przedstawienie hierarchiczne (któremu poświęcony został podrozdział w dalszej części tej książki) jest właśnie typem klasteryzacji dla której powstało wiele różnych algorytmów wykonujących to zadanie. Właśnie dzięki jednej z takich metod, udało się pokazać medialne oddalenie na osi PiS – PO po katastrofie Smoleńskiej.

### 3 Wprowadzenie do metodologii, filozofii oraz historii nauk obliczeniowych w kontekście społecznym

Patrząc na ten problem z punktu widzenia filozofii nauki, można dostrzec, że historycznie fizyka jako stosowana matematyka włożyła najtrwalszy wkład w rozwój naszej cywilizacji praktycznie od czasów starożytnej Grecji, począwszy od Talesa z Miletu czy Arystotelesa, który właśnie fizykę uznał za najważniejszą z dziedzin, którymi się zajmował. Pozytywiści XIX wieku, a przede wszystkim neopozytywiści, umieścili fizykę na pierwszym miejscu listy najbardziej naukowych dziedzin jako dziedzinę najlepiej osadzoną w doświadczeniu i logicznym wnioskowaniu, co cenili m.in. empiryści brytyjscy z Ianem Hume na czele. Fizyka znalazła się przed chemią, biologią i socjologią, która w tamtych czasach praktycznie była psychologią społeczną i balansowała na krawędzi naukowości. To właśnie neopozytywiści bowiem, bazując na osiągnięciach logiki i bacznie obserwując wielkie osiągnięcia naukowe swoich czasów – ze szczególną teorią względności i mechaniką kwantową na czele, które potwierdziły się doświadczalnie, ostatecznie ustalili kryteria naukowości. Otóż teoria, aby być naukowa, musi być falsyfikowalna i weryfikowalna, czyli zgadzać się z doświadczeniem. Fizyka oczywiście spełnia oba te warunki. Co więcej, istniejące w niej alternatywne teorie, jak np. mechanika fal i cząstek, jako równouprawnione doskonale uzupełniają się wzajemnie. Dlatego właśnie neopozytywiści jako pierwsi postulowali wyrażenie innych nauk w języku fizyki.



Rys. 12 Schemat stosowalności i ewolucji metod obliczeniowych w naukach społecznych. Źródło: (Gilbert, Troitzsch, 2005)

Socjofizyka wydaje się być całkiem nową dziedziną nauki, a przynajmniej tak jest przedstawiana przez niektórych. Nie ma w tym nic dziwnego, obecnie mamy do dyspozycji symulacje komputerowe oraz aparat matematyczny, wywodzące się wprost z fizyki statystycznej, które są używane do modelowania zachowań społecznych. Warto zadać sobie pytanie czy to symulacje komputerowe układów fizycznych zainspirowały naukowców do komputerowego odtwarzania społeczeństw? A może jednak jajko pojawiło się wcześniej niż kura? Nie wszyscy jednak zgadzają się z jego wersją i za początek ery symulacji agentowych uznają prace Schellinga (Schelling, 1971), a później Galama (Galam et al., 1982).

### 3.1 Jednostka a społeczeństwo

Najprawdopodobniej cała historia zaczęła się w epoce oświecenia. Istotny wkład w ideę mechanistycznego ujęcia miał filozof przyrody Thomas Hobbes (1588–1679). Postrzegał on człowieka, a także społeczeństwo jako mechanistyczny twór balansujący między wolnością/brakiem bezpieczeństwa a podporządkowaniem/bezpieczeństwem. Uważał, że bez ustanowienia państwa i prawa każdy będzie walczył z każdym, egoistycznie dążąc do spełnienia swoich osobistych celów, zagrażając przy tym bezpieczeństwu innych. Społeczeństwa składają się z jednostek, a prawa odnoszą się do nas jako konkretnych jednostek i regulują pośrednio funkcjonowanie społeczeństwa. Hobbes z Johnem Locke pytał się jak można zorganizować życie społeczne o ład społeczny. Stworzył w ten sposób załączek koncepcji, która parę wieków później „wynaleziona” przez amerykańskiego matematyka i ekonomistę Johna Nasha została nazwana „równowagą Nasha”. Hobbes uważał, że jedyną drogą do dobrego, stabilnego i bezpiecznego społeczeństwa jest wprowadzenie monarchii absolutnej. Przekonaniu temu dał wyraz w swoim najbardziej znanym dziele z 1651 roku zatytułowanym: „Lewiatan”. Był twórcą koncepcji teoretycznego aparatu opisu społeczeństwa, który bez większych trudności można by było obecnie zastosować jako model sieciowy. Czasowa odległość między ową koncepcją a jej wdrożeniem, oraz nowatorstwo pomysłu, stawia go wysoko w hierarchii myślicieli. Następną ważną postacią w tej historii był Sir William Petty (1623-1687). Ten wychowanek Hobbesa związany był zarówno z Racjonalistami jak i Empirykami. Uważał, że nauka o społeczeństwie może być dokładna tylko wówczas gdy będzie ujmowała zjawiska ilościowo. Tę dyscyplinę badawczą nazwał arytmetyką polityczną (ang. political arithmetic). Petty był również fizykiem, anatomem, muzykiem, wynalazcą oraz członkiem parlamentu. Tak wszechstronne zainteresowania pozwoliły mu na szerokie i nowatorskie spojrzenie na problematykę reguł rządzących społeczeństwem. Prace Petty'ego zainspirowały Johna Graunta (1620 – 1674) do stworzenia statystyk urodzin i śmierci. Uważał, że efektywne zarządzanie państwem jest niemożliwe bez takich danych. Graunt zajmował się również epidemiologią. Petty badał dane demograficzne posługując się wartościami średnimi. Taka metoda, nazwana później ekonomią, zwała się wówczas „political economy” - termin „economy” został wprowadzony dopiero w XIX wieku. Owe studia prowadziły go często do nieco bezdusznych i nieludzkich wniosków, które Jonatan Swift nieco satyrycznie opisywał, „racjonalnie” sugerując, że irlandzcy biedacy utrzymują się przy życiu dzięki temu, że zjadają swoje dzieci. Dane dotyczące narodzin i zgonów były w owym czasie bardzo modnym tematem badań wśród naukowców z różnych dziedzin. Jednym z takich naukowców był Edmund Halley. Uważano wówczas, że tak jak z obserwacji astronomicznych wyłaniają się pewne ściśle prawa, tak wyłonią się one z danych dotyczących społeczeństwa. W tym miejscu pojawia się nawet wątek wrocławski. Otóż Halley, korzystając z tablic narodzin i zgonów opracowanych przez wrocławskiego pastora Caspara Neumanna z burmistrzem Justellem, opracował algorytm obliczania składek emerytalnych. Wrocław, zdaniem Halleya, był miastem wzorcowym, nieposiadającym portu i leżącym we względnie spokojnym miejscu kontynentu. Innymi słowy uważał je za zupełnie typowe i niczym nie wyróżniające się miasto, doskonale nadające się jako model przeciętności. Pisał o modelowym mieście:

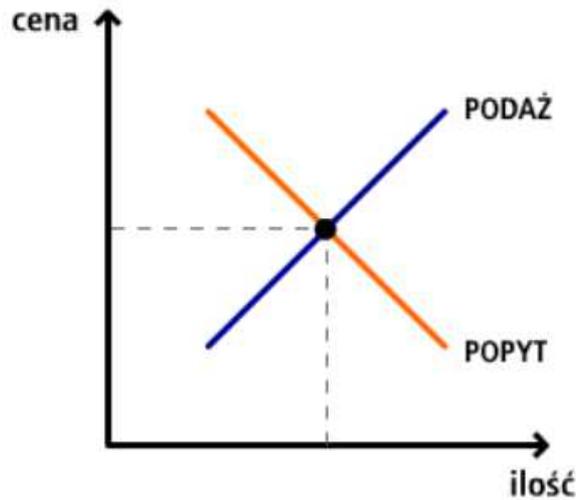
„This *Defect* seems in a great measure to be satisfied by the late curious Tables of the Bills of *Mortality* at the City of *Breslaw*, lately communicated to this Honorable Society by Mr. *Justell*, wherein both the *Ages* and *Sexes* of all that die are monthly delivered, and compared with the number of the *Births*, for Five Years last past 1687, 88, 89, 90, 91, seeming to be done with all the Exactness and Sincerity possible. (...) This City of *Breslaw* is the Capital City of the Province of *Silesia*; or, as the *Germans* call it, *Schlesia*, and is situated on the Western Bank of the River *Oder*, anciently

called *Viadrus*; near the Confines of *Germany* and *Poland* (...) It is very far from the Sea, and as much a *Mediterranean* Place as can be desired, whence the confluence of Strangers is but small, and the Manufacture of Linnen employs chiefly the poor People of the place, as well as of the Country round about; whence comes that sort of Linnen we usually call your Schlesie *Linnen*; which is the chief, if not the only Merchandize of the place. For these Reasons the People of this City seem most proper for a *Standard*; and the rather, for that the *Births* do, a smaller matter, exceed the *Funerals*. The only thing wanting is the Number of the whole People, which in some measure I have endeavored to supply by comparison of the *Mortality* of the People of all Ages, which I shall from the said Bill traces out with all the Accuracy possible."

Dane Halley o Wrocławiu pozwoliły na stworzenie pierwszego modelu epidemiologicznego w historii przez Bernoulliego (Bernoulli, 1766). Bardzo szczegółowe rejestry – tabele śmiertelności, pozwoliły na oszacowanie oczekiwanej długości życia dla różnych grup wiekowych. Sam model dotyczył ustalenia długości życia bez Ospy (o czym więcej w podrozdziale Modelowania epidemiologicznego o Bernoullim). Niektóre towarzystwa ubezpieczeniowe korzystały ze stworzonego przez niego algorytmu nawet do końca XIX wieku. Wielu oświeceniowych myślicieli uważało naukę za klucz do osiągnięcia powszechnej wolności, równości i szczęśliwości. Dziś między innymi dzięki socjologii wiemy, że niestety wcale tak nie musi być, a historia pokazała że nauka to miecz obosieczny i trzeba jednak zachować ostrożność aby się nie skaleczyć. Właśnie też w oświeceniu różne teorie socjologiczne zaczynały się kształtować, pojawiła się idea nauki o społeczeństwie. Po okresie starożytnym – kiedy wyodrębniony został porządku ludzkiego z porządku przyrody, średniowieczu zdominowanym prawem naturalnym, barokiem kiedy odróżniano już państwo od społeczeństwa przyszła pora w XVIII w. na uczynienie społeczeństwa przedmiotem systematycznej refleksji. Wraz z romantyzmem powstała koncepcja umowy społecznej, a indywidualizm i samorealizacja jednostek jako przyczynę do idealnego społeczeństwa znalazły się w głównym nurcie rozważań, np. J.J. Rousseau. Pojawia się w tych koncepcjach emergencja socjologiczna. Społeczeństwo to coś więcej niż suma jednostek. Które są w stanie utworzyć warunki życia, które nie wypływają z natury jednostek a są skutkiem interakcji między nimi.

### 3.2 Fizyka społeczna i fizyka ekonomiczna

W kontekście powyższych rozważań, a także tematyki książki, warto zauważyć, że ścisłe związki pomiędzy fizyką i matematyką a teoriami ekonomicznymi oraz socjologicznymi i sięgają ponad dwustu lat. Można tu wskazać A. Smitha, ojca współczesnej ekonomii, który inspirował się w swoich badaniach metodologią fizyki i astronomii, a w szczególności podstawowym dla tych nauk tworzeniem teorii na bazie badania prawidłowości empirycznych. W jego modelach ekonomicznych egoizm jednostek jest motorem rozwoju ekonomicznego. Jest to przykład modelu agentowego w którym efekt makroskopowy (gospodarka) nie jest reprezentowany przez działanie składowych tego modelu (ludzi i przedsiębiorstw) oraz reguł mikroskopowych (zasada maksymalizacji zysku). Tak zwana „niewidzialna ręka rynku” to również efekt niekoniecznie widoczny z obserwacji pojedynczych przedsiębiorstw, a właśnie interakcji między nimi i konsumentami. Wielu XIX-wiecznych ekonomistów i społeczników, takich jak L. Walras i V. Pareto, starało się przenieść idee i formalizm matematyczny na grunt ekonomii i, co doprowadziło do powstania wielu ważnych dla ekonomii pojęć i teorii, takich jak np. pojęcie równowagi rynkowej, inspirowane fizyką statystyczną. Z drugiej strony, w 1900 L. Bachelier starał się modelować ruchy cen akcji, wyrażając je przez proces stochastyczny, z dzisiejszego punktu widzenia równoważny ruchowi Browna, co wyprzedziło nawet dokonania Smoluchowskiego i Einsteina.



Rys. 13 Graficzne ujęcie najważniejszego prawa ekonomiki zachowań ludzkich, czyli prawa popytu i podaży

Termin fizyka społeczna (ang. social physics) został po raz pierwszy wykorzystany w książce Augusta Comte „System of positive Philosophy” napisanej w latach 1830-1842. Jego zdaniem fizyka społeczna miała uzupełnić i zarazem zamknąć naukowy opis świata. Zaproponowany opis społeczeństwa wywodził się bezpośrednio z opisu przyrodniczego. Socjologia dzieliła się na statykę społeczną – budowa i funkcjonowania społeczeństwa oraz na dynamikę społeczną – procesy rozwojowe. Tak jak w naukach przyrodniczych panuje przekonanie, że królową nauk jest matematyka, tak Comte był zwolennikiem imperializmu socjologicznego (socjologia zna odpowiedzi na pytania, które poszukują inne nauki i tak psychologia, historia itp. są redukowane do dyscyplin pomocniczych socjologii).

W tamtych czasach wydawało się że, wszystko lub prawie wszystko w fizyce zostało już wyjaśnione i pozostało już tylko kilka, mało istotnych, drobiazgów. Fizyka społeczna doskonale nadawała się do roli tego ostatniego elementu układanki, ponieważ łączyła świat fizyczny ze światem człowieka. W tym samym czasie socjologia nako nauka również stawiała swoje pierwsze kroki. Dokonało się unaukowanie refleksji o społeczeństwie przez narzucenie refleksji socjologicznej rygorów formalnych, poddanie kontroli danych empirycznych, oczyszczenie socjologii z filozoficznego i religijnego myślenia.

W pewnym momencie dokonano odkrycia, które zdawało się potwierdzać słuszność mechanistycznego podejścia do nauki o społeczeństwie. Dziś może nie brzmi to zbyt sensacyjnie, ale wówczas było uważane za wielkie odkrycie. Okazało się, że błędy w wyznaczaniu pozycji gwiazd, wahania w proporcjach płci podczas narodzin, rzuty monetą i cały szereg innych zjawisk ma pewną intrygującą cechę. Otóż można te wszystkie obserwacje opisać jedną uniwersalną krzywą nazwaną „krzywą błędu”. Najprawdopodobniej, pierwszym człowiekiem, który natknął się na ową tajemniczą krzywą był francuski astronom Pierre-Simon Laplace. W 1781 roku, podczas badania proporcji płci noworodków urodzonych w kolejnych latach, natknął się na przypadkowość, która jednak oscylowała wokół pewnej wartości średniej. Stwierdził, że najprawdopodobniej jest to przejaw procesu losowego, a nie jak wówczas uważano wola boża, która układa proporcje płci dokładnie po równo. Wykazał on, że owe błędy mogą być opisane za pomocą tak zwanej „krzywej błędu” odkrytej w 1733 przez matematyka Abrahama De Moivre'a. Opisał on przy jej pomocy statystykę rzutów monetą. Obecnie owa tajemnicza funkcja jest znana jako krzywa Gaussa. Odkrycie to mocno poruszyło Adolfa Queteleta, gdy przybył w 1823 roku do Francuskiego Obserwatorium Astronomicznego aby uczyć się od Laplace'a i Poissona. Quetelet przechrzczył „social physics” na „mecanique sociale” i zajął się popularyzacją badań Laplace'a w tej dziedzinie. Była to mechanistyczna nauka osadzona mocno w statystyce matematycznej. Jego działania wywarły ogromny wpływ na wielu współczesnych i późniejszych myślicieli. John Hershel bardzo pozytywnie wyrażał się o jego pracy, Florence Nightingale zaproponowała żeby uczyć tej nowej nauki na Oxfordzie, a Karol Marx wykorzystał jego teorie w swoich dziełach. W ten sposób powoli wyłaniał się obraz społeczeństwa, które w skali mikro może zachowywać się

kompletnie nieprzewidywalnie-losowo, ale w skali makro przestrzega pewnych, bardzo ścisłych (niczym w fizyce) reguł. Początek XX wieku przyniósł ogromne zainteresowanie procesami społecznymi a socjologia jako nauka niekwestionowanie zajęło swoje miejsce na katedrach uniwersyteckich. Współczesny obraz tej dziedziny zawdzięczamy głównie Emilowi Durkheimowi, ponieważ uznał on, że społeczeństwo istnieje ponad jednostką jako system działających sił. Tajemnice życia społecznego przyrównał do organizmu. Zaobserwował, że jednostki zgromadzone razem zachowują się inaczej niż w odosobnieniu (*psychologia tłumów*). Rzeczywistość społeczna jest dla jednostki rzeczywistością zastaną (człowiek nie tworzy języka czy religii), co wiąże się ze świadomością fałszywą. Jako pierwszy empirycznie na podstawie rejestru samobójstw zbadał rozbieżność między zamierzeniami jednostek a skutkami ich działalności, społeczeństwo rozwija się według własnych praw, w których obliczu jednostki mogą być bezsilne. Od Durkheima nie ma już naukowych podstaw, aby kwestionować podstawową zasadę kryjącą się za modelami iż społeczeństwo nie jest „sumą” jednostek. Znany fizyk Otto Wiener na początku XX wieku przedstawił swoje matematyczne modele rozwoju społecznego korzystając z zagadnień rozwijanej w tym czasie teorii ewolucji.

Jednak rozwój modelowania w naukach społecznych został później zahamowany przez rozkwit humanistycznej i neopozytywistycznej filozofii nauki lat XX-lecia międzywojennego oraz osłabiony przez kolejne odkrycia mechaniki kwantowej, które częściowo zrywały z klasyczną logiką. Następnie zainteresowanie socjologów indywidualistycznym nurtem postmodernizmu przez postawy typu „anything goes” Feyerabenda i Kuhna doprowadziło do sporu o metody nauk społecznych, czego rezultatem na gruncie socjologii polskiej (co nie mało miejsca w krajach o wyższej kulturze naukowej, np. w Skandynawii, czy krajach anglosaskich) było zerwanie z technikami nauk przyrodniczych (za S. Ossowskim i P. Sztompką). Być może konsekwencją tego stanu rzeczy, jest kryzys polskiej socjologii, którego największym wyznacznikiem było toż rola ruchu „Solidarność” została zdecydowanie niedoszacowania. Paradoksalnie jednak, dzięki rozwojowi komputerowych technik obliczeniowych, metodologia stworzona przez neopozytywistów doskonale sprawdziła się w zastosowaniu fizyki w innych dziedzinach, czasem mocno odległych. Analiza statystyczna danych finansowych zaczęła być prowadzona odkąd tylko dane te mogły być zapamiętywane, czyli od roku 1950 gdy dostępna stała się komputerowa obróbka.

### 3.3 Metody obliczeniowo - statystyczne zdobywają naukę

Późniejsze odkrycie Mechaniki Kwantowej ukazało statystyczną naturę rzeczywistości nawet w najmniejszej najbardziej fundamentalnej skali. Oczywiście trzeba tu nadmienić, że owa natura jest wyraźnie innego rodzaju niż ta spotykana w układach złożonych. Ostatecznie w 1918 roku Marian Smoluchowski stwierdził że prawdopodobieństwo jest centralnym problemem współczesnej nauki. Gdyby nie sukcesy socjologów nie wiadomo jak długo jeszcze fizyka trwałaby w deterministyczno-redukcyjnej utopii. Do przenikania się dziedzin przyczynił się także rozwój techniki obliczeniowej, przez co możliwe stało się analizowanie dużych zbiorów danych. Na przykład impuls do rozwoju geometrii fraktalnej dał B. Mandelbrot, który w 1963 r. badał ceny bawełny i zauważył, że rozkład ich fluktuacji ma charakter potęgowy. Geometria fraktalna znalazła następnie zastosowanie w fizyce, do opisu struktury wielu układów naturalnych i procesów fizycznych. Olbrzymi impuls do rozwoju ekonomii, a zwłaszcza ekonometrii, dali w końcowych dekadach XX w. fizycy zatrudnieni w instytucjach finansowych, którzy dzięki znajomości fizyki statystycznej mieli odpowiednie przygotowanie, by analizować olbrzymie zbiory danych finansowych, zawierających zapisy miliardów transakcji giełdowych, często dostępnych w czasie rzeczywistym (dla porównania, w latach 60-tych B. Mandelbrot miał do dyspozycji tylko około 2 tysięcy danych). Od lat 90-tych XX wieku rynki finansowe mogły być badane na podstawie milionów a nawet miliardów danych zarejestrowanych z dokładnością nawet do kilku sekund. W efekcie już nie tylko socjologia, ale także ekonomia, którą neopozytywiści traktowali na równi z matematyką (a więc dziedziną, która była falsyfikowalna, ale nieweryfikowalna doświadczalnie) stały się według ich kryterium naukami. Bowiem nawet tak klasyczne pojęcia jak hipoteza rynku efektywnego, dzięki socjofizyce i ekonofizyce – czyli zastosowaniu metod matematycznych w socjologii i ekonomii – zostały lepiej zdefiniowane, zwłaszcza jeśli

chodzi o granice ich stosowalności. W ten sposób ekono- i socjo- fizyka, która tak naprawdę powstała w latach 90-tych XX wieku, stała się w ostatnich latach pełnoprawną dyscypliną naukową, najpierw docenioną przez fizyków, a z czasem także przez ekonomistów, których ujęła choćby przewidując krachy finansowe a na socjologów dopiero przyjdzie czas. Wiek XX należał też do badań na „żywym organizmie”, oraz symulacji komputerowych. W roku 1967 psycholog społeczny Stanley Milgram przeprowadził swój słynny eksperyment, poświęcony roli autorytetu. Nurtowało go pytanie o przyczyny ślepego posłuszeństwa wobec rozkazów zbrodniarzy II wojny światowej, które doprowadziły do ludobójstwa w obozach koncentracyjnych. Eksperyment dał szokujące rezultaty, okazało się bowiem, że ludzie pod wpływem autorytetów potrafią dokonywać rzeczy okrutnych, nawet gdy czują do nich wstręt. Ten sam człowiek w tym samym roku, przeprowadził również eksperyment, w którym pokazał, że dowolną parę ludzi na świecie dzieli średnio zaledwie 6 stopni separacji. Oznacza to, że aby przekazać między nimi informację, potrzeba zaledwie sześciu pośredników. Często rozmawiamy z kimś i okazuje się, że mamy z nim jakiegoś wspólnego znajomego. Zazwyczaj wydaje wtedy, że mamy do czynienia z niecodziennym zbiegiem okoliczności. Od tamtej pory stało się jasne, że to raczej reguła niż rzadki przypadek. Eksperymenty na żywych ludziach, czy realnych społeczeństwach to tylko jedna z dostępnych obecnie metod badawczych. Oczywiście takie badania mają swoją niewątpliwą zaletę, jaką jest realizm. Z drugiej jednak strony trudno jest o przeprowadzenie wielu powtórzeń tego samego eksperymentu, albo eksperymentowanie na bardzo dużych liczbach prawdziwych ludzi. Na szczęście postęp w dziedzinie elektroniki włożył w ręce naukowców narzędzia o jakich wcześniejsze pokolenia nie śmiały nawet marzyć. Nowoczesna technologia lat 70 pozwalała na tworzenie prostych, tak zwanych "zabawkowych" modeli komputerowych, próbujących odtwarzać zjawiska społeczne. Jedną z pierwszych takich prób był powstały 1971 roku model Schellinga (Schelling, 1971), mający tłumaczyć przestrzenną segregację rasową w grupach społecznych. Symulacje wykazały, że w grupie złożonej z dwóch typów osobników (np. czarni i biali) nawet stosunkowo słabe preferencje dotyczące sąsiedztwa (biali preferują białych sąsiadów, ale akceptują pewną liczbę czarnych sąsiadów i odwrotnie) powodują całkowitą segregację na dwie grupy. W ten sposób Thomas Schelling pokazał jak z mikroskopowych reguł rodzą się makroskopowe struktury. Późniejsze lata to okres powstawania wielu podobnych modeli. W latach 80 spory wkład w dziedzinę wniósł Serge Galam (Galam, 2012). Jak wynika ze wstępu tego rozdziału, nie powinno się szukać początków tej dziedziny w tym okresie, gdyż ta dziedzina powstała już dawno, a jedynie nazwa „sociophysics” jest nowa. Jednak nikt przed Galamem nie zaproponował przeniesienia wprost metod i modeli teorii przejść fazowych na grunt socjologii. Wykorzystał on znany z fizyki statystycznej – wykorzystywany do opisu ferromagnetyków – model Isinga do modelowania wyników wyborów. Trzeba jednak wspomnieć, że mimo postępu w metodach badawczych nawet lata 80 nie były okresem gwałtownego rozwoju socjofizyki. W tych czasach zajmowanie się tą dziedziną nadal należało do rzadkości. Prawdziwy rozkwit nastąpił dopiero w okolicach roku 2000, kiedy pojawiło się dużo nowych modeli, starających się odtworzyć całe spektrum zjawisk społecznych. Również wtedy zaczęły się odbywać pierwsze konferencje poświęcone socjofizyce. Narodziły się też pierwsze sekcje towarzystw fizycznych, oraz działają czasopisma, zorientowane na tę tematykę.

Na przełomie wieków socjofizyka wzbogaciła się o cały szereg różnych modeli próbujących odtwarzać bądź opisywać wiele różnych zjawisk. Niektóre zagadnienia określiły główny nurt, a inne pozostały niszowe. Jednym z jej najbardziej popularnych tematów jest modelowanie dynamiki opinii. Służy ono między innymi do symulacji: wyników głosowania, zmiany poglądów w społeczeństwie, rozprzestrzeniania się innowacji itp. Niektórymi modelami dynamiki opinii zajmę bardziej szczegółowo w dalszej części niniejszej rozprawy. Badania dynamiki opinii wydają się szczególnie ważne w epoce Internetu, mediów społecznościowych, wzmożonego przepływu danych, koncepcji, poglądów.

Dzięki doskonale rozwiniętej technologii obliczeniowej w oparciu o metody fizyki statystycznej ta młoda nauka na ogół jest weryfikowalna. Oczywiście są w socjo- i ekono- fizyce teorie, które w praktyce sprawdzają się zadziwiająco doskonale, ale są mało falsyfikowalne i wymagają jeszcze wyjaśnienia.

### 3.4 Wspomagane komputerowo modelowanie i myślenie systemowe

Inteligentne metody w opisie zjawisk i w praktyce społecznych mają też swoją genezę informatyczno-matematyczną. Komputery stworzyły nową jakość. Prawidłowości, poszukiwane przez nauki czy społeczne, dają się opisać algorytmami określającymi zachowanie się systemów. Programy komputerowe pozwalają na zbadanie konsekwencji zakładanych praw, symulację rozwoju komplikowanych systemów i określanie własności systemów. Komputery pozwalają na robienie doświadczeń w sytuacjach zbyt skomplikowanych, by możliwa była uproszczona analiza teoretyczna.

„Transformacja społeczeństwa przez rewolucję naukową XIX i XX wieku zostanie wkrótce przyćmiona przez jeszcze dalej idące zmiany, wyrastające z naszych rosnących możliwości zrozumienia złożonych mechanizmów, które leżą w centrum zainteresowania człowieka. Bazą technologiczną tej nowej rewolucji będą niewyobrażalnie potężne komputery razem z narzędziami matematycznymi i eksperymentalnymi oraz oprogramowaniem, które jest niezbędne by zrozumieć układy złożone... Przykładami adaptujących się, złożonych systemów jest ewolucja biologiczna, uczenie się i procesy neuronalne, inteligentne komputery, chemia białek, znaczna część patologii i medycyny, zachowanie się ludzi i ekonomia.”

Lata 90 XX wieku, Murray Gell-Mann, tłum. W. Duch

O rewolucji komputerowej więcej pojawi się w rozdziale o informatyzacji, ale należy zaznaczyć, że informatyzacja ma wpływ na społeczeństwo jak i na metody badawcze opisujące to społeczeństwo. Nauki informatyczne są kompleksem wielu technik matematycznych i komputerowych. Metody komputerowe są wspólne dla wielu dziedzin jak programowanie, analiza numeryczna, techniki wizualizacji danych, symulacji, modelowania, nienumeryczne metody komputerowe. Komputery o ogromnej mocy są obecnie nadzieją na wiele problemów człowieka i choć wiele udało się już dokonać (np. sekwencjonować genom człowieka), to wiele wyzwań nie ruszyło w stopniu istotnym (np. zwalczanie chorób genetycznych). Jednak początek nauk informatycznych to **cybernetyka** (XIX-XX w.), czyli systemy sterowania i automatycznego przetwarzania informacji w m.in. mechanice, biologii. Następnym krokiem była **sztuczna inteligencja** (lata 50. XX w.) – dziedzina nauki, której celem jest zaprojektowanie inteligentnych maszyn. Sukcesy modelowania w zakresie optymalizacji i teorii podejmowania decyzji doprowadziły do pierwszych udanych prób wykorzystania komputerów w pracy społecznej w postaci **systemów eksperckich** (lata 70. XX w.). Wykorzystują one logikę matematyczną do imitacji wnioskowania człowieka. Zainteresowanie sieciami neuronowymi oraz fascynacja mózgiem narodziła **nauki kognitywne** (lata 70. XX w.). Poznanie i modelowanie zjawisk dotyczących działania umysłu stało się motorem do opisu zjawisk. Następnie algorytmy genetyczne, drzewa decyzyjne stały się bazą **uczenie maszynowe** (lata 90. XX w.). Stosuje się modele statystycznych i metody optymalizacji do projektowania algorytmów uczenia maszyn. Obecnie stosuje się wiele nazw i wykorzystuje różnorodne metody w **Inżynierii wiedzy** (początek XXI w.). Koncentruje się na sposobach reprezentowania wiedzy, procesach pozyskiwania wiedzy, opracowaniu systemów informatycznych wykorzystujących wiedzę. Jest to dyscyplina techniczna, która dotyczy przetwarzania wiedzy w systemach informatycznych w celu rozwiązywania problemów zazwyczaj wymagających wiedzy eksperckiej (człowieka-specjalisty). Wiedza jest podstawowym składnikiem wnioskowania i działania i składa się z faktów z danej dziedziny, reguł, heurystyk, teorii i ogólnie przyjętych schematów postępowania (komputerowe wspomaganie decyzji). Rozwój technologii informatycznych pozwolił, iż systemy komputerowe zaczęły pełnić istotną rolę w procesach decyzyjnych, szczególnie tam, gdzie do podjęcia decyzji konieczne jest szybkie przetworzenie ogromnych ilości danych poprzez przeszukiwanie i analizę hurtowni danych za pomocą wyuczonych reguł. Obecnie spory procent transakcji, zwłaszcza na rynku walutowym jest obsługiwana przez automaty. Popularnym narzędziem wykorzystywanym przez przedsiębiorstwa są systemy klasy business intelligence. Uznaje się, że jedną z przyczyn światowego kryzysu ekonomicznego 2007 było używanie bez kontroli komputerowych algorytmów decyzyjnych zwłaszcza w zakresie instrumentów pochodnych (kontraktów finansowych pozwalających na duże zarobki przy małych inwestycjach, ale objętych dużym ryzykiem). Według jednej z dyrektyw unijnej w celu zapobieżeniu przyszłym kryzysom, opracowanej na podstawie pracy matematyków i przyrodników, zakazano używania automatów powyżej pewnego poziomu ryzyka (ang. Value at risk).

Myślenie w kategoriach przetwarzania wiedzy za pomocą modelu jest nazywane myśleniem systemowym i jest od dawna stosowane w inteligentnych rozwiązaniach życia społecznego. Może to zdziwić polskich socjologów, gdzie wciąż dominującym paradygmatem jest „czynnik humanistyczny” Znanieckiego, ale na skandynawskich uczelniach studenci nauk społecznych zazwyczaj uczestniczą w kursach tego typu. Jedną z wartych podkreślenia inicjatyw zmierzającą w kierunku „myślenia systemowym” jest SZOPA<sup>4</sup> przy Uniwersytecie Techniczno-Przyrodniczym w Bydgoszczy. Jest to projekt skierowany do przyszłych przedsiębiorców, jednak bazowa analityczna struktura oparta na modelu jest mocno eksponowana. Proponowany jest przez autorów schemat:

ZROZUMIENIE – zebranie podstawowych zasobów informacji, które posłużą do zrozumienia zagadnienia i określenia co jest wyzwaniem w całym procesie.

OBSERWOWANIE – rozwinięcie odczucia empatii i utożsamia się z klientem ostatecznym.

DEFINIOWANIE – ustalenie właściwego punktu widzenia, który nada odpowiedni bieg dalszej pracy.

WYMYŚLANIE – budowanie skojarzeń po to, żeby w końcu wybrać rozwiązanie jedyne w swoim rodzaju.

PROTOTYPOWANIE – stworzenie modelu, który pomoże w dalszej komunikacji, wyeliminowaniu podstawowych błędów i dalszej inspiracji

TESTOWANIE – obserwacja w jaki sposób prototyp spełnia oczekiwania.

Niejako równoległe z myśleniem systemowym idzie wykorzystanie metod statystycznych i komputerowych w analizie danych (analiza zawartości/treści, analizy trendu, testy hipotez statystycznych itp). Opis statystyczny ma sumaryczny charakter, co oznacza, że dotyczy on całej zbiorowości generalnej bądź próbnej, a nie poszczególnych jednostek statystycznych. Wiąże się to z klasyczną ilościową analizą treści (budowanie klucza kodowego, wybór jednostek analizy) jaki i z elementami jakościowej analizy treści (analiza pola semantycznego kodowanie materiałów). Po odpowiednim przygotowaniu i zebraniu materiału (unikając błędów losowych i systematycznych), dane społeczne poddawane są po procesowi czyszczenia. Wykorzystywana na tym etapie jest teoria prawdopodobieństwa czyli wyrażenie szansy wystąpienia jakiegoś zdarzenia. Znając algebrę zdarzeń istotne jest wykorzystanie prawdopodobieństwa warunkowego. W statystyce dominują dwa nurty Bayesowski (gdzie dominują warunkowe struktury) i częstotliwościowy (gdzie obiekty opisywane są przez częstości występowania). W podejściu Bayesowskim korzystamy z jakiejś apriorycznej wiedzy o zdarzeniach oraz metodami zazwyczaj symulacyjnymi dochodzimy do wyniku. Natomiast w podejściu częstotliwościowym zlicza występowanie cech a następnie najczęściej szuka się w nich zależności. Trzy podstawowe metody służą do tego: regresja, klasyfikacja i klasteryzacja.

**Regresja** jest najczęściej stosowaną metodą i jednocześnie najprostszą. Najczęściej wykorzystuje się regresje liniową, ale jeżeli za zależnością kryje się bardziej skomplikowana funkcja, to ją również możemy wykorzystać. Celem regresji jest skonstruowanie modelu na podstawie danych (zmienna niezależna – wyjaśniająca i zmienna zależna – wyjaśniana) aby zmierzyć siłę zależności. Ponadto model konstruowany jest tak, aby możliwe było przewidywanie nowych obserwacji. Tak samo, gdy mamy do czynienia z więcej niż jednym czynnikiem wyjaśniającym, to wtedy wykorzystamy regresję wieloraką. W regresji standardowo zakłada się, że dysponujemy obserwacjami z odpowiadającymi im wartościami liczbowymi. W przypadku zmiennych nominalnych typu płeć stosuje się regresje z predyktorami jakościowymi. Tak samo zmienna zależna wyjaśniana gdy nie jest w postaci liczbowej, to wymagane są odpowiednie transformacje. Przykładem jest regresja stosunku liczby rozwodów do nowo-zawieranych małżeństw w zależności od siły norm społecznych (mierzonymi poparciem partii konserwatywnych) będąca pobocznym wynikiem badań z

---

<sup>4</sup> <http://shopa.utp.edu.pl/>

rozdziału o wiązaniu się w pary (Tabela 1). Zależność jest ujemna (to znaczy im wyższe poparcie dla PiS, tym mniej stosunkowo mniej rozwodów), oraz istotna statystycznie (błąd mniejszy od współczynnika regresji, a p-Value jest mniejsze większość progów istotności jak 0,05 lub 0,01). Więc wzrost poparcia PiS o jeden procent zmniejsza występowalność rozwodów o 0,11%.

N=17 (Województw)	Współczynnik regresji	Niepewność regresji.	p-Value
Procent poparcia "PiS" w województwie	0,11	0,06	<10 <sup>-3</sup>

Tabela 1 Regresja dla znormalizowanej liczby rozwodów w latach 2010-2012, której wskaźnikiem jest poparcie konserwatywnej partii "PiS" wyborów parlamentarnych w 2011 roku zgrupowane według województw Polski

Wnioskowanie statystyczne pozwala na ustalenie na ile dana zależność bądź jej zmiana jest istotna statystycznie. Same stwierdzenie zależności nie wiąże się bezpośrednio z ustaleniem przyczynowości, na co są dodatkowe metody, np. bazujące na następowności czasowej. W tym celu stawia się hipotezę zerową, np. czy poziom poparcie PiS wiąże się z prawdopodobieństwem rozwodu, przeciw hipotezie alternatywnej, że takiego związku nie ma. Hipotezę weryfikuje się testem i w zależności od poziomu wartości p-Value przyjmuje się hipotezę zerową (nie ma podstaw do odrzucenia hipotezy zerowej), albo się ją odrzuca (stwierdza się, że nie ma podstaw do przyjęcia hipotezy zerowej). Do oceny zależności wykorzystuje się iloraz szans (ang. odds ratio - OR), albo prawdopodobieństwo względne (ang. relative risk - RR).

$$y = a + bx$$

$$a = \frac{\sum y - b \sum x}{n}$$



$$b = \frac{n \sum (xy) - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2}$$



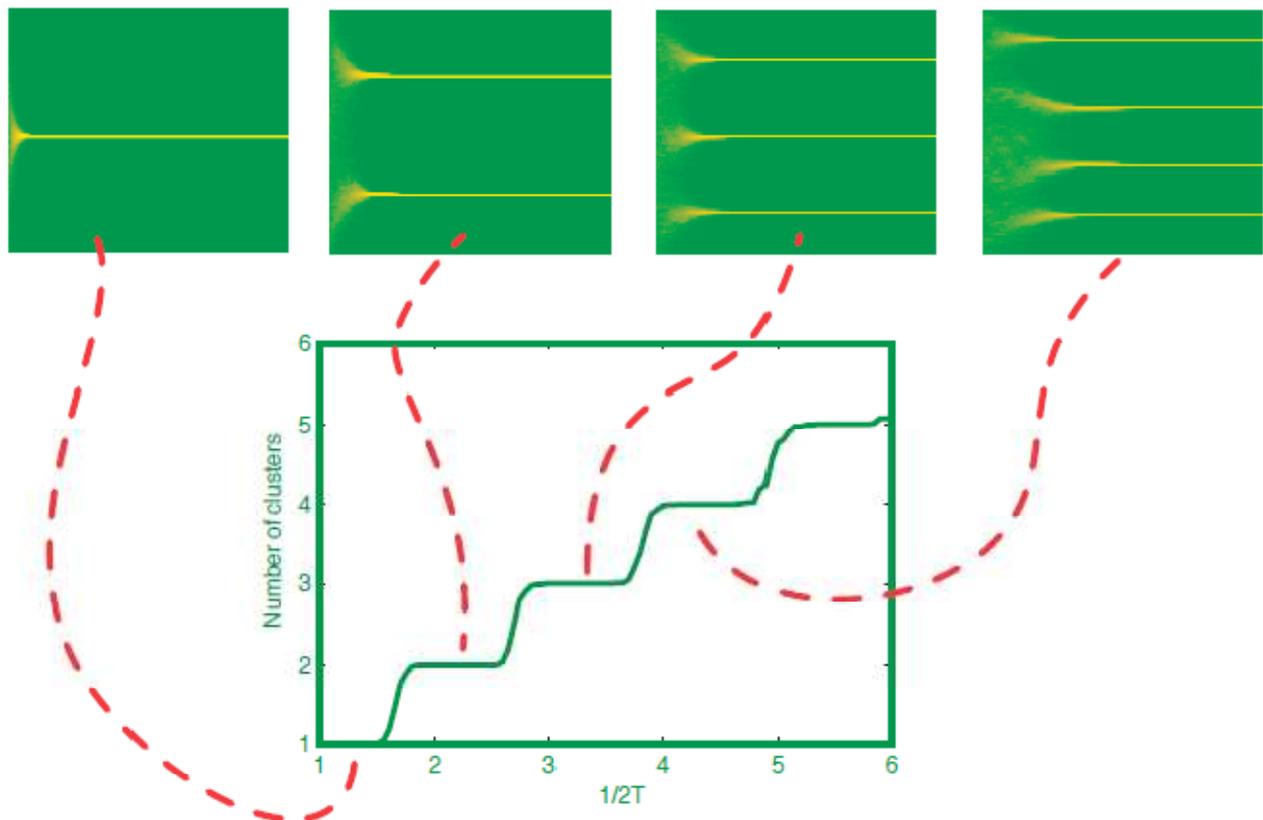
Rys. 14 Przykład wykorzystania regresji liniowej (z estymacją parametrów metodą najmniejszych kwadratów)

**Klasyfikacja** jest bardziej skomplikowaną metodą klasyfikacji danych do pewnych kategorii. Załóżmy, że dysponujemy obserwacjami z etykietami (klasami), czyli posiadamy wiedzę ekspercką. Celem algorytmów I metodologii jest skonstruowanie klasyfikatora separującego obiekty należące do różnych klas. Problem bardzo często pojawia się przy automatycznym kodowaniu danych. Analizując wypowiedzi w Internecie, np. chcielibyśmy znać atrybuty autora, np. przynależność partyjną. W takim przypadku możemy ręcznie/automatycznie otagować autorów wedle ustalonych reguł klasyfikacji. Czasami potrzebujemy, żeby klasyfikator konstruowany jest tak, aby możliwe było przewidywanie klas nowych, niesklasyfikowanych obserwacji. Z metodami klasyfikacji wiążą się pojęcia **precyzji** (czy wyniki badań dają podobne rezultaty) oraz **dokładności** (czy estymowana wartość zgadza się z rzeczywistością) oraz dla testów binarnych określa się **falszywie pozytywne** i **falszywie negatywne** proporcje<sup>5</sup>. Załóżmy, że chcemy określić, czy respondent pójdzie na wybory: tak/nie. Jeżeli wśród wszystkich faktycznych uczestników wyborów wszyscy zostali zaklasyfikowani na tak, to mówimy o braku wyników fałszywie pozytywnych. Jeżeli wśród wskazanych na tak wszyscy poszli na wybory, to mamy brak wyników fałszywie negatywnych. Czasami zależy nam na **czułości** (dopełnienie proporcji fałszywie negatywnych), kiedy chcemy zaklasyfikować jak najwięcej respondentów, którzy faktycznie pójdą na wybory na tak. Czasami na **specyficzności** (dopełnienie proporcji fałszywie pozytywnej), kiedy chcemy żeby wśród zaklasyfikowanych na tak jak najwięcej było takich, którzy

<sup>5</sup> Niestety dokładność i precyzja w niektórych naukach jak informatyka nachodzi znaczeniowo na fałszywie pozytywne/negatywne stosunki.

faktycznie pójdą na wybory. Idealna klasyfikacja to 100 % specyficzności i czułości, ale w rzeczywistości społecznej jest to praktycznie nie możliwe. Ogólną moc testu wyznacza się krzywą ROC, gdzie na jednej osi odkładamy wartość recall (czułości), a na drugiej wartość (1-specyficzność). We wnioskowaniu statystycznym należy uważać na wiele czynników, np. niewłaściwego zaklasyfikowania jakiejś kategorii. Może też wynikać ze stronniczego doboru próby do badań z populacji. Zakłócenie (ang. confounding) stanowi błąd wywołany przez inną zmienną pozostającą w związku statystycznym między przyczyną a skutkiem.

**Klasteryzacja**<sup>6</sup> za to jest metodą oddzielania i grupowania danych bez uprzedniej wiedzy na temat ich klasyfikacji. Dysponujemy obserwacjami bez etykiet. Celem uczenia jest znalezienie grup (klastrow), w których skupione są dane. Przykładem są algorytmy detekcji społeczności z rozdziałów o sieciach społecznych. Obserwacje z danej grupy charakteryzują się podobieństwem. Możliwe są alternatywne nazwy: grupowanie, analiza skupień.



Rys. 15 Wizualizacja wzrostu liczby klastrow w modelowanym układzie

**Redukcja Wymiarów** jest metodą wyszukania zbioru cech o największym wpływie na badane zjawisko. Wracając do kwestii rozwodów, weźmy szczegółowe rejestry z informacjami (dane demograficzne, społeczne, ekonomiczne itp.) o osobach które się rozwiodły. Gdy wymiarów jest dużo, ciężko jest stwierdzić, które cechy mają największy wpływ na decyzję o dochodach. Możliwe jest skorelowanie tych cech, np. majątek będzie związany z dochodami. Dodatkowo różne zmienne, będą wchodzić w interakcje, np. praca w budżetówce i zarobki (osoby o wysokich zarobkach z sektor państwowego często pełnią ważne funkcje publiczne, i jako że podlegają kontroli społecznej, toteż rozwodzą się rzadziej niż to by wynikało z ich charakterystyk społeczno-ekonomicznych). Metody wykorzystujące redukcje wymiarów mają na celu

<sup>6</sup> Klasteryzacja jako miara usieciowienia jest inaczej zdefiniowana, stąd często stosuje się jej zamienną nazwę – gronowanie

zgrupowanie zmiennych w agregaty zmiennych oraz wychwycenie dominujących składowych wyjaśniających zmienność cechy objaśnianej.

**Sieci neuronowe i algorytmy genetyczne:** Dzięki wdrożeniu typowego dla istot żywych jak ludzi zjawiska adaptacji komputer się uczy. Sztuczne sieci neuronowe stały się nowym paradygmatem obliczeń w których mózg jest wzorcem ideowym sztucznej sieci neuronowej. Elementem budulcowym takiej sieci jest perceptron – jednostka połączona w sieci neuronowej z wejściami i wyjściem. Z wykorzystaniem perceptronu można tworzyć bramki logiczne co daje rozwiązywać te same problemy co klasyczne topologie komputerów. Sieć Hopfielda pozwala na wyszukiwanie informacji za pomocą modelu pamięci asocjatywnej. Oddziałując bodźcami uczymy sieć neuronową do rozwiązywania konkretnych problemów poprzez odpowiednie dobieranie połączeń między elementami i ich wag. Inną metodą adaptacji są algorytmy genetyczne. Tworzy się funkcje celu dzięki i pozwala się na jej ewolucje przez mutacje, krzyżowanie i selekcja. Następnie z rodziny funkcji wybierana jest najlepsza. Najprostszy algorytm genetyczny pozwala na poszukiwanie maksimum. Tego typu metodologie wywodzą się z obserwacji przyrody i ludzi i trafiają do wzmocnienia metod ścisłych.

**Meta-analiza** jest metodą porównania różnych wyników w celu wyciągnięcia zbiorczego wniosku. W książce „Historia rocka według krytyków” użyliśmy aparatu matematycznego aby określić, które albumy w danym roku były obiektywnie najlepsze według opinii krytyków które traktowaliśmy jako pojedyncze subiektywne obserwacje. Inną nazwą tej metody jest głosowanie, ponieważ ostateczna decyzja jest podejmowana według pewnych reguł ze zbioru decyzji podrzędnych. W metodzie zakłada się, że poszczególne pomiary są obciążone błędem, to jednak średnia z nich powinna być najlepszym estymatorem szukanej wielkości, co ma swoje matematyczne uzasadnienie w teorii wielkich liczb oraz centralnym twierdzeniu granicznym.

### 3.5 Układy złożone jako inspiracja do modelowania procesów społecznych

Zjawiska społeczne, a zwłaszcza te dotyczące nieracjonalnych zachowań ludzkich są niezwykle trudne do opisanie i wyjaśnienia mimo posiadania wciąż rozwijającego się aparatu statystycznego i informatycznego. To, co obserwuje się w układach, czy – szerzej – w cywilizacjach, jest efektem oddziaływania na te układy ogromnej liczby czynników wewnętrznych i zewnętrznych. Do tych pierwszych zaliczyć można przepływ informacji pomiędzy uczestnikami życia społecznego, brak znajomości warunków początkowych, mikroskopowego stanu układów i precyzyjnie wyznaczonych praw nimi rządzących czy, ogólniej, praw społecznych. Do czynników zewnętrznych można z kolei zaliczyć zdarzenia polityczne, czynniki naturalne (katastrofy, zjawiska klimatyczne) i technologiczne oraz wpływ innych społeczeństw. Co więcej, uczestnikami rynków są istoty inteligentne, które potrafią się uczyć i dostosowywać swoje zachowanie do zmiennych okoliczności. Olbrzymia liczba tych czynników i ich wzajemne powiązania powodują, że rynki finansowe są jednymi z najbardziej skomplikowanych układów istniejących we Wszechświecie. Wszystko to razem sprawia, że poprawny i całościowy opis tych rynków wymaga wzięcia pod uwagę informacji dostępnej na różnych poziomach organizacji: na poziomie pojedynczych jednostek i innych instytucji społecznych, poziomie branż, geografii oraz zjawisk makroekonomicznych. Potrzebę takiego całościowego opisu spotyka się w przypadku wielu innych układów występujących w przyrodzie i strukturach społecznych. Układy takie zaliczane są do klasy, która jest przedmiotem zainteresowania stosunkowo nowej dziedziny - nauki o układach złożonych. Układy tego typu charakteryzują się dużą liczbą oddziałujących ze sobą elementów składowych, a nieliniowy charakter tych oddziaływań sprawia, że całość zachowuje się inaczej niż układ, będący prostą sumą elementów składowych.

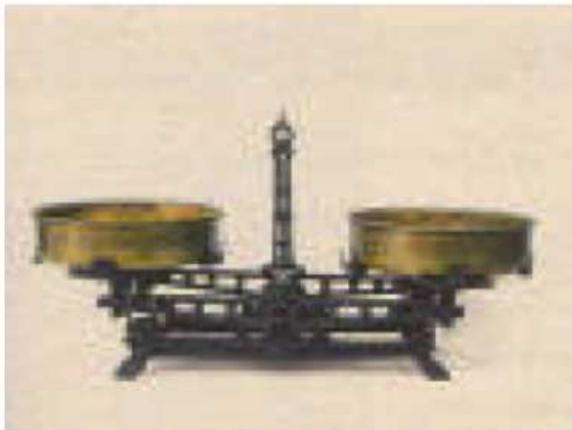
Koncepcja układów złożonych zaczęła się klarować od lat 50-tych XX w. dzięki pracom biologa Ludwiga Bertalanffy, który zajmował się teorią układów otwartych. Skrytykował próby opisu żywych organizmów na gruncie termodynamiki układów zamkniętych, będąc zdania, że układy biologiczne wymagają zupełnie nowego podejścia, uwzględniającego ich otwartość i wymianę energii i informacji z otoczeniem. Z punktu widzenia dynamiki takich układów bardzo ważne jest, że Bertalanffy, wbrew stanowisku neopozytywistów,

twierdził, że zmiany zachodzące w procesie ewolucji układów biologicznych zachodzą skokowo, co według dzisiejszego stanu wiedzy wydaje się być prawdą. Za rozwinięcie tej koncepcji można uznać np. późniejszą teorię *punctuated equilibria*, opisującą ewolucję jako ciąg długotrwałych stanów równowagi przerywanych gwałtownymi zmianami. Od czasów prac Bertalanffy'ego pojęcie złożoności układów uległo rozwinięciu, jednak do dziś nie ma jednej uznanej powszechnie definicji, która nie budziłaby żadnych wątpliwości. Wiąże się to przede wszystkim z obecnością tego terminu w wielu odrębnych dziedzinach nauki, na przykład w psychologii, socjologii, cybernetyce, biologii, fizyce, matematyce i teorii informacji, a – co za tym idzie – wielu różnym próbom jego formalizacji (mimo, że intuicyjnie idea złożoności może być w każdej dziedzinie rozumiana podobnie). Nauka o złożoności określa układy złożone jako układy nieredukowalne, czyli takie, w których na poziomie makroskopowym występują zjawiska, których nie da się wytłumaczyć jedynie na podstawie znajomości ich struktury mikroskopowej. Są to tzw. zjawiska emergentne. Teoria systemów definiuje te własności jako (Poczobut, 2006):

- własności układów, których nie obserwuje w ich podukładach;
- własności, które są rezultatem dynamicznej organizacji ich elementów;
- własności globalne, które mogą być efektem mieszanym czynników wewnętrznych i zewnętrznych, związanych z przeszłością układu i warunkami środowiska;
- własności funkcjonalne na wyższym poziomie organizacji mają potencjalnie wielorakie wyjaśnienie na niższych poziomach organizacji (brak jednoznaczności pomiędzy poziomami organizacji).

W przyrodzie zjawiska emergentne są powszechnie spotykane, np. skomplikowane formy ukształtowania terenu, takie jak niektóre typy wybrzeży o strukturze fraktalnej (fiordowe, riasowe, szkieletowe), łańcuchy górskie, wydmy piaszczyste, fraktalna struktura systemów rzecznych, wielkoskalowe zjawiska atmosferyczne, takie jak tornada, niektóre formacje chmur, przepływy turbulentne cieczy, życie i formacje organizmów żywych (kolonie bakterii, ławice ryb, gromady ptaków). Podobnie, zjawiska emergentne są cechą układów społecznych: pojawienie się specyficznych cech w psychologii tłumu, istnienie ruchów społecznych, struktur państwowych, wojen, pieniądza, rynków finansowych, literatury, sztuki itp. Zjawiskiem emergentnym na pograniczu biologii i zjawisk społecznych jest z kolei język naturalny i związane z nim efekty wyższego rzędu, takie jak gramatyka czy styl literacki. Z drugiej strony emergentność zjawiska, które może się wydawać najbardziej złożonym spośród nietrywialnych cech organizmów żywych, czyli świadomości, budzi kontrowersje, ponieważ jest ciągle słabo poznane (Poczobut, 2006). Cechą układów złożonych, która prowadzi do pojawienia się zjawisk emergentnych, jest samoorganizacja. Jest to długotrwały proces ewolucji układu, który na skutek oddziaływania z otoczeniem poprzez wymianę informacji (energii), zmienia swą strukturę wewnętrzną w celu adaptacji do nowych warunków. Ta zmiana struktury, a właściwie ciąg następujących po sobie zmian, ma charakter nieodwracalny. O nieodwracalności ewolucji układów pisał już L. Dollo w latach 90-tych XIX w., który odnosił ją do układów biologicznych, ale jest to prawdą dla dowolnych układów złożonych. Z punktu widzenia fizyki, nieodwracalność ewolucji jest skutkiem drugiej zasady termodynamiki oraz własności nieliniowych układów o dynamice dysypatywnej. W tym kontekście warunki zewnętrzne i ich wpływ na układ mają charakter parametru kontrolnego, który zmieniając się, zmienia dynamikę układu opisywanego pewnymi równaniami różniczkowymi. Zmiany parametru powodują, że w pewnym momencie układ traci równowagę i zamiast jednego rozwiązania stabilnego pojawiają się dwa lub więcej rozwiązań, z których jedno, któremu odpowiada bieżący stan układu, jest niestabilne. Pojawia się stan nierównowagi i układ musi przejść do nowego stanu równowagi, który odpowiada jednemu z rozwiązań stabilnych. Występuje wtedy tzw. spontaniczne złamanie symetrii. Ciąg takich zmian stanów odpowiada więc ewolucji poprzez sekwencję stanów metastabilnych. Jeśli historia układu jest wystarczająco długa i wystąpiło wiele zdarzeń tego typu, to struktura układu może stać się dowolnie złożona. W konsekwencji samoorganizacja może występować w układach, które są zarazem

elastyczne, by mogły zmieniać swą strukturę, ale również stabilne, by mogły istnieć przez dłuższy okres czasu. Zmienność zachowania układów samoorganizujących się jest widoczna w zjawiskach kolektywnych, które wyłaniają się z pozornego chaosu i w takim chaosie później zanikają. Do takich zjawisk w układach fizycznych można zaliczyć np. reakcję Biefousowa-Żabotyńskiego w układach dyfuzyjnych, agregację ograniczoną dyfuzją, czy komórki Benarda w cieczy. Podobne efekty występują w świecie przyrody ożywionej. Na przykład ptaki mogą przez dłuższy okres czasu prowadzić indywidualny tryb życia, by w pewnym momencie zacząć grupować się w wielkie stada i odlatywać wspólnie w jednym kierunku, po czym w innym momencie stada ulegają dezintegracji i ptaki powracają do indywidualnego stylu życia. Ta łatwość przechodzenia z nieuporządkowania do uporządkowania i odwrotnie jest jedną z najbardziej charakterystycznych i uniwersalnych cech złożoności. Na rynkach finansowych i towarowych, o których mowa będzie w dalszych rozdziałach, można także zaobserwować takie kolektywne efekty. Bańki spekulacyjne, które są fazami uporządkowanego zachowania inwestorów i wyłaniają się ze nieuporządkowanego stanu równowagi rynkowej, typowo kończą się jeszcze bardziej kolektywnymi krachami, po których równowaga i nieuporządkowanie zostają przywrócone. To samo dotyczy okresów wzmożonej liczby rozwodów, wywołanych przez stadne zachowania ludzi w obliczu zmieniających się norm. Przyczyny takich zachowań zostaną wyjaśnione w tej książce.



Rys. 16 Ilustracja Samoorganizującej się krytyczności w postaci lawin ze stanów nierównowagowych

Jednym z ciekawszych zjawisk spotykanych w układach złożonych jest tzw. samoorganizującej się krytyczności (ang. *self-organized criticality*). Jest to forma samoorganizacji, w której ewolucja układu dąży do osiągnięcia lub przywrócenia stanu krytycznego. Stanem krytycznym w tym rozumieniu jest stan, w którym układ wykazuje gotowość do dowolnie silnej odpowiedzi na zewnętrzne zaburzenie. Jest to stan pokrewny stanowi układu w okolicach punktu krytycznego przejścia fazowego drugiego rodzaju, gdzie obserwuje się pojawianie się struktur o dowolnym rozmiarze, od najmniejszych, do porównywalnych z rozmiarem całego układu (np. w zjawisku perkolacji). Mechanizm samoorganizującej się krytyczności opisuje model stożka piasku, tworzonego przez dodawanie pojedynczych ziaren. W modelu tym stożek, początkowo płaski, stopniowo zwiększa kąt swojej tworzącej aż osiągnie on taką wartość, że stożek ulega osypaniu i spłaszczeniu, po czym proces zaczyna się od nowa. Stan krytyczny odpowiada tu takiemu kątowi, w którym pojedyncze ziarno piasku może spowodować zarówno małą lokalną lawinę, jak i globalną katastrofę. Istotą procesu jest jednak to, że kąt krytyczny jest swoistym atraktorem, do którego stan stożka nieustannie dąży. Model samoorganizującej się krytyczności został wykorzystany do wyjaśnienia powszechności szumu  $1/f$ , występującego w przyrodzie np. w zjawisku trzęsień ziemi, wylewach rzek, muzyce, rytmie ludzkiego serca, sieciach neuronowych, a nawet w funkcjonowaniu mózgu.

Szum  $1/f$  jest przykładem jednego z typów zależności potęgowej, często obserwowanych w układach naturalnych. Mimo że zależności potęgowe mogą mieć bardzo różne przyczyny, w tym także dość trywialne,

to charakteryzują także zjawiska krytyczne – np. wiele wielkości fizycznych skaluje się potęgowo w pobliżu punktu krytycznego. Dlatego też ich obecność w danych empirycznych może być przesłanką, że odpowiedni układ znajduje się w stanie krytycznym. Z tego m.in. powodu panuje pogląd, że prawa potęgowe są jedną z cech układów złożonych. Z drugiej strony zależności potęgowe są obecne w strukturach fraktalnych, stąd wiele układów uważanych za złożone przejawia strukturę tego typu.

Fraktale, które są obiektami samopodobnymi, posiadają strukturę hierarchiczną. W przypadku układów złożonych taka hierarchia struktur może wynikać bądź z ich fraktalnej struktury na danym poziomie organizacji, bądź z wielopoziomowej struktury organizacyjnej. Oba typy hierarchiczności są świetnie widoczne na układach społecznych. Cywilizacja dzieli na społeczeństwa w poszczególnych częściach świata i w pojedynczych krajach, społeczeństwa krajowe dzielą się na społeczności nieformalne (te wśród nich wyróżnia się subkultury) i społeczeństwa formalne, te dzielą się na grupy społeczne, a te z kolei na niewielkie sieci powiązanych ludzi. Z innego punktu widzenia uczestnicy rynków finansowych, którzy są elementami mikroskopowej struktury układu, mogą być organizacjami (firmami, instytucjami itp.), składającymi się z ludzi. Ludzie, jako organizmy żywe, są zbudowani z komórek, które to komórki mają wewnętrzną strukturę organelli, zbudowanych z molekuł. To tworzy wspomnianą wielopoziomową strukturę, w której jedne poziomy są pierwotnymi elementami wyższych poziomów.

### 3.6 Redukcjonizm operacyjny a rzetelność badań

W świecie symulacji komputerowych danej cesze można przypisać obiekt matematyczny, który w większości przypadków sprowadza się do pojedynczej liczby. Opisywanie czegoś tak skomplikowanego jak np. opinia jakiejś osoby za pomocą tylko jednej liczby, może z początku wydawać się karkołomnym i wręcz niedopuszczalnym uproszczeniem. Okazuje się jednak, że w wielu zastosowaniach w zupełności wystarcza, a zalety takiego podejścia nierzadko górują nad wadami. Warto zaznaczyć, iż mówiąc o opinii opisanej jedną liczbą zazwyczaj nie mamy na myśli całego światopoglądu danej osoby, a jedynie jej pogląd dotyczący jednej konkretnej sprawy, czyli niejako wycinek całości. Z pewnością do opisu całego światopoglądu bardziej odpowiedni byłby wektor opinii, jednak skupmy się na pojedynczych opiniach, opisywanych za pomocą skalara.

Opisując modelowanie opinii można posłużyć się analogią do koloru. Teoretycznie, jeśli chcemy opisać barwę docierającego do nas światła, powinniśmy podać informację dotyczącą całego jego widma, wraz ze wszystkimi szczegółami, liniami widmowymi itp. Jednakowoż nasze oczy mają pewne ograniczenia i nie widzą tych wszystkich szczegółów. Mamy trzy rodzaje czopków, wrażliwych na trzy różne części widma, tak zwane kolory podstawowe (czerwony, zielony i niebieski). Zatem do pełnego opisania naszego wrażenia wizualnego wystarczą zaledwie trzy liczby, reprezentujące natężenia tych trzech kolorów podstawowych. Co prawda możemy rozróżnić niezliczoną liczbę kombinacji, ale jeśli zaczniemy je nazywać to okaże się, że nazw znamy kilkanaście, najwyżej kilkadziesiąt. Jeśli więc będziemy opisywać komuś kolor który widzimy, posłużymy się właśnie jedną z takich nazw. Można iść jeszcze dalej i ograniczyć liczbę opisywanych kolorów do kilku, a w skrajnym przypadku nawet do dwóch: czarnego i białego. Pozostając przy analogiach wizualnych, zastanówmy się jak takie uproszczenie wpływa na widziany/rejestrowany/opisywany obraz. Kolor może być opisany trójwymiarowym wektorem, w którym wartość każdej z trzech składowych RGB jest opisywana 16 bitami. Tak więc może przyjąć jedną z 65536 wartości, które będą całkiem niezłym przybliżeniem ciągłej skali. Spróbujmy teraz ograniczyć głębię koloru do 24 bitów (tzw. truecolor), a więc 8 bitów (256 różnych wartości) na każdy kolor podstawowy. W tym przypadku przybliżenie ciągłej skali już nie jest tak dobre jak poprzednio, ale mimo to jedynie nieliczne osoby są w stanie zauważyć różnicę i to jedynie na dobrym, profesjonalnym monitorze. Możemy zrobić krok dalej i skorzystać z trybu HiColor który korzysta jedynie z 15 bitów - 5 bitów (32 wartości) na składową, co daje w sumie 32768 różnych kombinacji. Tutaj już nawet niewprawione oko dostrzeże pewne niedoskonałości obrazu, które będą szczególnie widoczne w obszarach łagodnych przejść tonalnych na dużych powierzchniach. Dalej mamy tryb 256 kolorów, w którym niedoskonałości są już bardzo widoczne. Idąc jeszcze dalej, napotkamy tryby: 16 i 4 kolorów, aby na końcu

dojść do zaledwie dwóch kolorów: czarnego i białego. Podczas tej drogi widać wyraźną i coraz szybszą degradację jakości obrazu. Jednak nawet w trybie dwukolorowym możemy rozróżnić zarysy obrazu i jego ogólny kształt. I to jest właśnie sedno rozumowania. Taki binarny obraz jest o wiele łatwiejszy do opisanie. Co prawda nie dostrzeżemy na nim wielu subtelności ani koloru, ale nadal jednak jesteśmy w stanie uchwycić jego sens. Jeśli nie zależy nam na subtelnościach, a właśnie na owym ogólnym sensie, to nie ma potrzeby zajmowania się nadmierną liczbą detali. Takie upraszczanie nie jest praktyką stosowaną wyłącznie w socjofizyce. Robi się tak w najróżniejszych sondażach, ankietach itp., sprowadzając skomplikowaną opinię do pojedynczej liczby, ewentualnie zestawu liczb (wektora). Przy odrobinie dobrej woli można nawet dostrzec dalekie kwantowe analogie. Otóż proces odpowiedzi na pytanie postawione w ankiecie przypomina nieco proces pomiaru znany z mechaniki kwantowej. Oto skomplikowany światopogląd - transpozycja stanów, jest sprowadzany do konkretnego wyniku, w akcie zapytania-pomiaru. Oczywiście należy traktować to porównanie z dużym przymrużeniem oka i myśleć o nim raczej jak o niewielkiej dygresji niż czymś poważniejszym. W każdym razie wiadomo już, że można opinie opisywać liczbowo. Teraz pora na nieco więcej szczegółów, warto zastanowić się jakie mogą być rodzaje cech i odpowiadające im przykłady ze świata realnego.

- Przykładem cech ciągłych są poglądy polityczne. Określając poszczególne partie zazwyczaj posługujemy się pojedynczą cechą określającą najczęściej lewicowość/ prawicowość danej partii. Można je zatem reprezentować na osi liczbowej gdzie skrajnie lewicowym partiom przypiszemy  $-1$ , centrowym  $0$ , a tym skrajnie prawicowym  $1$ . Zatem poglądy polityczne danego agenta możemy opisać liczbą rzeczywistą z przedziału  $< -1, 1 >$ .

- Cechą dyskretną, może być wskazanie konkretnej partii. Jeśli agent chciałby głosować w wyborach, to musiałby wybrać jedną z kilku partii. Jeśli byłoby 5 partii a ich nazwami byłyby numery, to wtedy opinia agenta byłaby liczbą ze zbioru  $[1, 2, 3, 4, 5]$  czyli zmienną dyskretną.

- Cechy binarne – czyli takie, które mogą przyjmować tylko jedną z dwóch wartości np. ze zbioru  $[0, 1]$ . Z takimi opiniami mielibyśmy do czynienia gdyby były tylko dwie partie, lub jeśli byśmy zadali jakiegokolwiek pytanie na które możliwe są tylko dwie odpowiedzi, np. Czy jesteś za karą śmierci czy przeciw, lub PC vs. Mac, Canon vs. Nikon, C64 vs. Atari...

- Kiedy chcemy opisać kilka różnych cech naraz, możemy się posłużyć wektorem opinii, w którym każda z cech będzie opisana z osobna.

Mając układ agentów oraz definicje cech brakuje nam jeszcze tylko dynamiki. Jak samo słowo wskazuje będziemy mieli do czynienia z jakąś zmiennością cech w czasie. W rozpatrywanych modelach taka zmiana następuje bądź w wyniku oddziaływania agentów między sobą, bądź pod wpływem jakiegoś innego czynnika wpływającego na nich, najczęściej takiego jak szum lub pole zewnętrzne.

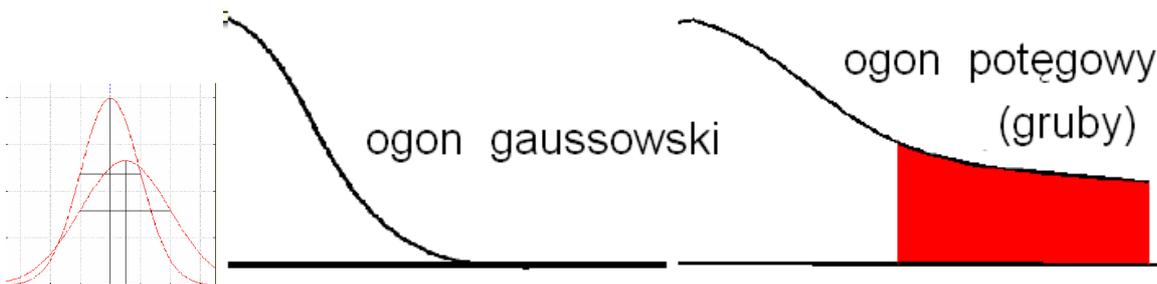
### 3.7 Model a rzeczywistość

Wiele procesów zachodzących w przyrodzie jak i w społecznościach ludzkich posiada pewne uniwersalne właściwości dlatego. Autor jest stosunkowo sceptycznie nastawiony do generalizujących wniosków dotyczących jedności praw fizycznych rządzących wszechświatem, aczkolwiek wiele rozważań na ten temat miało wpływ na obecny kształt obliczeniowych nauk społecznych.

Przykładem jest chaos deterministyczny, którego spektakularnym przykładem "efekt motyla", a odnosi się do możliwości wywołania huraganu pod wpływem machnięcia skrzydełkami motyla po drugiej stronie kuli ziemskiej. Chodzi tu o czułość układu na warunki początkowe modelu (niewielka zmiana warunków zmienia w istotny sposób wynik) i właśnie w meteorologii w połowie XX wieku taki efekt został zaobserwowany przez Lorenza. Tak samo można interpretować zdarzenia społeczne, dlatego często niewielkie zdarzenia jak błąd maklera w złożeniu zamówienia powodował krach na giełdzie.

Innym zjawiskiem jest samoorganizują się krytyczność. Występuje jeżeli ograniczoność jakiejś zmiennej związana jest istnieniem pewnego progu. Uwalnianie zgromadzonych zasobów następuje w postaci lawin (sporadycznie a nie ciągle). Takim przykładem może być Arabska wiosna 2012, która z niewielkiego lokalnego protestu skończyła się przemianami politycznymi w wielu krajach Afryki Północnej i Bliskiego Wschodu.

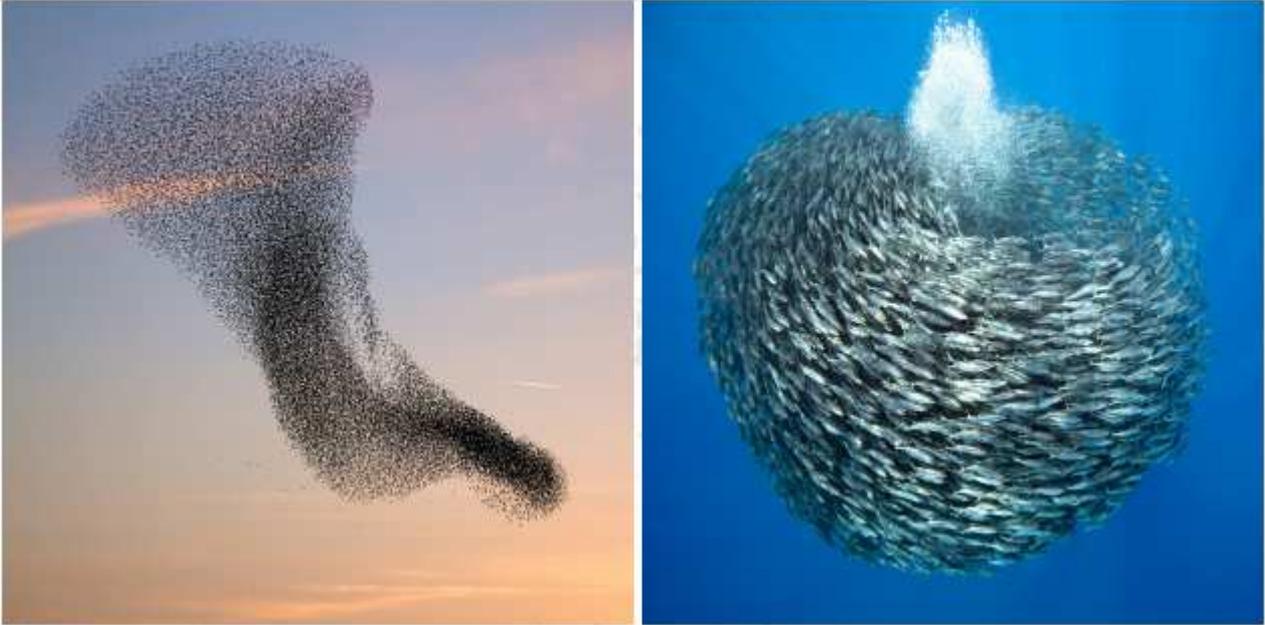
Często w opisie różnych zjawisk wykorzystuje się prawa potęgowe. W klasycznych opisach statystycznych zmienne obserwowane powstałe w wyniku procesu przypadkowego opisuje się rozkładem normalnym. Prawa potęgowe przykład takich rozkładów (z istotną z punktu widzenia fizyki teoretycznej cechą skalowania), które posiadają grube ogony (zdarzenia ekstremalne występują częściej niż to wynika z rozkładów normalnych). Okazuje się, że wiele układów podlega takim prawom: Trzęsienia Ziemi, Lawiny (piasku, śnieżne), Katastrofy naturalne (deszcze, pożary), Wymierania gatunków, Krachy na giełdzie.



Rys. 17 Różnica między rozkładem normalnym a potęgowym dominuje w ogonie (zdarzenia rzadkie – dalekie od wartości oczekiwanej)

Za układami z dynamiką progową, bądź z regułami preferencyjnymi, stoi ten sam mechanizm (prawa potęgowe), co dotyczy zdarzeń o dowolnym charakterze. Z prawami potęgowymi wiążą się sieci bezskalowe o czym będzie mowa w rozdziale o sieciach.

Przyroda dostarcza nam wiele ciekawych przykładów zjawisk, które występować mogą u ludzi. Weźmy model skoordynowanego ruchu zwierząt (ang. flocking, swarm): stada ptaków, czy ławice ryb. Zwierzęta poruszające się w takich schematach nie mają zazwyczaj lidera, a wiedza jednostek jest lokalna, jednak w większości przypadków jako stado zmiernają we właściwym kierunku. Taki model może oddawać właściwości stadne w przypadku zbiorowisk ludzkich jeżeli uwzględnimy inną zmienną cechą niż położenie jak u zwierząt.



Rys. 18 Stadne zachowania ryb i ptaków (flocking, swarm). Źródło: <http://www.flickr.com/photos/eosgreg/6965121429/>, <http://swantalks.blogspot.com/>

### 3.8 Hierarchiczność

Tematem jest dynamika i struktura korelacji w układach hierarchicznych przy wykorzystaniu metod statystycznych i sieciowych. Na podstawie korelacji można ustalić odległości między podmiotami, stworzyć dla nich strukturę hierarchiczną i zbudować drzewo minimalnej rozpiętości.

Niech  $Y_i(t)$  oznacza zmianę cechy  $i$  w punkcie czasu  $t$ .

$$\rho_{ij} = \frac{\langle Y_i Y_j \rangle - \langle Y_i \rangle \langle Y_j \rangle}{\sqrt{(\langle Y_i^2 \rangle - \langle Y_i \rangle^2)(\langle Y_j^2 \rangle - \langle Y_j \rangle^2)}}$$

Zatem odległość między elementami układu można zdefiniować jako odległość  $d(i,j)$ :

$$d_{ij} = \sqrt{2(1 - \rho_{ij})}$$

która spełnia aksjomaty metryki euklidesowej, analogicznie jak długości boków w trójkącie.

Dla przestrzeni metrycznej, w której połączonych ze sobą jest  $n$ -elementów, subdominująca ultrametryka może być otrzymana za pomocą drzewa minimalnej rozpiętości (powiązań o minimalnej rozpiętości – minimal spanning tree MST). Dla ważonego grafu złożonego z  $n$  wierzchołków, MST jest drzewem o  $n-1$  krawędziach, które minimalizuje łączną odległość. Subdominująca ultrametryczna przestrzeń powiązana z przestrzenią metryczną wprowadza dobrze zdefiniowany układ topologiczny, z którym jednoznacznie wiąże się zindeksowana hierarchia rozważanych  $n$  obiektów.

### 3.9 Pomiar złożoności

Jak już to zostało wspomniane, układy złożone wymagają równoległego opisu na kilku poziomach organizacji struktury. Choć teoretycznie możliwe, wyprowadzenie opisu takich układów z wiedzy o budowie i oddziaływaniach pomiędzy pojedynczymi ich elementami nie jest wykonalne w praktyce. Dzieje się tak z powodu nieznaności warunków początkowych, co w przypadku zjawisk nieliniowych wyklucza dokładne prześledzenie ewolucji układu w czasie (tzw. czułość na warunki początkowe), a wielka liczba stopni

swobody układu wyklucza użycie współcześnie dostępnej mocy obliczeniowej. Te problemy sprawiają, że opis układów złożonych jest albo jedynie jakościowy, albo do badań trzeba wykorzystać te ich aspekty, które poddają się łatwo analizie. Dużą rolę w badaniu układów złożonych odgrywa analiza danych empirycznych. Dane te w postaci sygnałów (serii czasowych) są dostępne po wyborze odpowiednich zmiennych, które poddają się pomiarom. Własności układów są w związku z tym określane w sposób pośredni poprzez własności tych sygnałów. Złożoność układów przy tym może, choć nie musi, oznaczać złożoności sygnałów. Pojawia się tu w ogóle problem, jak zdefiniować stopień złożoności ciągu liczb. W przeciwieństwie do definicji złożoności układów, złożoność sekwencji liczbowych można definiować w sposób ilościowy i ścisły. Niestety, można to zrobić na wiele sposobów, bez jednej uniwersalnej definicji. Pierwszą z zaproponowanych miar złożoności była złożoność algorytmiczna, która opierała się na długości minimalnego algorytmu, który daną sekwencję może odtworzyć. Jej wadą było przypisywanie nieskończonej złożoności sekwencjom przypadkowym, w związku z czym konieczne było wprowadzenie innych wielkości, które sobie z tym potrafiły poradzić: np. złożoności efektywnej czy głębokości logicznej. Niestety, wielkości te zazwyczaj są trudne do wyliczenia i przez to ich stosowalność bardzo ograniczona. O wiele prościej wylicza się miary oparte na pojęciu entropii informacyjnej:

$$S = - \sum p_i \ln p_i$$

gdzie  $p_i$  oznacza, że prawdopodobieństwo, że sygnał przyjmie wartość  $i$  z  $N$  dozwolonych wartości (możliwe jest też zdefiniowanie entropii w sposób ciągły). I choć entropia jest użyteczną miarą, to – podobnie jak złożoność algorytmiczna – może osiągać najwyższą wartość dla sygnałów, które są wynikiem serii losowej. Ważną zaletą entropii i metod określania złożoności ją wykorzystujących (np. złożoności efektywnej) jest to, że może być stosowana nie tylko do sygnałów, ale też do układów. Można wprowadzić rozróżnienie między "informacjami opisującymi" a "informacjami identyfikującymi" wyrażonych właśnie przedstawionym wzorem Shannona. Kwestie związane z teorią informacji były postawą na równi z inspiracją bioinżynierią cybernetyki społecznej<sup>7</sup>. W tym wypadku zamiast prawdopodobieństwa wartości sygnału używa się prawdopodobieństwa, że układ jest w danym stanie.

Z faktu, że obiekty fraktalne mogą być intuicyjnie odbierane jako obiekty o dużej złożoności strukturalnej, wielkości charakteryzujące fraktale mogą też służyć jako ilościowe miary złożoności, wykorzystywane zarówno do obiektów, jak i sygnałów. Takimi wielkościami są wymiary fraktalne i wykładniki Hursta. W przypadku multifraktali lub sygnałów multifrakalnych, które są splotami wielu różnych jednorodnych fraktali, a przez to bardziej złożone od jednorodnych fraktali, miarą ich złożoności może być też tzw. spektrum multifrakalne. Miary te stosuje się powszechnie w analizie danych pochodzących z układów przyrodniczych. W przypadku socjologii wydaje się, że mają one jednak mniejsze zastosowanie. Złożoność układu można też opisywać poprzez identyfikację efektów kolektywnych rozumianych jako redukcja efektywnej liczby stopni swobody układu. Na poziomie analizy sygnałów empirycznych można to zrobić, badając korelacje pomiędzy sygnałami związanymi z różnymi stopniami swobody. Im więcej sygnałów jest od siebie zależnych i im silniej są one zależne, tym kolektywność jest większa. W tym kontekście takimi miarami mogą być omówione w dalszej części książki współczynniki korelacji Pearsona i Spearmana oraz informacja wzajemna, wraz z ich rozkładami.

---

<sup>7</sup> <http://socjocybernetyka.pl/>

## 4 Nauka, technika i zmiana społeczna w kontekście informatyzacji

Pojawienie się technologii cyfrowej, a uprzednio elektronicznej i elektrycznej, w znaczącym stopniu wpłynęło na obraz naszego świata. Rewolucja technologiczna, skoncentrowana wokół technik informacyjnych tworzących nowe poziomy komunikacji, a co za tym idzie — współzależności i współbieżności, przeobraziła z dużym przyspieszeniem materialne podstawy społeczeństwa. Ten rozdział jest poświęcony specyfice życia społecznego w świecie wirtualny w kontekście wykorzystania danych (zwłaszcza Big Data) coraz bardziej i w większej ilości dostępnych. W konsekwencji gospodarczego globalnego ujednoczenia oraz rozwoju nowego medium – Internetu, obserwujemy społeczeństwo rządzące się innymi prawami. Nie oznacza to jednak, że technika determinuje społeczeństwa. Również społeczeństwo nie pisze scenariusza zmiany technologicznej, ponieważ w okresie odkryć naukowych, technicznych innowacji i społecznych zastosowań interweniuje wiele czynników, włączając w to indywidualną przedsiębiorczość i innowacyjność. Dlatego też efekt końcowy jest wynikiem złożonego procesu dialektycznej interakcji. W badaniach naukowych sprawa wygląda jednak prościej, gdyż uznaje się paradygmat, że technika jest społeczeństwem, a społeczeństwo nie może być rozumiane lub przedstawiane bez swych technicznych narzędzi. Postaram się w dalszej części rozdziału pokazać najważniejsze zjawiska społeczno-kulturowe współistniejące z pojawieniem się „techniki XXI-go wieku”, z perspektywy dostępnej wiedzy socjologicznej (Castells, 2000), jak również własnych obserwacji człowieka nie tylko wykorzystującego nowoczesne technologie, ale też pracującego w *High-Tech*. Przedstawię również wycinki z przemiany procesu edukacji, nauki i kultury wynikające z digitalizacji współczesnego świata.

### 4.1 Postęp technologiczny naszej epoki

W ciągu 250 lat, od 1750 do 2000, technologia stworzyła nową, światową cywilizację. Wynalazki techniczne stały się czymś powszechnym od Zachodu po Wschód. Transformacja ta dokonała się przez radykalną zmianę w znaczeniu wiedzy. Za Peterem Drucknerem (Druckner, 1969) podzieliłem ten okres na trzy etapy:

1) Rewolucję przemysłową: Pierwsze 150 lat (mniej więcej od powstania silnika parowego Watta), kiedy to rzemiosło przekształcało się w technologię, doświadczenie w wiedzę, wtajemniczenie w metodologię, a działanie w stosowanie wiedzy. Powstawały też uczelnie techniczne (pierwszy dyplom inżyniera we Francji ok. 1800) i coraz mniej cenione było *Allgemeine Bildung*, tylko wysoka specjalizacja. Przykładem może być główny bohater powieści Marka Twaina *Jankes na dworze Króla Artura* (1889), nie znał łaciny, nie czytał Szekspira, ale potrafił budować wszelkie maszyny, w tym generator prądu. Uprzemysłowienie odbywało się nierównomiernie. Zupełnie inaczej na chociażby niemieckim Śląsku i carskiej Rosji. Podstawą ekonomii wg Adama Smitha jest „Ziemia, kapitał i zasoby ludzkie”.

2) Rewolucję produktywności: Szybkie bogacenie się uprzemysłowionych społeczeństw poprzez większą efektywność pracy przez pierwsze 50 lat XX wieku. Poziom życia nawet w najbiedniejszych dzielnicach przemysłowych miast był wyższy niż na wsiach, z których pochodziła napływowa ludność (śmiertelność noworodków w favelas Rio de Janeiro była dwukrotnie niższa niż w otaczającej miasto prowincji). Dane pochodzące z produkcji, sprzedaży itd. podlegać zaczęły szczegółowej naukowej analizie. Taylor zapoczątkował w USA profesjonalne zarządzanie. Rozpoczęcie taśmowej produkcji (np. Ford T). Przemysł wciąż był motorem gospodarki, w tym wojennej (dwie wojny światowe).

3) Rewolucję w zarządzaniu: Menagerowie wprowadzają w życie wiedzę wyniesioną ze szkół. Proces podejmowania decyzji podlega naukowym wyznacznikom. Coraz większa automatyzacja produkcji. Dzięki temu mniej ludzi musi być związanych bezpośrednio z przemysłem, a może się zająć usługami. Wymiana i posiadanie informacji ma ogromne znaczenie na wszystkich poziomach działalności. Twierdzenie Adama Smitha w tych czasach nabiera nowego wymiaru — „Wiedza, kapitał i zasoby ludzkie”. Powstają coraz to

nowe technologie informacyjne w tym komputery i Internet, które nadają jeszcze szybsze tempo przemianom społecznym.

Właśnie powstanie powszechnych narzędzi ułatwiających komunikację i pozwalających na niemal natychmiastowy dostęp do informacji ma ogromny wpływ na społeczeństwo XXI wieku.

## 4.2 Rewolucja internetowa

Początków Internetu można doszukiwać się różnie m.in. w sieci komputerów stworzonej przez fizyków jądrowych z laboratorium w Los Alamos we wczesnych latach siedemdziesiątych dwudziestego wieku (Castells, 2000). Dzisiaj sieć obejmuje ok. miliarda użytkowników na całej Ziemi. Taki szybki rozrost mógł być możliwy dzięki wręcz wykładniczemu przyrostowi mocy obliczeniowej współczesnych komputerów (Prawo Moora głosi, że co roku moc obliczeniowa komputera za tą samą wartość bieżącą podwaja się) jak również dzięki wielu innym czynnikom, jak unitaryzacja kodowania danych (tu także fizycy tym razem z CERN-u mają swój wkład, tworząc język HTML, który stał się standardem). Komputeryzacja dotyczy w znacznie szerszym stopniu naszego życia (Capra 1998). Używając chociażby kart płatniczych czy rozmawiając przez telefon komórkowy korzystamy z nowoczesnych rozwiązań do przekazywania informacji w trybie niemal natychmiastowym. Mogąc robić zakupy w Internecie, możemy rzadziej chodzić do *realnego* sklepu. Korzystając z komunikatorów internetowych (jak polskie Gadu-Gadu, czy ogólnoświatowe IRC) zmieniają się relacje międzyludzkie. Portale społecznościowe (jak polska Nasza-Klasa, czy międzynarodowy Facebook) pochłaniają niewyobrażalną jeszcze kilka lat wcześniej ilość czasu ich użytkownikom. Dzięki tym środkom możliwym stało się przesyłanie informacji między kontynentami i utrzymywanie kontaktów z ludźmi mieszkającymi w dowolnym miejscu na naszej planecie.



Rys. 19 Usiecienie współczesnego świata. Źródło: [www.facebook.com](http://www.facebook.com)

Ogólnodostępność informacji (wyszukiwarka internetowa Google, wolna encyklopedia Wikipedia, poczta elektroniczna, czy ogromna ilość portali tematycznych), telefonia komórkowa sprawiły, że ludzie stali się niecierpliwi, liczy się dla nich głównie „Tu i teraz”. Nie jeden internauta zdenerwował się obserwując na swojej przeglądarce internetowej *błąd 404* (świadczący o tym, że strona pod danym adresem już nie istnieje, bądź nie ma do niej w tej chwili dostępu). W związku z tym często razem z adresem internetowym w publikacjach dodaje się datę, kiedy ostatnio był on aktywny.

Również bardzo źle widziane jest zatrzymywanie wymiany informacji. Pierwsi hakerzy nie włamywali się dla zysków, tylko po to, żeby ujawnić to, co ktoś chciał utajnić. Równie mocno krytykowane jest prawo autorskie i patentowe. Odpowiedzią na to są społeczności wolnego oprogramowania (szteandarowymi projektami są systemy operacyjne typu Linux na licencji GNU) czy wymiany wirtualnych dzieł artystycznych (muzyczny MySpace, filmowy YouTube).

Wystarczyło umożliwić internautom, żeby mogli między sobą kupować i sprzedawać przedmioty na akcjach internetowych (eBuy, czy polski odpowiednik — Allegro), żeby zaobserwować jak tworzą się relacje między

licytującymi. Dzięki Internetowi ludzie mogą posmakować życia w *wirtualnym* świecie odwzorowującym w wielu szczegółach *realny* świat (jak w grze „Second Life”, gdzie użytkownicy mogą tworzyć awatary-postacie, którymi się posługują udając czynności wykonywane w rzeczywistości). Daje to możliwość bycia innym, choć przez chwilę. W Internecie można być anonimowym, co pozwala na odgrywanie takich ról, na które ma się w tej chwili ochotę, często okłamując innych użytkowników. Z drugiej strony Internet pozwala na uzewnętrznienie, które inaczej mogłoby zostać wyhamowane. Świadczą o tym liczne blogi (pamiętniki internetowe) i wypowiedzi w postaci postów na forach internetowych.

Zaangażowanie w taką działalność oraz gry internetowe często przenosi się na życie rodzinne, pracę i stosunki z przyjaciółmi. Coraz częściej słyszy się o uzależnieniu od Internetu i innych zaburzeniach życia społecznego. Portal Nasza-klasa jest w wielu firmach zablokowany, bo pracownicy spędzali na nim zbyt wiele czasu (podobna sytuacja ma miejsce w krajach zachodnich z Facebookiem). Młode pokolenie wyrastające z nowymi technologiami staje w konflikcie ze starszym pokoleniem, dla którego wszystkie te nowości nie są tak automatycznie przyswajane i opanowywane. Z drugiej strony brak dostępu do komputera jest traktowany jako wyznacznik marginesu społecznego, a powszechna komputeryzacja szkół ma na celu wyrównywanie szans młodzieży. W związku z tym, że dostęp do sieci jest powszechny możemy się spotkać z przeniesieniem przemocy z ulic do przestrzeni Internetu. Wystarczy przeczytać komentarze do jakiegoś poczytniejszego artykułu z serwisu internetowego, aby przekonać się, jak agresywne, bądź na jak niskim poziomie zdarzają się wypowiedzi. Z podobnym problemem stykają się twórcy wolnej encyklopedii Wikipedia, która bazuje na możliwości edycji haseł przez wszystkich internautów, którzy wytworzyli cały zestaw sposobów walki z różnego rodzaju wandalizmem. O młodych Polakach, którzy spędzali najmłodsze lata przed gramami komputerowymi, surfującymi po Internecie, a jednocześnie nieczytającymi książek, nieuprawiającymi sportu, mówi się wartościująco ujemnie „Dzieci Neostrady” (Neostrada – czyli Internet dostarczany przez Telekomunikację Polską).

Z powyższych przykładów łatwo zauważyć jak skomplikowany i wielowymiarowy jest obraz współczesnego społeczeństwa informacyjnego. Tak ogromna ilość innowacji (będąca kontynuacją cywilizacyjnej przemiany rozpoczętej rewolucją przemysłową) ma rację bytu dzięki ukierunkowaniu działań społecznych na rozwój. Właśnie orientacja na naukę i zdobywanie wiedzy pozwoliła nam najpierw zaprojektować urządzenia i systemy, w ramach których one pracują. Promowanie pomysłowości stało się kołem zamachowym rozwoju technik wirtualnych. Rynek technologii informacyjnych jest o tyle szczególnym rynkiem, w którym kapitał schodzi na drugi plan, a na pierwszy pomysł, czyli pewien rodzaj wiedzy. Niejednokrotnie można się spotkać z biografiami dziś znanych i bogatych ludzi, którzy zaczęli od pomysłu gdzieś na amerykańskich uniwersytetach. Na polskiej scenie wystarczy tylko wspomnieć o trójce wrocławskich studentów, którzy stworzyli portal Nasza-klasa i czerpią z niego obecnie spore (jak na nasze warunki) zyski.

Nigdy jeszcze w historii ludzkości technologia nie miała takiego wpływu na społeczeństwo. W obrotach danymi w formie elektronicznej to technologia stanowi prawo, to od niej tylko zależy, co użytkownik zrobić może albo nie. Jeszcze nie tak dawno temu, wszyscy cieszyliśmy się z wolności, jaką przyniosła informatyzacja społeczeństwa: swoboda wymiany danych, dzielenia się twórczością, możliwością dostosowania sprzętu do swoich potrzeb. Dziś coraz częściej spotykamy się z próbami narzucenia przez korporacje sposobu korzystania z urządzeń i obrotu danymi za pomocą nowych technologii. Technologia stała się narzędziem kontroli, narzucenia de facto prawa społeczeństwu, które nigdy (jako suwereni) tego prawa nie wprowadzili. Uderzyło to również w rynek, ograniczyło konkurencyjność i rozwój nowych technologii — stąd paląca potrzeba spojrzenia na zjawiska takie jak patenty na oprogramowanie, czy zmuszanie użytkownika do używania oprogramowania wskazanego przez producenta sprzętu. Owe prawa patentowe są sporym zagrożeniem dla młodych rynków informatycznych, jakim jest m.in. polski rynek. Pewne reguły wypracowane przez amerykańskich prawników, wydające się dla nas zupełnie bezsensowne, trzymają w ryzach całą gałąź gospodarki. Podstawową kwestią są prawa patentowe do często banalnych procedur, jak zakup w sklepie internetowym przez pojedyncze kliknięcie (Szymański, 2008). Jakby uczciwy

internetowy przedsiębiorca chciał rozwiązać problem zakupu inaczej, przez podwójne kliknięcie, to niestety patent na podwójne kliknięcie też jest już zarejestrowany. Patenty są założone na tak proste rozwiązania, które pierwsze przychodzą na myśl przy projektowaniu strony WWW, a ich ominięcie jest często niemożliwe. Korporacje posiadające patenty szantażują się wzajemnie tym, że każda ze stron nielegalnie korzysta z czyichś rozwiązań. Na najgorszej pozycji stoją małe firmy, które pozwane do sądu za naruszenie patentu, mają niewielkie szanse na przetrwanie (nawet jak uda im się dowieść, że do wykorzystania opatentowanego rozwiązania doszły niezależnie, to i tak koszty sądowe mogą okazać się nie do udźwignięcia). Póki co w Europie nie obowiązują w pełni amerykańskie patenty, ale nie wiadomo jak długo UE będzie opierać się naciskom amerykańskiego lobby.

### 4.3 Rzeczywistość Wirtualna (VR)

Wirtualna rzeczywistość (Zacher, 2005), czyli to co wygląda realnie, jest odczuwane realnie, działa lub zachowuje się realistycznie, aczkolwiek w rzeczywistości realne nie jest. Sztuczną rzeczywistość tworzy się przy wykorzystaniu technologii informatycznej (Benkler, 2006). Polega na multimedialnym kreowaniu komputerowej wizji przedmiotów, przestrzeni i zdarzeń nieistniejących w naturze. Uzyskuje się to poprzez generowanie obrazów, efektów akustycznych, a nawet zapachowych. Podstawowym przykładem są gry komputerowe, w których dzieci uczą się wielu użytecznych umiejętności jak zdolność kojarzenia faktów, refleks, rozwiązywanie zagadek oraz kształcą pamięć wzrokową. Programy edukacyjne pomagają w uatrakcyjnianiu procesu nauczania. Trudno sobie obecnie wyobrazić współczesną technikę bez symulacji jazdy, czy architektonicznych programów wspierających budownictwo. Powstanie VR łączy się z wejściem w trzeci etap rozwoju cywilizacyjnego, kiedy skończyła się modernistyczna rewolucja produktywności. VR otrzymana przez nowoczesne technologie informatyczne pozwoliła na przynajmniej częściową realizację pomysłów takich ludzi jak Stanisław Lem. W eseju *Summa technologiae* w roku 1964 napisał, że wejście do rzeczywistości wirtualnej jest jakby podróżą w jedną stronę, z której nie ma powrotu. Bo wychodząc z "maszyny fantomatycznej" (jak to nazwał Lem), nigdy nie możemy być pewni, czy to wyjście nie jest tylko elementem symulacji. Może być przecież tak, że kiedy już będziemy wychodzić z maszyny, odłączać elektrody, zamykać oprogramowanie itd., wszystko to dalej będziemy robić tylko na niby — w wirtualnej rzeczywistości.

Abstrahując nawet od tego paradoksu, ale pozostając blisko *science fiction*, warto wspomnieć o kulturze *cyberpunk* dla której VR jest niezmiernie istotnym elementem. Obecnie łatwiej niż kiedyś (za pomocą choćby książek fantastycznych) oderwać się od otaczającej rzeczywistości. W mgnieniu oka można z szeregowego pracownika biurowego stać się kimś o niesamowitych zdolnościach, włączając odpowiednią grę. Cyberpunk koncentruje się najczęściej na mrocznej stronie wpływu technologii na człowieka. Istotnym jest połączenie człowieka z maszyną, od cybernetyki, przez całkowicie sztuczne ciała, po zwykłe uzależnienie człowieka od techniki w podstawowych czynnościach dnia powszedniego. Jest też postmodernistycznym rajem, (Fisher, 1996) gdzie np. według Wiliama Gibsona autora *Neuromancera*, w celu zwiększenia możliwości ludzkiego organizmu stosowane są cybernetyczne rozszerzenia (tzw. *cyberdecks*) zdolne m.in. wymieniać informacje ze światową siecią komputerową. Bohater poruszający się w ten sposób w sieci, znajdował się w cyberprzestrzeni. Wizja Gibsona stała się inspiracją dla innych autorów oraz wielu sieciowych społeczności. Choć przykłady podane przeze mnie są technologicznymi absurdami jak na dzień dzisiejszy, ale środowisko odbiorców cyberkultury nie ogranicza się już tylko do wąskiego grona wielbicieli cyberpunka. Świadczy o tym choćby komercyjny sukces filmu *Matrix* braci Wachowskich. Cyberkultura obecnie jest utożsamiana z Internetem i tak właśnie odnosi swoje sukcesy. Nie mogłyby się jednak rozwinąć podobne formy kultury, gdyby nie postęp technologiczny, który jest związany z kształceniem.

### 4.4 Społeczne sieci Internetu a nauka

W obecnych czasach lokalne sieci więzi społecznych ulegają przeobrażeniu. Od niedawna spotykamy się ze zjawiskiem „małego świata”. Jak kiedyś ludzie utrzymywali kontakt tylko ze swoimi geograficznymi sąsiadami, tak teraz dochodzą relacje z ludźmi zamieszkującymi odległe miejscowości. Internet jest takim

medium, które umożliwia podtrzymywanie znajomości mimo znacznych odległości przestrzennych między ludźmi. Ponadto poprzez Internet coraz częściej nawiązywane są nowe kontakty, które czasami przenoszone są również na świat *realny*. Badania Milgrama z lat 60. XX-go wieku pokazujące, że średnia ilość pośredniczących ludzi między dwiema osobami wynosi 7, mogą już być trochę nieaktualne. Owa droga może być już krótsza, co dobrze opisuje określenie współczesnego świata jako „globalnej wioski”. Na przykład dla społeczności internetowej jednej z najpopularniejszych gier sieciowych (*Massive Multiplayer Online Role Playing Game* (MMORPG)) stopień oddzielenia szacuje się na 4,7 (Grabowski & Kosinski, 2010).

Pojęcie lokalności w sensie geograficznym zostało zdecydowanie wykluczone. Przyjrzyjmy się dokładniej środowisku naukowemu fizyków, aby zobaczyć, jak zmieniła się struktura tej społeczności. Przypomnę, że na samym początku szybkiego rozwoju fizyki, czyli w pierwszych dekadach XX-go wieku fizycy skupiali się wokół „szkół” i pewnych osobowości jak np. „szkoła kopenhaska” w mechanice kwantowej przy Nielsie Bohrze czy nasza „szkoła lwowska” przy Stefanie Banachu w przestrzeniach metrycznych. Publikowane prace często w różnych językach nie trafiały w ogóle albo z dużym opóźnieniem do naukowców z różnych ośrodków akademickich. Spotkania były rzadkością ze względu na długie podróże, które musieliby odbyć naukowcy, jak również przez problemy z przekraczaniem granic w podzielonej konfliktami Europie. Jeżeli teraz chcielibyśmy badać powiązania między fizykami, zauważymy znaczne rozproszenie (Hirsch 2005). Przykładem mogą być publikacje, a właściwie ich autorzy. Budując sieć powiązań wystarczy połączyć ze sobą fizyków, którzy napisali wspólną pracę. Tak zbudowana sieć wykazuje zdecydowanie cechy „małego świata”. Inną metodą pokazującą taką zależność jest wyznaczenie, kto kogo cytuje, pisząc artykuł. Taka struktura nie charakteryzuje tylko fizyków. Dostęp do dorobku naukowego jest teraz zdecydowanie szybszy. Na serwerze arXiv.org umieszczone są praktycznie wszystkie wyniki badań (nawet te, które nie zostały jeszcze zaakceptowane przez wydawnictwa naukowe). Można w kilka minut dowiedzieć się, kto na całym Globie jeszcze pracuje nad podobnym problemem i jaki jest postęp w pracach. Konferencje stały się codziennością i dzięki liniom lotniczym mobilność fizyków jest ogromna. Współczesny naukowiec może być przywiązany do jakiejś jednostki badawczej a pracować w różnych miejscach na świecie. Listy elektroniczne docierają niezależnie gdzie się on w tej chwili znajduje. Mamy tu do czynienia z nowym znaczeniem czasu i miejsca. Internet jest właśnie taką przestrzenią, w której można wymieniać informacje niezależnie od godziny i miejscowości. To sprawia, że praca może być wykonywana nie w wyznaczonych z góry ramach czasowo-przestrzennych, ale w wybranym momencie i miejscu. W każdej dziedzinie nauki i techniki, gdzie istotna jest bardzo wysoka specjalizacja, sieci społecznych zależności wyglądają podobnie. Obecnie praktycznie nie da się uprawiać nauki bez korzystania z Internetu. Co prawda pozostali jeszcze profesorowie starej daty, którzy nie dali się przekonać nowościom technologicznym, ale z chwilą ich odejścia na emeryturę chyba niemożliwym będzie awans na szczeblach kariery naukowej bez wspierania się Internetem.

#### 4.5 Społeczeństwo informacyjne i e-styl życia

Społeczeństwo informacyjne, to jest takie, które nie tylko posiada rozwinięte środki przekazu, lecz przekazywanie informacji jest podstawą tworzenia dochodu narodowego i dostarcza źródeł utrzymania większości społeczeństwa (Salimowicz, 1999). Wielość sposobów wykorzystania Internetu nadaje mu szczególnie ważne znaczenie w naszym społeczeństwie. Coraz powszechniejsza jest praca zdalna przez Internet. Dlatego też wielu badaczy podjęło problem zachowań użytkowników Internetu. Z drugiej strony Internet daje nam odpowiedzi na temat naszego społeczeństwa (np. na podstawie częstotliwości zapytań w wyszukiwarce Google można wnioskować, jak popularne są pewne zagadnienia, postacie itp. w danej chwili). W wyniku większości przeprowadzonych badań wynika, że zmiany społeczne, jakie nastąpiły, jakkolwiek są istotne, to jednak nie musimy się poki co obawiać, że następnego dnia obudzimy się w świecie cyberpunkowym. Internet jest przestrzenią wolności, ale większość użytkowników traktuje ten obszar swojego życia z dystansem. Internet nie jest miejscem unikania odpowiedzialności, ale tylko jej ograniczenia. Zwracając jednak większą uwagę na istotę zmiany społecznej na podstawie badania na

krakowskich studentach, zauważamy, że Internet jest teraz głównym miejscem spędzania wolnego czasu (80% respondentów wybrało taką opcję spędzania wolnego czasu). Przeciętny student spędza 20 godzin tygodniowo przed komputerem z czego 10 godzin z Internetem. Marian Niezgódka wyróżnia 4 kategorie zmiany społecznej spowodowanej upowszechnieniem Internetu:

- Ograniczenie kontaktów z ludźmi poznanymi w świecie rzeczywistym i związanymi z miejscem zamieszkania, pracą, szkołą itp.
- Zmiana form kontaktów, wykorzystanie nowoczesnych technologii – sms, telefon komórkowy, jak również komunikatory internetowe i poczta email.
- Zmiana przyzwyczajeń – zakupy w Internecie zamiast w sklepie osiedlowym.
- Zmiana form uczestnictwa w kulturze – oglądanie filmów na kinie domowym zamiast w zwykłym kinie.

Spółeczeństwo Internetu nie jest jednak równoznaczne ze społeczeństwem choćby polskim. Już demograficzna struktura dużo różni się od tej w świecie rzeczywistym. Tylko 27% internautów to kobiety. Poza tym występuje w tej społeczności duża nadreprezentacja uczniów i studentów. Również zamieszkanie na wsi i niskie wykształcenie ma negatywny wpływ na pojawienie się w środowisku internautów. Z tych samych badań wynika, że aż 80% populacji nie ma w domu dostępu do Internetu. Co prawda obecnie sytuacja się poprawiła w Polsce, ale nie jest jeszcze tak dobra, jak w krajach zachodnich.

#### 4.6 Podsumowanie zmian w nauce i technologiach wirtualnych

Generalizując, nie da się nie zauważyć, jak wiele zmieniło się w ostatnich latach zwłaszcza w poprzedniej dekadzie. Informacja stała się dobrem niematerialnym, ale niezwykle cennym dla gospodarki kultury i polityki. Wykorzystując dane w oparciu o Internet, trzeba brać pod uwagę wszystkie obciążenia związane ze specyfiką aktywności społecznej w tym medium. Należy zdawać sobie sprawę z etycznych i prawnych konsekwencji wynikających z prowadzonych badań (jak choćby przytoczony przeze mnie casus patentów na oprogramowanie). Wiele spraw w urzędach, dziekanatach i innych instytucjach można załatwić nie wychodząc z domu. W Wielkiej Brytanii głosowanie, czyli podstawa demokracji, odbywa się na życzenie obywatela zdalnie bez konieczności przychodzenia do lokalu wyborczego. Nauka również zmieniła swoje oblicze, wiedza stała się coraz bardziej powszechna, jak również jest przekazywana w coraz przyjaźniejszy sposób. Zmieniły się relacje uczeń-nauczyciel. Upowszechniła się nauka na odległość jeszcze bardziej komplikując relacje lokalne. Platformy edukacyjne zindywidualizowały proces nauczania. Naukowcy musieli również przyzwycząć się do nowych warunków mobilności, które przyniosły niewiarygodny przyrost osiągnięć zwłaszcza w naukach ścisłych. Niestety dla tych, którzy z różnych przyczyn nie opanowali obsługi nowoczesnych technologii, ten brak umiejętności staje się przyczyną wykluczenia. Dotyczyć to może również kadry naukowej opierającej się nowym formom komunikacji, jak e-mail. Nie uczestniczenie w wyścigu technologicznym powoduje zejście na boczny tor życia. Należy być ostrożnym w entuzjastycznym przekonaniu o zlikwidowaniu barier politycznych i ekonomicznych między biednymi i bogatymi państwami. Otworzenie się nowych technologicznych możliwości spowodowało również pojawienie się zagrożeń (pojawiających się w cyberpunku). Mimo wszystko warto jest wykorzystywać zdobycze techniki i dostosowywać do nich prawo i po części samych siebie. Dlatego też informatyzacja powinna dotknąć również metodologię nauk społecznych i to nie poprzez na przykład cyfrową obróbkę zebranych danych, czy przeniesienie znanych technik do Internetu, ale zmiana musi dotyczyć stworzenia nowych narzędzi badawczych.

## 5 Rodzaje modelowania wykorzystywanego w naukach społecznych obliczeniowych

Matematyczne modele, które traktują o zachowaniu zbiorowisk ludzkich, już od lat początku XX-go opisywane są równaniami różniczkowymi (deterministycznie). Dużym zainteresowaniem cieszą się również modele stochastyczne, gdzie opis probabilistyczny zjawisk w których przejścia między poszczególnymi stanami odbywają się na podstawie zadanych rozkładów prawdopodobieństwa. Jednak w nowszych rozważaniach pojawiają się częściej modele agentowe (ang. agent based models - ABM). Symulacje komputerowe zaczęły grać znaczącą rolę od kiedy ilość i zróżnicowanie interakcji społecznych we współczesnym świecie skomplikowało się, a zachowanie jednostek nie jest homogeniczne (modele sieciowe – empiryczne i teoretyczne). Poza samymi technikami matematycznymi w formułowaniu i rozwiązywaniu problemów, opisane tu metody różnią się metodami mapowania współrzędnych przestrzennych (geograficznych) na układ dla którego da się coś policzyć, czy coś wysymulować. Z grubsza możemy je podzielić ze względu na wiele czynników/

Technikę matematyczną:

- równania różniczkowe i różnicowe (deterministyczne);
- modele agentowe, automaty komórkowe itp. (stochastyczne).

typ zmiennych opisujących agenta/zmienną:

- pojedyncza cecha przyjmująca stany dyskretne - np. różne rodzaje modelu wyborcy
- pojedyncza cecha ciągła - np. modele ograniczonego zaufania
- kilka cech tego samego typu ciągłych lub dyskretnych - np. wektor zmiennych dyskretnych w modelu Axelroda (Axelrod, 1997)
- kilka cech różnego typu - np. modele wpływu społecznego

typ aktualizacji stanu układu:

- ciągłe
- dyskretne
- losowa sekwencyjna - agenci są wybierani losowo, a ich stan aktualizowany pojedynczo
- synchroniczna - jak w automatach komórkowych - wszyscy agenci są aktualizowani jednocześnie

reguły oddziaływania:

- siły fizyczne jak np. nacisk sąsiadów i tłumy w opisie procesu ewakuacji
- hamiltonian – czyli funkcja energii i zasada podążania układu w stronę minimalnej energii
- reguły dynamiczne zadane w postaci algorytmicznej lub za pomocą macierzy przejścia

Oczywiście każdy z powyższych modeli może być rozpatrywany na dowolnej topologii. Najczęściej spotykanymi w symulacjach dynamiki opinii są:

- sieć regularna - zwykle kwadratowa np. model Axelroda (Axelrod, 1997), Schellinga, wyborcy

- sieć złożona - np. Barabasięgo-Alberta (bezskałowa), Erdosa Renyiego (prosta), Wattsa-Strogatza (małego świata)
- graf zupełny
- graf losowy
- sieć rzeczywista

Mimo wszystko, to dzięki rozwiązaniom równań różniczkowych nauczyliśmy się „rozumieć” procesy i dlatego, z edukacyjnego punktu widzenia, warto od nich zacząć. Rozróżnienie między modelami i zapis formalny zostanie zaprezentowany na przykładzie zjawiska przenoszenia patogenu z osoby na osobę. Właśnie bowiem modelowanie epidemiologiczne przyniosło najwięcej praktycznych zastosowań w życiu codziennym ze wszystkich dziedzin modelowania procesów społecznych.

## 5.1 Modelowanie deterministyczne

Założmy, że populacja jest podzielona na trzy klasy: podatni (S), którzy mogą zarazić, zarażający (I), którzy mogą przekazywać chorobę i usunięci (R), który przeżyli chorobę i są zdrowi (z odpornością na patogen) lub zostali wyłączeni (zostali odizolowani, wypisani ze szpitala lub po prostu zmarli). Jest jeden z najprostszych, a zarazem najchętniej stosowany model epidemiologiczny. Schemat przejść może być reprezentowany przez:



Rys. 20 Standardowy model ewolucji choroby

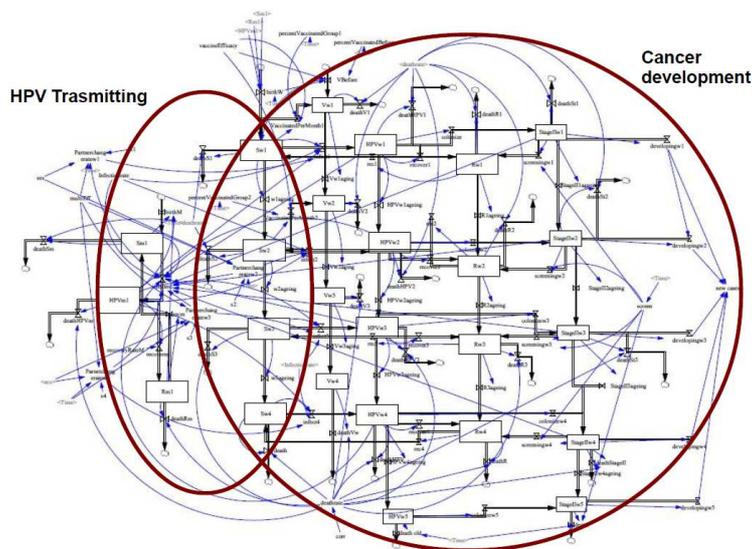
Opiszmy najważniejszy etap tego procesu, czyli transmisję patogenu. Rozważmy jedną podatną osobę (S) w populacji homogenicznej (jednorodnie wymieszanej). Indywidualne kontakty z innymi członkami społeczeństwa oznaczmy  $C$  (z jednostką  $\text{czas}^{-1}$ ) i proporcją  $I/N$  kontaktów z zarażonymi. Jeśli szansa przeniesienia zakażenia na jeden kontakt jest  $\beta$ , wtedy tempo przenoszenia patogenu na podatnych (S) jest  $\beta CI/N$ , a tempo stawania się przez nich zarażającymi (I) jest  $\beta CS/N$ . Częstość ( $C$ ) może być funkcją gęstości zaludnienia, co odzwierciedla fakt, że zakładamy jednorodne cechy wszystkich jednostek w subpopulacji. Można wyobrazić sobie sytuację, w których  $C$  mogłoby być proporcjonalne do  $N$  lub  $C$  może być po prostu stałą. Wyrażenia takie jak  $\beta SI$  i  $\beta SI/N$  często pojawiają się w literaturze. Żeby zmniejszyć liczbę parametrów zapiszmy:  $r = \beta C/N$ . Pełny system z uwzględnieniem pozostałych procesów może zostać zapisany w postaci układu równań różniczkowych:

$$\begin{aligned} \frac{dI}{dt} &= rSI - aI, \\ \frac{dS}{dt} &= -rSI, \\ \frac{dR}{dt} &= aI \end{aligned}$$

Wybór parametrów modelu oraz warunków początkowych  $r$ ,  $a$ ,  $S_0$ ,  $I_0$  powoduje, że czasami mamy do czynienia z epidemią, a czasami ona nie wybucha. Dla warunku początkowego na  $S$ , że  $S_0 < a/r$ , wtedy  $dI/dt < 0$  i w każdym momencie  $I_0 > I(t)$  oraz  $I$  zmierza do 0 z  $t$  dążącym od nieskończoności. Z drugiej strony, jeśli  $S_0 > a/r$ , wtedy  $I(t)$  wzrasta i epidemia wybucha. Mamy tu do czynienia z pewnym progami epidemiologicznym:  $R_0$ , czyli stopą reprodukcji epidemii (parametrem wykorzystywanym w codziennej pracy epidemiologa). Ta stopa jest kluczowa w kontroli epidemiologicznej, np. w szczepieniu populacji. Akcja jest potrzebna dla  $R_0 > 1$ , ponieważ poniżej tej wartości epidemia naturalnie wygasa. Właśnie

określenie wartości  $R_0$  jest standardowym zadaniem epidemiologa. Znając tę wartość rysują się różne ścieżki zatrzymania epidemii. W obecnych czasach nie tylko zdrowie społeczeństwa jest ważne, ale również koszt, jaki musi zostać poniesiony przez służbę zdrowia. Przykładem ekonomicznego zastosowania różniczkowego modelu SIR jest ustalenie procentowego pokrycia społeczeństwa szczepieniami. Wystarczy w modelu wyznaczyć minimalną część populacji szczepionej, dla jakiej  $R_0 < 1$  z odpowiednim zapasem, to otrzymamy wtedy informację, że nie trzeba kupować szczepionek dla wszystkich, tylko dla tej zoptymalizowanej subpopulacji. Na powyższym przykładzie pokazaliśmy, że równania różniczkowe cały czas dają nam wiele informacji i lepiej pozwalają zrozumieć zjawisko, ale w przypadku ogromnej ilości wiedzy, jaką posiadamy o społeczeństwie, nie dają się już tak łatwo stosować. Największym mankamentem jest potrzeba budowy kolejnego równania dla każdej wydzielonej podgrupy i tak w nieskończoność.

W związku z często bardzo złożonymi, nieliniowymi postaciami funkcji składających się na takie równania dla rzeczywistych problemów, nie daje się opracować rozwiązania analitycznego (możliwe jest to często tylko dla asymptotyk, bądź ustalania ograniczeń na wartości rozwiązania). Dlatego też stosuje się metody numeryczne, bądź (zdecydowanie częściej) rozwiązania różnicowe. Koncepcja równań różnicowych pozwala epidemiologom na budowanie intuicyjnych schematów z wykorzystaniem teorii systemów, co zostało bardzo szybko wykorzystane przez instytucje zdrowia publicznego krajów rozwiniętych, ponieważ poza umiejętnościami myślenia analitycznego, nie wymaga to znajomości twardej matematyki. Ponadto epidemiolodzy mogą korzystać z gotowych rozwiązań (Brouwers, 2009) w postaci dedykowanego im oprogramowania (np. Vensim).

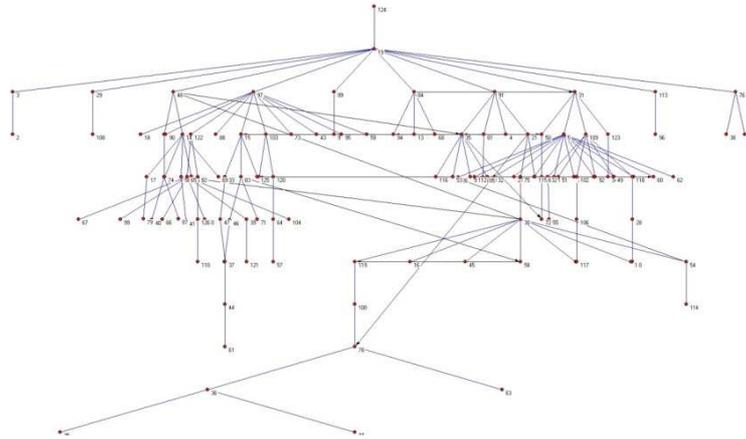


Rys. 21 Schemat potoków między kategoriami i zależności między nimi w programie Vensim na przykładzie chorób przenoszonych drogą płciową w Polsce

Innym przykładem zastosowań równań różniczkowych (dokładnie cząstkowych) jest geograficzne rozprzestrzenianie się chorób. I tak dynamika rozprzestrzeniania się choroby w przestrzeni geograficznej (dyfuzji) została wykorzystana do zmapowania epidemii dżumy, która od pojawienia się w portach Morza Śródziemnego na początku XIV w. trafiła do Polski 5 lat później, żeby zawędrować dalej na północ i wygasnąć.

## 5.2 Modelowanie stochastyczne

Podejściem konkurencyjnym do modelowania deterministycznego epidemii jest wykorzystanie metod probabilistycznych. W kwestii opisu procesu zmiana następuje w znaczeniu parametrów modelu, ponieważ zamiast stóp (ang. rates) mamy do czynienia z prawdopodobieństwami. W związku z tym np. stopa zarażenia (przepływ z kategorii S do I) jest reprezentowana czystym prawdopodobieństwem. W modelu stochastycznym można posługiwać się koncepcją procesu gałęziowego (Jacob, 2010).



Rys. 22 Gałęziowy proces transmisji patogenu od źródła infekcji (górze) aż do wygaśnięcia epidemii (dół). Poszczególne poziomy oznaczają pokolenia epidemiczne (od 0 na górze do 12 na dole)

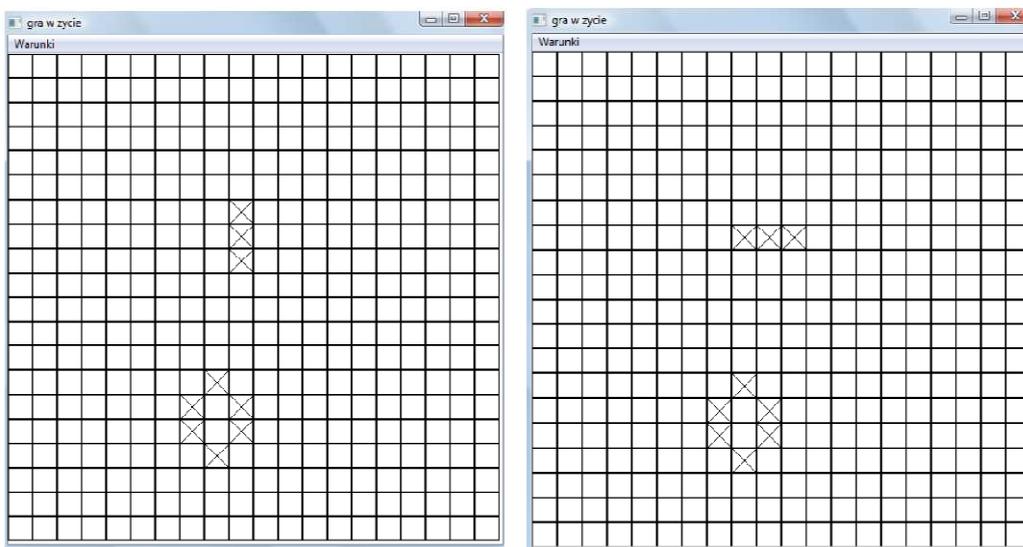
W modelu stochastycznym  $R_0$ , nie rozgranicza już w sposób definitywny wybuchu, bądź nie epidemii. W tym ujęciu dla  $R_0 < 1$  możliwy już jest start epidemii, czego nie obserwujemy w modelu deterministycznym.

Poza procesami gałęziowymi stosuje się bardzo często automaty komórkowe (siatka komórek ze stanami regułami przejść) w stochastycznym modelowaniu epidemiologicznym. Głównym celem tych badań jest analiza statystyczna modeli typu SIR (gdzie stan komórki odpowiada kategorii medycznej), gdzie epidemia rozprzestrzenia się poprzez sąsiadów zarażonych komórek zgodnie z ustalonymi probabilistycznymi regułami przejść. Poza standardowymi modelami na siatce (najwięcej zastosowań w epidemiologii roślin, gdzie układ grządek często przypomina taką właśnie strukturę) stosuje się często różnego rodzaju skróty odpowiadające wektorom zarażenia.

W związku z często bardzo skomplikowanymi procedurami statystycznymi związanymi np. z dopasowywaniem parametrów, epidemiolodzy korzystają z gotowych rozwiązań dostępnych w postaci aplikacji wspierających jak pakietów Epi w środowisku R.

### 5.3 Modelowaniu przy wykorzystaniu sieci społecznych (rzeczywistych i przypadkowych)

Wyobraźmy sobie ulokowanie osobników w jakimś układzie przestrzennym. Najprostszym może być np. siatka. Na takiej siatce komórki mogą zmieniać swoje stany oraz stany okolicznych komórek według pewnych reguł o nazywamy taki układ automatem komórkowym.



Rys. 23 Wizualizacja „Gry w życie 2005” autorstwa A. Jarynowskiego jak przykład automatu komórkowego

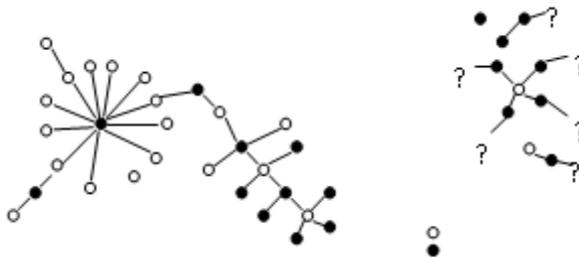
Reguły gry w mogą być opisane w dwuwymiarowej siatce z możliwościami stanów: żywa lub martwa komórka. Przyjmując sąsiedztwo typu von Neumanna (sąsiadami są 4 komórki – z góry, dołu lewej i prawej) obowiązują zasady:

1. Każda żywa komórka z mniej niż dwiema żywymi sąsiadami umiera, jakby spowodowane to było przez niedoludnienie.
2. Każda żywa komórka z więcej niż trzema żywymi sąsiadami umiera, jakby z przeludnienia.
3. Każda żywa komórka z dwoma lub trzema żywymi sąsiadami przechodzi do następnego pokolenia.
4. Wszelkie martwe komórki z dokładnie trzema żywymi sąsiadami stają się żywe.

Pierwsza generacja tworzą nową stosując powyższe zasady jednocześnie do każdej komórki a następnie do każdej martwej komórki.

Jeżeli zamiast na siatce jak w automacie komórkowym, umieścimy układ na sieci teoretycznej o konkretnych właściwościach, wtedy mamy do czynienia z siecią złożoną.

Rozwój technik komputerowych ułatwiających gromadzenie i analizowanie dużych ilości danych, umożliwił dokładne zbadanie struktury sieci rzeczywistych kontaktów (Grabowski & Rosinska, 2012) oraz procesów epidemiologicznych występujących na nich (Green & Bossomaier, 2000). Ponieważ główną drogą rozprzestrzeniania się chorób zakaźnych są kontakty interpersonalne, struktura sieci tych kontaktów ma istotny wpływ na dynamikę oraz rezultaty epidemii.

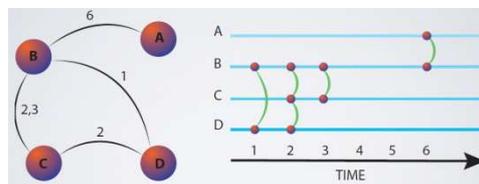


Rys. 24 Modelowe kontakty między w przypadku 2 subgroup (np. mężczyźni i kobiety). Znając topologię sieci można wnioskować o prognozie epidemicznej, szybkości rozprzestrzenienia się epidemii. Położenie jednostek w sieci pomaga również w ocenie ryzyka zarażenia

Cechą charakterystyczną większości sieci kontaktów jest potęgowa postać dystrybucji prawdopodobieństwa ilości sąsiadów i relatywnie duża liczba węzłów posiadających wiele połączeń (Rocha L. , 2011). Obecność takich węzłów istotnie wpływa na dynamikę procesu propagacji wzbudzeń w sieci, gdyż łączą one bardzo dużą liczbę innych węzłów, często znajdujących się przestrzennie daleko od siebie (w epidemiologii są to tzw. Superspreaders). Inną cechą charakterystyczną sieci rzeczywistych jest wysoka wartość współczynnika klasteryzacji - można wyróżnić grupy węzłów (klastery), pomiędzy którymi występuje duża liczba połączeń. Tak więc sieci rzeczywiste symuluje się często za pomocą sieci przypadkowych spełniających założenia sieci „małego świata” oraz sieci bezskalowych (Watts & Strogatz, 1998). Oddziaływania interpersonalne pomiędzy jednostkami należącymi do tej samej grupy są silniejsze niż oddziaływania pomiędzy jednostkami z różnych grup społecznych. Osoby należące do małych społeczności spędzają ze sobą więcej czasu niż osoby i średnia intensywność kontaktów pomiędzy członkami nimi jest większa. Z punktu widzenia rozprzestrzeniania się infekcji to właśnie kontakty w ramach najmniejszych grup są najbardziej efektywne jako drogi przenoszenia patogenu. Jednakże, nie można pominąć przypadkowych kontaktów pomiędzy osobami (np. w wyniku przemieszcza się z miejsca na miejsce – wektory). Niestety żaden układ (nawet szpitale) nie jest zupełnie oddzielony od świata zewnętrznego z tego względu istnieje niezerowe prawdopodobieństwo transmisji patogenu spoza badanej sieci. W celu rozróżnienia efektywności interakcji pomiędzy jednostkami znajdującymi się na różnych poziomach hierarchii, wprowadzone są różne

prawdopodobieństwa zainfekowania podatnej jednostki. Najczęściej najwyższy (najniższy) poziom interakcji interpersonalnych stanowią przypadkowe kontakty pomiędzy dowolnymi dwoma jednostkami należącymi do populacji generalnej. Ciekawe obserwacje zostały poczynione w badaniu procesu jednoczesnego rozprzestrzeniania się dwóch różnych infekcji. Z ewolucyjnego punktu widzenia wiemy, że wyniku przebycia choroby ludzie nabywają czasami odporności krzyżowej na patogeno podobnego typu oraz zazwyczaj tylko jeden typ dominuje w danym momencie. W przypadku konkurencji między patogenami, proces rozprzestrzeniania się epidemii może być traktowany jako proces rozprzestrzeniania się wzbudzeń w układzie, co jest potwierdzone przez powstawanie, charakterystycznych dla ośrodków aktywnych (Grabowski & Kosiński, 2006). Okazuje się, że w takim układzie dla pewnego zakresu parametrów kontrolnych obecność jednostek chorych na pierwszą chorobę uniemożliwia rozprzestrzenianie się drugiej choroby.

Tak więc, charakterystyka kontaktów w postaci czasowej (większość dotychczasowych analiz wykorzystywała sieci statyczne) i sieciowej struktury (kiedy i z kim) wiadomo że ma wpływ na rozprzestrzenianie się patogenów. Odtworzenie dynamicznej, czasowej siatki kontaktów jest podstawą do dalszej analizy (Holme, October 2012).



Rys. 25 Model dynamicznej sieci kontaktów (z uwzględnieniem momentu kiedy kontakt nastąpił oraz jak długo trwał)

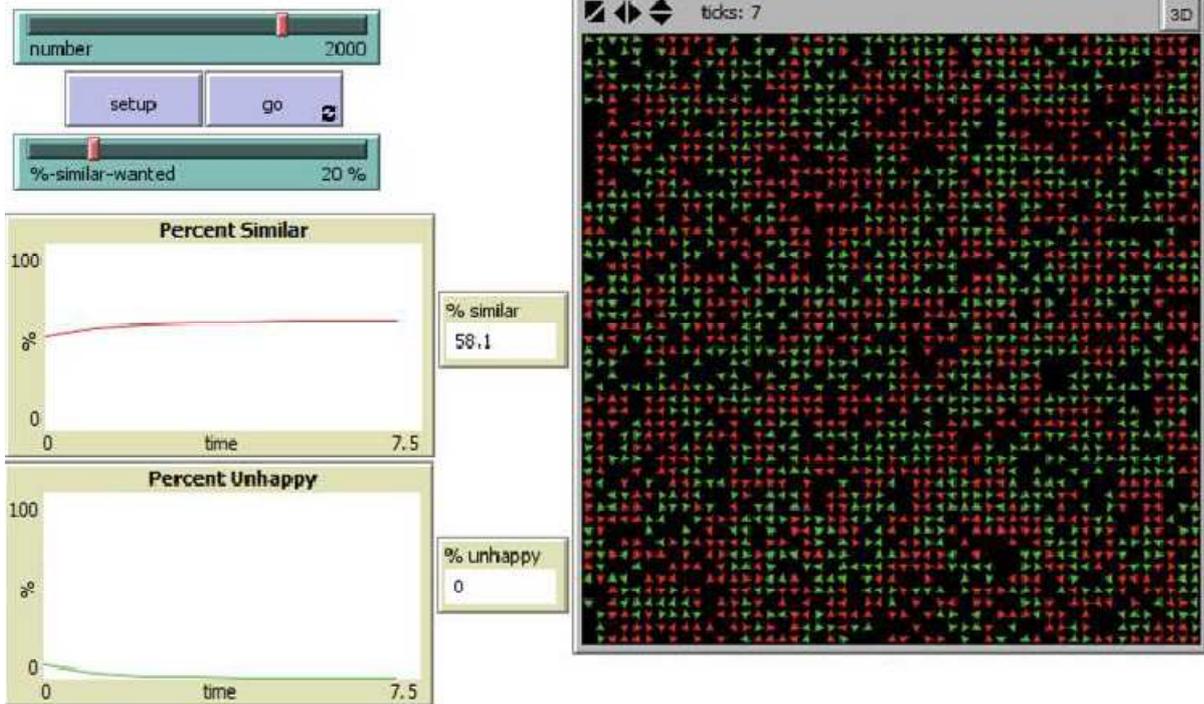
W ramach zaimplementowanej sieci (czy to rzeczywistej, czy złożonej-przypadkowej) uruchamia się modele zwane agentowymi bądź indywidualnymi, ponieważ atrybuty są zdeterminowane jednostkowo oraz czasami nawet przestrzennie. Dzięki rozwiązaniom inżynierskim typu GIS (ang. Geographic Information Systems), można lokować jednostki (ludzi, miasta itp.) na rzeczywistych mapach (Rocha, Liljeros, & Holme, 2011) i modelować transmisję patogenów z uwzględnieniem topologii (np. przeszkód naturalnych w postaci gór).

## 5.4 Modelowanie agentowe

Motywacją do symulacji i modelowania jest złożoność systemów społecznych nieliniowych, wielopoziomowych, dynamicznych interakcji. Modelowania różnego rodzaju funkcji społecznych, jest często to niewykonalne innymi metodami obliczeniowymi. Do modelowania agentowego potrzebna jest największa ilość informacji. Trzeba zdobyć wiedzę, w jaki sposób interakcje na poziomie mikro-makro zachodzą. Jednak jeśli posiadamy takie zbiory danych, to właśnie zaletą modelowania agentowego jest właśnie możliwość ich wykorzystania co daje okazję do kalibracji modeli, zwłaszcza, że obecna moc obliczeniowa komputerów pozwala na symulacji na dużą skalę.

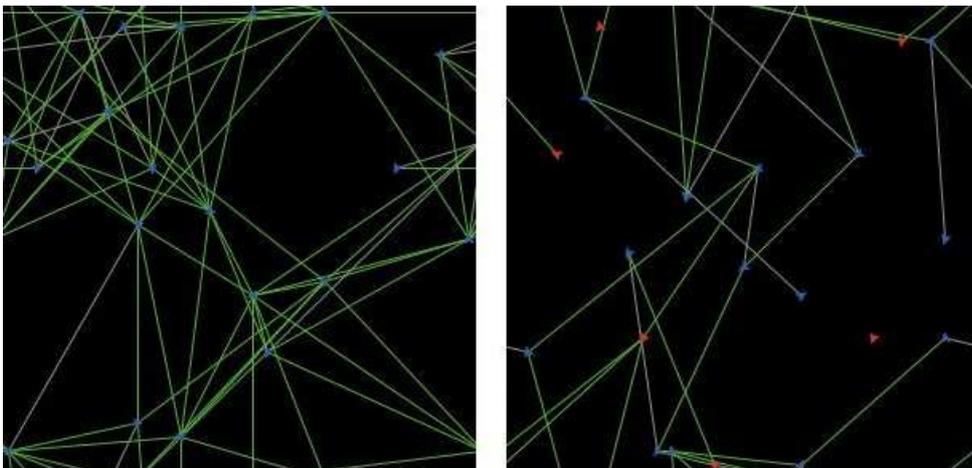
Dla poprawnego przeprowadzenia analiz symulacyjnych zachowań systemu podstawowe znaczenie mają następujące kwestie:

1. zdefiniowanie systemu podlegającego modelowaniu,
2. wybór adekwatnego narzędzia analizy – modelu systemu,
3. organizacja i pozyskanie zbioru danych wejściowych,
4. ustalenie warunków początkowych analizy i kalibrowanie parametrów modelu,
5. ustalenie wiązki scenariuszy symulacyjnych i eksperymentowanie z modelem,
6. przeprowadzenie obliczeń i analiza wrażliwości wyników modelowania na zmiany parametrów symulacji,
7. zasadna interpretacja wyników obliczeń wykonanych z pomocą modelu.



Rys. 26 Wizualizacja interfejsu oprogramowania Netlogo modelu Schellinga (Schelling, 1971)

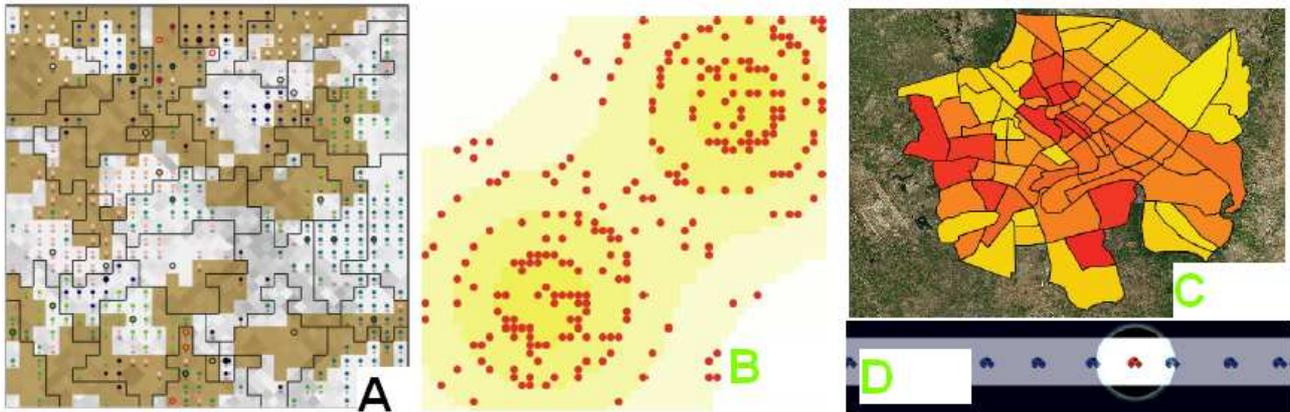
Jednym z pierwszych modeli agentowych o dużej popularności był model segregacji (Schelling, 1971). Agenci mogą być tylko dwóch typów (co ma odpowiadać białym oraz Afroamerykanom) i początkowo rozmieszczeni są losowo w układzie typu kratka (Rys. 26). Reguły są proste: agent jest niezadowolony jeżeli ma w otoczeniu większość innych niż on agentów i w każdym kroku symulacji jeden losowy niezadowolony, jest przesuwany do losowo wybranej wolnej komórki. W efekcie obserwowana jest segregacja nawet dla mały preferencji (reguł kiedy agent jest niezadowolony). Jednocześnie nie otrzymujemy ścisłej segregacji.



Rys. 27 Wizualizacja sieci modelu agentowego w oprogramowaniu Netlogo

Niestety, poziom skomplikowania reguł a modelowaniu agentowym pociąga za sobą wiele matematycznych konsekwencji. W mikrosymulacjach (modelowaniu stochastycznym, czy w automatach komórkowych) proces można opisać w kategoriach stochastycznego równania różniczkowego na prawdopodobieństwach przejścia (tak zwane równania fundamentalne lub typu Master). Nie istnieje takie ogólne równanie dla ABM. Agenci oddziałują w symulacjach a ta interakcja jest istotnym elementem w ABM. Standardowe modelowanie stochastyczne i deterministyczne zakłada, że układ ma własność Markowa: prawdopodobieństwo bycia konkretnym w stanie zależy tylko od stanu poprzedniego. W ABM, układ może

mieć pamięć i zdolność uczenie się. ABM jest bliższe empirii, natomiast standardowe typy modelowania są bardziej oparte na teorii, dzięki czemu są zapewne „piękniejsze” z matematycznego punktu widzenia oraz pozwalają na więcej generalizacji. ABM za to daje konkretne odpowiedzi na konkretne pytania/scenariusze, dzięki czemu jest to obecnie najczęściej stosowany typ modelu w opisie rzeczywistości społecznej, zwłaszcza że możliwa jest implementacja technik analizy sieci społecznych, czy rozprzestrzeniania się wzorca w przestrzeni.



Rys. 28 Przykłady symulacji agentowych w środowisku Netlogo. A – geosymulacje wojny (z wyglądu podobne do gier planszowych, jednak stosowane w praktyce przy prawdziwych konfliktach zbrojnych). B – sztuczne populacje, czyli rozwinięcie gry w życie dla bardziej realistycznych reguł. C – symulacje dla rzeczywistego miasta, w tym wypadku konfliktów w Bagdadzie. D – liniowy model ruchu. Źródło: (Weidmann, 2014)

Tematyka możliwości stosowania modeli agentowych jest bardzo szeroka, bo te modele dają największe spectrum zastosowań.



Rys. 29 Stopień skomplikowania modelu a jego poziom odzwierciedlenia rzeczywistości w liniowym schemacie. Źródło (Komosiński, 2011)

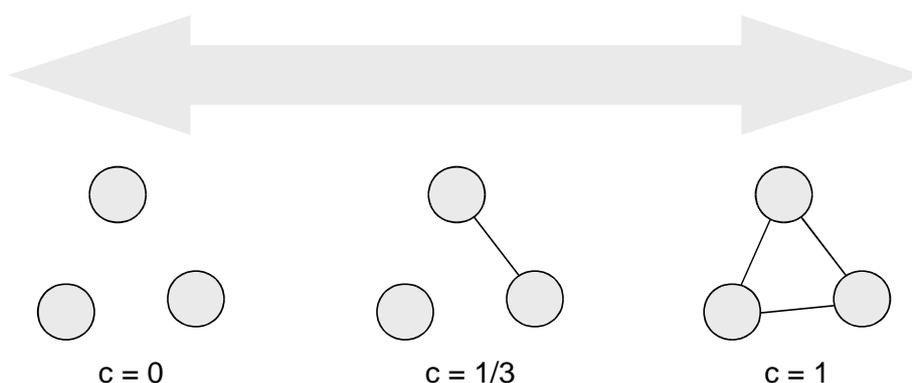
W większości przypadku opisywane tu modele dotyczą obszaru: ograniczonych modeli badawczych (Rys. 29), ze względu na problemy z kontrolą parametrów piętrzących się wraz z komplikacją modelu.

## 6 Analiza sieci społecznych i sieci złożonych

Analiza sieci społecznych - SNA (ang. social network analysis), to dziedzina, w której nauki ściśle włożyły bardzo dużo, ale również wiele zyskała. Ponieważ, zdaniem autora, jest to najbardziej obiecująca i mająca najwięcej praktycznych zastosowań dziedzina przenikania się nauk ścisłych i społecznych, powinna zostać najdokładniej opisana. Dyscyplina ta czerpie z teorii grafów na gruncie nauk społecznych. Jej przedmiotem są sieci tworzone przez różnego rodzaju „obiekty” społeczne – są nimi najczęściej ludzie, ale także grupy, firmy czy nawet państwa. Mamy cały zestaw poziomów od interakcji diad – czyli par, przez struktury trójkątne (różne motywy) po sieci wielopoziomowe oraz hipersieci (sieć sieci). Pewne charakterystyczne role sieciowe jak most (wierzchołek łączący dwie społeczności), czy jednostki izolowane zostały dokładniej opisane w zastosowaniach analizy sieciowej. Sieciami społecznymi w Polsce zajmuje się obecnie wiele ośrodków naukowych (reprezentujących fizykę, informatykę, matematykę, socjologię, psychologię, ekonomię, biologię czy medycynę) skupionych głównie w Warszawie i we Wrocławiu, ale należy wspomnieć, że Uniwersytet Jagielloński należał kiedyś do światowych centrów pionierskiej wtedy socjologicznej teorii sieci (Szmatka, 1997). Przedstawimy te techniki dwuetapowo, zaczynając od empirycznego opisu analizy sieci społecznych po matematyczny opis sieci złożonych.

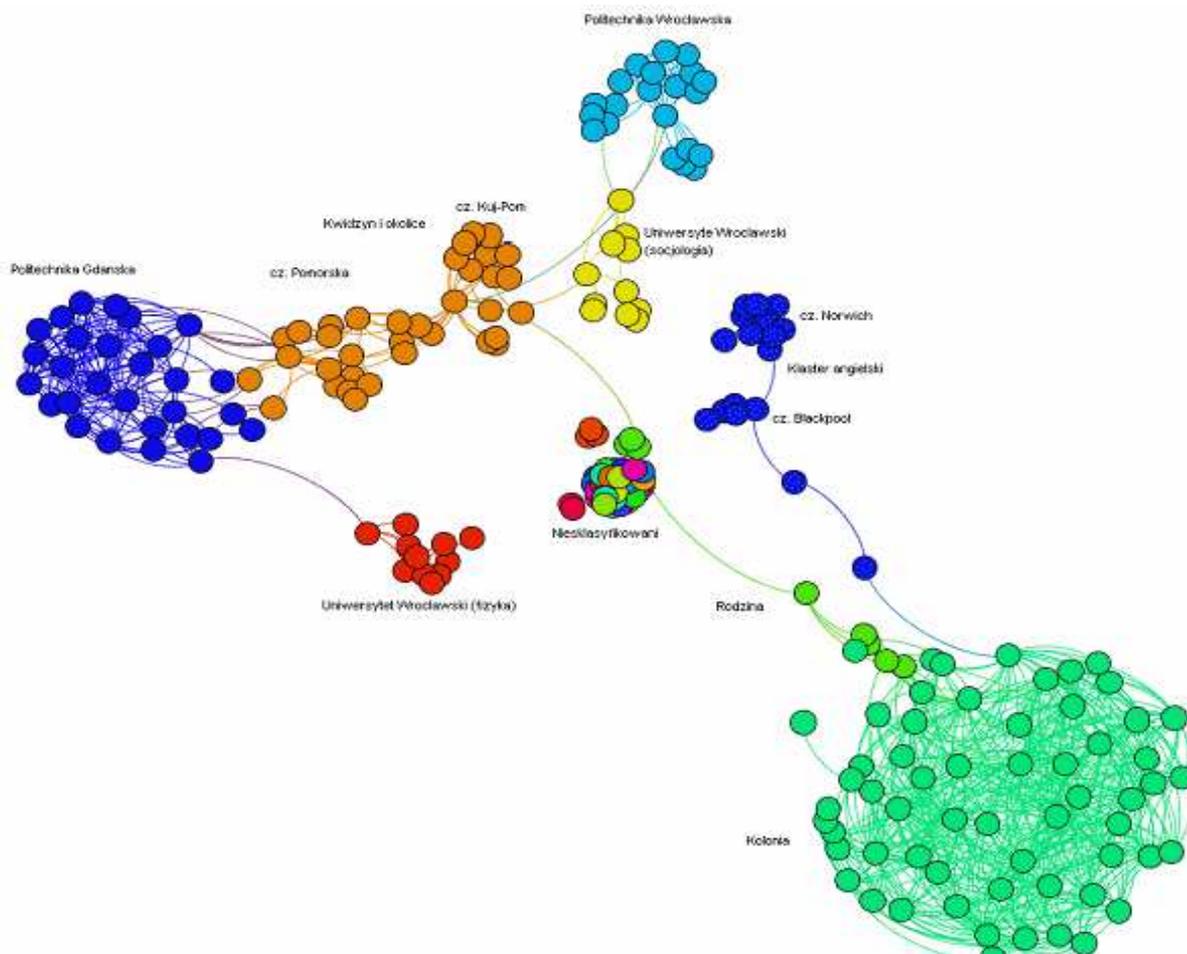
### 6.1 Podstawowe pojęcia w analizie sieci społecznych

Małe struktury społeczne są w socjologii znane od II Wojny Światowej i były obiektem badań przez Roberta Mertona (który wyróżniał wymiary władzy, komunikacji i interakcji) oraz Jacoba Moreno (ojca socjometrii w wymiarze interakcji – więzi społecznej). Przez wiele lat socjometria była stosowana w ograniczonym zakresie, ale odkąd algorytmy komputerowe pozwalają wizualizować całe układy automatycznie. Wcześniej robiono ręcznie wizualizacje i sam autor pamięta jak zaczynając studia socjologiczne męczył się z rysowaniem grafów, według przepisów, które i tak nie gwarantowały czytelności sieci. Opisując sieć używa się różnych współczynników oraz metod rozpoznawczych (Rocha, 2011), które często pod swoimi nazwami kryją szereg różnych algorytmów. W celu identyfikacji wzorców zachowań w sieci opracowano bardzo wiele schematów, ale większość bazuje na kilku podstawowych obserwacjach.



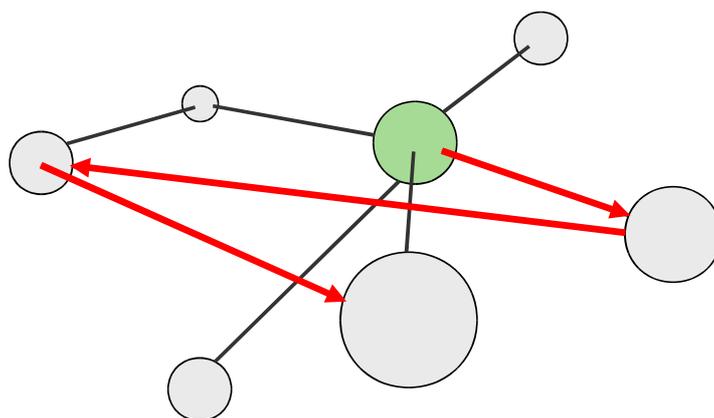
Rys. 30 Klasteryzacja i wykrywanie społeczności

**Sposób grupowania (ang. Clustering (C)),** interpretowany jako rozmieszczenie blisko siebie obiektów w jakiś sposób ze sobą powiązanych, a powstała struktura fizyczna określana jest jako klaster lub grono. Współczynnik gronowania (klasteryzacji) służy do szacowania, ilu sąsiadów danego wierzchołka jest połączonych każdy z każdym. Jest to iloraz liczby krawędzi pomiędzy sąsiadami danego wierzchołka oraz liczby krawędzi, jaki miałby graf pełny składający się z tych sąsiadów.



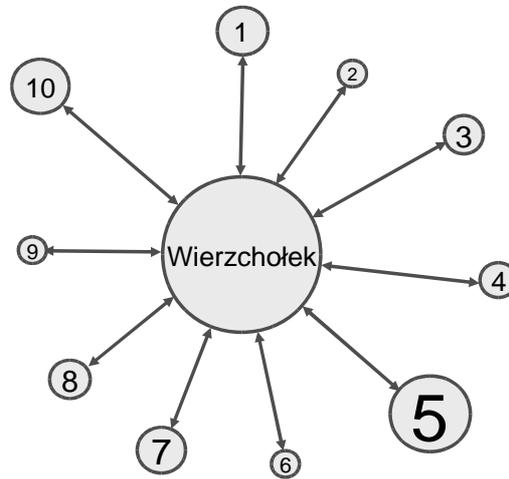
Rys. 31 Społeczność to zespół wierzchołków połączona ze sobą zdecydowanie częściej niż z resztą sieci

**Detekcja społeczności** polega na wyszukiwaniu podzbiorów sieci, które kontaktują się ze sobą częściej niż resztą sieci. Jest ogromna liczba algorytmów poszukujących takich podzbiorów, ale większość z nich wymaga ustalenia przez analityka liczby tych społeczności jako wielkości kontrolnej.



Rys. 32 Średnia odległość między węzłami

**Średnia długość ścieżki (L)** w sieci mówi, przez ile wierzchołków należy średnio przejść, aby dojść od jednego wierzchołka do dowolnego innego.

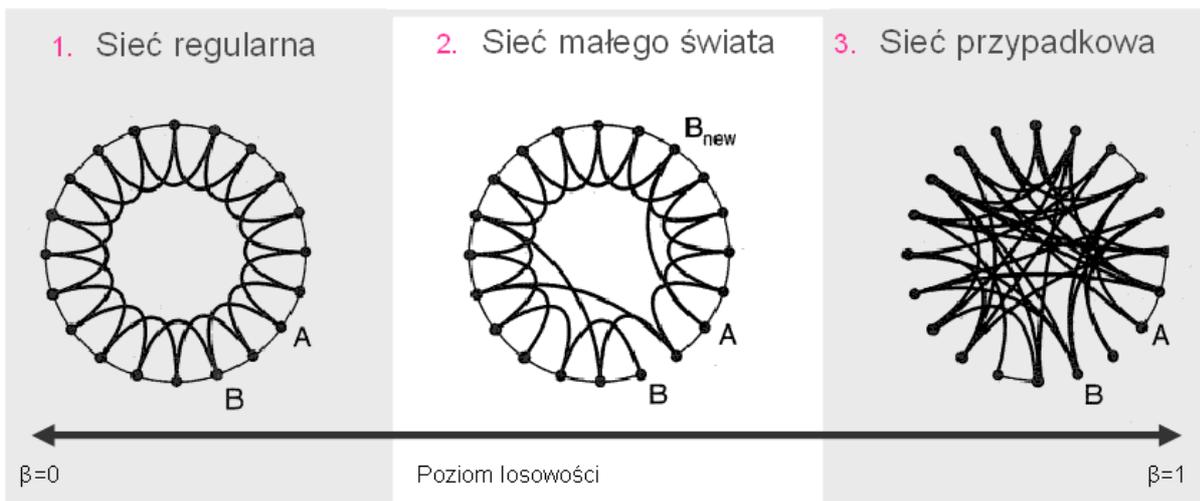


Rys. 33 Krotność wierzchołka i centralność wierzchołka

**Krotność** to ilość połączeń danego wierzchołka. W przypadku sieci skierowanej (uwzględniającej osobno połączenia przychodzące i wychodzące) można wyróżnić dodatkowo krotność *in* oraz *out*.

**Centralność** to atrybut opisujący pozycję każdego wierzchołka w sieci. Stanowi miarę ważności, znaczenia i wpływu na inne obiekty w tej samej sieci. Wyróżnia się wiele typów centralności, ale najczęściej stosowane to środek ciężkości i centralność „Betweenness”.

**Wizualizacja sieci** jest istotnym aspektem analizy sieciowej. Reprezentacja graficzna położenia i odległości między krawędziami nie ma zazwyczaj odzwierciedlenia w rzeczywistych odległościach i położeniu w sensie np. geograficznym. Wiąże się to z tym, że aby zmapować odległości w sensie matematycznym w ogólności dla  $N$  wierzchołków potrzebujemy  $N-1$  wymiarów. To znaczy, że na obrazku 2D, który ma 2 wymiary odniesiemy jedynie 3 wierzchołki. Są algorytmy wielowymiarowego skalowania, które z pewnym przybliżeniem pozwalają na zachowanie rzeczywistych odległości, ale ich skuteczność jest różna. To co się robi zazwyczaj, to prezentacja (wybór współrzędnych) sieci w sposób arbitralny. Najprostszym sposobem jest ulokowanie wierzchołków na obwodzie jakiejś figury, np. koła. W takim przypadku żaden z wierzchołków nie jest uprzywilejowany, a krawędzie mogą się w takim przypadku przecinać. Aby uniknąć właśnie przecinania się krawędzi oraz lokowania węzłów o dużej krotności w odpowiednich miejscach używa się różnych algorytmów bazujących np. na środku ciężkości, albo metodzie wartości własnych. W efekcie otrzymujemy bardziej czytelne przedstawienie sieci.



Rys. 34 Losowość od sieci regularnej (np. kratki, czy plastra miodu) przez jej zaburzenie do sieci przypadkowej prostej

**Poziom losowości** również jest istotny. Bardzo często sieci rzeczywiste są porównywane z teoretycznymi sieciami złożonymi, ponieważ ich własności są dobrze zbadane. Społeczeństwo, na które spojrzymy jak na zbiór osób powiązanych znajomościami, przyrównywane jest do znanej z przyrody struktury regularnej (jak w kryształach) lub do w pełni losowej, (którą można wygenerować z procesu przypadkowego). Jeśli nasza wygenerowana w pełni losowa sieć będzie miała rozkład stopni wierzchołków spełniający o określone prawo, to możemy ją odpowiednio sklasyfikować. Wyróżnia się 2 podstawowe typy sieci przypadkowych (Barabasi, 2002):

- **sieć przypadkową prostą lub inaczej Erdos'a-Renyiego (ER)**, kiedy każde połączenie jest równoprawdopodobne (tak jakby rozdzielić połączenia zupełnie przypadkowo).

- **sieć bezskalową lub inaczej Barabasiego-Albert (BA)**, kiedy rozkład krotności wierzchołka jest potęgowy, czyli długoogonowy (występują tak zwane huby, czyli wierzchołki o ogromnej licznie połączeń – rezultat efektu św. Mateusza, nazywane często preferencyjnym dołączaniem).

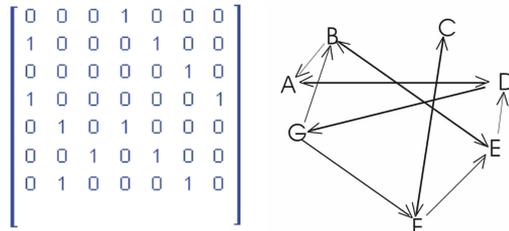
Sieci losowe można przeciwstawić sieciom o strukturze regularnej jak siatka czy plaster miodu (dobrze zbadanymi przez fizyków teoretyków). Pomiędzy skrajnymi przypadkami (regularnymi i losowymi) plasuje się sieć małego świata, często nazywana Watts i Strogatza (WS), gdzie na sieć regularną nałożone są przypadkowe połączenia (Sznajd-Weron, 2004). Odzwierciedlać to ma zjawisko znane w socjologii już od 1967 roku, po eksperymentach Milgrama, jako proces sześciu uściśnień ręki. Wspomniany amerykański psycholog społeczny rozesłał do kilkuset losowo wybranych ludzi z Nebraski i Kansas przesyłki z prośbą, by przekazali je dalej komuś ze swych znajomych, tak by możliwie najszybciej dotarły do pewnej osoby mieszkającej w Bostonie. Milgram śledził los przesyłek. Okazało się, że łańcuch dzielący osoby, które losowo wybrał, od celu, miał średnio sześć ogniw.

Generowanie sieci losowych jest ważnym elementem analizy sieciowej. Rzeczywiste sieci społeczne posiadają struktury, które nie występują w żadnej z wymienionych sieci przypadkowych. Znając podstawowe cechy więzów społecznych powinniśmy obserwować mniejsze bądź większe natężenie konkretnego wzorca topologicznego. W związku z tym losowe sieci społeczne generuje się w sposób bardziej skomplikowany, modyfikując odpowiednio proces generowania sieci dynamicznie. W rzeczywistości obserwuje się takie struktury jak: gwiazdy i kliki. Istnieją systemy losowania Siecie typu Siena, czy  $P^*$ , które bazują na grafie typu ER, ale przetasowują połączenia tak, aby uzyskać pożądane struktury. Takie sieci nazywane są grafami wykładniczymi i jest to metoda konkurencyjna do metod znanych z sieci złożonych prezentowanych w następnym rozdziale. W związku z konfliktem między tymi metodami, przedstawiciele paradygmatu sieci złożonych obwiniają grafy wykładnicze za brak uniwersalnych własności. Z drugiej strony głównym zarzutem jest niedostateczna porównywalność sieci złożonych z sieciami rzeczywistymi.

## 6.2 Sieci społeczne jako sieci złożone

W układach, które nie są zbiorem przypadkowo dobranych, niezwiązanych ze sobą elementów, do opisu struktury i oddziaływań między tymi elementami wykorzystuje się teorię grafów (sieci), która jest częścią matematyki dyskretnej. Analiza sieci społecznych (przedstawiona wcześniej) jest szczególnym przykładem sieci złożonych, w których zapis matematyczny jest zapożyczony z nauk przyrodniczych. Przedmiotem zainteresowania tej teorii jest graf  $G$ , będący zbiorem węzłów  $v_i$  i połączeń między nimi  $e_{jk}$ , gdzie  $i, j, k = 1, \dots, n$ . Koncepcja grafów ma już długą historię, która sięga XVIII w., gdy L. Euler rozpatrywał tzw. problem mostów miasta Królewiec. Została ona znacznie rozwinięta w latach 50-tych XX w., gdy węgierscy matematycy P. Erdős i A. Rényi rozważali własności grafów o przypadkowych połączeniach. Dopiero jednak pod koniec ubiegłego stulecia nastąpiła prawdziwa eksplozja zainteresowania wykorzystaniem grafów do analizy danych empirycznych, zwłaszcza pochodzących z badań układów złożonych. Podstawową cechą grafów jest kierunkowość połączeń; jeśli połączenia są symetryczne, to graf jest nieskierowany, natomiast w

przeciwnym razie mamy do czynienia z grafem skierowanym, w którym połączenia mają swój początek i koniec. Oba typy grafów, choć pozornie wydają się obiektami geometrycznymi, mogą w istocie być łatwo wyrażone w postaci algebraicznej przy pomocy tzw. macierzy sąsiedztwa. Dla grafów nieskierowanych macierz ta jest symetryczna, natomiast dla skierowanych – asymetryczna. Inną ważną cechą grafów jest rodzaj połączeń, które mogą być ważone lub nieważone (binarne). Przykładowy graf skierowany o binarnych połączeniach został zaprezentowany poniżej w formie macierzy i w formie graficznej.



Rys. 35 Sieć skierowana w wersji macierzowej oraz zwizualizowana

Element macierzowy (1,4) (pierwszy rząd, czwarta kolumna) odpowiada połączeniu pomiędzy węzłami A i D. Gdyby macierz z rysunku była symetryczna, odpowiadałaby sieci nieskierowanej.

Podstawowym parametrem każdego węzła jest liczba połączeń (stopień, krotność)  $k$  z innymi węzłami, która w przypadku sieci skierowanych dzieli się na liczbę połączeń przychodzących i wychodzących. Sieć skierowana jest wygodna do opisu jednostronnych oddziaływań między elementami np. sieci wpływów, natomiast w mojej pracy ograniczam się do użycia grafów nieskierowanych ze względu na symetrię problemów, którymi się zajmuję. W zależności od własności statystycznych rozkładu krotności węzłów i mechanizmu tworzenia się połączeń wyróżnić można kilka ważnych typów sieci (termin ten będzie używany wymiennie z grafem).

Historycznie pierwszym rozważanym w fizyce typem sieci były sieci regularne, w których każdy węzeł był charakteryzowany przez taką samą wartość  $k$ . Sieci tego typu były wykorzystywane zwłaszcza w fizyce ciała stałego do opisu struktury krystalicznej. Kolejnym typem grafów były tzw. sieci przypadkowe, wprowadzone przez Erdösa i Rényiego (tzw. model ER), w których połączenia między węzłami tworzyły się w sposób losowy przez wybór pary węzłów.

Modyfikacją modelu ER jest sieć, która tworzy się nie przez wypełnianie połączeniami istniejących węzłów, ale poprzez dodawanie nowych do już istniejących węzłów, losując za każdym razem węzeł, z którym łączony jest nowy. Rozkład połączeń w takiej sieci jest także dany rozkładem wykładniczym. Przez długi czas uważano, że sieci typu ER dobrze opisują własności sieci rzeczywistych. Tymczasem z upływem lat okazało się, że sieci rzeczywiste posiadają własności mocno niekiedy odbiegające od tego modelu. Sieci typu ER są jednorodne i żaden węzeł nie jest statystycznie wyróżniony (mówi się w tym wypadku o sieciach demokratycznych), natomiast w wielu sieciach rzeczywistych występuje silna niejednorodność i pewne węzły są faworyzowane. Prowadzi to do potęgowych rozkładów krotności. A.-L. Barabási i R. Albert jako pierwsi pokazali, że sieć WWW, do tej pory uważana za sieć typu ER, ma strukturę bezskalową (potęgową). Potęgowy rozkład  $k$  implikuje istnienie niewielkiej liczby węzłów o dużej krotności (tzw. hubów) oraz dużej liczby węzłów o niewielkich krotnościach. Implikuje także istnienie korelacji strukturalnych w sieci, o bardzo dużych hubach. Okazuje się, że tego typu topologię wykazuje wiele innych sieci rzeczywistych, takich jak: fizyczne sieci komputerowe (hardware), sieci transportowe (połączenia lotnicze), sieci metaboliczne, sieci opisujące rynki finansowe, a zwłaszcza sieci powiązań społecznych: sieci współpracy artystów, kolaboracji naukowych, cytowań publikacji naukowych, sieci rozprzestrzeniania się epidemii, czy sieci znajomości w mediach społecznościowych. Rozpowszechnienie sieci o potęgowych rozkładach opisujących bardzo odmienne od siebie układy spowodowało potrzebę wprowadzenia nowego modelu rozwoju sieci, który w efekcie dawałby tego typu topologię. Najbardziej popularnym modelem jest model wprowadzony przez Barabásiego i Albert (tzw. model BA), w którym mechanizmem wzrostu sieci jest preferencyjne

przyłączanie. Polega ono na przyłączaniu nowych węzłów do sieci za pomocą przypadkowych połączeń z istniejącymi węzłami, przy czym prawdopodobieństwo przyłączenia do konkretnego węzła dane jest przez krotność tego węzła. Taki mechanizm wydaje się być intuicyjny, ponieważ np. w przypadku cytowań publikacji naukowych, szansę aby otrzymać kolejne cytowania mają przede wszystkim te artykuły, które już zdobyły duży rozgłos i dużą liczbę cytowań. Sieci bezskalowe posiadają inne ciekawe cechy, które sprawiają, że naturalne układy tak często preferują tę właśnie topologię. Przede wszystkim mała liczba węzłów kluczowych dla struktury sieci powoduje, że przypadkowy wyeliminowanie takiego węzła jest mało prawdopodobne (najwięcej jest węzłów peryferyjnych i mało istotnych). Ponieważ przyroda nie ma informacji o tym, które węzły są hubami, ten typ topologii dobrze zdaje egzamin. Niestety, problem staje się poważny, gdy atak na sieć przeprowadzi istota inteligentna, która ma informację o położeniu hubów, wtedy taką sieć bezskalową łatwo zniszczyć; przykładem są blackouty elektryczne i ataki hackerskie w Internecie. Drugą istotną własnością takich sieci jest ich optymalność pod kątem zasobów potrzebnych do ich skonstruowania – posiadają krótkie drogi łączące węzły między sobą, nawet jeśli te węzły są od siebie odległe w przestrzeni.

Oprócz długości ścieżek sieci, opisujących zawartość ich struktury, do ważnych cech należy własność gronowania (klasteryzacji). Polega ona na tym, że sąsiadujące ze sobą węzły tworzą w pełni połączone trójki. Miarą opisującą gronowanie jest tzw. współczynnik gronowania (czasami nazywany klasteryzacją ( $C$ )), który jest równy prawdopodobieństwu, że pomiędzy dwoma sąsiadami danego węzła także istnieje połączenie. Nazwa sieci małego świata ma swoje źródło w eksperymencie S. Milgrama, w którym określona została średnia odległość pomiędzy przypadkowo wybranymi ludźmi w sieci znajomości. Okazała się ona zaskakująco mała:  $L \approx 6$ . Dla porównania, współcześnie ta długość może być jeszcze mniejsza i wynosi 4,74 (wg serwisu społecznościowego Facebook). Tzw. problem małego świata był przez długi czas uważany za ciekawostkę, aż do momentu, gdy D.J. Watts i S. Strogatz opublikowali model, w którym własności sieci małego świata można otrzymać w niezwykle prosty sposób, przełączając losowo drobny ułamek istniejących połączeń w sieci regularnej. Model ten, zwany modelem WS, daje sieć o małej wartości  $L$  i dużej wartości  $\langle C \rangle$ . Obecnie wiadomo, że sieci małego świata są powszechne w przyrodzie i strukturach społecznych.

Jest wiele innych miar charakteryzujących sieci, np. asortatywność, która odnosi się do korelacji między węzłami różnych typów. Dla różnych układów wielkość ta waha się od przypadku dysasortatywnego, gdy węzły o dużej krotności łączą się najchętniej z węzłami o małej krotności (np. Internet, sieci zależności pokarmowych w ekosystemach, sieci reprezentujące rynki finansowe), do asortatywnego, gdy węzły tych samych typów grupują się razem (sieci społecznościowe, sieci współautorstwa publikacji naukowych). W przypadku rynku fonograficznego spodziewać się należy *mieszania asortatywnego*, ponieważ największe gwiazdy, o dużej liczbie współpracowników, często współpracują z artystami o podobnej randze. Inną miarą jest rozmiar sieci lub, w przypadku sieci składających się z oddzielnych fragmentów, największego jej spójnego podzbioru (ang. *giant cluster*). Jest to istotna miara w przypadku sieci, w których mamy do czynienia z pojawianiem się lub znikaniem połączeń, gdyż można dzięki niej określić, czy w sieci ma miejsce odpowiednik fizycznego zjawiska perkolacji. Samo zjawisko perkolacji ma olbrzymie znaczenie dla problemu funkcjonowania sieci, ponieważ przejście z sieci spójnej do rozłącznych jej fragmentów w przypadku pewnych układów powoduje destrukcję sieci i zaburzenie procesów odbywających się na tych sieciach.

W klasyfikacji sieci najważniejszym kryterium jest podział na sieci przypadkowe i deterministyczne, a następnie na statyczne i ewoluujące. W tej pracy przedmiotem badań będą głównie sieci ewoluujące, powstałe w sposób indeterministyczny. Współcześnie obserwuje się niebywały rozwój teorii sieci, co pozwoliło zarówno na odkrycie wielu interesujących zjawisk w układach rzeczywistych, wyrażanych przez reprezentacje sieciowe, jak i opracowanie nowych metod badania tych układów. Jednym z najbardziej spektakularnych osiągnięć analizy sieciowej jest wydobywanie informacji, które nie są bezpośrednio dostępne obserwacjom. Chodzi tu o identyfikację pewnych ukrytych poziomów struktury układów, odtwarzanie nieznanymi, ale istniejącymi fragmentów struktury oraz przewidywanie przyszłych zdarzeń.

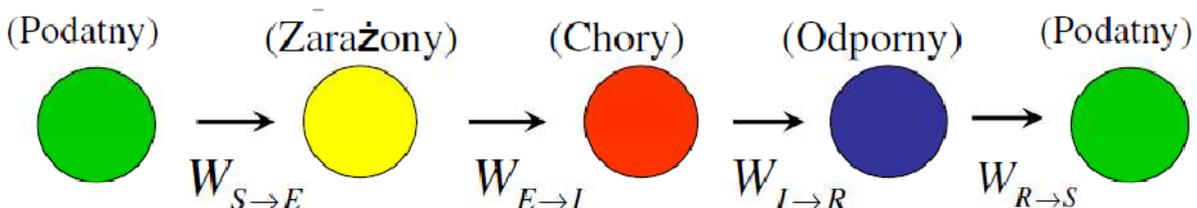
Przykładem są tu sieci kontaktów w organizacjach przestępczych, przewidywanie miejsca i czasu przestępstw, przewidywanie wyroków sądów na podstawie znajomości składów sędziowskich i przewidywanie decyzji konsumenckich na podstawie historii zakupów.

## CZĘŚĆ II

# MODELOWANIE PROCESÓW SPOŁECZNYCH W PRAKTYCE (SOCJOFIZYKA)

## 7 Modelowanie epidemiologiczne

Modele przenoszenia patogenów w wyniku kontaktów społecznych są największą gałęzią modelowania procesów społecznych oraz obszarem największych sukcesów w praktycznym zastosowaniu modelowania i analizy danych. Sam fakt istnienia kontaktu jest warunkiem koniecznym do przekazania choroby, której dalszy rozwój jest już opisywany parametrami medycznymi. Zaprezentujemy teraz wybór modeli epidemiologicznych, które w swojej naturze są dedykowane Polsce (parametry dostosowane na podstawie danych o polskim społeczeństwie oraz problem badawczy dotyczący polskiej epidemiologii). Taki model dla miasta Wrocławia powstał już w XVIII w. (jako pierwszy na świecie).



Rys. 36 Rozwinięty schemat rozwoju choroby typu SEIRS

### 7.1 Model Bernoulli'ego – pierwszy opis epidemii (Bernoulli, 1766)

Pierwszy taki model powstał prawdopodobnie już w XVIII w. Znany szwajcarski matematyk, Bernoulli, był pierwszą osobą, która wyraziła odsetek osób podatnych na zakażenie endemiczne pod względem siły zakażenia i długości życia i w tym celu przyjął dane opisujące epidemię czarnej ospy we Wrocławiu. W Polsce ospa pojawiła się po raz ostatni również we Wrocławiu w 1963 roku, ale została zatrzymana m.in. przez działania rządu i epidemiologów. Bernoulli oparł się w swojej pracy na Edmundzie Halley'u – znanym brytyjskim astronomie (Halley, 1692/1693). W XVII wieku angielskie *Breslaw* oznaczało *Breslau* – niemiecką nazwę Wrocławia. „Tabele List Śmiertelności miasta *Breslaw*” (Halley, 1692/1693) zostały mu przekazane przez pana Justell'a, gdzie wiek i płeć każdego zmarłego w danym miesiącu zostały zapisane i porównane z liczbą urodzin z pięciu lat (od 1687 do 1691) i które zostały zrobione z zachowaniem wszystkich wymagań dokładności i szczerości. Głównym celem Bernoulli'ego było policzenie skorygowanej tabeli życia,

Na podstawie swoich rozważań Bernoulli skonstruował tabelę opisującą zmiany populacji we Wrocławiu i otrzymał średnią długość życia w stanie naturalnym jako 26 lat i 7 miesięcy, a dla stanu bez ospy 29 lat i 9 miesięcy. Parametry wprowadzone przez Bernoulli'ego są stosowanych do dziś: siła zakażenia  $1/n$  (roczna zapadalność na osobę), w przypadku zgonu  $1/m$  (śmiertelność).

### 7.2 Projekt System Informatyczny Redukcji Szpitalnych Zakażeń (SIRS-Z)

Ze względu na zamkniętą sieć kontaktów międzyludzkich oraz występowaniem patogenów (HAI, czyli szpitalne zakażenia) specyficznych tylko dla tego typu miejsc, modelowanie zakażeń szpitalnych daje potencjalnie możliwości niedostępne przy chorobach dotyczących ogółu społeczeństwa. Autorzy książki jest

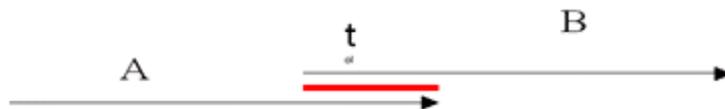
w trakcie realizacji projektu, którego celem jest zmniejszenie liczby zakażeń szpitalnych do których dochodzi w polskich szpitalach poprzez stworzenie bezpłatnej aplikacji desktopowej (na licencji wolnej i otwartej) wspierającej pracę epidemiologa szpitalnego w obszarze zakażeń szpitalnych (bakteryjnych, wirusowych, grzybiczych i pasożytniczych). Docelowym przeznaczeniem modelu jest odtworzenie najbardziej prawdopodobnych ścieżek zakażeń HAI (ang. hospital acquired infections) oraz klasyfikacja jednostek do różnych grup ryzyka. Problem jest powszechny, bo np. w roku 2012 liczba pacjentów szpitali wyniosła 7,9 mln według GUS z czego 5-10% uległo zakażeniu szpitalnemu. Dotychczasowa sprawozdawczość podmiotów medycznych w dziedzinie zakażeń szpitalnych była bardzo wybiórcza. Nawet jeśli są rejestrowane, często są ukrywane (80% szpitali nie przyznało się do ani jednego takiego zdarzenia w roku 2011). Korzystając w opracowań dostępnych raportów szpitalnych została wygenerowana mapa wiedzy dotycząca tych ryzyka zakażeń.



Rys. 37 Sieć wiedzy Bayesa wygenerowana dla projekcji [czynniki ryzyka występowania zakażenia] – [typy zakażeń związanych]. Źródło (Zygmunt & Valenta, 2005)

Ze względu na drogę wykorzystuj się w badaniu najważniejsze sieci kontaktów społecznych: pacjent-pacjent; pacjent – personel; pacjent – personel – pacjent, personel – pacjent; oraz dodatkowo w postaci poza zewnętrznego uwzględnimy środowisko – pacjent. Wszystkie patogeny przenoszone tymi ścieżkami kontaktu mogą być badane proponowanym oprogramowaniem, choć model testowany był pierwotnie na szpitalnych szczepach odpornych na antybiotyki szczepach bakterii MRSA (multi-odporny gronkowiec złocisty).

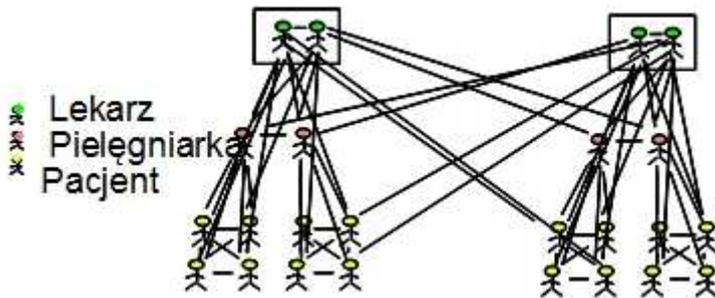
Największym problemem w budowaniu modeli w zakresie zdrowia publicznego jest brak indywidualnych informacji na poziomie pacjentów. Leczenie szpitalne jest kwestią indywidualną w związku z tym bardzo trudno o dostęp i administrację danych, w tym wypadku szczególnie wrażliwych. Tak więc jest wiedza o kontaktach społecznych i liczbie zarażeń, co akurat w miarę precyzyjnie jest rejestrowane. Natężenie kontaktów międzyludzkich w środowisku szpitalnym jest mierzone czasem takiego kontaktu (przebywanie w tej samej jednostce szpitalnej określoną ilość czasu jest podstawą do modelowania).



Rys. 38 Modelowe kontakty wewnątrzszpitalne: czas przebywania wspólnego w tej samej sali (klinice) jako miara siły kontaktu.

Do modelowania zakażeń szpitalnych wykorzystujemy rejestr przyjęć i wypisów poszczególnych pacjentów zawierających dokładne miejsce pobytu (sala) wraz z bazą rejestrowanych przypadków testów na posiadanie interesującego nas patogenu. Istnienie kontaktu między jednostkami daje możliwość przekazania patogenu, ale tego nie gwarantuje. W celu określenia warunków zarażenia należy zbudować mechanizm transmisji. W przypadku HAI dokładne dotyczące kontaktów między pacjentami (wyznaczonymi na podstawie przebywania w tej samej sali) są podstawowym źródłem dla funkcjonowania naszego

algorytmu. Kontakty z personelem, będą stanowić uzupełnienie, ponieważ wiemy iż pracownicy szpitali ze względów prywatnych nie są skłonni na rejestrację w systemie i w zależności poziomu dostępu do danych epidemiologa ta informacja może zostać wprowadzona. Aby uzupełnić pośrednią rolę kadry (w przypadku braku twardych danych) dodaliśmy „wirtualne” kontakty na wyższym poziomie sąsiedztwa, czyli na poziomie klinik – w których zatrudniona jest kadra.



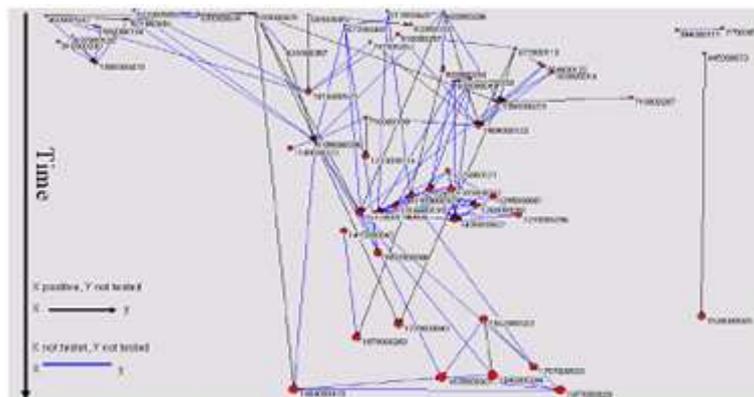
Rys. 39 Modelowe kontakty wewnątrzszpitalne wskazujące na potencjalne ścieżki zarażeń między pacjentami oraz personelem. Źródło: (Jarynowski A. , 2013)

Do modelowania zakażeń szpitalnych została zaimplementowana metodologia MCqMC (Monte Carlo quasi – Markov Chain) dzięki której możemy symulować sztucznie epidemie na empirycznym zbiorze danych. Jest to model typu SI (zgodnie z klasyfikacją modeli epidemiologicznych z rozdziału o teorii modelowania), ponieważ większość typów zakażeń szpitalnych jest odporna na leczenie. Prawdopodobieństwo przejścia ze stanu  $S$  do  $I$  jest dla agenta (reprezentującego pacjenta)  $i$  w chwili  $t$  opisywane jest funkcją:

$$p_i(t) = 1 - \exp\left(-\sum \sqrt{P_i(i,j)}s - \sum \sqrt{Q_i(i,j)}k - \sum \sqrt{P'_i(i,j')}m\right)$$

gdzie  $P$  jest macierzą kontaktów między zarażającymi na poziomie sal,  $Q$  na poziomie klinik a  $P'$  uwzględnia wszystkich potencjalnych pacjentów, a  $s, k, m$  są parametrami skalowania intensywności kontaktów na poszczególnych poziomach sąsiedztwa.

Wynikiem takiej symulacji są nie tylko wirtualne epidemie, ale również ścieżki zakażeń (z wielu realizacji należy wybrać najczęściej powtarzające się). Algorytm wymaga jednak uzupełnienia o nowe estymatory parametrów (dotychczasowa metoda najmniejszych kwadratów jest zbyt naiwna). Poza głównym algorytmem działają dodatkowe – pomocnicze, takie jak dynamiczne sieci Bayesa, sieci neuronowe i automaty komórkowy oraz możliwość przewidywania zakażeń przez modele uczenia maszynowego.



Rys. 40 Przykładowe najbardziej prawdopodobne ścieżki zakażeń przy uwzględnieniu strzałki czasu. Źródło: (Jarynowski A. , 2010)

### 7.3 Badanie sieci kontaktów: Polymod

W celu przygotowania modeli sieci społecznej można wykorzystać dane eksperymentalne zebrane w Polsce (Grabowski & Rosinska, 2012) w ramach projektu POLYMOD (Mossong & Rosinska, 2008). Badania ankietowe były przeprowadzone przez Narodowy Instytut Zdrowia Publicznego - Państwowy Zakład Higieny

(NIZP-PZH). W przeprowadzonych w Polsce badaniach ankietowych wzięło udział 1012 osób, które zarejestrowały dane o 16501 kontaktach. Uczestnicy badań ankietowych byli rekrutowani przez profesjonalnych ankieterów, którzy odwiedzali losowo wybrane gospodarstwa domowe. Osoby do badania były dobrane w taki sposób, aby uwzględnić strukturę demograficzną Polski. Każdy z uczestników otrzymał dzienniczek w którym odnotowywał informacje na temat bezpośrednich kontaktów interpersonalnych, które miały miejsce jednego, losowo wybranego dnia. Ochotnicy biorący udział w badaniu odnotowywali takie informacje jak: całkowity czas kontaktu z daną osobą w ciągu dnia oraz częstość tych kontaktów.

Przeprowadzona analiza wyników powyżej opisanych badań ankietowych pokazała, że badana sieć ma trójpoziomą strukturę kontaktów interpersonalnych: (a) kontakty codzienne (z ludźmi, których spotykamy niemal każdego dnia; stanowią one 72.4% wszystkich kontaktów); (b) kontakty sporadyczne (z ludźmi, których spotykamy kilka razy w miesiącu; stanowią one 16.3% wszystkich kontaktów) oraz (c) kontakty przypadkowe (z ludźmi których spotykamy po raz pierwszy; stanowią one 11.3% wszystkich kontaktów).

Wartości liczbowe rozkładów prawdopodobieństwa  $P(w)$  są przedstawione w Tabeli 2. Uzyskane wyniki pokazują, że czas trwania kontaktu jest silnie skorelowany z częstością jego występowania.

Lp.	(waga połączenia)	Wszystkie kontakty	Kontakty codzienne	Kontakty sporadyczne	Kontakty przypadkowe
1	(5/(24 x 60))	0.14	0.10	0.20	0.31
2	(10/(24 x 60))	0.18	0.14	0.26	0.34
3	(40/(24 x 60))	0.19	0.17	0.30	0.21
4	(2.5/24)	0.19	0.20	0.20	0.10
5	(4/24)	0.30	0.40	0.03	0.04

Tabela 2 Rozkłady prawdopodobieństwa opisujące wagi połączeń dla kontaktów o różnej częstości występowania. Każdy element macierzy zawiera odsetek kontaktów o określonej częstości występowania i wadze (czasie trwania), np. 14% wszystkich kontaktów trwało krócej niż 5 minut. Źródło: (Grabowski & Rosinska, 2012)

Na podstawie tego badania ułatwione będzie modelowanie procesów w których kontakt fizyczny jest potrzebny (jak przenoszenie chorób), dzięki rozkładowi natężenia poszczególnych kontaktów.

## 7.4 Modele zagrożeń bio-terrorystycznych

Na potrzeby bezpieczeństwa narodowego zostały opracowane proste modele i symulacje (Najgebauer, 2009) rozprzestrzeniania się mocno zakaźnych patogenów (jak czarna ospa, czy wąglik) z uwzględnieniem kierowanego zainicjowania łańcucha epidemiologicznego (bioterroryzm). Ze względu na możliwy celowany charakter ataku, wierzchołki centralne sieci kontaktów społecznych są szczególnie interesujące, ponieważ pełnią one kluczowe role i stanowią swego rodzaju katalizatory epidemii. Jednym z wyników tych analiz było oszacowanie 10 krotnie większej skuteczności ataku terrorystycznego w przypadku celowania w węzły sieci. Za to efektem praktycznym są gotowe aplikacje do wspomaganiania zarządzania kryzysowego typu SARNA, czy CARE. W zaimplementowanej realizacji można:

- symulować epidemie na sieciach złożonych;
- wizualizować możliwe scenariusze rozprzestrzeniania się choroby (z wbudowanymi parametrami medycznymi dla potencjalnie niebezpiecznych patogenów);
- zidentyfikować tzw. supers-spreaders;
- ocenić potencjalne skutki epidemii.

Duże skupiska ludzi, zwłaszcza aglomeracje jak Warszawa są szczególnie narażone na atak (Brouwers, 1995), dlatego poza modelami zarażenia w wyniku interakcji człowiek-człowiek rozważono skażenia o zasięgu terytorialnym na rzeczywistej topologii miasta, np. chlorem.

## 7.5 Grypa pandemiczna – zwłaszcza przykład H1N1, tzw. Świńskiej grypy w 2009

Model grypy pandemicznej są chlubnym wyjątkiem, gdzie wiele instytucji z całej Polski podjęło trud budowy dedykowanego modelu: CIOP (Grabowski & Kosinski, 2009), WAT (Kasprzyk, Najgebauer, & Pierzchała, 2010), ICM UW (Rakowski, Gruzziel, Bieniasz-Krzywiec, & J, 2010) (Rakowski, Gruzziel, Krych, & Radomski, 2010), AGH (Porzycki, 2013), PWr (Wdowiak, 2007), (Wywrot, 2004).

W związku z tym, że kontakty między podatnymi i zainfekowanymi są nielosowe można pokusić się o symulacje dla społeczeństwa całego kraju –Polski na podstawie danych z GUS-u:

- krajowy rejestr populacji (informacje na temat struktury zatrudnienia, bądź pobierania nauki, wielkości gospodarstwa domowego, wielkości zakładów pracy itp.);
- geograficzna baza gmin.

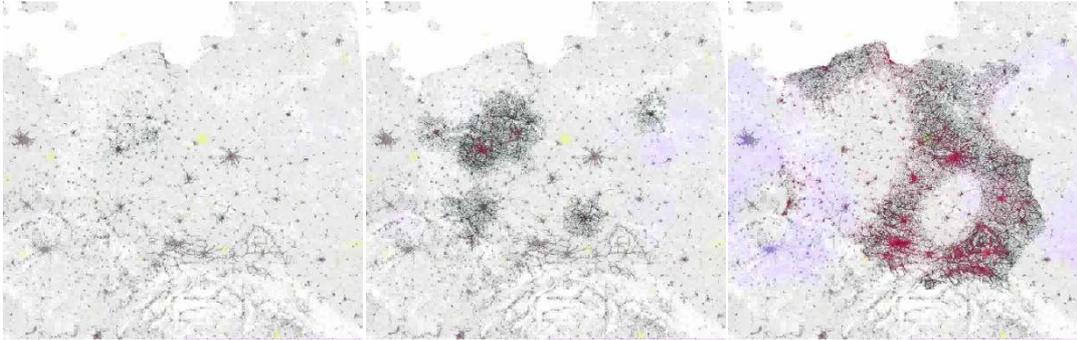
Przeniesienie choroby może odbywać się przynajmniej dwa razy dziennie w domu i głównym miejscu przebywania (np. praca i szkoła). Dodatkowo uwzględnia się możliwość zarażenia poprzez wektory (transport), przy czym prawdopodobieństwa zostały oszacowane na podstawie modelu grawitacyjnego (ICM) lub na podstawie potoków (AGH). W realizacji AGH dodatkowo transmisja patogenu ma miejsce przypadkowo, a w realizacji ICM poprzez sąsiedztwo. Program sprawdza symuluje potencjalne kontakty. Dodatkowe kontakty (i zakażenie) poza-domowe są zawarte w modelu na dwa sposoby przez: miejsce pracy/nauki, sąsiedztwo, podróż. Schemat zakażeń zależy od stopnia zachorowań. Parametry epidemiologiczne takie jak zakaźność, intensywność choroby, czy czas trwania infekcji zostały dopasowane do rozkładu statystycznego (konkretnie rozkładu Weibulla dla jednej z realizacji modelu). Przebieg choroby gwarantuje odporność w danym sezonie. Po przeprowadzeniu kalibracji odpowiedzi immunologicznej wyznaczono różne scenariusz, w których  $R_0$  (współczynnik reprodukcji epidemii opisany w rozdziale o modelowaniu deterministycznym) mieści się w przedziale od 1,1 do 2,1. Symulacje były prowadzone przy następujących założeniach: wybuch pandemii grypy w Polsce nastąpił w czerwcu lub we wrześniu (w różnych miejscach kraju), wartość  $R_0$  odpowiada w przybliżeniu 1,4 w modelu głównym, ponieważ taką wartość wyznaczono w Nowej Zelandii epidemii zimowej w południowej hemisferze. Wykorzystano estymator  $R_0$  w postaci:

$$R_0 = \frac{-\ln\left(\frac{A}{S_0}\right)}{1 - \frac{A}{S_0}}$$

Gdzie:  $S_0$  – łączna ilość osób podatnych przed wybuchem epidemii;  $A$  – łączna ilość osób podatnych po wybuchu epidemii.

Poza reprodukcją epidemii została podjęta dodatkowo próba oszacowania kosztów epidemii (wyniki niepublikowane). Aby porównać koszty społeczne dla różnych scenariuszy, zostały wykorzystane: koszt jednego dnia nieobecności w pracy, koszt leczenia przez lekarza pierwszego kontaktu, koszt jednego dnia leczenia szpitalnego, koszt szczepionki i administracji szczepień. Np. koszt szczepionki na poziomie 100PLN od osoby został oparty na założeniu, że dla całej potencjalnej populacji szczepionych (w sumie 76 milionów dawek) koszty są podzielone pomiędzy koszt szczepionki i administracji. Model zakłada, że pracownicy

opiekują się chorymi dziećmi (tak jakby rodzic był chory). Następujące zostały porównane scenariusze: brak szczepień oraz szczepienia na różnych procentowych poziomach objęcia.

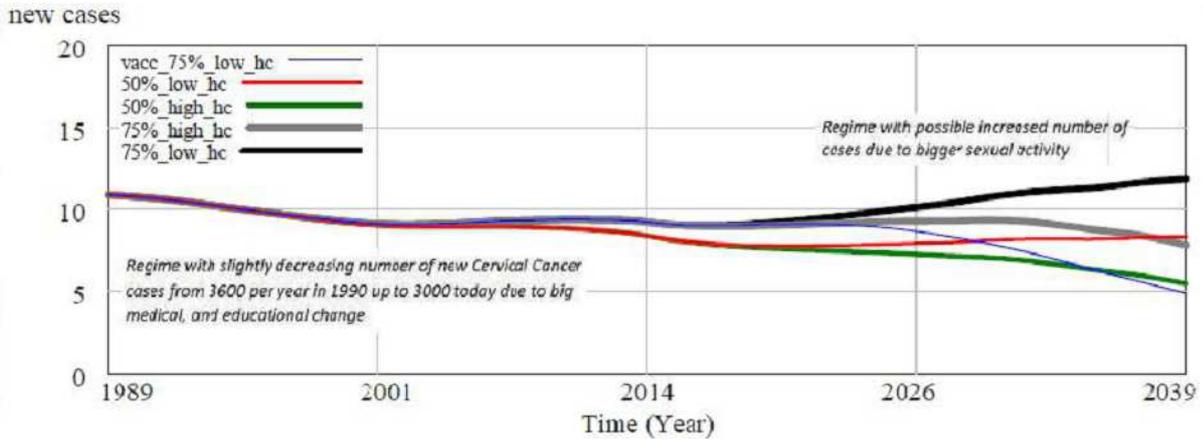


Rys. 41 Wizualizacja rozprzestrzeniania się grypy ze źródłem w Wielkopolsce. Źródło (Rakowski, 2010)

Model ten był pierwszym w Polsce wykorzystującym indywidualne dane o kontaktach na poziomie całego kraju. Należy więc być ostrożnym przy interpretowaniu wyników (na podstawie wstępnych analiz wynikało, że należy zakupić szczepionki i zaszczepić nimi 70% populacji). Jednak wiedza zdobyta na tym poziomie zaowocuje zapewne tym, że kolejne iteracje modelu staną się bardziej realistyczny.

## 7.6 Rozprzestrzenianie się chorób przenoszonych drogą płciową – wirus HPV

Życie seksualne Polaków, mimo dużego zainteresowania medialnego, jest cały czas słabo poznane. Z punktu widzenia epidemiologii również mało wiemy o stanie zdrowia Polaków (WHO klasyfikuje nas w dziedzinie chorób przenoszonych drogą płciową na poziomie krajów trzeciego świata). Programy profilaktyczne, czy edukacja seksualna oparte często są na nienaukowych stereotypach i nie przygotowują na nowe fale patogenów endemicznych w krajach sąsiednich. Już wstępna analiza dotycząca wirusa HPV pozwala wyciągnąć wnioski dotyczące złej profilaktyki raka szyjki macicy w Polsce (wciąż promowanej przez Państwowy Zakład Higieny, który nie uwzględnia w swoich ekspertyzach zmian epidemiologiczno-demograficznych). W celu uzupełnienia tej luki powstał (i jest cały czas rozwijany) zweryfikowany, działający model rozprzestrzeniania się chorób przenoszonych drogą płciową (ang. sexually transmitted infections: STI) dedykowany polskiej społeczności, docelowo uwzględniający najważniejsze ścieżki zakażeń (główna populacja, prostytutka oraz mężczyźni uprawiających seks z mężczyznami) dla najczęstszych patogenów (wirusy: HPV, HSV, HIV, HB(C)V; bakterie: Chlamydie, Rzeżączki, Krętki; grzyby: Drożdżaki; inne drobnoustroje: Rzęsistki). Obecnie dostępna krajowa literatura badawcza, bądź edukacyjna skupia się na rozprzestrzenianiu HIV, mimo że endemiczne zarażenia (poza grupami wysokiego ryzyka) praktycznie zanikają w krajach UE, dlatego w naszych rozważaniach stawiamy nacisk na potencjalne problemy przyszłości, a nie przeszłości. W konsekwencji przemian, a w szczególności zaobserwowania w Polsce tak zwanego Drugiego Przejścia Demograficznego, permanentnie wzrasta aktywność seksualna (w kontekście liczby nowych partnerów) oraz społeczeństwo się starzeje. Dodatkowo należy pamiętać o epidemiologicznych konsekwencjach członkostwa Polski w EEA i swobodnej migracji ludzi (również tych zarażonych). Dotyczy to Polaków powracających z pobytów turystycznych, czy zarobkowych w krajach o dużo większej występowalności patogenów .



Rys. 42 Liczba nowych przypadków raka szyjki macicy w Polsce dziennie w poszczególnych latach (po lewej dane historyczne po prawej wyniki symulacji dla różnych scenariuszy). Źródło: (Jarynowski & Sefimovic, 2014)

Wstępne badania dotyczyły zarażenia wirusem brodawczaka ludzkiego (ang. Human Papilloma Virus, HPV). Wirus może zakażać okolice narządów płciowych wywołując brodawki - kłykciny kończyste (głównie typy 9 i 11). Zakażenie HPV wiąże się z występowaniem raka szyjki macicy (głównie typy 16 i 18). Jego epidemiologia została najlepiej opisana i zamodelowana w ostatnich latach na całym świecie. Na podstawie wstępnych analiz wynika, że należy ponownie rozważyć obowiązkowe szczepienia, które są jedynym gwarantem zastopowania wzrostu przypadków raka w przyszłości. Na podstawie naszych analiz wnioskujemy, że jeżeli tempo wzrostu ilości nowych partnerów utrzyma się na tym samym poziomie, przy tym samym poziomie świadczeń medycznych, nastąpi ponowny wzrost zachorowań na raka w perspektywie najbliższych 20 lat. Najważniejszą techniką wykorzystywaną było testowanie różnych scenariuszy, to znaczy szczepień, programów przesiewowych, czy akcji profilaktycznych.

## 8 Ewolucyjna Teoria gier

Współpracować, czy zdradzać: na te pytania można odpowiedzieć wedle pewnych matematycznych reguł. Wyodrębnia się wiele rodzajów gier od jednokrotnych po powtarzalne (w których można np. adaptować strategię na podstawie historii). Ciekawym przypadkiem problemu początkowo rozważanego przez fizyków jest gra mniejszościowa, gdzie wypłata przy wyborze jakiejś konkretnej strategii drastycznie spada, gdy zbyt wielu innych graczy wybierze taką strategię. Jednak podstawową grą jaką wymienia się w teorii gier jest dylemat więźnia i tym zagadnieniem zajmiemy się szerzej w tym rozdziale. Mamy dwóch. Wedle klasycznej teorii gier wzajemna współpraca nigdy nie jest optymalna z indywidualnego punktu widzenia, ale zawsze opłacalna dla społeczeństwa.

Jeśli więzień A się przyzna (zdradzi swojego współpracownika), a więzień B zaprzeczy (będzie współpracował), to więc więzień A będzie występował w roli świadka przeciwko B i pójdzie na krótko do więzienie ( $k$  lat), natomiast wówczas B więzień dostanie pełny wyrok ( $p$  lat) więzienia (i vice versa). Jeżeli obaj zaprzeczą (będą współpracować), nie zostaną skazani (0 lat). Jeżeli obaj się przyznają (zdradzą się) obaj otrzymają ten sam średni wyrok ( $s$  lat).

Kolejność wypłat dla takiej gry będzie w następującej kolejności od najbardziej opłacalnej  $0 \text{ lat} < k \text{ lat} < s \text{ lat} < p \text{ lat}$ .

Taki układ posiada punkt równowagowy, skąd zdradzanie jest najbardziej optymalną strategią.

Teoria gier poza swoim matematycznym rodowodem, posiada swój załazek fizyczny. Układ wypłat i kar stosuje się z powodzeniem w opisie procesów interakcyjnych. W tym wypadku jest to gra wieloosobowa, gdzie pewne kategorie cząstek mogą preferować pewne strategie. W zjawisku krystalizacji, wyróżnia się jądro krystalizacji i agenci w jego obrębie mają silne powinowactwo do stanu stałego (wypłata za taki wynik jest dodatnia) oraz negatywne do płynnego (wypłata ujemna, jeżeli taki jest wynik gry). Odpowiednio manewrując między wielkościami wypłat i proporcjami poszczególnych frakcji otrzymujemy opis takiej gry (Gubceac, Gutu, Paladi, 2013).

### 8.1 Próba zamodelowania korupcji w piłce nożnej – teoria gier

Spróbujmy do rozgrywki ligowej podejść w ten sposób, że kolejne mecze są sekwencjami gier (nie tylko sportowych, ale również ekonomiczno-politycznych). W pojedynczym spotkaniu drużyna może wybrać 2 strategie: zostać donorem (D) albo akceptorem (A) łapówki. Przypadek meczu uczciwego opisywałaby gra kiedy zawsze obie drużyny, chciały by być akceptorem łapówki i nie dochodziłoby do transakcji, a o wyniku spotkania decydowały by tylko czynniki sportowe. W przypadku wyboru strategii A przez jedną drużynę oraz D przez drugą dochodzi do transakcji. Kiedy natomiast obie drużyny chciały by dać łapówkę może dojść do sytuacji kiedy jedna drużyna kupuje mecz np. u sędziego i transakcja nie zachodzi między obiema drużynami, a tylko między jedną z nich a np. sędzią (dla ułatwienia przyjmijmy, że tylko gospodarz może przekupić sędziego). Przypadkowi D-D można by było przypisać również sportowe rozstrzygnięcie (obie drużyny, chcą wygrać) i wtedy interpretowany byłby tak samo jak A-A. Dla konsystencji modelu przyjęto jednak, że jeśli drużyny wybierają strategie dania łapówki, to i tak ją dadzą jak nie innej drużynie, to sędziemu. Skonstruowana wedle powyższych założeń macierz wypłat bazuje na mierze użyteczności różnych efektów. Dlatego musimy sprowadzić do wspólnych jednostek, takie czynniki jak: sportowa ilość strzelonych bramek z modelu stochastycznego (w zależności od pozycji drużyny w lidze), kwota pieniężna łapówki, zysk lub strata w punktach, czy koszt moralny uczestniczenia w procederze korupcyjnym.

$\frac{1}{2}$	D	A
D	$(W1-L, W2+G-C-M)^*$	$W1+G-C-M, W2-L+C-M$
A	$W1-L+C-M, W2+G-C-M$	$W1, W2$

$$w2 = \Pi(a r_{ij} + b m_i + c)$$

$$w1 = \Pi(a r_{ji} + c)$$

Tabela 3 Macierz wypat \* (dla ułatwienia przyjmijmy, że tylko gospodarz może przekupić sędziego, choć w ogólności może to być drużyna, której bardziej może zależeć na zwycięstwie, bądź ma więcej pieniędzy)

Rozwijając symbole (Tabela 3):

W1 i W2- Użyteczna wartość stochastycznych funkcji zdobycia bramek w meczu w1-gospodarz i w2-gość (zmienne wewnętrzne: różnica w sile drużyn  $r_{ij}$  ze swoistym współczynnikiem wpływu  $\alpha$ , przewaga gry u siebie  $m_i$  ze swoistym współczynnikiem wpływu  $b$  oraz czynnik losowy  $c$ )

G- Użyteczna wartość zysku w punktach

L- Użyteczna wartość straty w punktach

C- Użyteczna wartość kwoty pieniężnej łapówki

M- Użyteczna wartość kosztu moralnego

Rozumowanie jest zgodne z wytycznymi teorii gier, której autorstwo przypisuje się Johnowi von Neumanowi w roku 1928, aczkolwiek równolegle problem opisał polski matematyk związany z Wrocławiem i Lwowem: Hugo Steinhaus. W tym modelu graczami są kluby, a strategiami gry są D i A. Cały sezon odpowiadałby ewolucyjnym grom, a wtedy parametry modelu zależały by od czasu.

Niestety, powyższa analiza ma tylko charakter orientacyjny i jej przeprowadzenie dokładniejsze nie mieści się w zakresie tej książki. Gdyby jednak udało się znaleźć sposób na sprowadzenie wszystkich czynników do jednostek użyteczności (o czym w ekonomii klasycznej i neoklasycznej dużo zostało powiedziane), można by było poszukać równowag (typu Nasha lub Min-Max) i porównać otrzymane strategie. Zaznaczamy, że teoria gier wielokrotnie przydała się już do opisu zjawiska korupcji, a umiejętne jej wykorzystanie dało by więcej świat na same mechanizmy korupcyjne, a stąd już tylko krok do zwalczania tego procederu.

## 8.2 Gry ewolucyjne i mniejszościowe

W przypadku gier powtarzanych gracze często dynamicznie ustalają reguły. Stosując równowagowe strategię niestety dla powtarzalnych gier najbardziej racjonalną zasadą jest konsekwentna zdrada. W związku z zyskiem osiąganym przy obopólnej współpracy pewne mieszane strategie o pewnych cechach (np. przebacząca) również spisują się dobrze (Axelrod, 1981). Próbę klasyfikacji obserwowanych podstruktur sieciowych gdzie łamana jest permanentnie równowagowa strategia została podjęta przez antropologów (Nowak, 2006). Ze względu na funkcjonowanie społeczeństwa w relacjach pokrewieństwa jak i wewnątrz grupy/klastra współpraca jest nadreprezentowana. Konkurencja między grupami powoduje wzrost liczby zdrad na styku oddzielnych grup. Ze względu na wzajemność, jednostka często przybiera postawę współgracza. Reputacja jednostki również wpływa jak inni będą z nią grać.

Jednym z ciekawych zastosowań adaptacyjnych układów są tak zwane gry mniejszościowe. Załóżmy, że chcemy iść do dobrego baru. Jeżeli będzie on spełniał oczekiwania użytkowników, to z pewnością będzie się cieszył dużą popularnością. Gdy jednak zacznie przychodzić do niego za dużo ludzi, to komfort będących w środku będzie spadał. Dlatego optymalną sytuacją dla klienta jest taka, gdy nie będzie on jeszcze zbyt popularny, a gdy przekroczy pewien próg popularności, to należy zaprzestać chodzenia tam. Stąd nazwa gra mniejszościowa, ponieważ bycie w tej mniejszej grupie która ma dostęp do powszechnego dobra jest właśnie celem wielu jednostek (Brian, 1994). W tej grze stosowanie czystej strategii (np. podążanie za tłumem) nie jest optymalne. W związku z tym wyróżnia się szereg kategorii ludzi w zależności od ich ról w danym procesie: Innowatorzy, wczesni naśladowcy, wcześniejsza większość, późniejsza większość i na końcu maruderzy.

## 8.3 Teoria gier dla graczy postępujących wedle norm społecznych

Odnosząc się do opozycji archetypów graczy wyróżniamy 2 podstawowe koncepcje:

- Homo Economicus – istota racjonalna kierująca się własną korzyścią

- Homo Sociologicus – istota postępująca według norm społecznych

Dylemat więźnia rozumiany jak zwykle według teorii racjonalnego wyboru (ekonomicznie), lecz my rozważymy ten wybór normatywnie (socjologicznie). Formalny zapis norm jest stosowany w literaturze nauk społecznych, dla przykładu:

„Mówimy że R jest w populacji P normą społeczną, jeśli istnieje wystarczająco duża część populacji  $P_c$  należąca do P taka, że dla każdej jednostki „i” należącej do  $P_c$  zachodzi:

okoliczność (contingency): „i” wie, że reguła R istnieje i stosuje się do sytuacji typu S preferencja warunkowa (conditional preference): „i” woli przestrzegać R w sytuacji S pod warunkiem że

(a): spodziewanie empiryczne (empirical expectations): „i” wierzy że wystarczająco duży podzbiór P stosuje R w sytuacji typu S

(b): spodziewanie normatywne (normative expectations): „i” wierzy że wystarczająco duży podzbiór P spodziewa się, że „i” zastosuje się do R w sytuacji typu S

(b’): spodziewanie normatywne z sankcją (normative expectations with sanctions): „i” wierzy że wystarczająco duży podzbiór P spodziewa się, że „i” zastosuje się do R w sytuacji typu S, woli aby „i” się zastosował, i może wyrzucić sankcję.”

[Fragment (Bicchieri, 2006): The Grammar of Society (tłum. K. Kułakowski)]

Zmiana społeczna w wyniku zmiany normy jest znanym zagadnieniem. W latach 90tych XX wieku w Nowym Jorku problem przestępczości był bardzo dokuczliwy dla mieszkańców. Jednym z zastosowanych narzędzi psychologicznych w celu zapobieżenia takim zdarzeniom było wprowadzenie zakazu (a właściwie respektowanie tej normy) malowania graffiti w pociągach metra. W efekcie przestępczość spadła (zaobserwowaliśmy zmianę społeczną). Władze miasta zaingerowały w normę dotyczącą przyzwolenia na wandalizm (w postaci umiarkowanej, czyli graffiti), co poskutkowało zmianą społeczną w ważnym dla społeczności wymiarze bezpieczeństwa. W przeciwieństwie do typowych analiz społecznych jesteśmy zainteresowani dynamiką procesu a nie jedynie stanem końcowym. W tym celu prezentujemy zerowy model z minimalnymi założeniami - zmienne charakterystyki graczy: reputacja i altruizm.

Rozważmy grę, gdzie agenci grają w parach i za każdym razem mają 2 możliwości: współpracować (C), bądź oszukiwać (D). Prawdopodobieństwo współpracy zależy liniowo od altruizmu gracza ( $i$ ) oraz reputacji współgracza ( $j$ ) w oczach ( $i$ ):  $P(i,j) = W_j(i) + \epsilon_i$

Jeżeli  $P(i,j) > 1$  to przyjmujemy 1, a kiedy  $P(i,j) < 0$  to 0. Reputacja  $W$ - jest w przedziale  $[0,1]$ . Altruizm  $\epsilon$ - jest w przedziale  $[-1/2, 1/2]$ . Ogólna zasada:

- reputacja gracza rośnie (spada) jeśli on współpracował (oszukiwał);

- altruizm gracza rośnie (spada), jeżeli jego współgracz współpracował (oszukiwał).

Szybkość zmian jest definiowana przez  $x_w/x_\epsilon$  jako procentowa zmiana reputacji / altruizmu.

$$(C) \quad \epsilon := (0.5 - \epsilon)x_\epsilon + \epsilon$$

$$(D) \quad \epsilon := \epsilon + (-0.5 - \epsilon)x_\epsilon$$

$$(C) \quad W := (1 - W)x_w + W$$

$$(D) \quad W := W - W x_w$$

W grze zakładamy, że bierze udział 100 graczy (czasami 1000) z pewnym początkowym rozkładem  $W$  i  $\epsilon$ . Będą to rozkłady jednorodnie z ustaloną szerokością połówkową 0,5. Wyróżniamy możliwe strategie graczy:

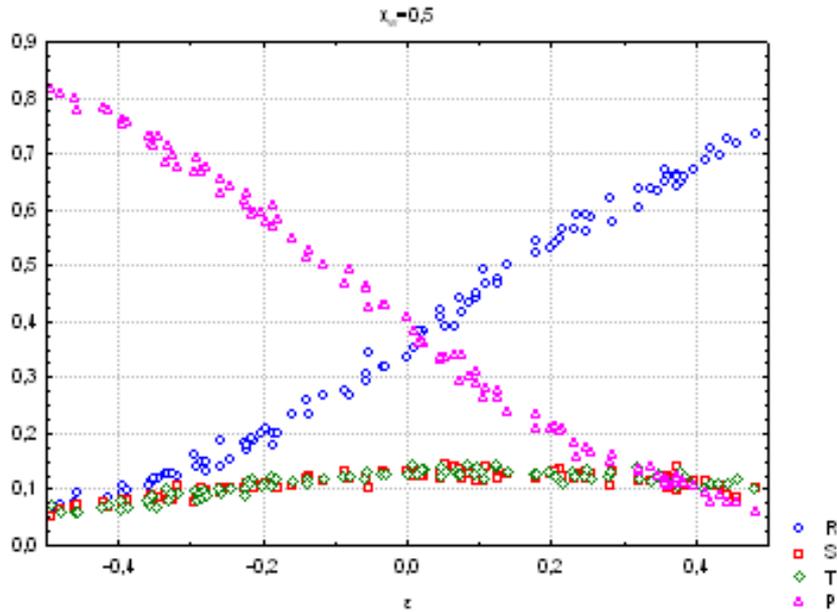
R- obaj współpracują

S- współgracz oszukiwał, kiedy gracz współpracował

T- gracz oszukiwał, kiedy współgracz współpracował

P- obaj oszukują

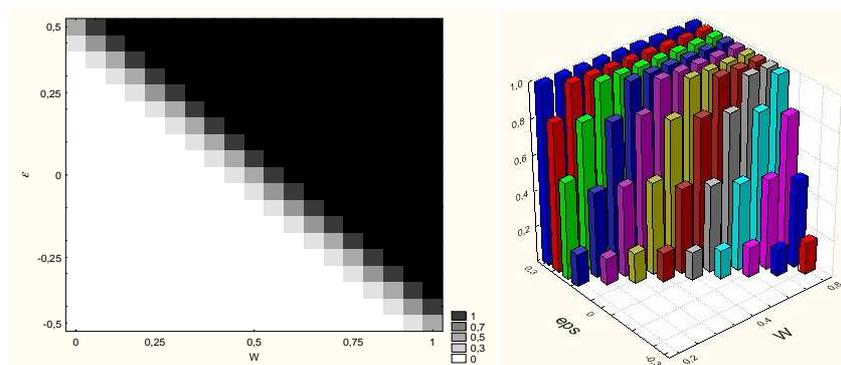
Założmy wstępnie, że altruizm jest stały, a jedynie w trakcie gry zmienia się reputacja graczy zgodnie z przyjętymi wyżej zasadami. Dotyczy to sytuacji, kiedy altruizm jednostki jest jego cechą wrodzoną i nic nie może na nią wpłynąć ani jej zmienić. Dodatkowo reputacja gracza jest jego globalną wielkością (jest taka sama w oczach wszystkich graczy). Dla grupa przypadków dla stałego altruizmu mamy do czynienia z agregacją w formy strategii.



Rys. 43 Grupa przypadków dla stałego altruizmu  $x_w=0,5$  (zmiana reputacji o połowę)

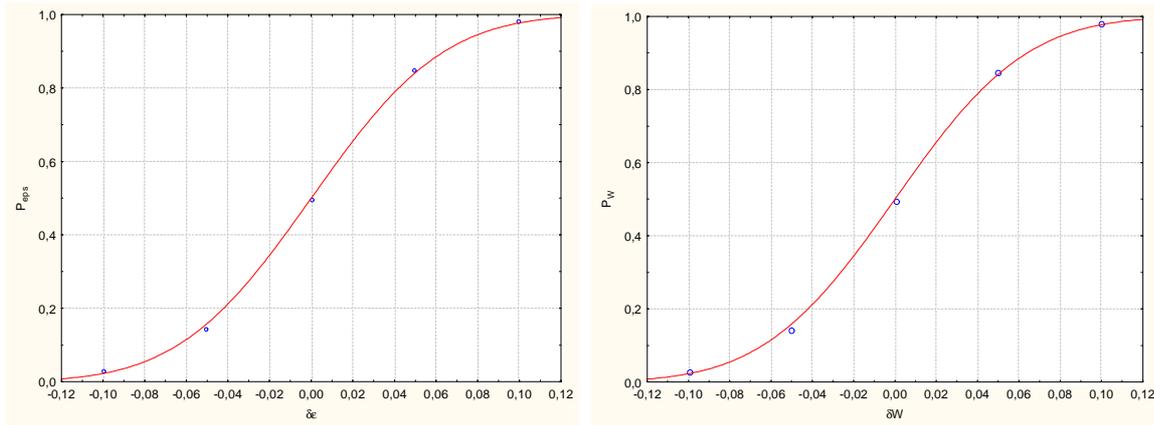
Jakkolwiek dominujące strategię to strategię czyste (P dla osobnika o słabym altruizmie oraz R dla takiego o wysokim altruizmie), jednak strategię mieszane (T, S) ciągle stanowią około 20% wszystkich decyzji oraz są stosunkowo niezależne od altruizmu jednostki.

Rozważmy możliwość adaptacji altruizmu, w związku z tym, że altruizm w rzeczywistości nie jest cechą stałą i zależy od wielu czynniki, np. czy skierowany jest do członków swojej społeczności, czy na zewnątrz oraz jednostki dopierają swoją postawę moralną do zachowań otoczenia. Kiedy  $x_w$  i  $x_e$  są niezerowe to zanikają różnorodne strategię. Zaczynając od przypadku, kiedy po każdej kolejce następuje zmiana reputacji i altruizmu o  $\frac{1}{2}$  otrzymamy:



Rys. 44 Rozdzielenie faz czystych współpracujących i zdradzających dla zadanych warunków początkowych, z fazą mieszaną w okolicy antydiagonalnej

Praktycznie dla wszystkich warunków początkowych układu w stanie końcowym przybiera w całości formę, że wszyscy zdradzają lub współpracują -białe i czarne (Rys. 44).



Rys. 45 Granica faz opisywana dystrybuantą rozkładu Gaussa dla zmian reputacji oraz zmian altruizmu.

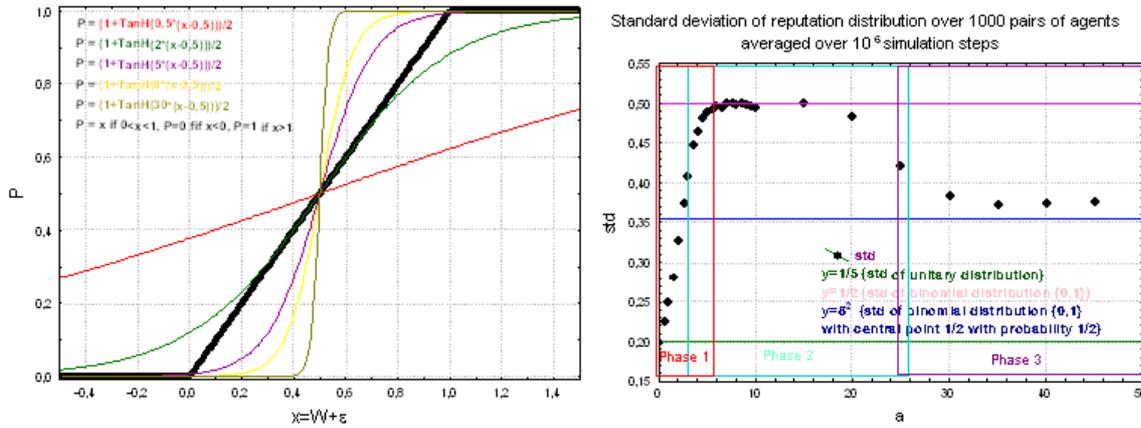
W obszarze granicznym mamy do czynienia z tak zwanym przejściem fazowym (Rys. 44, Rys. 45). Układ znajduje się w tym obszarze, kiedy średnie wartości początkowe reputacji i altruizmu znajdują się na antydiagonalnej (Rys. 44), i strategie wybierane przez graczy nie są jednorodne przez dłuższy czas, aż w nieskończoności układ stabilizuje się do jednej z 2 dozwolonych strategii czystych. Właśnie pobudzając układ, bądź poprzez samoistną ewolucję (Rys. 45) układ posuwa się w stronę jednorodnego.

Próbując socjologicznie zinterpretować wyniki musimy, zacząć że przy tak postawionym problemie strategie zdradzające i współpracujące są równouprawnione i nie został wdrożony żaden dryft w żadną stronę. „Uzmiennienie” altruizmu stało się przyczyną rozdzielenia się faz współpracujących i oszukujących. Dynamika w obszarze o niedeterministycznym stanie końcowym (w okolicy granicy faz) jest odzwierciedleniem rzeczywistych procesów społecznych. W związku z tym, socjotechniczne zabiegi w przypadku oscylowania układu w okolicy stanu granicznego, mogą dawać efekty i regulacja układu jest możliwa.

Stochastyczne własności w rodzaju przeskoków stanów jednorodnych wydają się ciekawe z naukowego punktu widzenia (na ile prawdopodobna jest zmiana zachowania społecznego). Proponujemy wprowadzenie innej funkcji prawdopodobieństwa współpracy. Prawdopodobieństwo współpracy zależy jak tangens hiperboliczny (funkcja sigmoidalna) od altruizmu gracza ( $i$ ) oraz reputacji współgracza ( $j$ ) dla altruizmu neutralnego (równego zero):

$$P_i(W_j, a) = \frac{1 + \tanh(a(W_j - 1/2))}{2}$$

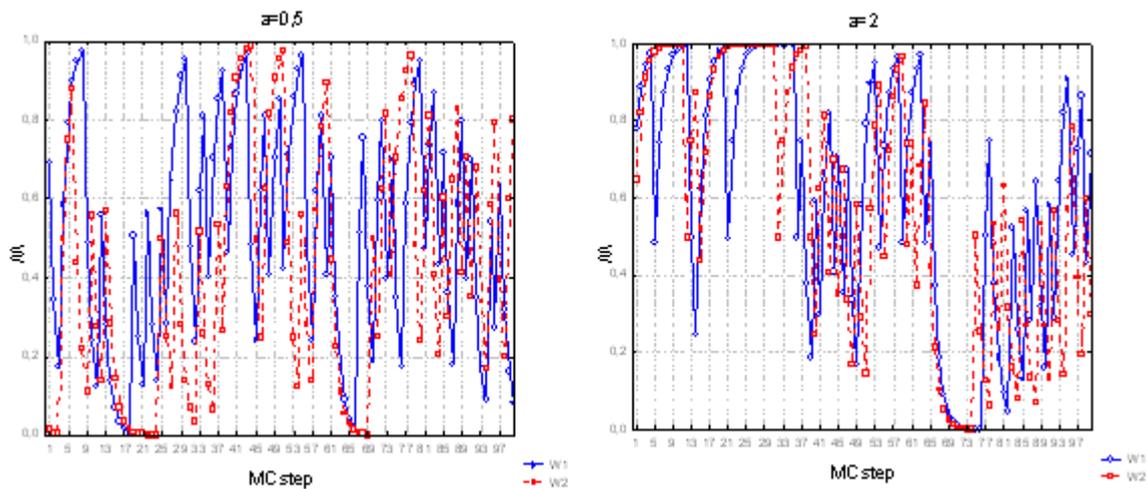
Gdzie  $a$  jest parametrem krzywizny funkcji aktywacji współpracy. Zmianom zachowania systemu dla tego parametru zajmiemy się w tym dziale.



Rys. 46 Przebieg funkcji tanh dla różnych parametrów  $a$  i standardowej funkcji liniowej (lewy) i odchylenie standardowe dla różnych wartości parametru  $a$

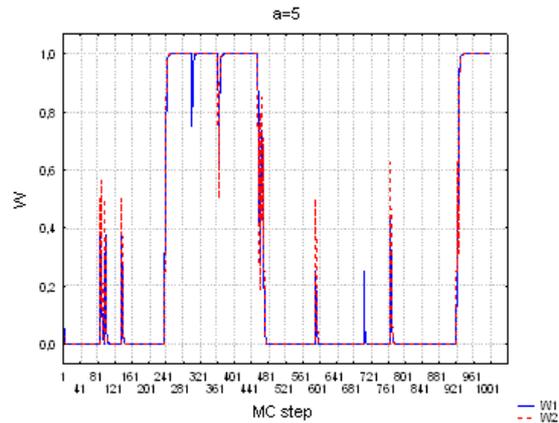
Wydzielone zostały różne obszary zachowań funkcji aktywacji (prawdopodobieństwa) współpracy i dokonana została klasyfikacja tych faz. Należy pamiętać jedynie zwycięskie w sensie Axelroda, ponieważ nie odzwierciedlają one wszystkich obserwowanych rzeczywistych sytuacji, a powyższa funkcja pozwala na obserwację większej liczby potencjalnych ludzkich zachowań. Stąd odniesienie funkcji prawdopodobieństwa do funkcji aktywacji oraz parametru  $a$  jako szumu informacyjnego. Przedstawiono tutaj schematy zachowań dla pary graczy o zmieniających się reputacjach  $W_1$  i  $W_2$ .

**Faza 1** dla  $a < 5$  Strategie przypadkowo (schizofreniczne zachowanie lub brak pamięci jak u dzieci). Z jednego kroku w drugi następuje często biegunowa zmiana zachowania (stąd porównanie do schizofrenii).



Rys. 47 Dryf przypadkowy strategii

**Obszar między fazami 1 i 2** dla  $a \sim 5$ , krótkotrwałe ustalanie strategii (częste zmiany spójnych strategii typu Dr. Jekyll and Mr. Hyde). W tym wypadku nasi aktorzy mają okresy w których ze sobą współpracują, po czym zmieniają strategię, która znowu utrzymuje się jedynie przez krótki okres czasu.



Rys. 48 Strategie przeskokowe, gdzie przez jakiś czas gracze się zdradzają, by później współpracować dla parametru  $a=5$

**Faza 2** dla  $5 < a < 25$  Wspólne strategie dominują z rzadkimi przeskokami (jak w przypadku przewartościowania moralności niemieckich nazistów).

**Faza 3** dla  $a > 30$  Układ staje się deterministyczny.

Została więc nadana interpretacja faz dla różnych wartości parametru  $a$  jako pewnych kontekstów społecznych. W tym modelu nacechowanie współpracy i zdrady nie występuje. Jeżeli któryś ze stanów jest premiowany, to jest nim współpraca (jako norma społeczna).

Na gruncie prostych gier wiele społecznych zjawisk jest niewytłumaczalnych. Układ graczy, w którym jeden permanentnie oszukuje a drugi współpracuje jest bardzo trudny do zamodelowania a występuje np. w rzeczywistych układach z powinowactwem. Spójrzmy na model, w którym reputacja jest częściowo indywidualna, a częściowo wspólna (rozszerzona), a dokładniej gracz  $i$  ma możliwość modyfikowania wizji gracza  $j$  według pewnych cech swoistych dla gracza  $i$ . Taki zabieg został wprowadzony ze względu na jedynie pośrednią możliwość zmiany reputacji gracza  $j$  przez gracza  $i$ . W klasycznym modelu „wykorzystywanie” może mieć miejsce w niszach społecznych, w których interesujący nas gracze muszą być „zanurzeni w kąpiel termicznej”, każdy w przeciwnym kierunku. Pomysł „rozszerzenia” reputacji nie wymaga takiego mechanizmu, ale zakłada wprowadzenia czegoś takiego jak „pole zewnętrzne”. Działa to w ten sposób, że na gracza  $i$  „coś” wywiera wpływ tak, że modyfikowana jest reputacja gracza  $j$ . Kontakt z rezerwuarem współpracy czy zdrady można również rozumieć w ten sposób. Cała zmiana sposobu myślenia polega na zróżnicowaniu między realnym a wirtualnym oddziaływaniem. Tak więc gracz  $j$  zanurzony w kąpiel termicznej reprezentuje w kontakcie z rezerwuarem odpowiedniego usposobienia tylko wirtualne oddziaływanie, a realnie ów rezerwuar działa na gracza  $i$ . Dlatego rezerwuar zanurzenia gracza  $j$  łączymy więzią rzeczywistą z graczem  $i$  (w końcu reprezentują te same usposobienia). Oddziaływanie z rezerwuarem może przyjmować różne formy. Może to być normalna gra, z tym że gracze nie mają wpływu na rezerwuar. Tak więc częstotliwość kąpienia gracza  $i$  w rezerwuarze oznacza tak naprawdę podatność gracza  $j$  na zewnętrzne czynniki. Przykład. Niech gracz  $i$  jest osobą ze wspólnoty religijnej, w której propagowana jest wizja miłości do bliźniego. Ten sam gracz  $i$  ma do czynienia w życiu z graczem  $j$ , którego sposobem na życia jest oszustwo. Wynika z tego opisu, że gracz  $i$  ma wysoki, a  $j$  niski altruizm, więc w normalnej grze wzajemna strategia wyzysku zdarzałaby się maksymalnie z 50% częstością. Pozwalając graczowi  $i$  być podatnym na naukę „kościół” reputacja gracza  $j$  może wzrosnąć poprzez zanurzenie w kąpiel, co zwiększy naiwność gracza  $i$ , a co za tym idzie zwiększy częstość występowania wyzysku nawet do 95% (przy założeniu, że kąpiel jest też grą).

## 9 Modelowanie ruchu

Ludzie przemieszczają się, co jest podstawowym czynnikiem tworzącym cywilizację. Istotę modelowania w tej dziedzinie wprowadziliśmy już we wstępnej liście najważniejszych zagadnień obliczeniowych nauk społecznych, zwłaszcza że wyniki tych modeli jako jedne z pierwszych znalazły praktyczne zastosowania. Mamy tu do czynienia z literalnym wykorzystaniem praw fizyki (zapisanych wprost, a nie jako analogie) do opisu zjawisk społecznych. W wielu krajach na etapie przyjmowania budynków użyteczności publicznej od dewelopera sprawdzana jest zgodność z zadanymi normami (obliczenia czy dane parametry korytarza są wystarczające, do przeprowadzenia ewakuacji sumy maksymalnej pojemności danej części budynku). Sygnalizacja świetlna jest konfigurowana pod kątem przepustowości ruchu. Omówimy tu przykłady ruchu pieszego (np. w kontekście ewakuacji obiektów) jak również ruchu ulicznego ( np. nawiązując do korków. W takich modelach wykorzystuje się kilka podstawowych założeń. Jednostki zachowują bezpieczną odległość od sąsiadów, posiadają docelową prędkość i kierunek, ale dopasowują się do w pewnych warunkach do prędkości i kierunku ruchu do sąsiadów. Ruch jest jednym z niewielu przypadków, kiedy eksperymenty społeczne są wykonywane na szeroką skalę, ponieważ nie ma tu zbyt wiele problemów natury etycznej, bo większość takich badań polega na umieszczeniu grupy ludzi w odpowiednich warunkach topologicznych . Tam mogą poruszać się permanentnie (zazwyczaj ochotnicy proszeni są o zachowywanie się -jazdę lub chodzenie- w sposób naturalny), bądź rozpoczynają ruch ze stanu spoczynku (co ma odzwierciedlać ewakuację). Tylko w niektórych eksperymentach element paniki jest wprowadzony w postaci rywalizacji (nagrody dla uczestników badań, którzy dotrą najszybciej do celu).



Rys. 49 Wizualizacja ruchu ludzi. Źródło: Robert Hillman/Bigstockphoto

### 9.1 Modele ewakuacji

Do opis zachowania ludzi w czasie ewakuacji, niezwykle istotnego z punktu widzenia życia ludzi, zostały zaproponowane różne modele oraz algorytmy symulacji, np. prawa fizyki opisujące dynamikę płynów poprzez analogię pomiędzy zachowaniem ludzi oraz cząstek w cieczech. Wzrost wiedzy o przebiegu procesów ewakuacji i jej upowszechnienie będzie miało efekt na ograniczenie ryzyka w procesie ewakuacji i zmniejszenie jej czasu (Matuszak, 2010).

Na początku można zapoznać się z opisem nierównowagowych układów wielocząsteczkowych z siłami kierującymi do czego wprowadza się siłę samonapędzającą. Wynikiem jest równanie różniczkowe ruchu podobne do równań ruchu znanych z licealnych kursów dynamiki:

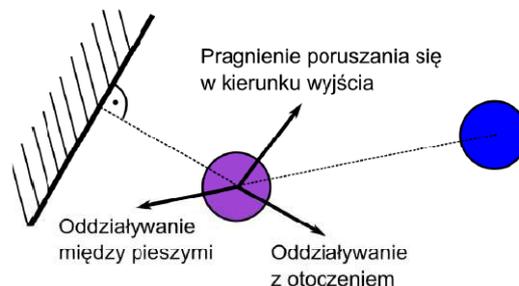
$$m_i \ddot{x}_i = F_0 - \gamma \dot{x}_i + \sum F_j + \xi_i$$

gdzie  $i$  – oznacza interesującego nas agenta a  $j$  jego sąsiadów,  $m$ - masa,  $x$ - położenie,  $x'$ - prędkość oraz  $x''$  przyśpieszenie,  $\gamma$ - tarcie,  $F_\sigma$ - siła zewnętrzna,  $F_f$ - siła nacisku od sąsiada  $j$ ,  $\xi$  – przypadkowa siła.



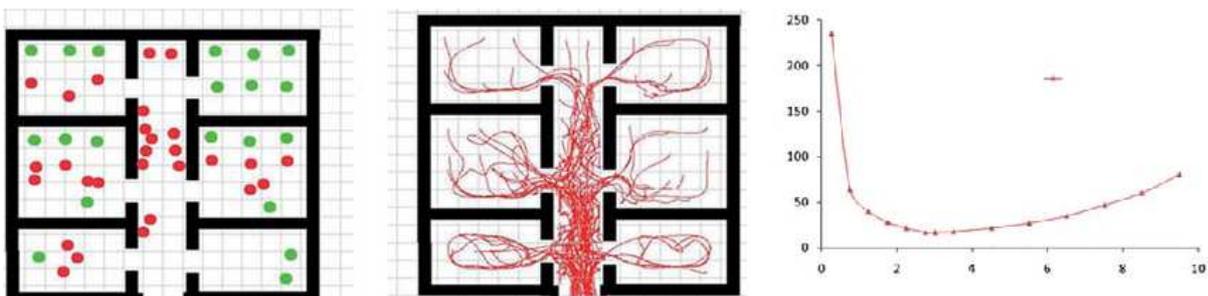
Rys. 50 Ewakuacja obiektu sportowego w warunkach eksperymentalnych. Źródło: <http://www.fz-juelich.de/.../PedestrianDynamics/node.html>

Wiedza wynikająca z modelowania oparta jest w głównej mierze na modelu Helbinga (Helbing, 1994). Zakładając połączenie oddziaływań o charakterze socjologicznym oraz fizycznym (jak wzór powyżej): każda z  $N$  cząsteczek (o indeksie  $i$ ) mająca masę  $m$  dąży, aby poruszać się z wcześniej znaną prędkością docelową  $v_{oi}$ , w określonym kierunku. Będąc w grupie  $i$ -ta cząstka stara się dopasować swoją prędkość do prędkości sąsiadów (poprzez oddziaływania z otaczającymi cząsteczkami). Jednocześnie stara się utrzymać stałą odległość od sąsiadów oraz ścian (siła socjologiczna). Siła socjologiczna to nic innego jak efekt naruszenia przestrzeni osobistej przez otoczenie (sąsiadów, bądź ściany). Wspomniana wcześniej „siła kierująca” zdefiniowana jako kierunek w którym jednostki się poruszają. Prędkość docelowa, jest to maksymalna prędkość jaką może osiągnąć dany człowiek. Każdy krok czasowy symulacji zwiększa (lub zmniejsza) aktualną prędkość poprzez oddziaływania oraz zmienia kierunek.



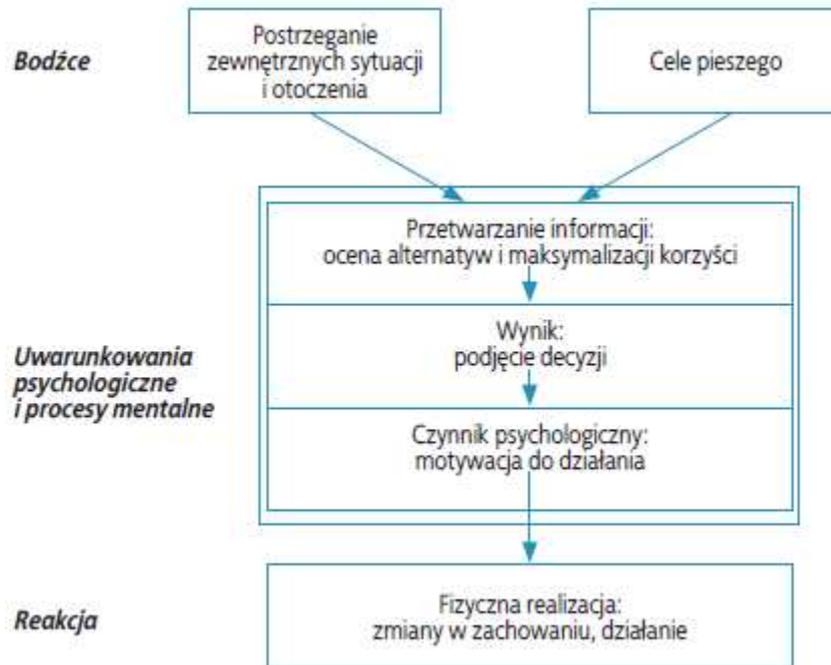
Rys. 51 Przykładowy układ sił społecznych w modelu ewakuacji. Źródło: (Lubas, Wąs, Porzycki, & Mycek, 2013)

Tak zdefiniowany model pomaga wizualizować empiryczne efekty obserwowane w ewakuacjach oraz pozwala wstępnie oszacować czasy ewakuacji dla różnych topologii.



Rys. 52 Ewakuacja biura ze prędkość docelowa  $v_0 = 2$  m/s. Przedstawiono rozmieszczenie pieszych i trajektorie w czasie ewakuacji oraz czasy ewakuacji w funkcji prędkości docelowej. Źródło: (Grabowski & Kosiński, 2013)

Istotną kwestią jest tu tak zwana spójność, czyli sterowanie w kierunku uśrednionego środka lokalnej grupy. Obserwuje się tłumie szeregu zjawisk kolektywnych, które inżynierskim opisem dają się odtworzyć modelami. Zakleszczenia mają miejsce przy zwięzieniach oraz przy napieraniu grup pieszych poruszających się w przeciwnych kierunkach. Ciekawym efektem jest tworzenie się ścieżek ruchu. Dla mas ludzkich poruszających się, tą samą drogą spontanicznie formują się ścieżki. Ten efekt obserwacyjny został potwierdzony eksperymentalnie.



Rys. 53 Schemat postępowania pieszego w czasie ewakuacji. Źródło: (Grabowski & Kosiński, 2013)

## 9.2 Modele ruchu drogowego

Takie czynniki jak szybki rozwój motoryzacji, nasilenie się ruchu na drogach, bariery rozrostu sieci drogowej spowodowały zainteresowanie inteligentnymi systemami wspomagającymi zarządzanie transportu drogowego. Podstawowym modelem wykorzystywanym w opisie ruchu ulicznego jest automat komórkowy (Nagel & Schreckenberg, 1992). Założyli oni, że każdy pojazd ma taką samą, określoną długość, droga jest jednokierunkowa i jednopasmowa, jedna komórka jest zajęta (wartość = aktualnej prędkości samochodu) przez jeden pojazd lub wolna (wartość = 0). Zmiany prędkości i przemieszcza odbywa się według ustalonych reguł. Pojazdy, zwiększają prędkość o jeden o ile jest to zgodne z przepisami ruchu drogowego (ustalona dozwolona prędkość). Jedną jeżeli liczba wolnych komórek przed autem jest mniejsza niż aktualna prędkość to pojazd hamuje. Prędkość pojazdu jest zmniejszana z pewnym prawdopodobieństwem o jeden, co oznacza warunki losowe na drodze. Tak postawiony model daje w efekcie spontanicznie powstające korki. Obserwujemy istnienie dwóch faz ruchu pojazdów. W fazie swobodnej jest zapewniona ciągłość ruchu (dla gęstości pojazdów mniejszej od wartości krytycznej). Druga to faza ruchu zagęszczonego (dla gęstości pojazdów większej od wartości krytycznej) i charakteryzuje się lokalnymi korkami (ciągami unieruchomionych aut).

## 10 Formowanie się opinii

Modelowanie dynamiki opinii jest jednym z centralnych problemów Socjofizyki.

W rzeczywistym świecie możemy obserwować zmiany opinii społeczeństwa dotyczące jakiegoś zadanego tematu. Podejmowane są próby stworzenia modeli odtwarzających owe zmiany. Dzięki nim możliwe staje się zrozumienie procesów za nimi stojących. Próbuje się ich używać w celu opisanie wyników wyborów, preferencji konsumenckich, spraw społecznych itp. Choć problemem jest odniesienie wielu wyników teoretycznych do rzeczywistych danych, to w ostatnich latach sytuacja się powoli zmienia, między innymi dzięki portalom społecznościowym.

Najczęściej wykorzystywane są modele agentowe, w których pojedynczy „agent” jest odpowiednikiem osoby, a cały układ składa się z wielu takich agentów. Agent może mieć przypisane pewne cechy, choć najczęściej ma tylko jedną – opinię. Opinia jest liczbowym przedstawieniem poglądu agenta, w większości przypadków może być:

- liczbą całkowitą – w skrajnym przypadku 0 lub 1
- liczbą ciągłą
- wektorem

Dotychczas powstało bardzo wiele modeli dynamiki opinii. Można wręcz powiedzieć, że mamy do czynienia z prawdziwym zoo. I choć łatwo utonąć w tym gąszczu to na szczęście istnieją pewne proste fundamentalne modele, które pozwalają zrozumieć najważniejsze zjawiska i uchwycić ogólny sens problemu bez tonięcia w szczegółach.

Pewien ogólny schemat można przedstawić następująco:

- Aranżujemy spotkania pojedynczych agentów, lub ich grup – algorytm wybierania agentów do spotkań zależy od modelu
- Następuje dyskusja, w wyniku której agenci mogą zmienić opinię – tutaj również szczegóły interakcji zależą od konkretnego modelu
- Powtarzamy procedurę

Załóżmy, że zaczynaliśmy od stanu, w którym każdy agent miał losowo przydzieloną opinię. W wyniku powyższego postępowania, po wielu powtórzeniach przedstawionej procedury, otrzymamy stan absorbcyjny, czyli taki, w którym nic więcej się nie zmienia. Znaczna część badań koncentruje się wokół badania owego stanu absorbcyjnego, dochodzenia do niego i czasu w którym to się dzieje. Można jednakże nasz schemat rozszerzyć o zaburzenie. To znaczy można zmieniać opinię niektórych agentów w sposób losowy. Wówczas będzie to wyglądało tak:

- Aranżujemy spotkania pojedynczych agentów, lub ich grup – algorytm wybierania agentów do spotkań zależy od modelu
- Następuje dyskusja, w wyniku której agenci mogą zmienić opinię – tutaj również szczegóły interakcji zależą od konkretnego modelu
- Wprowadzamy zaburzenie – szczegóły zależą od konkretnego modelu
- Powtarzamy procedurę

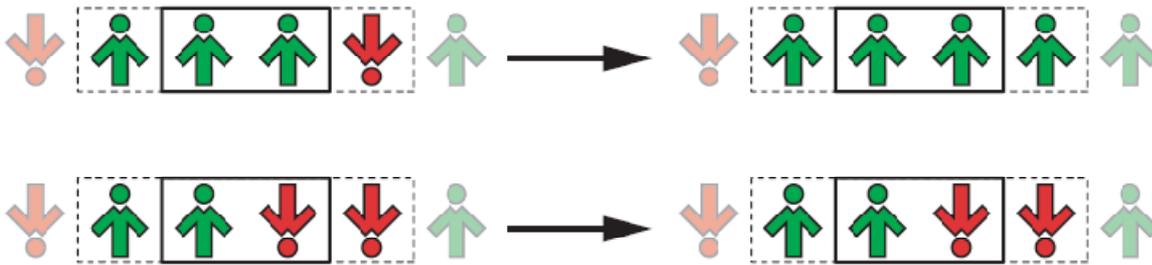
Układ do którego wprowadzono zaburzenie nigdy nie dojdzie do stanu absorbcyjnego, ponieważ zawsze będzie zeń wytrącany. Może za to osiągać stany równowagowe i może również w pewnych sprzyjających okolicznościach przełączać się między różnymi stanami równowagowymi.

### Dwa modele

W dalszej części pokrótce przedstawimy dwa modele, będące już w zasadzie kanonem modelowania dynamiki opinii: model Deffuanta oraz model Sznajdów. Wybraliśmy akurat te dwa ponieważ są proste, ciekawe i prezentują odmienne podejście do tematu.

### Model Sznajdów

Koncepcja modelu Sznajdów jest oparta na fakcie znanym w psychologii społecznej jako społeczny dowód słuszności. Otóż, jeśli jakaś osoba o zdaniu odmiennym od naszego próbuje nas przekonać do swoich racji, to ma przed sobą trudne zadanie. Lecz gdy jednocześnie dwie lub więcej osób, które zgadzają się ze sobą, ale mają inne zdanie niż my i próbują wyrzucić na nas presję, to może im się udać bez porównania łatwiej. Opierając się na tym spostrzeżeniu, Katarzyna Sznajd- Weron wraz z Jozefem Sznajdem stworzyli prosty model dynamiki opinii, który okazał się ogromnym sukcesem. Jest badany do dziś i w dalszym ciągu dostarcza wielu ciekawych wyników. Znalazł zastosowanie w modelowaniu rezultatów wyborów, konkurencji produktów na rynku (o czym podrozdział o oligopolu), oraz tworzeniu nowego rodzaju sieci złożonych. W swojej pierwotnej wersji symulacje były realizowane na zamkniętym łańcuchu jednowymiarowym. Agenci w modelu Sznajdów mogą przyjąć tylko jedną z dwóch możliwych opinii -1 lub +1. Postępujemy zgodnie ze schematem opisanym powyżej, czyli aranżujemy spotkania grup agentów i sprawdzamy czy na drodze dyskusji coś się zmieniło.

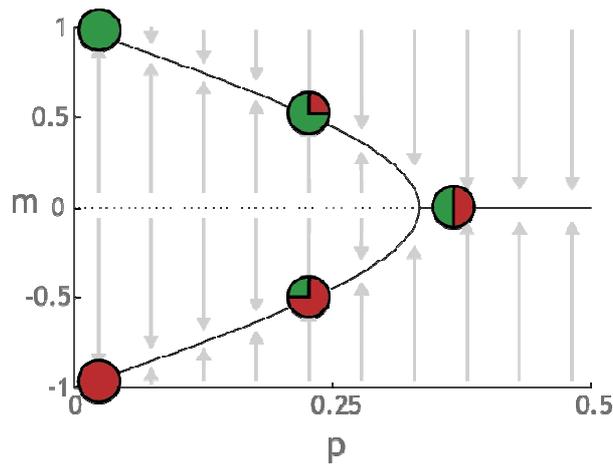


Rys. 54 Jednowymiarowa dynamika w modelu Sznajdów. Agenci zakreśleni pełną ramką wpływają na agentów zakreślonych ramką przerywaną, ale tylko wówczas gdy mają to samo zdanie (górny rysunek). Jeśli agenci należący do grupy wpływu nie są ze sobą zgodni, nic się nie dzieje (dolny rysunek). Na podstawie : (Nyczka, 2015)

- Najpierw wybieramy grupę dwóch sąsiadujących ze sobą agentów (Rys. 54).
- Jeśli są oni ze sobą zgodni, to ich sąsiedzi przyjmują podzielaną przez nich opinię. Jeśli nie, nic się nie dzieje.
- Powtarzamy procedurę

Dynamika taka prowadzi do tak zwanego konsensusu, czyli stanu w którym wszyscy agenci podzielają tę samą opinię.

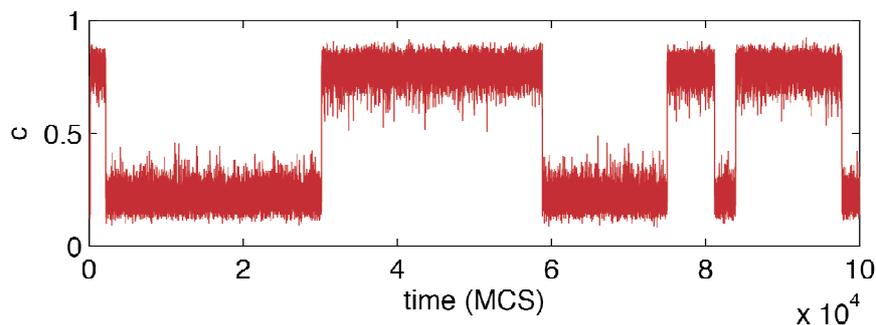
Jeśli jednak wprowadzimy do układu zaburzenie symulujące bądź nie zależność bądź nieposłuszeństwo agentów, sytuacja się zmieni. Dla małych wartości zaburzenia w układzie będzie dominowała jakaś większość opinii, natomiast dla dużych zaburzeń ta większość zanika i opinie rozkładają się mniej więcej po połowie. Zależność asymetrii poglądów od prawdopodobieństwa zaburzenia ilustruje rysunek (Rys. 55). Jak widać asymetria maleje wraz ze wzrostem prawdopodobieństwa zaburzenia  $p$ , aby zniknąć zupełnie po przekroczeniu pewnej wartości  $p^*$  zwanej wartością krytyczną. Mamy więc do czynienia z ciągłym przejściem fazowym.



Rys. 55 Asymetria poglądów od prawdopodobieństwa zaburzenia  $p$

Ponadto w skończonych układach, dla  $p < p^*$ , możemy się spodziewać spontanicznych przeskoków między stanami o przeciwnych znakach (Rys. 56). Te „rewolucje” są tym radsze im większy jest układ i im mniejsze jest zaburzenie.

Widać stąd, że nawet tak prosty zabawkowy model zachowuje się w nietrywialny sposób, występuje w nim samoorganizacja, spontaniczna zmiana stanu oraz przejście fazowe.



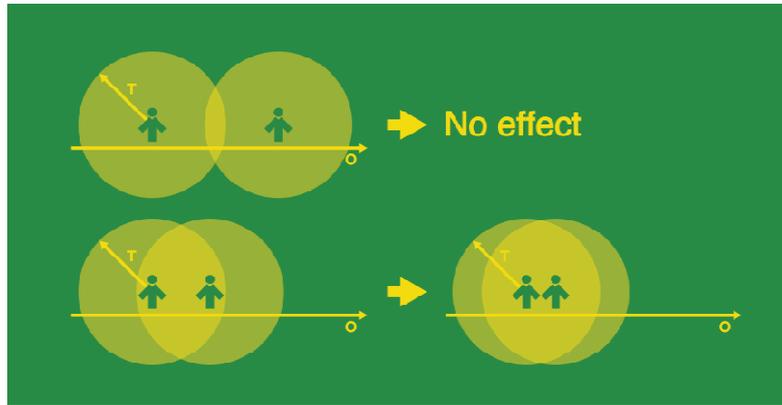
Rys. 56 Ponadkrytyczna ilość zaburzeń powodująca przeskoki

### Model Deffuant

Tutaj koncepcja jest nieco odmienna. Każdy agent może przyjmować opinię z przedziału  $-1, +1$  ale tym razem jest to liczba rzeczywista.

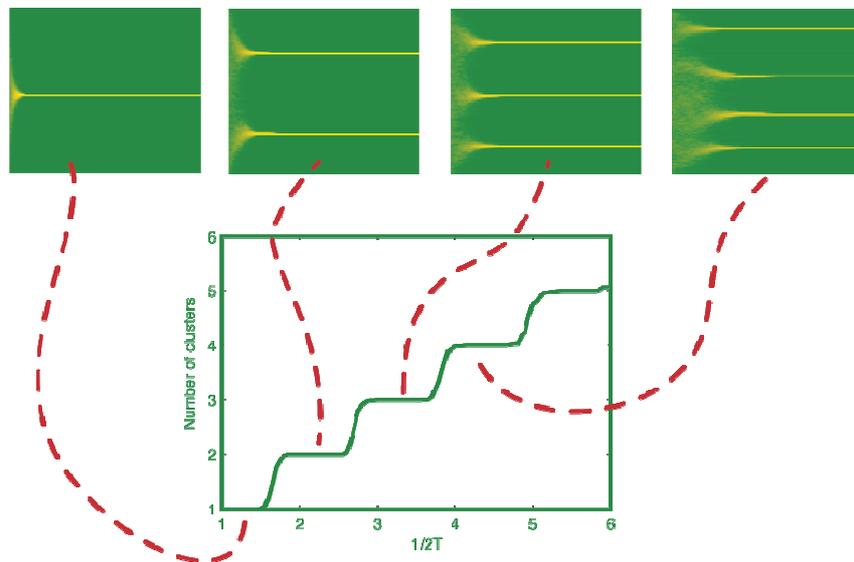
- Aranżujemy spotkania między losowo wybranymi parami agentów.
- Jeśli ich odległość w przestrzeni opinii jest mniejsza niż pewien próg tolerancji  $T$ , to wówczas mogą ze sobą rozmawiać i w wyniku tej rozmowy ich poglądy zbliżają się do siebie wzajemnie (Rys. 57).
- Powtarzamy procedurę

Wprowadzenie parametru  $T$  uzasadniane jest faktem, że osoby o poglądach bardzo odmiennych od naszych, nie są w stanie nas przekonać, natomiast ci nieco bliżsi już tak.



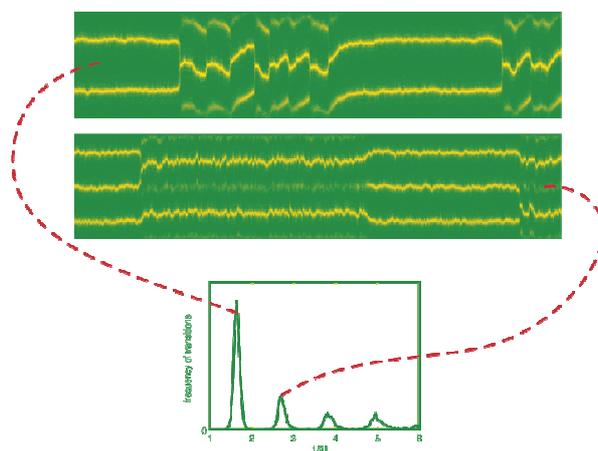
Rys. 57 Zbliżająca się odległość między poglądami

Jeśli zaczynamy od stanu w którym agenci mają opinie wylosowane z jednostajnego rozkładu z przedziału  $<-1, 1>$  to wskutek takiej dynamiki osiągany jest równowagowy stan z jednym, dwoma lub większą liczbą klastrów opinii. Ich liczba zależy od  $T$  i jest proporcjonalna do  $1/2T$  (Rys. 58). Po osiągnięciu tego stanu już nic więcej się nie dzieje.



Rys. 58 Zależność liczby klastrów od wielkości progno tolerancji ( $T$ )

Wprowadzenie losowego zaburzenia nieco zmienia sytuację. Raz na jakiś czas zmieniamy opinię losowego agenta na zupełnie nową losowaną z przedziału  $<-1,1>$ , ma to symulować pewien stopień niezależności agentów. Wskutek tej zmiany, w niektórych obszarach parametru  $T$  możemy się spodziewać spontanicznych przeskoków układu między stanami z liczbą klastrów 1 i 2 (Rys. 59), 2 i 3 (Rys. 59), itd. Ten model również pokazuje że prosta konstrukcja może prowadzić do skomplikowanych zachowań. Niestety nawet w tak prostych przypadkach uzyskiwanie analitycznych wyników jest trudne lub nawet niemożliwe.



Rys. 59 Charakterystyki poszczególnych klastrów

Jak już było wspomniane modeli dynamiki opinii jest o wiele więcej, jednak uważamy że te dwa (jako klasyki) są godne szczególnej uwagi.

## 10.1 Model oligopolu

Ciekawym przykładem pobocznym jest model oligopolu (Pilarczyk, 2009). Konkurencja w ogólności polega na: różnicowaniu produktów, wprowadzaniu innowacji, różnicowaniu cen, jakości warunków sprzedaży oraz reklamie. W przypadku rynków zamkniętych z ograniczoną liczbą graczy jak oligopol produkty są podobne i główny obszar działań producentów to wizerunek (czyli reklama). Podkreśla się wtedy różnice między produktami - czasem tylko pozorne, co ma ułatwić konsumentowi podejmowanie decyzji,

Przykładem mogą być wojny konkurentów typu „Cola Wars” między Pepsicola i Coca-cola. Opisana tu została sytuacja na polskim rynku telefonii komórkowej (Sznajd-Weron, Weron, & Włoszczowska, 2008), dzięki pozyskaniu przez Maję Włoszczowską (złota medalistka olimpijska i matematyk z zawodu) od firmy Expert Monitor danych dotyczących efektywnego wystawienie reklam w mediach oligopolistów. Rynek telefonii komórkowej w Polsce charakteryzował się obecnością tylko 3 operatorów (Polkomtel, PTK Centertel, Polska Telefonia Cyfrowa), wysokimi barierami wejścia, podobnymi cenami, oraz była to najsilniej reklamującą się branżą.

Model właściwy bazuje na modelu Sznajdów dynamika wpływu (typu na zewnątrz) na siatce z trzema stanami (1, 2, 3) gdzie wymiana opinii z otoczeniem grupy oraz przez pole zewnętrzne wkładana jest reklama. Symulacje przebiegają w ten sposób, że losujemy firmę, tworzymy jej docelową grupę konsumentką, wyliczamy opinię średnią tej grupy. Jeżeli nie ma jednogłośności, to możliwe są 2 mody.

1) Konformizm i reklama równoległe: Z prawdopodobieństwem  $p$  stosujemy konformizm, a z prawdopodobieństwem  $(1 - p)$  działa reklama.

2) Konformizm pierwszą: Jeżeli grupa jest jednomyślną przyjmuje się, że wszyscy członkowie tej grupy są konsumentami wybranej firmy (konformizm). W przypadku braku jedności konformizm nie działa. Wtedy z prawdopodobieństwem  $(1 - p)$  działa reklama.

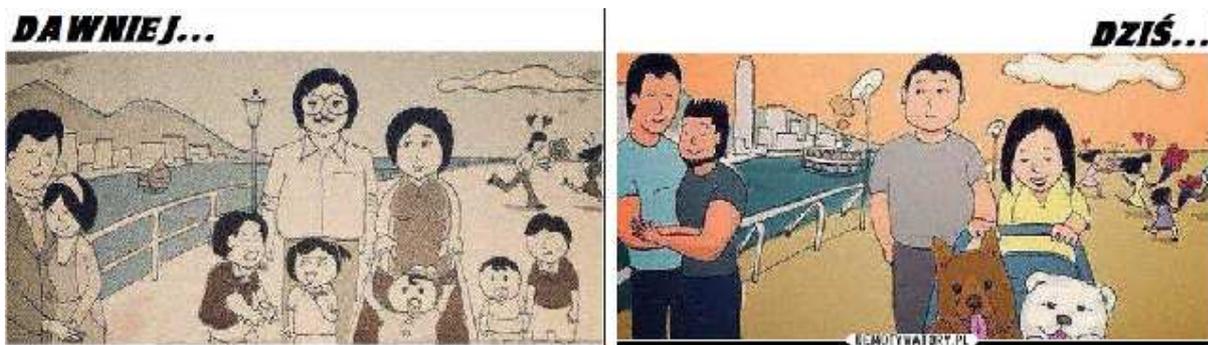
Wybór operatora następuje proporcjonalnie do siły reklamy danego operatora. W wynikach zauważono (niejako wkładając te zmienne do modelu), że na decyzje konsumenta wpływają czynniki lokalne (wpływ społeczny) i globalne (reklama - zewnętrzne pole), a reklama ma w zależności od modu bardzo duży albo decydujący wpływ na wynik firmy.

## CZĘŚĆ III

# ANALIZA DANYCH SPOŁECZNYCH W UJĘCIU MODELOWO - OBLICZENIOWYM

### 11 Modelowanie wejścia i wyjścia z małżeństwa

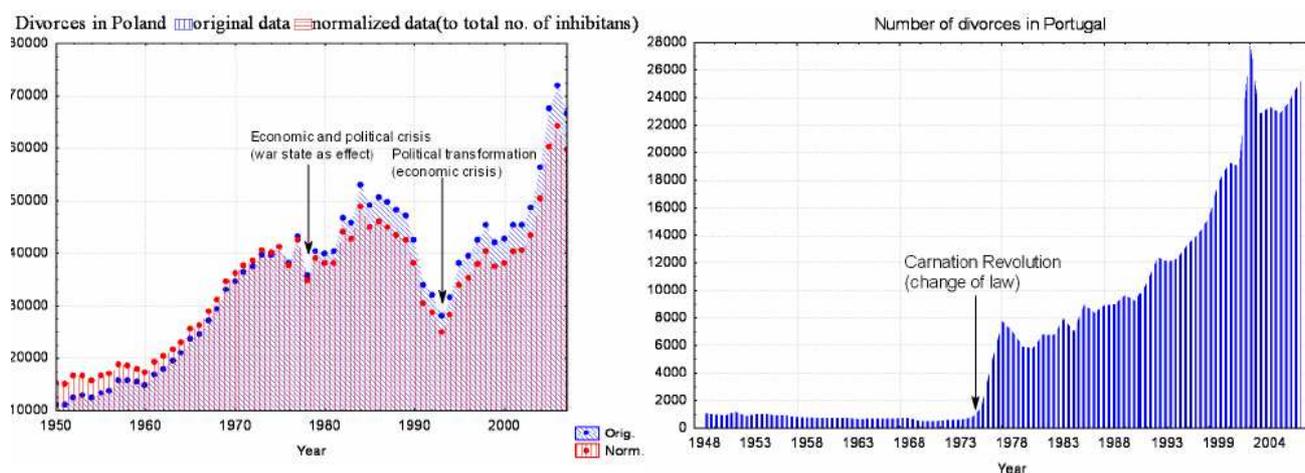
Decyzja jednostek o wstąpieniu bądź wystąpieniu ze związku małżeńskiego zależy od wielu czynników. W tym rozdziale omówimy 2 modele w których podstawowym motorem działań ludzkich jest norma społeczna wyrażana przez presję lub pęd owczy.



Rys. 60 Alegoria zmiany społecznej. Źródło: [www.demotywatory.pl](http://www.demotywatory.pl)

#### 11.1 Zachowania kolektywne i wpływ zmiany norm na przykładzie rozwodów

Punktem wyjścia do analizy była hipoteza, że rozwody jako zjawisko społeczne będzie można badać przy użyciu kolektywnego modelu Bouchauda (bazującego na fizycznej teorii magnetyzmu). W modelu kolektywne zachowania (poprzez imitację i presję społeczną) będą generować zmianę społeczną (w zależności od siły „pędu owczego” zmiana może być łagodna bądź gwałtowna). Podejście kolektywne bardzo dobrze sprawdza się przy opisie podejmowania przez jednostki decyzji przy dużej liczbie połączeń w społeczeństwie. W Europie dla kupna telefonów komórkowych, czy chęci urodzenia dziecka dokonano analizy porównawczej zmiany ilościowej dla poszczególnych państw (liczby sprzedanych telefonów czy współczynnika płodności). W obu przypadkach mamy do czynienia ze zmianą normy, raz nakazującą zakup telefonu w celu uniknięcia wykluczenia, a z drugiej strony osłabiający się nakaz posiadania potomstwa ze względu na zmieniający się tryb życia.

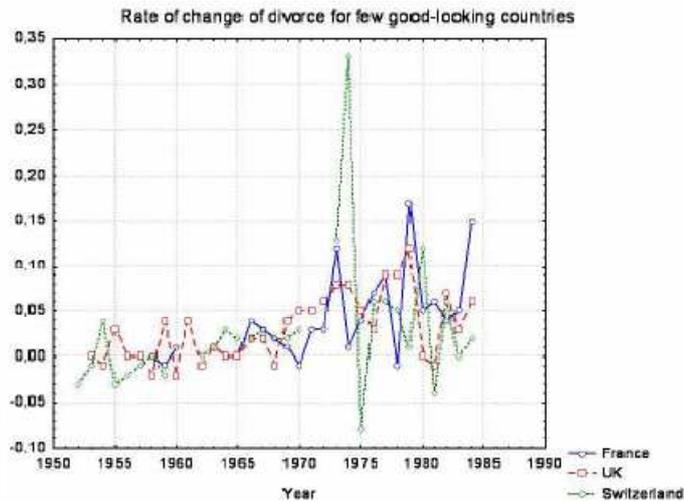


Rys. 61 Zmiana liczby rozwodów w Polsce i w Portugalii, spowodowana odpowiadającymi czasowo przemianami

**Czynniki społeczno-ekonomiczne w kontekście demograficzno-politycznym:** Liczba rozwodów również zależy od zmieniających się norm, gdyż pary decydują się na rozstanie zdając sobie sprawę z możliwych konsekwencji społecznych z tego czynu wynikających. Decyzja jednostek o wystąpieniu ze związku małżeńskiego zależy od bardzo wielu zmiennych. Z mierzalnych czynników makroskopowych wyróżnić można 2 główne składowe: ekonomiczne i społeczne. Na podstawie pojedynczych trajektorii dla konkretnych państw można wskazać okresy, w których zmiana ilości rozwodów w poszczególnych latach może być wytłumaczona konkretnym wydarzeniem ekonomiczno-politycznym. Czułość na lokalne zjawiska, nazywana idiosynkratycznością, będzie utrudniała analizę porównawczą różnych państw. Tłem ilości rozwodów jest również struktura demograficzna społeczeństwa, która raz, że zmienia się w czasie, a dwa, jest różna dla każdego kraju. Gdzieś w głębi przemian tkwi efekt zmian norm społecznych. Gdyby udało się wyeliminować szumy ekonomiczno-demograficzne, to powinniśmy zaobserwować jak zmieniająca się norma piętnująca rozwodników, przekształcającą się w przyzwolenie społeczne, skutkuje gwałtownym bądź spokojnym przyrostem liczby rozwodów. Do wspomnianej analizy wykorzystano bazę danych ONZ o ilości rozwodów w poszczególnych latach, dla większości krajów świata. Na samym wstępie odrzucono małe kraje (dla których fluktuacje roczne mocno zaburzałyby analizę), jak również kraje o krótkich szeregach czasowych. Później wybrano okres 1948-1982, jako przedział czasu, kiedy dla większości krajów mamy do czynienia z trendem wzrostowym ilości rozwodów. Powodem tej selekcji jest poprawka na strukturę demograficzną społeczeństw dotkniętych II Wojną Światową. Po roku 1982 (dla różnych krajów ta data może się różnić o kilka lat) nastąpił spadek ilości rozwodów dla większości krajów świata. Można to tłumaczyć dominacją związków małżeńskich zawartych przez ludzi urodzonych w czasie niżu demograficznego, który nastąpił zaraz po wyżu powojennym. Tak samo trzeba uwzględnić, że końcówka rozpatrywanego okresu jest zawyżona rozwodami ówczesnych 30 latków, będącymi nad-reprezentowanymi w społeczeństwie.

Analizy wykazała brak korelacji między dochodem ekonomicznym a predyspozycją do rozwodu. Pomimo niemal ciągłego trendu wzrostowego liczby rozwodów, natrafiono na pewne odchylenia. Mianowicie, w przypadku dużego przyrostu liczby rozwodów w skali roku, w roku następnym występuje korekta (liczba się zmniejsza). Zamiast oczekiwanego zwolnienia tempa wzrostu, co wcale nie musi oznaczać odwrócenia trendu. Innymi słowy, po maksymalnej zmianie zanotowanej wartości liczby rozwodów na przełomie dwóch lat, w następnym roku powinniśmy zaobserwować spadek (stąd użycie słowa korekta w znaczeniu zjawiska występującego na giełdzie). W socjologii znane są przypadki, kiedy działa sprzężenie zwrotne ujemne, pomimo przewagi dodatnich efektów. Można przypuszczać, że układy społeczne stawiają opór (przynajmniej chwilowo i lokalnie) polu zewnętrznemu (Zeitgeist). Zrozumienie przemian norm społecznych, wymaga wielu działań i analiz, jednak już na tym etapie można postawić kilka prawdopodobnych hipotez o kierunku przemian norm, przeze mnie omawianych. Tak więc można przyjąć, że lokalnie pojawi się przeciwna norma, interferująca z normą główną. Mogłoby to być zjawisko analogiczne do norm dotyczących aborcji, tylko na mniej medialną skalę, gdzie bez poparcia większości społeczeństwa, działają „ruchy obrony życia” walczące z aborcją.

Kolejną nasuwającą się hipotezą, jest pojawieniu się sił społecznych usiłujących zatrzymać zmieniającą się normę. Przykładem mogłyby być regulacje prawne utrudniające udzielanie rozwodów. Prawdopodobnym zjawiskiem powstałym w wyniku przemian norm byłyby także znamiona samoorganizacji społeczeństwa, kiedy to samoistnie odpowiadałoby ono na duże zmiany przeciwną reakcją.



Rys. 62 Przyrosty liczby rozwodów dla kilku wybranych państw

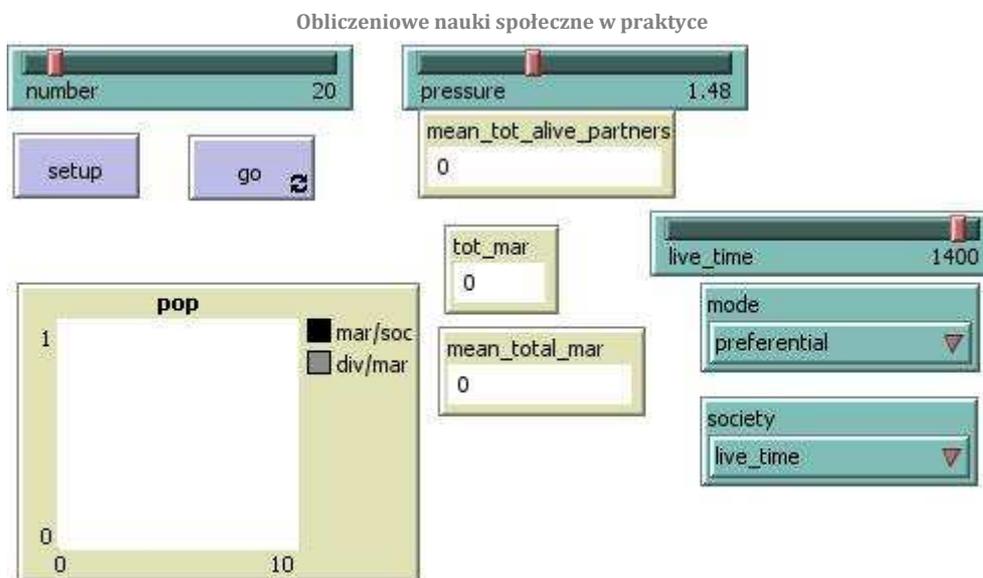
## 11.2 Dynamiczny model wiązania się w pary

Kontynuując zagadnienia demograficzne z poprzedniego rozdziału rozważmy model dopierania się w pary (Jarynowski, Nyczka, 2014). Niech agenci posiadają listę swoich atrybutów i cech preferowanych u potencjalnego partnera. Zgodnie z teorią kapitału i powiązanej z nim atrakcyjności międzyludzkiej (na podstawie Colemana i Bourdieu) można wyznaczyć wartości liczbowe odpowiadające cechom jednostki w wymiarach: kapitał społeczny, kapitał ekonomiczny, kapitał kulturowy oraz kapitał fizyczny. Poza parametrami indywidualnymi wyróżniamy właściwości globalne zinternalizowane przez członków społeczności, a mianowicie aktualne wartości wolności społecznej (siła norm, religijność, system prawny) oraz ekonomicznej (stosunek dochodu do zgromadzonego/pożyczanego kapitału). Jedynym parametrem omawianego modelu decydującym o wejściu bądź wyjściu z małżeństwa jest wolność tudzież presja (suma wolności/presji społecznej i ekonomicznej). W celu odzwierciedlenia cech „żywego społeczeństwa” w modelu następuje wymiana agentów w populacji aktywnej seksualnie (narodziny/śmierć). Dobór partnerów w celu porównania ich dopasowania (w celu sprawdzenia możliwości zawiązania, bądź rozpadu związku) następuje w sposób losowy z powinowactwem osób wolnych (nieprzebywających w związkach). Powyższy prosty model pozwala na klasyfikację społeczeństw na podstawie rejestrów demograficznych oraz na śledzenie zmian norm.

David Gale i Lloyd Stowell Shapley: "Stability marriage problem" w 1962 zaproponowali model matematyczny za który otrzymali nagrodę im. Nobla z ekonomii. W ich metodzie optymalizacyjnej szukamy skojarzeń na grafie dwudzielnym (kobiety, mężczyźni). Optymalne rozwiązanie wymaga wiedzy globalnej (np. każda kobieta posiada listę rankingową wszystkich mężczyzn). Zakłada się małżeństwo stabilne- czyli takie, że nie ma możliwości zdrady. Z modelu wychodzi, że małżeństwo optymalne dla np. kobiety jest nie gorsze od dowolnego małżeństwa stabilnego. Problem stabilnego małżeństwa w wersji podstawowej – algorytm:

1. Każdy mężczyzna oświadcza się najlepszej kobiecie na swojej liście preferencji.
2. Jeśli jakiejś z kobiet oświadczył się więcej niż jeden mężczyzna, to kobieta ta wybiera najlepszego z nich, a pozostałym odmawia.
3. Mężczyźni, których oświadczyły nie zostały przyjęte przez jakąś z kobiet, skreślają tę kobietę ze swojej listy preferencji. Przechodzimy do kroku 1.

Model deterministyczny można zamienić symulacjami. Zamiast informacji globalnej – lokalna. Każdą osobę charakteryzuje lista cech własnych i preferowanych. Ustala się dystans między potencjalną parą (miara Manhattan albo iloczyn skalarny). Pozwala się na wyjście z małżeństwa.



Rys. 63 Okno dialogowe symulacji dynamiki populacyjnej małżeństw/ rozwodów w Netlogo

Ustalamy, że agenci posiadają listę swoich atrybutów i cech preferowanych u potencjalnego partnera. Wykorzystujemy teorię kapitału i powiązanej z nim atrakcyjności międzyludzkiej (Coleman, Bourdieu) gdzie kapitał posiada wiele wymiarów: Kapitał Społeczny, Kapitał Ekonomiczny, Kapitał Kulturowy, Kapitał Fizyczny. Parametrem modelu Just Aktualna wartość wolności/presji społecznej (siła norm, religijność, system prawny) oraz ekonomicznej (stosunek dochodu do zgromadzonego/pożyczonego kapitału).

Jedynym parametrem omawianego modelu decydującym o wejściu bądź wyjściu z małżeństwa jest wolność tudzież presja (suma wolności/presji społecznej i ekonomicznej). W trakcie symulacji następuje wymiana agentów w populacji aktywnej seksualnie (narodziny/śmierć). W każdym kroku ma miejsce losowy dobór partnerów w celu porównania ich dopasowania (kroki Monte Carlo w znaczeniu czasu życia) z powinowactwem do osób wolnych (nieprzebywających w związkach).

Dynamika wiązania się w pary następuje według reguł.

1) Wybieramy agenta męskiego i żeńskiego oraz sprawdzamy:

- Jeżeli oboje są wolni, towiążą się pod warunkiem, że ich dopasowanie jest lepsze niż krytyczna wartość (parametr presji – im wyższa presja, tym mniej dopasowani agenci sięwiążą)
- Jeżeli choć jedno z agentów jest w związku, to aby powstał nowy związek, dopasowanie wylosowanych agentów musi być lepsze od dotychczasowego wzmocnionego o wartość siły związku (parametr presji – im wyższa presja, tym rzadziej dochodzi do wymiany partnerów)

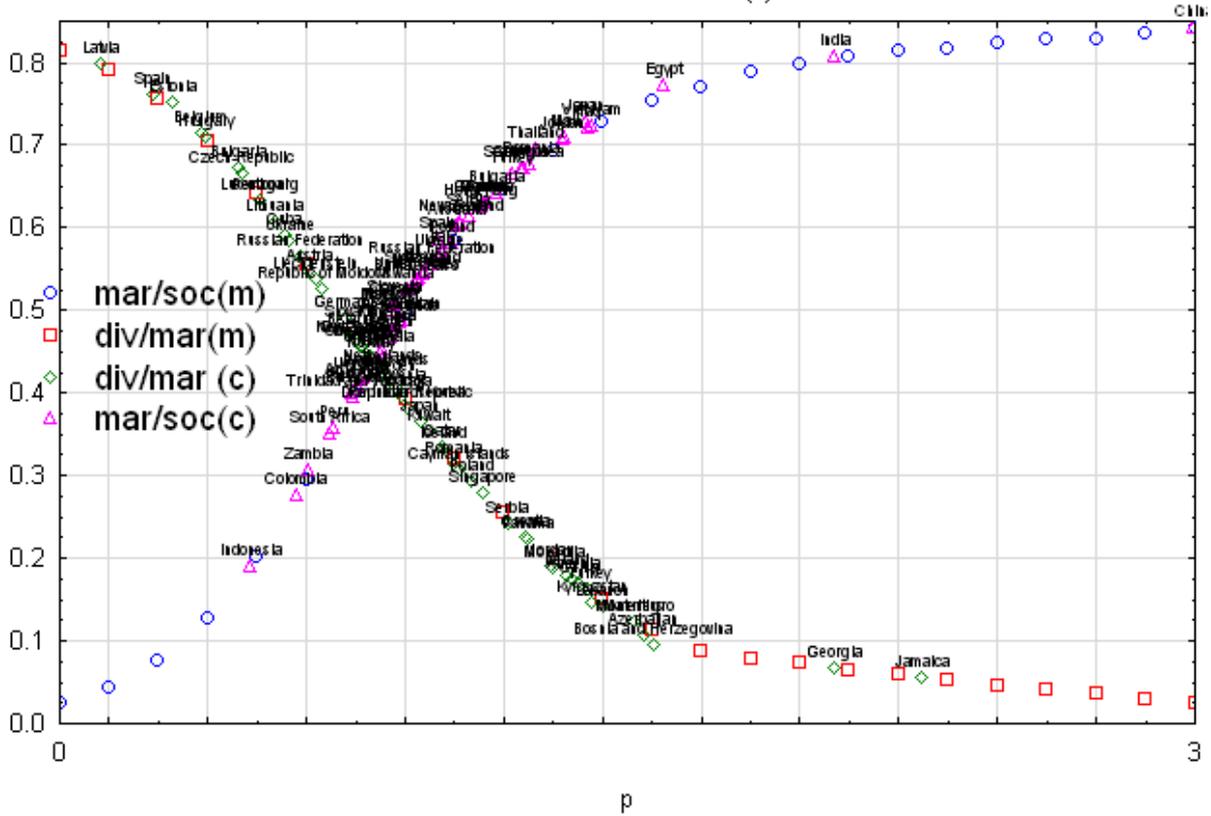
2) Dodatkowo z pewnym prawdopodobieństwem (odwrotnie proporcjonalnym do presji) małżeństwo agenta męskiego lub żeńskiego rozpada się spontanicznie.

W deterministycznej części modelu agenci dobierają się w pary jeżeli dopasowanie między nimi jest lepsze (= mniejsze) niż wartość krytyczna wejścia w nowy związek, bądź wymieniają się partnerami jeżeli dopasowanie między nowymi parami będzie lepsze niż wspomniana wartość krytyczna pomniejszona o siłę związku. Zakładamy przy tym, że wartość krytyczna wejścia w związek rośnie wraz ze wzrostem wolności, a siła związku wtedy maleje.

Porównaliśmy makro- charakterystyki wyników symulacji z danymi ONZ dla lat 2010-2012 oraz badania sondażowego World Value Survey dla tego samego okresu (Rys. 64). Wskaźnikami są stosunek liczby rozwodów do zawartych małżeństw w danym roku (div/mar) oraz stosunek liczby osób w związku małżeńskim do seksualnie aktywnej kohorty populacji (mar/soc). Poczyniliśmy pewne obserwacje odnośnie

pewnych generalnych wyników modelu. Ewolucyjna presja do prokreacji (wartość krytyczna dopasowanie par-presja w granicy zera oznacza brak możliwości zawarcia związku, czyli w konsekwencji wymieranie populacji). Obszar poniżej presji wartości krytycznej niefizyczny ze względu również na zbyt mały procent związanych agentów (brak możliwości wychowania potomków). Nie istnieje i nie może więc istnieć w pełni "wolne" społeczeństwo.

Divorce/marriage ratio (div/mar) and married/population ratio (mar/soc) for model model (m) and empirical data from different countries (c)



Rys. 64 Makro- charakterystyki wyników modelu dla różnych wartości parametru presji (p) oraz dla rzeczywistych populacji w latach 2010-2012

Odtworzyliśmy makro-charakterystyki populacyjne małżeńsko-rozwodowe za pomocą mikro-reguł wiązania się w pary opisanych jednym parametrem. Obszary niefizyczne (dla niskich wartości parametru  $p$ ) również są nieobserwowane w empirii. Dynamika modelu będzie rozważana pod kątem sieciowym (Sztokholm). Nastąpi rozdział presji na składowe społeczną i ekonomiczną. Dokonana zostanie obserwacja zmiany presji na podstawie szeregów czasowych danych historycznych

## 12 Modele rynku fonograficznego

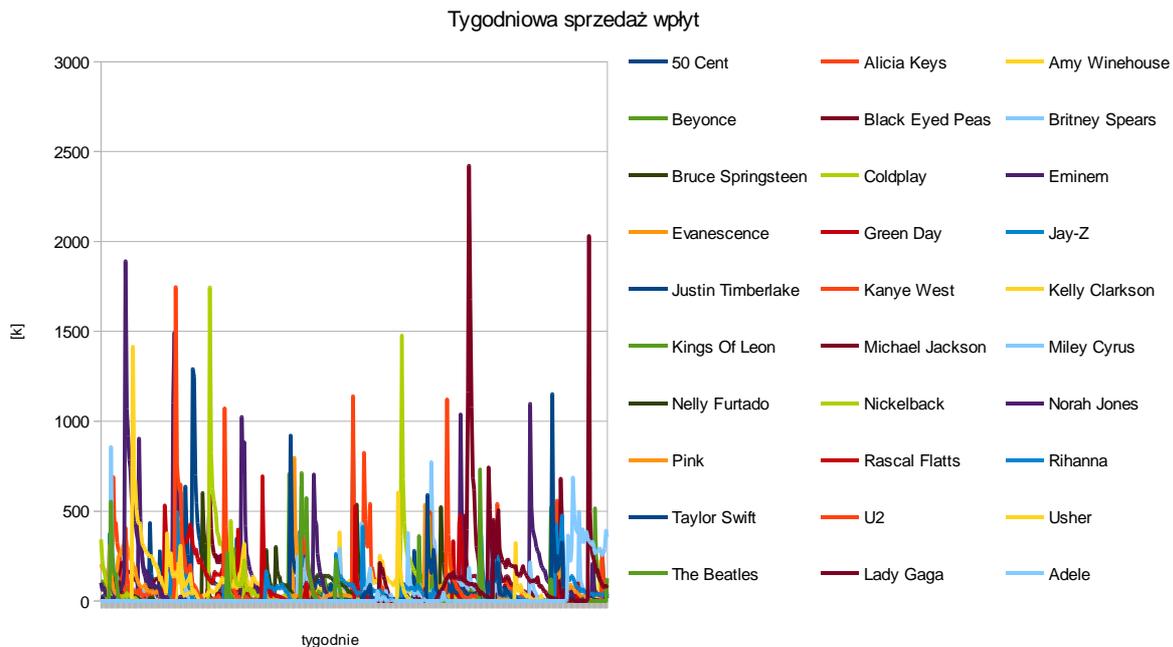
Rynek fonograficzny tworzą artyści, których wartość rynkowa mierzona jest notowaniami na cotygodniowych listach sprzedaży płyt (winyle, CD, mp3 etc.). W dobie informatyzacji i szybkiej wymiany informacji przez Internet specyfika sprzedaży przybrała gwałtowne i szybko ewoluujące formy (ekonomiczny efekt Eppsa). Muzyka jako dobro kulturowe z punktu widzenia ekonomii jest towarem, stąd analogia do rynku towarowego (Alexander, 1994). Odkąd szybki dostęp do informacji – promocja internetowa, wymiana wiedzy między użytkownikami - Internet stał się dominującym czynnikiem kreującym rynek, opis zmian jest konieczny. Nowa społeczna przestrzeń Internetu, to z jednej strony przestrzeń wymiany, ale też przestrzeń produkcji (Batorski, 2006). Dlatego też, podstawowym celem naukowym naszego projektu jest zrozumienie funkcjonowania i konstrukcja układu producentów (firm fonograficznych), produktów (artystów) i konsumentów, powiązanych i oddziałujących zarówno pomiędzy sobą, jak i wewnątrz swoich grup. Rynek fonograficzny może być traktowany jako przykład układu, który oprócz kilku charakterystycznych własności (takich jak np. nietypowa periodyczność sprzedaży czy nadzwyczaj silne uzależnienie jej wyników od historii) posiada cechy typowe dla rynków towarowych (konkurencja między wytwórcami, heterogeniczność konsumentów, cykl życia produktu). Struktura i dynamika rynku muzycznego jest konsekwencją oddziaływania czterech płaszczyzn: biologicznej (funkcjonowanie mózgu i mechanizm tworzenia się preferencji muzycznych), psychologicznej (emocje, zob. Perlovsky, 2010), społecznej (komunikacja interpersonalna, mass media) i czysto ekonomicznej (relacja producent-konsument). Dzięki silnej złożoności większości z tych elementów i nieliniowości oddziaływań między nimi, także na rynku fonograficznym obserwuje się szereg zjawisk, takich jak zjawiska kolektywne, których istnienie ma charakter emergentny, a Internet stał się katalizatorem tych procesów. Dodatkowo nasz artykuł obejmuje zagadnienia związane z socjologią muzyki szeroko rozumianą jako całokształt zagadnień zawartych w sformułowaniu: muzyka a społeczeństwo. Do tej pory większość badań muzykologów polskich analizujących sferę społeczną ograniczała się do jakościowego ujęcia, jak estetyka muzyki, społeczna historia muzyki czy studia kulturowe (Mika, 2004). Badania ilościowe w kontekście muzyki popularnej i jej konsumentów są w stadium początkującym w Polsce (Mysliwski, 2011) a nasz artykuł jest właśnie ilościowym obrazem rynku fonograficznego. Wkraczamy na nurt badań spokrewniony z antropologią kulturową współczesnego globalnego społeczeństwa (ponieważ analizujemy dane sprzedażowe zebrane z całego świata). Rynek fonograficzny stał się bowiem globalny właśnie dzięki Internetowi. Wybór jedynie najlepiej sprzedających się muzyków (wszyscy ukształtowani w obrębie kultury anglo-amerykańskiej) ukierunkowuje badania w stronę ekonomii behawioralnej, w rozumieniu konsumpcji kultury. Sztuka zredukowana do muzyki popularnej staje się więc artystycznym produktem. W tym paradygmacie uznajemy muzykę popularną jako pełnoprawny element kultury współczesnej (Rogers, 2005), ale też jako istotną branżę gospodarki. Przedstawiona praca jest wynikiem współpracy socjologa analitycznego oraz dziennikarza muzycznego doświadczonego w ilościowych badaniach nad muzyką popularną obecnie oraz jej historią (Buda A. , 2006).

### 12.1 Szczegółowa analiza zmian na obecnym rynku fonograficznym

Badany przez nas rynek sprzedaży albumów stał się globalny pod koniec lat 60-tych XX wieku. 70% światowego rynku jest pod kontrolą czterech wielkich koncernów – EMI, Universal, Warner i Sony BMG. Mają one w swoich szeregach 30 najpopularniejszych na świecie artystów uwzględnionych w badaniach.

Jednak wówczas ze względu na wolniejszą wymianę informacji cykl życia produktu – płyty długogrającej – przechodził ewolucję. W efekcie album osiągał najwyższe pozycje na listach bestsellerów dopiero kilka tygodni po premierze po wyczerpaniu promocji, po czym pamięć o nim zazwyczaj powoli ginęła, wraz z wyczerpaniem wytłoczonego nakładu. Aby stosować analogię do dobrze zbadanych rynków finansowych, musieliśmy dostosować rynek fonograficzny do statystycznej analizy. Dlatego założyliśmy, że przedmiotem

rynku fonograficznego są artyści, których wartość (notowania) mierzone są cotygodniową sprzedażą płyt długogrających, czyli albumów (Winyle, CD, mp3, etc.). Założenie to jest uzasadnione, bowiem na rynku towarowym cena płyty jest na ogół stała, ustalona z góry przez wytwórnę i dystrybutora. W szczególności stała cena dotyczy płyt nowych i premierowych (bowiem ich przeceny w sklepach są dokonywane dopiero, gdy płyta spadnie już z list przebojów, a więc jej sprzedaż dąży do zera i nie jest ujmowana w naszych danych). Oczywiście zdarza się, że popularni artyści mają więcej niż jedną płytę na liście najlepiej sprzedawanych albumów. Wówczas notowania artysty oznaczają sumę sprzedaży jego płyt w danym tygodniu. Staje się tak, gdy w okolicach premiery na fali promocji konsumenci sięgają po wcześniejsze tytuły swojego idola. Bowiem na rynku fonograficznym walorami są artyści.



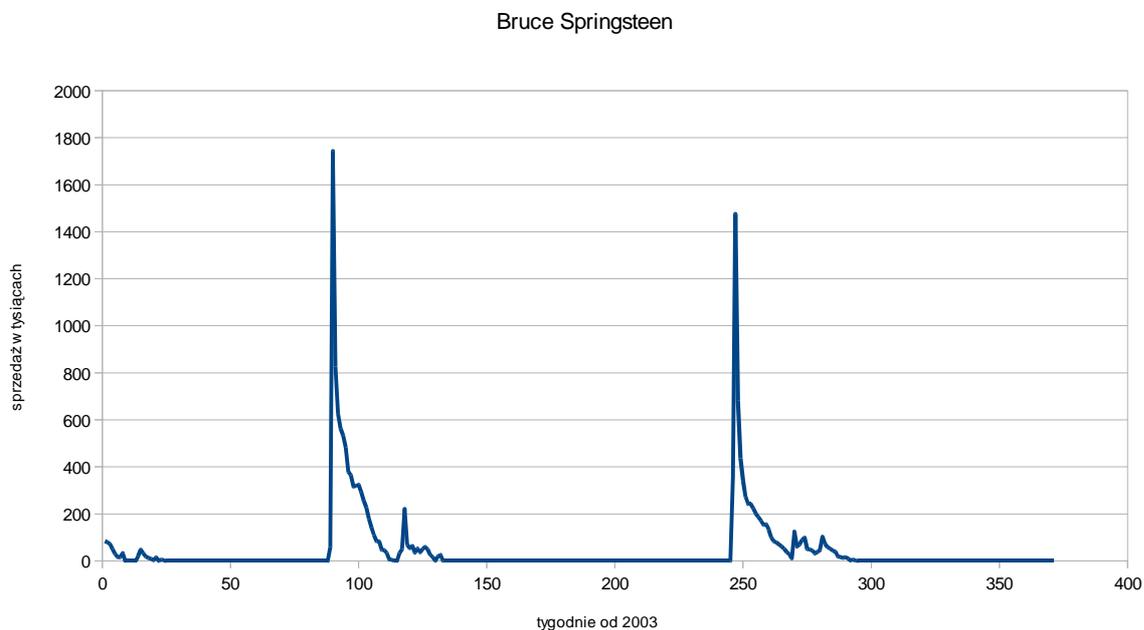
Rys. 65 Najlepiej sprzedający się artyści ostatniej dekady

W naszych badaniach bierzemy pod uwagę cotygodniowe dane o sprzedaży albumów na świecie w okresie od września 2003 do września 2013 (źródło: [www.musiccharts.net](http://www.musiccharts.net)), bowiem dopiero od dekady cotygodniowa globalna sprzedaż płyt może być uzyskana w sposób tak dokładny. Do analizy powiązań między artystami wybraliśmy trzydziestkę najpopularniejszych wykonawców (Rys. 65), którzy w latach 2003-2013 sprzedali co najmniej 11 milionów płyt. Najpopularniejsi z nich – Eminem i Black Eyed Peas osiągnęli sprzedaż prawie 30 milionów swoich albumów. W rzeczywistości jednak artyści często sprzedają sporo płyt występując solo poza macierzystym zespołem. Z tego powodu w sprzedaży płyt zespołu Black Eyed Peas zawarliśmy także sprzedaż solowych płyt jego członków – Fergie i Will I Ama. Podobnie w przypadku The Beatles uwzględniliśmy także sprzedaż płyt Johna Lennona i Paula McCartneya, a do sprzedaży płyt Beyonce także te, które nagrała ze swoim macierzystym zespołem Destiny's Child. Ma to uzasadnienie, bowiem konsumenci identyfikują danego artystę i kupują jego muzykę nawet, jeśli występuje on pod innymi nazwami. Działo się to jednak o wiele rzadziej, niż u wykonawców, którzy wydawali płyty pod inną nazwą ze względów artystycznych, lub prawnych.

Obserwacje empiryczne – widoczne choćby na Rys. 65– pokazują, że sprzedaż płyt danego wykonawcy osiąga maksimum dokładnie w tygodniu premiery jego nowej płyty. W praktyce premiery płytowe odbywają się w pierwszy dzień roboczy nowego tygodnia – a więc w poniedziałki, lub wtorki – bowiem sprzedaż za dany tydzień tradycyjnie zliczana jest od poniedziałku do soboty (lub w niektórych krajach do niedzieli) i publikowana w sobotę następnego tygodnia. Oczywiście na części rynków lokalnych wyniki sprzedaży znane są już wcześniej. Na przykład w Wielkiej Brytanii zliczone są już kilka godzin po zakończeniu tygodnia handlowego, a dla wybranych wykonawców możliwe jest nawet oszacowanie sprzedaży ich płyt

dzień po dniu. Dane te są jednak tajemnicą handlową wytwórni i nie są publikowane, bo służą głównie kierowaniu promocji w taki sposób, aby całkowita sprzedaż była największa. Znane są jednak otwarte wojny handlowe jak ta między The Rolling Stones a The Beatles (1967), między Blur i Oasis (1995), oraz z sierpnia 2001 między singlami Victorii Beckham „Out Of Your Mind” a „Groovejet” Spillera, gdy do piątku prowadziła Victoria Beckham, a Spiller od soboty objął ostateczne prowadzenie, bo pod koniec tygodnia do zakupów włączyli się bywalcy nocnych klubów, gdzie singiel Spillera był częścią grany.

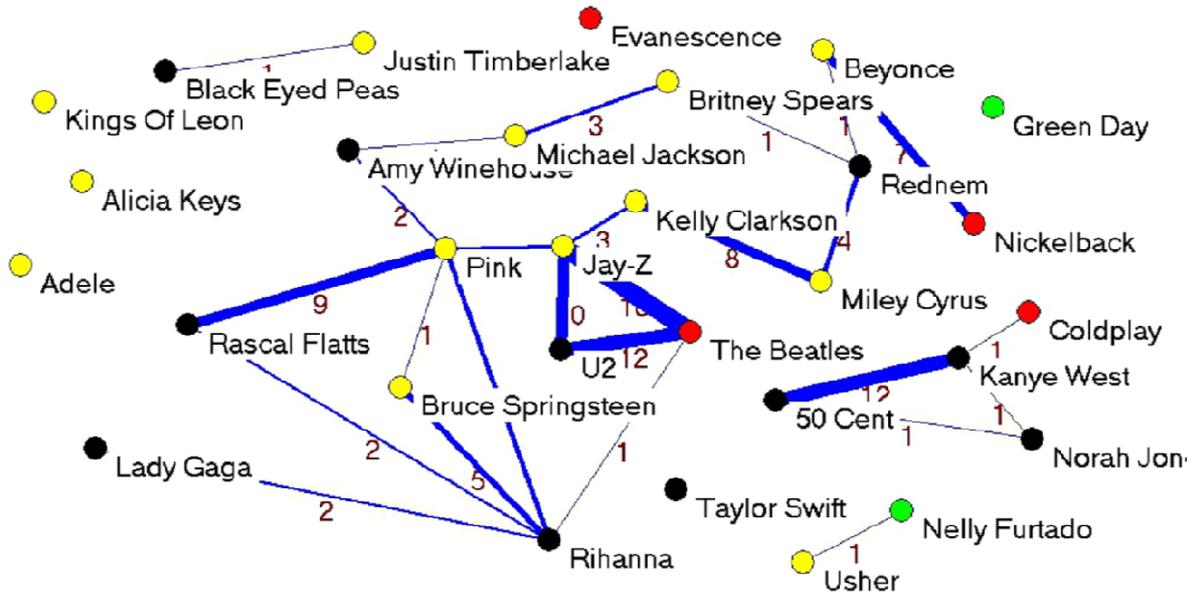
Wojny handlowe najczęściej dotyczą singli, bo jeśli chodzi o sprzedaż albumów, to „wielka czwórka” koncernów (EMI, Universal, Warner, Sony BMG) dążąc do maksymalnej sprzedaży woli jednak ustalać daty premier płytowych odgórnie, nawet w porozumieniu między sobą, często z wielomiesięcznym wyprzedzeniem. Premierze albumu najczęściej towarzyszy promocyjny singiel, którego obecność w mediach podnosi sprzedaż. Dlatego łatwiej ją kontrolować, bowiem tuż po premierze obserwujemy spadek sprzedaży, podobny do wykładniczego (Rys. 66) podobnie jak np. w trzęsieniach ziemi. Oczywiście wytwórnie chcą mu zapobiec, więc wydając kolejne single wykrojone z albumu podnoszą choć na chwilę jego sprzedaż – co również widać na wykresie (Rys. 66) w kształcie opadania po premierach płytowych.



Rys. 66 Większość sprzedaży w pierwszych tygodniach od premiery

Mimo, że rynek fonograficzny jest dużo bardziej przewidywalny niż rynki finansowe jednak i tu na sprzedaż płyt mają wpływ zdarzenia losowe takie jak śmierć artysty. Bowiern zapotrzebowanie na nagrania zmarłych artystów jak Michael Jackson (2009) czy Amy Winehouse (2011) było tak duże, że wywindowało ich stare płyty na sam szczyt list sprzedaży tuż po ich śmierci. Było to możliwe, bo wielkie koncerty mają dziś technologię pozwalającą dotłoczyć setki tysięcy nowych egzemplarzy w ciągu tygodnia – a zatem zdążyć na następne notowanie cotygodniowej listy sprzedaży. Dla porównania przed internetyzacją rynku: w roku 1977 gdy umierał Elvis Presley czy w roku 1980, gdy zastrzelono Johna Lennona, cykl produkcyjny płyt winylowych trwał dłużej i aby sprostać zapotrzebowaniu Wielka Brytania musiała zlecać pracę tłoczniom we Francji, a i tak zmarli artyści dotarli na pierwsze miejsca z trwającym parę tygodni poślizgiem.

Prosta analiza korelacji między walorami na tych rynkach – to jest ilości korelacji silnych, słabych i ujemnych – pokazuje, że rynek fonograficzny zdecydowanie różni się od rynków finansowych, bo dominują na nim korelacje ujemne. Oznacza to, że jeden artysta zyskuje wysoką sprzedaż kosztem drugiego, a zatem mamy jak na każdym rynku towarowym zjawisko konkurencji między produktami (czyli artystami).



Rys. 67 Istotne ( $\alpha = 0.05$ ) dodatnie korelacje między artystami (Buda & Jarynowski, 2013)

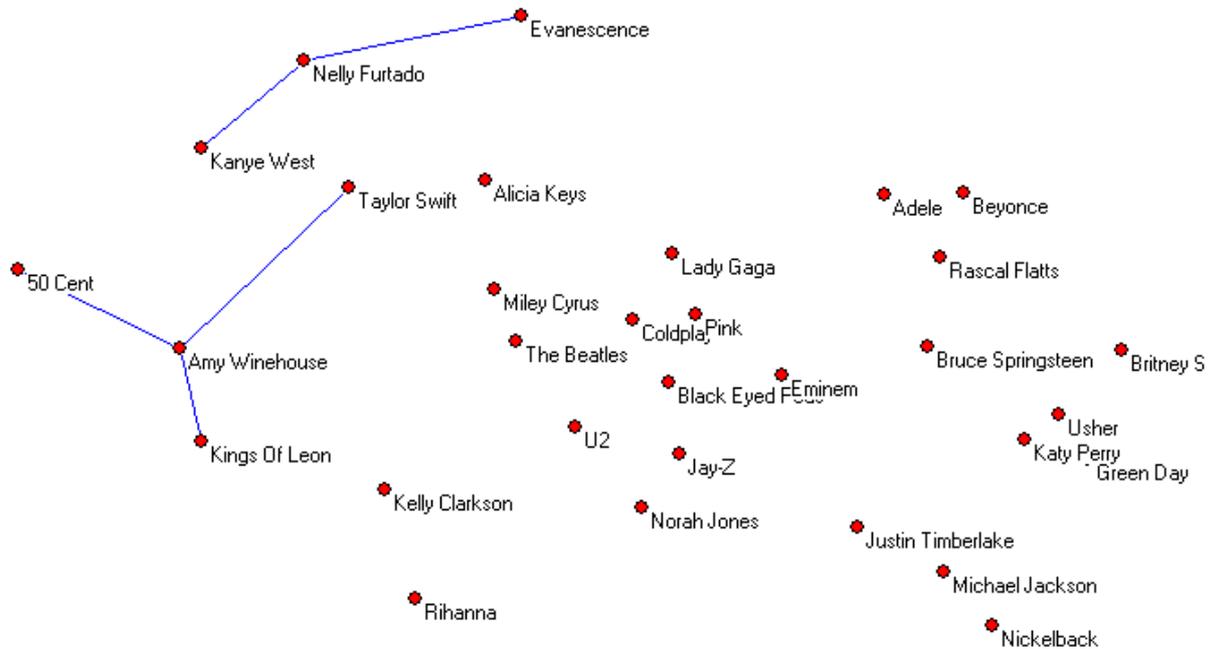
Na rynkach finansowych sieć powiązań korelacyjnych odzwierciedlała strukturę rynku – czyli sektory i podsektory gospodarki. Czy zatem rynek towarowy jakim jest fonografia także pokaże nam stylistyczne związki między artystami? Okazuje się, że nie do końca. Bowiem sieć (Rys. 67) otrzymana dla „portfela” 30 najpopularniejszych artystów odzwierciedla wszystkie style muzyczne, jednak z wyjątkiem muzyki pop.



Rys. 68 Drzewo minimalnej rozpiętości uzyskane dla 30 najpopularniejszych artystów na podstawie sprzedaży ich płyt w latach 2003-2011.

W reprezentacji drzewiastej więcej informacji możemy odzyskać, bo wszystkie wierzchołki muszą być połączone. Artyści rockowi są połączeni (Coldplay, Nickelback, Kelly Clarkson, Kings Of Leon, Green Day, Evanescence, U2). Obserwujemy również sektor związany z rapem (50 Cent, Kanye West, Usher, Eminem), z którym sąsiadują artyści soul i r'n'b (Adele, Beyonce, Britney Spears, Alicia Keys, Michael Jackson). Jednak brakuje sektora dla muzyki pop. Zamiast niego w środku drzewa sąsiadują ze sobą gwiazdy różnych gatunków – rocka (The Beatles, Bruce Springsteen), rapu (Jay-Z) r'n'b (Rihanna) i pop (Lady Gaga).

Zgodnie bowiem z definicją gatunek „pop” jest skrótem od słowa „popularny” . Oznacza to, że tak naprawdę to wszyscy artyści mimo tworzenia estetycznie różnych odmian muzyki należą do kręgu pop i próba tworzenia estetycznie nowego gatunku mija się z celem. Dlaczego? Otóż klasyczne gatunki jak rock, soul czy r'n'b zostały zdefiniowane 50-60 lat temu, a ich granice estetyczne praktycznie się nie zmieniły, podobnie jak rap, który istnieje w przemyśle płytowym od 35 lat . Z kolei pop dziś oznacza zupełnie co innego, niż 30, 40 czy 50 lat temu, a wyrazem różnicy jest zupełnie inna estetyka tego stylu obowiązująca dziś i w przeszłości. Zatem nasze odkrycie, na podstawie drzewa minimalnej rozpiętości, które gatunki istnieją rzeczywiście – w sensie ekonomicznym dla konsumenta – a które nie, jest wynikiem nietrywialnym.



Rys. 69 Istotne ( $\alpha = 0.05$ ) ujemne korelacje między artystami (Buda & Jarynowski, 2013)

Ponieważ na rynku płytowym dominują korelacje ujemne postanowiłem sprawdzić, czy ich przyczyną jest jedynie konkurencja między produktami. Okazuje się, że spośród 375 ujemnie skorelowanych par znacząco ujemnych jest jedynie kilka (Rys. 69), przy czym są one skupione głównie wokół piosenkarki Amy Winehouse. Aby znaleźć przyczynę istotnie ujemnych korelacji wokół Amy Winehouse warto sięgnąć do historii rynku i cofnąć się do 2011 roku, gdy Amy Winehouse zmarła. Było to w lipcu, a więc w czasie, gdy wytwórnie unikają premier płytowych, a sprzedaż jest niska. Wówczas to śmierć artystki wywindowała jej nagrania z powrotem do czołówek list najlepiej sprzedawanych płyt, kosztem innych artystów. Śmierci Amy Winehouse nie towarzyszyła żadna inna premiera płytowa, a jej sprzedaż płyt gwałtownie rosła, w sytuacji gdy reszta rynku miała powolny spadek sprzedaży. Stąd ujemnych korelacji u Amy Winehouse jest najwięcej, bo śmierć artysty wyprowadza rynek z równowagi.

Dla porównania: gdy w 2009 roku umierał Michael Jackson – a więc największa gwiazda ostatnich 30 lat – był czerwiec, a zatem jego śmierć nałożyła się na obecne jeszcze przed wakacjami premiery płytowe, stąd w otoczeniu Jacksona brak istotnych ujemnych korelacji. I to mimo że w pierwszym tygodniu po śmierci „króla popu” sprzedano kilka milionów jego płyt.

## 12.2 Sformułowanie modelu rynku fonograficznego

Gwałtowny wzrost sprzedaży w chwili premiery, czyli dostarczenia nowego towaru na rynek, jest charakterystycznym zjawiskiem na rynku fonograficznym. W miarę upływu czasu krzywa sprzedażowa zanika wykładniczo z okresowo pojawiającymi się wzmocnieniami, odpowiadającymi chwilom wydawania promocyjnych singli. Daty premiery – a więc pików sprzedaży czyli „wstrząsów głównych” – są ustalane odgórnie przez wytwórnię płytową wiele tygodni przed premierą. Na ogół najpopularniejsi artyści wydają kolejne nowe albumy co 2 lata, jednak 40 lat temu powszechne było wydawanie albumów co roku. Ta historyczna zmiana ma związek z pojemnością nośnika, bowiem do połowy lat 80-tych na rynku dominowała płyta winylowa, która w wersji długogrającej zawierała do 40 minut muzyki. Natomiast wprowadzone 30 lat temu na rynek płyty kompaktowe mogą pomieścić do 80 minut nagrań, co przy założeniu, że artyści są tak samo płodni, w naturalny sposób wydłużyło odstęp między premierami płytowymi. W efekcie wytwórnie aby skrócić czas oczekiwania na nowy album wydają w międzyczasie składanki największych przebojów artysty, bądź jego nagrania koncertowe, najczęściej pod koniec roku.

Ilość singli przypadająca na jeden album oczywiście wpływa na czas trwania promocji, który w praktyce równa się ilości tygodni jakie album danego wykonawcy spędza na liście sprzedaży od chwili premiery (Ordanini A. , 2006). Bowiem wytwórnice wychodzą z założenia, aby promować te produkty, które się dobrze sprzedają, stąd płyty niszowych, awangardowych artystów nie są promowane w ogóle, a płyty Michaela Jacksona „Thriller” (1982) i „Bad” (1987) doczekały się odpowiednio 7 i 9 przebojowych singli, powodując wzrosty sprzedaży macierzystych albumów i przedłużając ich obecność na listach do kilku lat.

Zbudowany przez nas model stochastyczny sprzedaży został uzupełniony o agentową implementację zachowań konsumenckich. Jako udoskonalenie modelu ujmującego zachowania konsumentów zaproponowaliśmy kolejne założenia:

- a) zależność od sezonowości jeśli chodzi o częstotliwość wydawania albumów - tak aby uwzględnić okresy szczytowej sprzedaży w ciągu roku;
- b) markowowski rozmiar maksimum sprzedaży w chwili premiery, proporcjonalny do całkowitej sprzedaży poprzedniej płyty - tak, aby uwzględnić, że zakupu albumu w chwili premiery dokonują przede wszystkim dotychczasowi fani, posiadający poprzednie płyty artysty;
- c) zależność częstotliwości premier płytowych od pory roku wynika z kulturowo -przyrodniczego cyklu rocznego, dlatego prawdopodobieństwo wydania albumu zostanie przeskalowane.

Drugim nowym udoskonaleniem jest wprowadzenie założenia, że ilość sprzedaży płyt w dniu premiery jest proporcjonalna do całkowitej sprzedaży poprzedniego albumu. Zamiast dotychczasowej zmiennej opisującej skok sprzedaży w momencie wydania albumu, wprowadzamy nową zmienną zależną od całkowitej sprzedaży poprzedniego albumu. Jednocześnie przyjmujemy, że sprzedaż pierwszego (debiutanckiego) albumu artysty jest niezależną zmienną losową. Założenie to ma uzasadnienie w małych budżetach, które koncerty fonograficzne przeznaczają na promocję pierwszych płyt. Bowiem debiutanci zanim podpiszą kontrakt płytowy najczęściej promują swoją twórczość na własną rękę i dopiero jeśli spodoba się publiczności, trafiają pod skrzydła wielkich koncernów płytowych.

W wyniku porównania trajektorii modelowych i empirycznych przez obserwowanie zachowań kolektywnych konsumentów muzyki, zapostulowane zostanie wyodrębnienie dwóch kategorii kupujących – **świadomych fanów muzyki** (zwracających uwagę na jakość produktu i przywiązanych do konkretnych gatunków muzycznych) i **klienteli ‘świętecznej’** (podążającej za trendami).

- Ci święteczni klienci kupują oryginały jedynie na prezent dla kogoś;
- Ci świadomi przed zakupem płyty przesłuchują pirackie wersje aby podjąć decyzję o ewentualnym zakupie.

Pokazaliśmy mechanizmy rządzące rynkiem fonograficznym (zmiany wynikające z pojawienia się Internetu). Dokonaliśmy analizy zależności na gwiazdorskim rynku fonograficznym (w tym wykorzystanie e-ryнку przez wytwórnice). Przedstawiliśmy model stochastyczny sprzedaży – klasyfikacja konsumentów muzyki (w tym wykorzystanie piractwa przez konsumentów). Dynamika cotygodniowych list pozwala na dwa sposoby uchwycić świat muzyczny. Po pierwsze wydaje się, że pozycja na liście ma związek z jakością. Do tego mnogość muzycznych gustów wielu ludzi, powinna pozwalać na wiele interpretacyjnych układów tak, że w obrębie swojej grupy dobrzy artyści zostaną docenieni. Po drugie jednak obserwujemy zachowania kolektywne (stadne), gdy większość kupujących, reaguje w sposób kaskadowy. Muzyka jest dość często pojmowana jako, produkt konsumpcyjny. Ludzie nie koniecznie słuchają muzyki żeby się nią cieszyć, a celem jest zakup muzyki, który jest postrzegana “hip” (Syrda, 2013), czym można się dzielić na mediach społecznościowych.

Fani jak i wytwórnice płytowe przenieśli swoją główną działalność do Internetu. Konsumenty skorzystali przez dostęp do tańszej (darmowej?) muzyki, ale czy tej samej jakości? Wytwórnice utrzymały stały poziom dochodów przez zmniejszenie kosztów (reedycje na e-nośnikach jako przykład).

W dobie zmieniających się warunków na rynku fonograficznym różne strategie graczy (firm fonograficznych i samych artystów) są obserwowane:

- Czy lepiej iść na fali gatunku i wydawać płyty o podobnej tematyce zaraz po sobie, czy przesuwac premiery, aby nie robić sobie konkurencji;
- Zdobywać 'gwiazdy' czy promować młodych artystów;
- Fokus na rynkach dojrzałych, czy rozwijających się (walka z piractwem).

### 12.3 Analiza innych układów społeczno-ekonomicznych, rynków towarowych i pieniężnych

Metody obliczeniowe w ekonomii mają dużo dłuższą tradycję niż w naukach społecznych. Wymiana handlowa, czy zaspokajanie potrzeb konsumentów jest czynnością społeczną, więc wiele matematycznych modeli/analiz danych mikro- i makro-ekonomicznych ma podstawy socjologiczne. Model stabilnego małżeństwa (rozwinęty w tej książce), model wyboru sekretarki, modele skoringowe wyceny ryzyka, modele podejmowania ryzyka, to podstawowe narzędzia pracy współczesnego ekonomisty. Ekonomia, ma w pewnym sensie sporą przewagę nad socjologią – ilość danych ekonomicznych jest olbrzymia. Dlatego to w tym obszarze nastąpił ogromny rozwój ekstrakcji i analizy tak zwanej nowej dyscypliny naukowej „Big Data Science”.

Na rynkach kapitałowych podstawą do określenia przyszłej wartości danego waloru jest wykorzystywana analiza fundamentalna w połączeniu z analizą techniczną. Pierwsza z nich służy do określenia bieżącej wartości danego waloru, np. produktu czy akcji spółki, na podstawie przesłanek wynikających z oceny sytuacji na rynku, kondycji przedsiębiorstwa, jakości zarządzania itd. i porównywania tej wartości z aktualną wyceną tego waloru przez rynek giełdowy, natomiast druga z nich ma charakter czysto statystyczny i pozwala na przewidywanie ruchu cen  $P(t)$  na podstawie ich historii bez wnikania w ich relację z rzeczywistą (pozagiełdową) wartością waloru. Analiza techniczna jest wykorzystywana przez inwestorów znacznie częściej niż analiza fundamentalna, ponieważ jej stosowanie nie wymaga aż takich kompetencji i dostępu do wielkiej ilości danych, często niejawnych, jak ma to miejsce w przypadku analizy fundamentalnej. Ponadto w przypadku danych o cenach i transakcjach giełdowych mamy do czynienia lepszą ich jakością i mniejszym zaszumieniem niż w przypadku tzw. danych fundamentalnych. Ze względu na statystyczny charakter analizy technicznej, inwestor dysponuje też w tym przypadku nieporównywalnie większym zestawem narzędzi analitycznych o dobrym umocowaniu formalnym.

## 14 Analiza procesów ewolucyjnych, dyfuzyjnych, korupcji i hierarchiczności

Dane o procesach społecznych analizowane są różnymi metodami informatycznymi jak sztuczna inteligencja, uczenie maszynowe, sieci neuronowe czy analiza semantyczna (Szostek, Jaszuk, & Walczak, 2012). Poprzedni rozdział o rynku fonograficznym stał już na pograniczy socjofizyki (gdzie modelem jest opis matematyczny procesu) a ekstrakcją informacji (gdzie dominują statystyczne narzędzia do wyciągania wiedzy o układzie na podstawie modelowych zasad interakcji). W tym i następujących rozdziałach będzie dominować analiza danych w oparciu o metody obliczeniowe.

### 14.1 Modele ewolucji społecznej

Modele ewolucyjne wyjaśniają wiele zjawisk biologicznych (jak wymieranie gatunków), a algorytmy genetyczne z powodzeniem znalazły swoje zastosowania w obliczeniach komputerowych. Tego typu podejście wykorzystuje się również do procesów społecznych których skala czasowa to dziesiątki, bądź setki lat. Przekazywanie kolejnym pokoleniom pewnych uwarunkowań odpowiedzialne jest powstawanie różnych struktur populacyjnych (o pewnych cechach charakterystycznych. Biologia wyjaśnia takie zjawiska specjacji poprzez rywalizację dwóch sił natury - doboru naturalnego (selekcja) i przepływu genów. W zależności od modelu różne analogie są stosowane, ale przykładowo przepływowi genów odpowiada transmisja kulturowa, a selekcja naturalna jest również obserwowana społeczeństwie poprzez uprzywilejowanie najlepiej dostosowanych jednostek. Przy czym należy zaznaczyć, że podobnie jak w przypadku środowiska naturalnego obserwowana jest zmienność warunku spowodowana różnymi czynnikami jak moda, bądź zewnętrznymi jak wojna. Dobór naturalny powoduje różnicowanie się sąsiadujących społeczeństw, ponieważ każda grupa społeczna przystosowuje się do swoich lokalnych warunków. W zależności od cywilizacji transmisja kulturowa dąży do zatarcia różnic między populacjami. Dodatkowo poza wertykalną transmisją od rodziców do potomków mamy w społeczeństwie transmisję horyzontalną (media, szkoły, itp.) której zadaniem jest utrzymanie struktury społecznej (przykład Francji i Wielkiej Brytanii odtwarzania struktury P. Boudieu). Na podstawie analizy układów biologicznych możemy się spodziewać zależności średniej wartości cechy metapopulacji od warunków zewnętrznych. Czasami niewielkie zmiany warunków zewnętrznych prowadzą do drastycznych zmian metapopulacji. Poza stosowaniem analogii do zachowań społecznych można do genetycznej populacji ludzi włożyć pewien dryft poprzez społeczne mechanizmy. Na prostym modelu populacji poddanej selekcji naturalnej zbadano jaki wpływ na jej rozwój ma stosowanie eugeniki (Pękalski & Cebrat, 1999). Porównano dynamikę dwu populacji znajdujących się w zmiennym środowisku - jednej, w której pulę genetyczną jest bez ingerencji i dwóch z interwencją, w której usuwa osobniki źle dostosowane do zmiany środowiska (eugenika) lub celowe wprowadzanie do układu osobniki uprzednio z niego wyeliminowane (jak poprzez zapłodnienie *In vitro*). W modelu z eugeniką pomimo zwiększenia dopasowania do środowiska i coraz lepiej funkcjonującego społeczeństwa, mogą się zdarzyć katastrofy, w przypadku zmiany warunków tego środowiska (obserwowane zdecydowanie rządu w układzie bez interwencji). Z kolei wprowadzanie uprzednio wyeliminowanych osobników powoduje permanentny dryft słabnącego dopasowania. Dla zilustrowania modelu ewolucji populacji, będącym społeczeństwem, wykorzystajmy zmianę wartości pewnej wybranej cechy, a mianowicie dzietności przełomie wieków. Z badań demograficznych historii dziejów dzietność utrzymywał się mniej więcej na podobnym poziomie (umieralność podlegała dużo większym wahaniom) aż do rewolucji przemysłowej, kiedy trend się odwrócił. Obecnie cywilizacja zachodnia stoi u progu wymierania (dzietność spadła w wielu krajach poniżej progu reprodukcji społeczeństwa). Ekstrapolując zmniejszająca się już od wielu lat dzietność można by przewidywać, że społeczeństwa zachodnie będą wymierać. W niektórych badaniach wykorzystujących obecne tempo spadku dzietności wyznacza się czas zaniku populacji endemicznych (punkt czasu, kiedy etniczni mieszkańcy danego państwa będą stanowić mniejszość) i dla Francji jest to już mniej niż 50 lat od pierwszego wydanie tej książki. Jednak wielu

naukowców, wykorzystujących modelowanie ewolucyjne, twierdzi, że wcale tak nie musi być. Zgodnie z teorią Simmela, pewne cechy w jednostkach są aktywowane bądź nie w zależności od warunków zewnętrznych. Jeżeli transmisja cech ukrytych odbywa się drogą wertykalną, a aktywacja horyzontalną, bo możliwe są scenariusze zmiany trendu dzietności już w przyszłym pokoleniu. Otrzymane wyniki pokazały, że różnicowanie postaw (ukrytej – biologicznej, chęci posiadania potomstwa) wpływa na średnią wartość obserwowaną w populacji. Zmiana proporcji subpopulacji wyrażających chęć posiadania potomstwa, bądź nie jest jednak obserwowana bezpośrednio, a jednak po zadziałaniu funkcji aktywacji w postaci dzietności (Kolk, Cownden, & Enquist, 2014).

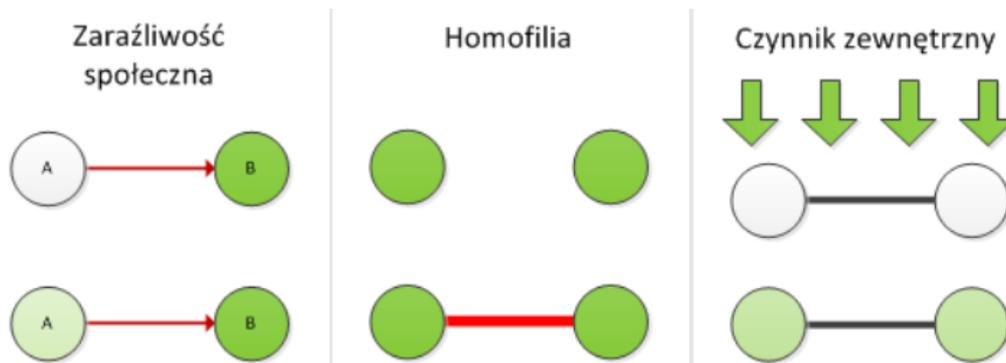
## 14.2 Modele pseudo-epidemiologiczne

Poza modelami medycznymi rozprzestrzenianie się poprzez zaraźliwość stosuje się do opisu zachowań dyfuzyjnych (Leskovec, Adamic, & Huberman, 2007) i innowacyjnych (Rogers, 2005) społeczeństwa oraz rozprzestrzeniania się złośliwego oprogramowania (Kasprzyk R. , 2010). W stosunku do ruchów masowych stosuje się podobne analogie jak do cyklu życia produktu (opisuje się różne fazy produktu, jak też różne kategorie użytkowników). Zazwyczaj początek tworzenia się ruchu ma charakter epidemiczny. Tak jak większość epidemii nie wybucha (stopa reprodukcji poniżej progu), większość ruchów społecznych również tłamsi się w zarodku. Jeżeli jednak patogen krąży w społeczeństwie (tak jak idea) mogą zajść warunki wystarczające do powstania ruchu masowego.



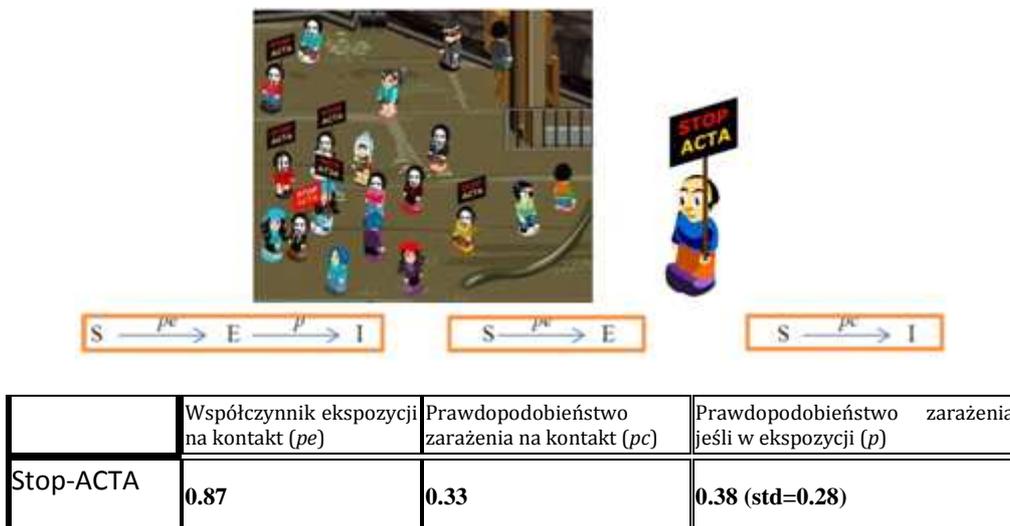
Rys. 70 Ewolucja poparcia/liczby członków ruchu w czasie

W ramach przeprowadzonych prac zbadano wpływ dynamiki ludzkich zachowań na proces formowania się opinii społecznej. Podobnie jak w modelach zasięgowych (np. odnośnie bioterroryzmu) wpływ pola zewnętrznego (np. wpływ mass mediów) musi być brany pod uwagę. Na sieciach społecznych klasyfikuje się działanie trzech mechanizmów: homofilii, zaraźliwości społecznej i czynnika zewnętrznego.



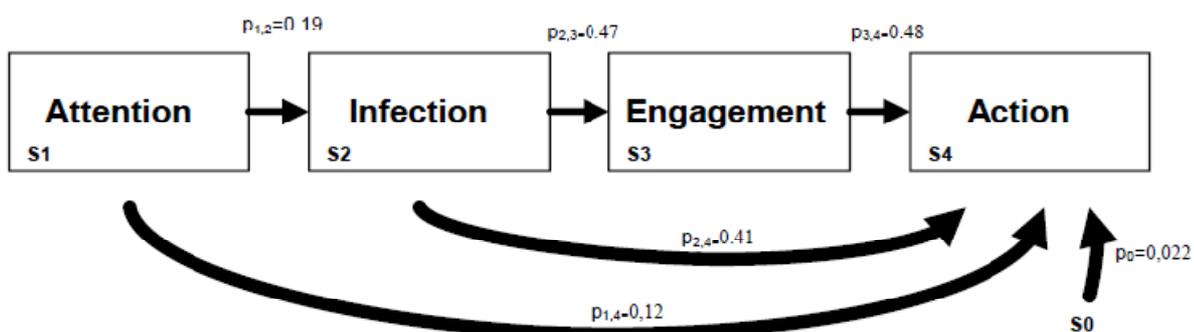
Rys. 71 Wizualizacja społecznych fenomenów na relacjach. Źródło: <http://www.lome.pl/model/>

Okazuje się, że dla każdego z mechanizmów (Rys. 71) węzły w grafie upodobniają się do siebie, jednak za każdym razem w inny sposób. Aby wnioskować o tym, który mechanizm miał największy wpływ na podobieństwo bliskich znajomych, używa się wskaźnika korelacji postaw pomiędzy Ego i Alter. Opinie, postawy czy normy to nie grypa i to w jaki sposób rozprzestrzeniają się wśród ludzi jest dużo bardziej skomplikowane i długotrwałe. Już nawet uproszczony model prezentowany w tym rozdziale oprócz zaraźliwości społecznej uwzględnia jeszcze dwa inne mechanizmy społecznego rozprzestrzeniania się opinii, a jest ich najprawdopodobniej więcej. Niemniej jednak rozprzestrzenianie się wirusa epidemii lub innowacji odbywa się według pewnego schematu, który obrazuje Rys. 70. I tak otyłość, czy palenie papierosów można traktować jako chorobę zakaźną (pomimo braku medycznego mechanizmu przenoszenia patogenu), ponieważ natężenie wspomnianej dysfunkcji (Alter) ma istotny statystycznie wpływ na stopień owej dysfunkcji u badanej jednostki (ego). Przykładem badań nad zaraźliwością społeczną dotyczących polskiego społeczeństwa jest analiza ruchu Stop-ACTA (Anti-Counterfeiting Trade Agreement, traktat podpisany przez rząd polski w październiku 2011 w celu ustalenia międzynarodowych standardów ochrony własności intelektualnych) w społeczności internetowej. W dostępnej literaturze zaproponowano schemat rozprzestrzeniania się idei oparty na modelu SEI (Susceptible, Exposed, Infective) (Jarynowski, Jankowski, & Zbieg, 2013) oraz na modelu adoptowanym *attention, infection, engagement and action* (Jankowski, Ciuberek, Zbieg, & Michalski, 2012).



Rys. 72 Modele zaraźliwości społecznej. Źródło: (Jarynowski, Jankowski, & Zbieg, 2013)

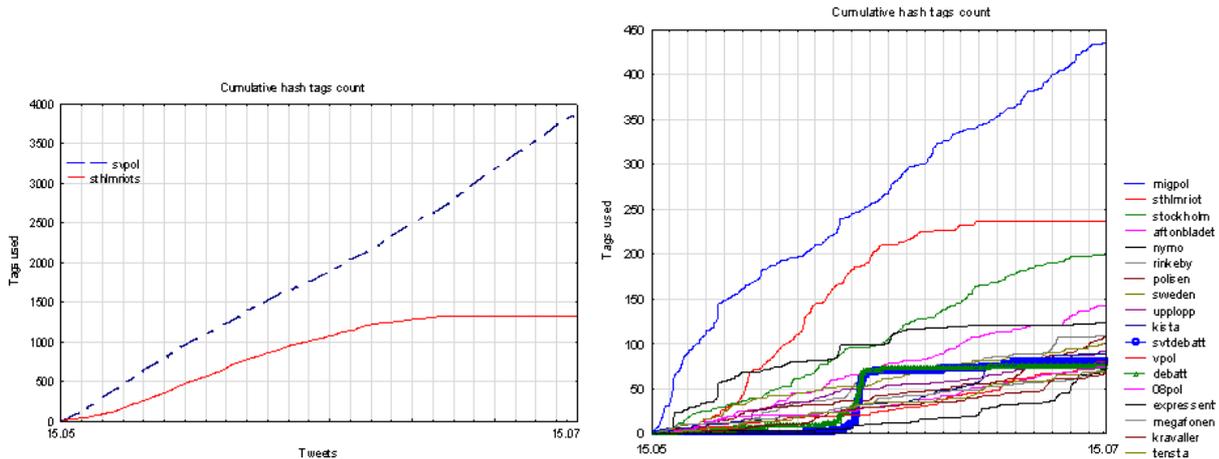
W obu przypadkach patogenem jest poparcie ruchu z tym że to poparcie jest podzielone na kategorie ze względu na zaangażowanie (podobnie dzieli się chorych na różne kategorie ze względu na ich stan zdrowia i poziom zaraźliwości). Na podstawie obserwowanej zaraźliwości społecznej i rozprzestrzeniania się ruchu sprzeciwu wobec ACTA dopasowano parametry modelu pseudo-epidemiologicznego.



Rys. 73 Model rozprzestrzeniania się ruchu społecznego z parametrami pseudo-epidemiologicznymi. Źródło: (Jankowski, Ciuberek, Zbieg, & Michalski, 2012).

### 14.3 Analiza mediów społecznościowych w czasie krytycznych wydarzeń

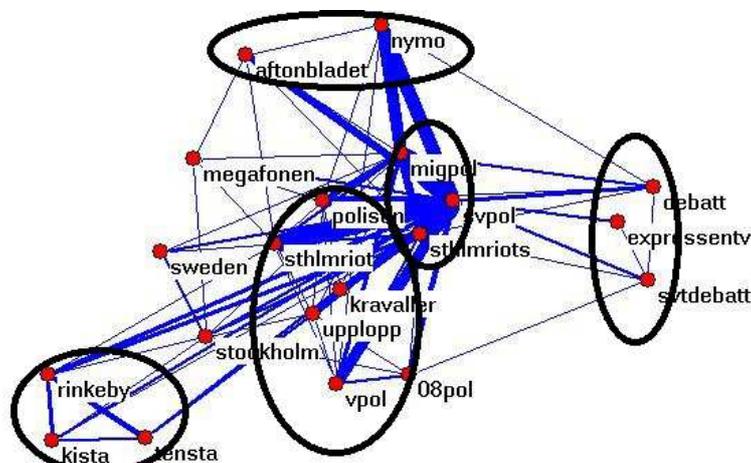
Twitter, Facebook, czy Fora Internetowe dostarczają wiele informacji odnośnie aktualnych nastrojów społecznych. W dużej mierze właśnie w tej sferze komunikacji społecznej kształtowane są opinie i rozprzestrzeniana jest informacja. W poprzednim podrozdziale wspomnieliśmy o obserwacji rozwoju akcji Stop-ACTA w Polsce. Teraz przyjrzyjmy się procesowi tworzenia się medialnej otoczki wokół zamieszek w Sztokholmie w maju 2013, które mimo że wcale nie odbiegały (według statystyk policyjnych poniesionych strat, przestępczości protestujących, itp.) rozmiarem od podobnych zajęć w roku poprzednim, zostały jako jedyne zauważone przez media krajowe i zagraniczne. Obserwacja słów kluczowych na Twitterze wskazuje na pewne uniwersalne zachowania w takich ekstremalnych przypadkach jak zamieszki.



Rys. 74 Skumulowane występowania 20 najczęściej pojawiających się hashtagów w sekwencji czasowej, po lewej 2 dominujące; Po prawej 18 kolejnych

Przede wszystkim warto zaobserwować, że „svpol”, czyli tager odnoszący się do szwedzkiej policji pojawia się ze stałą częstotliwością, mimo że same zamieszki się skończyły (o czym świadczy wygaszenie się krzywej „sthlmriots”, właśnie oznaczającej tager zamieszek). Ponadto wśród najczęstszych tagów znajdują się takie odzwierciedlające media, co łatwo poznać po schodkowym przebiegu krzywych skumulowanych występowalności tagów jak „svtdebatt” (debata w szwedzkiej telewizji).

Powiązania między tagami –współwystępowanie w obrębie jednego Tweeta również tworzą struktury dające się łatwo zinterpretować.

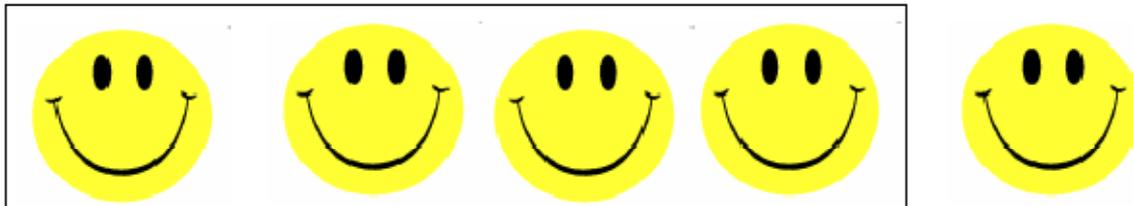


Rys. 75 Połączenie jako współwystępowanie tagów w jednej wiadomości z grubością połączenia jako jej częstotliwość.

W lewym dolnym rogu klastrują się dzielnice miasta objęte zamieszkami, po prawej stronie media telewizyjne, na górze media prasowe, a w środku określenia dotyczące policji i samych zamieszek.

### 14.4 Dynamika emocji

Badając wątku na forach dyskusyjnych i forach można zauważyć pewne prawa analizując automatycznie wypowiedzane frazy. Przyciąganie się emocji: obecność długiego grona o stałych emocjach zwiększa szanse następnego komentarza o tej samej emocji. Ponadto wątki o negatywnym nacechowaniu trwają dłużej od innych, co oznacza że podgrzanie atmosfery, przez negatywne wypowiedzi może być zasadne z socjotechnicznej perspektywy. Lojalni użytkownicy charakteryzują różnego rodzaju światów internetowych się znacznym negatywnym nastawieniem emocjonalnym.



Rys. 76 Emocje grupują się w jednolite ciągi. Prawdopodobieństwo, że nowo dołączony element będzie podobny do poprzednich wzrasta z liczbą poprzednich (rodzaj emocjonalnego społecznego dowodu słuszności)

### 14.5 Korupcja w piłce nożnej

Statystyka sportowa jest dziedziną bardzo mocno eksploatowaną przez media, zakłady i hobbistów. Jednak moim zadaniem nie jest przewidywanie wyników meczów (a przynajmniej nie podstawowym), co za tym idzie zamierzam wykorzystać inne metody, niż te wypracowane przez statystyków i probablistów. Pierwszym krokiem w tym kierunku było potraktowanie ligi piłkarskiej jako układu złożonego - czyli takiego, którego działanie jest skomplikowane jak czarna skrzynka do którego wkładane są drużyny z ich charakterystykami, a na wyniku jest tabela na zakończenie sezonu. To że nie można *a priori* powiedzieć, która z drużyn, będących elementami tego układu, zostanie mistrzem polski w następnym sezonie świadczy o powyższym. Ponadto fakt, że układ drużyn grających ze sobą niesie ze sobą emergencje, więc metody dobrze opisujące właściwości poszczególnych drużyn nie sprawdzają się na poziomie całej ligi. W związku z tym techniki jakich będę używał zostały wymyślone na potrzeby innych układów złożonych-jak giełda, czy układów oddziaływujących części.



Rys. 77 Współzależność występowania wyników w ramach rozgrywek Pierwszej Ligi Piłkarskiej w Polsce w sezonie 2006-2007 (kiedy wiele klubów zostało ukaranych za działania korupcyjne w celu ustawiania wyników)

Liga Piłkarska jako układ oddziaływujących ze sobą drużyn, poprzez rozgrywanie spotkań ligowych jest przykładem układu złożonego. Działanie takiego układu nie jest dokładnie poznane i wygląda trochę jak

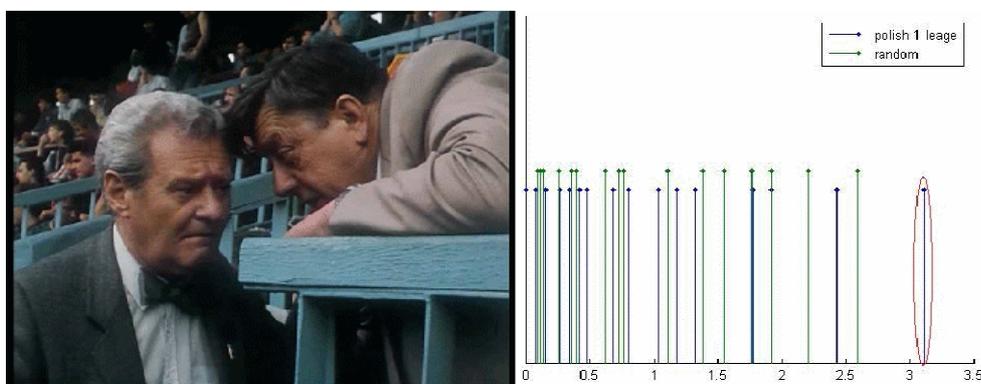
czarna skrzynka, do której wkładane są drużyny z ich charakterystykami, jako wynik otrzymujemy tabelę na zakończenie sezonu. Na brak przewidywalności i pełnej kontroli wpływa fakt, że nie można *a priori* powiedzieć, która z drużyn, będących elementami tego układu, zostanie mistrzem Polski w następnym sezonie. Dodatkowo obserwuje się tu emergencję<sup>8</sup>, więc metody dobrze opisujące właściwości poszczególnych elementów – zespołów piłkarskich nie sprawdzają się na poziomie całej ligi. W związku z tym można tu używać technik opracowanych wcześniej na potrzeby innych układów złożonych – jak giełda czy układy oddziaływujących cząstek. Dla przykładu, autor próbował wytropić ślady korupcji polskiej piłce nożnej (Jarynowski, 2010). Jeżeli przyjmiemy, że korupcja w polskiej piłce rzeczywiście występuje, to nasuwa się pytanie „jak ją mierzyć?”. Samo zjawisko korupcji jest na tyle skomplikowane, że w pierw wypadłoby ją zdefiniować, przynajmniej operacyjnie (konceptualizacja aczkolwiek ciekawa nie mieści się w obszarze badań fizyka). Pierwszym etapem analizy, było wyciągnięcie informacji o systemie z obserwacji anomalii. Mogłoby się wydawać, że czynnik losowy w sporcie jest decydujący. Dlatego właśnie występowanie anomalii, będzie świadczyło o ingerencji czynników niesportowych. W związku z tym, postawiono hipotezę zerową (zaprezentowaną już we wstępie pracy):

***H0a: Jeżeli wiele spotkań zostało „ustawionych”, to pewne statystyki polskiej ligi powinny się różnić od innych sportowych lig jak Bundesliga.***

***H0b: Powinny również występować anomalie (zdarzenia nieprzewidywalne) w porównaniu do stochastycznego wzorca ligi.***

W toku analizy okazało się, że polska piłka jest bardziej przewidywalna niż wzorcowa liga niemiecka, co świadczy na korzyść hipotezy H0a. Jednakże jednocześnie należy odrzucić hipotezę H0b (ponieważ dużo rzadziej niż przypuszczano na podstawie danych niemieckich, czy modelu stochastycznego obserwuje się zdarzenia ekstremalne, np. zwycięstwo ostatniej drużyny w tabeli z pierwszą). Dlatego też zdecydowano się przyjąć hipotezę alternatywną wobec H0b na poczet, której wysuwają się następujące interpretacje:

- częstsze zwycięstwa dobrych drużyn walczących o mistrzostwo bądź prawo gry w Europejskich Pucharach, jak również awans do wyższej klasy rozgrywek;
- podkładanie się przeciętnych drużyn, pewnych utrzymania w danej klasie rozgrywek, w meczach z drużynami walczącymi o mistrzostwo bądź Europejskie Puchary, jak również o awans do wyższej klasy rozgrywek;
- innego rodzaju przekupienia sędziów, piłkarzy czy działaczy, skutkujące zapewnieniem spodziewanego wyniku spotkania.



Rys. 78 Możliwe pozasportowe rozgrywki (lewy). Źródło: Klatka z filmu Piłkarski Poker Odstającej wartości własne dla Pierwszej Ligi Piłkarskiej

<sup>8</sup> Właściwości układu są czymś więcej niż sumą właściwości elementów tworzących ten układ.

Powyższe stwierdzenia świadczyłyby o przewadze zjawisk korupcyjnych wynikających z dążenia na zachowania równowagi – to jest zespoły zadowolone ze swojego miejsca w tabeli mogłyby używać wszystkich dostępnych środków, aby zachować swoją pozycję.

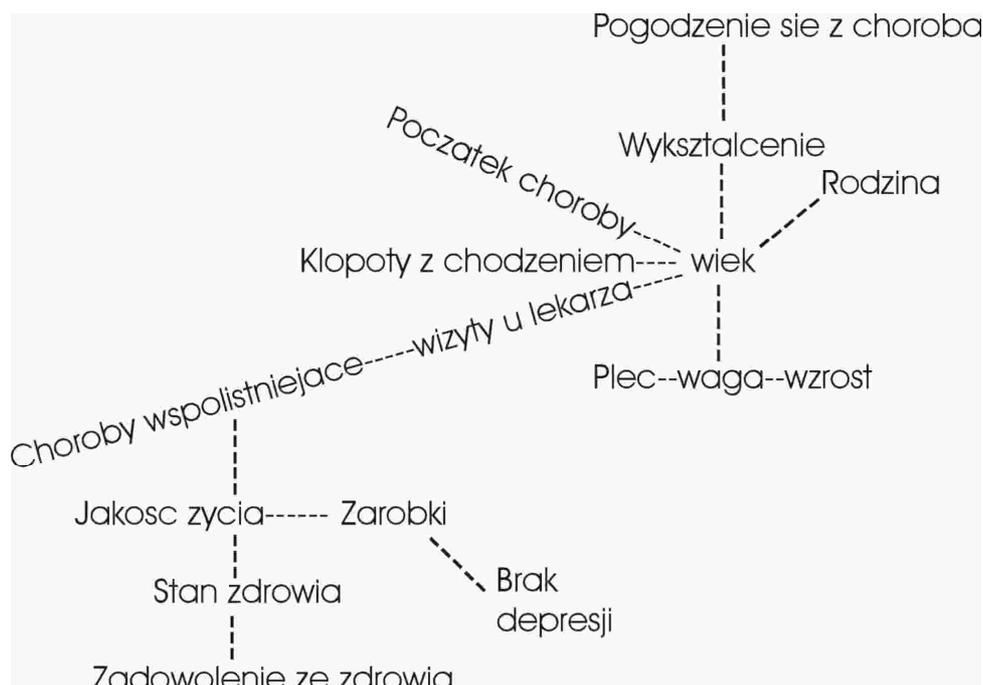
### 14.6 Zastosowanie technik wizualizacji sieciowych zależności na przykładzie oceny jakości życia osób ze zdiagnozowaną dysfunkcją tarczycy

W związku ze wzrostem zapotrzebowania na cyfrową analizę danych epidemiologicznych przedstawiamy przykład wykorzystania technik MST (ang. Minimal Spanning Tree) oraz SNA (ang. Social Network Analysis) w medycynie kontekście związków o skomplikowanej naturze przyczynowo-skutkowej.

Zgodnie z tradycją medycyny zależności te opisywane i sprawdzane są na ogół poprzez testowanie hipotez statystycznych. Ta metoda nie zawsze się jednak sprawdza, zwłaszcza gdy mamy do czynienia ze słabszymi korelacjami i nie jesteśmy w stanie stwierdzić, które cechy w sposób istotny wpływają na inne, które wpływają bezpośrednio, a które za pośrednictwem innych cech. Taki właśnie słabsze korelacje występują choćby w badaniach samooceny jakości życia osób ze zdiagnozowaną dysfunkcją tarczycy.

Do prezentacji przewagi technik nieliniowych wykorzystaliśmy przykład badania oceny jakości życia przeprowadzony w Poznaniu wśród 86 pacjentów za pomocą kwestionariusza (WHOQOL-BREF) składającego z 26 pytań w zakresie czterech dziedzin: somatycznej (fizycznej), psychologicznej, socjalnej oraz środowiskowej i dodatkowych pytań dotyczących samooceny stanu zdrowia i subiektywnej percepcji jakości życia. Wyniki zostały unormowane tak, aby móc policzyć współczynniki korelacji między badanymi cechami.

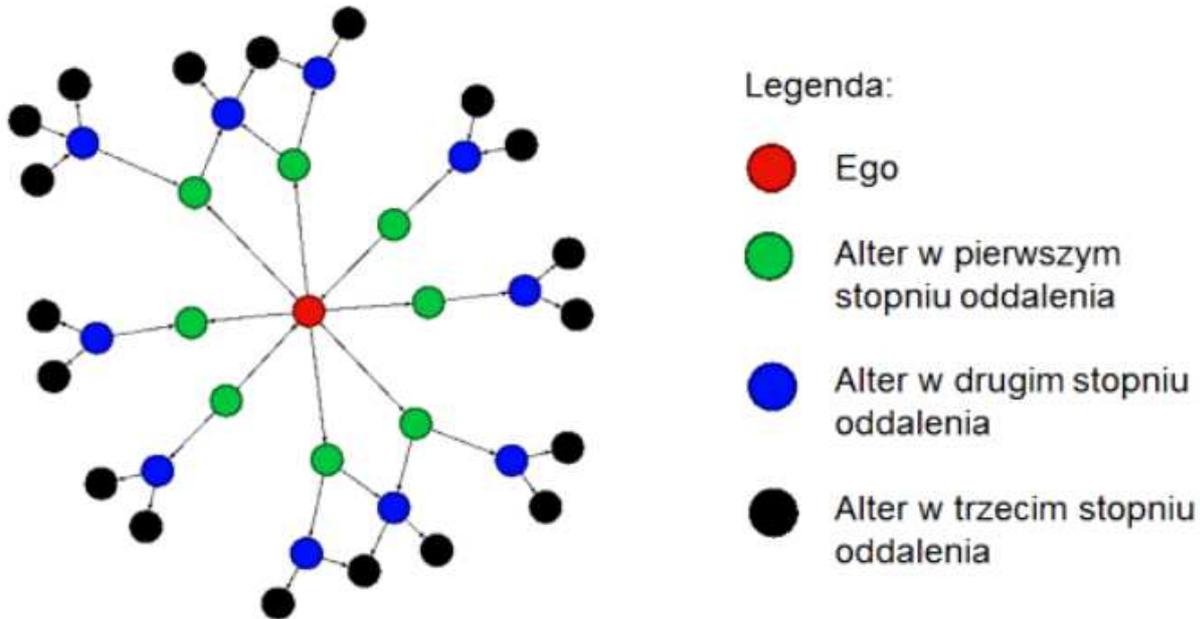
Okazało się, że szukając związku badanych cech z jakością życia od strony tradycyjnej w medycynie metody testowania hipotez otrzymujemy ledwie kilka istotnych statystycznie korelacji. To jest trójkąt: płeć, waga i wzrost oraz stan zdrowia połączony z jakością życia. Takie przedstawienie cech wpływających na jakość życia jest mało zadowalające poznawczo, a wizualizacja MST oraz SNA daje czytelne z rozpoznaniem, które cechy wpływają na jakość życia bezpośrednio, a które pośrednio – co z pewnością jest nową jakością dla badań epidemiologicznych.



Rys. 79. Drzewo minimalnej rozpiętości dla badanego układu korelacje między cechami wpływającymi na jakość życia. Istotne ( $\alpha = 0.05$ ) korelacje jedynie w obrębie trójka waga-wiek-płeć

## 15 Analiza sieciowa w zastosowaniach

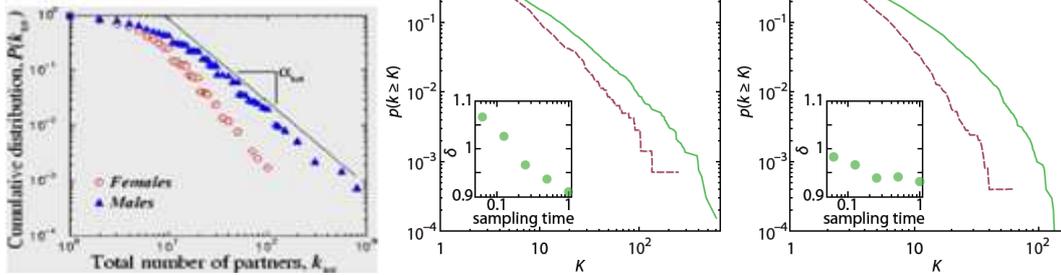
W badaniach społecznych, bądź marketingowych bardzo często odtwarza się sieci społeczne w zastanej społeczności. Po odpowiedniej przeróbce danych, wizualizacji i analizie można wyciągać wiele informacji o właściwościach danej społeczności. Fizycy jednak nie poprzestają na tym. Na sieciach empirycznych i losowych (zazwyczaj budowanych według wspomnianych wyżej schematów) modeluje się ogromną ilość zjawisk od rozprzestrzenia się chorób zakaźnych, idei jak choćby w masowych protestach przeciwko ACTA (Jarynowski, Zbieg, 2013), po zdarzenia kryminalne (Jarynowski, Rostami, 2013). Sama sieć, ze względu na swój abstrakcyjny charakter, może być nośnikiem informacji o kontaktach telefonicznych, personalnych, mailowych, instytucjonalnych, itp. Wierzchołkami mogą być ludzie, grupy ludzi, instytucje itp.



Rys. 80 Koncepcja sieci Ego, czyli najbliższego otoczenia każdego węzła w sieci relacji. Źródło: <http://www.lome.pl/model/>

### 15.1 Sieci seksualne

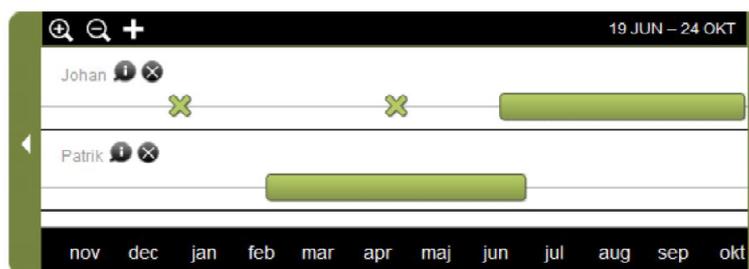
Kolejną przestrzenią do badań dla socjofizyki jest zagadnienie dotyczące sieci kontaktów seksualnych, w kontekście rozprzestrzeniania się chorób. Ludzka seksualność należy do zachowań intymnych, dlatego też dostęp do danych na temat jest utrudniony. Autor opierał się w większości na danych szwedzkich, ze względu na wyjątkowo wysoki stopień ingerencji państwa w życie obywateli, co skutkuje ogromną wiedzą o społeczeństwie. Jednocześnie daje to możliwość pracy w oparciu o informacje, które nie są dostępne nigdzie indziej. Zatem podstawą do modelowania jest wiedza o kontaktach społecznych i liczbie zarażeń, co jest przedmiotem precyzyjnej rejestracji. W przypadku stosunków seksualnych dostępnych jest niewiele zbiorów danych, a najbardziej reprezentatywne wydają się być opracowania statystyczne zachowań seksualnych mieszkańców wyspy Gotlandii oraz prostytutki w Brazylii. W pierwszym przypadku, w ramach spisu powszechnego w Gotlandii, liczącej 50 tysięcy mieszkańców, zostało zadane pytanie o liczbę partnerów seksualnych. Sieci prostytutki zostały zbudowane na podstawie postów zamieszczanych przez Brazylijczyków na forum internetowym dotyczącym płatnego seksu.



Rys. 81 (lewy) Kontakty seksualne mieszkańców wyspy Gotlandii Źródło (Liljeros, 2001): Kontakty seksualne prostytutek (A) i ich klientów (B) Źródło (Rocha, Liljeros, & Holme, 2011):

Do każdego z prezentowanych rozkładów krotności wierzchołków (Rys. 81) można orientacyjnie dopasować prawo potęgowe (linia prosta w skali podwójnie logarytmicznej). Oznacza to, że empiryczne sieci epidemiologiczne można zaklasyfikować jako bezskalowe w pierwszym przybliżeniu. Należy zwrócić uwagę na jednostkom o skrajnie dużej ilości kontaktów, bo mogą być oni w przypadku zarażenia 'super-roznościcielami'.

Jeżeli chodzi o choroby przenoszone drogą płciową, to rodzaj patogenu będzie miał tu kluczowe znaczenie. Wirusa HIV (dla którego prawdopodobieństwo zarażenia w czasie pojedynczego stosunku jest małe 0,08-1,2%) można przeciwstawić bakteriom Chlamydia (gdzie to prawdopodobieństwo jest wysokie 7-40%). Właśnie analiza sieciowa pozwala stwierdzić, że społeczeństwa zachodnie są bardziej podatne na zarażenia Chlamydia, ze względu na dużą liczbę kontaktów przygodnych, natomiast społeczeństwa z dopuszczalnymi poligamicznymi związkami, jak na przykład społeczeństwo arabskie, są bardziej narażone na zarażenie wirusem HIV, ze względu na intensywność współżycia seksualnego z wieloma partnerami. Najcenniejszym środkiem do wykonywania analiz byłaby się baza prawdziwych kontaktów, która od początku byłaby zorientowana sieciowo. W tym celu zaprojektowano badanie, dzięki któremu udało się odtworzyć sieć kontaktów seksualnych (która wymagałaby później stochastycznego uzupełnienia ze względu na braki danych wynikające z wrażliwości tematyki). Badanie zostało przygotowane w postaci anonimowej ankiety (Rys. 82), w której każdy respondent miałby stworzyć listę partnerów seksualnych (za pomocą aliasów, a nie prawdziwych imion) i wskazać na osi czasu, kiedy dochodziło do zbliżeń.



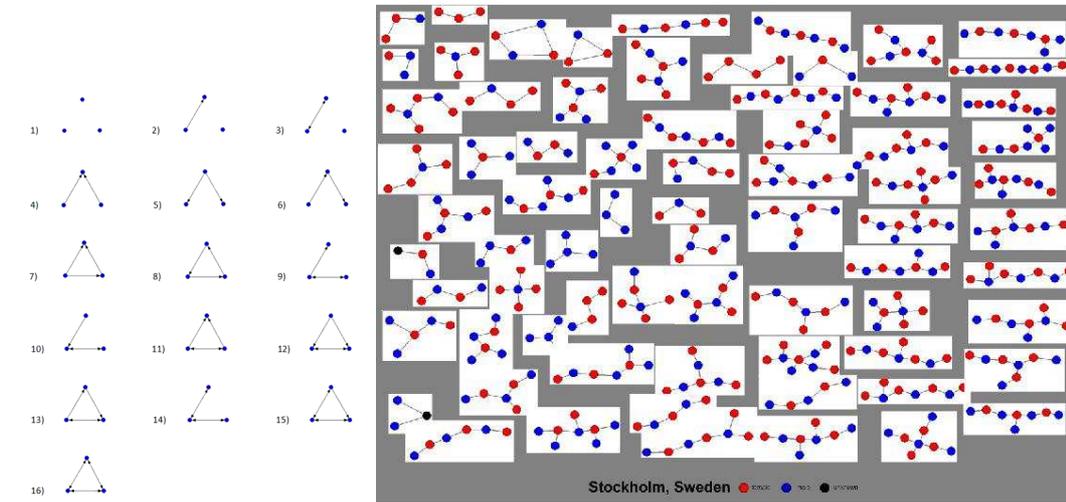
Rys. 82 Layout ankiety opracowanej do anotacji partnerów seksualnych

Zbudowany model, który opisywałby rozprzestrzenianie patogenów, pozwoliłby przygotować skuteczne narzędzia w walce z chorobami zakaźnymi, mógłby także pomóc w podejmowaniu skutecznych akcji przez nadzór epidemiologiczny. Ponieważ problem wymaga wiedzy z zakresu różnych dziedzin nauki, jedynie współpraca naukowców z różnych środowisk naukowych, nie tylko lekarzy, ale również: fizyków, matematyków, statystyków, informatyków i socjologów może przyczynić się do rozwoju nauki. Dzięki tak daleko idącej interdyscyplinarnej współpracy możliwym stałoby się zreplikowanie symulacyjne epidemii i budowanie modeli predykcyjnych, które po wdrożeniu, z pewnością przyczyniłyby się do poprawy jakości życia wielu ludzi, którzy mogliby uniknąć zakażenia. Już wstępna analiza dotycząca wirusa HPV pozwala wyciągnąć wnioski dotyczące optymalnej profilaktyki raka szyjki macicy w Polsce i wskazuje na konieczność zmiany strategii, ze względu na zmieniające się warunki epidemiologiczne i demograficzne i potencjalny

ponowny wzrost liczby nowych przypadków raka. Niestety, należy oczekiwać pogorszenia się sytuacji epidemiologicznej chorób przenoszonych drogą płciową mimo wzrostu świadomości społecznej i medycznej w ostatnich latach w naszym kraju.

## 15.2 Analiza motywów sieciowych – sieci rosnące i malejące

W schematach komunikacyjnych na trójkącie wyróżniamy 16 podstawowych motywów. Częstotliwość ich występowania została zbadana dla wielu układów społecznych.



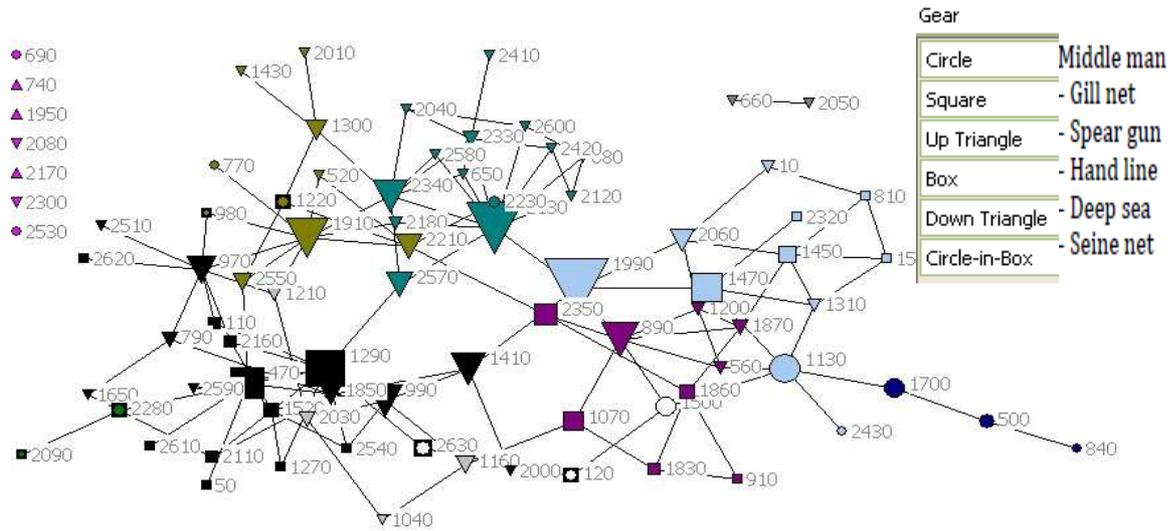
Rys. 83 Motywy sieciowe, często wykorzystywane w opisie schematów komunikacyjnych, gdzie kierunek relacji ma znaczenie (lewy). Motywy nieskierowane związków małżeńskich w Sztokholmie (prawy). Źródło: Xin Lu

Popularna stała się analiza maili instytucjonalnych pod tym kątem. Innym ciekawym zagadnieniem jest rozrost i kurczenie się sieci. Przykładowo jeżeli A i B oraz B i C są połączeni ze sobą, to z dużym prawdopodobieństwem połączenie między A i C zostanie stworzone (zasada balansu). W przypadku zmniejszania się sieci często degradacja zostaje przyspieszona jeżeli doszło do wydzielenia się więcej niż pary rozdzielonych sieci.

## 15.3 Sieci geograficzne w układach człowiek - przyroda

Typowa prezentacja sieci nie reprezentuje jej przestrzennego układu i jest często wynikiem działania algorytmu. Czasami jednak taka sieć powinna odzwierciedlać informacje geograficzne, ze względu na specyfikę badań i tak często jest w przypadku modelowania interakcji człowiek-przyroda. Proszę zwrócić uwagę, że modelowanie człowiek – środowisko (reprezentowane przez przyrodę, ale również przemysł, infrastrukturę itp.) jest jedną z najszybciej rozwijających się w XXI w. obszarów praktycznych obliczeniowych nauk społecznych.

Weźmy przykład układu społeczno-ekologicznego rybaków z Madagaskaru. W wyniku nadmiernego połowu jednej z ryb ekosystem się załamał i wielu z rybaków musiała opuścić swoje miejsce zamieszkania z powodu braku ryb. Przeanalizujemy sytuację takiej wioski rybackiej, tak aby w przyszłości nie doszło do takiej sytuacji.



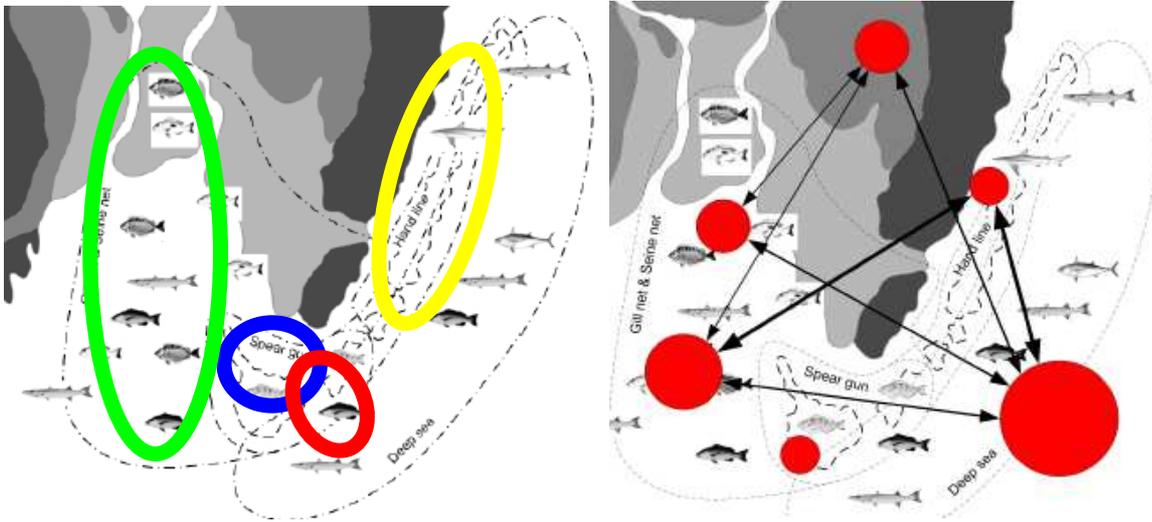
Rys. 84 Zależności między rybakami z uwzględnieniem techniki rybackiej

Przedstawia sieć (Rys. 84) dotyczy transferu wiedzy i powstała z wywiadów z rybakami w oparciu o pytania typu: „od kogo czerpiesz branżowe informacje”. W związku z ograniczonymi zasobami rybnymi i terytorialnymi rybacy reprezentujący różne typy połowów konkurują ze sobą o miejsce i konkretne gatunki ryb.

Ryba	Sposób połowu	Obszar konfliktu
Baracuda	Gill, hand, deep, saine	Yellow
Snapper	Gill, spear, deep, saine	Yellow, Red
Kawakawa	Hand, deep, saine	
Rabbitfish	Gill, saine	Green
Emperor	Gill, saine	Green

Tabela 4 Zestawienie możliwych konfliktów terytorialnych, o zasoby oraz ze względu na stosowaną technikę połowu

Patrząc na sieci powiązań w ramach zarządzania zasobami naturalnymi (ang. Natural Resources Management - NRM), można zaobserwować pewne wspólne struktury (najsilniejszy związek między “gill” z “saine” i silne powiązane między “deep” z “gill&saine”). Można założyć, że sieć relacji społecznych opartych na wiedzy (SNA) będzie związana z siecią wspólnych zasobów i obszarem połowów.



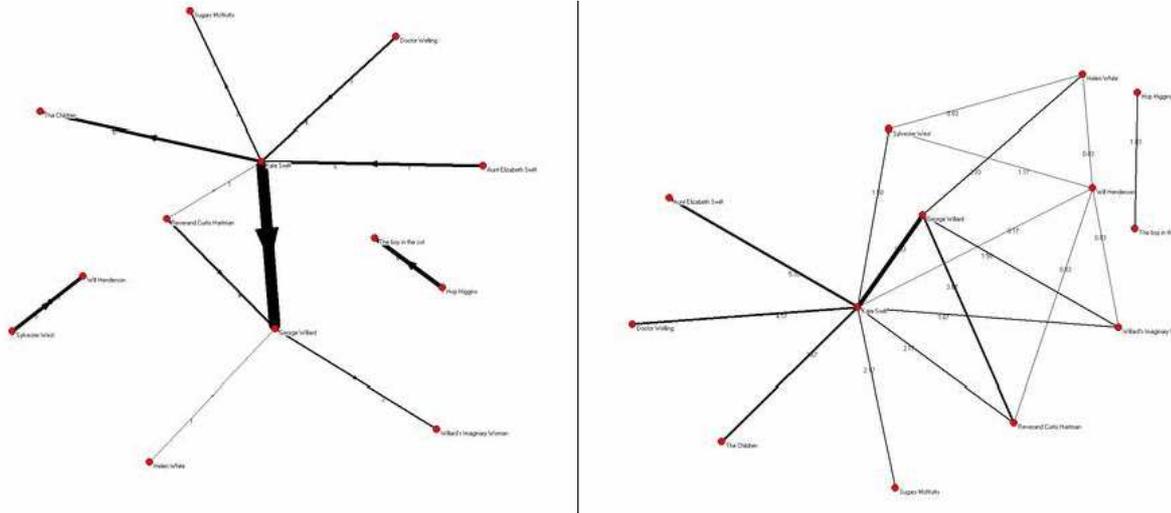
Rys. 85 Wizualizacja obszarów oddziaływań metodą zarządzania zasobami naturalnym – NRM (lewy) oraz klasycznej aniazy sieci społecznych – SNA (prawy). Na podstawie: (Bodin & Crona, 2014)

Możemy porównać sieci społeczne (SNA) i ekologiczne (NRM) i zobaczyć dowody, że główna rola w społeczeństwie z obu perspektyw grają głęboko-morscy rybacy „deep”.

Na podstawie powyższej analizy pewne wnioski odnośnie możliwości lepszego gospodarowania zasobami można wyciągnąć. Należy poinformować większość osób centralnych w sieci (Rys. 84) w celu przyspieszenia rozprzestrzeniania informacji na temat ograniczeń limitów połowowych na poszczególne gatunki ryb oraz miejsc. Dalsze instrukcje wymagają uwzględnienia większej ilości informacji ekologicznych, co wykracza poza ramy tej książki, ale już wstępna analiza pokazuje możliwości analizy sieci społecznych.

#### 15.4 Sieci w literaturze

Odkrywanie sieci społecznych w literaturze jest pasjonującym tematem ciągle o charakterze pionierskim. Zostały jednak już opracowane pewne metody pozyskiwania sieci na podstawie ankiet czytelników (Jarynowski, Boland, 2013). Pierwotnie badania zostały przeprowadzone na zbiorze opowiadań: *Winesburg, Ohio*. Już na podstawie uzyskanych wyników można wyciągnąć pewne wnioski i postawić nowe hipotezy badawcze. W omawianej pracy sposób zbierania informacji od czytelnika został uproszczony do minimum, w celu zachowania możliwości porównania z również wypracowanymi równolegle metodami komputerowymi. M.in. dlatego nie umieszczono respondenta w żadnym konkretnym kontekście czytelniczym. Już przedwstępna analiza pokazała, że sposób zadawania pytań ma znaczenie. Dlatego zostawiono mu pole do interpretacji interakcji między bohaterami w opowiadaniu wedle jego rozumienia (*reader world*). Zamiast pytania o interakcje odczytywane przez czytelnika (*reader world*), można by zastanowić się nad kwestią znaczenia interakcji dla fabuły (*story world*). W ten sposób percepcja czytelnicza została by ukierunkowana na świat przedstawiony w utworze, nie zaś na jego subiektywną wizję (Welsh, 2011). Być może zmieniłoby to uzyskane wyniki i zbliżyłoby metody ankietowe do komputerowych. Pozostaje to jednak otwartym zagadnieniem, podobnie jak pytanie o związek między interakcjami między bohaterami opowiadania typu *implicite* i *explicite*.



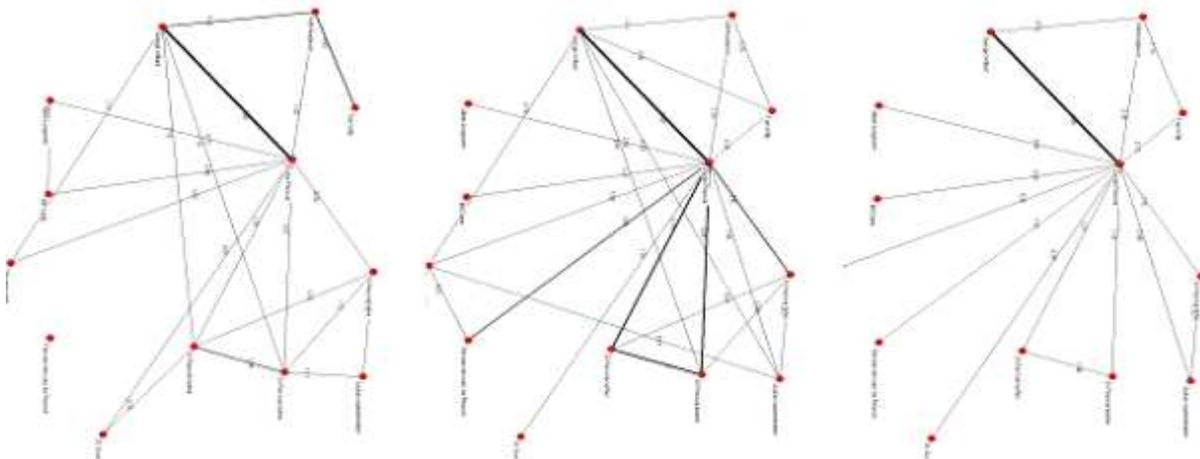
Rys. 86 Sherwood Anderson "Winesburg, Ohio" *The Teacher*. Ekstrakcja ręczna (ludzka) sieci explicite i implicite

W celu porównania powyższych metod (dwóch ludzkich: pierwszej komunikacyjnej i drugiej interakcyjnej oraz dwóch komputerowych: akapitowej i zdaniowej) dokonano normalizacji wyników. Uzyskane sieci można przedstawić graficznie (Rys. 87). Widać, że wyglądają one podobnie. Najważniejsze połączenie z punktu widzenia fabuły (pomiędzy dziennikarzem a doktorem) w każdej metodzie jest uwypuklone. Zauważalną różnicą jest za to różna ilość połączeń w sieciach. Zadanie pierwsze-komunikacyjne wykazuje najmniej połączeń (16), a drugie-interakcyjne najwięcej (23), podczas gdy algorytm akapitowy daje wynik pośredni (21).

	Algorytm akapitowy	Algorytm zdaniowy
AD 1 (komunikacyjna)	0,84	0,91
AD 2 (interakcyjna)	0,70	0,58

Tabela 5 Macierz korelacji pomiędzy sieciami otrzymanymi od respondentów i za pomocą algorytmu.

Jednak, odnosząc się do uzyskanych wyników, najważniejsze wydają się korelacje między rezultatami ankiet oraz algorytmów. Niezmiernie ciekawym aspektem jest inny poziom korelacji pomiędzy odpowiednimi metodami komputerowymi a różnymi zadaniami czytelnickimi, co wskazuje na niejednoznaczność w sposobach opisywania interakcji i wrażliwość na czynniki zewnętrzne.



Rys. 87 Uzyskane sieci społeczne opowiadania *The Philosopher* z zadania pierwszego-komunikacyjnego (lewo), z zadania drugiego-interakcyjnego (prawe) oraz algorytmu komputerowego - akapitowego (w środku).

Oznaczać to może, że nie istnieje jedna uniwersalna metoda ukazująca sieć społeczną w utworze literackim, co jest paradygmatem wciąż rozwijającej się dziedziny nauki, jaką jest NLP (ang. natural language processing). Interesujące wydaje się również zaprezentowanie istotnego statystycznie związku między sposobem postrzegania sieci a wykształceniem (reprezentanci nauk ścisłych budują sieci bardziej powtarzalne niż reszta). Na koniec sprawdzono, że mierząc intensywność interakcji nie można przełożyć tego bezpośrednio na budowanie napięcia w utworze. Podsumowując, wydaje się, że analiza sieci społecznych w literaturze, która, ze względu na charakter ilościowy, może być narzędziem uzupełniającym w warsztacie współczesnego humanisty.

Wykształcenie	Parametr	p-Value
Artystyczne i humanistyczne	0,05	0,37
Nauki społeczne	-0,03	0,71
Nauki ścisłe i medyczne	0,16	<0,01

**Tabela 6** Regresja logistyczna zmiennych nominalnych pokazująca relację między wykształceniem ścisłym a zgodnością w uzupełnianiu zadania pierwszego-komunikacyjnego oraz drugiego-interakcyjnego

Porównajmy teraz korelacje pomiędzy sieciami uzyskiwanymi przez tych samych respondentów w zadaniu pierwszym-komunikacyjnym i drugim-interakcyjnym. Zgodność odpowiedzi spróbowaliśmy wyjaśnić za pomocą zmiennych niezależnych, podanych przez respondenta w metryczce za pomocą regresji logistycznej zmiennych o wymiarze nominalnym (Tabela 6). Okazało się, że tylko rodzaj wykształcenia ma istotny statystycznie wpływ na te korelacje: tylko respondenci o wykształceniu ścisłym odpowiadali w wyraźnie podobny sposób. I tak: osoba z wykształceniem ścisłym wykazuje się większą powtarzalnością w obu zadaniach.

Opisana metodologia może pomóc w weryfikowaniu hipotez odnoszących się do konkretnych tekstów, autorów czy też ogólnie – literatury. Stworzone tu narzędzie jest na tyle uniwersalne, że może zostać użyte – po dokonaniu niewielkich modyfikacji – w analizie innych tekstów literackich, tworzonych w różnych językach (choć po polsku tylko częściowo, gdyż cyfrowy korpus języka polskiego dopiero od niedawna daje się stosować w badaniach *stricte* ilościowych (Głowacka, 2012)).

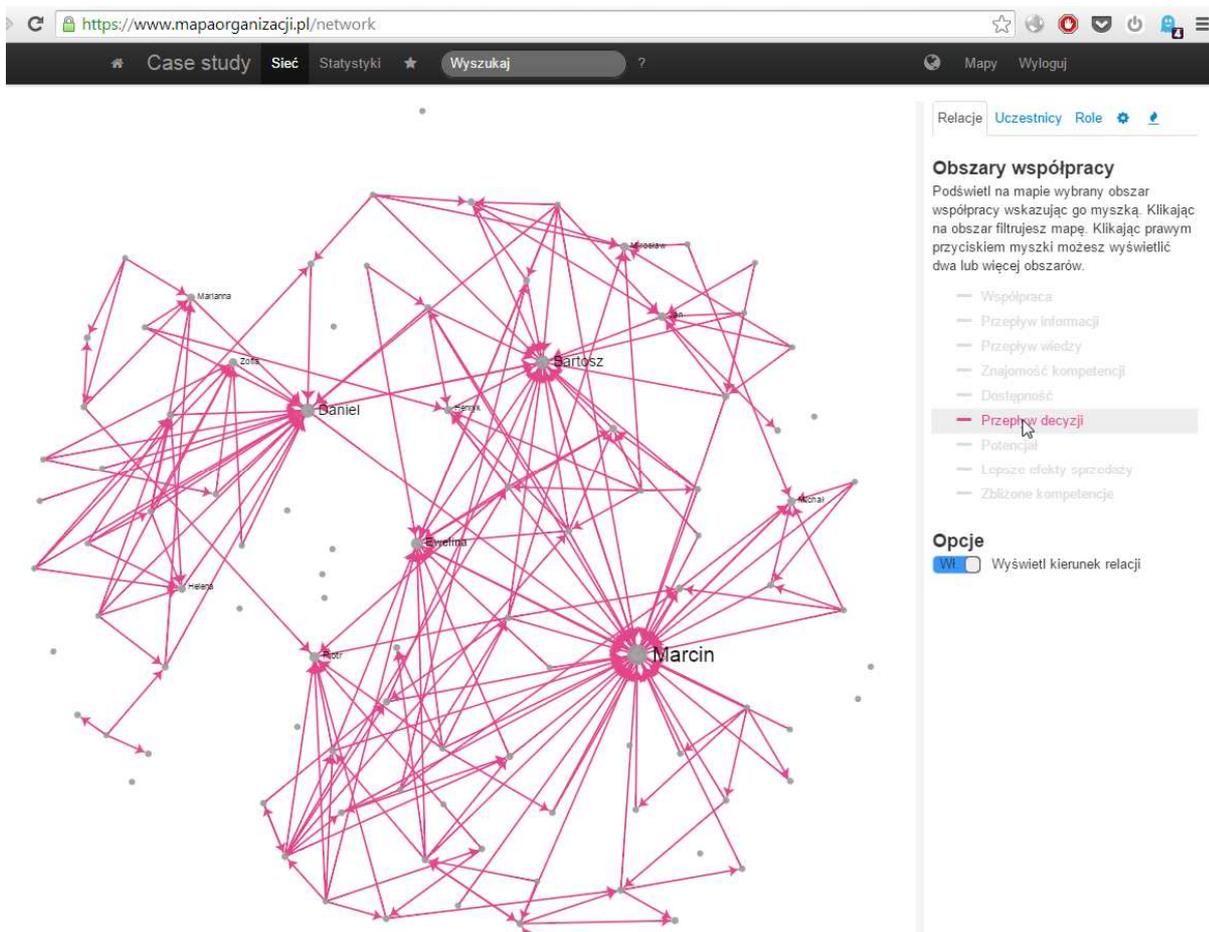
## 15.5 Analiza sieciowa organizacji

Analiza sieci jest wykorzystywana także w ekonomii gdzie funkcjonuje pod pojęciem analizy sieci organizacyjnych (*ONA - Organizational network analysis*). Sieci organizacyjne są najczęściej rozumiane dwojako. Sieć może reprezentować zbiór relacji łączących pracowników danej firmy (tzw. sieć wewnątrzorganizacyjna), lub zbiór relacji pomiędzy organizacjami (tzw. sieci międzyorganizacyjne). Może więc służyć do analiz struktury zespołów, firm, holdingów, klastrów lub całych sektorów gospodarki. W badaniach teoretycznych można spotkać każdą z powyższych form analiz sieciowych (Borgatti, Foster, 2004), natomiast w praktyce analizę sieci stosuje się zazwyczaj w celu optymalizacji struktury firmy względem takich jej wymiarów jak: współpraca i przepływ zadań, przepływ decyzji, komunikacja i przepływ informacji oraz przepływ wiedzy (Cross, Parker, Cross, 2004).

Formalne schematy organizacyjne są punktem odniesienia dla większości decyzji w zarządzaniu, jednak prezentowany przez nich obraz jest bardzo uproszczony i statyczny. Analiza sieci pozwala na uzyskanie wglądu w rzeczywiste sposoby organizowania się pracowników, które zazwyczaj opierają się na doświadczeniach pracowników i praktyce dynamicznie wypracowywanej w kontakcie z rynkiem. Jest to wiedza potoczna, często niejawna (*tacit knowledge*) i trudna do zobrazowania, czy zwerbalizowana przy użyciu dotychczas stosowanych analiz, np. analizy SWOT pracownika czy oceny 360 stopni (Zbieg, Żak, et al. 2014).

Sieć organizacyjna to struktura firmy przedstawiona jako zbiór relacji łączących pracowników (Rys. 88). Węzłem na grafie jest pracownik, połączenie to np. fakt współpracy pomiędzy pracownikami. W celu odzwierciedlenia sieci organizacyjnej stosuje się właśnie techniki statystyczno – wizualizacyjne SNA. Jednym z dostępnych narzędzi na w Polsce jest Mapa Organizacji, narzędzie przydatne dla naukowców, konsultantów i menadżerów pracujących nad optymalizacją funkcjonowania lub rozwojem organizacji (Żak, Zbieg, et al, 2014) . Ponadto, ta partycypacyjna platforma badań sieci organizacyjnych stworzona przez polskich badaczy sieci umożliwia udostępnienie wszystkim pracownikom widoku i podstawowych parametrów sieci, zgodnie z badaniami empirycznymi wskazującymi na pozytywny efekt informacji zwrotnej i wiedzy na temat całej sieci względem zwiększenia jakości koordynacji pracy w sieci (Enemark, McCubbins, Weller, 2014).

Przykładowy zrzut ekranu aplikacji (Rys. 88) prezentuje efekt odwzorowania sieci powstałej w wyniku przeprowadzenia wśród pracowników automatycznego badania kwestionariuszowego, dzięki któremu można ujawnić strukturę sieci w oparciu o takie parametry jak przepływowość informacji (średnia długość ścieżki), podział na grupy rzeczywistego współdziałania (wykrywanie społeczności i klastrowanie), czy koncentrację zasobów czy istotność jednostek (centralność węzłów).



Rys. 88 Przykładowa sieć współpracy w firmie. Źródło: <https://www.mapaorganizacji.pl/>

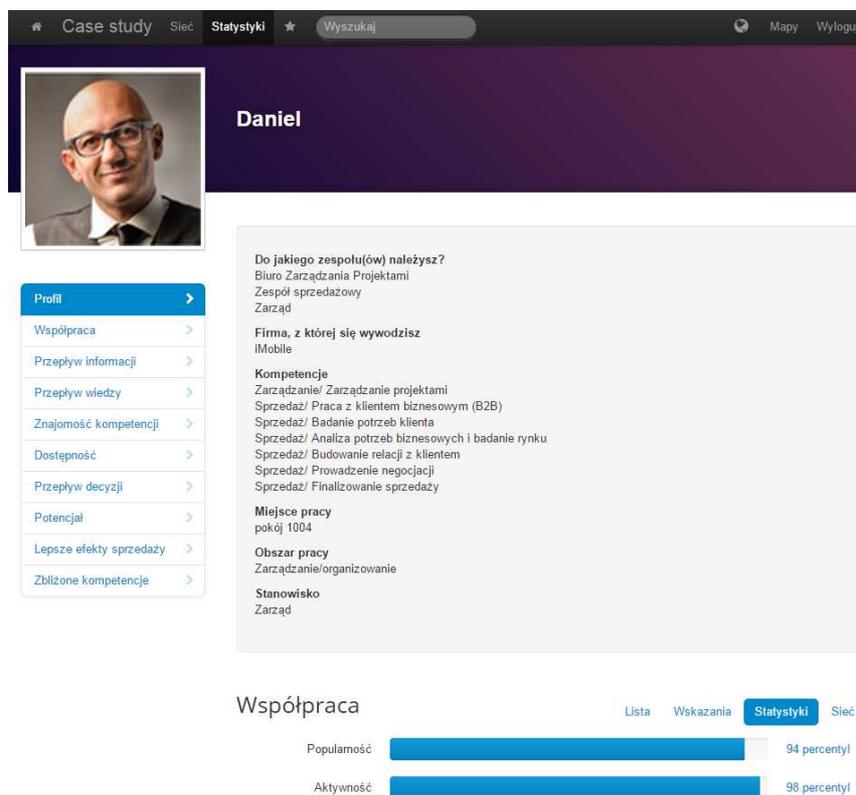
Wykorzystanie analizy sieci w zarządzaniu i doskonaleniu struktury firmy dostarcza informacji stosunkowo trudno dostępnych przy użyciu innych metod, a następnie zaplanowanie na ich podstawie odpowiednich działań:

- Zidentyfikowanie osób istotnych dla procesu współpracy, przepływu informacji, decyzji czy dzielenia się wiedzą (np. osób łączących wiedzę z różnych obszarów, osób nadmiernie koncentrujących władzę, osób posiadających szybki dostęp do informacji, itp.) oraz zreorganizowanie i/lub korekta ich obowiązków w taki sposób, by mogli pełnić swoje nieformalne funkcje,

- Zlokalizowanie barier w przepływie zasobów (współpracy, informacji, wiedzy, władzy) i określenie źródła ich powstania (np. staż pracy, przynależność do grupy, miejsce pracy, różnice w pozycji formalnej, itp.), co umożliwia podjęcie działań zmierzających do zmniejszenia blokad, np. przeorganizowanie miejsca pracy pracowników,

- Możliwość trafnego zaplanowania działań zwiększających stopień wykorzystania potencjału (współpracy, przepływu informacji, władzy i wiedzy) tkwiącego w firmie, jak np. nawiązywanie kluczowych połączeń pomiędzy specjalistami z danych dziedzin czy zmniejszenie tendencyjności nawiązywania relacji z osobami podobnymi pod względem posiadanych informacji,

W procesie prowadzenia zmiany organizacyjnej ważna jest także autodiagnoza własnego zaangażowania pracowników w proces dzielenia się zasobami (pracy, informacji, władzy i wiedzy) dokonana przez każdego z pracowników oraz zwrócenie uwagi na potrzebę wymiany zasobów oraz promowanie pożądaných postaw, możliwy dzięki spersonalizowanym statystykom wbudowanym w Mapę organizacji (Żak, Zbieg, et al, 2014), co przedstawia Rys. 89.



Rys. 89 Przykład statystyk sieciowych spersonalizowanych dla każdego pracownika. Źródło: <https://www.mapaorganizacji.pl/>

I wreszcie, w analizie sieci organizacyjnych ważny jest element monitorowania zmian w czasie. I tak, w prezentowanym narzędziu analitycznym możliwe jest monitorowanie procesu zarządzania w czasie przepływem zasobów, np. śledzenie przepływu wiedzy między działami czy szczeblami hierarchii służbowej, lub liczba relacji na pracownika i gęstość sieci dla każdej z obserwowanych relacji co prezentuje Rys. 90.

Case study Sieć Statystyki Wyszukaj

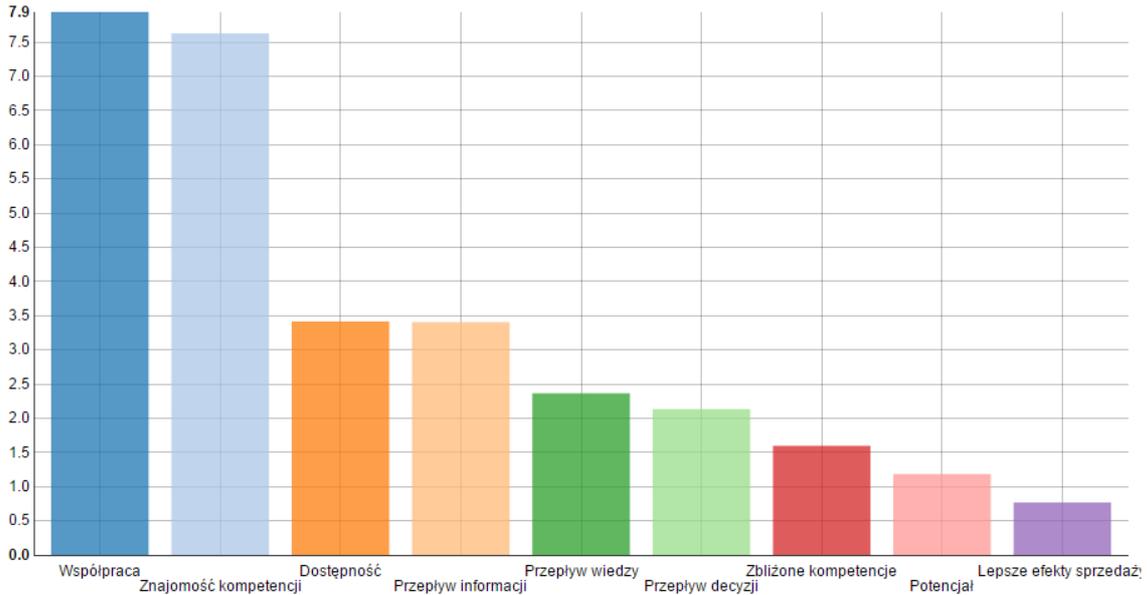
# Relacje i gęstość sieci

Porównaj: z poprzednim kwartałem

- z poprzednim dniem
- z poprzednim tygodniem
- z poprzednim miesiącem
- z poprzednim kwartałem
- z poprzednim półroczem
- z poprzednim rokiem
- z dowolną datą

Relacje Gęstość sieci Interpretacja

Wykres przedstawia typy relacji i częstość ich występowania w sieci. Pionowe słupki prezentują średnią liczbę powiązań na jednego użytkownika sieci.



Rys. 90 Możliwość śledzenia zmian występujących na sieci w czasie. Źródło <https://www.mapaorganizacji.pl/>

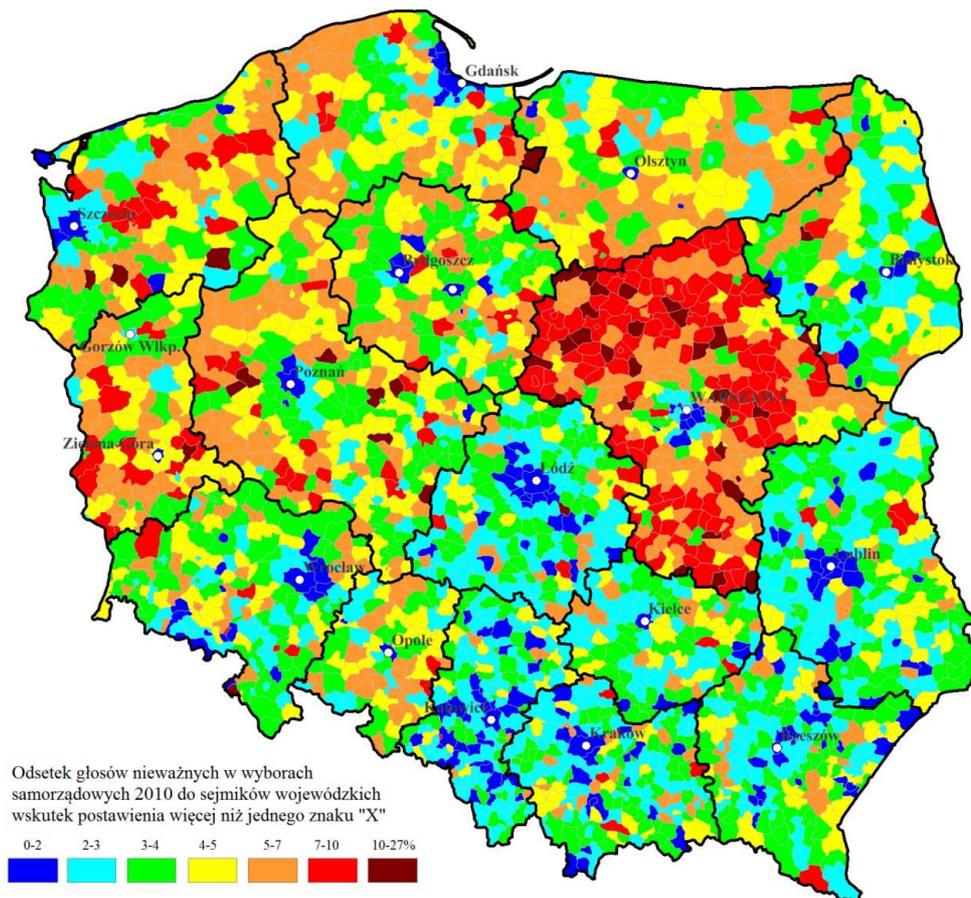
Opisane techniki są szybko adoptowane przez przedsiębiorstwa i organizacje w celu wizualizacji i analizy relacji interpersonalnych. Jednak wymaga to czasu pracowników – wypełniających kwestionariusze ręczne przetwarzania danych ankietowych, jak również konieczność analizy rezultatów, oraz umiejętności interpretacyjnych przez menadżerów średniego i wyższego stopnia, przez co jest to stosunkowo nowa branża na polskim rynku.

## 16 Dynamika wyborcza

Modelowanie wyborów ma bardzo długą tradycje matematyczną, ze względu na żywotny interes naukowców i ich zainteresowanie polityką. Już w rozdziale o formowaniu opinii zagadnienia dotyczące wyborów zastały dotknięte z teoretycznego punktu widzenia. Różnej maści naukowcy regularnie badają i sprzecają się, które systemy wyborcze jak proporcjonalny, czy jednoosobowe są lepsze z punktu widzenia funkcjonowania systemu. Kwestie dotyczące parytetów również były dobrze badane, czego przykładem jest „Kompromis Jagielloński”, którym liczba reprezentantów jest proporcjonalna do pierwiastka wielkości populacji (Życzkowski et al., 2014). Tym razem skoncentrujemy się na empirii w kontekście wykrywania anomalii oraz zmian na scenie politycznej.

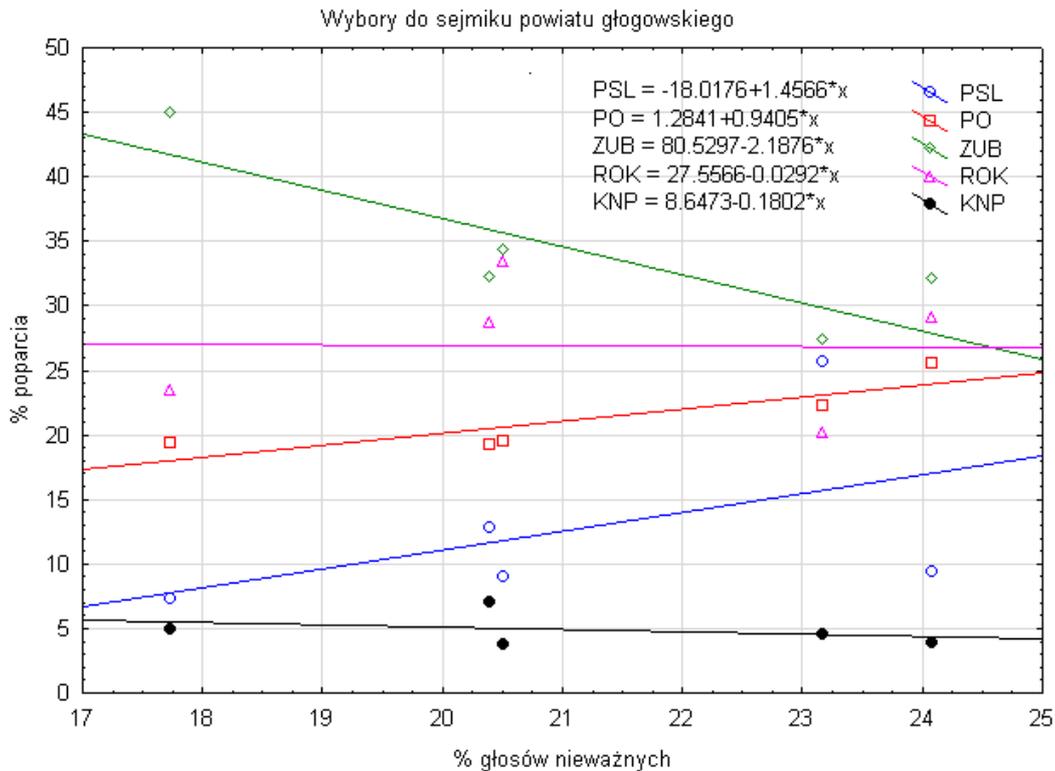
### 16.1 Czy można sfałszować wybory w Polsce?

Statystyczne anomalie wyborcze dają pojawiają się często tam, gdzie wynik wyborczy był poddany manipulacji. Statystyka daje narzędzia znajdowania manipulacji, co było między innymi tematem podrozdziału o korupcji. Przykładem jest łamanie prawa Benforda. Chodzi tu o nierównomierny rozkład prawdopodobieństw występowania pierwszej cyfry. Dla procesów naturalnych prawdopodobieństwo, że pierwszą cyfrą jest 1 jest największe, a całkowity rozkład jest dobrze znany. Dodatkowo ostatnia cyfra powinna być równomiernie reprezentowana, a w przypadku fałszerstw dominują okrągłe końcówki jak zero. Odstępstwa od tego rozkładu mogą pochodzić z ręcznej manipulacji wynikami. Łamanie prawa Benforda oraz dziwne związki między wynikami a frekwencją i liczbą głosów nieważnych są często obserwowane w nie do końca demokratycznych państwach trzeciego świata, ale również np. w Rosji.



Rys. 91 Odsetek głosów nieważnych w wyborach samorządowych 2010 do sejmików wojewódzkich wskutek postawienia więcej niż jednego znaku X. Źródło: <http://uczciwe-wybory.pl/>

Jednak, czy w Polsce takie anomalie wstępują? Przyjrzyjmy się wyborom samorządowym. Już w 2010 roku zauważono dziwną anomalie statystyczną na Mazowszu związaną z liczbą głosów nieważnych przez skreślenie więcej niż jednego kandydata. Jakkolwiek mogło to wynikać ze sposobu oddawania głosu za pomocą książeczki a nie karty wyborczej, jednak rozkład głosów był bardzo nierównomierny, po często w sąsiadujących komisjach oddano mniej niż 1% a w drugiej ponad 20% głosów nieważnych w ten sposób.



Rys. 92 Korelacje między wynikiem partii a % głosów nieważnych, pozytywne współczynniki dla partii rządowych PO- PSL oraz ujemne dla opozycji

W wyborach samorządowych 2014 problem głosów nieważnych ogarnął całą Polskę. Jedną z miar zafałszowania wyborów jest liczba głosów nieważnych. 16 listopada 2014 w wyborach do powiatów i sejmików wojewódzkich sięgnęła ona aż 20-40%, podczas gdy we Francji, Wielkiej Brytanii a nawet przedwojennej II RP wynosi ona w okolicach 1% i to mimo analfabetyzmu. Dla porównania: gdyby fizyk lub chemik otrzymał wynik z błędem rzędu 20-40% to powtórzyłby pomiar i określił przyczyny błędu. Państwowa Komisja Wyborcza, której sędziowie zarabiają po 20 tysięcy miesięcznie, aż dwukrotnie dokonała rażącego naruszenia prawa. Po pierwsze: złamała kodeks wyborczy, bo do sejmów i powiatów zamiast kart do głosowania dała wyborcom książki z listami komitetów wyborczych. Drugie złamanie prawa to błędne pouczenie wyborców (zawarte choćby w oficjalnych spotach Komisji Wyborczej), że mają stawiać po jednym krzyżyku na każdej z tych list. W efekcie ogromna część wyborców zgodnie z tym pouczeniem stawiała krzyżyki na każdej liście, czyniąc głos nieważny, co przy wysokiej jak na Polskę frekwencji bliskiej 50% sprawiło, że 3 miliony Polaków zostało w praktyce pozbawione swojego prawa wyborczego poprzez błędne pouczenie, a następnie unieważnienie głosów. Ale co tam, jak mawiał Józef Stalin: „nieważne kto głosuje, ważne kto liczy głosy” i w tym celu PKW zamówiła 3 miesiące przed wyborami niesprawny system informatyczny, pozwalający dopisywać głosy a nawet kandydatów także osobom spoza PKW logującym się do niego z zewnątrz. W ten sposób według systemu PKW prezydentem Szczecina został prezydent Szczecinka, a w Głogowie i Gdyni PSL zdobył 10 razy więcej głosów, niż w poprzednich wyborach. Dlatego przedstawiciele wszystkich partii politycznych opowiedzieli się za powtórzeniem wyborów – Paweł Kukiz (bezpartyjny), Jarosław Kaczyński (PiS), Leszek Miller (SLD). Eugeniusz Kłopotek (PSL), Cezary Grabarczyk (PO), Janusz Korwin-Mikke (Nowa Prawica), Andrzej Rozenek (Ruch Palikota) czy Krzysztof Bosak (Ruch Narodowy). Zwłaszcza, że nawet rodzina i znajomi wskazywali lokale, gdzie głosowali sami na siebie, a

dostawali 0 (słownie: zero!) głosów. Aby sprawdzić, czy w anomaliach wyborczych była jakaś reguła, policzyliśmy współczynnik korelacji między liczbą głosów nieważnych, a poparciem dla partii. Innymi słowy: kto skorzystał, a kto stracił na głosach nieważnych wyborów. Wyniki przedstawiamy na powyższym wykresie na przykładzie wyników do Rady Powiatu Głogowa. Okazuje się, że w Głogowie podobnie jak w całej Polsce było 25% głosów nieważnych, a na sfałszowaniu wyborów najbardziej skorzystało PSL i PO, a najwięcej potencjalnych głosów unieważniono komitetowi obecnego prezydenta Głogowa Jana Zubowskiego (koneksje z PiS), oraz Nowej Prawicy Janusza Korwina-Mikke. Co ciekawe, komitetowi Rafaela Rokaszewicza (koneksje z SLD) nie głosy nieważne nie zmieniły ani na plus ani na minus.

## 16.2 Polaryzacja partyjna po katastrofie smoleńskiej - hierarchiczna struktura partii politycznych w Polsce

Na podstawie zmiany dynamiki sondaży będziemy budować mapy semantyczne na podstawie stenogramów z obrad sejmiku badając zmianę w polaryzacji na linii PO-PiS przed i po katastrofie smoleńskiej. Poprzez m.in. szczegółową analizę tekstów będziemy badać diagramy hierarchiczne przed i po katastrofie. Przedstawiony rozdział będzie składał się ostatecznie z empirycznych badań; polaryzacji medialnej:

- sondaże
- transfery międzypartyjne
- wypowiedzi posłów PO-PiS

polaryzacji realnej:

- głosowania sejmowe

Zmiany w społeczeństwie demokratycznym mogą zajść na wiele sposobów a katastrofa smoleńska jest tylko incydentem w porównaniu z doświadczonymi w Polsce przemianami :

- Oddolne ruchy społeczne (ruch narodowy II RP, Solidarność, Stop-ACTA)
- Zdarzenia zewnętrzne: (II Wojna Światowa, Transformacja ustrojowa)

Hipoteza związana z katastrofą smoleńską: polaryzacja jedynie w sferze medialnej (bez zmian w sferze działań faktycznych). Będziemy budować mapy semantyczne na podstawie stenogramów z obrad sejmiku badając zmianę w polaryzacji na linii PO-PiS

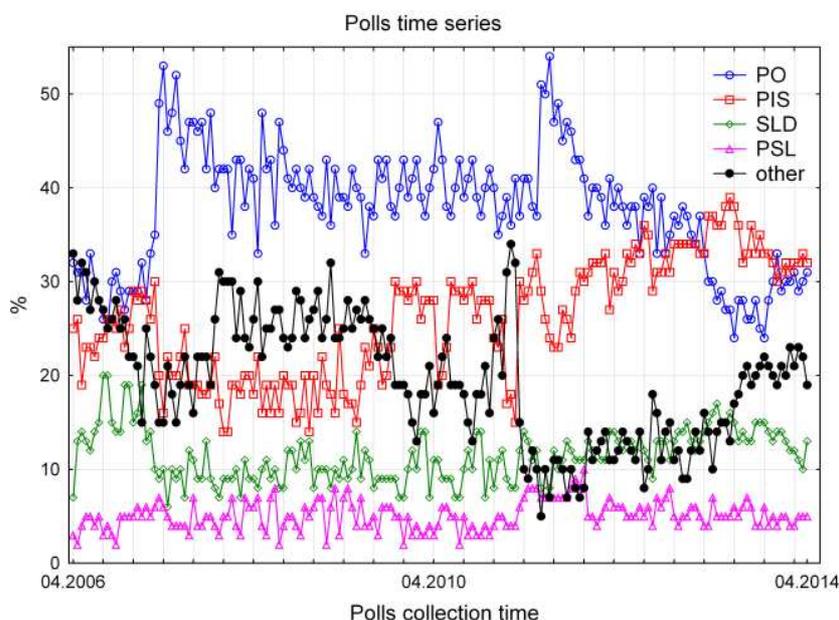
Hipoteza związana z II RP: zniszczenie elit przez okupantów pociągnęło za sobą kolaps polityczny – organizacyjny – badawczy PRL . Poprzez analizę usieciowienia przed i po WWII można zbadać stabilność multi-sieci przez analizę roli i utraty hubów (elit społecznych)

Jak pokazują dzieje Polski ostatnich 150 lat, polskie partie polityczne przeżywają zawirowania, więc w najnowszej historii uwzględniliśmy 7 najważniejszych sił politycznych.

Dla uproszczenia przyjęliśmy, że partie startujące razem w wyborach stanowią jedną siłę polityczną, stąd druga najstarsza partia - SLD zawiera w sobie także poparcie dla innych partii lewicowych - Unii Pracy, Partii Demokratycznej czy Socjaldemokracji RP. Trzecia najstarsza partia, liberalna Unia Polityki Realnej, jest rozpoznawalna w sondażach głównie za sprawą lidera – Janusza Korwina-Mikke, stąd w jej wyniku uwzględniłem Platformę Janusza Korwina-Mikke, Wolność i Praworządność, oraz Kongres Nowej Prawicy. Z kolei do wyników powstałej w 2001 roku Platformy Obywatelskiej doliczałem punkty procentowe dla Ruchu Poparcia Palikota i Twojego Ruchu, bowiem polityk ten do końca 2010 był formalnie posłem PO.

Analogicznie - do sondażowych wyników Prawa i Sprawiedliwości doliczałem poparcie dla partii utworzonych przez jej byłych członków – Polska Jest Najważniejsza, Polska Plus i Solidarna Polska, której politycy mają podobny, konserwatywny program i często przed wyborami wracali na łono macierzystej partii. Podobnie uwzględniliśmy, że konserwatywno-liberalna Liga Polskich Rodzin w 2009 roku startowała do wyborów do Parlamentu Europejskiego jako Libertas, w wyborach do Sejmu 2007 z list LPR startowali politycy Prawicy RP i UPR. Stąd sumowałem poparcie LPR i UPR, zwłaszcza że lider LPR Roman Giertych w 1993 roku startował do Sejmu właśnie z list UPR. Najbardziej stabilną i niepodzielną partią okazał się najstarszy PSL, ale także Samoobrona – a więc partie ludowe i rolnicze, posiadające głównie wiejski elektorat.

Zakres danych sondażowych – od kwietnia 2006 do kwietnia 2010 (około 90 następujących po sobie badań) i zbiór uzupełniający – od kwietnia 2010 do kwietnia 2014 (również około 90 sondaży) pozwolił na ustalenie korelacji Pearsona między partiami.



Rys. 93 Zestawienie wyników sondaży 4 głównych partii

O ile dla pomiarów wielkości fizycznych dane są zazwyczaj obarczone małym błędem, to sondaż jest obarczony błędem pomiaru rzędu 2-3 %, a więc praktycznie na poziomie poparcia dla niektórych mniejszych partii. Stąd do tak uzyskanej macierzy korelacji należy podchodzić ostrożnie.

	PO	PiS	SLD	PSL	Samoobrona	LPR-UPR
PO	1					
PiS	0,1	1				
SLD	-0,03	-0,27	1			
PSL	0,04	0,16	-0,11	1		
Samoobrona	0	0,04	0,05	-0,04	1	
LPR-UPR	-0,1	-0,1	-0,23	-0,09	-0,05	1

Tabela 7 Macierz korelacji między partiami politycznymi w Polsce na podstawie sondaży z lat 2006-2010

Dominują korelacje słabe i ujemne. Wynika to z faktu, że na scenie politycznej partie konkurują ze sobą o wiele bardziej, niż dobra na rynkach finansowych lub towarowych. Bowiem inwestor może kupować wiele rodzajów akcji, konsument – wiele płyt na raz, a wyborca – ma tylko jeden jedyny głos. Dlatego scena polityczna jako rynek partii jest w praktyce ograniczona do kilku największych sił.

Najsilniejsza korelacja zachodzi między PiS i PSL (wynosi 0,16). Zatem łatwo obliczyć, że odległość między ludowcami a PiS wynosi:  $d = 1,29$ .

Druga najmniejsza odległość:  $d = 1,31$  odpowiada korelacji między PiS a Samoobroną. W ten sposób metodą MST znalazłem 2 pierwsze strefy koordynacyjne. Kontynuując to postępowanie, znajduję trzecią strefę odpowiadającą korelacji PiS-PO ( $d = 1,34$ ). Kolejnej strefy  $d = 1,37$  odpowiadającej parze PO-Samoobrona nie uwzględniam, bo obie te partie są już połączone za pośrednictwem znalezionych wcześniej związków z PiS. W podobny sposób dołączamy do drzewa kolejne partie i otrzymujemy przedstawione drzewo minimalnej rozpiętości MST dla polskiej sceny politycznej. W przeciwieństwie do typowych danych naukowych, wyniki te są już na starcie obarczone błędem pomiaru, który w przypadku LPR i UPR jest równy poparciu dla tych partii. Dlatego w macierzach korelacji i odległości obie te partie występują razem jako LPR-UPR. Ma to uzasadnienie polityczne (wspólne starty w wyborach), ale także ilościowe. Bowiem nawet w hierarchii partii (obarczonej niepewnością pomiaru) obie partie ze sobą sąsiadują.

Wyniki opierają się wyłącznie na czystych wynikach sondaży opinii, ale bez pytania wyborcy np. o partie drugiego wyboru, czy analizy przepływu elektoratów – a więc tradycyjnych metod socjologów. Z drugiej strony wybrana przeze mnie analiza - drzewa minimalnej rozpiętości, hierarchiczność, przestrzenie ultrametryczne – opiera się na sprawdzonej metodzie, która odniosła wielki sukces na rynkach finansowych, a wcześniej okazała się przydatna do badania szkieł spinowych. Bowiem chciałem sprawdzić, czy partie polityczne też zachowują się kolektywnie wobec zewnętrznych oddziaływań. Jak więc moje wyniki pasują do już znanych z innych badań faktów?

Po pierwsze: drzewo minimalnej rozpiętości MST dobrze odzwierciedla lewicę (SLD, Samoobrona) i prawicę. Z tym, że pokazuje, iż mamy w Polsce dwie prawice. Gałąź na której w podobnych odległościach znalazły się PSL-PiS-PO można nazwać centroprawicą. Z kolei LPR-UPR jest skorelowana ujemnie z praktycznie wszystkimi pozostałymi partiami – zwłaszcza, że ich program jest o wiele bardziej konserwatywny i liberalny od pozostałych. Oznacza to, że polski wyborca właśnie w taki sposób wyobraża sobie scenę polityczną i postrzega różnice między partiami. Jak więc nasza scena wygląda na tle innych demokracji? Warto dodać, iż podobne zjawisko (dwie prawice) obserwujemy choćby w parlamencie europejskim, gdzie obok lewicowych socjalistów mamy centroprawicową chadecję i prawicowych konserwatystów i reformatorów. Także w Stanach Zjednoczonych, gdzie od lat na przemian rządzą Demokraci i Republikanie, występuje to zjawisko.

Warto zauważyć, że na scenie politycznej dominują korelacje ujemne i słabe Tabela 7. Oznacza to, że partie ze sobą konkurują, a wyborcy raczej nie zmieniają poglądów (raczej nie w skali czasowej pół miesiąca). Dlatego zdolność koalicyjna partii nie zawsze jest na drzewie minimalnej rozpiętości dobrze odzwierciedlona. Bo o ile połączone linkami LPR-Samoobrona-PiS tworzyły razem koalicję w latach 2005-2007, to kolejna koalicja PO-PSL nie jest bezpośrednio połączona, a planowany od 2011 roku związek PO-SLD plasuje się po przeciwnych stronach drzewa. Zwłaszcza, że obie partie różnią się programem zarówno gospodarczym jak i obyczajowym. Oznacza to, że koalicje na polskiej scenie rządzą się swoimi prawami i niekoniecznie pokrywają się z wolą wyborców. Gdyby zatem prawa znane z rynków finansowych przedłużyć na scenę polityczną, można by postulować, że - podobnie jak w koszyku akcji – koalicja partii ze skrajnie odmiennych części drzewa wydaje się najmniej ryzykowna. Tyle, że taki pomysł nie znalazł dotąd realizacji w praktyce.

Kolejną bardzo ciekawą sprawą jest interpretacja połączeń między partiami w drzewie minimalnej rozpiętości MST. Najbardziej zaskoczył nas fakt – nieprzewidywalny wcześniej – iż przedstawione na drzewie  $n-1$  połączeń (Rys. 94) w całości pokrywa się... z przejściami polityków między partiami! Aby podkreślić ten fakt, przygotowałem tabelę 22 najbardziej głośnych zmian barw politycznych w latach 2001-2010 (Tabela 8). Oczywiście następowały one głównie tuż przed wyborami, dlatego były badane również przez fizyków i uzasadniane dzięki zastosowaniu teorii przejść fazowych ferromagnetyków i

antyferromagnetyków. Z porównania drzewa minimalnej rozpiętości MST (Rys. 94 lewy) i przedstawionej Tabela 8 jednoznacznie widać, iż każde połączenie w drzewie znalazło swoją personalną realizację w praktyce. Bo nawet najbardziej nieprawdopodobne przejście byłego premiera Leszka Millera z SLD na listy wyborcze Samoobrony (i z powrotem) zdarzyło się naprawdę.

Warto dodać, że - zdaniem socjologów - przejścia między partiami są w oczach wyborców jednoznacznie negatywnie postrzegane i kosztują polityków utratę zaufania elektoratów obu konkurujących partii. Co ciekawe, również przed 2005 rokiem premier Miller rozważał koalicję z Samoobroną, a poseł Janusz Maksymiuk w 2001 roku zamiast SLD wybrał Samoobronę.

lp	Polityk	partia	rok
1	Roman Giertych	UPR → LPR	2001
2	Tomasz Tomczykiewicz	UPR → PO	2001
3	Janusz Maksymiuk	SLD → Samoobrona	2001
4	Bogdan Pęk	PSL → LPR	2002
5	Julia Pitera	UPR → PO	2004
6	Wojciech Mojsesowicz	Samoobrona → PiS	2005
7	Maciej Płażyński	PO → PiS	2005
8	Ryszard Bugaj	PSL → PiS	2005
9	Zbigniew Religa	PO → PiS	2006
10	Zyta Gilowska	PO → PiS	2006
11	Zygmunt Wrzodak	LPR → Samoobrona	2007
12	Ryszard Bender	LPR → PiS	2007
13	Antoni Mezydło	PiS → PO	2007
14	Radek Sikorski	PiS → PO	2007
15	Nelly Rokita	PO → PiS	2007
16	Leszek Miller	SLD → Samoobrona	2007
17	Antoni Mezydło	PiS → PO	2007
18	Ryszard Czarnecki	Samoobrona → PiS	2008
19	Leszek Miller	Samoobrona → SLD	2008
20	Janusz Wojciechowski	PSL → PiS	2009
21	Danuta Hubner	SLD → PO	2009
22	Tomasz Nałęcz	SLD → PO	2010

Tabela 8 Przejścia polityków między partiami w latach 2001-2010

Oczywiście scena polityczna tak jak rynki finansowe, rynek fonograficzny czy walutowy zmieniają się w czasie, dlatego drzewa minimalnej rozpiętości MST również podlegają ewolucji – co dokładnie opisałem w poprzednich rozdziałach. Dlatego Tabela 8 oprócz ścisłego odzwierciedlenia drzewa minimalnej rozpiętości MST zawiera także inne przejścia. Zwłaszcza przed 2005 rokiem, oraz ostatnich latach 2009-2010 (przejścia polityków lewicy na listy wyborcze PO i do kancelarii Prezydenta Bronisława Komorowskiego z Platformy Obywatelskiej). Tyle, że przejścia te znajdują się w skrajnych okresach (przed 2005 i po 2009) i mogą zwiastować kierunki ewolucji drzewa MST. Z tego powodu postanowiliśmy powtórnie przebadac serię czasową danych z sondaży, ale przedłużoną do listopada 2013, a więc do 170 pomiarów.

Okazało się, że w wydłużonym okresie zaczęły jeszcze bardziej dominować korelacje ujemne, a największa z nich między PO i PiS wyniosła aż -0,46, podczas gdy przez pierwsze lata istnienia obu partii była ona dodatnia i wynosiła 0,09. Zmiana ta wynika z polaryzacji sceny politycznej po katastrofie smoleńskiej 10 kwietnia 2010, w której oprócz Prezydenta RP z małżonką zginęli m.in. posłowie największych partii. Zjawisko to można porównać do zmiany struktury rynku fonograficznego wywołanej śmiercią artysty, bądź bankructwem spółek na rynkach finansowych. Na przykład niewyjaśniona śmierć Andrzeja Leppera w

sierpniu 2011 spowodowała chwilowy wzrost poparcia dla Samoobrony, a następnie zniknięcie jego partii ze sceny politycznej.

	PO	PiS	SLD	PSL	Samoobrona	LPR-UPR
PO	1					
PiS	-0,46	1				
SLD	-0,15	-0,16	1			
PSL	-0,32	0,05	0,09	1		
Samoobrona	0	-0,03	-0,11	-0,01	1	
LPR-UPR	-0,04	-0,04	-0,24	-0,09	-0,09	1

Tabela 9 Macierz korelacji między partiami politycznymi w Polsce na podstawie sondaży z lat 2006-2013

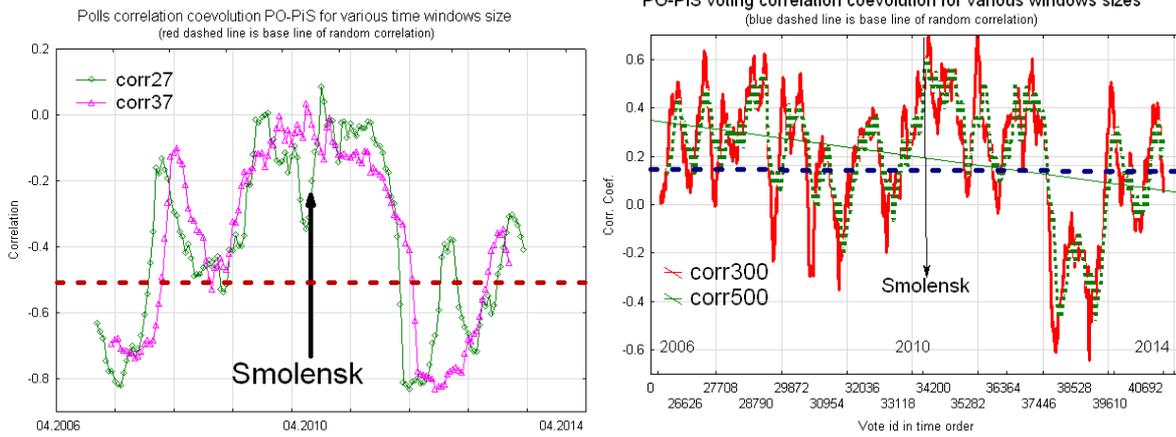
Co ciekawe, wydłużenie serii czasowej na okres po katastrofie smoleńskiej zasadniczo zmienił strukturę drzewa minimalnej rozpiętości, które dla serii danych z lat 2006-2013 zamiast jednej lewicy i dwóch prawic odzwierciedla już jednoznacznie scenę polityczną od strony lewej do prawej (Rys. 94). Oczywiście można nad tym dyskutować, zwłaszcza że wskutek coraz większej dominacji korelacji ujemnych odległości między partiami uległy zwiększeniu, jednak samo drzewo ma już postać rzędu, w którym po lewej stronie lewej mamy lewicę (SLD, Samoobronę, PSL), centrum (PiS) i prawicę (UPR/LPR, PO). Przypomina ono zatem coraz bardziej klasyczny podział na partie lewicowe – czyli socjalne, ale z wolnością obyczajową, oraz partie prawicowe – konserwatywne obyczajowo, ale wolnorynkowe. Na pytanie czy świadczy to o dojrzałości polskiej demokracji celowo unikamy odpowiedzi, bo zagadnienie to należy już do innej dziedziny.



Rys. 94. Drzewo minimalnej rozpiętości (MST) zbudowane dla partii politycznych w Polsce na podstawie sondaży przeprowadzonych przez Gemius a następnie Homo Homini w latach 2006-2013, przed (lewe) i po (prawe) katastrofie smoleńskiej

W analizie nie braliśmy pod uwagę “partii drugiego wyboru” i przepływu elektoratu, bo ośrodki sondażowe jak częściowo państwowy OBOP czy niezależne Homo Homini i Polska Grupa Badawcza wykonują takie badania zbyt rzadko, by uzyskać z nich prawidłową statystykę przydatną do liczenia np. współczynników korelacji. Ośrodki te przy okazji wyborów z 2007 roku podały, że o ile PO przejmuje część elektoratu PiS, to partia Jarosława Kaczyńskiego zachowuje poparcie dzięki skonsumowaniu Samoobrony i LPR – co pokrywa się z drzewem MST. Zwłaszcza, że w 2005 roku to właśnie Samoobrona zajęła 3 miejsce w wyborach, przejmując pokaźny elektorat SLD. Słabością naszych badań są też statystyczne niepewności pomiarów w sondażach rzędu 2-3%. W efekcie ograniczają one badaną scenę polityczną, bo w ostatnich 100 latach zawsze tylko kilka partii wznosiło się ponad próg błędu statystycznego. Ponadto drzewa minimalnej rozpiętości nie odzwierciedlają tak istotnych zjawisk jak silna ujemna korelacja między PO i PiS o wartości -0,46, która w przypadku partii ma znaczenie, zwłaszcza że pozostałe korelacje są słabe, a zbiór danych, czyli punkty procentowe po zsumowaniu nie przekraczają wartości 100%, a więc ich wkład we współczynniki korelacji jest ograniczony, co skutkuje na ogół korelacjami słabymi.

Patrząc na zmianę korelacji między sondażami głównych partii PO-PiS, a zmianą korelacji ich głosowań można zauważyć podobieństwa w kształcie krzywych. Linią kreskowaną zaznaczony jest poziom odniesienia wynikający z właściwości procesów przypadkowych – bazowych.



Rys. 95 Korelacje między sondażami (lewy) korelacje między wynikami głosowań (prawy) dla różnych okien czasowych głównych partii PO-PiS, 4 lata przed i po Smoleńsku

Warto zatem kontynuować podobne badania w kierunku większej dokładności i innych parametrów socjologicznych, bo otrzymane wyniki potwierdziły, że metody znane z fizyki i rynków finansowych sprawdzają się także na rynku konkurujących ze sobą partii politycznych, jednak przez specyfikę danych ich skuteczność jest jednak ograniczona. W ramach tematyki z następnego rozdziału o języki będziemy szukać metody wydzielenia różnic w dyskursie sejmowym, np. poprzez słownictwa znaczącego. Już w wstępna analiza na pełnych listach frekwencyjnych wydaje się interesująca już z porównania list rankingowych słów dla PO i PiS z korpusami tła - ogólnymi (nazwane random, bo losowane).

## 17 Techniki wykorzystania języka naturalnego

Mimo że komputerowe narzędzia językowe od właściwie początków algorytmiki istnieją w świecie informatyki i lingwistyki, to ich zastosowanie w naukach społecznych jest wciąż znikome. W związku z rozwojem metod komputerowych w przetwarzaniu informacji pojawiło się pole do wykorzystania automatów ukazujących korelację między różnymi jednostkami analizy. Umożliwiły to techniki przetwarzania języka naturalnego (ang. Natural language processing - NLP) lub coraz częściej zastępują ją nazwa: inżynieria języka naturalnego, które pozwoliły analizować treści ustrukturyzowane składniowo i gramatycznie w ramach określonego kodu językowego. Metody i techniki NLP<sup>9</sup> zostały opracowane przez informatyków w celu ułatwienia przetwarzania informacji zapisanych w sposób zrozumiały dla człowieka, a nie dla systemów analitycznych. Dopiero przetransformowanie takich informacji pozwala analitykom na zautomatyzowanie swojej pracy. Początkowo techniki NLP służyły dokonywaniu analizy niestrukturyzowanych danych w postaci zwykłego tekstu. Na przykład bez konieczności czytania tekstów dokonywano oceny nacechowania emocjonalnego wypowiedzi na (podstawie) doboru słów oraz ich struktury. Datamining w przestrzeni języka naturalnego na dobre już zagościł jednak w marketingu i biznesie (Lazer, et al., 2009). Częściowo o NLP już wspomniano przy ekstrakcji sieci społecznych z literatury (tam autor wykorzystywał metody automatyczne do wyławiania interakcji między bohaterami w tekście). Dostępne są programy do analizy znaczenia i informacji wyrażonej w tekście. Np. automatyczny monitoring mediów na temat wydarzeń na podstawie słów kluczowych istotnych dla ekonomii jest już chlebem powszednim inwestorów skoncentrowanych na pewnej konkretnej branży, czy regionie. Bardzo ważnym wskaźnikiem zadowolenia klientów z produktu jest sentyment, czyli ocena nastawienia emocjonalnego i rozpoznanie opinii, również jako narzędzie automatyczne. Niestety barierami do wykorzystania tych narzędzi w socjologii są wymagane umiejętności programistyczne oraz wymagana wiedza z zakresu inżynierii języka naturalnego, stąd ten rozdział. Część zagadnień dotyczących jakościowej analizy zwłaszcza dużych zbiorów danych daje się w sposób znaczący zautomatyzować. Badania nad automatyczną klasyfikowaniem tekstów opierają się na założeniu, że istnieją uniwersalne wzorce komunikowania poszczególnych treści. Mody, trendy społeczne zgodnie z teorią Simmela można są utożsamiać z wytworami kultury. Takim wskaźnikiem jest częstotliwość występowania wyznaczników danej mody w zasobach np. książkowych danego okresu. Możliwe jest też obserwowanie określonej frazy w zapytaniach, jak np. robi to Google Trends. Techniki opisane powyżej znalazły wiele komercyjnych zastosowań i są coraz silniej rozwijane w celu jak najdokładniejszej klasyfikacji tekstów, jakkolwiek jednoznaczna klasyfikacja jest utrudniona, ponieważ najczęściej występują zmiana tła (tematyki) i współwystępowanie różnych znaczeń pojęciowych. Nie jest to jednak ich jedyne zastosowanie, gdyż badania ilościowe w literaturze zagościły już na dobre w literaturoznawstwie czy językoznawstwie.

---

<sup>9</sup> Należy uważać na możliwy konflikt nazewniczy z projekcjami neurolingwistycznymi, które używają tego samego akronimu –NLP



Rys. 96 Diagram 50 najczęściej stosowanych słów w tej książce wygenerowany przez <http://tagcrowd.com/>

Prawo Zipfa jest dobrze znaną empiryczną prawidłowością dotyczącą statystyki języka. Jest to związek częstości występowania słów i ich rang. Ranga to miejsce w rankingu najczęściej występujących słów. Mamy tu do czynienia z rozkładem tłusto-ogonowym. Dla języka angielskiego, gdzie przeciętny użytkownik wykorzystuje do 10000 słów, a słowniki zawierają ich około miliona, 10 najczęściej występujących słów wypełnia 15% przeciętnego tekstu (Kulig, Kwapien, 2012). Sama częstość występowania słów to mało, więc przydałby się model, który równocześnie uwzględniałby dwie główne składowe języka: statystykę poszczególnych słów oraz relacje pomiędzy nimi. Takie sieci lingwistyczne mogą mieć charakter sieci semantycznej (znaczeniowej), bądź współwystępowania (generatywnej).

Tradycyjne socjologiczne ujęcie przetwarzania tekstu jest jedną z podstawowych technik badawczych. Począwszy od analiz pamiętników polskich imigrantów w USA Znanieckiego z początku XX wieku, a skończywszy na współczesnych analizach blogów, jest to podstawowe narzędzie jakościowej analizy. Zważywszy na to, że np. analizę transkrypcji wywiadów socjologicznych również można poddać automatyzacji, okazuje się, że takie wspomaganie pracy socjologa jest istotne z praktycznego punktu widzenia. Powstało wiele teorii wykorzystujących takie techniki jak interakcjonizm symboliczny, etnografia, analiza konwersacyjna i krytyczna analiza dyskursu. Bardzo popularnym obszarem badań socjologicznych jest zrozumienie działania społecznego (uwarunkowania i konsekwencje) wykorzystując schematy komunikacyjne. W najprostszym ujęciu w każdej komunikacji możemy wskazać nadawcę i odbiorcę, a sam przekaz odbywa się poprzez określony kanał za pomocą ustalonego kodu językowego. Identyfikuje się przekaz w teorii informacji: nadawca – komunikat – odbiorca. Oznacza to istnienie kierunku w kanale komunikacji, choć czasami trudne wydaje się ustalenie, kto jest kim w tej konstrukcji, przy pewnym rygorze jest to możliwe (mniej lub bardziej obiektywnie). Należy mieć na uwadze, że komunikat nie jest tylko wypowiedzią ustną bądź pisemną. Rozróżniamy również komunikaty niewerbalne, a także różne fizyczne i nie-fizyczne formy interakcji. Założyliśmy, że nadawca to osoba, która ma wpływ na odbiorcę. Przykładowe zdarzenie: jeśli myślę teraz o mojej dziewczynie i jej słowach z poprzedniego dnia, to ma ona wpływ na mnie w tym momencie. Należy rozumieć „myślenie” jako komunikację, w procesie której odbieram informacje, uprzednio wysłane przez moją dziewczynę. Konsekwencją będą tu interakcje bezpośrednie (*explicite*) i pośrednie (*implicite*). Tak więc ja, myśląc o mojej dziewczynie, będę odbierał wiadomość *implicite*, w przeciwieństwie do komunikatu *explicite*, wysłanego przez nią wczoraj. W komunikacji mamy do czynienia z redundancją czyli to, co jest się powtarza lub jest przewidywalne lub konwencjonalne w przekazie. Pomaga ona w dekodowaniu ze względu na wysokie prawdopodobieństwa występowania istotnych informacji, co wiąże się z entropią czyli stopniem uporządkowania i przewidywalność informacji. Zaobserwowanie takich frekwencyjnych zależności jest głównym motorem metod komputerowych. Ważne

jest sprzężenie zwrotne, czyli reakcji odbiorcy powracająca do nadawcy, o czy wiem wiedzą wszyscy specjaliści od marketingu. Analiza związków przyczynowo-skutkowych pozwala określić psychosocjologiczne aspekty komunikatu (dlaczego? Po co?). Wyznacza się przepustowość kanału, gdzie poza informacją znajdują się zakłócenia. Tu również metody znane z nauk technicznych mają zastosowania. Specjaliści od PR wiedzą jaka mowa publiczna będzie najefektywniejsza ze względu na oddziaływanie (funkcje retoryczne). Sam kontekst komunikatu wiąże się z regulującymi go instytucjonalnymi procedurami. Komunikowanie jest wymiarem odtwarzania rzeczywistości społecznej, a język kształtuje percepcję tej rzeczywistości. W tekstach literackich jest odwrotnie, bo świat fikcyjny jest rzutowany czytelnikowi za pomocą komunikatów. Tego typu badania zaliczają się do cyfrowej humanistyki (ang. Digital Humanities). Każdy z zespołów podejmujących tą tematykę koncentruje się na innym aspekcie badań, niemniej jednak ich wspólnym mianownikiem pozostaje problem automatycznej analizy tekstu. Pierwsze prace dotyczyły analizy utworów dramatycznych, gdzie połączeniem między bohaterami (postaciami) było ich występowanie w tej samej scenie. Franco Moretti ze Stanford, koordynujący działania, pokazał właściwości sieci komunikacyjne na przykładzie dzieł Szekspira, następnie porównał je z dramataми chińskimi. Jedną z najważniejszych poczynionych obserwacji była znacząca różnica w ilości połączeń społecznych (liczba znajomych) centralnych postaci z dzieł europejskich, wokół których koncentruje się cała sieć w porównaniu z chińskimi, gdzie postacie miały bardziej równomiernie rozłożoną liczbę znajomych w społeczności. Naukowcy z Harvardu, wśród nich Martin Nowak, szukali asocjacji pomiędzy różnymi hasłami pojawiającymi się w zbiorze książek zdigitalizowanych przez Google. Na Uniwersytecie Columbia wyekstrahowano sieci kontaktów konwersatoryjnych bohaterów dziewiętnastowiecznych powieści wiktoriańskich. W tym celu skonstruowano algorytm oparty na liczeniu maszynowym, który na początku ustala listę bohaterów. Celem takich badań wzorcowych było zrozumienie świata kreowanego w tekście przy wykorzystaniu metod analizy rzeczywistych społeczności.

Algorytmy przetwarzania języka naturalnego, czyli programy komputerowe do przetwarzania tekstu i mowy na różnych poziomach analizy języka pozwalają na automatyczną analizę struktur językowych, np. analiza składniowa czy na zastosowania użytkowe, np. rozpoznawanie i klasyfikacja. Dzięki zgromadzonym zasobom językowym z korpusami (duże zbiory) dokumentów tekstowych i nagrań mowy, przykłady użycia (fragmenty, wypowiedzi lub całe dokumenty) i po anotowaniu - opisanie pod względem lingwistycznym w sformalizowany sposób (np. pod względem gramatycznym, czy też znaczenia) można wykorzystać komputerowe narzędzia NLP do analizy. Dostępne są już systemy do analizy struktury semantycznej polskich tekstów na przykładowych korpusach jak dialogi telefoniczne rejestrowane w *call center*, baza danych polskich tekstów prasowych itp. Bardzo przydatne okazały się algorytmy do analizy tekstów potencjalnych samobójców. Dotychczasowa praca psychologa policyjnego, którego zadaniem były ustalenie na ile prawdopodobna jest realna próba samobójcza autora tekstu (np. listu pożegnanego) została usprawniona metodami komputerowymi. Teksty można opisać metadanymi socjolingwistycznymi w zależności o potrzeby badawczej. Systemy obsługi klienta są w stanie określić na podstawie właśnie takich socjolingwistycznych cech i sentymentu (nacechowania emocjonalnego) priorytet rozpatrywania zgłoszeń (czy konsument naprawdę potrzebuje pilnego rozwiązania swojej sprawy, czy może poczekać). W praktyce bardzo dużym utrudnieniem jest subiektywność i różnice indywidualne zarówno w ekspresji jak i percepcji komunikatu. Na etapie etykietowania tekstów koder nie zapewnia pełnej obiektywności klasyfikacji. W związku z tym poziom odbioru i oceny emocji zależy od indywidualnej cech za pomocą metody ludzkiej i komputerowej często obarczone są podobnym błędem, po zostało zilustrowane na przykładzie ekstrakcji interakcji w podrozdziale o sieciach w literaturze, wrażliwości, empatii i inteligencji emocjonalnej.

Przedmiotem dotychczasowych badań był w większości nie język naturalny, a dokładniej jego zapisana forma – teksty literackie. Wybór takiej reprezentacji języka był zdeterminowany przez, że jest to jego forma „zmaterializowania” w formie próbki, mogącej być przedmiotem ilościowych analiz. Analiza języka używanego na co dzień, wymagałaby znacznie większego nakładu pracy związanej np. z

przetransformowania sygnału akustycznego (mowa) w sygnał, podatny ilościowej analizie (pismo). Stąd analizuje się transkrypcje języka mówionego i dołącza się je do korpusów w celu lepszego odwzorowania rzeczywistego obrazu języka współczesnego. Mimo że sama materia języka posiada specyficzny charakter, to znalezienie odpowiedniej reprezentacji języka i przetransformowanie jego próbek na sieciowe – obiektywne struktury daje szerszy wgląd w taki układ. Analizy statystyczne jednoznacznie wskazały na uniwersalizm wielu cech częstotliwościowych różnych narodowych języków oraz typów zapisu. Zipfowski rozkład występowania słów w danym korpusie języka, świadczy nie tylko o oparciu struktury języków europejskich o ten sam mechanizm generatywny, ale również wskazuje na uniwersalny charakter typowy dla układów złożonych (Kulig, Kwapien, 2012). Słownictwo charakterystyczne dla epoki, w której zostały stworzone, indywidualnego stylu autora oraz poruszanej tematyki pozostaje w jedności z niezmienną strukturą statystyczną. Opis języka, uwzględniający nie tylko częstość występowania różnych słów, ale również związek jaki istnieje pomiędzy nimi, był możliwy dzięki zastosowaniu analizy sieciowej. Modelowanie może być związana z sąsiedztwem słów w tekście, ich powiązaniem semantycznym, czy znaczeniowym. Takim przykładem zastosowania jest leksykalna sieć semantyczna dla języka polskiego, opisuje znaczenia leksykalne za pomocą relacji semantycznych. Jest nią Słowosieć składająca się z ponad 200 tys. lematów (haseł) – jeden z największych słowników języka polskiego w historii. Poza samą funkcją słownikową Słowosieć jest wzbogacona o zasoby leksykalne relacji znaczeniowych. Możliwa jest dzięki temu płytko, ograniczona analiza składniowo- semantycznej tekstu. Dla socjologów istotne mogą być jednostki wielowyrazowe, czyli wyrażenia języka składające się z 2 lub więcej wyrazów. Dzięki nim możliwa jest budowa sieci semantycznych czy znaczeniowych wokół interesujących badacza pojęć (Van Dijk, 1985), ich ewolucja w czasie oraz zróżnicowanie w obrębie kohort. Zorientowany dziedzinowo tokenizator (podział na podstawowe, niepodzielne części) umożliwia podstawową analizę lingwistyczną tekstów dziedzinowych. W obrębie tych tekstów próbuje się wykryć i sklasyfikować:

<i>obiekty</i> (osoby, organizacje, miejsca, przedmioty, itp.):	Może mieć to znaczenie w celu wyszukiwania częstotliwości wytypowania agentów społecznych (np. jak często mówi się o pewnych politykach) lub zdarzeń (np. jak często pojawiają się odniesienia do Katastrofy Smoleńskiej)
<i>związki</i> (np. tożsamość nazw, przynależność osoby do organizacji, autorstwo, położenie przestrzenne obiektów):	Jakie relacje społeczne występują między obiektami naszej analizy, np. czy można przyporządkować blogerów do poszczególnych partii?
<i>sytuacje</i> (czyli „kto co komu zrobił, gdzie i kiedy?”)	Można się dowiedzieć co myślą blogerzy o np. ruchu smoleńskim

Tabela 10 Analiza semantyczna tekstów z socjologicznej perspektywy

Pogłębia taka klasyfikacja analizę semantyczną tekstu i jest krokiem w kierunku rozumienia tekstu. Możliwa jest analiza całych kolekcji dokumentów ze względu na podobieństwo, bądź zróżnicowanie zawartości, co może mieć zastosowanie w weryfikacji hipotez społecznych o różnicach między wytworami tekstowymi badanych podmiotów (np. poglądy partii politycznych Tabela 11).

**Funkcja podobieństwa znaczeniowego jednostek leksykalnych**

**Hiperonimy**

**działalność 1**   
zorganizowany zespół aktywności podjęty...

**Meronymy (części)**

**dyplomacja 2**   
domena: związek między ludźmi, rzeczami i...

**Synonimy międzyjęzykowe**

**policy 2**   
domena: związane z porozumiewaniem się

**Potencjalne odpowiedniki międzyjęzykowe**

**policy 2**   
domena: związane z porozumiewaniem się

Wynik:  [ Szukaj ]

**polityka**

podobieństwo	jednostka leksykalna
0.167	sekcja
0.155	działanie
0.137	tematyka
0.120	pragmat
0.117	charakter
0.113	konkrecja
0.113	proces
0.110	postępowanie
0.105	rozwiązanie
0.103	podjęcie
0.105	postępu
0.103	system
0.103	między
0.099	kierunek
0.099	medycyna
0.096	działalność
0.095	dotyczy
0.094	przebiegający
0.093	nauczyciel
0.090	metoda

Rys. 97 Możliwości wykorzystania Słownosieci w praktyce. Źródło: <http://plwordnet.pwr.wroc.pl/wordnet/>

Korpus PiS pre	Korpus PO pre	Korpus PiS post	Korpus PO post
subst:działanie;0.147956	subst:działanie;0.15547	subst:działanie;0.1301	subst:obrad;0.0692369
subst:działalność;0.14090	6	11	ger:uprawiać;0.0492651
3	subst:działalność;0.1407	adj:ten;0.1299	imps:sprawdzić;0.046780
adj:ten;0.138978	04	adj:który;0.126568	1
subst:proces;0.133851	adj:ten;0.138319	subst:osoba;0.123256	ger:posiadać;0.0250339
adj:który;0.131678	subst:proces;0.133304	subst:rok;0.122144	fin:zakreślać;0.0202764
subst:pracownik;0.128008	adj:który;0.127588	subst:działalność;0.119	imps:zdefiniować;0.0198
ppas:zapisać;0.120989	subst:rok;0.123059	833	915
praet:przyjść;0.120838	subst:dyskusja;0.1208	subst:pracownik;0.119	subst:działalność;0.0189
subst:życie;0.120237	subst:ustawa;0.118847	553	619
subst:rok;0.119951	subst:pracownik;0.1181	subst:środek;0.112253	ger:wejść;0.018882
subst:osoba;0.117869	54	subst:życie;0.111894	subst:osoba;0.0186368
subst:ustawa;0.117594	subst:debata;0.117411	subst:zatrudnienie;0.1	subst:nieruchomość;0.01
subst:postępowanie;0.117	subst:posiedzenie;0.116	09862	83257
21	678	subst:to;0.106784	subst:tendencja;0.01812
subst:projekt;0.116233	subst:projekt;0.116402	subst:sprawa;0.105284	7
subst:komisja;0.11606	subst:realizacja;0.11600	subst:ustawa;0.105094	subst:przypadek;0.01714
subst:nieprawidłowość;0.1	6	subst:rząd;0.104112	06
14889	subst:rząd;0.11453	subst:projekt;0.104015	subst:rynek;0.0167068
praet:ratyfikować;0.11362	subst:komisja;0.113419	adj:finansowy;0.10377	adj:globalny;0.0166164
2	subst:to;0.112473	subst:czas;0.102698	subst:ograniczenie;0.015
subst:to;0.113338	subst:sprawa;0.111636	subst:komisja;0.10213	2412
adj:publiczny;0.111117	subst:postępowanie;0.1	3	subst:aparatus;0.0151571
subst:szużba;0.111029	10735	subst:państwo;0.1016	subst:metoda;0.0150217
	subst:zdrowie;0.110692	01	adj:taki;0.0148862
	subst:zadanie;0.108298	fin:chodźć;0.101532	subst:możliwość;0.01480
			77
			subst:rok;0.0144897

Tabela 11 Przykładowa analiza relacji znaczeniowych dla kategorii: praca w korpusach sejmowych

## CZĘŚĆ IV

# DYSKUSJA NAD SKUTECZNOŚCIĄ I ROZWOJEM METOD OBLICZENIOWYCH

## 18 Krytyka metod obliczeniowych

Socjofizyka została już uznana za naukę (w ramach fizyki), o czym świadczy dołączenie jej do oficjalnie zatwierdzonych kategorii poddyscyplin w wielu instytucjach naukowych. Jednak jako dziedzina nauk ukształtowała się dopiero w XXI wieku, więc jak każda młoda dyscyplina socjofizyka ma swoje słabe strony. Dopiero pojawienie się Internetu pozwoliło na w miarę obiektywny pomiar wielu zmiennych społecznych na masową skalę. Jak wspomniano na wstępie, socjofizyka różni się do innych modeli matematycznych próbą odnalezienia analogii fizycznych i wykorzystaniu metod wypracowanych w ten sposób. Ciężko mówić w tym momencie o sprawdzalności modeli socjofizycznych, bo nie udało się do tej pory w sposób spektakularny przewidzieć żadnego fenomenu społecznego dzięki socjofizyce. Jednakże to, co jest największą chyba porażką polskich socjologów (nieumiejętność przewidzenia znaczenia ruchu "Solidarności"), można bardzo łatwo wyjaśnić modelami socjofizyki (przynajmniej *post factum*). Książka została napisana w optymistycznym nastroju (gdyż autor jest zwolennikiem tej metodologii), lecz należy przestrzec przed ograniczeniami i niebezpieczeństwami. Jak każda metodologia, modelowanie ma swoje słabe strony. Wiele modeli jest słabych od samego początku, bądź po prostu błędnych, co niestety bardzo ciężko zweryfikować. Np. do tej pory trwa dyskusja czy czarna śmierć niszcząca średniowieczną Europę to choroba wirusowa (dżuma krwiotoczna), czy bakteryjna (dżuma dymienicza) i mimo zupełnie innych mechanizmów opisujących to zjawisko, ciągle oba typu modelowania pojawiają się równolegle. Warunkiem stosowalności jest znajomość parametrów społecznych. Czułość na zmianę parametrów i inne efekty chaotyczne, czy numeryczne również może mieć wpływ na wyniki, zwłaszcza że modele często mają setki parametrów. Tak samo modelowanie populacji we współczesnym świecie jest również problematyczne (zmiennosc społeczna, czy mobilność). Na zjawiska społeczne ma wpływ wiele niekontrolowanych czynników jak pogodowe, których zazwyczaj nie da się ująć w modelu nie komplikując go diametralnie. Jednakże największym problemem w Polsce jest brak specjalistów i centrów badawczych w tej dziedzinie. Do budowy i analizy tak skomplikowanych modeli potrzebne są umiejętności matematyczne, programistyczne oraz wiedza socjologiczna (rzadka kombinacja). W związku z tym matematyczne modelowanie i obliczenia symulacyjne nie są jeszcze stosowane w pracy i szkoleniu specjalistów do spraw życia publicznego, czy w podejmowaniu decyzji rządowych. Dlatego warto też przedstawić krytyczne ujęcie socjofizyki w dużej mierze w duchu Ossowskiego (Ossowski, 2005).

### 18.1 Brak zdolności predykcyjnej

Główny zarzut odnosi się do braku zdolności predykcyjnej ze względu na zbyt dużą liczbę zmiennych dotyczących ludzkiej działalności. Techniki układów złożonych, które stanowią podstawę socjofizyki, działają dobrze (w sensie czytelności i zrozumiałości modelu) dla niewielkiej liczby parametrów. Cała idea polega więc na wyborze głównych czynników wyjaśniających mechanizm zachodzący w społeczeństwie. Dlatego reszta zmiennych (zwłaszcza tych niekontrolowanych) jest zaniebywana, bądź traktowana jako szum. Z punktu widzenia predykcyjności ma to niestety ogromne znaczenie (negatywne), ale co do zrozumienia zjawiska już niekoniecznie.

Rozwiązywanie modeli metodami numerycznymi, bądź symulacyjnymi może prowadzić do zakłamania wyniku a nawet do zupełnie błędnych rozwiązań. Weźmy np. „efekt motyla” czyli - chaos deterministyczny

(niewielkie zmiany warunków początkowych prowadzą do znacznych zmian efektu końcowego). Nawet posiadając stosunkowo precyzyjne dane badany układ może być zbyt czuły i otrzymane rozwiązanie może być dalekie od rzeczywistości. Tego przykładem były równania Lorentza mające opisać układy meteorologiczne, afektem stał się motyl Lorenza (Białynicki-Birula, Białynicka-Birula, 2002).

Zdolność do przewidywania przyszłości jest obecnie głównym kryterium przyznawania grantów badawczych i jest podstawą finansowania nauki. Zdolność do przewidzenia efektów społecznego eksperymentu będzie zawsze ograniczona a co dopiero całego rzeczywistego społeczeństwa. Przyszłość tak złożonego układu jak np. miasto czy kraj zależy nie tylko od opisujących go zmiennych, tak obserwowalnych jak nieobserwowalnych, ale także od zmiennych opisujących otoczenie. Zaburzenia zewnętrzne w modelowaniu można go zaniedbać, w innych zredukować w teoretycznych problemach. W praktyce raczej się to nie udaje.

## 18.2 Problem pomiaru, przetwarzanie danych

Kolejną sprawą budzącą kontrowersje jest pomiar, gdyż modele działają przy wykorzystaniu danych empirycznych. Brak powtarzalności warunków w układach społecznych jak i wpływ badacza na ten układ (za Ossowskim) są niestety rzadko brane pod uwagę przy modelowaniu. Tu ścierają się dwa paradygmaty nauk ścisłych o intersubiektywnej sprawdzalności (kantyzm) i nauk społecznych z współczynnikiem humanistycznym (w sensie Znanieckiego). Klasyccy fizycy nie będą uznawać jako dane takich badań jakościowych jak wywiady, czy kwestionariusze otwarte (chyba, że przy użyciu metod dataminingowych), a socjologowie wykształceni we wspomnianej kulturze nie będą korzystać z metod matematycznych innych niż opracowania statystyczne. Wystarczy spojrzeć na konflikt między socjologią analityczną, a socjologią humanistyczną.

Zgromadzone dane, często personalne, trafiają do analizy mimo iż osoby, których działanie zostało zarejestrowane nie mają o tym pojęcia. Często, gdyby je zapytano, nie wyraziłyby na to zgody. W związku z tym, że autor ma dostęp do wrażliwych danych, wiedziałby jak je wykorzystać w sposób nieetyczny. Jednak nikt tego (w otoczeniu autora) nie robi. Mimo wszystko zagrożenie istnieje i warto o nim pamiętać.

## 18.3 Nauki ścisłe a socjologia: kontrowersje

Wiele naukowych modeli matematycznych jest z miejsca odrzucanych przez przedstawicieli nauk społecznych ze względu na niechęć do „pozytywizmu”. Takim mianem określa się obecnie w środowisku socjologicznym podpieranie się modelami i poleganie na empirii w badaniach.

Badania interdyscyplinarne są bardziej niż inne narażone na odejście od rygorów naukowych, dlatego często czasopisma odrzucają prace z zakresu społecznych modeli obliczeniowych. Można publikować jednak artykuły w czasopismach branżowych, ale wtedy model publikowany w czasopiśmie fizycznym/matematycznym/informatycznym nie będzie już modelem społecznym, ponieważ mechanizmy społeczną zostaną sprowadzone jedynie do funkcjonalnych procedur, a badanie modelu będzie dotyczyło jedynie jego cech znanych przez rozumianych przez recenzentów. Alan Sokal przedstawił prowokacyjny artykuł, używając naukowego żargonu, dotyczący związków pomiędzy koncepcjami rozwoju społecznego i teoriami fizyki współczesnej: Transgresja granic: ku transformatywnej hermeneutyce kwantowej grawitacji. Został on opublikowany w prestiżowym czasopiśmie socjologicznym pomimo tego, że praca polegała na doborze przypadkowym układów słów z korpusów tekstów postmodernistycznych, typu:

„Podobnie jak liberalne feministki często zadowolają się minimalnym zakresem prawnej i społecznej równości kobiet, a w kwestii aborcji - wolnością wyboru, tak również liberalni (a nawet niektórzy socjalistyczni) matematycy często zadowolają się pracą pod hegemonistycznym rygorem Zermelo-Fraenkela (który, w obliczu swych liberalnych dziewiętnastowiecznych korzeni, już zawiera aksjomat równości) uzupełnionym jedynie aksjomatem wyboru.”

## 18.4 Socjocybernetyka i formułowanie modelu

Większość badań obliczeniowych nad społeczeństwem opiera się na symulacjach komputerowych. Jest dużo wyników, dużo modeli, ale brakuje bardziej ogólnego i analitycznego podejścia do badań. Zastana sytuacja skłoniła do podjęcia próby uporządkowania i usystematyzowania przynajmniej pewnego wycinka owego modelarskiego zoo i zastanowienia się nad pytaniem: "które modyfikacje są tak naprawdę ważne, a które są raczej mało istotnym szczegółem i pewną nadmiarową ekstrawagancją?". Głównym problemem, jest odpowiedź na pytanie, jak różne typy mikroskopowego wpływu społecznego, wpływają na zachowanie układu w skali makroskopowej. Ponieważ uwzględnienie wszystkich możliwości, jeśli w ogóle możliwe, znacznie wykraczałoby poza ramy owej rozprawy, musiałem dokonać pewnych wyborów, zawężających pole poszukiwań, jednak na tyle rozsądnych, aby objąć swoim zasięgiem jak największą liczbę zagadnień i warto i można zawęzić to do jedynie dobrze znanych z socjologii i psychologii społecznej. Kolejnym problemem jest wybór rozsądnego modelu, który będzie na tyle elastyczny, żeby objąć swoim zasięgiem jak największy obszar zagadnień, ale z drugiej strony będzie dosyć prosty i będzie miał niedużą liczbę parametrów. Następnym dylematem, który staje przed twórcą modelu, był wybór topologii. Niestety istnieje ich bardzo wiele. Autorom powinno zależeć na wyborze takiej, która pozwoli na możliwie najbardziej ogólne i jednoznaczne podejście, oraz uzyskanie jak największej ilości wyników w formie analitycznej. Najczęściej wybierana jest topologia grafu zupełnego, w której każdy agent jest połączony z każdym innym, bądź układ regularny (jak automat komórkowy). Topologia grafu zupełnego dobrze nadaje się do symulowania procesów zachodzących w małych grupach ludzi, ale im grupa większa tym bardziej struktura kontaktów przypomina automat komórkowy. Z drugiej strony, dzięki zastosowaniu grafu zupełnego, definicja reguł dynamicznych badanych przez mnie modeli staje się jednoznaczna. Pytaniem jakie sobie należy postawić jest kwestia dotycząca tego czy są jakieś istotne makroskopowe różnice pomiędzy dwoma głównymi typami zachowań, znanymi z psychologii społecznej/socjologii, a obecnymi w takich modelach. Jak się okazuje, owe różnice mogą się zmanifestować w pewnych warunkach, a w pewnych nie.

Socjocybernetycy krytykują formułowanie modelu wychodząc od obserwacji rzeczywistości i proponują wyjście od aksjomatów. Argumentują to koniecznością zmiany modelu przy każdej nowej obserwacji empirycznej. Wysyp modeli, w kręgach krytyków nazywana „modelarskim zoo” lub „biegunką modelową” jest konsekwencją indukcyjnej i płynnej metodologii nauk (od obserwacji do modelu i z powrotem). Twarda metodologia nauk cybernetyki wzorująca się na matematyce zakłada mechanizm dedukcyjny (Kossecki, 1996) od założeń przez twierdzenia do wniosków. W związku z powyższym aksjomaty elementów układu i relacji między nimi są podstawą do budowy każdego naukowego produktu i tworzą autonomiczny system (Mazur, 1966). Cybernetyk społeczna również stosuje metody matematyczne i statystyczne oraz analizę jakościową zachowań społecznych różnorodnych podmiotów. Tak samo dąży do określenia możliwości, warunków i skutecznych sposobów sterowania zjawiskami i procesami społecznymi. Jednak kolejna różnica leży w dynamice układu, ponieważ cybernetyk wszędzie widzi cel działania (kwestie sterowania), a tradycja przyrodnicza nawiązuje do obserwacji historycznych. To znaczy, że w modelu socjofizycznym przyszły stan układu będzie wyznaczany na podstawie historycznych obserwacji podobnych zachowań, a w produkcie cybernetycznym, na podstawie mechanizmu celu (odgórne sterowanie). Cybernetyka nie uznaje pojęcia „samoorganizacji”. Przykładem różnych interpretacji rzeczywistych może być ruch reformatorski 2013/2014 na Ukrainie. Socjofizyk przede wszystkim będzie w nim widział protest wynikający z „przegrzania układu”, a cybernetyk skoncentruje się na tym komu (w sensie ważnych politycznych graczy) może zależeć na tym konflikcie i wskaże główną przyczynę jako agenturę USA.

### Brak głębi poznania danego tematu

W tym momencie socjofizycy w zasadzie nie publikują w czasopismach 'czysto' socjologicznych. Częściowo właśnie z powodu powierzchowności prowadzonych badań. Jest bardzo mało ludzi pracujących rzeczywiście na pograniczu, bo fizycy zazwyczaj zamykają się w swoich laboratoriach i nie słuchają co socjologowie mają

do powiedzenia. Chyba najsilniejsza strona fizyków – czyli umiejętność wyłowienia decydujących czynników rządzących dynamiką procesu, obraca się niestety w wadę, ponieważ parametrów tak uproszczonych modeli nie da się wykalibrować na danych empirycznych. Niestety fizyków interesują w większości właśnie jedynie zjawiska a nie zastosowanie modeli w praktyce. Znajdują w układach społecznych interesujące ich własności i skupiają się na nich, a nie na opisie rzeczywistości społecznej. Zdecydowanie za mało uwagi poświęca się wykorzystaniu rezultatów w projektowaniu metod pozwalających na uzyskanie korzyści społeczeństwu. Często fizycy podążają ścieżkami będącymi ślepyimi zaułkami kopiowania przyrody. Zbyt mocno akcentują holizm nauki i powszechność występowania pewnych uniwersalnych praw. Przykładem niech będzie rozkład potęgowy w sieciach (opisanych w rozdziale o losowości w sieciach społecznych), bardzo lubiany przez fizyków przez swoje uniwersalne własności. W środowisku fizyków panuje przekonanie, że rozkłady potęgowe występują powszechnie, również w systemach społecznych, gdyż wiele z nich zostało tam wstępnie odkrytych. Jednakże po głębszej analizie niektórzy z odkrywców, zmienili zdanie na temat częstotliwości występowania praw potęgowych w społecznościach ludzkich (Liljeros, 2001). Często te same zjawiska społeczne daje się dokładniej opisać metodami łamiącymi uniwersalność praw potęgowych.

### 18.5 Konsekwencje wprowadzenia modeli w życie

Obserwując bogatsze kraje świata, w których metody modelowania stosuje się na co dzień (w Szwecji żadna ważniejsza decyzja rządu nie obejdzie się bez prognozy modelowej) można łatwo zauważyć niebezpieczeństwa. Chociażby efekt sprzężenia zwrotnego.



Rys. 98 Abstrakcyjna wizualizacja planu zagospodarowania przestrzennego z Wrocławia autorstwa A. Jankowskiej wraz z kolegami ze studiów

W Nadrenii od kilku lat kierowcy mogą korzystać z aplikacji przewidującej natężenie ruchu drogowego opracowanej przez fizyków z Kolonii. Jednakże kierowcy, widząc zagrożenie korkami zaczęli zmieniać swoje trasy podróży. W związku z czym korki zaczęły się pojawiać w innych miejscach. Wiele modeli zostało źle zbudowanych i w wyniku dają błędne prognozy o co bardzo łatwo. Niestety bardzo często na podstawie raportów z nich sporządzonych zostały podjęte konkretne decyzje polityczne. Choćby przypadek pandemii świńskiej grypy na świecie, kiedy to podjęte środki ostrożności zostały mocno przeszacowane ponad realne

zagrożenie. W zaawansowanych technologicznie społeczeństwach azjatyckich już od dawna ludzie stracili wiele praw obywatelskich w wyniku stosowania się do optymalnych strategii społecznych. Jednak konsekwencje, gdy karny naród stosuje się do zaleceń wynikających z modelu, mogą być przerażające, zwłaszcza jeśli model jest błędny. Na szczęście u nas (w Polsce) i tak nie da się stosować większości modeli wprost, bo nie ma dostępu do danych. Pośrednio wprowadza się jednak pewne projekty w życie typu inteligentne miasto, stąd można spodziewać się zielonej fali oraz otrzymywać na bieżąco informacje na telebimach o czasach przejazdu na głównych trasach podróżując po Wrocławiu. Wyniki naukowe mogą być wykorzystane do różnych celów: z jednej strony można ją wykorzystać, aby usprawnić funkcjonowanie społeczeństwa, a z drugiej - w celu manipulowania ludźmi, czyli sterowania nimi wbrew ich interesom.

## 18.6 Metody matematyczne w naukach społecznych

Na szczęście wielu socjologów i humanistów jak Klemens Szaniawski (Szaniawski, 1971), uważa, że metody matematyczne są w socjologii niedostatecznie eksploatowane. Prymitywność metod matematycznych stosowanych w socjologii jest, zdaniem autora jak również wspomnianego Szaniawskiego, przyczyną zastoju w tej dziedzinie. Orientacja na analizę danych oraz wykorzystanie aparatu matematycznie uwalniają w dużej mierze z konkretnej wizji rzeczywistości społecznej. Matematyka jest neutralna w stosunku do problemów rzeczowych socjologii. Najszybciej matematyzuje się analiza struktury danej grupy społecznej dzięki metodom badania sieci społecznych. Również w dziedzinie opis statystycznego oraz w formalnej charakterystyce materiału empirycznego postęp jest widoczny, przez wspomaganie statystyczno-komputerowe. Zdaniem autora najwięcej jest do zrobienia w modelowaniu dynamicznym, czego nie kształci się na polskich uczelniach. Brakuje również użyciu komputera do imitacji procesów (badania *in silico*).

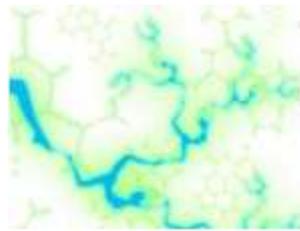
Wiele mitów niemożliwości zastosowania metod przyrodniczych zostało już obalonych w naukach społecznych. Co więcej te metody są pożądane. W socjologii współczesnej istnieją już teorie dające szansę nawiązania kontaktu z sciences (Szmątka, Sozański, 1993). Za Jackiem Szmatką (zdaniem autora najwybitniejszego polskiego socjologa drugiej połowy XX w.) można rozprawić się z mitami utrudniającymi postęp obliczeniowych nauk społecznych. Szmątka rozprawia się z czterema mitami socjologii i trzema generacjach teorii socjologicznych:

- O zasadniczej odmienności świata społecznego, o jego interpretatywności – Rozwiązanie problemów nie wymaga zastosowania środków właściwych socjologii interpretatywnej. Dyskusja staje się nadmiernie filozoficzna, a nierozwiązywalne abstrakcyjne dalekie od obserwacji kwestie lepiej pozostawić filozofii.
- O złożoności świata społecznego znacznie przekraczającej złożoność innych dziedzin rzeczywistości – Wskazuje się tu na sprzeczność, bo skąd wziął się pogląd o większej złożoności świata społecznego, jeżeli nie można wykorzystać metody naukowej w tym celu. Ponadto świat społeczny zawiera mniej informacji niż świat fizyczny, więc nie może być on niego bardziej skomplikowany.
- O czysto eksploracyjnym charakterze socjologii empirycznej – Mimo iż badacz zawsze będzie patrzył z własnego doświadczenia (np. społeczeństwo, w którym żyje) na obiekt, to z drugiej strony należy zachować proporcje badań podstawowych (teoretycznych) i stosowanych (opisowych), przy czym fundamentem pozostaje zbieranie faktów i stawianie hipotez, bez czego nie robi się nauki.
- O użyteczności ogólnych, wszechobejmujących teorii socjologicznych – (w socjologii rzadko uprawiają teorię, a w konsekwencji wiemy żenująco mało o uniwersum społecznym [...] kiedy zaś teoretycy wypowiadają twierdzenia teoretyczne, czynią to, by opisać empirycznie historyczne przypadki lub wznoszą się na skrzydłach filozoficznej wyobraźni do królestwa uniwersalnych funkcji).

Ta książka jest próbą dotarcia do szerszego grona z aktualnymi problemami nauk społecznych z wykorzystaniem metod obliczeniowych bez mityzowania socjologii ani nauk ścisłych, a ze wskazaniem praktycznych możliwości tych metod.

## 18.7 Metody obliczeniowe i co dalej

Na dzień dzisiejszy szczytem możliwości modeli socjofizycznych jest odtworzenie tego, co socjologia już wie i dopiero w niewielu przypadkach pozwoliły one odkryć nowe, nieznane zależności. Dużo większymi sukcesami praktycznymi mogą pochwalić się metody statystyczno-obliczeniowe, ponieważ inteligentne algorytmy nie są już domeną nauki, a wspierają procesy życia codziennego. Opisywana metodologia może jednak pomóc w weryfikowaniu hipotez odnoszących się do konkretnych problemów społecznych, ze względu na powtarzalność i ścisłość metodyki. Wspomniane w tym artykule stworzone przez fizyków narzędzia są na tyle uniwersalne, że mogą zostać użyte po dokonaniu niewielkich modyfikacji w analizie innych różnych zagadnień społecznych. Do niedawna weryfikacja socjofizycznych teorii była w zasadzie niemożliwa z powodu niedostępności danych i niemożliwości prowadzenia wielkoskalowych eksperymentów. Dotychczas jedynym źródłem naprawdę dużych zbiorów danych były wyniki wyborów, czy dane demograficzne. Po raz pierwszy zajęto się tą tematyką badając wyniki wyborów w Brazylii z roku 1998. Ostatnimi czasy prowadzi się eksperymenty na coraz większą skalę. Niektórzy robią je na grupach ochotników, można wtedy szczegółowo ustalić parametry eksperymentu, jednak dotychczas w ten sposób można badać jedynie około 1000 osób. Jednak dzięki dostępowi do danych z serwisów społecznościowych sytuacja się powoli zmienia i dostępne są naprawdę duże zestawy danych. Wiele wskazuje na to, że takie badania mają przed sobą wielką przyszłość i że to dopiero początek drogi. Pewne badania socjofizyczne prowadzą do ciekawych wniosków i niektóre kierunki wydają się być bardzo obiecujące.



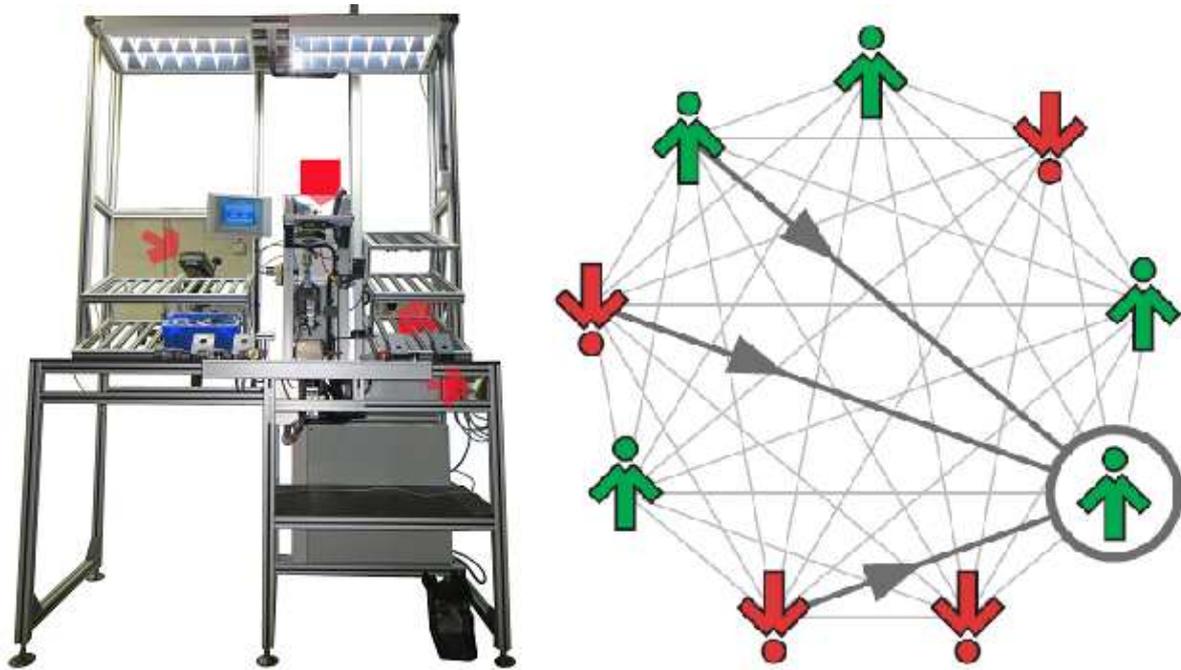
Rys. 99 Ludzie i wzory – fragment grafiki P. Nyczki

Absolwenci polskich uczelni w obszarze modelowania, układów złożonych, czy analizy danych mogą znaleźć pracę w wielu miejscach jak firmy komputerowe, firmy telekomunikacyjne, e-firmy, banki, instytuty naukowe w Polsce i za granicą oraz wszędzie tam, gdzie potrzebna jest interdyscyplinarna wiedza na temat modelowania dynamiki układów społecznych. Wierzmy, że niedługo taka możliwość otworzy się w skostniałych instytucjach administracji publicznej i szeroko pojętej sfery budżetowej służb społecznych jak policja, czy samorządy terytorialne. Potrzebne są nowe umiejętności aby:

- prowadzić obserwacje zdarzeń ekstremalnych jak korupcja w układach stochastycznych;
- tworzyć wizualizacje i analizować zjawiska społeczne w rzeczywistych układach;
- badać izolację i integrację grup społecznych;
- tworzyć algorytmy genetyczne do predykcji zmian opinii;
- analizować sieci transakcji finansowych i bankowych metodami;
- symulować dynamikę opinii społecznej za pomocą automatów komórkowych.

Socjofizyka coraz częściej staje się częścią współczesnych nauk społecznych. Dyrektorami instytutów socjologii zostają fizycy z wykształcenia jak Fredrik Liljeros na Uniwersytecie w Sztokholmie, czy Dirk Helbing na ETH w Zurychu. Nawet w Polsce m.in. na Uniwersytecie Warszawskim fizyką inspirują się psychologowie społeczni (Nowak, 2010), czy socjologowie (Batorski, 2001), a zainteresowanych socjofizyką zachęca się do zapoznania się z przytoczonymi pozycjami książkowymi. Na koniec odnosząc się do jednej z dyrektyw nauk społecznych: "Socjologia jako służba społeczna" (Sitek, 2007), socjofizykę również można traktować jako

służbę społeczną, gdyż docelowo wyniki symulacji powinny przyczynić się dla dobra tegoż społeczeństwa. Modelowanie jest potencjalnym krokiem w stronę tworzenia lepszego społeczeństwa. Nic dziwnego, że np. główną tematyką Instytutu Badań nad Przyszłością (Hedstrom, 2009) jest modelowanie procesów społecznych (Rys. 100).



Rys. 100 Czy mamy szansę zbudować model sprawnie działającego społeczeństwa?

## 19 Spis ilustracji i tabel

Rys. 1 Mały świat powiązań społecznych. Źródło: obraz Idahly Stanley .....	2-8
Rys. 2 Wzory I ludzie, Grafika A. Jankowskiej .....	2-9
Rys. 3 Modele makroskopowe liczebności społeczności w matematycznym opisie dynamiki populacji ...	2-10
Rys. 4 Przykłady zastosowań metod obliczeniowych układów złożonych opracowane przez Agatę Fronczak: Źródło <a href="http://www.if.pw.edu.pl/~agatka">www.if.pw.edu.pl/~agatka</a> .....	2-11
Rys. 5 Wybór opinii pod wpływem interakcji z innymi ludźmi .....	2-11
Rys. 6 Model rozprzestrzeniania się opinii ze źródłami rozprzeczonymi skoncentrowanymi na Ukrainie (stan przed A I po B oraz ścieżki rozchodzenia się opinii C). Źródło: (Dybiec, 2014).....	2-12
Rys. 7 Standarde techniki symulacji metodami dynamiki molekularnej. Jeżeli zamiast cząstek wstawimy ludzi, to możemy mówić o dynamice społecznej. Źródło: <a href="http://udel.edu">udel.edu</a> .....	2-13
Rys. 8 Literalne znaczenie masy krytycznej (lewy). Źródło: <a href="http://longorshortcapital.com/critical-mass-supplier.htm">http://longorshortcapital.com/critical-mass-supplier.htm</a> , ugruntowane społecznie znaczenie (prawy). Źródło: <a href="http://bikeblognyc.com">bikeblognyc.com</a> .....	2-14
Rys. 9 Model przejścia układu z jednego stanu do drugiego w sposób gwałtowny.....	2-14
Rys. 10 Schemat analogii przejścia fazowego między społeczeństwem preferującym stan wolny a stanem związanym .....	2-15
Rys. 11 Klasyfikacja różnych modeli sieci. Źródło (Fronczak, Fronczak, 2009) .....	2-16
Rys. 12 Schemat stosowalności I ewolucji metod obliczeniowych w naukach społecznych. Źródło: (Gilbert, Troitzsch, 2005) .....	3-17
Rys. 13 Graficzne ujęcie najważniejszego prawa ekonomiki zachowań ludzkich, czyli prawa popytu i podaży .....	3-20
Rys. 14 Przykład wykorzystania regresji liniowej (z estymacją parametrów metodą najmniejszych kwadratów).....	3-25
Rys. 15 Wizualizacja wzrostu liczby klastrów w modelowanym układzie .....	3-26
Rys. 16 Ilustracja Samoorganizującej się krytyczności w postaci lawin ze stanów nierównowagowych....	3-29
Rys. 17 Różnica między rozkładem normalnym a potęgowym dominuje w ogonie (zdarzenia rzadkie – dalekie od wartości oczekiwanej).....	3-32
Rys. 18 Stadne zachowania ryb i ptaków (flocking, swarm). Źródło: <a href="http://www.flickr.com/photos/eosgreg/6965121429/">http://www.flickr.com/photos/eosgreg/6965121429/</a> , <a href="http://swantalks.blogspot.com/">http://swantalks.blogspot.com/</a> .....	3-33
Rys. 19 Usieciwienie współczesnego świata. Źródło: <a href="http://www.facebook.com">www.facebook.com</a> .....	4-36
Rys. 20 Standardowy model ewolucji choroby.....	5-42
Rys. 21 Schemat potoków między kategoriami i zależności między nimi w programie Vensim na przykładzie chorób przenoszonych drogą płciową w Polsce.....	5-43
Rys. 22 Gałęziowy proces transmisji patogenu od źródła infekcji (górze) aż do wygaśnięcia epidemii (dół). Poszczególne poziomy oznaczają pokolenia epidemiczne (od 0 na górze do 12 na dole).....	5-44
Rys. 23 Wizualizacja „Gry w życie 2005” autorstwa A. Jarynowskiego jak przykład automatu komórkowego .....	5-44
Rys. 24 Modelowe kontakty między w przypadku 2 subgrup (np. mężczyźni i kobiety). Znając topologię sieci można wnioskować o progę epidemiczną, szybkości rozprzeczzenia się epidemii. Położenie jednostek w sieci pomaga również w ocenie ryzyka zarażenia .....	5-45
Rys. 25 Model dynamicznej sieci kontaktów (z uwzględnieniem momentu kiedy kontakt nastąpił oraz jak długo trwał) .....	5-46
Rys. 26 Wizualizacja interfejsu oprogramowania Netlogo modelu Schellinga (Schelling, 1971).....	5-47
Rys. 27 Wizualizacja sieci modelu agentowego w oprogramowaniu Netlogo .....	5-47
Rys. 28 Przykłady symulacji agentowych w środowisku Netlogo. A – geosymulacje wojny (z wyglądu podobne do gier planszowych, jednak stosowane w praktyce przy prawdziwych konfliktach zbrojnych). B – sztuczne populacje, czyli rozwinięcie gry w życie dla bardziej realistycznych reguł. C – symulacje dla	

rzeczywistego miasta, w tym wypadku konfliktów w Bagdadzie. D – liniowy model ruchu. Źródło: (Weidmann, 2014).....	5-48
Rys. 29 Stopień skomplikowania modelu a jego poziom odzwierciedlenia rzeczywistości w liniowym schemacie. Źródło (Komosiński, 2011).....	5-48
Rys. 30 Klasteryzacja i wykrywanie społeczności .....	6-49
Rys. 31 Społeczność to zespół wierzchołków połączona ze sobą zdecydowanie częściej niż z resztą sieci	6-50
Rys. 32 Średnia odległość między węzłami.....	6-50
Rys. 33 Krotność wierzchołka i centralność wierzchołka.....	6-51
Rys. 34 Losowość od sieci regularnej (np. kratki, czy plastra miodu) przez jej zaburzenie do sieci przypadkowej prostej .....	6-51
Rys. 35 Sieć skierowana w wersji macierzowej oraz zwizualizowana .....	6-53
Rys. 36 Rozwinięty schemat rozwoju choroby typu SEIRS .....	7-56
Rys. 37 Sieć wiedzy Bayesa wygenerowana dla projekcji [czynniki ryzyka występowania zakażenia] – [typy zakażeń związanych]. Źródło (Zygmunt & Valenta, 2005).....	7-57
Rys. 38 Modelowe kontakty wewnątrzszpitalne: czas przebywania wspólnego w tej samej sali (klinice) jako miara siły kontaktu. ....	7-57
Rys. 39 Modelowe kontakty wewnątrzszpitalne wskazujące na potencjalne ścieżki zarażeń między pacjentami oraz personelem. Źródło: (Jarynowski A. , 2013) .....	7-58
Rys. 40 Przykładowe najbardziej prawdopodobne ścieżki zakażeń przy uwzględnieniu strzałki czasu. Źródło: (Jarynowski A. , 2010).....	7-58
Rys. 41 Wizualizacja rozprzestrzeniania się grypy ze źródłem w Wielkopolsce. Źródło (Rakowski, 2010) .	7-61
Rys. 42 Liczba nowych przypadków raka szyjki macicy w Polsce dziennie w poszczególnych latach (po lewej dane historyczne po prawej wyniki symulacji dla różnych scenariuszy). Źródło: (Jarynowski & Sefimovic, 2014).....	7-62
Rys. 43 Grupa przypadków dla stałego altruizmu $x_w=0,5$ (zmiana reputacji o połowę) .....	8-66
Rys. 44 Rozdzielenie faz czystych współpracujących i zdradzających dla zadanych warunków początkowych, z fazą mieszaną w okolicy antydiagonalnej .....	8-66
Rys. 45 Granica faz opisywana dystrybuantą rozkładu Gaussa dla zmian reputacji oraz zmian altruizmu....	8-67
Rys. 46 Przebieg funkcji tanh dla różnych parametrów $a$ i standardowej funkcji liniowej (lewy) i odchylenie standardowe dla różnych wartości parametru $a$ .....	8-68
Rys. 47 Dryf przypadkowy strategii .....	8-68
Rys. 48 Strategie przeskokowe, gdzie przez jakiś czas gracze się zdradzają, by później współpracować dla parametru $a=5$ .....	8-69
Rys. 49 Wizualizacja ruchu ludzi. Źródło: Robert Hillman/Bigstockphoto .....	9-70
Rys. 50 Ewakuacja obiektu sportowego w warunkach eksperymentalnych. Źródło: <a href="http://www.fz-juelich.de/.../PedestrianDynamics/_node.html">http://www.fz-juelich.de/.../PedestrianDynamics/_node.html</a> .....	9-71
Rys. 51 Przykładowy układ sił społecznych w modelu ewakuacji. Źródło: (Lubas, Wąs, Porzycki, & Mycek, 2013).....	9-71
Rys. 52 Ewakuacja biura ze prędkość docelowa $v_D = 2$ m/s. Przedstawiono rozmieszczenie pieszych i trajektorie w czasie ewakuacji oraz czasy ewakuacji w funkcji prędkości docelowej. Źródło: (Grabowski & Kosiński, 2013).....	9-71
Rys. 53 Schemat postępowania pieszego w czasie ewakuacji. Źródło: (Grabowski & Kosiński, 2013) .....	9-72
Rys. 54 Jednowymiarowa dynamika w modelu Sznajdów. Agenci zakreśleni pełną ramką wpływają na agentów zakreślonych ramką przerywaną, ale tylko wówczas gdy mają to samo zdanie (górny rysunek). Jeśli agenci należący do grupy wpływu nie są ze sobą zgodni, nic się nie dzieje (dolny rysunek). Na podstawie : (Nyczka, 2015) .....	10-74
Rys. 55 Asymetria poglądów od prawdopodobieństwa zaburzenia $p$ .....	10-75
Rys. 56 Ponadkrytyczna ilość zaburzeń powodująca przeskoki.....	10-75

Rys. 57 Zbliżająca się odległość między poglądami .....	10-76
Rys. 58 Zależność liczby klastrów od wielkości proggu tolerancji (T).....	10-76
Rys. 59 Charakterystyki poszczególnych klastrów.....	10-77
Rys. 60 Alegoria zmiany społecznej. Źródło: <a href="http://www.demotywatory.pl">www.demotywatory.pl</a> .....	11-78
Rys. 61 Zmiana liczby rozwodów w Polsce I w Portugalii, spowodowana odpowiadającymi czasowo przemianami.....	11-78
Rys. 62 Przyrosty liczby rozwodów dla kilku wybranych państw .....	11-80
Rys. 63 Okno dialogowe symulacji dynamiki populacyjnej małżeństw/ rozwodów w Netlogo.....	11-81
Rys. 64 Makro- charakterystyki wyników modelu dla różnych wartości parametru presji (p) oraz dla rzeczywistych populacji w latach 2010-2012 .....	11-82
Rys. 65 Najlepiej sprzedający się artyści ostatniej dekady .....	12-84
Rys. 66 Większość sprzedaży w pierwszych tygodniach od premiery.....	12-85
Rys. 67 Istotne ( $\alpha = 0.05$ ) dodatnie korelacje między artystami (Buda & Jarynowski, 2013).....	12-86
Rys. 68 Drzewo minimalnej rozpiętości uzyskane dla 30 najpopularniejszych artystów na podstawie sprzedaży ich płyt w latach 2003-2011.....	12-86
Rys. 69 Istotne ( $\alpha = 0.05$ ) ujemne korelacje między artystami (Buda & Jarynowski, 2013) .....	12-87
Rys. 70 Ewolucja poparcia/liczby członków ruchu w czasie .....	14-91
Rys. 71 Wizualizacja społecznych fenomenów na relacjach. Źródło: <a href="http://www.lome.pl/model/">http://www.lome.pl/model/</a> .....	14-91
Rys. 72 Modele zaraźliwości społecznej. Źródło: (Jarynowski, Jankowski, & Zbieg, 2013) .....	14-92
Rys. 73 Model rozprzestrzeniania się ruchu społecznego z parametrami pseudo-epidemiologicznymi. Źródło: (Jankowski, Ciuberek, Zbieg, & Michalski, 2012). .....	14-92
Rys. 74 Skumulowane występowania 20 najczęściej pojawiających się hashtagów w sekwencji czasowej, po lewej 2 dominujące; Po prawej 18 kolejnych.....	14-93
Rys. 75 Połączenie jako współwystępowanie tagów w jednej wiadomości z grubością połączenia jako jej częstość.....	14-93
Rys. 76 Emocje grupują się w jednolite ciągi. Prawdopodobieństwo, że nowo dołączony element będzie podobny do poprzednich wzrasta z liczbą poprzednich (rodzaj emocjonalnego społecznego dowodu słuszności).....	14-94
Rys. 77 Współzależność występowania wyników w ramach rozgrywek Pierwszej Ligi Piłkarskiej w Polsce w sezonie 2006-2007 (kiedy wiele klubów zostało ukaranych za działania korupcyjne w celu ustawiania wyników.....	14-94
Rys. 78 Możliwe pozasportowe rozgrywki (lewy). Źródło: Klatka z filmu Piłkarski Poker Odstającej wartości własne dla Pierwszej Ligi Piłkarskiej .....	14-95
Rys. 79. Drzewo minimalnej rozpiętości dla badanego układu korelacje między cechami wpływającymi na jakość życia Istotne ( $\alpha = 0.05$ ) korelacje jedynie w obrębie trójka waga-wiek-płeć .....	14-96
Rys. 80 Koncepcja sieci Ego, czyli najbliższego otoczenia każdego węzła w sieci relacji. Źródło: <a href="http://www.lome.pl/model/">http://www.lome.pl/model/</a> .....	15-97
Rys. 81 (lewy) Kontakty seksualne mieszkańców wyspy Gotlandii Źródło (Liljeros, 2001): Kontakty seksualne prostytutek (A) i ich klientów (B) Źródło (Rocha, Liljeros, & Holme, 2011):.....	15-98
Rys. 82 Layout ankiety opracowanej do anotacji partnerów seksualnych.....	15-98
Rys. 83 Motywy sieciowe, często wykorzystywane w opisie schematów komunikacyjnych, gdzie kierunek relacji ma znaczenie (lewy). Motywy nieskierowane związków małżeńskich w Sztokholmie (prawy). Źródło: Xin Lu .....	15-99
Rys. 84 Zależności między rybakami z uwzględnieniem techniki rybackiej.....	15-100
Rys. 85 Wizualizacja obszarów oddziaływań metodą zarządzania zasobami naturalnym – NRM (lewy) oraz klasycznej anlaizy sieci społecznych – SNA (prawy). Na podstawie: (Bodin & Crona, 2014) .....	15-101
Rys. 86 Sherwood Anderson "Winesburg, Ohio" <i>The Teacher</i> . Ekstrakcja ręczna (ludzka) sieci explicite i implicite .....	15-102

Rys. 87 Uzyskane sieci społeczne opowiadania <i>The Philosopher</i> z zadania pierwszego-komunikacyjnego (lewe), z zadania drugiego-interakcyjnego (prawe) oraz algorytmu komputerowego - akapitowego (w środku).....	15-102
Rys. 88 Przykładowa sieć współpracy w firmie. Źródło: <a href="https://www.mapaorganizacji.pl/">https://www.mapaorganizacji.pl/</a> .....	15-104
Rys. 89 Przykład statystyk sieciowych spersonalizowanych dla każdego pracownika. Źródło: <a href="https://www.mapaorganizacji.pl/">https://www.mapaorganizacji.pl/</a> .....	15-105
Rys. 90 Możliwość śledzenia zmian występujących na sieci w czasie. Źródło <a href="https://www.mapaorganizacji.pl/">https://www.mapaorganizacji.pl/</a> .....	15-106
Rys. 91 Odsetek głosów nieważnych w wyborach samorządowych 2010 do sejmików wojewódzkich wskutek postawienie więcej niż jednego znaku X. Źródło: <a href="http://uczciwe-wybory.pl/">http://uczciwe-wybory.pl/</a> .....	16-107
Rys. 92 Korelacje między wynikiem partii a % głosów nieważnych, pozytywne współczynniki dla partii rządowych PO- PSL oraz ujemne dla opozycji .....	16-108
Rys. 93 Zestawienie wyników sondaży 4 głównych partii .....	16-110
Rys. 94. Drzewo minimalnej rozpiętości (MST) zbudowane dla partii politycznych w Polsce na podstawie sondaży przeprowadzonych przez Gemius a następnie Homo Homini w latach 2006-2013, przed (lewe) i po (prawe) katastrofie smoleńskiej .....	16-113
Rys. 95 Korelacje między sondażami (lewy) korelacje między wynikami głosowań (prawy) dla różnych okien czasowych głównych partii PO-PiS, 4 lata przed i po Smoleńsku .....	16-114
Rys. 96 Diagram 50 najczęściej stosowanych słów w tej książce wygenerowany przez <a href="http://tagcrowd.com/">http://tagcrowd.com/</a> .....	17-116
Rys. 97 Możliwości wykorzystania Słownosieci w praktyce. Źródło: <a href="http://plwordnet.pwr.wroc.pl/wordnet/">http://plwordnet.pwr.wroc.pl/wordnet/</a> .....	17-119
Rys. 98 Abstrakcyjna wizualizacja planu zagospodarowania przestrzennego z Wrocławia autorstwa A. Jankowskiej wraz z kolegami ze studiów .....	18-123
Rys. 99 Ludzie i wzory – fragment grafiki P. Nyczki .....	18-125
Rys. 100 Czy mamy szansę zbudować model sprawnie działającego społeczeństwa? .....	18-126
Tabela 1 Regresja dla znormalizowanej liczby rozwodów w latach 2010-2012, której wskaźnikiem jest poparcie konserwatywnej partii "PiS" wyborów parlamentarnych w 2011 roku zgrupowane według województw Polski .....	3-25
Tabela 2 Rozkłady prawdopodobieństwa opisujące wagi połączeń dla kontaktów o różnej częstotliwości występowania. Każdy element macierzy zawiera odsetek kontaktów o określonej częstotliwości występowania i wadze (czasie trwania), np. 14% wszystkich kontaktów trwało krócej niż 5 minut. Źródło: (Grabowski & Rosinska, 2012).....	7-59
Tabela 3 Macierz wypłat *(dla ułatwienia przyjmijmy, że tylko gospodarz może przekupić sędziego, choć w ogólności może to być drużyna, której bardziej może zależeć na zwycięstwie, bądź ma więcej pieniędzy).....	8-64
Tabela 4 Zestawienie możliwych konfliktów terytorialnych, o zasoby oraz ze względu na stosowaną technikę połowu.....	15-100
Tabela 5 Macierz korelacji pomiędzy sieciami otrzymanymi od respondentów i za pomocą algorytmu.....	15-102
Tabela 6 Regresja logistyczna zmiennych nominalnych pokazująca relację między wykształceniem ścisłym a zgodnością w uzupełnianiu zadania pierwszego-komunikacyjnego oraz drugiego-interakcyjnego .....	15-103
Tabela 7 Macierz korelacji między partiami politycznymi w Polsce na podstawie sondaży z lat 2006-2010 .....	16-110
Tabela 8 Przejścia polityków między partiami w latach 2001-2010.....	16-112
Tabela 9 Macierz korelacji między partiami politycznymi w Polsce na podstawie sondaży z lat 2006-2013 .....	16-113
Tabela 10 Analiza semantyczna tekstów z socjologicznej perspektywy .....	17-118
Tabela 11 Przykładowa analiza relacji znaczeniowych dla kategorii: praca w korpusach sejmowych ....	17-119

## 20 Bibliografia

- Alexander, P. (1994). New technology and market structure: Evidence from the music recording industry. *Review of Industrial Organization*, 9, 85-98.
- Axelrod, R, et al. (1981). The evolution of cooperation. *Science*, Vol. 211, 4489, s:1390-1396
- Axelrod, R. (1997). Dissemination of culture: a model with local convergence and global polarization. *J. of Conflict Resolution* 41
- Barabási, AL, et al. (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.* 74: s:47-97.
- Ball, P. (2008). *Masa krytyczna*. Tłum. Witold Turopolski, Warszawa: Insignis.
- Batorski, D, et. al. (2001). *Wprowadzenie do dynamiki społecznej*. Warszawa: Wydawnictwo SWPS Academia.
- Batorski, D. (2006). *Społeczna przestrzeń Internetu*. W M. Marody, & A. Nowak. Warszawa: SWPS.
- Benkler, Y (2006), *Bogactwo sieci*, Warszawa, WaiP, s. 45-50
- Bernoulli, D. (1766). "An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it". *Mem Math Phy Acad Roy Sci Paris*, (the english translation by Sally Blower).
- Bertrandt, J., & Lasocki, K. (2012). *Modelowanie i symulacja procesów oraz określenie komputerowo wspomaganých procedur w zakresie zarządzania ryzykiem bezpieczeństwa żywności i żywienia*, redakcja naukowa: Jerzy Bertrandt i Krzysztof Lasocki, Warszawa: Wojskowy Instytut Higieny i Epidemiologii.
- Białynicki-Birula, I, Białynicka-Birula I, (2002) *Modelowanie Rzeczywistości*, WNT: W-wa.
- Bicchieri, C. (2006). *The Grammar of Society*. Cambridge UP.
- Bodin, Ö., & Crona, B. (2014). Social networks in a landscape context: Uncovering social--ecological (mis)matches in heterogeneous landscapes. *in progress*.
- Borgatti S. P., Foster P. C. (2003) *The Network Paradigm in Organizational Research: A Review and Typology*, *Journal of Management* 29(6) 991-1013
- Brian, A. (1994). Inductive Reasoning and Bounded Rationality. *American Economic Review (Papers and Proceedings)* 84.
- Brouwers, L. (2009). *MicroSim: Modeling the Swedish Population*.
- Brouwers, L. (1995). *Microsimulation Models for Disaster Policy Making*. PhD Thesis, Stockholm University.
- Buda, A. (2006). *Historia rocka, popu i hip-hopu według krytyków 1974-2006*. Wrocław: WN.
- Buda, A., & Jarynowski, A. (2010). *Lifetime of correlation*. Wrocław: WN.
- Buda, A., & Jarynowski, A. (2013). Network structure of phonographic market with characteristic similarities between artists. *Acta Physica Polonica A*, 123 (3).
- Buda, A. (2015). *Dynamika i struktura na rynkach finansowych i fonograficznych* (doktorat w przygotowaniu IFJ PAN)
- Business L.O (organizacja). (2013). The forms of promotion on the international music market. Pobrano z lokalizacji <http://ladderofbusiness.wordpress.com/2013/09/24/the-forms-of-promotion-on-the-international-music-market/>
- Castells, M. (2000). *Społeczeństwo sieci*. Warszawa: PWN.
- Camitz, M. (2010). *Computer Aided Infectious Disease Epidemiology - Bridging to Public Health*. Praca doktorska, Karolinska Institutet.
- Capra, F (1996), *Web of Life*, Berkeley, Berkeley Press, s. 82-83.
- Cross R. L., Parker A., Cross R. (2004), *The Hidden Power of Social Networks: Understanding How Work Really Gets Done in Organizations*, Harvard Business Review Press (June 2, 2004)
- Czaplicka, A, Sloot, P, et al. (2013). Noise enhances information transfer in hierarchical networks. *Nature Sci. Rep.* 3, 1223.
- Drucker, P. (1969), *The Age of Discontinuity*, New Brunswik, Transaction Publishers, s.134-145.
- Dybiec, B., Kleczkowski, A., & Gilligan, C. (2009). Modelling control of epidemics spreading by long-range interactions. *Journal of The Royal Society Interface*, 941-950.
- Dybiec, B. (2014). *Dynamiczne sieci społeczne: powstawanie grup społecznych, hierarchii i ich trwałość*, CZM: <https://drive.google.com/viewerng/viewer?url=http://www.czm.mif.pg.gda.pl/wp-content/uploads/baza+wiedzy/ukl-zlo/Dybiec.pdf>
- Enemark, D., McCubbins, MD., Weller, N. (2014). Knowledge and networks: An experimental test of how network knowledge affects coordination, *Social Networks* 36, 122-133.
- Fisher, J (1996), *The Postmodern Paradise*
- Fronczak, A., Fronczak P. (2009). *Świat sieci złożonych. Od fizyki do Internetu*. PWN
- Galam, S, Gefen, Y, Shapir J., (1982) Y, *Journal of Math. Sociology* 9, 1-13

- Galam, S. (2012) *Sociophysics: A Physicist's Modeling of Psycho-Political Phenomena*, Springer, New York
- Ganczarek, W. (2013). Finite size effects in epidemic spreading: the problem of overpopulated systems. *Central European Journal of Physics* , 662-1673.
- Gilbert, N., & Troitzsch, K. (2005). *Simulation for the social scientist*. McGraw-Hill International
- Gubceac, G., Gutu, R., & Paladi, F. (2013). A New Formula for Partitions in a Set of Entities into Empty and Nonempty Subsets, and Its Application to Stochastic and Agent-Based Computational Models. *Applied Mathematics*, 4, 14.
- Głowacka, K. (2012). Seks w 2005 miał kryzys, ale poprawia wyniki..., czyli jakich słów używa polska prasa [online]. TOK FM, Dostęp: [http://www.tokfm.pl/Tokfm/1,102433,11794560,\\_Dusza\\_\\_przegrywa\\_z\\_\\_cialem\\_\\_\\_czyli\\_jakich\\_slow\\_uzywa.html](http://www.tokfm.pl/Tokfm/1,102433,11794560,_Dusza__przegrywa_z__cialem___czyli_jakich_slow_uzywa.html) [27.07.2013].
- Grabowski, A., & Kosinski, R. (2010). Life span in online communities. *Phys. Rev. E* , 82, 066108.
- Grabowski, A., & Kosinski, R. (2004). Epidemic spreading in a hierarchical social network. *Physical review. E* , 031908.1-031908.7.
- Grabowski, A., & Kosinski, R. (2009 ). Modelowanie epidemii. *Bezpieczeństwo Pracy. Nauka i praktyka* , No. 5 (452) .
- Grabowski, A., & Kosiński, R. (2006). Ising-based model of opinion formation in a complex network of interpersonal interactions. *Physica A* , 651-664.
- Grabowski, A., & Kosiński, R. (2013). Matematyczne modelowanie i badania symulacyjne zachowania się ludzi podczas ewakuacji z budynków. *CIOP* (1) .
- Grabowski, A., & Rosinska, M. (2012). The relationship between human behavior and the process of epidemic spreading in a real social network. *Eur.Phys.J D* 83 .
- Green, D., & Bossomaier, E. (2000). *Complex systems*. Cambridge University Press.
- Halley, E. (1692/1693). Estimate of the Degrees of the Mortality of Mankind, drawn from curious. *Tables of the Births and Funerals at the City of Breslaw*. *Philosophical Transactions* , 196.
- Hedström, P, et al. (2009). *The Oxford Handbook of Analytical Sociology*. Oxford: Oxford University Press.
- Helbing, D. (1994). A mathematical model for the behavior of individuals in a social field, . *Journal of Mathematical Sociology* 19 (3).
- Hirsch, JE (2005), An index to quantify an individual's scientific research output , *Proc. Nat. Acad. Sci. (PNAS)*, vol. 102, nr 46.
- Holme, P. (October 2012). Temporal networks. *Physics Reports* , 519 (3).
- Jacob, C. (2010). Branching Processes: Their Role in Epidemiology. *International Journal of Environmental Research and Public Health* , 7, 1186–1204.
- Jankowski, J., Ciuberek, S., Zbieg, Z., & Michalski, R. (2012). Studying paths of participation in viral diffusion process. *Social Informatics* , 503-618.
- Jarynowski, A. (2007). Zjawiska krytyczne w przyrodzie. [online]. Wrocław, Uniwersytet Wrocławski. Dostęp: <http://th.if.uj.edu.pl/~gulakov/kracylin.pdf> [27.07.2013].
- Jarynowski, A. (2013). Modelowanie epidemiologiczne przy wykorzystaniu analizy tymczasowych sieci społecznych. *W Postępy inżynierii biomedycznej*. Rzeszów: Inprona.
- Jarynowski, A. (2010). Anomalous interactions in network of Polish Football League. In *Life time of correlation*. Wrocław: WN.
- Jarynowski, A. (2010). Contact networks and the spread of MRSA in hospitals. conference poster.
- Jarynowski, A, Gawroński, P, Kułakowski, K. (2012). How the competitive altruism leads to bistable homogeneous states of cooperation or defection. *LNCS*, 7204 s: 543.
- Jarynowski, A, Kliś, M. (2012). Socio-economic models of divorces in different societies. W: *Proceedings of the XVIII National Conference Applications of Mathematics in Biology and Medicine*, Gdańsk.
- Jarynowski, Andrzej, Nyczka, P. (2013). Dynamic network approach to marriage/divorces problem, ENIC, DOI 10.1109/ENIC.2014.24
- Jarynowski, A, Boland, S. (2013). Rola analizy sieci społecznych w odkrywaniu struktury fikcji literackiej. *Buletyn Instytutu Systemów informatycznych*. 12, 35-42
- Jarynowski, A., & Sefimovic, A. (2014). Studying Possible Outcomes in a Model of Sexually Transmitted Virus (HPV) Causing Cervical Cancer for Poland. *Advances in Intelligent Systems and Computing* , 129-141, 229.
- Jarynowski, A., Jankowski, J., & Zbieg, A. (2013). Viral spread with or without emotions in online community. *arXiv:1302.3086*.
- Kosiński, Robert, Grabowski, Andrzej. (2010). Langevin Equations for Modeling Evacuation Processes. *Acta Physica Polonica B*, 3, 365.

- Kasprzyk, R. (2010). Symulator rozprzestrzeniania się złośliwego oprogramowania w sieciach komputerowych. *Symulacja w Badaniach i Rozwoju*, 139-150.
- Kasprzyk, R., Najgebauer, A., & Pierzchała, D. (2010). Creative Application to Remedy Epidemics. *Risk Analysis VII and Brownfields V*, 545-562.
- Kolk, M., Cownden, D., & Enquist, M. (2014). Correlations in fertility across generations: can low fertility persist? *Proceedings of the Royal Society B: Biological Sciences*.
- Komosinski, M. (2011). Życie w komputerze: symulacja czy rzeczywistość?. *Nauka 2*, PAN
- Kossecki J., (1996) *Cybernetyczna analiza systemów i procesów społecznych*. Wyd. WZiA WSP, Kielce.
- Kulig, A, Kwapien, J, et al. (2012). Complex network analysis of literary and scientific texts. *International Journal of Modern Physics C*, 23, 1250051.
- Kułakowski, K. (2012). Od fizyki do socjologii i z powrotem. W: A. Chmielewski, M. Dudzikowa i A. Grobler (red.), *Interdyscyplinarnie o interdyscyplinarnosci*, Kraków: Impuls, s: 99-111
- Kwapien, J., & Drożdż, S. (2012). Physical approach to complex systems. *Physics Reports*, 515 (3), pp. 115-226.
- Leskovec, J., Adamic, L., & Huberman, B. (2007). The dynamics of viral marketing. *ACM Trans. Web*, May.
- Liljeros, F. et al (2001). The web of human sexual contacts. *Nature*, 411 (6840), 907-908.
- Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., ... & Van Alstyne, M. (2009). Life in the network: the coming age of computational social science. *Science (New York, NY)*, 323(5915), 721.
- Lubas, R, Wąs, J, Porzycki, J, Mycek, M. (2013) *Wspomaganie projektowania budynków pod kątem optymalnego przepływu ludzi*, ISI AGH Kraków.
- Mantegna, R. (1999). Hierarchical Structure In Financial Markets. *Eur. Phys. J. B*.
- Matuszak, M. (2010). Zastosowanie modelu Helbinga w symulacji ewakuacji. *Praca mgr (UWr)*.
- Mazur M., (1966). *Cybernetyczna teoria układów samodzielnych*. PWN, Warszawa/
- Merton, R. (1968). The Matthew effect in science. *Science*, Jan 5; 159 (3810), s: 56-63.
- Mezard, M., Parisi, M., & Virasoro, A. (1987). *Spin glass theory and beyond*. Singapore: World Scientific.
- Mika, B. (2004). *Dopominając się o kontekstowy wymiar muzyki*. De musica, Vol. VII.
- Mossong, J., & Rosinska, M. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *Plos Medicine Vol. 5 No. 3*.
- Murray, J. (2002). *Mathematical Biology. I. An Introduction (chapter 10: Dynamics of Infectious Diseases: Epidemic Models and AIDS)*. Springer (dostępne jest polskie tłumaczenie Urszuli Forys).
- Mysliwski, M. (2011). A microeconomic analysis of album sales success in the Polish music market. *Warsaw: SGH (No 54)*.
- Nagel, K., & Schreckenberg, M. (1992). A cellular automaton model for freeway traffic. *Journal de Physique I 2 (12)*.
- Najgebauer, A. (2009). *Modele zagrożeń aglomeracji miejskiej wraz z systemem zarządzania kryzysowego na przykładzie miasta stołecznego Warszawy*. Warszawa: Wojskowa Akademia Techniczna.
- Najmrodzka, S. (2012). *Badanie jakości życia pacjentów z chorobami tarczycy*. Poznań: Medical University.
- Niezgódka, M. (2006). E-styl życia w społeczeństwie informacyjnym. W M. Niezgódka, *Spółeczeństwo informacyjne. Aspekty funkcjonalne i dysfunkcjonalne*. Kraków: WUJ.
- Nowak, M. (2006). Five rules for the evolution of cooperation. *Science*.
- Nowak, A, et al. (2010). *Układy złożone w naukach społecznych. Wybrane zagadnienia*. Warszawa: Scholar.
- Nyczka, P. (2015). *Modelowanie q-p wyborcy*. UWr (doktorat w przygotowaniu).
- Oleś, K., Gudowska-Nowak, E., & Kleczkowski, A. (2012). Understanding disease control: influence of epidemiological and economic factors. *PloS one*, e36026.
- Ordanini, A. (2006). Selection models in the music industry: How a prior independent experience may affect chart success. *Journal of Cultural Economics*, 30, 183-200.
- Ossowski, S. (2005). *Wzory nauk przyrodniczych wobec osobliwości zjawisk społecznych*. W: Marek Kucia (red.), *Socjologia. Lektury*, Kraków: Znak, s: 28-34.
- Oswiecimka, P., Kwapien, J., & Drożdż, S. (2011). Computational approach to multifractal music. *arXiv:1106.2902*.
- Pabjan, B. (2004). The use of models in sociology. *Physica A: Statistical Mechanics and its Applications*, Volume 336, Issues 1–2, s: 146–152.
- Percino, G., & all., e. (2014). Instrumentational complexity of music genres and why simplicity sells. *arXiv:1405.5057*.
- Pękalski, A., & Cebrat, S. (1999). Model of population evolution with and without eugenics. *The European Physical Journal B*.

- Pilarczyk, A. (2009). Modelowanie oligopolu metodami fizyki statystycznej. praca mgr UW (pod kierunkiem K. Sznajd-Weron) .
- Porzycki, J. (2013). Model epidemiologiczny dla Krakowa. Kraków: AGH (praca mgr).
- Qiao, M., Liu, A., & Foryś, A. (2013). Qualitative analysis of the SICR epidemic model with impulsive vaccinations. *Mathematical Methods in the Applied Sciences* , 695-706.
- Rakowski, F. (2010). Influenza epidemic spread simulation for Poland — a large scale, individual based model study. Pobrano z lokalizacji [http://www.icm.edu.pl/\simmagd/filmy/epidemic\\_f0.6.avi](http://www.icm.edu.pl/\simmagd/filmy/epidemic_f0.6.avi)
- Rakowski, F., Gruziel, M., Bieniasz-Krzywiec, L., & J, R. (2010). Influenza epidemic spread simulation for Poland — a large scale, individual based model study . *Physica A* 389(16) , 3149-3165.
- Rakowski, F., Gruziel, M., Krych, M., & Radomski, J. (2010). Large Scale Daily Contacts and Mobility Model - an Individual-Based Countrywide Simulation Study for Poland. *JASSS* 01/13 .
- Rocha, L. (2011). Exploring patterns of empirical networks. PhD Thesis, Umea University.
- Rocha, L., Liljeros, F., & Holme, P. (2011). Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts. *PLoS Comp. Bio* , 7 (3).
- Rocha, LEC, Liljeros, F, et al. (2010). Information dynamics shape the sexual networks of Internet-mediated prostitution. *PNAS*, 107 (13), s: 5706-5711.
- Rogers, E. (2005). *Diffusion of Innovations*. New York: Free Press.
- Rzazewski, L, Słomczynski, W, Życzkowski, K. (2014). *Każdy głos się liczy*, Warszawa
- Saliowicz. P. (1999), *Spółeczeństwo informacyjne*
- Schelling, T. (1971) *The American Economic Review*, Vol. 59, No. 2 Sitek, W. 2007. *Paradoksy prognoz socjologicznych. W: Socjologia jako służba społeczna*, Kraków: WUJ.
- Szaniawski, K . (1971) *Metody matematyczne w socjologii*, Ossolineum.
- Szmatka, J, et al. (1997). *Status, Network, and Structure: Theory Development in Group Processes*, Nowy Jork: Stanford University Press.
- Sztompka, P. (1973). O osobliwościach nauk społecznych raz jeszcze. *Studia filozoficzne*, nr 8 (105).
- Syrda, J. (2013). *The Economics of Music Charts*. Come. Warszawa.
- Sznajd-Weron, K. (2001), grudzień. *Opowieści o fizyce egzotycznej*. *Wiedza i Życie*.
- Sznajd-Weron, K. (2002), kwiecień. *Seks według wzoru*. *Wiedza i Życie*.
- Sznajd-Weron, K. (2004), luty. *W sieci małego Świata*. *Wiedza i Życie*.
- Sznajd-Weron, K., Weron, R., & Włoszczowska, M. (2008). Outflow dynamics in modeling oligopoly markets: the case of the mobile telecommunications market in Poland. *J. Stat. Mech.* P11018 .
- Szostek, G., Jaszuk, M., & Walczak, A. (2012). "Automatyczna budowa semantycznego modelu objawów chorobowych na bazie korpusu słownego." . *Biuletyn Instytutu Systemów Informatycznych* , 9, 35-43.
- Szymański, P. (2008), *Technologia, a wolność społeczeństw* , Wrocław, SunRise, s. 1-2
- Van Dijk, T. A. (1985). *Semantic discourse analysis*. *Handbook of discourse analysis*, 2, 103-136.
- Walsh, R. (2011). *Emergent Narrative in Interactive Media*. *Narrative*, vol. 19, nr 1, s: 72-85.
- Watts, D., & Strogatz, S. (1998). *Collective dynamics of small worlds networks*. *Nature* , 393 (440).
- Weidmann, N. (2014). *Agent-Based Modeling in the Social Sciences*, Vienna: ECPR
- Wdowiak, R. (2007). *Identyfikacja struktur sieci złożonych (model ptasiej grypy H5N1)*. Wrocław: PWr (praca dyplomowa).
- Wiener, N. (1950). *The Human Use of Human Beings*, London: The Riverside Press.
- Wywrot, R. (2004). *Modelowanie rozprzestrzeniania się epidemii SARS*. Wrocław: PWr (praca dyplomowa).
- Zacher, L. (2005), *Transformacje społeczeństw od informacji do wiedzy*
- Zbieg, A, Żak, B, Zaręba P. (2014). *Analiza Sieci w Badaniach Struktury Organizacji*, W: *Nauki o Zarządzaniu* 2 (19), Wydawnictwo Uniwersytetu Ekonomicznego, Wrocław, s. xx-xx – w druku
- Zygmunt, A., & Valenta, M. (2005). *Pozyskiwanie wiedzy probabilistycznej dla modelu zakażeń szpitalnych*. *Prace Naukowe Akademii Ekonomicznej we Wrocławiu* , 290-301.
- Żak, B, Zbieg, A, Możdżyński D. (2014). *Mapaorganizacji.pl – partycypacyjna platforma badań sieci organizacyjnych*, W: *Nauki o Zarządzaniu*, 1 (18), Wydawnictwo Uniwersytetu Ekonomicznego, Wrocław, s. 100-110.

## 21 Nota o autorach

Pomysł na tę książkę zrodził się w głowach absolwentów fizyki Uniwersytetu Wrocławskiego, którzy pracując w instytucjach publicznych oraz prywatnych zauważyli brak wspólnego języka między praktykami społecznymi, a twórcami modeli.

### **Andrzej Jarynowski**

Andrzej studiował na Politechnice Gdańskiej (informatyka stosowana), Politechnice Wrocławskiej (matematyka przemysłowa), Uniwersytecie Wrocławskim (ekonofizyka i fizyka komputerowa oraz socjologia), Uniwersytecie Sztokholmskim (socjologia i rozwiązania systemowe), czy Karolinska Institutet (epidemiologia). Pochodzi z Pomorza – IN MARI VIA TUA. Obecnie próbuje obronić doktorat z fizyki na Uniwersytecie Jagiellońskim w Krakowie i jest stażystą w Pracowni Wirtualnej Rzeczywistości, CIOP-Państwowy Instytut Badawczy w Warszawie oraz specjalizantem epidemiologicznym w Województwie Mazowieckim przy WIHiE. Początkowo zajmował się bioinformatyką i biomatematyką, aż odnalazł swoją pasję w modelowaniu epidemiologicznym i w socjologii analitycznej. Jest autorem około 20 artykułów (w dziedzinie fizyki, informatyki, matematyki, medycyny, czy nauk społecznych, ale również archeologii, czy prawa). W swojej pracy badawczej łączy polską szkołę matematyczno – fizyczną z nordyckimi technikami innowacyjnymi w zakresie zdrowia publicznego i życia społecznego. W kręgu jego zainteresowań są głównie zastosowania technik układów złożonych i modelowania w naukach społecznych, a w szczególności odkrywanie powiązań sieciowych w epidemiologii, czy kryminologii, a ostatnio również w muzyce i literaturze. Został wyróżniony przez ONE BSR (One Baltic Sea Region) za uwzględnienie wyników jego badań w decyzjach podejmowanych przez rząd Szwecji w dziedzinie wspieranej komputerowo epidemiologii. Pisząc tę książkę opracowuje model rozprzestrzeniania się chorób przenoszonych drogą płciową w Mołdawii, a od roku 2015 Andrzej zostanie kierownikiem projektu NCBI „Innowacje Społeczne” którego zadaniem będzie implementacja algorytmów sieci społecznych w profilaktyce zakażeń szpitalnych w Polsce.

### **Andrzej Buda**

Polski fizyk, aktywista pro-wolnościowy, dziennikarz muzyczny, autor Historii kultury hip-hop w Polsce, pracy doktorskiej o zastosowaniu fizyki na rynkach finansowych i książek o rynku płytowym. W wolnym czasie żegluje i podróżuje zwiedzając kraje tropikalne w poszukiwaniu słońca.

### **Piotr Nyczka**

Młody, choć nie tak bardzo jak by chciał naukowiec, niespokojny duch łąpczywie rzucający się na wszystko, co go interesuje. W dzieciństwie pragnął zostać astronautą, po to aby w dorosłym życiu przemierzać pustkę, zamknięty samotnie w rakiecie mknącej do którejś z bliższych gwiazd. Niestety niespodziewany koniec zimnej wojny pokrzyżował te ambitne plany. Obecnie jego zainteresowania naukowe oscylują gdzieś pomiędzy twardą fizyką a naukami społecznymi. W tej chwili nie można dokładnie przewidzieć jakie będą owoce tej szamotaniny. Interesuje się sztuką, wzornictwem, fotografią, namiętnie projektuje i wykonuje nadruki na ubraniach ekologicznych materiałów, bądź z odzysku, nie jest mięsożerny, ale uwielbia gotować, czasami dłubie przy elektronice i rowerach, ponadto aktywnie uczestniczy w podziemnym świecie ciężkich odmian muzyki elektronicznej jako grajek i producent. Jednakże wszystko to blednie w obliczu cechującej go, niespotykanej wręcz, skromności.