

MODELOWANIE EPIDEMIOLOGICZNE: KLASYCZNIE I SIECIOWO

Andrzej Jarynowski^{(a),(b)}

^(a) Instytut Fizyki, Uniwersytet Jagielloński

^(b) Instytut Socjologii, Uniwersytet Sztokholmski

Charakterystyka kontaktów społecznych w postaci czasowej i sieciowej struktury (kiedy i z kim) ma wpływ na rozprzestrzenianie się chorób zakaźnych. W związku z rozwojem stosowanych metod komputerowych pojawiła się perspektywa wykorzystania ich w analizie sieci społecznych. Fizycy ze swoimi modelami starają się zrozumieć zachodzące zjawiska i w konsekwencji pomóc w kontroli epidemiologicznej. W tej pracy postaram się przybliżyć podstawową metodykę działania w celu zarekomendowania służbom epidemiologicznym najkorzystniejszych działań.

1 Potrzeba modeli epidemiologicznych

Zgodnie z klasycznym trójkątem epidemiologicznym [1], choroba zakaźna może się rozprzestrzeniać tylko w obecności źródła, podatnych jednostek i odpowiedniego środowiska. Na elementy trójkąta (osobno lub w połączeniu) kierowane są działania prewencyjne oraz interwencyjne. Zazwyczaj najważniejszym etapem działania epidemiologicznego jest identyfikacją czynnika chorobotwórczego. Całą trudnością jest ustalenie źródła na podstawie posiadanych danych. W moich badaniach będę się skupiał na innym elemencie trójkąta epidemiologicznego – na łańcuchu zakażeń. Znając historię transmisji zarażeń, służby epidemiologiczne mają szansę znaleźć miejsca, obszary, bądź ludzi lub inne istoty żywe – rezerwuariat patogenu. Nie jest to zadanie łatwe i do tej pory tylko na małą skalę realizowane. Techniki komputerowe pozwoliły na automatyzację w procesie zbierania i analizy informacji jak również na postulowanie i symulowaniu różnych scenariuszy zarażeń. Epidemiologiczne modele, które traktują o transmisji z człowieka

na człowieka, od lat 20-tych XX-go opisywane są równaniami różniczkowymi. Jednak w nowszych modelach pojawiają się częściej modele agentowe (agent based models). Symulacje komputerowe zaczęły grać znaczącą rolę od kiedy ilość interakcji społecznych jest ogromna, a jednostki nie są homogeniczne. Mimo wszystko, to dzięki rozwiązaniom równań różniczkowych nauczyliśmy się „rozumieć” procesy epidemiologiczne i dlatego, z edukacyjnego punktu widzenia, warto od nich zacząć.

2 Klasyczny model epidemii

Matematyczny opis zjawisk wymaga określenia założeń w celu zrozumienia rzeczywistej sytuacji [2]. Załóżmy, 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 ze społeczeństwa (zostali odizolowani lub po prostu zmarli). Taki system może zostać zapisany w postaci układu równań różniczkowych:

$$\frac{dI}{dt} = rSI - aI, \frac{dS}{dt} = -rSI, \frac{dR}{dt} = aI \quad (1)$$

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 progiem. Zdefiniujmy zatem:

$$R_0 = \frac{rS_0}{a} \quad (2)$$

gdzie R_0 stopą reprodukcji epidemii. 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. Z powyższego przykładu wynika, że równania różniczkowe pozwalają zrozumieć zjawisko i pomóc w kontrolowaniu epidemicznym, ale w przypadku ogromnej ilości wiedzy, jaką posiadamy o społeczeństwie, nie dają się już tak łatwo stosować, stąd rozwój modeli sieciowych.

3 Analiza sieciowa

Aby móc symulować epidemię na sieciach, najpierw wprowadzę SNA (*social network analyse*), prężnie rozwijającą się dziedzinę nauki. 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. Opisuując sieć używa się różnych współczynników oraz metod rozpoznawczych.

1. Klasteryzacja: sposób grupowania (clustering), 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.
2. Średnia odległość między węzłami: średnia długość ścieżki w sieci mówi, przez ile wierzchołków należy średnio przejść, aby dojść od jednego wierzchołka do dowolnego innego.
3. Rozkład krotności wierzchołka (który będzie miał kluczową rolę w klasyfikacji sieci): rozkład ilości kontaktów posiadanych przez poszczególne jednostki.
4. Losowość: 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 [3]:

- (a) rozkład krotności wierzchołków spełnia prawo potęgowe (oczekiwana ilość wierzchołków o stopniu k będzie proporcjonalna do $k^{-\gamma}$, wtedy klasyfikujemy tę sieć jako bezskalową lub inaczej Barabasięgo-Alberta (BA);
- (b) rozkład krotności wierzchołków spełnia prawo wykładnicze (oczekiwana ilość wierzchołków o stopniu k będzie proporcjonalna do $\exp(-k)$), wtedy klasyfikujemy tę sieć jako przypadkową prostą lub inaczej Erdos'a-Renyiego (ER).

Pomiędzy skrajnymi przypadkami (regularnymi i losowymi) plasuje się sieć małego świata, często nazywana Watta i Strogatza (WS), gdzie na sieć regularną nałożone są przypadkowe połączenia.

Wykorzystanie aspektu sieciowego pozwala (jeżeli dysponujemy odpowiednimi danymi) znacznie lepiej przewidzieć skutki epidemii oraz pomaga w wyborze optymalnej strategii walki z epidemią o czym można więcej przeczytać na przykład w moim artykule o szacowaniu kosztów grypy [4], czy modelowaniu zakażeń szpitalnych i chorób przenoszonych drogą płciową [5] oraz na mojej stronie internetowej (<http://th.if.uj.edu.pl/gulakov>).

Literatura

- [1] W. Jedrychowski, *Podstawy Epidemiologii. Metody badań oraz materiały ćwiczeniowe.*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków, 2002.
- [2] J. Murray, *Mathematical Biology. I. An Introduction.*, 2002.
- [3] AL. Barabási *et al.* , Rev. Mod. Phys. 74: 47–97 2002.
- [4] A. Jarynowski, *Human-Human interaction: epidemiology w: Lifetime of Correlations*, Wrocław, 2010.
- [5] A. Jarynowski, SiMIS, Tom 4, Nr 4, 2012.