

Recenzja rozprawy doktorskiej

Autor: mgr Frederic Grabowski

Tytuł: Decoding viral infections and cellular signalling via mathematical modelling

Promotor: prof. dr hab. Tomasz Lipniacki

Promotor pomocniczy: dr Marek Kochończyk

Ogólna charakterystyka rozprawy

Przedłożona do recenzji rozprawa doktorska jest napisana w języku angielskim, liczy 253 strony. Jest bezpośrednio powiązana z 10 oryginalnymi publikacjami naukowymi, których współautorem jest Doktorant. W tekście pracy doktorskiej 44 strony na początku stanowią omówienie tych artykułów naukowych, a dalsze strony, 45 – 253, zawierają kopie publikacji. Literatura naukowa cytowana za omówieniem obejmuje 48 pozycji.

Tematyka rozprawy dotyczy modeli matematycznych oraz numerycznych związanych ze scenariuszami rozprzestrzeniania się infekcji wirusowych, wirusa COVID oraz wirusa grypy. Rozwijanie się infekcji wirusowych, w obrębie organizmu gospodarza, a także w populacji ma złożone, wieloaspektowe podłoże biologiczne oraz charakteryzuje się wieloskalową, skomplikowaną dynamiką wymagającą używania różnorodnych modeli matematycznych i numerycznych. Jest to tematyka bardzo szeroka i decyzję Doktoranta o poświęceniu swojej dysertacji doktorskiej jej rozwijaniu należy uznać za bardzo ambitną i interesującą. Część wyników / publikacji bazuje na eksperymentach biologicznych *in vitro* wykonanych w laboratorium / zespole, którego członkiem jest Doktorant (obrazowanie immunofluorescencyjne, ELISA, dPCR i Western blot), co należy uznać za silną stronę pracy.

Doktorant dzieli publikacje wchodzące w skład swojej pracy na dwie grupy, pierwsza obejmuje modele dynamiki rozwijania się infekcji wirusowej, druga dotyczy teorii informacji i

sygnalizacji komórkowej. Pierwsza grupa obejmuje 6 publikacji, a druga 4. Cztery publikacje należące do drugiej grupy można jeszcze dodatkowo podzielić na dwie podgrupy. Jedną to dwie prace związane z zastosowaniem teorii informacji do badania sygnalizacji komórkowej, a druga to dwie ostatnie prace (opublikowane w innym niż pozostałe zespole autorskim) poświęcone numerycznym aspektom badania informacji wzajemnej pomiędzy parami zmiennych losowych.

W opisie (str. 38) Doktorant wspomina także dwa dodatkowe, współautorskie artykuły, których tematyka nie jest bezpośrednio związana z obszarem badań w pracy.

Praca zawiera 10 oryginalnych tez / hipotez, sformułowanych na stronach 7 – 9. Każda z tez jest bezpośrednio powiązana z jedną z publikacji naukowych wchodzących w skład rozprawy.

Omówienie publikacji wchodzących w skład rozprawy

Publikacja, „Grabowski, F., Kochańczyk, M., Korwek, Z., Czerkies, M., Prus, W., & Lipniacki, T. (2023). Antagonism between viral infection and innate immunity at the single-cell level. *PLoS Pathogens*, 19(9), e1011597”, (**Antagonism**), jest studium scenariusza infekcji komórek gospodarza (hodowli komórkowej A549, human alveolar cell line) przez wirusa RSV (respiratory syncytial virus). Praca bazowała zarówno na technikach eksperymentalnych obrazowania pojedynczych komórek jak też na modelowaniu stochastycznym. Wiadomo, że po zakażeniu wirusem komórki mogą wydzielać interferony (IFN), które pobudzają pobliskie komórki do przygotowania się na nadchodzącą infekcję. Z drugiej strony, białka wirusowe często zakłócają syntezę IFN i sygnalizację indukowaną przez IFN. RSV blokuje wrodzone odpowiedzi komórkowego układu odpornościowego poprzez hamowanie aktywacji IRF3, hamowanie syntezy IFN i hamowanie aktywacji STAT1/2. Z kolei odpowiedź układu odpornościowego białka kodowane przez geny stymulowane IFN (aktywowane przez STAT1/2) hamują syntezę wirusowego RNA i białek wirusowych. Sytuacja ta nazywana jest antagonizmem pomiędzy działaniem wirusa a odpowiedzią komórkową. Tezą tej publikacji oraz rozprawy doktorskiej jest, że istnienie tego antagonizmu może być adekwatnie zbadane / wykazane przy użyciu techniki eksperymentów na pojedynczych komórkach. Przy wykorzystaniu obrazowania mikroskopii immunofluorescencyjnej, wykazano, że wzajemny antagonizm między wirusem RSV a zakażonymi komórkami A549 prowadzi do dychotomicznych odpowiedzi na poziomie pojedynczej komórki i złożonych wzorców przestrzennych stanów sygnalizacji komórkowej. Przeprowadzone równoległe symulacje, z wykorzystaniem stochastycznego systemu agentowego potwierdziły scenariusze oddziaływania wirusa z układem odpornościowym komórek A549.

Publikacja, „Prus, W., Grabowski, F., Koza, P., Korwek, Z., Czerkies, M., Kaczyńska, P., ... & Lipniacki, T. (2024). Type III interferons may suppress viral infections by triggering cell death. *bioRxiv*, 2024-09”, (**Lambda**) poświęcona była analizie scenariuszy infekcji komórek

A549 przez wirusa grypy typu A (IAV). Główną tezą tej pracy jest, że interferony typu III (IFN- $\lambda 1$ – $\lambda 4$) komórkowego układu odpornościowego pełnią inną rolę w odpowiedzi na infekcję wirusa IAV niż interferony typu I (IFN β). Ograniczają one rozprzestrzenianie się wirusa grypy typu A poprzez promowanie obumierania komórek zakażonych wirusem grypy typu A. Teza ta jest również tezą drugą doktoratu. Teza ta jest wspierana przez wykonanie eksperymentów *in vitro* polegających na dokonaniu „knock outu” funkcjonalności białek interferonu INF λ (IFNLR1 KO). Eksperymentalnie wykazano, że stosunek komórek obumierających do zakażonych jest niższy w komórkach IFNLR1 KO w porównaniu z komórkami typu dzikiego. Badania eksperymentalne uzupełniono przez komputerowe modelowanie dynamiki zakażenia hodowli komórek A549 przez wirusa grypy typu A. Użyto deterministycznego modelu SIR, którego analiza pozwoliła na replikację obserwacji w eksperymentach z komórkami A549.

W publikacji, „Kochańczyk, M., Grabowski, F., & Lipniacki, T. (2020). Dynamics of COVID-19 pandemic at constant and time-dependent contact rates. *Mathematical Modelling of Natural Phenomena*, 15, 28”, (**Dynamics**), skonstruowano model rozprzestrzeniania się epidemii COVID-19, oparty na równaniach SEIR (Susceptible–Exposed–Infectious–Removed, Podatny–Narażony–Zakażony–Usunięty), parametryzowany przez średni okres inkubacji, τ , oraz dwa parametry tempa zakażeń oraz wyzdrowień/wykluczeń: wskaźnik kontaktów, β , i wskaźnik wykluczenia, γ . Wykazano, że wprowadzenie kaskady stanów narażenia (E1, E2, E3, E4, E5) pozwala na odtworzenie rozkładów okresu inkubacji i czasu od zakażenia do zarażenia zgodnych z danymi epidemiologicznymi. Na bazie wprowadzonego modelu rozważano także scenariusz, w którym β jest modulowane przez zmiany zachowań społecznych: najpierw β spada w odpowiedzi na gwałtowny wzrost liczby nowych przypadków dziennie, wywierający presję na ludzi, aby poddali się samoizolacji, a następnie, w dłuższej perspektywie, β rośnie, w miarę jak ludzie stopniowo akceptują ryzyko. W tym scenariuszu, po początkowym gwałtownym rozprzestrzenianiu się epidemii następuje plateau i powolna regresja.

W pracy, „Kochańczyk, M., Grabowski, F., & Lipniacki, T. (2020). Super-spreading events initiated the exponential growth phase of SARS-COVID-19 with \mathcal{R}_0 higher than initially estimated. *Royal Society open science*, 7(9), 200786”, (**Super-spreaders**), rozważano sytuację, w której tempo zakażeń w modelu SEIR rozprzestrzeniania się epidemii SARS-COVID nie jest stałą liczbą, lecz może zmieniać się w czasie. Mechanizm zmian parametru β w czasie wynika z hipotezy o istnieniu superroznośców (super-spreaders), którzy przez swoje wyjątkowo liczne kontakty społeczne przyczyniają się do modyfikacji dynamiki rozwoju epidemii SARS-COVID. Hipoteza / teza o istnieniu superroznośców i ich roli w epidemii SARS-COVID jest także tezą 4 doktoratu. Teza ta jest wykazywana na podstawie rozbieżności pomiędzy początkowymi szacunkami R_0 (podstawowa liczba odtwarzania, basic reproduction number) dla COVID-19 (między 2 a 4) a znacznie szybszym rozprzestrzenianiem się epidemii w różnych krajach, w tym w krajach europejskich (R_0 w zakresie 4,7–11,4). Obliczenia i szacowanie parametrów modeli były wspierane przez symulacje stochastyczne z

zastosowaniem algorytmu Gillespiego, w których zakładano różne wartości proporcji subpopulacji superroznośców w stosunku do całej populacji, oraz przez deterministyczne trajektorie wariantów modelu SEIR.

Praca, „Grabowski, F., Preibisch, G., Giziński, S., Kochańczyk, M., & Lipniacki, T. (2021). SARS-CoV-2 variant of concern 202012/01 has about twofold replicative advantage and acquires concerning mutations. *Viruses*, 13(3), 392”, (**Alpha**), poświęcona była studiom dynamiki epidemii wariantu wirusa SARS-CoV-2, nazywanego / oznaczanego jako alfa, lub też innymi określeniami / nazwami, PANGO B.1.1.7, albo VOC (wariant budzący obawy, variant of concern). Wariant ten zaobserwowano po raz pierwszy pod koniec roku 2020 w Wielkiej Brytanii. W następnym roku, 2021, zaobserwowano jego szybkie rozprzestrzenienie się i wypieranie innych wariantów wirusa SARS-COVID w Wielkiej Brytanii oraz w Europie. Autorzy pracy podsumowali podstawy biologiczne funkcjonowania i agresywności wariantu alfa, oraz wykorzystali swoje doświadczenia w modelowaniu epidemii do oszacowań pozwalających na przewidywanie dalszego scenariusza rozwoju epidemii SARS-COVID z wariantem alfa. Na podstawie dostępnych danych obserwacyjnych oszacowano wartości przewagi replikacyjnej wariantu alfa (na około 2.0). Tezami tej publikacji oraz pracy doktorskiej (hipoteza 5) są znacząca przewaga replikacyjna wariantu alfa nad poprzednimi szczepami oraz duże prawdopodobieństwo dalszego rozprzestrzeniania się tego wariantu, a także możliwość pojawienia się w przyszłości nowych wariantów związana z kumulacją mutacji w sekwencjach wariantu alfa.

W pracy, „Grabowski, F., Kochańczyk, M., & Lipniacki, T. The spread of SARS-CoV-2 variant omicron with a doubling time of 2.0-3.3 days can be explained by immune evasion. *Viruses*. 2022; 14 (2): 294.”, (**Omicron**), analizowano rozprzestrzenianie się wariantu Omikron wirusa SARS-COVID i płynące z tej analizy wnioski. Sformułowano model matematyczny rozwoju epidemii SEIR, dla dwóch szczepów wirusa, kierując się empirycznymi dowodami, że wariant Omikron ma potencjał do zakażenia także osób już zakażonych poprzednimi wariantami. Oszacowano, że tygodniowy wzrost stosunku szczepu Omicron do szczepu Delta mieści się w zakresie 7,2–10,2, co jest wartością znacznie wyższą niż wzrost stosunku szczepu Delta do szczepu Alfa (szacowany na 2,5–4,2) oraz szczepu Alfa do szczepów wcześniej istniejących (szacowany na 1,8–2,7). Na podstawie zbudowanego modelu SEIR dla dwóch szczepów sformułowano hipotezę (jest to także hipoteza 6 pracy), że przewaga wzrostowa szczepu Omicron może wynikać z unikania odpowiedzi immunologicznej.

Praca, „Grabowski, F., Czyż, P., Kochańczyk, M., & Lipniacki, T. (2019). Limits to the rate of information transmission through the MAPK pathway. *Journal of the Royal Society Interface*, 16(152), 20180792.”, (**Limits**), badano przepustowość informacyjną ścieżki sygnałowej MAPK/ERK (Mitogen-Activated Protein Kinase / Extracellular Signal-Regulated Kinase). Badano transmisję informacji w szlaku MAPK/ERK, w dwóch wariantach stałej i pulsacyjnej stymulacji EGF (Epidermal Growth Factor) prowadzącej do generowania

impulsów aktywności ERK. W pracach wykorzystywano wcześniej opracowany w zespole badawczym model obliczeniowy szlaku MAPK/ERK. Model ten przewiduje pulsacyjne sygnały aktywności ERK co jest konsekwencją istnienia (hierarchicznie zbudowanych) dodatnich i ujemnych pętli sprzężenia zwrotnego. Z przeprowadzonych obliczeń / symulacji wynika, że gdy sygnał wejściowy składa się z sekwencji impulsów EGF, przesyłana informacja rośnie liniowo w czasie co prowadzi do wniosku, że kodowanie fazowe pozwala na przekazanie większej ilości informacji niż amplitudowe. Jest to także teza 7 pracy.

Praca, „Grabowski, F., Nałęcz-Jawecki, P., & Lipniacki, T. (2023). Predictive power of non-identifiable models. *Scientific Reports*, 13(1), 11143”, (**Predictive**), najpierw podkreśla rozróżnienie pomiędzy nieidentyfikowalnością strukturalną oraz praktyczną. Praca koncentruje się na tej drugiej. Istniejące podejścia do praktycznej nieidentyfikowalności polegają albo na wykonaniu dodatkowych eksperymentów pozwalających na uzyskanie rozmiaru danych wystarczających do wiarygodnych ocen parametrów albo na redukcji modelu, którą autorzy nazywają niealgorytmiczną. Słuszną oceną sformułowaną w pracy jest, że oba te podejścia mają wiele wad. Jako alternatywę proponuje się podejście bayesowskie powiązane z iteracjami MCMC (Markov Chain Monte Carlo). Jako przykłady zastosowania używa się dwa modele, kaskady sygnałowej z czterema czynnikami oraz układu mechanicznego zbudowanego z trzech mas i dwóch sprężyn.

Praca, „Czyż, P., Grabowski, F., Vogt, J., Beerenwinkel, N., & Marx, A. (2023). Beyond normal: On the evaluation of mutual information estimators. *Advances in neural information processing systems*, 36, 16957-16990”, (**Beyond**), poświęcona jest konstrukcji homeomorfizmów / dyffeomorfizmów do transformacji par zmiennych losowych w celu uzyskania różnorodnych rozkładów prawdopodobieństwa, o znanej informacji wzajemnej, (mutual information, MI). Wykorzystano transformacje wielowymiarowych rozkładów Gaussa i Studenta zachowujące informację wzajemną MI, do zbudowania benchmarków do testowania różnych estymatorów informacji wzajemnej. Na bazie zdobytych doświadczeń omawia się ogólną stosowalność i ograniczenia estymatorów klasycznych i neuronowych w środowiskach obejmujących dużą liczbę wymiarów, rzadkie interakcje, rozkłady długoogonowe i dużą ilość informacji wzajemnej. W pracy przedstawia się także zalecenia dotyczące wyboru odpowiedniego estymatora, dostosowanego do stopnia trudności rozpatrywanego problemu i kwestii, które należy wziąć pod uwagę, stosując estymator do nowego zbioru danych. Tezą pracy związaną z tą publikacją jest, że zaproponowana konstrukcja rozkładów prawdopodobieństwa par zmiennych losowych jest wystarczająca dla wiarygodnych porównań pomiędzy różnymi estymatorami MI.

W pracy, „Czyż, P., Grabowski, F., Vogt, J. E., Beerenwinkel, N., & Marx, A. (2025). On the properties and estimation of pointwise mutual information profiles. *Transactions on Machine Learning Research*”, (**Mixtures**), wykorzystuje się profile punktowej informacji wzajemnej oraz mieszaniny rozkładów, aby poszerzyć benchmark skonstruowany w poprzedniej publikacji. Wprowadza się także bayesowski estymator informacji wzajemnej.

Punktowy profil informacji wzajemnej, jest zdefiniowany jako rozkład punktowej informacji wzajemnej dla danej pary zmiennych losowych. Jego wartość oczekiwana jest wzajemną informacją między tymi zmiennymi losowymi. W artykule proponuje się także operacje Bend i Mix, dyffeomorfizmów realizacji zmiennych losowych oraz tworzenia mieszanin zmiennych losowych. Pokazuje się, że operacja mieszania rozkładów może prowadzić zarówno do tworzenia jak i anihilacji informacji wzajemnej. Jednak można znaleźć oszacowania na górne i dolne ograniczenia tej wartości. Co jeszcze ważniejsze, w pracy wykazuje się że dla mieszanin można zbudować stochastyczny algorytm MCMC, który z dowolną dokładnością pozwala na szacowanie wartości informacji wzajemnej. Proponowane metody wykorzystuje się do badania ograniczeń istniejących estymatorów informacji wzajemnej, badania neuronowych estymatorów wariacyjnych oraz zrozumienia wpływu obserwacji odstających na estymację informacji wzajemnej. Wprowadzona technika Bend and Mix znacznie poszerza możliwości algorytmów z poprzedniej publikacji (Beyond), co jest także tezą 10 pracy.

Ocena rozprawy

Recenzowana rozprawa jest oryginalnym dziełem naukowym. Publikacje naukowe wchodzące w jej skład są dobrze dobrane tematycznie, stanowią spójny zbiór studiów obejmujący wiele aspektów modelowania matematycznego i budowania systemów obliczeniowych użytecznych w badaniach rozprzestrzeniania się infekcji wirusowych. Ukazały się w renomowanych czasopismach naukowych. Są także już doceniane w środowisku naukowym. Wiele z nich było już wielokrotnie cytawne.

Wszystkie prace są wieloautorskie. Na podstawie deklaracji współautorskich oraz dość szerokiego omówienia treści publikacji w początkowej części pracy nie ma jednak wątpliwości co do bardzo istotnego wkładu Doktoranta w powstanie publikacji oraz decydującego wkładu w aspekty modelowania matematycznego oraz zastosowanych technik obliczeniowych. Większość artykułów (1-8) wchodzących w skład rozprawy ma walor interdyscyplinarny. Przedstawienie wyników publikacji w doktoracie pozwala także stwierdzić, że praca Doktoranta sprostowała interdyscyplinarnemu aspektowi wszystkich badań. Jego opisy i uwagi nie są jedynie wycinkowe. Widać, że Doktorant docenia wagę powiązania ze sobą wyników / faktów biologicznych, wiedzy epidemiologicznej oraz modelowania matematycznego i programowania dla rozwinięcia użytecznych technik i uzyskania nowych wyników naukowych. Praca jest ładnie i przejrzysto zredagowana.

Warsztat naukowy Doktoranta jest szeroki i różnorodny. Doktorant potrafi dopasować metody i modele matematyczne do różnego typu danych. Ma także kompetencje w zakresie budowania nowatorskim modeli statystycznych oraz ich weryfikowania.

Praca dokumentuje umiejętność Doktoranta posługiwania się warsztatem matematycznym i obliczeniowym obejmującym opracowanie i implementacje symulacji stochastycznych według algorytmu Gillespiego, symulacji stochastycznych w wersjach

agentowych, opracowywania różnych wersji modeli dynamiki ewolucji infekcji wirusowych, optymalizacji stochastycznej oraz implementacji potoków analizy danych biomedycznych różnych typów. Opracowanie stabilnych implementacji modeli, ich przetestowanie i udostępnienie dowodzi dysponowania przez Doktoranta dobrym warsztatem badawczym w zakresie technik informatycznych.

Powyższe uwagi pozwalają na bardzo pozytywną ocenę recenzowanej pracy.

Uwagi i pytania

Teza 3 pracy, zamiast tak jak obecnie powinna raczej brzmieć: „Wprowadzenie modyfikacji do modelu SEIR, polegającej wprowadzeniu założenia, że okres inkubacji jest zgodny z rozkładem Erlanga pozwala na urealnienie przewidywań tego modelu w aspekcie znanych obserwacji, a także na badanie zależności między dynamiką epidemii a zachowaniami populacji”.

Do pracy **Antagonism**: Jak uzasadnić konieczność zastosowania modelu/podejścia agentowego w symulacjach prowadzonych na potrzeby tej publikacji. Czy nie wystarczyłoby zastosowanie algorytmu Gillespiego z listą możliwych zdarzeń i odpowiednio zdefiniowanymi stanami?

Czy modele SEIR rozwijane w pracy w różnych wersjach („Dynamics”, „Super-spreaders”, „Alpha”, „Omikron”) mogłyby mieć zastosowanie do badania infekcji innymi wirusami (HCV, HPV, HIV)? Czy dostępne są jakieś dane eksperymentalne do takich badań?

Do pracy **Predictive**: Proponowana idea podejścia bayesowskiego z iteracjami MCMC jest zilustrowana na stosunkowo wąskim zbiorze przykładów. O ile dobrze rozumiem bazuje jedynie na danych symulowanych. Nasuwa się na przykład takie pytanie: Czy proponowane podejście można zastosować do modeli SEIR z prac „Dynamics”, „Super-spreaders”, „Alpha”, „Omikron”? Modele te zawierają rozkład czasu inkubacji, który jest nieznan, nie wiadomo ile ma parametrów. Zastosowano podejście, które zgodnie z nomenklaturą pracy „Predictive” powinno się chyba nazwać niealgorytmiczną redukcją (suma założonej liczby rozkładów wykładniczych o jednakowych czasach oczekiwania). Czy zastosowanie podejścia bayesowskiego z iteracjami MCMC byłoby tu użyteczne?

Do pracy **Beyond** oraz **Mixtures**: Byłoby interesujące i chyba bardzo pożyteczne aby badania dotyczące wpływu wymiarowości przestrzeni na dokładność oceny MI miały bardziej wyczerpujący oraz ciągły charakter. W praktycznych sytuacjach zawsze projektant systemu stoi przed decyzją jak duży powinien być zbiór danych. Potrzebna wielkość zbioru danych zależy od wymiarowości. Czy Doktorant mógłby bliżej opisać jaka była motywacja do stosowanych konkretnych postaci funkcji f (Mixtures, rys 5, ‘critic’)?

Czy wyniki prac **Beyond** oraz **Mixtures** dałoby się jakoś zastosować w pracy **Limits**?

Konkluzja

Praca stanowi podsumowanie oryginalnych, interdyscyplinarnych badań naukowych, w których sformułowano i weryfikowano oryginalne hipotezy badawcze. Dokumentuje wiedzę i kompetencje Doktoranta.

Osiągnięcia i oryginalne elementy pracy są na pewno wystarczające do jej ogólnej bardzo dobrej oceny. Stwierdzam, że rozprawa spełnia warunki stawiane pracom doktorskim i wnioskuję o jej dopuszczenie do publicznej obrony.

Ponadto, biorąc pod uwagę bardzo szeroki zakres badań, bardzo ładny, jasny i staranny styl pracy, wartościowe publikacje Doktoranta zawierające oryginalne wyniki naukowe, zarówno wchodzące w skład dysertacji jak też inne, wnioskuję o wyróżnienie pracy.

