

Paweł Paszek, PhD, dr hab. inż.
Division of Infection, Immunity and Respiratory Medicine
Lydia Becker Institute of Immunology
Faculty of Biology, Medicine and Health
The University of Manchester
Michael Smith Building
Oxford Road
M13 9PT
Manchester

| | |
|---------------------------|---------------|
| Biurowo Dzielakana | |
| Wpłynęło dnia 14.08.2019. | Nr 976 / zez. |
| RAU | |

Recenzja rozprawy doktorskiej mgr inż. Magdaleny Ochab

„Zastosowanie metodyki układów z przełączeniami do opisu i analizy układów biologicznych”

pod kierunkiem dr hab. inż. Krzysztofa Puszyńskiego, prof. nzw. Politechniki Śląskiej

Cell i charakterystyka pracy

Praca doktorska przedstawia możliwości zastosowania układów z przełączeniami do modelowania i analizy procesów biologicznych. Układy z przełączeniami stanowią klasę układów znaną z teorii automatyki, działających na zasadzie skokowych zmian parametrów i struktury modelu. Autorka *de facto* proponuje, że układy z przełączeniami mogą stanowić pożyteczne narzędzie do opisu układów biologicznych, swoistą alternatywę dla powszechnie stosowanych silnie nieliniowych modeli matematycznych.

Głównym osiągnięciem pracy jest zaproponowanie kompleksowych metod analizy układów z przełączeniami. Zaproponowane metody bazują na znanej w literaturze teorii układów kawałkami liniowych, jak również autorskich algorytmów symulacji i analizy heterogeniczności populacji. W rozprawie Autorka skupia się na analizie wewnątrzkomórkowych procesów regulacyjnych związanych z ekspresją genów oraz ścieżką sygnalizacji czynnika transkrypcyjnego p53. Metodyka oparta o układy z przełączeniami zilustrowana na powyższych przykładach pozwala na lepsze poznanie ich właściwości oraz sformułowanie nowych predykcji.

Wyniki bezpośrednio związane z pracą doktorską zostały opublikowane w jednym czasopiśmie figurującym na liście filadelfijskiej (Ochab *et al.* Biomedical Engineering Online 2017), pięciu artykułach w recenzowanych materiałach konferencyjnych, indeksowanych w *Web of Science* (w tym mgr inż. Ochab jest pierwszym autorem czterech publikacji) oraz około dziesięciu artykułach recenzowanych w innych materiałach konferencyjnych lub monografiach. Materiał zawarty w pracy doktorskiej jednocześnie stanowi podsumowanie tych publikacji, jak i w pewnych obszarach wykracza poza ich zakres.

Praca składa się z abstraktu (w języku angielskim) oraz sześciu rozdziałów: zwięzłego wprowadzenia określającego cele pracy, wprowadzenia do układów z przełączeniami w kontekście biologicznym, trzech rozdziałów z wynikami oraz podsumowania. Główne wyniki pracy zawierają:

1. Przedstawienie wybranej metodologii analizy układów z przełączeniami (Rozdział 3). Autorka przedstawia kompleksowy schemat analizy układów z przełączeniami, których podukłady są liniowe (PLDE). Na podstawie literatury, Autorka wyprowadza ogólne rozwiązania analityczne dla przełączeń w czasie (w przypadku, gdy znana jest sekwencja przełączeń) oraz opisuje stosowanie grafu tranzycji w celu przedstawiania przejść między

podukładami. W kolejnej części, Autorka wprowadza metody analizy punktów stacjonarnych na podstawie Mestla *et al.* (dla prostych układów z produkcją i degradacją). Autorka proponuje rozszerzenie powyższej metody dla układów bardziej złożonych (gdzie zmienne stanu mogą być zależne od zmiennych przełączających, a także bezpośrednio od innych zmiennych czasu) na podstawie macierzy Jacobiego oraz logoid-Jacobiego. Efektem tych rozważań jest ogólny algorytm poszukiwania punktów stacjonarnych w układach PLDE. Następnie, Autorka skupia się na analizie heterogeniczności populacji w układach z przełączeniami, gdzie proponuje dwa algorytmy; (i) uwzględnienie losowości wartości progowych na podstawie rozkładu normalnego; (ii) uwzględnienie losowości przełączeń przy użyciu funkcji Hilla zależnej od stanu układu. Wprowadzona metodologia zilustrowana jest na prostych i czytelnych przykładach.

2. Opis i analiza prostych układów regulacji białka (Rozdział 4). Autorka rozważa klasyczne modele produkcji białka opisujące transkrypcję genu, translację mRNA oraz degradację. Modele uwzględniają dodatnie i ujemne sprzężenie zwrotne oraz terapeutyczną interwencję w celu obniżenia poziomu białka. W modelu istnieją dwa rodzaje przełączeń; aktywacja/dezaktywacja genu zależy od poziomu białka, natomiast podanie leku (zwiększające skokowo degradację białka) jest przełączeniem zależnym od czasu.

Względna prostota rozważanych układów pozwala na szczegółowe zaprezentowanie metodologii oraz dokładnego przebadania własności układu. Autorka wyprowadza rozwiązania analityczne, a następnie wyznacza punkty stacjonarne (zwyčajne oraz osobliwe). W układzie ze sprzężeniem dodatnim, dla pośredniego zakresu leku współistnieją dwa zwyčajne punkty stacjonarne. W układzie ze sprzężeniem ujemnym jest jeden zwyčajny oraz jeden osobliwy punkt stacjonarny, który powoduje pojawienie się oscylacji w przebiegach czasowych wokół wartości progowych. W obu układach można osiągnąć cel terapeutyczny, przy czym dla układu ze sprzężeniem ujemnym jest on osiągnięty dla niższego stężenia leku.

W kolejnej części Autorka przeprowadza symulacje heterogeniczności populacji za pomocą losowego proggu aktywacji genu. W układzie ze sprzężeniem dodatnim powstają subpopulacje (wynikające z dwóch punktów stacjonarnych), natomiast w układzie ze sprzężeniem ujemnym dochodzi do zróżnicowania wartości białka w stanie ustalonym. Losowy czas podania leku wydłuża czas odpowiedzi, szczególnie w modelu ze sprzężeniem dodatnim. W ostatniej części Autorka analizuje modele z losowym przełączaniem, zakładając, że prawdopodobieństwo aktywacji zależy od stanu (ilości białka) opisanej funkcją Hilla. Różna „wrażliwość” komórki modelowana jest poprzez wartości parametru Hilla, n . Dla dużych wartości n , e.g. 5 symulacje heterogeniczności populacji są bardzo zbliżone do rozwiązania deterministycznego, natomiast dla mniejszych wartości dochodzi do rosnących fluktuacji poziomu białka.

Opis i analiza modułu regulatorowego p53 (Rozdział 5). Autorka proponuje uproszczony model regulacji białka p53 w konwencji PLDE. Model uwzględnia dwa sprzężenia zwrotne; (i) ujemne, gdzie p53 aktywuje białko MDM2, które następnie degraduje p53 oraz (ii) dodatnie, gdzie p53 aktywuje białko PTEN, które blokuje działanie białka MDM2. Wymuszenie R odpowiadające uszkodzeniu DNA (stresowi komórkowemu). W modelu uwzględniono 4 wartości progowe zależne od poziomów białek w układzie. Podstawowe symulacje uzyskanego modelu odwzorowują zachowanie się analogicznego silnie nieliniowego modelu nieliniowego, tj. oscylacje p53 dla średniego poziomu stresu (areszt cyklu komórkowego) oraz wysoki stan ustalony dla dużego stresu (odpowiadający apoptozie). Co ciekawe, model z przełączeniami produkuje bardziej realistyczne wartości zmiennych w układzie, aczkolwiek w samych przebiegach widać momenty przełączania.

Autorka wyznacza rozwiązania analityczne w poszczególnych domenach oraz analizuje punkty stacjonarne modelu na podstawie metodologii wprowadzonej w Rozdziale 3. Dla tak skomplikowanego układu jest to trudne zadanie, a występują w nim zarówno zwyczajne punkty stacjonarne, jak i osobliwe punkty stacjonarne, wokół których mogą występować oscylacje. Autorka pokazuje wykresy bifurkacyjne; w układzie występują stabilne oscylacje dla średniego poziomu stresu R , natomiast dla dużego R trajektorie stabilizują się lub oscylują w zależności od bliskości do różnych punktów stacjonarnych. Ciekawą predykcją wynikającą z modelu są różne zakresy oscylacji dla średniego poziomu stresu, które wynikają z różnych poziomów białka PTEN.

W kolejnej części Autorka bada efekty heterogeniczności populacji w przedstawionym modelu. Założenie rozkładu normalnego dla wartości progowych (aktywacji produkcji MDM2 lub PTEN) prowadzi do tworzenia się subpopulacji (odpowiadającym różnym punktom ustalonym), powstawaniu oscylacji o różnych (>2) zakresach (w zależności od poziomu MDM2) oraz różnicowania się czasu apoptozy. Zróżnicowanie wartości progowej transportu MDM2, wpływa raczej na czas podjęcia decyzji apoptotycznej, natomiast zróżnicowanie progu degradacji p53 powoduje tworzenie się subpopulacji. Zróżnicowanie aktywacji wszystkich procesów na raz prowadzi do subpopulacji dla wszystkich zakresów stresu, np. dla dużych wartości R występują wszystkie rodzaje odpowiedzi.

Na końcu następuje losowa zmiana przełączeń na podstawie funkcji Hilla, analizowana dla różnych wartości współczynnika Hilla (od 5 do 30). Dla mniejszych współczynników Hilla (a więc większej losowości) obecne są fluktuacje poziomów zmiennych w układzie, które prowadzi do wyraźnego tłumienia oscylacji dla średnich i małych zakresów R . Dla dużych współczynników Hilla trajektorie przypominają symulacje deterministyczne bądź wydłuża się czas oczekiwania do apoptozy (dla dużego poziomu stresu).

Bez wątpienia przedstawiona analiza modułu p53 pozwala na lepsze zrozumienie działania tego układu, jak również pokazuje skuteczność zaproponowanej metody.

Uwagi ogólne, krytyczne i dyskusyjne

Jak pokazuje przedstawiona w pracy analiza źródeł, obecne metody analizy procesów biologicznych opierają się o zastosowanie silnie nieliniowych równań różniczkowych lub symulacji stochastycznych odpowiednich reakcji biochemicznych. Wadami tych metod są trudności związane z otrzymaniem rozwiązań analitycznych, skomplikowaną (bądź tylko lokalną) analizę stabilności czy też duże ograniczenia obliczeniowe w przypadku symulacji stochastycznych. W pracy tej, Autorka bardzo przekonująco wykazuje, że zastosowanie układów z przełączeniami może stanowić nie tylko pewną alternatywę, ale bardzo skuteczne narzędzie matematyczne do analizy dynamicznych procesów biologicznych. Zaproponowany schemat analizy układów wychodzi poza obecny stan wiedzy w dziedzinie i bez wątpienia stanowi istotny wkład w rozwój narzędzi inżynierii biologicznej. Przedstawione przykłady, a w szczególności analiza modułu p53 w istotny sposób przyczyni się do zrozumienia podstaw działania tego ważnego układu biologicznego. Bez wątpienia są to wyniki, które powinny być opublikowane w rozsądnych czasopismach w dziedzinie inżynierii medycznej.

Praca jest starannie przygotowana oraz dobrze przemyślana. Wyniki pracy są oryginalne i zaprezentowane w przystępny sposób. Pojawiają się pewne niedociągnięcia w opisie, np. abstrakt zawiera błędy językowe, strona 49: j21 zamiast l21, l32 zamiast j32 w opisie; Rys. 45. nie jest jasny czas włączania leku na podstawie skali czasowych, ale nie umniejszają one wartości pracy. Interesujące prace zawsze prowokują pytania, najważniejsze są załączone poniżej.

- 1) *Sformułowanie prostych modeli produkcji białka*: Autorka rozważa proste modele produkcji białka, jednakże sformułowanie problemu odbiega od klasycznego opisu. Tzn. model zakłada dwa allele genu, z których jeden jest włączony na stałe, natomiast drugi allel podlega przełączeniom w zależności od sprzężenia zwrotnego. Wydaje mi się, że taki typ regulacji jest raczej mało spotykany, z reguły allele jednego genu zachowują się dosyć podobnie. Niejasności dodaje opis odnoszący się do regulacji dwóch genów, zamiast jednego genu o dwóch allelach. Proszę o lepszą motywację zaproponowanego modelu. Przedstawiony problem można sformułować przy pomocy konstytutywnej produkcji oraz dwóch alleli regulowanych przez sprzężenia.
- 2) *Analiza stabilności*: Brakuje dyskusji o ogólnym zastosowaniu metod analizy stabilności na podstawie macierzy logoid-Jacobiego. Metoda zaproponowana przez Autorkę jest dosyć heurystyczna (równania 3.31-3.38), macierz logoid-Jacobiego nie zawsze uwzględnia wszystkie zależności (patrz prosty model regulacji białka, strona 82), natomiast analityczne punkty stacjonarne nie są zawsze potwierdzone numerycznie (układ p53, strona 128). Z czego wynikają te rozbieżności i na ile ograniczają one zaproponowaną metodę?
- 3) *Analiza heterogeniczności progów*: Autorka proponuje dwa algorytmy, pierwszy opierający się na losowaniu wartości progowych, drugi na zastosowaniu rozkładu prawdopodobieństwa zależnego od stanu poprzez nieliniową funkcję Hilla. Pierwsza metoda jest bardzo intuicyjna, tj. losowość związana jest z poziomami białek regulacyjnych, etc. i prowadzi do heterogeniczności poziomów jak również tworzenia się subpopulacji wynikających z różnych stanów ustalonych. Np. w module regulacji p53 powoduje to powstawanie oscylacji o różnych zakresach oraz współistnienia nawet 3 stanów dla jednej wartości wymuszenia. Można by się zastanowić, czy zamiast rozkładu normalnego nie zastosować rozkładów z długim ogonem (np. logarytmicznie normalnego), które są bliższe parametrom komórkowym. Również, brakuje porównania symulacji progów do losowych zmian innych parametrów modelu, np. temp produkcji i degradacji, stosowanych w przypadkach klasycznych układów nieliniowych. Wydaje się, że w tym przypadku efekt będzie podobny. Można też się zastanowić czy badany układ z przełączeniem jest bardziej wrażliwy na wartości progowe czy na zmiany innych parametrów modelu. Należy również zastanowić się w jakim stopniu progi aktywacji są parametrami, które można identyfikować na podstawie danych biologicznych lub mierzyć eksperymentalnie. Jest to ważne dla ogólnej przydatności takiej kasy modeli.
- 4) *Analiza heterogeniczności populacji poprzez funkcje Hilla*: Zastosowany schemat symulacji jest modyfikacją schematu Gillespiego (a dokładniej z symulacji hybrydowych, gdzie część zmiennych układu jest deterministyczna) zakładającą funkcje Hilla. W przypadku prostego modelu produkcji białka, dla dużego n (czyli dużego małego „szumu”) widzimy rozwiązania bliskie deterministycznym, natomiast dla małych n występują fluktuacje poziomów. Nie jest do końca jasne (przynajmniej dla mnie) czy w układzie z dodatnim sprzężeniem widać subpopulacje odpowiadające bistabilności. W modelu p53 mamy raczej do czynienia z tłumieniem oscylacji (dla mniejszych n), w skrajnych przypadkach widzimy różne stany ustalone (Rys. 5.28, $n=30$, mały „szum”). Można by to interpretować, że losowość (tzw. *intrinsic noise*) ma raczej mały wpływ na przebiegi, i generalnie prowadzi do tłumienia trajektorii. Z drugiej strony, to duży szum powinien wzbudzać oscylacje dla niższego poziomu R (w porównaniu z modelem deterministycznym) lub ułatwiać „wchodzenie” w różne stany ustalone - co nie jest jasne na podstawie przedstawionych przebiegów. Dodatkowo, w niektórych przykładach pojawiają się duże różnice między rozwiązaniami deterministycznymi a średnimi z symulacji stochastycznych (Rys. 4.13). Raczej tak nie powinno być, szczególnie w prostym układzie i może świadczyć, że funkcja Hilla zmienia

charakterystyki modelu. Wreszcie, współczynnik Hilla, który Autorka definiuje jako „wrażliwość” komórek jest dosyć trudny do zinterpretowania biologicznego i zapewne do identyfikacji na podstawie danych. Czy w związku z przedstawionymi wynikami, zróżnicowanie odpowiedzi p53 obserwowane eksperymentalnie wynika raczej z losowości progów (tzw. *extrinsic noise*), niż stochastyczności przełączeń (tzw. *intrinsic noise*)? Wydaje się, że interpretacja metody wymaga dodatkowej dyskusji.

- 5) *Różne odpowiedzi w module p53*: Z perspektywy biologicznej jest to najciekawsza część pracy. W układach z losowymi wartościami progowymi, model przewiduje różne zakresy oscylacji dla heterogenicznej populacji (np. Rys. 5.9B, 5.13A oraz 5.15B). Może to świadczyć o istnieniu dodatkowych/nowych stanów w układzie (oprócz aresztu cyklu komórkowego oraz apoptozy). Z drugiej strony Chen et al. niejednoznacznie interpretuje oscylacje vs. wysoki poziom p53- oba przebiegi wydają się prowadzić do tych samych losów komórek. Nie jest jasne na podstawie symulacji czy różne zakresy oscylacji charakteryzują się różnymi częstotliwościami, tzn. czy na tej podstawie mogą być zidentyfikowane eksperymentalnie. (O ile pokazane są przebiegi oraz amplitudy oscylacji, analiza częstotliwości nie jest pokazana- mimo kilku odwołań w treści rozprawy). Czy w obecnych danych eksperymentalnych przy użyciu mikroskopii przyżyciowej można zauważyć podobne zachowania. Z drugiej strony, czy różne zakresy oscylacji obecne są w modelu nieliniowym (5.13-5.16), czy raczej wynikają z założeń o ilości progów w układzie?

Wniosek końcowy

Przedstawiona rozprawa doktorska wnosi istotny wkład w rozwój narzędzi do analizy procesów biologicznych, jak również w wiedzę na temat procesów regulacji genów oraz modułu p53. Praca świadczy o bardzo dobrej znajomości problemu przez doktorantkę oraz w pełni realizuje przedstawione cele. W związku z tym wnoszę o dopuszczenie mgr inż. Magdaleny Ochab do publicznej obrony. Przy satysfakcjonującym odniesieniu się do powyższych uwag oraz przebiegu obrony, wnioskuję o wyróżnienie pracy.



Paweł Paszek

Manchester, 7 Sierpnia 2019