

Marcin RADOM^{1,2}, Piotr FORMANOWICZ^{1,2}

¹Instytut Informatyki, Politechnika Poznańska

²Instytut Chemii Bioorganicznej, Polska Akademia Nauk

SYMULACJA STOCHASTYCZNYCH SIECI PETRIEGO I JEJ ZASTOSOWANIE W ANALIZIE SYSTEMÓW BIOLOGICZNYCH*

Streszczenie. Procesy zachodzące w systemach biologicznych są ze swojej natury stochastyczne. Na poziomie molekularnym zajście konkretnej reakcji chemicznej jest zależne od tego, jak blisko znajdują się obok siebie jej wszystkie substraty. Tak dokładne modelowanie nawet prostego systemu biologicznego, tj. określenie pozycji w trójwymiarowej przestrzeni każdej cząsteczki będącej potencjalnym substratem reakcji jest niemal niemożliwe z uwagi na ich liczbę nawet w najprostszym systemie. Dlatego też pozycje cząsteczek zastępuje się często ich stężeniem w modelach badanego systemu. Taką możliwość oferują między innymi stochastyczne sieci Petriego, na bazie których możliwe jest przeprowadzenie wielu rodzajów symulacji uwzględniających różne czynniki opisujące dany system. W niniejszym artykule przedstawiony zostanie schemat symulacji stochastycznych oraz przykład jego implementacji w programie oferującym możliwość tworzenia i analizy różnych rodzajów sieci Petriego, w tym omawianych sieci stochastycznych.

SIMULATION OF STOCHASTIC PETRI NETS AND ITS APPLICATION IN BIOLOGICAL SYSTEMS ANALYSIS

Summary. Processes occurring in biological systems are stochastic in their nature. On the molecular level activation of some specific reaction depends on how close to each other are all of its substrates. Modeling a biological system with such a precision, i.e., tracking positions of all molecules in three dimensional space is still almost impossible due to their numbers in even a relatively simple system. Because of that, the positions of molecules are replaced by their concentration in a model of a biological system. Creation of such models and their simulation can be performed using stochastic Petri nets. In this paper description of a stochastic simulation has been given, along with its implementation in a software tool providing means to create and analyse models using different Petri nets, including the stochastic ones.

*Badania przeprowadzone w ramach realizacji projektu finansowanego ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji DEC-2012/07/B/ST6/01537.

1. Wstęp

Systemy biologiczne, a precyzyjniej procesy w nich zachodzące, są ze swojej natury stochastyczne. Konkretna reakcja na poziomie molekularnych zachodzi wtedy i tylko wtedy, gdy wszystkie jej substraty znajdują się odpowiednio blisko w przestrzeni, tak aby ich wzajemne oddziaływanie doprowadziło do zajścia tejże reakcji w systemie. Takie modelowanie systemu biologicznego, choćby dla jednej komórki, zależnie od stopnia dokładności modelu, może okazać się bardzo trudne lub wręcz niemożliwe. Symulacja takiego systemu wymagałaby ogromnej mocy obliczeniowej i pamięci, realistycznie rzecz ujmując również przekraczających możliwości większości współczesnych komputerów. Z tych powodów zamiast mało praktycznych prób dokładnego modelowania wszystkich elementów badanego systemu biologicznego w trójwymiarowej przestrzeni, stosuje się podejście, w którym nacisk w ramach tworzenia modelu położony jest na jak najdokładniejsze określenie stężenia (koncentracji) wszystkich ważnych związków chemicznych wchodzących w skład badanego systemu. Za praktycznością takiego podejścia stoi założenie mówiące, że prawdopodobieństwo zajścia konkretnej reakcji jest zależne od ilości jej substratów obecnych w danej chwili w systemie. Innymi słowy, im takich cząsteczek jest więcej w obserwowanym systemie biologicznym, tym większe jest prawdopodobieństwo, że część z nich znajdzie się na tyle blisko siebie aby pojawić się mogły oddziaływania prowadzące do zajścia danej reakcji chemicznej.

Jednym z coraz popularniejszych narzędzi służących do modelowania systemów biologicznych są sieci Petriego [9]. Mają one strukturę skierowanego grafu dwudzielnego, a modele budowane za ich pomocą posiadają wszystkie zalety modeli grafowych. Należy również pamiętać, że umożliwiają one także opisywanie dynamiki systemu za pomocą obiektów zwanych tokenami, reprezentujących konkretne ilości związków chemicznych. Ich przepływ w sieci odpowiada przepływowi przez modelowany system informacji, substancji, itp. W zależności od rodzaju użytej sieci Petriego w procesie tworzenia modelu (np. klasycznej, ciągłej, czasowej lub stochastycznej) możliwe są różne rodzaje analizy sieci oraz symulacji jej działania bazującej na dynamice przepływu tokenów.

W niniejszym artykule opisany zostanie sposób modelowania systemu biologicznego oraz jego symulacji z użyciem stochastycznych sieci Petriego. Ich symulacja bazuje na prawdopodobieństwie zachodzenia poszczególnych reakcji w systemie, prawdopodobieństwo to z kolei zależy między innymi od ilości poszczególnych składników systemu, tj. substratów każdej reakcji chemicznej zachodzącej w ramach badanego systemu biologicznego. Dodatkowo przedstawiony zostanie przykładowy moduł symulacyjny zaimplementowany w narzędziu o nazwie Holmes [10], służącym do tworzenia, analizy i symulacji modeli biologicznych opartych o różne rodzaje sieci Petriego, w tym sieci stochastyczne. Układ pracy jest następujący. W rozdziale drugim przedstawione są najbardziej podstawowe pojęcia związane z klasycznymi sieciami Petriego oraz definicja sieci stochastycznej. W rozdziale trzecim omówiony został jeden z podstawowych algorytmów symulacji sieci stochastycznych. W rozdziale czwartym zostanie krótko przedstawiony przykład użycia modułu symulacyjnego programu Holmes. Praca kończy się krótkim podsumowaniem zawartym w rozdziale piątym.

2. Klasyczne i stochastyczne sieci Petriego

Zanim przedstawiona zostanie definicja sieci stochastycznej, należy rozpocząć od zdefiniowania klasycznej sieci Petriego jako bazy dla dalszych rozważań. Następnie, w minimalnym lecz niezbędnym zakresie, wprowadzone zostaną najważniejsze pojęcia związane z analizą sieci klasycznych. O wiele bardziej szczegółowe definicje i zagadnienia można odnaleźć m. in. w [6, 8].

2.1. Klasyczne sieci Petriego

Klasyczna sieć Petriego może zostać zdefiniowana w następujący sposób.

Definicja 1. *Sieć Petriego* [8]

Sieć Petriego jest zbiorem $N = \{P, T, F, W, m_0\}$, gdzie:

$P = \{p_1, p_2, \dots, p_n\}$ jest skończonym zbiorem miejsc,

$T = \{t_1, t_2, \dots, t_m\}$ jest skończonym zbiorem tranzycji,

$F \subseteq (P \times T) \cup (T \times P)$ jest zbiorem łuków,

$W : F \rightarrow \{1, 2, 3, \dots\}$ jest funkcją przypisującą wagi łukom,

$m_0 : P \rightarrow \{0, 1, 2, 3, \dots\}$ określa początkowy stan sieci,

$P \cap T = \emptyset$ oraz $P \cup T \neq \emptyset$.

Strukturą sieci Petriego jest dwudzielny graf skierowany, w którym występują połączone łukami miejsca i tranzycje, miejsca zaś mogą przechowywać *tokens*. W przypadku modelowania systemu biologicznego tranzycje są najczęściej odpowiednikami elementarnych lub bardziej złożonych reakcji chemicznych w nim zachodzących, miejsca natomiast reprezentują substraty i produkty (mogą one również reprezentować pewne stany systemu). Liczba tokenów w danym miejscu oznacza pewną liczbę jednostek danego składnika (np. substratu, produktu). W opisach sieci Petriego często stosuje się określenie miejsca/tranzycje wejściowe/wyjściowe. Mówimy, że p_i jest miejscem wejściowym tranzycji t_j jeżeli $(p_i, t_j) \in F$ (p_i jest miejscem wyjściowym dla t_j jeżeli $(t_j, p_i) \in F$). Zbiór miejsc wejściowych (wyjściowych) tranzycji t_j oznacza się jako $\bullet t_j$ ($t_j \bullet$). Analogicznie określa się tranzycje wejściowe i wyjściowe względem danego miejsca p_i (ich zbiory oznacza się odpowiednio jako $\bullet p_i$ oraz $p_i \bullet$).

Zmiany w rozkładzie tokenów w miejscach sieci regulowane są poprzez zasadę aktywacji i uruchomienia tranzycji, a rozkład tokenów w danej chwili nazywamy *oznakowaniem sieci* (odpowiada ono stanowi modelowanego systemu). Aktywna tranzycja to taka, dla której liczba tokenów we wszystkich jej miejscach wejściowych jest równa lub większa niż odpowiednie wagi łuków prowadzących do niej z jej miejsc wejściowych. Tranzycja aktywna w danym stanie sieci może, choć nie musi zostać uruchomiona. Uruchomienie tranzycji powoduje zmianę rozkładu tokenów w sieci. Z miejsc wejściowych tranzycji pobierana jest liczba tokenów równa wagom łuków łączących te miejsca z daną tranzycją. We wszystkich miejscach wyjściowych tranzycji pojawiają się nowe tokeny, których liczba również równa jest wagom łuków łączących daną tranzycję z tymi miejscami. Należy w tym miejscu zaznaczyć, że o ile zasada aktywacji tranzycji oraz efekty ich uruchomienia (tj. zmiana rozkładu tokenów) są takie same dla klasycznych sieci Petriego oraz dla sieci stochastycznych, to w przypadku tych drugich określenie, która tranzycja ma być uruchomiona jako następna podlega bardziej zaawansowanym zasadom i zostanie dokładniej opisane dalej.

Sieć Petriego może być reprezentowana przez macierz incydencji. W takiej macierzy $C = [c_{ij}]_{n \times m}$ dla sieci Petriego posiadającej n miejsc i m tranzycji, każda komórka $c_{i,j}$ określa zmianę liczby tokenów w miejscu p_i w wyniku uruchomienia tranzycji t_j . t -niezmiennik jest nieujemnym wektorem $x \in \mathbb{N}^m$ spełniającym równanie $C \cdot x = 0$.

Wsparciem t -niezmiennika x nazywamy zbiór tranzycji $\text{supp}(x) = \{t_j : x_j > 0, j = 1, 2, \dots, m\}$. Uruchomienie wszystkich takich tranzycji odpowiednią liczbę razy (x_j dla tranzycji $t_j \in \text{supp}(x)$) nie zmienia stanu sieci. Sieć jest *pokryta przez t -niezmienniki*, jeżeli każda tranzycja należy do wsparcia przynajmniej jednego t -niezmiennika. W sieci będącej modelem systemu biologicznego t -niezmienniki odpowiadają więc pewnym podstawowym procesom biologicznym tego systemu i często są obiektem dalszej analizy. Przykłady tego typu analizy dla systemu biologicznego znaleźć można między innymi w [2, 3].

2.2. Stochastyczne sieci Petriego

Z badaniami złożonych procesów biologicznych nieodłącznie związane jest zagadnienie prawdopodobieństwa zachodzenia poszczególnych reakcji chemicznych. W związku z tym metody modelowania oparte na sieciach Petriego, w których możliwe jest uwzględnienie takiego prawdopodobieństwa wydają się być bardziej odpowiednie do opisu systemów biologicznych. Sieci stochastyczne doczekały się wielu rozwinięć, w których między innymi definiowane są nowe rodzaje tranzycji [4]. Niniejszy artykuł koncentruje się na podstawowym rodzaju sieci stochastycznej, opisanej Definicją 2 (por. [7, 8]).

Definicja 2. Stochastyczna sieć Petriego

Stochastyczna sieć Petriego jest zbiorem $SPN = \{N, v\}$, gdzie:

$N = \{P, T, F, W, m_0\}$ jest klasyczną siecią Petriego,

$v = (\lambda_1, \lambda_2, \dots, \lambda_m)$ jest wektorem wartości tzw. częstości uruchamiania, przypisanych do odpowiednich tranzycji, gdzie m jest liczbą tranzycji.

Można powiedzieć, że stochastyczna sieć Petriego składa się ze wszystkich elementów sieci klasycznej oraz dodatkowego wektora o liczebności zbioru tranzycji, który dla każdej z nich określa pewną wartość liczbową mającą wpływ na prawdopodobieństwo uruchomienia aktywnej tranzycji. Jak dokładnie wartość ta wpływa na omawiane prawdopodobieństwo zostanie szczegółowo opisane w następnym rozdziale. W ogólności, dla każdej tranzycji obliczane jest tzw. opóźnienie uruchomienia, które określa czas, jaki musi upłynąć do momentu jej uruchomienia. Wartość tego opóźnienia jest zmienną losową, której rozkład określony jest poprzez funkcję gęstości prawdopodobieństwa, której parametrem jest wartość λ_j przypisana do tranzycji t_j [7].

Zagadnienia związane z symulacją zostaną przedstawione w kolejnym rozdziale. Należy tutaj po raz ostatni zaznaczyć, że dla sieci stochastycznych również możliwa jest analiza oparta na niezmiennikach, która została wspomniana w poprzednim podrozdziale. Sieci SPN umożliwiają jednak bardziej precyzyjną symulację zachowań modeli biologicznych niż ma to miejsce w przypadku modeli opartych np. na klasycznych sieciach Petriego.

3. Prawdopodobieństwo uruchomienia tranzycji w sieci stochastycznej

Opisany zostanie teraz jeden z popularnych trybów symulacji sieci SPN. Na wstępie należy zauważyć, że sieci stochastyczne są pewnym specyficznym rodzajem sieci czasowych, tj. po upływie pewnego czasu aktywna tranzycja będzie w nich uruchomiona. Celem jest określenie prawdopodobieństwa uruchomienia tranzycji, co przekłada się na fakt, że im większe jest to prawdopodobieństwo, tym szybciej dana tranzycja zostanie uruchomiona. Czas ten będzie tu nazywany czasem oczekiwania i oznaczany dalej jako τ . Czas oczekiwania dla tranzycji t_j jest zmienną losową X_{t_j} z wykładniczym rozkładem prawdopodobieństwa, dla której funkcja gęstości prawdopodobieństwa dana jest poniższym wzorem [1]:

$$f_{X_{t_j}}(\tau) = \lambda_j \cdot e^{-\lambda_j \cdot \tau} \quad (1)$$

gdzie dla tranzycji t_j , wartość $\lambda_j > 0$ określa jej częstość uruchomienia (por. Def. 2.). Przy takim sformułowaniu funkcji im większa będzie wartość λ_j tym mniejszą wartość przyjmować będzie funkcja $f_{X_{t_j}}(\tau)$. Ponieważ, jak zostało już napisane na wstępie niniejszego rozdziału, wartość ta jest traktowana jako czas opóźnienia uruchomienia tranzycji (τ), im jest on mniejszy tym szybciej (innymi słowy z większym prawdopodobieństwem) dana tranzycja zostanie uruchomiona.

Opisana funkcja gęstości prawdopodobieństwa nie uwzględnia aktualnego stanu sieci, jedynie wartość λ_j przypisaną do odpowiedniej tranzycji. W rzeczywistości biologicznej występuje jednak dodatkowy czynnik mający wpływ na częstość zachodzenia poszczególnych reakcji chemicznych, mianowicie chodzi tutaj o ilość poszczególnych molekuł/substratów dostępnych w danym momencie w systemie. Wpływ ten znany jest jako *prawo działania mas* (inaczej prawo Guldberga i Waagego, [11]). Prawo to można wyrazić stwierdzeniem, że szybkość zachodzenia reakcji chemicznej jest proporcjonalna do stężenia wszystkich uczestniczących w niej substratów. Jego respektowanie w ramach symulacji sieci SPN będącej modelem systemu biologicznego pozwala bardziej precyzyjnie odwzorowywać zachowanie się badanego systemu w rzeczywistości biologicznej.

Wzór funkcji gęstości prawdopodobieństwa określający czas uruchomienia tranzycji w sieci SPN może być rozszerzony tak, aby uwzględniać przytoczone tu prawo działania mas, co w stochastycznej sieci Petriego oznacza uwzględnienie wpływu liczby tokenów w miejscach wejściowych danej tranzycji na jej prawdopodobieństwo uruchomienia jako następnej w ramach symulacji. W takim przypadku czas uruchomienia tranzycji t_j jest zmienną losową X_{t_j} , dla której funkcja gęstości prawdopodobieństwa ma następującą, nową postać [1]:

$$f_{X_{t_j}}(\tau) = \lambda_j \cdot q(t_j) \cdot e^{-\lambda_j \cdot q(t_j) \cdot \tau}. \quad (2)$$

Tranzycja t_j uruchomi się z pewnym prawdopodobieństwem po upływie czasu τ , gdzie $\lambda_j > 0$ jest częstością uruchomień, $q(t_j)$ określa maksymalną możliwą liczbę uruchomień t_j ograniczoną przez aktualną liczbę tokenów w jej miejscach wejściowych. Wzór na obliczenie $q(t_j)$ ma następującą postać:

$$q(t_j) = \min_{\forall p_i \in \bullet t_j} \lfloor p_i / w_i \rfloor \quad (3)$$

gdzie w_i jest wagą łuku (p_i, t_j) .

4. Wybrane narzędzia informatyczne dla sieci SPN

Opisany sposób symulacji został zaimplementowany w wielu programach (np. Snoopy [5], MonaLisa [1], Holmes [10]) służących do tworzenia modeli systemów biologicznych przy użyciu sieci Petriego. W dalszej części niniejszego rozdziału zostaną dokładnie przedstawione możliwości oferowane przede wszystkim przez program Holmes. Na wstępie jednak zostaną teraz krótko opisane podstawowe i często wspólne możliwości oferowane użytkownikowi przez trzy wymienione programy. Wszystkie trzy narzędzia są darmowe, umożliwiają tworzenie, edycję i analizę wielu rodzajów sieci Petriego, nie tylko sieci klasycznych czy stochastycznych, ale także sieci czasowych (Snoopy, Holmes), funkcyjnych (Holmes) czy kolorowanych (Snoopy). Programy wspierają różne formaty zapisu sieci Petriego, programy MonaLisa i Holmes potrafią jednak odczytywać i zapisywać sieci między innymi w formacie natywnym dla programu Snoopy, który jest z tej grupy najstarszy, co zapewnia możliwość wymiany danych o sieci pomiędzy nimi bez potrzeby tworzenia modelu od nowa w edytorze każdego z omawianych programów. Wszystkie wymienione programy posiadają więcej niż jeden tryb symulacji, w każdym przypadku możliwa jest jednak symulacja oparta o schemat opisany w niniejszym artykule. W przypadku sieci klasycznych oraz takich, które poza dodatkowymi elementami posiadają wszystkie cechy określone w Definicji 1 (czego przykładem jest właśnie między innymi sieć stochastyczna) w programach MonaLisa i Holmes możliwa jest rozbudowana analiza oparta o niezmienniki miejsc i tranzycji, zbiory MCT czy podział t-niezmienników na t-klastry, której dokładniejszy opis wykracza poza ramy niniejszej pracy. W przypadku symulacji sieci stochastycznej podstawowym sposobem prezentacji wyników dostępnym w każdym z wymienionych programów są wykresy działania symulacji oddzielnie dla miejsc i tranzycji. Szczegóły dotyczące dalszych możliwości opisywanych programów (szczególnie programów Snoopy oraz MonaLisa) dostępne są w zacytowanych na początku niniejszego rozdziału publikacjach.

Program Holmes pozwala na tworzenie sieci stochastycznej zdefiniowanej wcześniej, należy jednak nadmienić, że możliwe jest w nim także tworzenie modeli opartych o rozszerzone sieci stochastyczne [4] zawierające pewne dodatkowe typy tranzycji (sieci takie dostępne są także w programie Snoopy). Zachowanie podstawowej sieci typu SPN w symulacji zależne jest zarówno od stanu początkowego jak i od częstości uruchomień tranzycji. Program Holmes umożliwia przechowywanie wielu wektorów stanów początkowych oraz wielu zbiorów v zawierających różne wartości λ_j . W ten sposób w ramach jednej struktury sieci SPN można mówić o kilku modelach na bazie sieci stochastycznej, których zachowanie w symulacji będzie się różnić. Domyślnie symulator sieci bierze pod uwagę prawo działania mas, można jednak z tej opcji w ramach symulacji zrezygnować. Standardową funkcją w tego typu oprogramowaniu jest możliwość określenia zarówno liczby kroków do wykonania w ramach jednej symulacji, jak i przeprowadzania wielu z nich przy tych samych warunkach początkowych, a następnie otrzymanie uśrednionych wyników zachowania się tranzycji oraz rozkładu tokenów w poszczególnych miejscach sieci. Opcja taka zapewnia zebranie bardziej wiarygodnych danych o zachowaniu się badanego systemu.

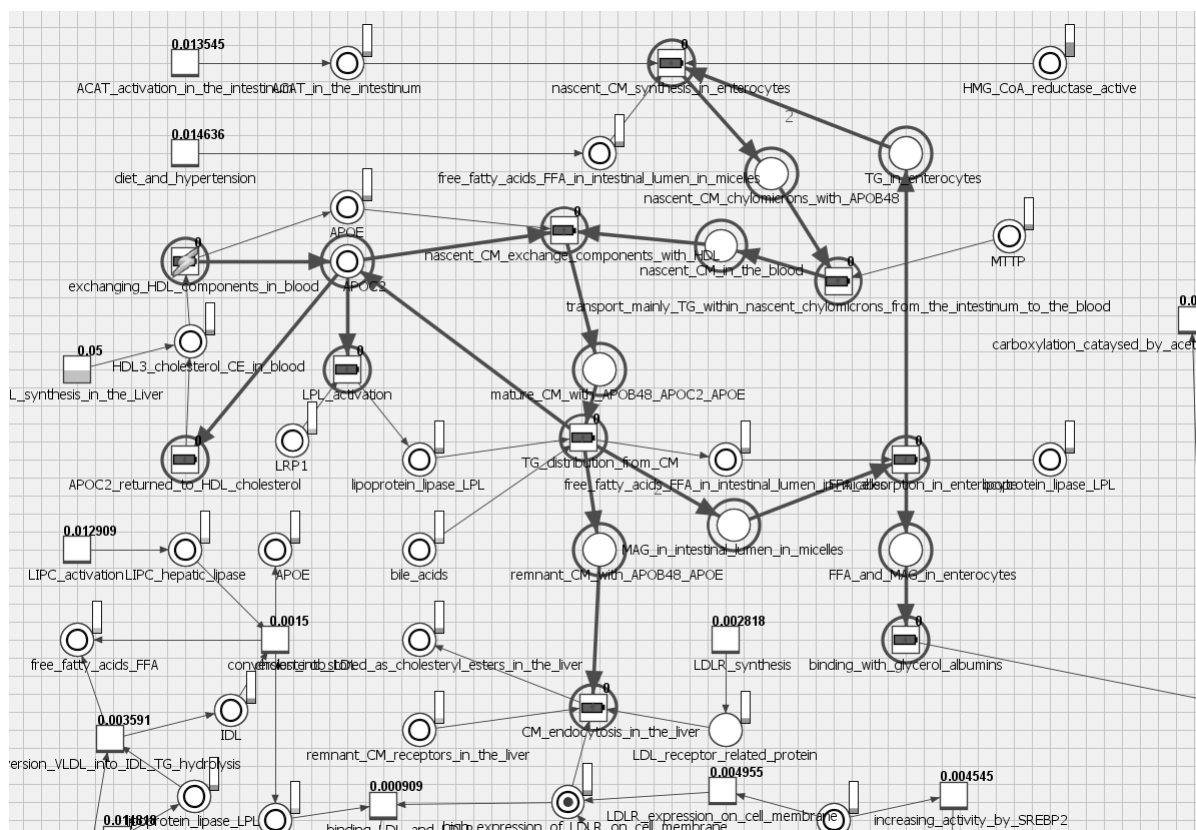
Symulacja zachowania się sieci stochastycznej przy konkretnych warunkach początkowych nie wyczerpuje listy możliwych opcji dostępnych w analizie sieci SPN. Interesującą dla tego typu analizy opcją jest tzw. *knockout symulacyjny*. W tym trybie

pewna reakcja lub zbiór reakcji (tranzycji) może być wskazana jako trwale nieaktywna. Oznacza to, że w czasie trwania symulacji tranzycja taka nie będzie nigdy uruchamiana, niezależnie od liczby tokenów gromadzących się w jej miejscach wejściowych. Dzięki temu, w przypadku badania zachowania się modelu systemu biologicznego możliwa jest analiza wpływu takiego wymuszonego wyłączenia na inne elementy systemu. Przykładem może być tutaj model choroby, oddający interakcje patogenu z elementami odpowiedzialnymi za prawidłowe działanie komórki. Knockout, czyli wyłączenie np. pewnych reakcji komórkowych, pozwala w ramach symulacji określić warunki prowadzące do zatrzymania progresji choroby oraz ich wpływ na inne elementy systemu biologicznego. Należy zauważyć, że analiza tego typu, tj. uwzględniająca trwałe wyłączenie pewnego podzbioru tranzycji, jest możliwa do stosowania w przypadku wielu rodzajów sieci Petriego, nie tylko stochastycznych. Szczególnie interesujące wydaje się jej połączenie z zasadami wynikającymi ze stosowania prawa działania mas. Wyłączona tranzycja nie tylko nie produkuje tokenów, ale także nie konsumuje tokenów ze swoich miejsc wejściowych, które gromadzą się w nich i mogą być potencjalnie wykorzystywane przez inne tranzycje. Co jednak jest tutaj bardziej istotne, gromadzące się tokeny (lub odwrotnie: ich mniejsza liczba) mają konkretny wpływ na prawdopodobieństwo (czas oczekiwania) uruchomienia aktywnych tranzycji. Należy tutaj uwypuklić różnicę pomiędzy knockoutem w sieci klasycznej oraz SPN, czyli umożliwiającej zastosowanie w symulacji danego modelu prawa działania mas. Wyłączenie jednej lub kilku tranzycji ma oczywiście wpływ na liczbę produkowanych tokenów. Jeśli liczba ta w pewnych miejscach spadnie do zera, pewne inne tranzycje przestają się uruchamiać. Jeśli jednak w sieci klasycznej liczba tokenów zmieni się, jednak wciąż pozostanie niezerowa, tranzycje których miejsc wejściowych taka sytuacja dotyczy będą uruchamiać się tak jak poprzednio. Jeszcze innymi słowy nie ma znaczenia, czy aktywna tranzycja ma jeden lub więcej tokenów w swoich miejscach wejściowych dla faktu jej uruchomienia. Inaczej ma się to w sieciach stochastycznych, które dzięki zastosowaniu wspomnianego prawa umożliwiają symulację lepiej oddającą rzeczywistość biologiczną którą modelują.

Tego typu symulacja w sieci SPN wykorzystująca mechanizm knockoutu, dostępna w programie Holmes, pozwala uzyskać bardziej precyzyjną wiedzę o zachowaniu się modelowanego systemu biologicznego niż ma to miejsce w przypadku sieci klasycznych. Program Holmes gromadzi szczegółowe dane z każdej wykonywanej symulacji, pozwala też na automatyzację ich przeprowadzania, np. wskazaną liczbę razy, osobno dla przypadku wyłączania każdej z tranzycji sieci. Zebrane dane to między innymi dokładna oraz średnia liczba gromadzonych tokenów w miejscach a także częstości uruchamiania każdej tranzycji w wybranych przedziałach czasu. Dane te mogą być reprezentowane zarówno w formie plików tekstowych jak i plików .csv (gdyby zaistniała konieczność dalszej analizy statystycznej za pomocą innych narzędzi), w specjalnie do tego dostosowanych oknach z tabelami w programie oraz w formie wykresów. W ten sposób łatwiej analizować można zależności pomiędzy prawdopodobieństwem zachodzenia poszczególnych reakcji reprezentowanych przez tranzycje, a także liczbą tokenów gromadzących się w poszczególnych miejscach sieci. Tego typu wizualizacja może bardzo pomóc w dogłębnej analizie zachowania się modelu systemu biologicznego.

Jedną z unikalnych możliwości programu Holmes w stosunku do innych narzędzi jest możliwość wizualizacji wyników symulacji np. bezpośrednio na obrazie sieci. Rysunek 1 pokazuje fragment tabeli przedstawiającej wpływ knockoutu poszczególnych

tranzycji na pozostałe tranzycje w pewnych obszarze większego modelu systemu biologicznego.



Rys. 1. Wizualizacja symulacji typu knockout na obrazie sieci. Tranzycja `exchanging_HDL_components_in_blood` została wyłączona, co pokazuje zarówno okrąg wokół niej, jak i symbol błyskawicy. Pozostałe tranzycje otoczone kręgiem przestały się w efekcie tego uruchamiać.

Na rysunku 1 liczby nad tranzycjami pokazują z jakim prawdopodobieństwem tranzycje były uruchamiane, np. liczba 0.05 oznacza 5% szans uruchomienia tranzycji w symulacji. Dodatkowo obok miejsc znajdują się pionowe, częściowo wypełnione kolorem paski, pokazujące średnią liczbę zgromadzonych tokenów w czasie symulacji. Należy tutaj nadmienić, że w programie wizualizacja wykorzystuje różne kolory opisanych pasków, czego nie daje się oddać na powyższym rysunku 1.

Rysunek 1 pokazuje efekty knockoutu symulacyjnego tylko dla jednej, wybranej tranzycji. Możliwe jest w programie Holmes, w sposób zautomatyzowany, przeprowadzenie bardzo dużej liczby symulacji porównującej efekty wyłączenia konkretnych tranzycji w sieci do przypadku, w którym nic nie podlega knockoutowi. Na początku, przy tych samych warunkach początkowych dla każdej symulacji wykonywana jest ich pewna określona liczba, przy założeniu, że każda tranzycja ma szansę na uruchomienie określone w sposób opisany wcześniej w niniejszym artykule (tj. brak knockoutu tranzycji). Użytkownik musi określić stan początkowy sieci, liczbę kroków symulacji oraz liczbę powtórzeń symulacji. Każda tranzycja w określonej liczbie kroków uruchamia się pewną liczbę razy, wartości te są uśredniane po liczbie symulacji i zapamiętywane do późniejszego porównania - tworzony jest więc tak zwany *zbiór referencyjny* danych sy-

mulacji. Następnie schemat ten powtarza się, przy czym w każdym kolejnym przypadku wyłączana jest jedna tranzycja w sieci (poprzez mechanizm knockoutu), a następnie dla każdego takiego wyłączenia przeprowadzana jest pewna określona wcześniej przez użytkownika liczba symulacji w celu późniejszego uśrednienia danych. Tego typu kolejne symulacje są już przeprowadzane automatycznie, dla tych samych warunków początkowych, które określił użytkownik na początku eksperymentu. Program Holmes gromadzi i uśrednia dane dotyczące tego, w jakim stopniu wyłączenie danej tranzycji wpływa na uruchamianie pozostałych. Różnice procentowe w liczbie uruchomień każdej tranzycji pomiędzy scenariuszem bez knockoutu a knockoutem każdej tranzycji z osobna (lub dla tranzycji wskazanych wcześniej przez użytkownika) są przedstawiane w formie tabeli.

Na sam koniec należy wspomnieć o dwóch ważnych cechach tego trybu w programie Holmes. Po pierwsze, dane są również zbierane dla miejsc sieci, tj. porównywana jest wtedy średnia liczba tokenów w poszczególnych miejscach w czasie symulacji. Po drugie, przeprowadzenie tego typu symulacji połączone ze zbieraniem wyników dla sieci z dużą liczbą tranzycji może być czasochłonne, jednak samo skonfigurowanie i rozpoczęcie obliczeń jest dla użytkownika bardzo proste, musi on mianowicie określić tylko stan początkowy sieci, wartości λ_j dla tranzycji sieci, liczbę kroków symulacji oraz liczbę powtórzeń symulacji (dzięki czemu możliwe jest uśrednienie wyników, co pozwala uzyskać bardziej wiarygodne dane o zachowaniu się symulowanego modelu systemu biologicznego).

5. Podsumowanie

Sieci stochastyczne są narzędziem modelowania systemów, które bardzo dobrze nadaje się do tworzenia modeli i analizy złożonych systemów biologicznych. Opisany tutaj schemat symulacji sieci SPN jest jednym z prostszych, choć równocześnie odpowiednio skutecznych podejść do symulowania zachowania się modelowanego systemu, zwłaszcza w porównaniu z ograniczonymi możliwościami w tej dziedzinie dostępnymi dla klasycznych sieci Petriego. W sieciach SPN możliwe jest uwzględnienie prawa działania mas, co jest szczególnie istotne w analizie modeli systemów biologicznych (w których np. liczba molekuł danego elementu systemu ma wpływ na szybkość przebiegania poszczególnych reakcji). Równie istotne jak sama symulacja jest odpowiednie gromadzenie i przetwarzanie jej wyników, zarówno dla poszczególnych miejsc jak i dla tranzycji. Opisane w ostatnim rozdziale niniejszej pracy narzędzie dedykowane modelowaniu przy użyciu sieci Petriego zawiera wiele przydatnych modułów pozwalających nie tylko na gromadzenie danych z wielu niezależnych symulacji, ale także ich wizualizację, która ułatwia dalszą interpretację i pełniejsze zrozumienie wyników badań. Same zaś sieci stochastyczne pozwalają na dokładniejsze badanie modeli systemów biologicznych w stosunku do sieci klasycznych.

LITERATURA

1. Balazki P., Lindauer K., Einloft J., Ackermann J., Koch I.: MONALISA for stochastic simulations of Petri net models of biochemical systems, *BMC Bioinformatics*, 16, 2015, p. 215-226.

2. Formanowicz D., Sackmann A., Kozak A., Błażewicz J., Formanowicz P.: Some aspects of the anemia of chronic disorders modeled and analyzed by petri net based approach. *Bioprocess and Biosystems Engineering*, 34, 2011, p. 581-595.
3. Formanowicz D., Radom M., Zawierucha P., Formanowicz P.: Petri net-based approach to modeling and analysis of selected aspects of the molecular regulation of angiogenesis, *PloS one*, 12(3), 2017, e0173020.
4. Heiner M., Lehrack S., Gilbert D., Marwan W.: Extended Stochastic Petri Nets for Model-Based Design of Wetlab Experiments, *Lecture Notes in Computer Science*, 5750, 2009, p. 138-163.
5. Heiner M., Herajy M., Liu D., Rohr C., Schwarick M.: Snoopy – a unifying Petri net tool, *Proc. PETRI NETS 2012, Lecture Notes in Computer Science*, Hamburg, Springer, 7347, 2012, p. 398–407.
6. Koch I., Reisig W., Schreiber F.: *Modeling in Systems Biology: The Petri Net Approach*, Springer-Verlag London, 2011.
7. Marsan M.A.: *Stochastic Petri nets: An elementary introduction*. *Lecture Notes in Computer Science*, 424, 1989, p. 1-29.
8. Murata T.: *Petri Nets: Properties, Analysis and Applications*, *Proceedings of the IEEE*, 77, 1989, p. 541-580.
9. Petri C.A.: *Communication with Automata* (in German), *Schriften des Instituts fur Instrumentelle Mathematik*, Bonn, 1962.
10. Radom M., Rybarczyk A., Szawulak B., Andrzejewski H., Chabelski P., Kozak A., Formanowicz P.: Holmes: a graphical tool for development, simulation and analysis of Petri net based models of complex biological systems, *Bioinformatics*, 33(23), 2017, p. 3822-3823.
11. Waage, P., Guldberg, C. M.: *Forhandlinger: Videnskabs-Selskabet i Christiana*, 35, 1864.