

Tomasz NYCZ

Politechnika Śląska, Instytut Informatyki

Tadeusz CZACHÓRSKI

Instytut Informatyki Teoretycznej i Stosowanej Polskiej Akademii Nauk

Politechnika Śląska, Instytut Informatyki

## **BADANIE SKALOWALNOŚCI MODELI SIECI KOMPUTEROWYCH WYKORZYSTUJĄCYCH APROKSYMACJĘ DYFUZYJNĄ WRAZ ZE ZWIĘKSZENIEM ROZMIARU MODELOWANEJ SIECI<sup>1</sup>**

**Streszczenie.** Zmienne natężenie ruchu generowanego przez aplikacje internetowe, np. multimedialne, a także przyjęta przez protokół TCP zasada regulacji natężenia transmisji w funkcji wykrytego obciążenia sieci, określanego na podstawie strat lub czasu transmisji, powodują, że natężenie transmisji internetowych jest permanentnie zmienne w czasie, a zmiany natężenia i dynamika tych zmian muszą być uwzględnione w ocenie algorytmów sterowania ruchem i unikania przeciążeń. Modelowanie struktur w Internecie wymaga z kolei dostosowania istniejących metod i modeli do analizy bardzo dużych konfiguracji sieciowych. Niniejsza praca bada pod kątem obliczeń numerycznych znany model matematyczny, umożliwiający analizę stanów nieustalonych w stanowisku obsługi i ich sieci, wykorzystujący metodę aproksymacji dyfuzyjnej. Zbadano wpływ poszczególnych parametrów modelu i rozmiaru sieci na całkowity czas obliczeń, a więc na skalowalność metody.

**Słowa kluczowe:** aproksymacja dyfuzyjna

## **SCALABILITY STUDY OF COMPUTER NETWORK MODELS USING A DIFFUSION APPROXIMATION WITH AN INCREASE IN THE SIZE OF THE MODELED NETWORK**

**Summary.** Varying intensity of traffic generated by Internet applications such as multimedia, together with the TCP protocol rules to regulate the intensity of the detected transmission as a function of network load which is determined on the basis of

---

<sup>1</sup> Praca współfinansowana ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego (nr umowy o dofinansowanie projektu: UDA-POKL.04.01.01-00-106/09).

loss or transmission time, cause that the intensity of Internet broadcasts is continuously variable in time, therefore intensity changes and dynamics of these changes must be included in the evaluation of algorithms for traffic control and congestion avoidance. Modeling of the structures on the Internet, in turn, requires adaptation of existing methods and models for the analysis of very large network configurations. This paper examines the numerical calculations for well-known mathematical model that allows analysis of transient states at the service station and their networks, using the diffusion approximation method. The influence of various parameters of the model and size of the network to the total computation time was investigated, and thus the scalability of the method.

**Keywords:** diffusion approximation

## 1. Wprowadzenie

Niniejsza praca bada możliwości dostosowania modelu sieci stanowisk obsługi, wykorzystującego metodę aproksymacji dyfuzyjnej do opisu bardzo dużych konfiguracji sieciowych, takich jakie spotyka się przy modelowaniu Internetu. Zaletą aproksymacji dyfuzyjnej jest możliwość opisu stanów nieustalonych kolejek – jest to cenne, ponieważ natężenie ruchu w sieciach telekomunikacyjnych jest zmienne, a momenty przeciążeń mogą powodować przepełnienie buforów w routerach i znacznie pogorszyć jakość usług transmisyjnych.

Aproksymacja dyfuzyjna jest metodą dokładniejszą niż stosowana często do opisu stanów nieustalonych w sieciach komunikacyjnych aproksymacja płynna (*fluid-flow approximation*), np. [8], wymaga jednak bardziej złożonych obliczeń i nie była, jak dotąd, stosowana do analizy sieci obejmujących więcej niż kilkadziesiąt węzłów, por. [2]. Inne teoretycznie możliwe podejścia, jak symulacja zdarzeń dyskretnych czy modele oparte na łańcuchach Markowa o bardzo dużej liczbie stanów, są jeszcze bardziej złożone obliczeniowo, por. [9].

Model stanowiska obsługi typu G/G/1/N, a więc, zgodnie z notacją Kendalla stanowiska z dowolnym rozkładem czasu pomiędzy nadchodzącymi klientami, dowolnym rozkładem czasu obsługi, z jednym kanałem obsługi i ograniczoną do  $N$  liczbą klientów w systemie, wykorzystujący aproksymację dyfuzyjną w rozpatrywanej tutaj postaci, zaproponował Gelenbe [3, 4], podając rozwiązanie modelu (rozkład liczby klientów w systemie) tylko dla stanu ustalonego, gdy prawdopodobieństwa stanów nie zależą od czasu. Rozwiązanie dla stanów nieustalonych zostało zaproponowane w pracy [1]. Model ten może być wykorzystywany do opisu kolejek pakietów w routerach IP, opisując dynamikę zmian tych kolejek przy nieustannie zmiennym natężeniu przychodzącego do routerów ruchu, którego charakter jest dodatkowo różny od strumienia Poissona, jak również do wyznaczania zmiennych w czasie prawdopodobieństw utraty pakietu w wyniku przepełnienia bufora – zadania nadchodzące w chwilach, gdy w stanowisku jest obecnych  $N$  klientów, są traczone.

W rozdziale 2 przedstawiono aproksymację dyfuzyjną jako metodę rozwiązywania modeli kolejkowych dla stanów nieustalonych, a w rozdziale 3 znajdują się uzyskane wyniki obliczeń w zależności od wybranych parametrów modelu i rozmiaru badanej sieci. Wszystkie obliczenia wykonano za pomocą specjalnego oprogramowania przygotowanego dla rozwiązywania – metodą zaproponowaną i opisaną w [1, 2] – modelu w postaci sieci stanowisk obsługi o dowolnej topologii. Oprogramowanie to umożliwia efektywne wykorzystanie metody i uzyskanie na jej podstawie wyników ilościowych, wymaga jednak odpowiedniego strojenia w postaci doboru parametrów – artykuł podsumowuje doświadczenia autorów w tej dziedzinie. W pracy [6] analizowano błąd metody aproksymacji dyfuzyjnej, wynikający z zastąpienia procesu dyskretnego – liczby klientów w stanowisku – ciągłym procesem dyfuzji; tutaj skoncentrowano się na zbadaniu błędów numerycznych związanych ze złożonymi obliczeniami modelu.

## 2. Aproksymacja dyfuzyjna stacji typu FIFO: model G/G/1/N

Poniżej jest omówiony klasyczny dyfuzyjny model stanowiska obsługi z jednym kanałem obsługi i kolejką. Niech  $A(x)$ ,  $B(x)$  określają rozkłady strumienia klientów przychodzących i czasu obsługi w stanowisku obsługi, a  $a(x)$  i  $b(x)$  będą ich funkcjami gęstości. Rozkłady są dowolne, ale nie sprecyzowane, metoda wymaga jedynie znajomości ich dwóch pierwszych momentów. Średnie rozkładów są określone jako  $E[A] = 1/\lambda$ ,  $E[B] = 1/\mu$ , a wariancje jako  $Var[A] = \sigma_A^2$ ,  $Var[B] = \sigma_B^2$ . Oznaczmy też kwadratowe współczynniki wariacji:  $C_A^2 = \sigma_A^2 \lambda^2$ ,  $C_B^2 = \sigma_B^2 \mu^2$ .  $N(t)$  przedstawia liczbę klientów obecną w systemie w momencie  $t$ . Dla pojedynczej kolejki typu FIFO zmiany  $N(t + \delta t) - N(t)$  mają w przybliżeniu rozkład normalny ze średnią  $(\lambda - \mu)\delta t$  i wariancją  $(\sigma_A^2 \lambda^3 + \sigma_B^2 \mu^3)\delta t$ , przy założeniu że czas  $\delta t$  jest wystarczająco długi i stacja działa bez przerwy. Aproksymacja dyfuzyjna zamienia proces  $N(t)$  na ciągły proces dyfuzji  $X(t)$ , którego infinitezymalne zmiany  $dX(t) = X(t + dt) - X(t)$  mają rozkład normalny o średniej  $\beta dt$  i wariancji  $\alpha dt$ , gdzie  $\beta$ ,  $\alpha$  są współczynnikami równania dyfuzji

$$\frac{\partial f(x, t; x_0)}{\partial t} = \frac{\alpha}{2} \frac{\partial^2 f(x, t; x_0)}{\partial x^2} - \beta \frac{\partial f(x, t; x_0)}{\partial x} \quad (1)$$

które określa funkcję gęstości prawdopodobieństwa procesu dyfuzji  $X(t)$ :

$$f(x, t; x_0) dx = P [x \leq X(t) < x + dx | X(0) = x_0].$$

Oba procesy  $X(t)$  i  $N(t)$  mają rozkład normalny zmian w czasie; wybór  $\beta = \lambda - \mu$ ,  $\alpha = \sigma_A^2 \lambda^3 + \sigma_B^2 \mu^3 = C_A^2 \lambda + C_B^2 \mu$  zapewnia, że parametry obu tych rozkładów rosną w tym samym tempie wraz ze wzrostem czasu obserwacji.

W [3] aproksymacja dyfuzyjna stacji G/G/1/N została przedstawiona jako proces  $X(t)$ , który jest określony na zamkniętym przedziale  $x \in [0, N]$ . Gdy proces dyfuzji dojdzie do  $x = 0$ , zachowuje tę wartość przez czas, który jest wielkością losową o rozkładzie wykładniczym z parametrem  $\lambda$ , a następnie powraca do  $x = 1$ . Czas, przez który proces pozostaje w  $x = 0$ , odpowiada okresowi beczynnemu stanowiska. Gdy dojdzie do  $x = N$ , zachowuje tę wartość przez czas, który jest wielkością losową o rozkładzie wykładniczym z parametrem  $\mu$ , a następnie powraca do  $x = N - 1$ . Czas, przez który proces pozostaje w  $x = N$ , odpowiada okresowi maksymalnego obciążenia stanowiska. Równanie dyfuzji jest uzupełnione przez równania równowagi dla  $p_0(t) = P[X(t) = 0]$ ,  $p_n(t) = P[X(t) = N]$  i ma postać

$$\begin{aligned}\frac{\partial f(x, t; x_0)}{\partial t} &= \frac{\alpha}{2} \frac{\partial^2 f(x, t; x_0)}{\partial x^2} - \beta \frac{\partial f(x, t; x_0)}{\partial x} + \lambda p_0(t) \delta(x-1) + \mu p_N(t) \delta(x-N+1) \\ \frac{dp_0(t)}{dt} &= \lim_{x \rightarrow 0} \left[ \frac{\alpha}{2} \frac{\partial f(x, t; x_0)}{\partial x} - \beta f(x, t; x_0) \right] - \lambda p_0(t) \\ \frac{dp_N(t)}{dt} &= -\lim_{x \rightarrow N} \left[ \frac{\alpha}{2} \frac{\partial f(x, t; x_0)}{\partial x} - \beta f(x, t; x_0) \right] - \mu p_N(t)\end{aligned}$$

Rozwiązanie, patrz [1], polega na rozpatrzeniu najpierw równania dyfuzji z barierami pochłaniającymi w  $x = 0$  i  $x = N$  za pomocą standardowych metod analitycznych i uzyskaniu funkcji gęstości prawdopodobieństwa  $\phi(x, t; x_0)$  tego procesu, a następnie wyrażeniu funkcji gęstości prawdopodobieństwa  $f(x, t; x_0)$  procesu dyfuzji z natychmiastowymi powrotami jako superpozycji funkcji  $\phi(x, t; x_0)$ :

$$f(x, t; x_0) = \phi(x, t; \psi) + \int_0^t g_1(\tau) \phi(x, t - \tau; 1) d\tau + \int_0^t g_{N-1}(\tau) \phi(x, t - \tau; N-1) d\tau$$

gdzie funkcje  $g_1(\tau)$ ,  $g_{N-1}(\tau)$  definiują intensywność rozpoczęcia nowych procesów (po skoku z bariery) w punktach  $x = 1$  i  $x = N - 1$ . Ich wartości są otrzymywane z układu równań równowagi dla przepływów prawdopodobieństw wchodzących i wychodzących z barier. Funkcję gęstości otrzymuje się analitycznie w postaci jej transformaty Laplace'a, której oryginał jest uzyskiwany numerycznie.

Powyższe rozwiązanie dla stanu nieustalonego jest uzyskane dla stałych parametrów. Aby wprowadzić  $\alpha(t)$ ,  $\beta(t)$  odzwierciedlające ewolucję strumieni wejściowych, oś czasu jest dzielona na małe przedziały, w których parametry są stałe, a rozwiązanie na końcu każdego przedziału daje warunek początkowy dla równania dyfuzji następnego przedziału o nowych parametrach. Czasami zachodzi konieczność wprowadzenia zależności parametrów dyfuzji od wartości samego procesu  $\alpha(x, t)$ ,  $\beta(x, t)$  – odzwierciedla to np. mechanizm kontroli reagujący na wielkość kolejki lub pozwala modelować serwery o wielu kanałach obsługi. W takim przypadku przedział dyfuzji  $x \in [0, N]$  jest podzielony na podprzedziały o określonej (np. jednostkowej) długości, w których parametry są stałe. Równania dla przedziałów przestrzeni

dyfuzji są rozwiązywane razem z równaniami bilansu równowagi dla przepływów prawdopodobieństwa pomiędzy sąsiadującymi przedziałami. Zasady tworzenia sieci stanowisk obsługi i obliczania zmiennych w czasie parametrów strumieni pomiędzy stacjami obsługi typu G/G/1/N są opisane np. w [2].

### 3. Wpływ parametrów modelu na czas obliczeń

Badanie wpływu parametrów na czas obliczeń zostało przeprowadzone dwustopniowo. W pierwszym kroku zbadano wpływ poszczególnych parametrów modelu dla pojedynczego stanowiska, a w drugim wpływ rozmiaru sieci (liczby węzłów) przy określonych parametrach wszystkich stanowisk na czas obliczeń.

Parametry modelu opartego na metodzie aproksymacji dyfuzyjnej, mające wpływ na czas obliczeń, można podzielić na dwie grupy. Pierwsza grupa to parametry ustawiane w pliku konfiguracyjnym modelu i należą do nich:

- całkowity czas, do którego odnosi się model *Time*,
- krok czasowy *Step*,
- maksymalny rozmiar kolejki *N*,
- liczba aktywnych kanałów obsługi *c*.

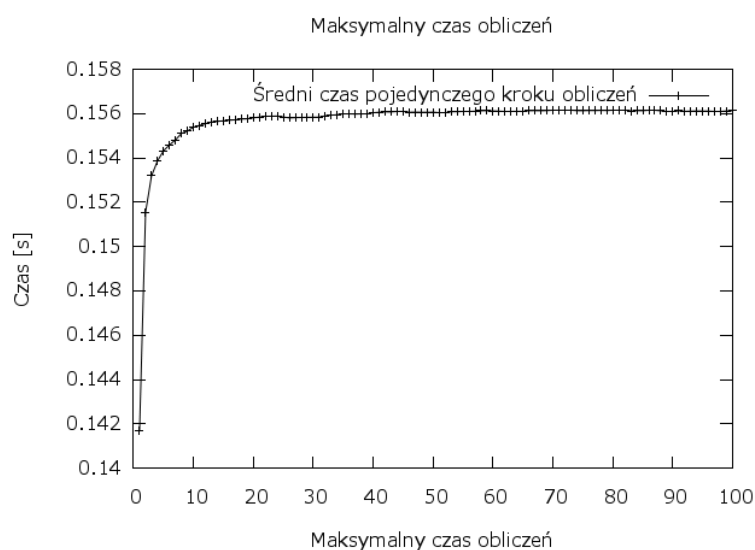
Druga grupa zawiera parametry, mające wpływ na czas i precyzję obliczeń, i są to parametry ustawione w aplikacji. Do tej grupy należą:

- precyzja algorytmu Stehfesta *N* [7], służącego do inwersji transformat Laplace'a, gdzie oryginał  $f(t)$  transformaty  $\bar{f}(s)$  oblicza się jako  $f(t) = \frac{\ln 2}{2} \sum_{i=1}^N V_i \bar{f}\left(\frac{\ln 2}{2} i\right)$ ,
- gęstość próbkowania całkowania numerycznego metodą Simpsona.

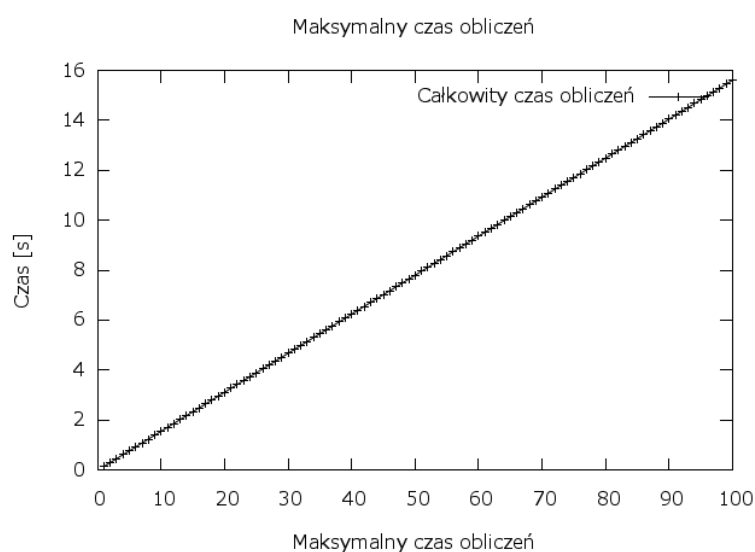
Analiza wpływu długości przedziału czasu, w którym jest rozpatrywana ewolucja modelu i kroku obliczeń na złożoność czasową metody, została przeprowadzona wspólnie, ponieważ oba te parametry mają jednakowy wpływ na czas obliczeń. Krok obliczeń oznacza, o ile zwiększa się upływ czasu metody zaczynając od 0 aż do momentu osiągnięcia przez metodę całkowitego czasu obliczeń. Obliczenia dla tego przypadku zostały przeprowadzone w następujący sposób: krok obliczeń został ustawiony na 1 sekundę, a zmieniał się maksymalny czas obliczeń od 1 sekundy do 100 sekund. Następnie zmierzono czasy dla kolejnych sekund, które następnie zostały uśrednione (przykładowo dla maksymalnego czasu 20 sekund zmierzono czasy obliczeń 0-1, 1-2, ..., 19-20 i uśredniono zebrane czasy poprzez podzielenie ich sumy przez 20), a także zmierzono całkowity czas obliczeń kolejno: 0-1, 0-2, 0-3, ..., 0-100. Maksymalna długość kolejki została ustawiona na  $N = 20$ , system ma jeden kanał obsługi, liczbę wyrazów  $n$  w teoretycznie nieskończonym szeregu obliczanym w algorytmie Stehfesta (okre-

śla to dokładność transformacji odwrotnej) na  $n = 18$ , a gęstość próbkowania całkowania numerycznego jest aktualnie uzależniona tylko od liczby aktywnych kanałów obsługi i w przypadku jednego kanału wynosi 2.

Rysunek 1 przedstawia średni czas obliczeń pojedynczego kroku uśredniony po liczbie kroków (zależnej od maksymalnego czasu obliczeń) i, jak widać, liczba kroków nie ma większego wpływu na czas trwania pojedynczego kroku. Rysunek 2 przedstawia całkowity czas obliczeń różniący się liczbą kroków wynikającą z większego maksymalnego czasu obliczeń (przyjęto stały krok obliczeń równy 1 sekundzie). Jak widać, wraz ze wzrostem liczby kroków mamy liniowy wzrost całkowitego czasu obliczeń.



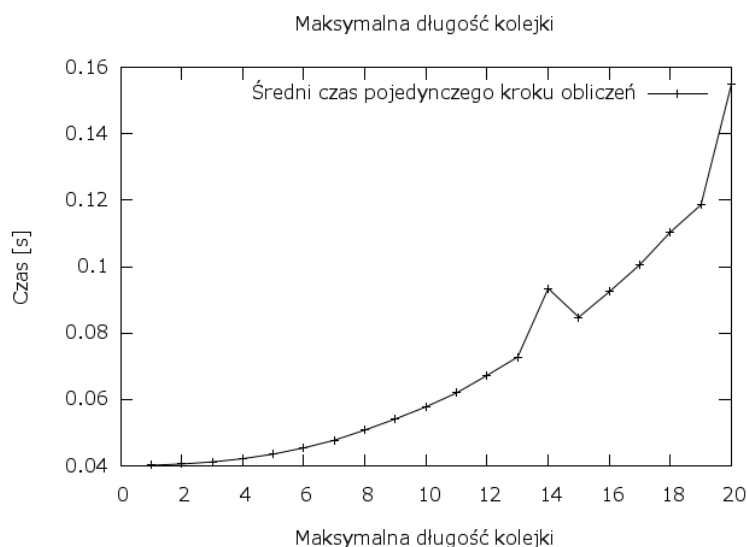
Rys. 1. Średni czas pojedynczego kroku obliczeń w funkcji maksymalnego czasu obliczeń  
Fig. 1. Average unitary calculations step as a function of total virtual calculations time



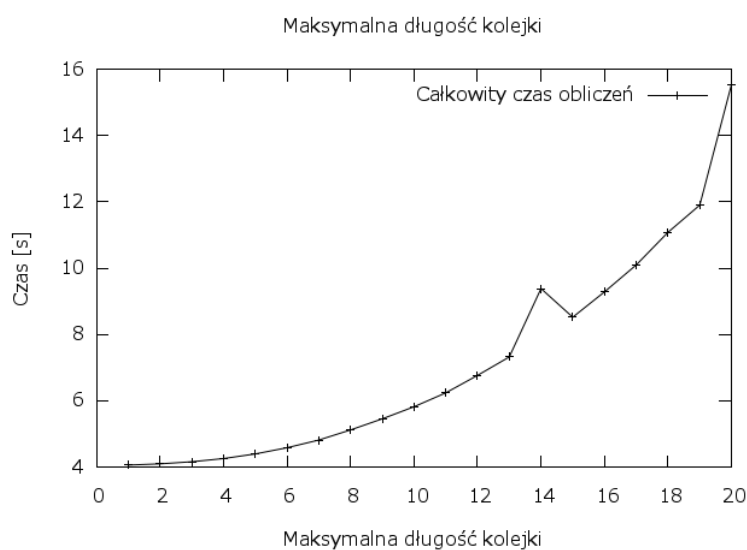
Rys. 2. Całkowity czas obliczeń w funkcji maksymalnego czasu obliczeń  
Fig. 2. Total calculations time as a function of total virtual calculations time

Kolejny parametr, który ma wpływ na złożoność czasową metody, to maksymalny rozmiar kolejki. Parametr ten został przebadany dla pojedynczego stanowiska o jednym kanale obsługi, precyzji algorytmu Stehfesta ustawionej na  $n=18$  i gęstości próbkowania całkowania numerycznego równej 2 (dla zachowania takich samych warunków obliczeń).

Jak widać na rysunkach 3 i 4, wraz ze wzrostem maksymalnej długości kolejki wzrasta średni i całkowity czas obliczeń. Wynika to ze wzrostu zakresu możliwych przepływów masy prawdopodobieństwa aktualnej liczby klientów w systemie, a więc wzrostu wielkości tablic je przechowujących.



Rys. 3. Średni czas pojedynczego kroku obliczeń w funkcji maksymalnej długości kolejki  
Fig. 3. Average unitary calculations step as a function of maximum queue length

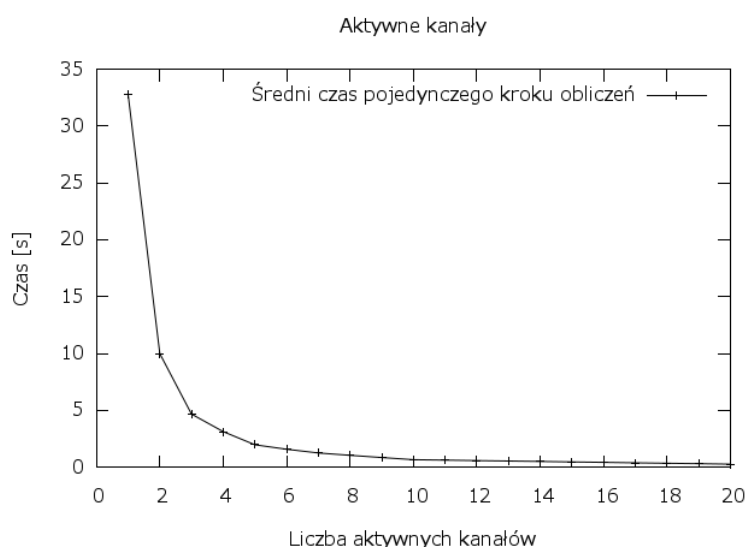


Rys. 4. Całkowity czas obliczeń w funkcji maksymalnej długości kolejki  
Fig. 4. Total calculations time as a function of maximum queue length

Kolejnym parametrem mającym wpływ na czas obliczeń jest liczba aktywnych kanałów obsługi, który może przyjmować wartości od 1 do maksymalnej długości kolejki. Parametr ten określa liczbę podprzedziałów metody (przepływ masy prawdopodobieństwa jest ograniczony do wielkości podprzedziałów i następuje zbilansowanie przepływów pomiędzy podprzedziałami). W efekcie gdy parametr ten ma wartość większą od 1, tablica zawierająca prawdopodobieństwa zmiany aktualnej liczby klientów w systemie z jednej wartości na drugą nie jest wypełniona w całości (wynika z ograniczeń podprzedziałami), a co za tym idzie czas obliczeń ulega skróceniu. Z drugiej strony, jeżeli podprzedział jest mały (w szczególności jednostkowy) i jeżeli gęstość próbkowania całkowania numerycznego jest mała, to funkcja aproksymująca aktualną liczbę klientów w systemie będzie miała małą dokładność, co negatywnie wpłynie na precyzję obliczeń. Aby przeciwdziałać temu zjawisku, aktualnie program przyjmuje dwie możliwe wartości gęstości próbkowania całkowania numerycznego uzależnione od liczby aktywnych kanałów obsługi:

- jeżeli jest jeden kanał obsługi, gęstość próbkowania wynosi 2 (przykładowo 0, 0.5, 1 dla przedziału 0-1),
- jeżeli jest więcej niż jeden kanał obsługi, gęstość próbkowania wynosi 16 (przykładowo 0, 0.0625, 0.125, 0.1875, ..., 0.9375, 1 dla przedziału 0-1).

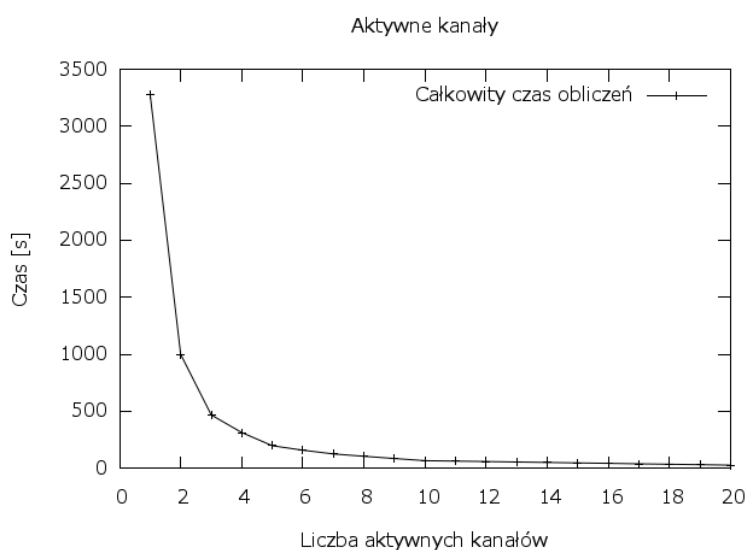
Ten negatywny problem dokładności obliczeń występuje również w przypadku bardzo małych maksymalnych długości kolejki (w szczególności  $N = 1$ ), ale ponieważ takie przypadki są znacznie rzadsze, to dla ich policzenia gęstość próbkowania jest ustawiana ręcznie w kodzie aplikacji. Aby zachować takie same warunki obliczeń (podobnie jak w przypadku badania wpływu maksymalnej długości kolejki), ustawiono taką samą wartość parametru gęstości próbkowania dla wszystkich przypadków równą 16.



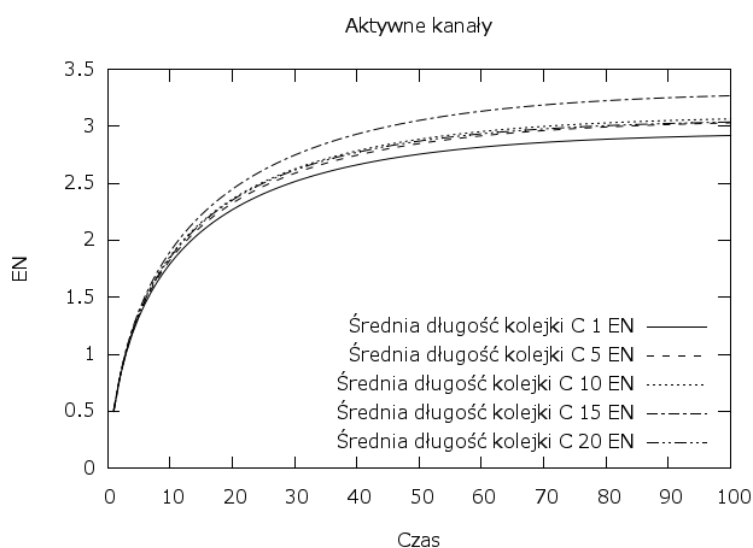
Rys. 5. Średni czas pojedynczego kroku obliczeń w funkcji liczby aktywnych kanałów obsługi  
 Fig. 5. Average unitary calculations step as a function of number of active channels



Jak widać na rysunkach 5 i 6, wzrost liczby kanałów obsługi przy stałej maksymalnej długości kolejki (w powyższych przykładach  $N = 20$ ) powoduje znaczne przyspieszenie obliczeń, wynikające z wypełniania tablicy prawdopodobieństw przejść tylko częściowo. Można również zauważyć, że przy dużej wartości parametru gęstości próbkowania numerycznego i jednym kanale obsługi obliczenia bardzo mocno się wydłużają, dlatego program w tym przypadku domyślnie korzysta z mniejszej wartości. Stosowanie podprzedziałów ma jednak wady. We wszystkich przedziałach parametry dyfuzji miały tę samą wartość, stąd można by oczekiwać, że średnia długość kolejki powinna być we wszystkich przypadkach taka sama. Niestety, tak nie jest, bilansowanie pomiędzy przedziałami wprowadza dodatkowy błąd, co ilustruje rysunek 7.



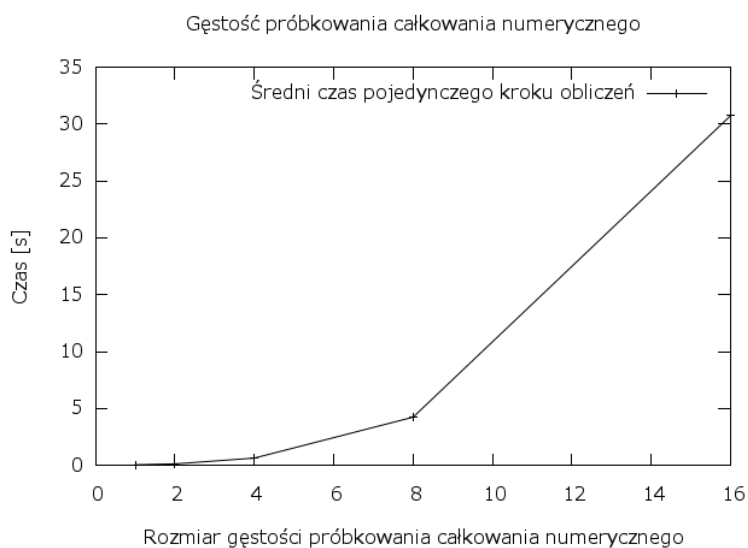
Rys. 6. Całkowity czas obliczeń w funkcji liczby aktywnych kanałów obsługi  
Fig. 6. Total calculations time as a function of number of active channels



Rys. 7. Średnia długość kolejki w czasie dla różnej liczby aktywnych kanałów obsługi  
Fig. 7. Mean queue length as a function of time for different number of active channels

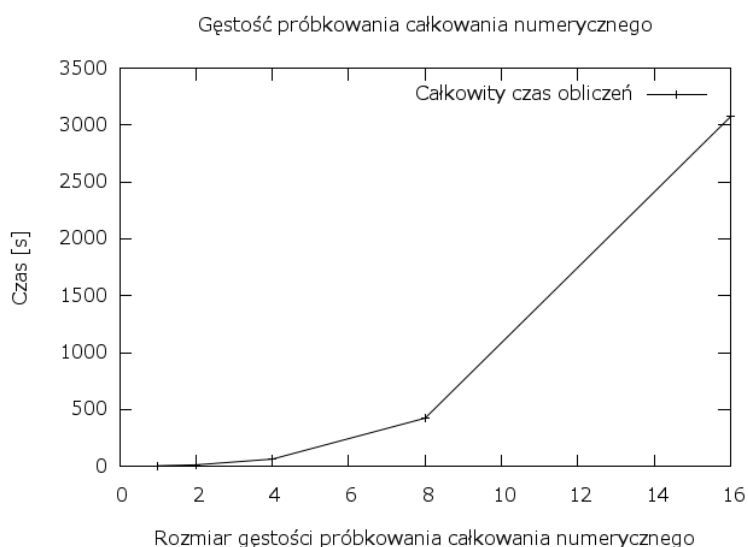
Jak widać, wzrost liczby aktywnych kanałów obsługi znacznie przyspiesza obliczenia, ale kosztem dokładności otrzymanych rezultatów.

Kolejny parametr wpływający na czas obliczeń to wspomniany już wcześniej parametr gęstości próbkowania całkowania numerycznego, który nie jest dostępny z poziomu pliku konfiguracyjnego modelu. Domyślnie przyjmuje dwie wartości: 2 dla jednego kanału obsługi i 16 dla większej ich liczby. Jak już zostało wspomniane, parametr ten odpowiada za dokładność całkowania numerycznego metodą Simpsona. Metoda Simpsona wymaga, aby parametr ten był parzysty, stąd w zależności od potrzeb były też stosowane wartości 4 i 8. Rysunki 8 i 9 przedstawiają wpływ wartości tego parametru na czas obliczeń.



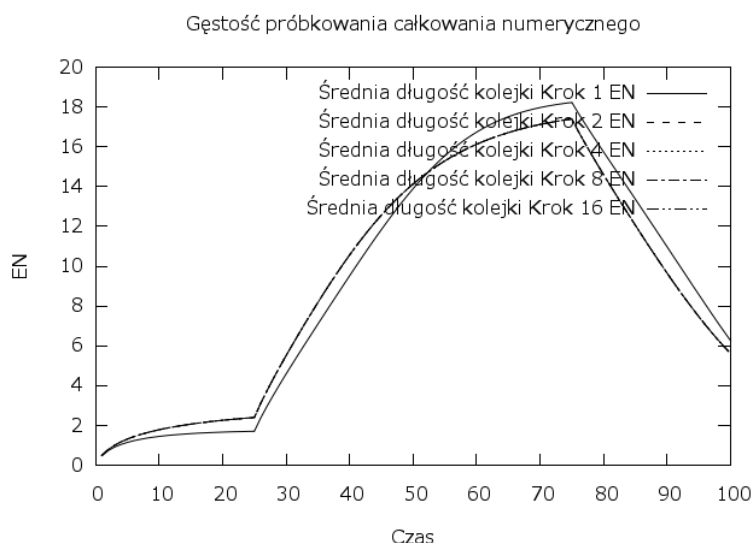
Rys. 8. Średni czas pojedynczego kroku obliczeń w funkcji gęstości próbkowania całkowania numerycznego

Fig. 8. Average unitary calculations step as a function of integration step size



Rys. 9. Całkowity czas obliczeń w funkcji gęstości próbkowania całkowania numerycznego

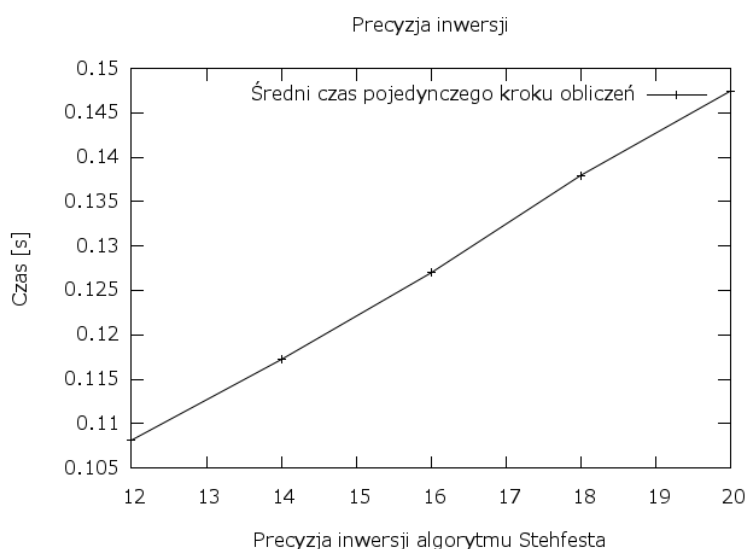
Fig. 9. Total calculations time as a function of integration step size



Rys. 10. Średnia długość kolejki w czasie dla różnych wartości gęstości próbkowania całkowania numerycznego

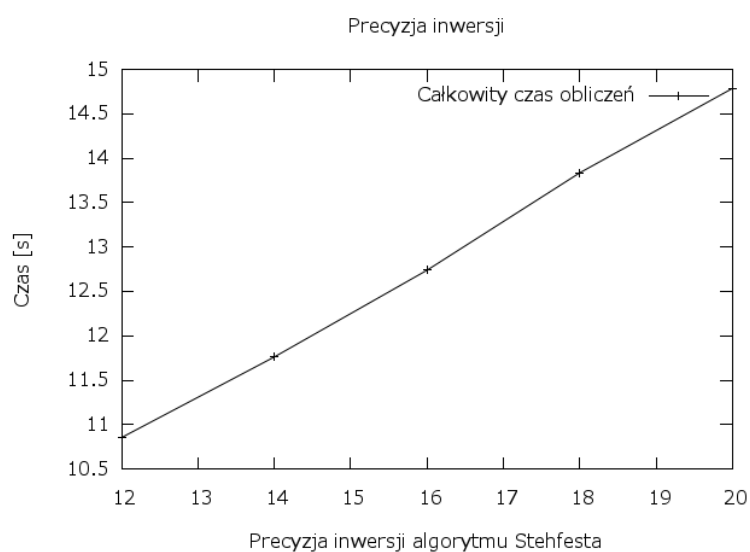
Fig. 10. Mean queue length as a function of time for different sizes of integration step

Rysunek 10 pokazuje wpływ tego parametru na dokładność obliczeń i, jak widać, jego wartości od 2 w górę dają bardzo dobre wyniki. Należy jednak pamiętać, że w tych przykładach mamy jeden kanał obsługi i maksymalną długość kolejki równą 20, co już dla wartości tego parametru równej 2 daje 41 punktów do aproksymacji. Gdyby przykład miał maksymalną liczbę aktywnych kanałów obsługi (20), wartość parametru równa 2 oznaczałaby aproksymowanie funkcji w przedziale zaledwie 3 punktami, co znacznie pogorszyłoby otrzymywane wyniki (stąd przy większej liczbie kanałów obsługi aplikacja stosuje wartość 16).

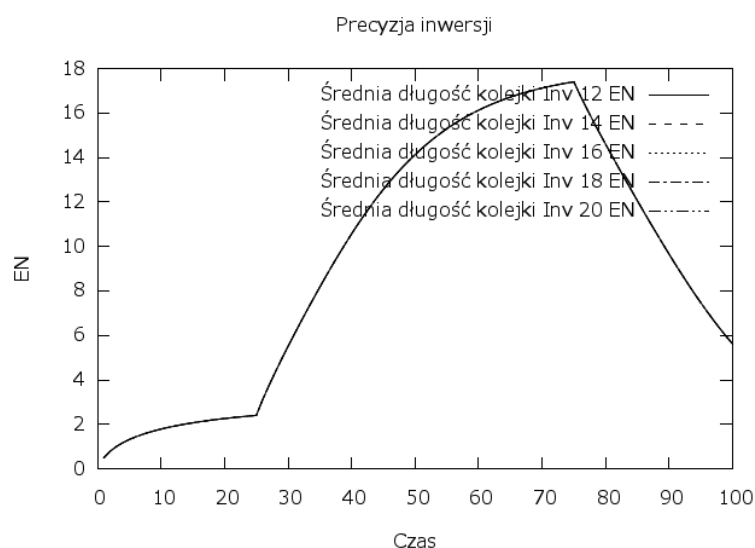


Rys. 11. Średni czas pojedynczego kroku obliczeń w funkcji parametru algorytmu Stehfesta

Fig. 11. Average unitary calculations step as a function of Stehfest inversion precision



Rys. 12. Całkowity czas obliczeń w funkcji parametru algorytmu Stehfesta  
 Fig. 12. Total calculations time as a function of Stehfest inversion precision



Rys. 13. Średnia długość kolejki w czasie dla różnych wartości parametru algorytmu Stehfesta  
 Fig. 13. Mean queue length as a function of time for different values of Stehfest inversion precision

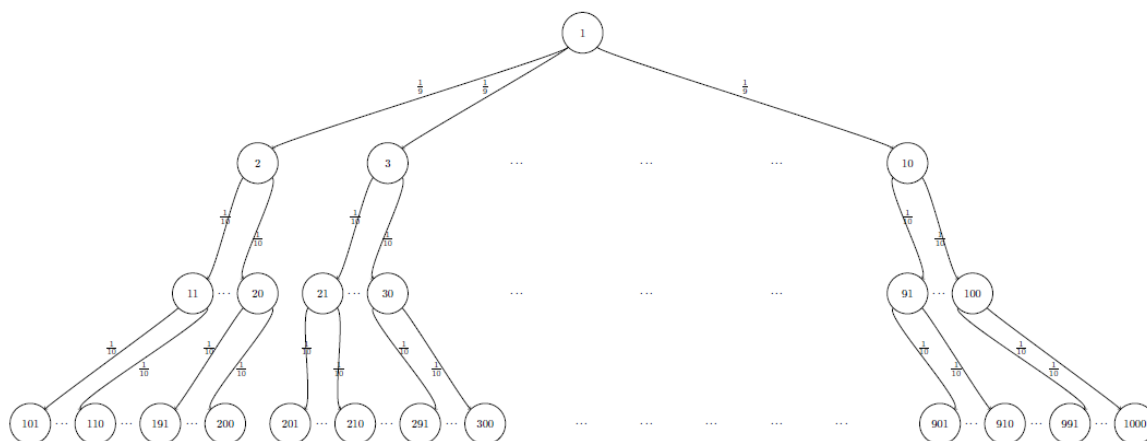
Ostatnim parametrem mającym wpływ na czas obliczeń jest precyzja algorytmu Stehfesta służącego do inwersji Laplace'a. Również z nią parametr jest ustawiony w aplikacji na sztywno na wartość  $n = 18$  i odpowiada on za jeden z wymiarów tablic prawdopodobieństw przejść (przy wartości 18 tego parametru z 18 wartości w dziedzinie Laplace'a otrzymujemy jeden wynik w dziedzinie czasu).

Wpływ tego parametru na czas obliczeń ilustrują rysunki 11 i 12, a rysunek 13 pokazuje wpływ na dokładność obliczeń, która w tym przypadku jest jednakowa dla wszystkich wartości od 12 do 20.

Kolejnym etapem prac było zbadanie wpływu rozmiaru sieci (liczba węzłów) na czas obliczeń. Przebadano sieć o topologii drzewa w 4 konfiguracjach:

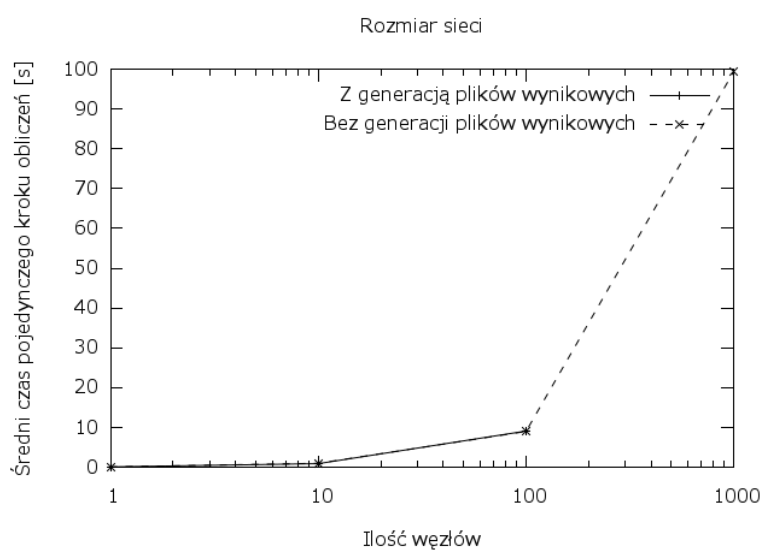
- 1 węzeł (I poziom),
- 10 węzłów (I i II poziom),
- 100 węzłów (I, II i III poziom),
- 1000 węzłów (I, II, III i IV poziom).

Parametry wszystkich węzłów były takie same, z wyjątkiem węzła pierwszego. Przyjęto przypadek w którym cały ruch jest kierowany do pierwszego węzła, a następnie z równym prawdopodobieństwem do kolejnych 9 węzłów (razem 10 węzłów). Na kolejnym etapie drzewa węzłów każdy z 9 wysyła swój ruch z równym prawdopodobieństwem do każdego ze swoich 10 dzieci (razem 100 węzłów). Na ostatnim poziomie mamy 900 węzłów skonfigurowanych podobnie jak węzły wyższej warstwy z tą różnicą, że one nie wysyłają już ruchu dalej. Schemat sieci i poszczególnych warstw ilustruje rysunek 14.

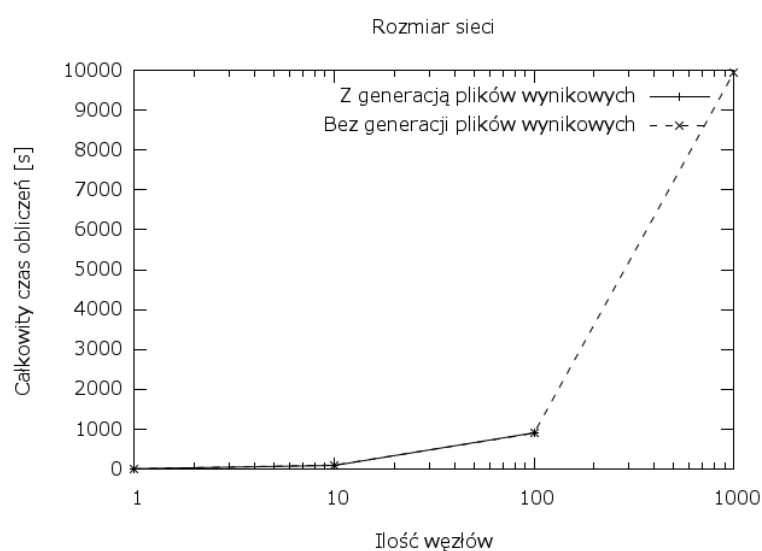


Rys. 14. Schemat sieci z podziałem na kolejne poziomy  
Fig. 14. Network diagram divided into layers

Badania wydajności dla tak skonstruowanej sieci zostały przeprowadzone dwukrotnie, ponieważ w przypadku domyślnej konfiguracji aplikacji nie udało się uzyskać wyników dla 1000 węzłów. Wynika to z ograniczeń systemu operacyjnego na maksymalną liczbę otwartych plików. W podanej konfiguracji (jeden kanał obsługi, maksymalna długość kolejki 20, całkowity czas obliczeń 100 sekund z krokiem co 1 sekundę) każdy z węzłów generuje 9 plików wynikowych. Niestety, system operacyjny nie był w stanie utworzyć więcej niż 1700 plików wynikowych i aplikacja została zamknięta z powodu braku dostępnych zasobów. Ponieważ badana jest wydajność czasowa metody, na potrzeby testu zapis plików wynikowych został wyłączony. Rysunki 15 i 16 przedstawiają otrzymane czasy obliczeń dla 4 rozmiarów sieci w przypadkach z zapisem i bez zapisu plików wynikowych.



Rys. 15. Średni czas pojedynczego kroku obliczeń w funkcji liczby węzłów sieci  
 Fig. 15. Average unitary calculations step as a function of number of nodes



Rys. 16. Całkowity czas obliczeń w funkcji liczby węzłów sieci  
 Fig. 16. Total calculations time as a function of number of nodes

#### 4. Podsumowanie

W pracy przedstawiono model matematyczny umożliwiający analizę stanów nieustalonych w stanowisku obsługi, wykorzystując metodę aproksymacji dyfuzyjnej, która pozwala na przyjęcie w modelu ogólnych założeń, a następnie zbadano, czy modele oparte na tej metodzie nadają się do badania dużych sieci poprzez wyznaczenie czasu potrzebnego do wyliczenia modelu sieci składającej się z czterech różnych wielkości. Jak widać z uzyskanych czasów, metoda ta pozwala na uzyskanie wyników znacznie szybciej niż odpowiednie symu-

latory (symulacja stanów nieustalonych jest niezwykle czasochłonna, bo wymaga wielokrotnego, rzędu setek tysięcy, powtórzenia przebiegów, tak by dla określonego momentu czasu wyznaczyć histogram rozkładu liczby klientów w stanowisku), jednakże nie nadaje się do zastosowań wymagających wyników w czasie rzeczywistym bez zastosowania równoległego modelu obliczeń. Zbadano też wpływ poszczególnych parametrów modelu i zastosowanego algorytmu obliczeniowego na czas i precyzję obliczeń. Parametrem mającym największy wpływ na złożoność czasową metody okazał się parametr odpowiadający za gęstość próbkowania całkowania numerycznego, który jest ściśle powiązany z liczbą aktywnych kanałów obsługi. Najlepszym rozwiązaniem byłoby uzależnienie tego parametru od wielkości podprzedziału. Okazało się również, że wartość  $n = 18$  elementów szeregu jest dla precyzji algorytmu Stehfesta zbyt wysoka, ponieważ już dla wartości  $n = 12$  uzyskuje się bardzo dobre rezultaty, przy jednoczesnym przyspieszeniu wynoszącym 33%.

## BIBLIOGRAFIA

1. Czachórski T.: A method to solve diffusion equation with instantaneous return processes acting as boundary conditions, Bulletin of Polish Academy of Sciences, Technical Sciences vol. 41 (1993), no. 4.
2. Czachórski T.: Modele kolejkowe w ocenie efektywności pracy sieci i systemów komputerowych. Wyd.: Pracowni Komputerowej Jacka Skalmierskiego, Gliwice 1999.
3. Gelenbe E.: On Approximate Computer Systems Models, J. ACM, vol. 22, no. 2, (1975).
4. Gelenbe E., Pujolle G.: The Behaviour of a Single Queue in a General Queueing Network, Acta Informatica, Vol. 7, Fasc. 2, s.123÷136, 1976.
5. Newell G. F.: Applications of Queueing Theory, Chapman and Hall, London 1971.
6. Nycz T., Czachórski T.: Analiza błędów modeli opartych na metodzie aproksymacji dyfuzyjnej, [w:] P. Pikiwicz (red.) Zastosowania Internetu, Wyższa Szkoła Biznesu, Dąbrowa Górnicza 2012, s. 135÷138.
7. Stehfest H.: Algorithm 368: Numeric inversion of Laplace transform, Comm. of ACM, vol. 13, no. 1, s. 47÷49 (1970).
8. Hollot K., Liu Y., Misra V., Towsley D., Gong W.-B.: Fluid methods for modeling large heterogeneous networks. NTIS, kwiecień 2005. AFRL-IF-RS-TR-2005-28
9. Czachórski T, Nycz M, Nycz T., Pekergin F.: Transient states of flows and router queues – a discussion of modelling methods, Proc. of International Conference on Networking and Future Internet (ICNFI 2012), Istanbul, April 25-27, 2012.

**Abstract**

The article presents investigation of numerical calculations for well-known mathematical model that allows analysis of transient states at the service station and their networks, using the diffusion approximation method. The influence of various parameters of the model and size of the network to the total computation time was investigated, and thus the scalability of the method.

**Adresy**

Tomasz NYCZ: Politechnika Śląska, Instytut Informatyki, ul. Akademicka 16,  
44-100 Gliwice, Polska, tomasz.nycz@polsl.pl

Tadeusz CZACHÓRSKI: Instytut Informatyki Teoretycznej i Stosowanej Polskiej Akademii  
Nauk, ul. Bałtycka 5, 44-100 Gliwice, Polska, tadek@iitis.pl