

ZDZISŁAW MILIAN*

DOKŁADNE ROZWIĄZANIA PROBLEMU PERT
Z MIESZANYMI ROZKŁADAMI CZASU REALIZACJI
ZADAŃ

EXACT SOLUTIONS OF PERT PROBLEM WITH MIXED
PROBABILITY DISTRIBUTIONS OF TASK DURATIONS

Streszczenie

W artykule zaprezentowano dokładne rozwiązania problemu PERT w przypadku mieszanych rozkładów czasu trwania zadań. Metoda ta polega na obliczaniu łącznych rozkładów terminów realizacji węzłów w sposób rekurencyjny, aby w końcowym kroku obliczyć rozkład terminu realizacji ostatniego węzła. Pokazano wiele przykładów ilustrujących własności rozwiązań dla rozkładów mieszanych. Metoda jest dostępna na serwisie internetowym, gdzie użytkownik może zapoznać się z bogatym zbiorem przykładów ilustrujących własności rozwiązań oraz otrzymać rozwiązania własnych przykładów.

Słowa kluczowe: dokładna metoda PERT, rozkłady mieszane

Abstract

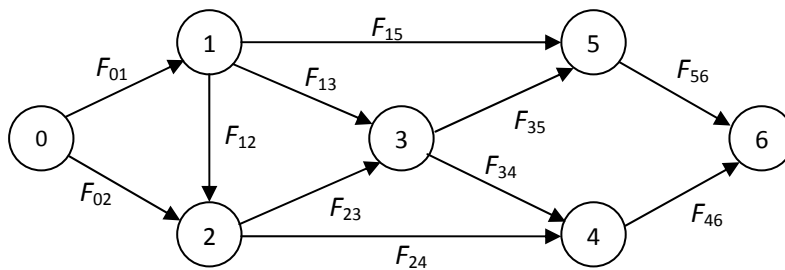
In the paper the exact solutions of PERT problem are presented. Task durations obey mixed probability distributions. The method evaluates the join mixed distribution of node realization dates in a recursive way from the start node to the terminal node. Many examples are presented to illustrate the properties of solutions. Moreover, the method is accessible on the web. A user may familiarize with collection of examples illustrating properties of solutions and obtain solutions of his own examples.

Keywords: the exact PERT method, mixed distributions

* Dr inż. Zdzisław Milian, Zakład Technologii i Organizacji Budownictwa, Wydział Inżynierii Lądowej, Politechnika Krakowska.

1. Wstęp

Sieć PERT określa graf skierowany, acykliczny i spójny wraz z przyporządkowanymi rozkładami czasu trwania zadań wchodzących w skład przedsięwzięcia. W przyjętej konwencji łuki grafu reprezentują zadania, natomiast węzły określają najwcześniejsze momenty zadań w następujący sposób. Zadania wychodzące z węzła mogą się rozpocząć tylko wtedy, gdy wszystkie zadania dochodzące bezpośrednio do węzła zostaną ukończone. Niech t_{ij} oznacza czas trwania zadania wychodzącego z wierzchołka i , a kończącego się w wierzchołku j . Czas t_{ij} trwania każdego zadania jest zmienną losową o zadanej funkcji dystrybuanty $F_{ij}(t)$. Zmienne losowe t_{ij} są wzajemnie niezależne. Przykładową sieć PERT ilustruje rys. 1.



Rys. 1. Przykład sieci PERT

Fig. 1. An example of PERT Network

Dokładny problem PERT polega na tym, aby na podstawie danych (graf + funkcje F_{ij}) określić rozkład czasu realizacji wszystkich zadań, innymi słowy, określić rozkład czasu realizacji przedsięwzięcia w postaci dokładnej formuły matematycznej lub zbioru matematycznych formuł. Sformułowanie to różni się zatem od znanej metody PERT w której poszukuje się jedynie rozwiązań przybliżonych w postaci krzywej Gaussa i ogranicza się rozkłady wejściowe do rozkładów beta. Choć odnośnie problemu PERT jest bardzo bogata literatura, to jednak na temat dokładnych metod jest jej bardzo mało. Bardziej szczegółowe omówienie literatury znajduje się w pracach [2, 7, 10].

Rozkładem mieszanym w sensie rachunku prawdopodobieństwa rozumie się rozkład który ma cechy rozkładu ciągłego i dyskretnego, a mianowicie dopuszcza skoki funkcji dystrybuanty. Przypomnijmy, funkcja dystrybuanty jest zawsze funkcją niemalejącą. Dla rozkładów ciągłych dystrybuanta wzrasta tylko w sposób ciągły a dla rozkładów dyskretnych tylko w dyskretnych punktach. Natomiast w rozkładzie mieszanym jej wzrosty następują zarówno w sposób ciągły, jak też skokowy. Przykład dystrybuanty rozkładu mieszanego przedstawiony jest na rys. 2.

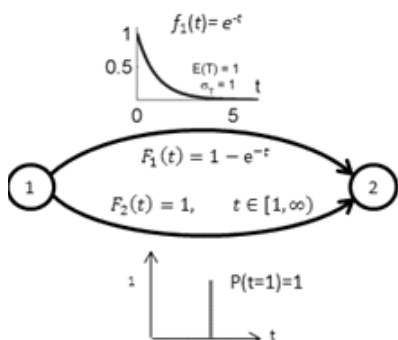
2. Sieci równoległe

Aby przedstawić mechanizm powstawania rozkładów mieszanych w kontekście problemu PERT, zaczniemy od najprostszego układu równoległego dwóch zadań. Niech $F_1(t)$ i $F_2(t)$ oznaczają dystrybuanty czasu realizacji zadania odpowiednio pierwszego

i drugiego. Wiadomo, że w tym przypadku dystrybuanta zastępcza dla sieci niezależnie od typu rozkładów $F_1(t)$ i $F_2(t)$ wyraża się poprzez iloczyn dystrybuant

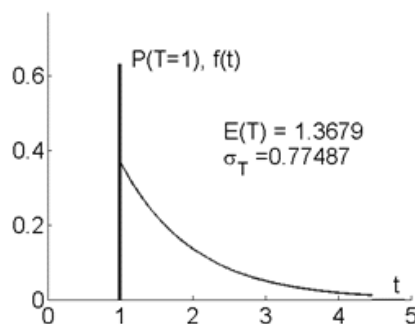
$$F(t) = F_1(t) \cdot F_2(t). \tag{1}$$

Przykład 1. W sieci równoległej składającej się z dwóch zadań, czas realizacji jednego z zadań ma rozkład ciągły, a czas drugiego zadania jest stały. Dane szczegółowo prezentuje rys. 2. Należy wyznaczyć rozkład dla danej sieci.



Rys. 2. Sieć PERT dla przykładu 1

Fig. 2. PERT Network of example 1



Rys. 3. Rozkład mieszany dla przykładu 1

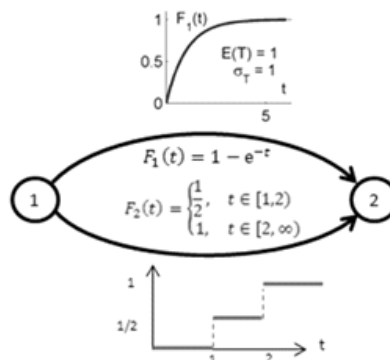
Fig. 3. Mixed distribution for example 1

W tym przypadku rozkład wynikowy jest rozkładem mieszanym. Jest to zrozumiałe, ponieważ wartość stała jest szczególnym przypadkiem rozkładu dyskretnego który jest skupiony tylko w jednym punkcie.

Przykład 2. W sieci równoległej składającej się z dwóch zadań, czas realizacji jednego z zadań ma rozkład ciągły, a drugiego dyskretny. Szczegóły dotyczące rozkładów podane są na rys. 4. Należy wyznaczyć rozkład dla danej sieci.

Na postawie (1) możemy obliczyć dystrybuantę zastępczą

$$F(t) = \begin{cases} 0 & t \in (-\infty, 1) \\ \frac{1}{2}(1 - e^{-x}) & t \in [1, 2) \\ 1 - e^{-x} & t \in [2, \infty). \end{cases}$$



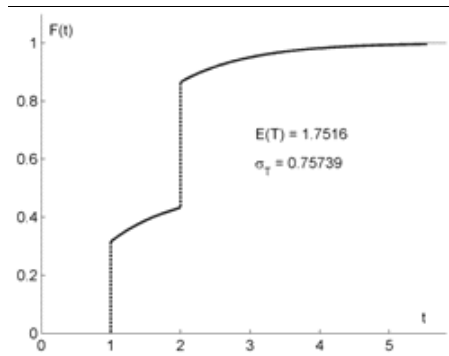
Rys. 4. Przykład sieci równoległej

Fig. 4. An example of parallel network

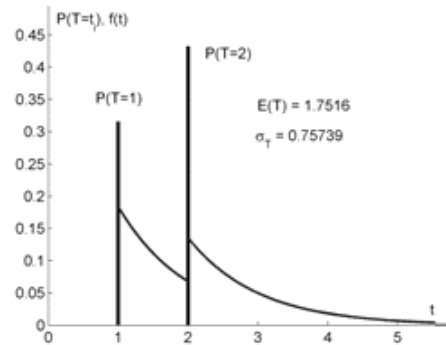
Przebieg tej funkcji został zilustrowany na rys. 5. Jak można na nim zauważyć, mamy dwa przedziały ciągłego wzrostu funkcji oraz dwa skoki w punkcie 1 i 2. Jest to zatem rozkład mieszany. Wielkość skoku w punkcie τ_i oznacza, podobnie jak dla rozkładów dyskretnych, prawdopodobieństwo, że czas realizacji sieci wynosi τ_i . Prawdopodobieństwo to jest zatem różnicą pomiędzy prawostronną i lewostronną granicą funkcji $F(t)$ w punkcie τ_i i będzie dalej oznaczane przez p_i . Możemy to zapisać następująco

$$P(T = \tau_i) = p_i = F^+(\tau_i) - F^-(\tau_i) \tag{2}$$

Na rys. 6 przedstawiona jest mieszana gęstość rozkładu odpowiadająca dystrybuancie z rys. 5. Dla rozkładów ciągłych gęstość rozkładu jest pochodną z dystrybuanty.

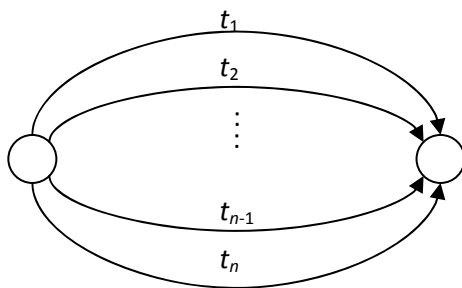


Rys. 5. Dystrybuanta czasu realizacji dla sieci 2
Fig. 5. The cdf function for example 2



Rys. 6. Rozkład mieszany dla przykładu 2
Fig. 6. Mixed pdf function for example 2

Rozważmy teraz przypadek realizacji n zadań, które mogą być realizowane jednocześnie. Czas realizacji ostatniego zadania jest czasem wykonania grupy n zadań. Tę sytuację ilustruje rys. 7.



Rys. 7. Równoległa sieć PERT z n zadaniami
Fig. 7. Parallel PERT Network with n tasks

Formułę (1) można łatwo uogólnić na przypadek n zadań, czyli

$$F(t) = \prod_{i=1}^{i=n} F_i(t). \tag{3}$$

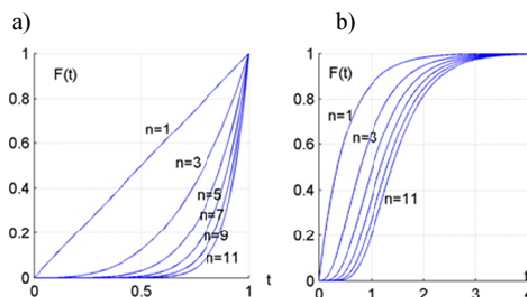
Na podstawie przykładu 2, w odniesieniu do sieci na rys. 7 nasuwa się następujący wniosek.

Wniosek 1. W równoległych sieciach PERT, jeśli występuje co najmniej jeden rozkład ciągły na odcinku i dyskretny (bądź mieszany) oraz rozkład ciągły rozpoczyna się wcześniej niż ostatni skok rozkładu dyskretnego (bądź mieszanego), to wynikowy rozkład dla całej sieci jest rozkładem mieszanym.

Wniosek 1 wyjaśnia nam zatem mechanizm powstawania rozkładów mieszanych.

Wniosek 2. W równoległych sieciach PERT, jeśli występuje rozkład, powiedzmy $F_i(t)$ który rozpoczyna się później niż kończą się wszystkie pozostałe rozkłady, to ten rozkład jest również rozkładem wynikowym dla całej sieci. Jeśli dodatkowo $F_i(t)$ jest rozkładem ciągłym, to rozkład wynikowy jest ciągły.

Przedstawimy teraz ważną tendencję zachowania się rozkładu wynikowego wraz ze wzrostem liczby n . Dla większej przejrzystości posłużymy się rozkładami ciągłymi. Rys. 8a przedstawia wykresy dystrybucji realizacji zadań w układzie równoległym o liczbie zadań $n = 1, 3, 5, 7, 9, 11$, gdy rozkład każdego z zadań jest równomierny na odcinku $[0, 1]$, natomiast rys. 8b odpowiada sytuacji, gdy każde z zadań ma rozkład wykładniczy z parametrem $\lambda = 2$. W obu przypadkach średni czas pojedynczego zadania wynosi $\frac{1}{2}$. Zgodnie z (3), w przypadku 8a wartość dystrybuanty wynikowej wynosi t^n dla $t \in [0, 1]$ i 1 dla $t > 1$. W przypadku 8b dystrybuanta wynikowa ma postać $(1 - e^{-2t})^n$, $t > 0$.



Rys. 8. Dystrybuanty sieci równoległej o liczbie zadań $n = 1, 3, 5, 7, 9, 11$

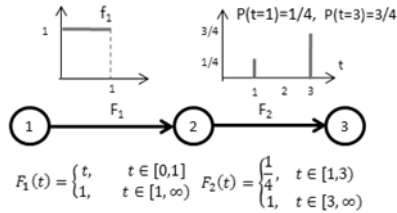
Fig. 8. The cdfs for the parallel network with n tasks, $n = 1, 3, 5, 7, 9, 11$, durations distributed: a) exponentially, b) uniformly

Jak widać z tych rysunków, przebieg wykresów zależy od rodzaju rozkładu, zadania jak również liczby zadań. Wraz ze wzrostem n dystrybuanta staje się coraz bardziej stroma, a odcinek wzrostu przesuwają się w prawo. Z tego wynika też fakt dużego spadku prawdopodobieństwa dla ustalonego czasu wraz ze wzrostem liczby zadań, np. prawdopodobieństwo wykonania 11-tu zadań w czasie średnim jednego zadania spada z wartości 0,5 do 0,0005 dla przypadku rozkładu jednostajnego oraz z wartości 0,6321 do 0,0064 dla rozkładu wykładniczego. Jest rzeczą oczywistą, że podobna tendencja dotyczy również rozkładów dyskretnych. Podobnie zachowują się duże sieci PERT w których jest wiele dróg krytycznych w sensie probabilistycznym ([7]).

3. Sieci szeregowe

O ile w sieciach równoległych mogą być generowane rozkłady mieszane, to w sieciach szeregowych rozkłady mieszane (lub dyskretne) mogą zniknąć. Rozważmy najprostszą sieć szeregową składającą się z dwóch zadań, w której pierwsze zadanie ma rozkład ciągły a drugie dyskretne.

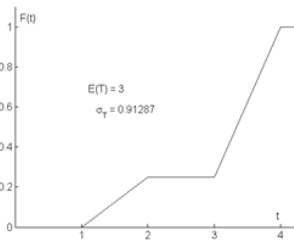
Przykład 3. Niech w sieci szeregowej składającej się z dwóch zadań, czas realizacji pierwszego zadania ma rozkład ciągły a drugiego dyskretny. Szczegółowe dane zamieszczone są na rys. 9. Należy wyznaczyć rozkład dla danej sieci.



Rys. 9. Przykład sieci szeregowej

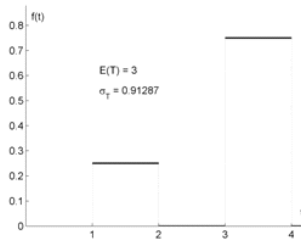
Fig. 9. An example of serial Network

$$F(t) = \begin{cases} 0, & t < 1, \\ \frac{1}{4}(t - 1), & 1 \leq t < 2, \\ \frac{1}{4}, & 2 \leq t < 3, \\ \frac{3}{4}(t - 3) + \frac{1}{4}, & 3 \leq t < 4, \\ 1, & t > 4, \end{cases} \quad f(t) = \begin{cases} 0, & t < 1, \\ \frac{1}{4}, & 1 \leq t < 2, \\ 0, & 2 \leq t < 3, \\ \frac{3}{4}, & 3 \leq t < 4, \\ 0, & t > 4. \end{cases}$$



Rys. 10. Dystrybuanta czasu realizacji dla sieci 3

Fig. 10. The cdf function for example 3



Rys. 11. Funkcja gęstości dla przykładu 3

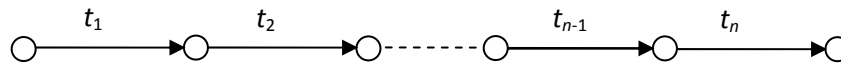
Fig. 11. The pdf function for example 3

Wykresy dystrybuanty i gęstości dla sieci z przykładu 3 przedstawione są na rys. 10 i 11. Zauważmy, że ta dystrybuanta nie ma skoków, a więc posiada odcinkami ciągłą gęstość. Ponadto widzimy, że gęstość wypadkowa przedstawiona na rys. 11 powstała przez przesunięcie gęstości $f_1(t)$ o wartości 1 i 3 oraz pomnożenie przez prawdopodobieństwo wystąpienia tych wartości, czyli przez $1/4$ i $3/4$.

W tym przypadku rozkład czasu trwania jest sumą zmiennych losowych, dlatego wykorzystamy operację splotu

$$F(t) = \int_0^t F_2(u) f_1(t - u) du \quad (4)$$

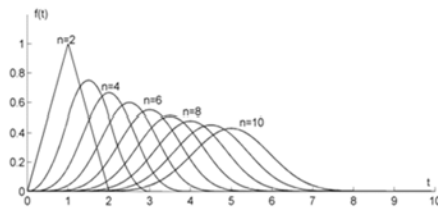
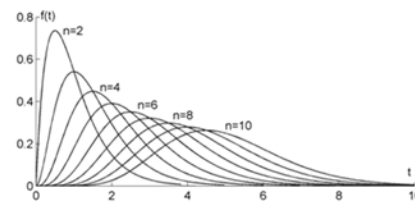
do obliczenia dystrybuanty czasu realizacji sieci. Łatwo sprawdzić, że na podstawie (4) otrzymujemy

Rys. 12. Szeregowa sieć n zadańFig. 12. Serial network of n tasks

W przypadku sieci szeregowych o większej liczbie zadań (rys. 12, $n > 2$) możemy wykorzystać formułę (4) w sposób rekurencyjny, tzn. mając dystrybuantę rozkładu czasu dwóch zadań dokonać splotu tej dystrybuanty z dystrybuantą czasu trzeciego zadania, następnie dokonać kolejnego splotu otrzymanej funkcji z dystrybuantą czasu zadania czwartego itd., aż otrzymamy dystrybuantę czasu n zadań. Oczywiście rekurencyjne stosowanie (4) wymaga aby zmienna t_1 posiadała gęstość. Ponieważ operacja splotu jest przemienne, to tylko w przypadku gdy żadna ze zmiennych nie posiada gęstości, całkę w operacji splotu (4) musimy zastąpić odpowiednią sumą. Na podstawie przykładu 3 oraz możliwości rekurencyjnego postępowania możemy wyciągnąć następujący wniosek.

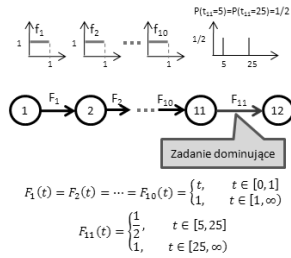
Wniosek 3. Rozkład wynikowy dla szeregowej sieci PERT jest rozkładem mieszanym tylko wtedy, gdy każda dystrybuanta zadania posiada co najmniej jeden skok, w przeciwnym przypadku rozkład wynikowy jest rozkładem ciągłym.

Zachowanie rozkładu sumy niezależnych zmiennych losowych ilustruje rys. 13.

a) rozkład zadania równomierny na odcinku $[0, 1]$ b) rozkład zadania wykładniczy z parametrem $\lambda = 2$ Rys. 13. Rozkład gęstości sieci równoległej o liczbie zadań $n = 2, 4, 6, 8, 10$ Fig. 13. The pdfs for the serial network with n tasks, $n = 2, 4, 6, 8, 10$, durations distributed: a) uniformly on $[0, 1]$, b) exponentially with $\lambda = 2$

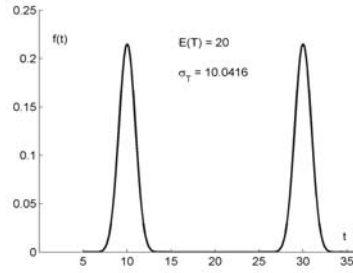
Rys. 13a i b przedstawiają zbieżność rozkładu sumy do rozkładu normalnego odpowiednio $N(n/2, \sqrt{n/12})$ oraz $N(n/2, \sqrt{n/4})$ co wynika z centralnego twierdzenia granicznego. Zbieżność rozkładu sumy do rozkładu normalnego nie zawsze występuje. Kolejny przykład pokazuje taki przypadek.

Przykład 4. Niech w sieci szeregowej składającej się z 11 zadań, czas realizacji pierwszych dziesięciu zadań ma rozkład ciągły a ostatniego dyskretny. Szczegółowe dane zamieszczone są na rys. 14. Należy wyznaczyć rozkład dla danej sieci.



Rys. 14. Sieć szeregową z dominującym zadaniem

Fig. 14. Serial network with dominating task



Rys. 15. Gęstość rozkładu dla przykładu 4

Fig. 15. The pdf for Example 4

Rozwiązanie tego przykładu ilustruje rys. 15. Ponieważ rozkłady pierwszych dziesięciu zmiennych losowych są równomierne na odcinku $[0, 1]$, rozkład zastępczy dla tego fragmentu sieci jest zbliżony do rozkładu normalnego $N(5, \sqrt{10/12})$ przedstawionego na rys. 13a. Następny splot tego rozkładu z rozkładem dyskretnym powoduje, podobnie jak w przykładzie 3, że rozkład normalny zostaje przesunięty o wartości 5 i 25. Skutkiem czego powstaje rozkład składający się z dwóch części podobnych do rozkładu normalnego. Oczywiście ten rozkład nie przeczy centralnemu twierdzeniu granicznemu, gdyż w tym przypadku nie są spełnione jego założenia. Zbieżność do rozkładu normalnego wymaga spełnienia warunku Linderberga [3]. Warunek Lindeberga oznacza, że poszczególne wariancje σ_k^2 są małe w porównaniu z ich sumą $s_n^2 = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2$ w tym sensie, że dla dowolnego $\varepsilon > 0$ i dostatecznie wielkich n

$$\frac{\sigma_k}{s_n} < \varepsilon, \quad k = 1, \dots, n. \tag{5}$$

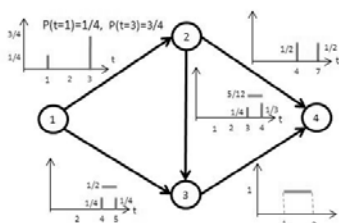
W naszym przypadku, wariancja rozkładów jednostajnych wynosi $1/12$ a rozkładu dyskretnego 100, dlatego stosunek σ_k/s_n wynosi $\sqrt{100/(100 + 10/12)}$ co nie spełnia (5). Dla praktyki modelowania czasu realizacji ciągu zadań, ważna jest następująca własność zmiennych losowych [3]: jeśli zmienne X_k są jednostajnie ograniczone, tzn. wszystkie rozkłady F_k są skoncentrowane na tym samym przedziale $[-a, a]$ oraz $s_n \rightarrow \infty$, wtedy warunek Lindeberga jest spełniony.

Celem tego przykładu było podkreślenie faktu, że nie tylko liczba zmiennych losowych decyduje o zbieżności ich sumy do rozkładu normalnego, ale również brak zmiennej, która dominuje nad innymi w sensie wariancji (takie znaczenie ma wyróżnienie zadania jedenastego na rys. 14). Ta własność pozwala również zrozumieć zalety i wady klasycznej metody PERT przyjmującej rozkład normalny jako przybliżenie rozkładu dla sieci.

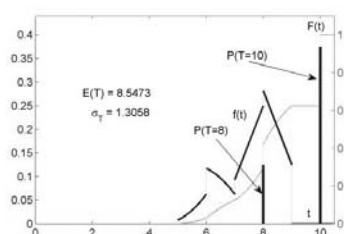
4. Sieci nieredukowalne

Przez sieci nieredukowalne rozumie się takie sieci, których nie można zredukować do jednego zastępczego zadania poprzez operację (3) i (4), czyli mnożenia i splotu dystrybuant. Przykład najmniejszej takiej sieci jest podany na rys. 16.

Przykład 5. Dana jest sieć PERT jak na rys. 16. Łukom przyporządkowane są rozkłady dyskretne i mieszane. Dane szczegółowe przedstawione są przy poszczególnych łukach. Należy wyznaczyć rozkład dla danej sieci.



Rys. 16. Przykład sieci nieredukowalnej



Rys. 17. Końcowy rozkład dla przykładu 5

Fig. 16. An example of irreducible Network

Fig. 17. Final distribution of example 5

Poszukiwany rozkład końcowy ma dystrybuantę, która jest nieciągła w punktach 8 i 10, poza tymi punktami ma gęstość, a więc gęstość mieszana ma postać

$$f(t) = \begin{cases} \frac{5}{192}t^2 - \frac{89}{384}t + \frac{33}{64} & 5 \leq t < 6 \\ \frac{-5}{192}t^2 + \frac{109}{384} * t - \frac{83}{128} & 6 \leq t < 7 \\ \frac{5}{32}t - 1 & 7 \leq t < 8 \\ \frac{1}{8} \cdot \delta(t - 8) & t = 8 \\ -\frac{5}{32} * t + \frac{49}{32} & 8 < t < 9 \\ \frac{3}{8} \cdot \delta(t - 10) & t = 10 \end{cases}$$

Szczegółowa metodologia dokładnego rozwiązywania problemu PERT znajduje się w pracach [6, 7]. Ograniczymy się tutaj do wyjaśnienia wartości dyskretnych rozkładu. Patrząc na postać rozwiązania widać, że mamy sześć obszarów z dodatnim prawdopodobieństwem, cztery odcinki i dwa dyskretne punkty. Dyskretne wartości pochodzą ze ścieżki 1–2–4, na tej ścieżce oba rozkłady są dyskretne, dlatego możliwe są skoki dystrybuanty. Na pozostałych dwóch ścieżkach 1–2–3–4 i 1–3–4 skoki nie wystąpią ze względu na ciągły rozkład na łuku 3–4, zgodnie z wnioskiem 3. Ze ścieżki 1-2-4 wynika, że jej czas trwania może mieć 4 wartości 5, 7, 8 i 10 z dodatnim prawdopodobieństwem. W rozwiązaniu występuje 8 i 10. Powstaje pytanie, dlaczego nie ma innych wartości. Wynika to stąd, że wartość 5 jako suma 1 + 4 gdyby miała być maksymalna, to wierzchołek 2 byłby osiągnięty po 1 jednostce a czas na ścieżkach 2–3–4 musiałby być mniejszy bądź równy 4 oraz 1–3–4

mniejszy bądź równy 5. Najmniejszy czas realizacji wierzchołka 3 wynosi 4, a prawdopodobieństwo tego zdarzenia jest dodatnie. Natomiast prawdopodobieństwo dotarcia z wierzchołka 3 do 4 w czasie krótszym niż 10, gdyż rozkład czasu na łuku 3–4 rozpoczyna się w 1 i jest rozkładem ciągłym. Prawdopodobieństwo osiągnięcia czasu realizacji 5 jest zerowe i nie występuje w rozwiązaniu. Podobnie zerowe prawdopodobieństwo ma czas równy 7. Krótko mówiąc, jeżeli na ścieżce 1–2–4 zrealizuje się wartość 5 lub 7, to na innych ścieżkach zrealizuje się wartość większa niż 5. Nietrudno zauważyć, że pozostałe potencjalne wartości dyskretne 8 i 10 mają dodatnie prawdopodobieństwo. Jeżeli na ścieżce 1–2–4 zrealizuje się wartość 8, to na podstawie analizy rozkładów na innych ścieżkach wynika, że nie można osiągnąć na nich wartości większej niż 8, stąd prawdopodobieństwo przyjęcia wartości 8 wynika tylko ze ścieżki 1–2–4. $P(T = 8) = P(t_{12} = 1) \cdot P(t_{24} = 7) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$. Podobnie $P(T = 10) = P(t_{12} = 3) \cdot P(t_{24} = 7) = \frac{3}{4} \cdot \frac{1}{2} = \frac{3}{8}$. Pozostałe przypadki rozwiązania wynikają z realizacji wartości 5 lub 7 na ścieżce 1-2-4 oraz wszystkich kombinacji pozostałych rozkładów.

Do rozwiązania problemu dla ogólnych sieci PERT opracowana została nowa metoda obliczania rozkładu wynikowego niezależna od topologii sieci i typów rozkładu [7]. Polega ona w skrócie na rekurencyjnym obliczaniu pośrednich *łącznych* rozkładów zbioru węzłów niezbędnych do wyznaczenia rozkładu węzła końcowego jako rozkładu granicznego.

Dla ciągłych rozkładów funkcje łącznej dystrybuanty i gęstości wektora V_k określone są następującymi formułami rekurencyjnymi

$$G_k(x_{j_1}, x_{j_2}, \dots, x_{j_{r+q}}) = P(T_{j_1} < x_{j_1}, T_{j_2} < x_{j_2}, \dots, T_{j_{r+q}} < x_{j_{r+q}}) = \quad (6)$$

$$= \int_{R^p} \int_0^{x_{j_{r+1}}} \dots \int_0^{x_{j_{r+q}}} \prod_{s=1}^r \left(\prod_{i < j_s} F_{i j_s}(x_{j_s} - u_i) \right) g_{k-1}(u_{i_1}, \dots, u_{i_{p+q}}) du_{i_1}, \dots, u_{i_{p+q}}$$

oraz

$$g_k(x_{j_1}, x_{j_2}, \dots, x_{j_{r+q}}) = \int_{R^p} \prod_{s=1}^r \frac{\partial}{\partial x_{j_s}} \left(\prod_{i \in I(j_s)} F_{i j_s}(x_{j_s} - u_i) \prod_{i \in L(j_s)} F_{i j_s}(x_{j_s} - x_i) \right) \cdot g_{k-1}(u_{i_1}, \dots, u_{i_p}, x_{j_{r+1}}, \dots, x_{j_{r+q}}) du_{i_1}, \dots, u_{i_p} \quad (7)$$

gdzie $u_0 = 0, g_0 = 1$. Szczegółowe objaśnienie tych wzorów znajduje się w (Milian, 2006). Ponadto, dopuszczalne są również dyskretne i mieszane rozkłady czasu trwania poszczególnych zadań, w tym też rozkłady skupione w jednym punkcie a więc wartości deterministyczne. Uogólnienie na rozkłady mieszane znacznie komplikuje problem, gdyż musi się uwzględniać przypadki gdy nie ma funkcji gęstości i trzeba wykorzystywać skupione miary prawdopodobieństwa. Analiza musi się odbywać w n -wymiarowej przestrzeni. W tym celu opracowane zostały specjalne procedury i algorytmy [6].

5. Praktyczne możliwości obliczeń dokładnych rozkładów dla sieci

Problem znajdowania dokładnych rozkładów dla sieci PERT jest bardzo złożony i praktyczne znalezienie rozwiązania wymaga wsparcia komputerowego do obliczeń symbolicznych (pochodne, całki, granice). Dlatego powstał autorski program komputerowy do rozwiązywania tego problemu dostępny w Internecie pod adresem <http://149.156.141.212:3003/~zmilian/index.php>. Jego głównym celem jest prezentacja nowej, analitycznej dokładnej metody określania rozkładu czasu realizacji przedsięwzięcia zamodelowanego przez sieć PERT, szczególnie prezentacja przykładowych rozwiązań oraz porównanie ich z wynikami otrzymanymi metodami przybliżonymi, np. metodą PERT (krzywa Gaussa) oraz metodą symulacji. Przygotowane przykłady prezentują różnorodność rozwiązań oraz ich własności w zależności od topologii sieci jak również rozkładów czasów zadań wchodzących w skład przedsięwzięcia. Przykłady mają na celu wskazanie kiedy można stosować przybliżone metody, a kiedy ich stosowanie obarczone jest dużym błędem. Różnorodność rozwiązań wynika z natury problemu oraz z tego, że metoda nie narzuca ograniczeń na topologię sieci ani na typ rozkładu. Dopuszcza się rozkłady ciągłe, dyskretne oraz mieszane, a także wartości deterministyczne.

Ponadto użytkownik tego portalu ma możliwość

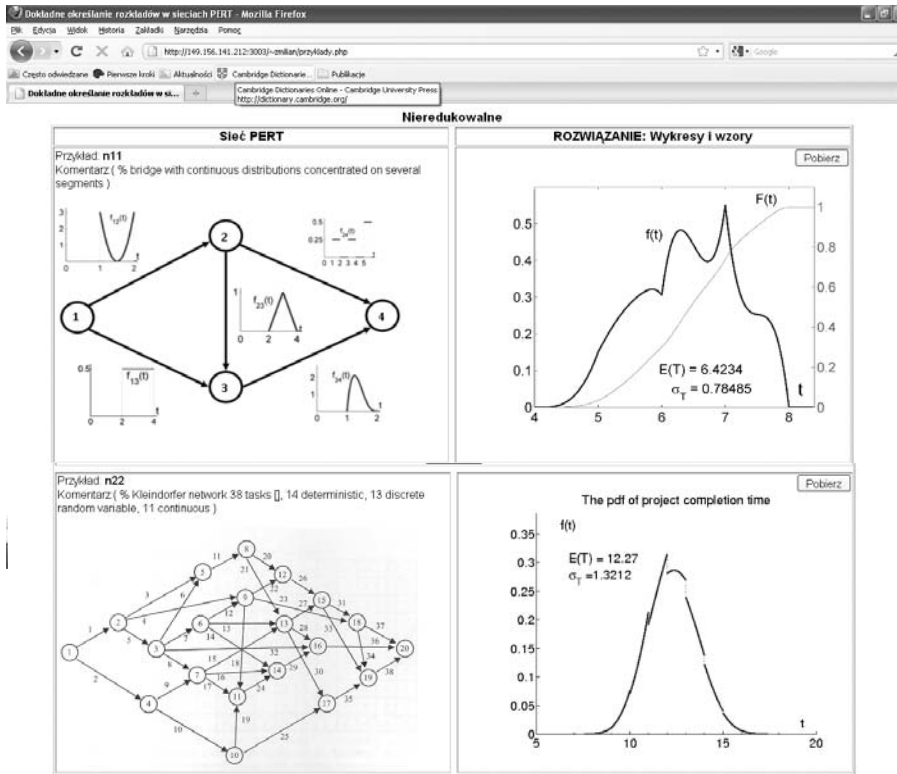
- definiowania i rozwiązywania własnych przykładów,
- importowania rozwiązań w postaci formuł matematycznych i wykresów.

Istnieją też ograniczenia wynikające z wykładniczego charakteru złożoności obliczeniowej problemu w zależności od liczby zmiennych losowych i ich charakteru, zdolności obliczeniowej serwera oraz z dydaktycznego przeznaczenia portalu. Na rys. 17 znajduje się przykładowy widok z ekranu.

Proces obliczeń z wykorzystaniem serwisu internetowego

Użytkownik, mając do dyspozycji tylko przeglądarkę internetową, może zdefiniować własną sieć i zlecić rozwiązanie zadania. Rozwiązanie w postaci analitycznej prezentowane jest automatycznie na stronie internetowej użytkownika w postaci graficznej i analitycznej. Obliczenia prowadzone są na serwerze Zakładu Technologii i Organizacji Budownictwa. W celu prezentacji własności rozwiązań, jak również dla ułatwienia użytkownikowi definiowania własnych sieci, serwis zawiera wiele przykładów, które może użytkownik przeglądać, pobierać i modyfikować jako własne przykłady.

Aby rozwiązać własny przykład, użytkownik musi się zalogować i przygotować dane określające jego sieć w postaci pliku definiującego funkcję w Matlabie. Wzorzec zawartości tego pliku opisany jest w instrukcji użytkownika dostępnej ze strony internetowej, a ponadto dostępne są przykłady. Opcja Więcej w przykładach powoduje odsłonięcie pliku danych i wzorów określających rozwiązanie. Użytkownik nie musi mieć zainstalowanego Matlab'a, gdyż obliczenia wykonywane są na serwerze skojarzonym ze stroną internetową, a edycję pliku danych można prowadzić bezpośrednio ze strony. Obliczenia inicjuje użytkownik, klikając opcję Rozwiąż. Po obliczeniach wyniki dostępne są na stronie. Czas oczekiwania na wyniki zależy od stopnia złożoności przykładu jak również obciążenia sieci i serwera. Użytkownik może nie tylko przeglądać wyniki, ale może również pobrać pliki wynikowe do własnych analiz. Dostępne są pliki z wykresami rozkładów oraz pliki ze wzorami matematycznymi.

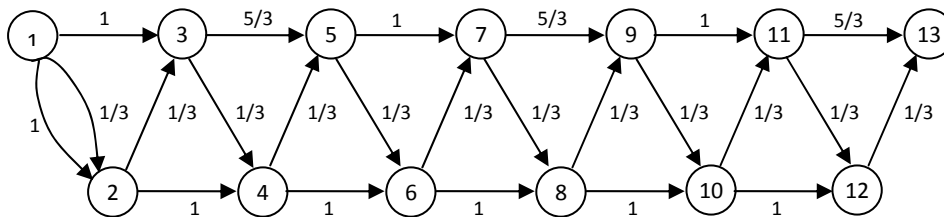


Rys. 18. Fragment ekranu prezentujący dokładną metodę PERT

Fig. 18. Screenshot presenting the exact PERT method

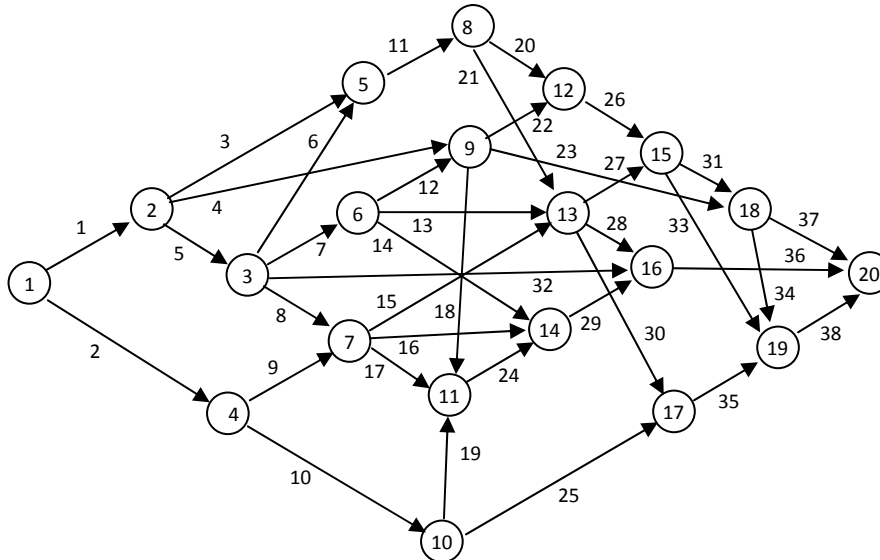
Własności obliczeniowe

Najkorzystniejsze z obliczeniowego punktu widzenia są sieci jak najbardziej zbliżone do szeregowych, wtedy łączne rozkłady, które występują w iteracyjnej procedurze mają małą liczbę zmiennych, np. w sieci na rys. 19 kolejne łączne rozkłady mają najwyżej trzy zmienne. Przykład dla tej sieci znajduje się w zbiorze przykładów.



Rys. 19. Sieć wymagająca 3-wymiarowej analizy

Fig. 19. A network that required 3-dimensional analysis



Rys. 20. Sieć wymagająca 11-wymiarowej analizy

Fig. 20. A network that required 11-dimensional analysis

Przeciwieństwem jest sieć na rys. 20, gdzie występuje konieczność obliczenia łącznego rozkładu o 11 zmiennych (związanych z wierzchołkami $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13\}$), co powoduje że wzory analityczne stają się długie, a to wiąże się z czasem obliczeń i zapotrzebowaniem na pamięć. Szczególny przypadek tej sieci został rozwiązany dokładnie i zamieszczony w zbiorze przykładów.

Ze względu na liczbę całkowań najkorzystniejsze są rozkłady wykładnicze lub Erlanga, gdyż obejmują całą półprostą. Jednak w tym przypadku mimo że wzory analityczne są na wejściu bardzo proste to wzory pośrednie na łączne rozkłady pomimo prowadzonej redukcji stają się długie, w efekcie rozkład końcowy również złożony. Gdy rozkłady są skupione na odcinku, jak w przypadku rozkładu jednostajnego, wówczas występuje konieczność rozwiązywania większej ilości całek na alternatywnych przedziałach.

Ograniczenia

Ograniczenia wynikają z dwóch powodów. Po pierwsze natura problemu narzuca wykładniczy charakter złożoności obliczeniowej zależny od liczby zmiennych losowych (liczba zadań) oraz liczby segmentów w funkcjach dystrybuanty, a bardziej dokładnie od liczby niezdominowanych segmentów każdej dystrybuanty przez segmenty innych dystrybuant. Nie trzeba dużej sieci, aby liczba koniecznych różniczkowań i całkowań wyrażała się w dziesiątkach czy setkach tysięcy. Z przeprowadzonych testów wynika, że więcej niż dwadzieścia zmiennych losowych może wiązać się z dłuższym czasem oczekiwania na rozwiązanie. Natomiast przekroczenie ponad trzydziestu zmiennych losowych może powodować przerwanie obliczeń przed znalezieniem rozwiązania. Trudno jest podać jasną regułę, określającą kiedy zasoby serwera są niewystarczające dla rozwiązania danego pro-

blemu. Drugi powód ograniczeń wynika z założenia, że interesuje nas rozwiązanie analityczne. W związku z tym:

- funkcja dystrybuanty musi mieć jawną postać matematyczną, tzn. musi istnieć analityczna możliwość obliczenia $\int_0^{\tau} f(u)du$ na odcinku w którym zmienna losowa ma charakter ciągły, gdzie $f(u)$ oznacza gęstość. Tego warunku nie spełnia np. rozkład normalny czy rozkład beta w ogólnym przypadku. Szczególny przypadek funkcji beta występuje w jednym z zamieszczonych przykładów.
- może się zdarzyć, że oprogramowanie, w tym przypadku procedura MatLab, nie potrafi znaleźć funkcji pierwotnej dla całki, która wynikała w trakcie obliczeń rekurencyjnych lub oprogramowanie jest w stanie znaleźć funkcję pierwotną, ale na komputerze o większych możliwościach obliczeniowych.

6. Uwagi końcowe

Przed wszystkim rozwiązania analityczne są dokładne. W metodach przybliżonych, takich jak PERT czy metoda Monte Carlo, nie jest możliwe dokładne poznanie osobliwości rozwiązania, tzn. skoków funkcji gęstości czy dystrybuanty, co ilustruje zamieszczone przykłady.

Ponadto rozwiązania analityczne dla dominujących podsieci pozwalają na konstruowanie bardzo efektywnych estymatorów dla rozwiązań przybliżonych dla pełnej sieci ([9]). Tę własność wykorzystuje się w metodzie symulacji ze zmiennymi kontrolnymi. Dlatego, zamiast powtarzać obliczenia w dziesiątkach tysięcy razy, można zredukować liczbę symulacji do rzędu kilkuset poprzez stosowanie estymatorów wykorzystujących wzory analityczne.

Wartość metody polega na rekurencyjnym obliczaniu rozkładów łącznych realizacji węzłów, co jest bardziej efektywne niż na podstawie wszystkich ścieżek, jak proponowano we wcześniejszych publikacjach [2, 10].

Literatura

- [1] Dodin B.M., *Reducibility of Stochastic Networks*, Omega, 13, No. 3, 1985, 223-232.
- [2] Elmaghraby S.E., *Activity Networks: Project Planning and Control by Network Models*, J.Wiley & Sons, New York 1997.
- [3] Feller W., *An Introduction to Probability Theory and Its Applications*, Vol. II, 3rd ed. Wiley, 1968.
- [4] Kleindorfer G.B., *Bounding Distributions for a Stochastic Acyclic Network*, Operations Research, 19, 1971, 1586-1601.
- [5] Malcolm D.G., Rosenboom J., Clark C., *Application of a Technique for Research and Development Program Evaluation*, Operations Research, 7, 1959, 646-669.
- [6] Milian Z., *A recursive method for computing the exact probability distribution of the project completion time in a stochastic PERT network* (w druku).

- [7] Milian Z., *Metody określania rozkładu czasu realizacji przedsięwzięć budowlanych w acyklicznych sieciach stochastycznych*, monografia 327, Wyd. Politechniki Krakowskiej, Kraków 2006.
- [8] Milian Z., *Notes on time buffers' estimation in CCPM*, Archives of Civil and Mechanical Engineering, Vol. 5, No. 1, 2005, 19-38.
- [9] Milian Z., *Monte Carlo simulation with exact analysis for stochastic PERT networks*, The 25th International Symposium on Automation and Robotics In Construction, ISARC-2008, 2008, 598-599.
- [10] Schmidt C.W., Grossmann I.E., *The exact overall time distribution of a project with uncertain task durations*, European Journal of Operational Research, 126, 2000, 614-636.