www.biswbis.pb.edu.pl
CIVIL AND ENVIRONMENTAL ENGINEERING 2 (2011) ISSN: 2081-3279 BUDOWNICTWO I INŻYNIERIA ŚRODOWISKA

ZASTOSOWANIE METODY LOSOWANIA LHS W BADANIACH SYMULACYJNYCH MODELI SIECIOWYCH

Sławomir BIRUK*, Piotr JAŚKOWSKI

Wydział Budownictwa i Architektury, Politechnika Lubelska, ul. Nadbystrzycka 40, 20-618 Lublin

Streszczenie: Metoda symulacji cyfrowej *Monte Carlo* jest popularnym narzędziem wykorzystywanym w analizie ryzyka oraz planowaniu przedsięwzięć budowlanych w warunkach losowych. Zaletą metody jest możliwość analizowania modeli sieciowych złożonych z czynności, których czas trwania jest opisany dowolnymi rozkładami prawdopodobieństwa, bez konieczności wprowadzania dodatkowych założeń upraszczających. Podejście to umożliwia także modelowanie dowolnych ograniczeń czasowych, zasobowych i kolejnościowych. W metodzie symulacji *Monte Carlo* zwiększanie liczby przebiegów symulacyjnych wpływa na dokładność estymowanych wielkości. Zmniejszenie „rozrzutu” wartości zmiennych wyjściowych można także uzyskać stosując jedną z metod redukcji wariancji (lub ich kombinację). W artykule przedstawiono wyniki pilotażowych badań symulacyjnych prowadzonych na testowych modelach sieciowych przedsięwzięć budowlanych. Podczas badań symulacyjnych w procesie generowania liczb losowych zastosowano metodę *Latin Hypercube Sampling*. Przeprowadzone eksperymenty mają na celu próbę oszacowania skuteczności redukcji wariancji średniej terminu realizacji przedsięwzięcia za pomocą metody *LHS* oraz zbadanie możliwości poprawy wyników poprzez zastosowanie liczb antytetycznych (losowania przeciwstawnego).

Słowa kluczowe: ryzyko realizacji przedsięwzięć budowlanych, symulacja komputerowa, metody redukcji wariancji.

1. Wprowadzenie

Analiza modeli sieciowych przedsięwzięć metodą symulacji cyfrowej *Monte Carlo* ma na celu najczęściej ustalenie średnich, wariancji lub typu i parametrów rozkładu terminów zaistnienia zdarzeń i ($i = 1, 2, \dots, m$) na podstawie obserwacji $x_{i,j}$ poczynionych w j ($j = 1, 2, \dots, n$) przebiegach symulacyjnych. Wartości $x_{i,j}$ można traktować jako realizacje zmiennych losowych X_i terminów zaistnienia zdarzeń (Platt, 1974).

Istota metody *MC* w analizie modeli sieciowych polega na losowym generowaniu w kolejnych przebiegach symulacyjnych czasów realizacji $t_{l,j}$ czynności ($l = 1, 2, \dots, w$, w jest to liczba czynności modelu sieciowego) i obliczaniu terminów zaistnienia zdarzeń jak w *Critical Path Method* (*CPM*). W każdej replikacji j czasy realizacji czynności są generowane zgodnie z przyjętym rozkładem prawdopodobieństwa, przy zastosowaniu niezależnych od siebie ciągów liczb losowych $u_{l,j}$ z przedziału (0, 1].

Nieobciążony estymator punktowy średniej $\mu_i = E(X_i)$ zmiennej losowej Xi terminu zaistnienia zdarzenia i ma postać (Platt, 1974):

$$\overline{\mu}_i = \frac{1}{n}\sum_{j=1}^{n} x_{i,j} \ , \tag{1}$$

natomiast jego wariancja dana jest wzorem:

$$D^2\left(\overline{\mu}_i\right) = \frac{\sigma_i^2}{n} \ , \tag{2}$$

gdzie: σ_i^2 jest wariancją zmiennej losowej X_i terminu zaistnienia zdarzenia i, tj. $\sigma_i^2 = D^2 \cdot (X_i)$**.** Nieobciążony estymator wariancji jest określony następująco:

$$\overline{\sigma}_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(x_{i,j} - \overline{\mu}_i\right)^2 . \tag{3}$$

W metodzie symulacji *MC* błąd estymacji średniej ($\overline{\mu}_i - \mu_i$) jest przedstawiany w postaci przedziału ufności. Wielkość przedziału ufności jest proporcjonalna do $\sigma_i / \sqrt{n}$. Jednym z podstawowych sposobów zawężania przedziału ufności (redukcji wariancji estymatora średniej) jest zwiększanie liczby obserwacji (przebiegów symulacyjnych), co jednak powoduje wydłużenie czasu badań symulacyjnych.

* Autor odpowiedzialny za korespondencję. E-mail: s.biruk@pollub.pl

Na dokładność wyznaczanych charakterystyk wpływa także sposób generowania zmiennych (danych) wejściowych do badań. Wywołując korelację między wartościami tych samych wielkości w eksperymentach symulacyjnych, uzyskuje się redukcję wariancji estymatora, a co się z tym wiąże zwiększa się dokładność oszacowania. Zawężenie przedziału ufności estymowanej charakterystyki można także uzyskać dokonując właściwego wyboru estymatora.

Postępowania mające na celu zmniejszenie „rozrzutu” obserwowanych wartości zmiennych wyjściowych, w literaturze określa się mianem metod redukcji wariancji (Law i Kelton, 1991; Tyszer, 1990).

W artykule podjęto próbę oceny efektywności metody losowania zgodnego ze schematem *Latin Hypercube Sampling* (*LHS*) oraz jej kombinacji z metodą losowania przeciwstawnego. Analizowane metody nie wydłużają czasu eksperymentów i są łatwe do zaimplementowania w komercyjnych językach symulacyjnych.

2. Metoda losowania *LHS*

Sposób planowania eksperymentów symulacyjnych *Latin Hypercube Sampling* (*LHS*), został zaproponowany w pracy McKay'a i in. (1979). Należy ona do grupy metod losowania (próbkowania) warstwowego, które mają na celu poprawę „równomierności” generowania liczb losowych. W metodach losowania warstwowego dystrybuanty czasu trwania wszystkich zmiennych wejściowych l (l = 1, 2, ..., w) są dzielone na s (l) rozłącznych przedziałów (warstw) $[a_{l.,k}, a_{l.,k+1}]$ takich, że $0 = a_{l,1} < a_{l,2} < ... < a_{l,s(l)-1} = 1$.

Liczba przedziałów nie musi być jednakowa dla każdej czynności (zmiennej czasu losowej trwania procesu), ale ich równa liczba bardzo ułatwia planowanie i prowadzenie badań symulacyjnych. Najczęściej dokonuje się podziału dystrybuanty czasu trwania wszystkich czynności na jednakową liczbę s równych części, tzn. $0 < 1 / s < 2 / s < ... < (s - 1) / s < 1$. Próbę taką nazywa się proporcjonalną, bowiem liczba wygenerowanych w danej warstwie wystąpień (czasów trwania czynności) jest proporcjonalna do prawdopodobieństwa ich wystąpienia.

W ogólnym przypadku, aby dokładnie odzorować zmienną losową będącą wynikiem symulacji, powinno się wylosować $s(1) \cdot s(2) \cdot ... s(w)$ (w jest to liczba czynności modelu sieciowego) liczb losowych, każdą z innej warstwy dla każdej zmiennej wejściowej. Jest to możliwe tylko przy stosunkowo niewielkiej liczbie czynności modelu sieciowego. Burt i Garman (1971) zalecają, aby przeprowadzić s skorelowanych ze sobą eksperymentów. Należy wylosować s niezależnych od siebie liczb losowych $u_{i,1}, u_{i,2}, ..., u_{i,s}$ z przedziału (0, 1] dla każdej ze zmiennych wejściowych i umieścić je w różnych warstwach $u_{l,1} / s < (u_{i,2} + 1) / s < ... < (u_{l,s} + s - 1) / s$. W każdym przebiegu symulacyjnym dla każdej czynności wybiera się jedną warstwę, a kolejność przypisywania warstw do czynności jest losowa. Stosując metodę losowania warstwowego zawsze uzyskuje się redukcję wariancji estymatora w stosunku do wariancji uzyskanej w klasycznej metodzie *Monte Carlo*.

Stosując schemat losowania *Latin Hypercube Sampling*, aby wygenerować s (j = 1, 2, …, s) skorelowanych ze sobą replikacji, liczby losowe należy określać są zgodnie z zależnością (Owen 1998):

$$U_{l,j} = \frac{\pi_l(j) - 1 + U^*_{l,j}}{s} \text{ dla } l = 1, 2, \dots, w \quad \text{oraz } j = 1, 2, \dots, s \tag{4}$$

gdzie: $\pi_1(\cdot)$, $\pi_2(\cdot)$, ..., $\pi_w(\cdot)$ są to permutacje liczb {1, 2, ..., s}, losowanymi z jednakowym prawdopodobieństwem ze zbioru s! takich permutacji, a $\pi_l(j)$ oznacza j-ty element (numer warstwy) w permutacji dla czynności l, $\{U^*_{l,j} : l = 1, 2, ..., s\}$są liczbami losowymi wzajemnie niezależnymi oraz niezależnymi od permutacji $\pi_1(\cdot), \pi_2(\cdot), ..., \pi_w(\cdot)$.

W każdym cyklu r (r = 1, 2, ..., k) badań symulacyjnych, obejmującym s replikacji ustalonych według procedury *LHS*, jest losowana tylko jedna liczba z każdej warstwy dla danej czynności.

Przeprowadzenie n przebiegów symulacyjnych wymaga wygenerowania $k = n / s$ niezależnych permutacji $\pi_1(\cdot), \pi_2(\cdot), ..., \pi_w(\cdot)$. W każdym powtórzeniu są obliczane wartości zmiennych losowych terminów zaistnienia zdarzeń w modelu sieciowym według metody *CPM*.

3. Metoda zmiennych antytetycznych

Metoda ta, nazywana również losowaniem przeciwstawnym, polega na przeprowadzeniu eksperymentu symulacyjnego w dwóch etapach (Law i Kelton, 1991; Tyszer, 1990). W pierwszym etapie stosuje się dla każdej czynności ciągi liczb pseudolosowych $u_{l,1}, u_{l,2}, ..., u_{l,n/2}$ (l = 1, 2, ..., w) o rozkładzie równomiernym na przedziale (0, 1], natomiast w drugim etapie ciągi liczb dopełniających $(1 - u_{l,1}), (1 - u_{l,2}), ..., (1 - u_{l,n/2})$. Jeżeli $\hat{X}_i^{(1)}$ oraz $\hat{X}_i^{(2)}$ są odpowiednio estymatorami parametru $\hat{X}_i$ odpowiednio w pierwszym i drugim etapie badań symulacyjnych, to nieobciążonym estymatorem tego parametru jest średnia:

$$\hat{X}_i = \frac{1}{2}\left(\hat{X}_i^{(1)} + \hat{X}_i^{(2)}\right). \tag{5}$$

Wariancję tego estymatora można obliczyć następująco:

$$D^2\left(\hat{X}_i\right) = D^2\left(\frac{1}{2}\left(\hat{X}_i^{(1)} + \hat{X}_i^{(2)}\right)\right) = \frac{1}{4}\left(D^2\left(\hat{X}_i^{(1)}\right) + D^2\left(\hat{X}_i^{(2)}\right)\right) + \frac{1}{2}\operatorname{cov}\left(\hat{X}_i^{(1)}, \hat{X}_i^{(2)}\right). \tag{6}$$

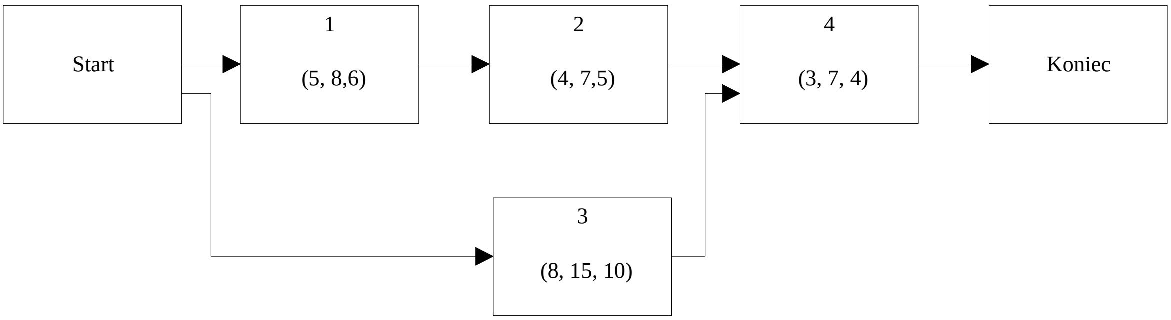
Stosując losowanie przeciwstawne, z ujemną korelacją między wynikami z następujących po sobie eksperymentów, uzyskuje się mniejszą wariancję niż w przypadku prowadzenia eksperymentów niezależnych.

Burt i in. (1970) stosowali ciągi liczb antytetycznych do symulacyjnego rozwiązywania prostych modeli z czasami trwania o rozkładzie wykładniczym. Następnie wykazali analitycznie, że wariancję oczekiwanego czasu realizacji modelowanych przedsięwzięć można zmniejszyć o ponad połowę w stosunku do wariancji estymatora stosowanego w bezpośredniej metodzie *Monte Carlo*. Rozważania te zostały uogólnione przez Sullivana i in. (1982), gdzie badano sieci złożone z czynności o różnych rozkładach. Stosując losowanie przeciwstawne, tę samą dokładność jak w metodzie *Monte Carlo* można uzyskać średnio przy 1/4 liczbie replikacji.

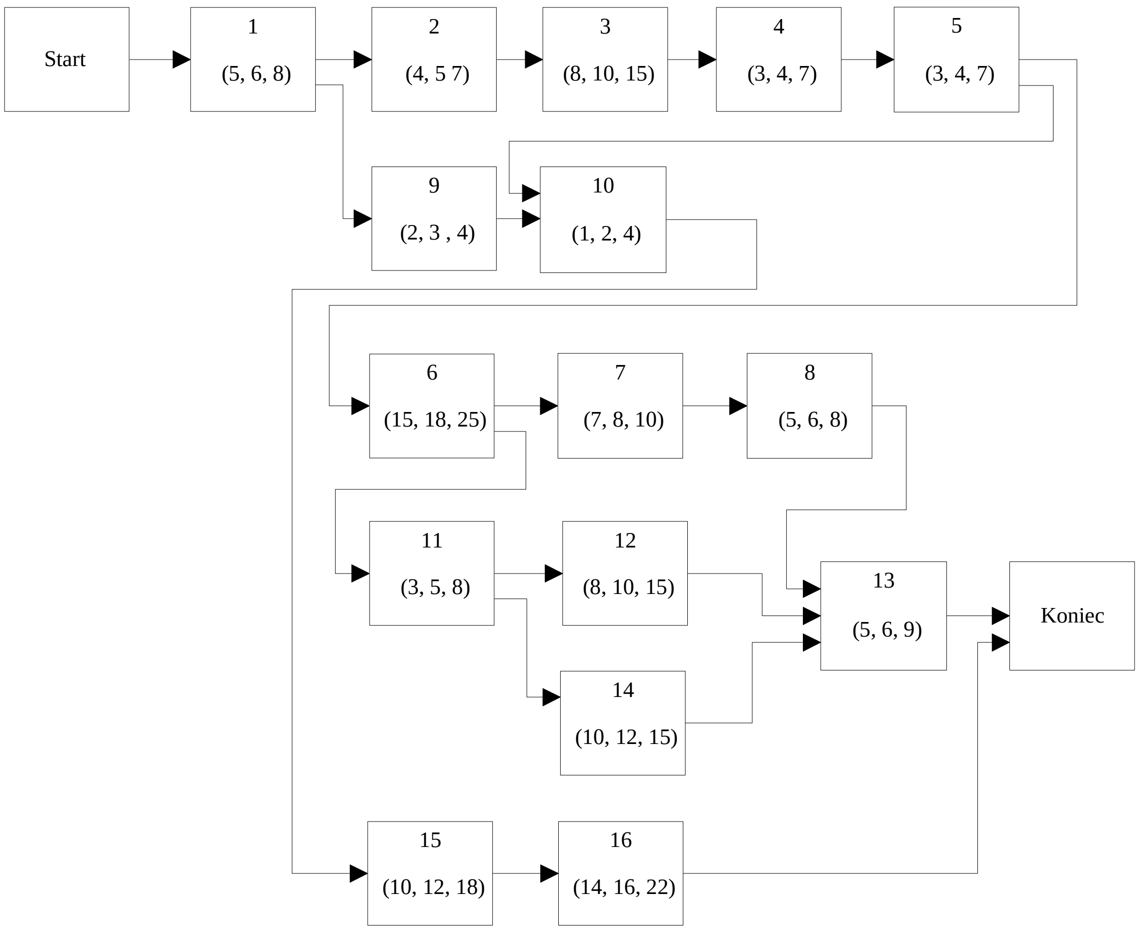
Metoda zmiennych antytetycznych może być również łączona z innymi metodami.

4. Badania symulacyjne

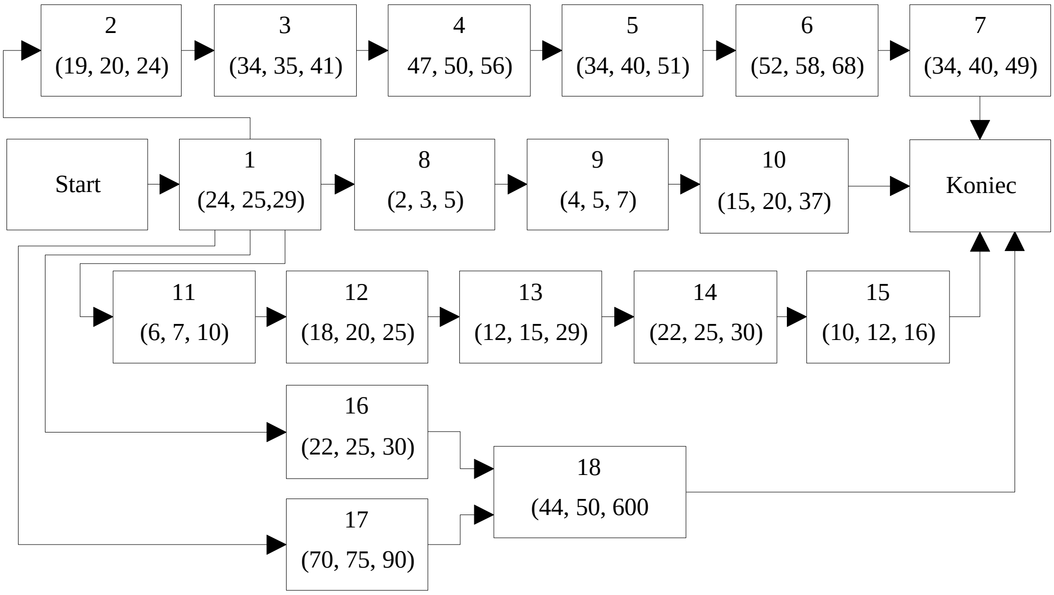
Przeprowadzono badania symulacyjne trzech modeli sieciowych (rys. 1-3), w których czasy wykonania czynności opisano za pomocą rozkładów trójkątnych (czasy podano na rysunkach przy numerze czynności w postaci (*a*, *c*, *b*), gdzie odpowiednio są to czasy optymistyczne, najbardziej prawdopodobne i pesymistyczne). Wszystkie programy symulacyjne zostały opracowane w języku symulacyjnym *GPSS World*. W celu porównywalności ocen skuteczności poszczególnych metod redukcji wariancji, dla każdego wariantu badań symulacyjnych przeprowadzono łącznie 1200 replikacji.



Rys. 1. Model sieciowy I



Rys. 2. Model sieciowy II



Rys. 3. Model sieciowy III

W pierwszym etapie badań symulacyjnych został sporządzony symulator wykorzystujący predefiniowane w języku *GPSS World* generatory rozkładu trójkątnego. Generatory te nie pozwalają jednak na bezpośrednie stosowanie niektórych metod redukcji wariancji, np. strumieni liczb antytetycznych. Dlatego w dalszych wariantach symulatorach zastosowano do generowania rozkładu trójkątnego powszechnie znaną metodę odwracania dystrybuanty.

W metodzie losowania *LHS*, zakres zmienności wszystkich zmiennych losowych został podzielony na 5 warstw o jednakowym prawdopodobieństwie wystąpienia.

Ostatnim etapem badań symulacyjnych było łączne zastosowanie metody losowania przeciwstawnego i *LHS*. Wykonano 600 replikacji zgodnie ze schematem *LHS* stosując ciągi liczb losowych $u_{l,1}$, $u_{l,2}$, ..., $u_{l,n/2}$ oraz 600 stosując ciągi liczb przeciwstawnych $(1 - u_{l,1})$, $(1 - u_{l,2})$, ..., $(1 - u_{l,n/2})$. Wyniki badań symulacyjnych zestawiono w tabeli 1.

Tabela 1. Wyniki uzyskane przy stosowaniu różnych metod wariancji

Metoda prowadzenia badań symulacyjnych	Model	Średnia	Wariancja estymatora
GPSS	Model I	16,758	0,001388
	Model II	76,007	0,010228
	Model III	276,746	0,034991
Metoda odwracania dystrybuanty	Model I	16,701	0,001450
	Model II	76,009	0,010588
	Model III	276,787	0,039716
Metoda zmiennych antytetycznych	Model I	16,752	0,000401
	Model II	75,951	0,000552
	Model III	276,706	0,000571
LHS	Model I	16,748	0,000239
	Model II	75,964	0,001069
	Model III	276,699	0,003244
Łączne stosowanie metody zmiennych antytetycznych i losowania LHS	Model I	16,758	0,000212
	Model II	75,952	0,000248
	Model III	276,668	0,000130

5. Podsumowanie

Metoda symulacji *Monte Carlo* jest efektywnym narzędziem analizy sieci zależności utworzonych z czynności o dowolnych rozkładach czasu ich trwania bez konieczności wprowadzania dodatkowych założeń upraszczających. Wykorzystanie metod redukcji wariancji może zarówno skrócić czas prowadzenia badań (zmniejszenie liczby replikacji), jak i poprawić wiarygodność oszacowania wybranych charakterystyk badanego modelu. Efektywność poszczególnych metod redukcji wariancji zależy przede wszystkim od konfiguracji sieci zależności i stosowanych typów oraz parametrów rozkładów czasu trwania czynności.

Na podstawie przeprowadzonych eksperymentów symulacyjnych można wnioskować, że zastosowanie metody *LHS* prowadzi do znacznej redukcji wariancji. Stosowanie łączne metod *LHS* i losowania przeciwstawnego prowadzi do dalszej redukcji wariancji estymatora średniej.

Literatura

Burt J. M., Garman M. B. (1971). Conditional Monte Carlo: A Simulation Technique for Stochastic Network Analysis. *Management Science*, Vol. 18, No. 3, 207-217.

Burt J. M., Gaver D. P., Perlas M. (1970). Simple Stochastic Networks: Some Problems and Procedures. *Naval Research Logistics Quartely*, Vol. 17, No. 4, 439-459.

Law A. M., Kelton W. D. (1991). Simulation Modeling & Analysis. *McGraw Hill, International Edition*.

McKay M. D,. Beckman R. J., Conover W. J. (1979). A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, Vol. 21, No. 2, 239-245.

Owen A. B. (1998). Latin Supercube Sampling for Very High-Dimensional Simulations. *ACM Transactions on Modeling and Computer Simulation*, Vol. 8, No. 1, 71-102.

Platt C. (1974). Problemy rachunku prawdopodobieństwa i statystyki matematycznej. *PWN*, Warszawa.

Sullivan R. S., Hayya J., C., Schaul R. (1982). Efficienty of the Antithetic Variate Method for Simulating Stochastic Networks. *Management Science*, Vol. 28, No. 5, 573-572.

Tyszer J. (1990). Symulacja cyfrowa. *Wydawnictwa Naukowo-Techniczne*, Warszawa.

ASSESSING EFFICIENCY OF LATIN SUPERCUBE SAMPLING METHOD IN CONSTRUCTION PROJECT NETWORK SIMULATION

Abstract: Monte Carlo simulation is a popular tool that supports planning projects affected by risk. Analysing the results of computer simulations enables the planner to formulate and verify hypotheses on distribution type and parameters of schedule event occurrence and the project duration. Accuracy of estimates obtained by means of simulations can be improved by increasing the number of replications, or by applying variance reduction methods. The latter may consist in change of the way the random numbers are generated. The paper analyses how the method of variance reduction affects simulation results in terms of standard error of estimated project duration mean value. The considered methods were: Latin supercube sampling and its combination with antithetic variates method. The object analysis was based on network models with task durations of triangular distribution. This type of distribution is commonly assumed in modelling the effect of random occurrencies on organisation of construction works.

Praca naukowa finansowana ze środków na naukę w latach 2009-2011 jako projekt badawczy