

Probabilistyczna ocena niezawodności systemu sieci infrastruktury z zastosowaniem estymacji jądrowej i w warunkach luki pomiarowej

Omówienie LEAD

W artykule przedstawiono probabilistyczną metodę oceny niezawodności w warunkach luki pomiarowej związanej z brakiem danych dotyczących awaryjności w długim okresie funkcjonowania sieci infrastruktury. Wykorzystano dane awaryjności magistral sieci wodociągowej. Zastosowano podejście nieparametryczne w postaci estymacji jądrowej z wykorzystaniem estymatora Parzena.

Słowa kluczowe: sieć wodociągowa, luka pomiarowa, niezawodność, estymacja jądrowa, estymator Parzena.

1. Wstęp

Ocena probabilistycznej niezawodności systemów sieci infrastruktury ma złożony charakter, między innymi na skutek luk pomiarowych awarii obejmujących często długi okres funkcjonowania systemu oraz problematycznej estymacji z uwagi na oczywistą nieznaną rozkładu funkcji gęstości prawdopodobieństwa w populacji generalnej.

Niniejsze opracowanie stanowi wstępną propozycję alternatywnego podejścia do oceny niezawodności uwzględniającego ograniczone informacje o awariach sieci infrastruktury. Na podstawie awaryjności magistral wybranej sieci wodociągowej, czyli awaryjności rurociągów o średnicy powyżej 250 mm, wykonano przykładowe obliczenia według opracowanej metody. Wnioski nasuwające się z pracy wskazują na możliwości badawcze, co stwarza szansę wymiernych skutków w aspekcie zastosowań praktycznych.

2. Niezawodność systemu

W literaturze pojęcie niezawodności odnosi się najczęściej do obiektu [Bobrowski 1985, Szopa 1999]. Pojęcie systemu jako zbioru obiektów oraz relacji zachodzących pomiędzy tymi obiektami i systemem jako całością, jest przedmio-

tem niniejszego opracowania. Można za Arystotelesem przyjąć tezę, że „całość” rozumiana dzisiaj jako system „...nie stanowi prostej sumy części, zaś o istocie tak rozumianej całości przesądzają relacje między jej częściami ...” [Wawrzyniak 1987]. System jest więc przedmiotem badań bez względu na to, czy stanowi fragment większego systemu, czy też jest systemem większym. Bez względu na to, czy systemem jest sieć wodociągowa, czy kanalizacyjna stanowiące podsystem infrastruktury budowlanej, sieć energetyczna czy też inny system, istnieje **pojęcie niezawodności systemu** oznaczające **zdolność danego systemu do funkcjonowania** bez niesprawności lub po prostu **bez awarii**. Przeciwnościem niezawodności jest oczywiście zawodność, czyli brak zdolności do bezawaryjnego funkcjonowania. Tak określaną niezawodność oraz zawodność określa się w kategoriach probabilistycznych obiektu [Bobrowski 1985, Szopa 1999].

Przyjmując umowny początek okresu działania dowolnego systemu jako chwilę to w pewnej chwili $t > t_0$, można zaobserwować wystąpienie awarii. Z uwagi na brak możliwości przewidzenia momentu awarii traktuje się go jako wielkość losową, wyrażoną w niniejszym opracowaniu jako zmienną losową T posiadającą wymiar czasu. Dystrybuanta $F(t)$ zmiennej losowej T :

$$F(t) = P(T \leq t) \quad (2.1)$$

określa prawdopodobieństwo zdarzenia, że awaria systemu nastąpi do chwili t charakteryzuje zawodność systemu i stanowi **funkcję zawodności**. Prawdopodobieństwo zdarzenia, że awaria nastąpi po chwili t określa tzw. **funkcja niezawodności** $p(t)$, czyli funkcja prawdopodobieństwa przewyższenia, przy czym zachodzi relacja:

$$p(t) = 1 - F(t) = P(T > t) \quad (2.2)$$

Zmiennej losowej T typu ciągłego i założonej niezawodności p odpowiada wartość t_p , która stanowi **okres bezawaryjnej pracy systemu**. Wartość t_p dla zadanego p otrzymuje się w wyniku rozwiązania równania:

$$\int_{t_p}^{\infty} f(t) dt = p, \quad (2.3)$$

w którym $f(t)$ jest funkcją gęstości prawdopodobieństwa/niezawodności. Z uwagi na fakt, że:

$$f(t) = \frac{dF(t)}{dt} = -\frac{dp(t)}{dt} \quad (2.4)$$

$f(t)$ interpretowana jest jako *spadek niezawodności lub wzrost zawodności w nieskończenie małym przedziale czasu*.

Intensywność awarii wyraża się w postaci [Bobrowski 1985, Szopa 1999]:

$$(t) \frac{f(t)}{p(t)} \quad (2.5)$$

Do określenia powyższych funkcyjnych charakterystyk niezawodności systemu, potrzebna jest znajomość spadku niezawodności lub wzrost zawodności w nieskończenie małym przedziale czasu, czyli znajomość funkcji gęstości prawdopodobieństwa w populacji generalnej, której reprezentację stanowi próba losowa zawierająca obserwacje długości okresu bezawaryjnego funkcjonowania rozważanego systemu sieci. Jednak postać funkcji gęstości prawdopodobieństwa w populacji generalnej nie jest znana. W związku z tym określa się przybliżoną postać powyższej funkcji, stosując jedno z dwóch alternatywnych podejść – podejście parametryczne lub nieparametryczne.

W przypadku parametrycznego podejścia, będącego podejściem apriorycznym, w odniesieniu do populacji generalnej zakłada się różne funkcje gęstości rozkładu prawdopodobieństwa i na podstawie próby losowej estymuje się parametry, a następnie w różny sposób bada się zgodność tych funkcji z rozkładem prawdopodobieństwa w próbie losowej [np. Karczmarek 1970] lub stosuje się kryterium selekcyjne [Akaike 1974] do wyboru funkcji gęstości najlepszej z przyjętej klasy.

Temat niniejszego artykułu związany jest z estymacją jądrową stanowiącą podejście nieparametryczne i nieaprioryczne, w ramach którego można zastosować jądrowy estymator funkcji gęstości prawdopodobieństwa [Feluch 1987, 1994]. Na podstawie próby losowej uzyskuje się przybliżenie rozkładu prawdopodobieństwa zdarzenia w populacji generalnej bez potrzeby dokonywania jakichkolwiek założeń co do kształtu funkcji gęstości prawdopodobieństwa ponad to, że próba losowa jest „dobrą” reprezentacją populacji generalnej. Można zauważyć, że gdyby próba losowa „źle” reprezentowała populację generalną, o czym najczęściej nie wiadomo i raczej dotychczas nie są znane metody badawcze pozwalające to ocenić, wówczas jakiegokolwiek podejście probabilistyczne nie miałyby sensu w kategoriach wiarygodnej oceny wszelkich charakterystyk statystycznych.

Z uwagi na problematyczność założeń dotyczących kształtu funkcji gęstości prawdopodobieństwa w przypadku oceny niezawodności sieci infrastruktury przy niepełnej ewidencji awarii sieci, podejście nieparametryczne wydaje się najbardziej uzasadnione.

3. Estymacja jądrowa

Zgodnie z powyższym, w estymacji jądrowej kształt funkcji gęstości prawdopodobieństwa nie jest zakładany a priori tak jak w metodach parametrycznych, lecz uzyskiwany w postaci estymatora na podstawie danych obserwacyjnych.

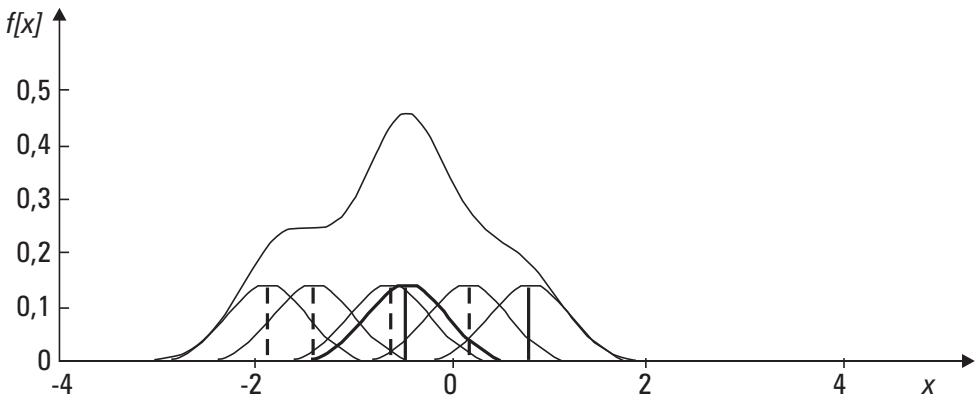
Każdej obserwacji N -elementowej próby losowej, przypisuje się określoną funkcję $K(y)$ zwaną funkcją jądra, spełniającą warunki [Parzen 1962, Feluch 1994]:

$$\sup_y |K(y)| \quad (3.1)$$

$$K(y) \geq 0 \quad (3.2)$$

$$\int K(y) dy = 1 \quad (3.3)$$

Powyższe warunki określają, że funkcją jądra może być każda funkcja przyjmująca wartości skończone (3.1), dodatnio określona (3.2) oraz całkowalna zgodnie z warunkiem (3.3). Estymator funkcji gęstości jest zależny od rozłożonych nad obserwacjami „jąder” (rys. 3.1)



Rys. 3.1. Ilustracja estymacji funkcji gęstości przy wykorzystaniu estymatora Parzena

Źródło: [Silverman 1986].

Typowym przedstawicielem estymatorów jądrowych jest estymator Parzena. Dla N -elementowej prostej próby losowej stanowiącej realizację zmiennej losowej X , estymator $\hat{f}(x)$ nieznanej funkcji $f(x)$ gęstości prawdopodobieństwa ma postać [Parzen 1962]:

$$\hat{f}(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x - X_i}{h}\right) \quad (3.4)$$

gdzie:

X_i – i -ty element N -elementowej próby losowej $\{X_i\}_{i=1}^N$

h – współczynnik gładkości zależny od liczebności próby (N) oraz od wartości jej elementów.

Bez względu na kształt funkcji jądra spełniającej warunki (3.1)–(3.3), estymator Parzena ma tę własność [Greblicki 1974], że przy wzrastającej liczebności (N) próby zachodzi relacja:

$$\lim_N E[\hat{f}(x)] = f(x) \tag{3.5}$$

gdzie: $E[\hat{f}(x)]$ – wartość średnia.

Powyższa własność oznacza asymptotyczną nieobciążoność estymatora. Z praktycznego punktu widzenia jest istotne, aby zbieżność ta była jak najszybsza.

Średni scałkowany błąd średniokwadratowy estymatora (3.4), umożliwiający globalną ocenę estymatora, ma postać [Parzen 1962]:

$$MISE = E[(\hat{f}(x) - f(x))^2] \tag{3.6}$$

Zgodnie z pracami Parzena [1962], Tapia i Thomsona [1978], a także Silvermana [1986] błąd MISE estymatora Parzena (3.4) ma postać:

$$MISE = \frac{1}{Nh} \int K^2(y) dy + \frac{h^2}{4} \int y^2 K(y) dy + f^{(2)}(x)^2 \tag{3.7}$$

Znak „ \approx ” oznacza równość asymptotyczną, tzn. z pominięciem reszty, zaś $f^{(2)}(x)$ jest drugą pochodną nieznannej funkcji gęstości prawdopodobieństwa odpowiadającej populacji generalnej.

W zależności (3.7) można zauważyć, że postać funkcji jądra rzutuje na wartości MISE. Na podstawie rachunku wariacyjnego minimalizując błąd MISE (3.7) uzyskuje się optymalną postać funkcji jądra [Epanechnikov 1969, Feluch 1994] w postaci:

$$K_{opt}(y) = \begin{cases} \frac{3}{4}(5b)^{\frac{1}{2}} \left(1 - \frac{y^2}{5b}\right) & \text{dla } |y| \leq 5b^{\frac{1}{2}} \\ 0 & \text{dla } |y| > 5b^{\frac{1}{2}} \end{cases} \tag{3.8}$$

przy czym b oznacza wariancję oraz:

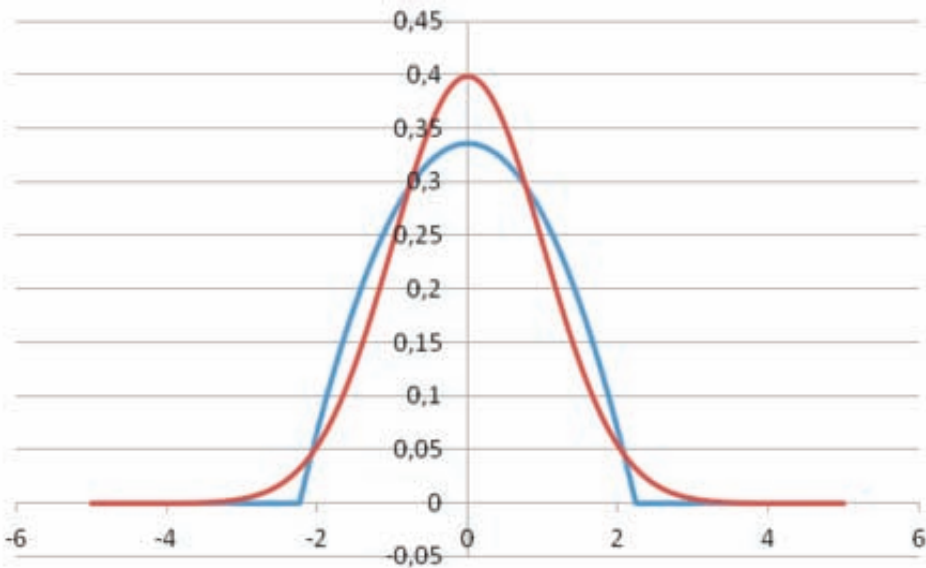
$$b = \int y^2 K_{opt}(y) dy \tag{3.9}$$

Powyższą postać (3.8) dla jednostkowej wariancji tj. $b = 1$ rozważał [Epanechnikov 1969].

W zastosowaniach praktycznych stosuje się również gaussowską funkcję jądra w postaci:

$$K_{\text{gauss}}(y) = \frac{1}{\sqrt{2}} e^{-\frac{y^2}{2}} \quad (3.10)$$

Powyższa gaussowska funkcja jądra różni się od optymalnej zaledwie pięć-procentową różnicą błędu MISE [Feluch 1994], co z praktycznego punktu widzenia jest wielkością możliwą do zaniedbania. Graficzna ilustracja optymalnej i gaussowskiej funkcji jądra (rys. 3.2) pozwala zauważyć niewielkie różnice w kształcie tych funkcji mało istotne z punktu widzenia rozważanej estymacji jądrowej.



Rys. 3.2. Ilustracja kształtów funkcji jądra:

- optymalnej tj. dla $b=1$ w wyrażeniu (3.8) – linia niebieska
- gaussowskiej (3.10) – linia czerwona.

Możliwość wykorzystania estymatora Parzena (3.4) stwarza potrzebę odpowiedniego doboru współczynnika gładkości. Pod tym kątem, analizując zależność (3.7) można zauważyć, że błąd MISE składa się z dwóch składników, z których pierwszy stanowi obciążenie estymatora, zaś drugi jego wariancję. Wzrost współczynnika gładkości spowoduje wzrost kwadratu obciążenia estymatora, tj. drugiego składnika zależności (3.7) oraz obniżenie wariancji, czyli pierwszego składnika (3.7). Objawiać się to będzie nadmiernym wygładzaniem przebiegu funkcji przybliżającej gęstość prawdopodobieństwa/niezawodności.

Spadek wartości współczynnika gładkości spowoduje obniżenie kwadratu obciążenia i wzrost wariancji estymatora, co z kolei spowoduje nadmierną oscylację przebiegu funkcji przybliżającej.

Łatwo więc wyobrazić sobie istnienie optymalnej wartości współczynnika gładkości, przy której wariancja i kwadrat obciążenia osiągają wartości, dla których MISE osiąga minimum.

Formalizacją powyższego opisu jest minimalizacja MISE względem współczynnika gładkości h . Zerując pierwszą pochodną wyrażenia (3.7) względem h , uzyskuje się minimum MISE [Parzen, 1962] dla:

$$h = (K) (f) N^{\frac{1}{5}} \tag{3.11}$$

gdzie:

$$(K) = \frac{\int_{-\infty}^{\infty} K^2(y) dy}{\int_{-\infty}^{\infty} y^2 K(y) dy} \tag{3.12}$$

oraz:

$$(f) = \int_{-\infty}^{\infty} f^{(2)}(x)^2 dx \tag{3.13}$$

Jakkolwiek dla danej funkcji jądra wyrażenie (3.12) jest jednoznacznie określone, niezależnie od populacji generalnej, a tym bardziej od próby losowej, to w zależności (3.13) występuje druga pochodna nieznannej gęstości prawdopodobieństwa odpowiadającej populacji generalnej $f(\cdot)$. Dla praktycznego wykorzystania równania (3.11) istnieje potrzeba określenia estymatora drugiej pochodnej $f^{(2)}(x)$.

Zakładając, że estymator Parzena (3.4) jest dwukrotnie różniczkowalny, za estymator drugiej pochodnej $f^{(2)}(x)$ można przyjąć wyrażenie [Feluch 1994]:

$$\hat{f}^{(2)}(x) = \frac{1}{Nh^3} \sum_{i=1}^N K^{(2)}\left(\frac{x - X_i}{h}\right) \tag{3.14}$$

gdzie: $K^{(2)}[\cdot]$ oznacza drugą pochodną optymalnej funkcji jądra.

Na podstawie powyższych równań można określać współczynnik gładkości dla optymalnej funkcji jądra h_{opt} w postaci [Feluch 1987; Adamowski, Feluch 1987]:

$$h_{opt} = \frac{w_n}{N} \frac{10}{3} \frac{1}{\sqrt{5b}} \tag{3.15}$$

gdzie:

$$w_n = \frac{1}{N} \sum_{i=1}^N (2i - N - 1)x_i \quad (3.16)$$

przy próbie losowej $\{X_i\}_{i=1}^N$ uporządkowanej w ciąg niemalejący.

Relacja (3.15) stanowi górne ograniczenie współczynnika gładkości, dlatego teoretycznie nie szacuje się błędu tego wyrażenia, a także błędu estymatora drugiej pochodnej $f^{(2)}(x)$, traktując je jako elementy estymatora gęstości $\hat{f}(x)$. Ponadto, w przypadku mało liczebnych prób losowych, przyjmując górne ograniczenie współczynnika gładkości zapewniana jest niwelacja ewentualnych oscylacji funkcji przybliżającej. Z uwagi na minimalizację błędu MISE, takie oszacowanie współczynnika gładkości należy oczywiście traktować jako quasi-optymalne.

W celu określenia współczynnika gładkości dla estymatora Parzena i gaussowskiej funkcji jądra, na podstawie zależności (3.11) można zauważyć, że:

$$h_{opt} = (K_{opt}) (f) N^{-\frac{1}{5}} \quad (3.17)$$

oraz:

$$h_{gauss} = (K_{gauss}) (f) N^{-\frac{1}{5}} \quad (3.18)$$

Stosunek obydwu wyrażeń (3.18) i (3.17) prowadzi do zależności:

$$\frac{h_{gauss}}{h_{opt}} = \frac{(K_{gauss})}{(K_{opt})} \quad (3.19)$$

stąd:

$$h_{gauss} = h_{opt} \frac{(K_{gauss})}{(K_{opt})} \quad (3.20)$$

Można wykazać, że:

$$(K_{gauss}) = \frac{1}{2\sqrt{e}}^{1/5} \quad (3.21)$$

oraz dla optymalnej funkcji jądra i jednostkowej wariancji ($b=1$)

$$(K_{opt}) = \frac{3}{5\sqrt{5}}^{1/5} \quad (3.22)$$

Ponadto, jeżeli przyjmie się, że równania (3.15) i (3.16) stanowią podstawę oceny h_{opt} , wówczas wyrażenie (3.20) z uwzględnieniem (3.21) i (3.22) przybiera postać:

$$h_{gauss} = \frac{1}{30\sqrt{}}^{1/5} \frac{w_n}{N N \frac{10}{3}} \tag{3.23}$$

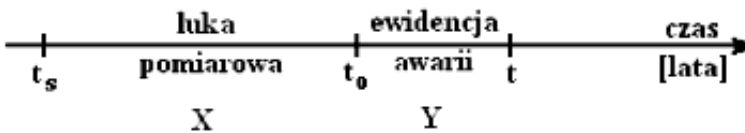
Wartość h_{gauss} uzyskana z wyrażenia (3.23) niewiele różni się od wartości h_{opt} uzyskanej z wyrażenia (3.15), gdyż po podzieleniu stronami relacji (3.23) i (3.15) uzyskuje się, że:

$$h_{gauss} = \frac{5}{6} \sqrt{\frac{5}{}}^{1/5} h_{opt} = 1,0100567 h_{opt} \tag{3.24}$$

Wartość współczynnika gładkości dla gaussowskiej funkcji jądra h_{gauss} uzyskana w powyższy sposób, jest większa o około 1% od wartości uzyskanej dla optymalnej funkcji jądra h_{opt} .

4. Luka pomiarowa w ewidencji awarii

W dziennikach pomiarowych, na przykład sieci wodociągowej czy kanalizacyjnej, oprócz okresu ewidencji awarii zauważa się czasami nawet znacznie dłuższy okres braku ewidencji, który można nazwać **luką pomiarową** (rys. 4.1). W okresie tym brakuje informacji o jakichkolwiek awariach. Może więc zaistnieć sytuacja, że w luce będącej pewnym odcinkiem czasu występowały awarie systemu sieci, które na skutek braku rejestracji nie zostały uwzględnione w ocenie niezawodności sieci. Od długości istniejącej luki i okresu ewidencji awarii zależy więc probabilistyczna ocena niezawodności sieci.



Rys. 4.1. Luka pomiarowa i okres ewidencji awarii; t_s – moment uruchomienia systemu, t_0 – moment rozpoczęcia ewidencjonowania awarii, t – moment awarii, X i Y – zmienne losowe odpowiadające odpowiednio luce pomiarowej i okresowi ewidencji.

Na rys. 4.1 zauważyć można, że dla danego systemu sieci infrastruktury jedyną stałą wielkością przy analizie awarii jest moment rozpoczęcia ewidencji awarii (t_0). Pozostałe wielkości, zarówno moment rozpoczęcia eksploatacji sieci (t_s), jak i moment wystąpienia awarii (t) mają charakter losowy. Można więc przyjąć, że luce pomiarowej odpowiadać będzie zmienna losowa X, której wartości stanowią odcinki czasu niezawodnej pracy systemu pomiędzy zmiennym (t_s)

i stałym t_0 , zaś okresowi ewidencji awarii – zmienna losowa Y o wartościach pomiędzy stałym t_0 i zmiennym t_s . Obydwie zmienne losowe można potraktować jako niezależne i posiadające ciągłe rozkłady. Podstawą oceny niezawodności może być suma tych dwóch zmiennych losowych, ponieważ sieć pracuje od momentu t_s do chwili t_0 :

$$Z = X + Y \quad (4.1)$$

przy czym:

X jest zmienną losową odpowiadającą luce pomiarowej o gęstości prawdopodobieństwa $f_x(x)$;

Y jest zmienną losową odpowiadającą okresowi ewidencji awarii o gęstości prawdopodobieństwa $f_y(y)$;

Z jest zmienną losową odpowiadającą okresowi bezawaryjnej pracy systemu o gęstości prawdopodobieństwa $f_z(z)$.

Dystrybuanta zmiennej losowej stanowiącej sumę dwóch niezależnych zmiennych losowych (4.1) w obszarze całkowania S ma postać:

$$F(z) = \int_S f_x(x)f_y(y)dx dy = \int_0^z f_x(x)dx \int_{z-x}^{\infty} f_y(y)dy = \int_0^z f_y(y)dy \int_{z-y}^{\infty} f_x(x)dx \quad (4.2)$$

zaś funkcja gęstości prawdopodobieństwa stanowi pochodną dystrybuanty co wobec powyższego, funkcja gęstości prawdopodobieństwa zmiennej losowej Z stanowi spłot dwóch funkcji gęstości:

$$f(z) = \int_0^z f_x(z-y)f_y(y)dy \quad (4.3)$$

lub:

$$f(z) = \int_0^z f_x(x)f_y(z-x)dx \quad (4.4)$$

Powyższe wyrażenia są tożsamościowe. Wybierając do oceny funkcji $f(z)$ wyrażenie (4.3) oraz stosując estymator Parzena (3.4) z wykorzystaniem niezależnych prób losowych odpowiadającym zmiennym X i Y w postaci:

$$X_i \sim N_x, \quad Y_j \sim N_y \quad (4.5)$$

uzyskuje się:

$$\hat{f}(z) = \frac{1}{N_x N_y h_x h_y} \prod_{i=1}^{N_x} K \left(\frac{z-y}{h_x} - \frac{X_i}{h_x} \right) \prod_{j=1}^{N_y} K \left(\frac{y}{h_y} - \frac{Y_j}{h_y} \right) dy \quad (4.6)$$

przy czym:

N_x, N_y – liczba lat pracy systemu odpowiednio w czasie luki pomiarowej i okresu ewidencji do zaobserwowanej awarii, tj. liczba elementów zbioru wartości zmiennych losowych X i Y ,

h_x, h_y – współczynniki gładkości odpowiadające luce pomiarowej (X) i okresowi ewidencji (Y).

Po przemnożeniu sum w wyrażeniu (4.6) można zapisać:

$$\hat{f}(z) = \frac{1}{N_x N_y h_x h_y} \prod_{i=1}^{N_x} \prod_{j=1}^{N_y} K \left(\frac{z - X_i}{h_x} \right) K \left(\frac{z - Y_j}{h_y} \right) dy \quad (4.7)$$

W przypadku gaussowskiej funkcji jądra (3.10) po scałkowaniu powyższe równanie (4.7) przybiera postać:

$$\hat{f}(z) = \frac{1}{N_x N_y \sqrt{2 h_x^2 h_y^2}} \prod_{i=1}^{N_x} \prod_{j=1}^{N_y} \exp \left(-\frac{(z - X_i - Y_j)^2}{2(h_x^2 + h_y^2)} \right) \quad (4.8)$$

co jest równoważne wyrażeniu:

$$\hat{f}(z) = \frac{1}{N_x N_y h} \prod_{i=1}^{N_x} \prod_{j=1}^{N_y} K_{gauss} \left(\frac{z - X_i - Y_j}{h} \right) \quad (4.9)$$

gdzie:

$$h = \sqrt{h_x^2 + h_y^2} \quad (4.10)$$

przy czym współczynniki gładkości h_x i h_y można wyznaczyć, wykorzystując zależność (3.23) w odniesieniu do prób losowych (4.5) odpowiadającym zmiennym X i Y , a następnie współczynnik h określony równaniem (4.10).

Estymator funkcji prawdopodobieństwa przewyższenia stanowiący **miarę niezawodności** przedstawia się następująco:

$$\hat{p}(z) = \int_z^{\infty} \hat{f}(y) dy \quad (4.11)$$

stąd:

$$\hat{p}(z) = \frac{1}{N_x N_y} \prod_{i=1}^{N_x} \prod_{j=1}^{N_y} \frac{X_i - Y_j - z}{h} \quad (4.12)$$

gdzie:

$$(u) = \frac{1}{\sqrt{2}} \exp \left(-\frac{u^2}{2} \right) \quad (4.13)$$

Estymator intensywności awarii zgodnie z równaniem (2.5) można zapisać w postaci:

$$\hat{\lambda}(z) = \frac{\hat{f}(z)}{\hat{p}(z)} \quad (4.14)$$

5. Przykład obliczeniowy

Na podstawie dostępnych z lat 1990–2002 rejestracji awarii magistral sieci wodociągowej jednego ze średniej wielkości miast polskich przygotowano dwa 75-elementowe zbiory danych, w których:

- pierwszy zawierał wydzielenie okresów dotyczących luk pomiarowych i ewidencji awarii zgodnie z rys. 4.1 co umożliwiło utworzenie dwóch prób losowych typu (4.5) i wykorzystanie zależności (4.12) do oceny niezawodności, (4.9) do oceny funkcji gęstości oraz (4.14) do oceny intensywności awarii; uzyskane wyniki obliczeń opisane nazwą „splot” (linie czarne) przedstawiono w powyższej kolejności na rys. 5.1, rys. 5.2 i rys. 5.3,
- drugi zbiór 75 danych stanowił jedną próbę losową w postaci:

$$T_{i \ i \ 1}^N \quad (5.1)$$

gdzie N stanowi liczbę lat do pierwszej awarii, licząc od początkowego roku uruchomienia magistral wyżej wspomnianej sieci wodociągowej. Zbioru (5.1) nie rozbito na dwie części i nie wydzielono okresów luk pomiarowych i ewidencji awarii. Do obliczeń przyjęto w przypadku oceny niezawodności zależność:

$$\hat{p}(t) = \frac{1}{N} \prod_{i=1}^N \frac{t - T_i}{h_{t \text{ gauss}}} \quad (5.2)$$

gdzie:

T_i – i -ty element N -elementowej próby losowej,

$h_{t \text{ gauss}}$ – współczynnik gładkości zależny od liczebności (N) próby $T_{i \ i \ 1}^N$ oraz od wartości jej elementów określony na podstawie równania (3.23).

Estymator funkcji gęstości prawdopodobieństwa miał postać:

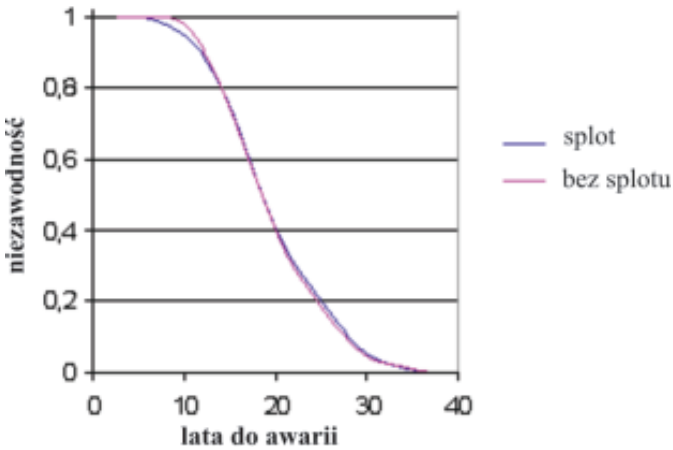
$$\hat{f}(t) = \frac{1}{N} \prod_{i=1}^N K_{\text{gauss}} \left(\frac{t - T_i}{h_{t \text{ gauss}}} \right) \quad (5.3)$$

gdzie:

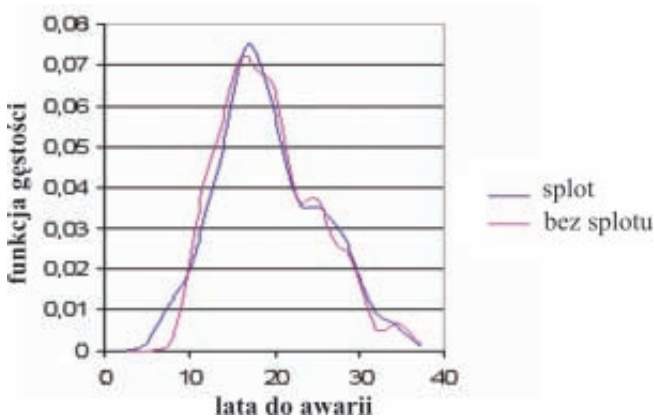
$K_{\text{gauss}} [.]$ – gaussowska funkcja jądra określona zależnością (3.10).

Estymator intensywności awarii przyjęty do obliczeń pozostał zgodny z postacią (4.14), ale w tym przypadku z wykorzystaniem zależności (5.2) i (5.3).

Uzyskane wyniki obliczeń opisane nazwą „bez splotu” przedstawiono w powyższej kolejności na rys. 5.1, rys. 5.2 i rys. 5.3.

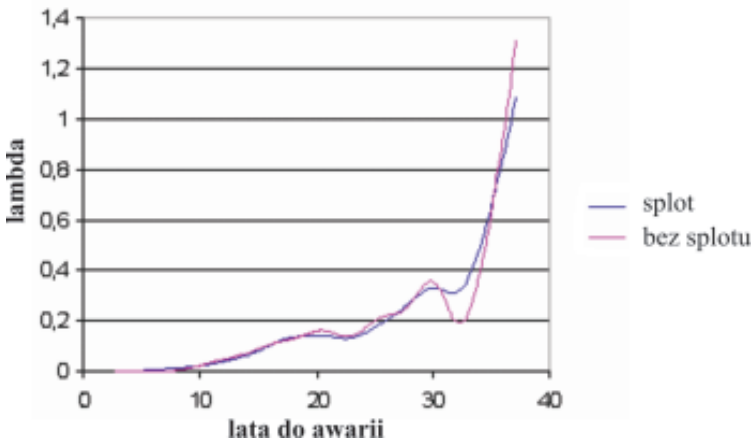


Rys. 5.1. Porównanie niezawodności w zależności od czasu pracy systemu magistral sieci wodociągowej do chwili zarejestrowania awarii w dwóch rozważanych przypadkach – z uwzględnieniem splotu i bez splotu



Rys. 5.2. Porównanie funkcji gęstości w zależności od czasu pracy systemu magistral sieci wodociągowej do chwili zarejestrowania awarii w dwóch rozważanych przypadkach – z uwzględnieniem splotu i bez splotu

Porównując uzyskane wyniki obliczeń niezawodności, funkcji gęstości i intensywności awarii dla dwóch probabilistycznych modeli (4.8 – 4.14) oraz (5.2, 5.3) i dwóch prób losowych (4.5) oraz (5.1) zauważa się bardziej „wyrównany” przebieg powyższych charakterystyk funkcyjnych w przypadku modelu splotowego.



Rys. 5.3. Porównanie intensywności awarii ($\lambda(t)$) w zależności od czasu pracy systemu magistral sieci wodociągowej do chwili zarejestrowania awarii w dwóch rozważanych przypadkach – z uwzględnieniem splotu i bez splotu

Rozbieżności obydwu podejść w ocenie niezawodności są niewielkie (rys. 5.1). Dotyczą głównie początkowego okresu lat do awarii oraz okresu pojawienia się wyraźnego przegięcia w funkcji gęstości (rys. 5.2). Różnice w ocenie niezawodności występują głównie w przypadku, gdy funkcja gęstości dla wariantu „bez splotu” osiąga wartości niższe niż w przypadku „splotu” (rys. 5.1).

Największe różnice występują w przypadku oceny intensywności awarii na końcu obszaru ewidencji awarii i w obszarze ekstrapolacji (rys. 5.3) – obszar powyżej trzydziestu lat.

6. Dyskusja

Przedstawiona metoda oceny niezawodności, a także charakterystyk funkcyjnych, takich jak funkcja gęstości oraz funkcja niezawodności, posiada wiele zalet wynikających z podejścia nieparametrycznego do estymacji, a mianowicie estymacji jądrowej. Najważniejszy przy tym jest fakt, że kształt funkcji gęstości prawdopodobieństwa/niezawodności nie trzeba przyjmować a priori tak jak w podejściu parametrycznym, tylko uzyskuje się na podstawie próby losowej, co do której milcząco przyjmujemy, że stanowi „dobrą” reprezentację populacji generalnej.

Wydzielenie początku okresu ewidencji awarii t_0 (rys. 4.1) i utworzenie dwóch prób losowych (4.5) pozwala na uwzględnienie dodatkowej informacji o awariach w porównaniu z podejściem podanym w przykładzie obliczeniowym (5.2 do 5.3). Metoda oceny niezawodności i rozważanych charakterystyk funkcyjnych bazuje na splotcie dwóch funkcji gęstości co jest zgodne z rzeczywistymi uwarunkowaniami pomiarowymi awarii przykładowej sieci infrastruktury. Dzięki temu można oczekiwać, że na podstawie prezentowanej metody (4.8–4.14) uzyska się większą adekwatność w stosunku do rzeczywistej i nieznannej niezawodności

przykładowej infrastruktury w porównaniu z przypadkiem opisanym wzorami (5.2 do 5.3). Tym niemniej wielkość luki pomiarowej ma nierozpoznany wpływ zawyżający ocenę niezawodności. Wynika to oczywiście z braku danych pomiarowych dotyczących awarii jakie mogły wystąpić w okresie luki pomiarowej.

Wobec powyższego, nasuwa się spostrzeżenie, że powyższe metody mogą być zweryfikowane techniką typu Monte Carlo podobnie jak w pracy [Feluch 1994]; zakładając określoną postać funkcji gęstości hipotetycznego systemu można wygenerować „próby” losowe okresów bezawaryjnej pracy i badać wpływ różnych długości luki pomiarowej na ocenę niezawodności. Podejście takie pozwoli na określenie zmienności błędu oceny niezawodności i innych charakterystyk niezawodności w zależności od wielkości luki pomiarowej.

Literatura

- [1] Akaike H.: A New look At the statistical model identification. *IEEE Transactions on Automatic Control* 1974 26(2): 358–375.
- [2] Adamowski K., Feluch W.: Porównanie metod parametrycznych i nieparametrycznych obliczania przepływów maksymalnych rocznych określonym prawdopodobieństwie przewyższenia. *Wiadomości IMGW* 1987, Zeszyt 2–3, Warszawa, 67–78.
- [3] Bobrowski D.: Modele i metody matematyczne teorii niezawodności w przykładach i zadaniach. WNT, Warszawa 1985.
- [4] Epanechnikov V.A.: Nonparametric Estimates of Multivariate Probability Density. *Theory of Probability and Its Applications* 1969, 14, 153–158.
- [5] Feluch W.: Nieparametryczna estymacja funkcji gęstości na przykładzie przepływów maksymalnych rocznych. Wydanie: VI Seminarium Naukowe Instytutu Inżynierii Środowiska PW, Warszawa 1987, 21–37.
- [6] Feluch W.: Wybrane metody jądrowej estymacji funkcji gęstości prawdopodobieństwa i regresji w hydrologii. *Prace Naukowe, Inżynieria Środowiska* 1994, z. 15. Oficyna Wydawnicza Politechniki Warszawskiej.
- [7] Greblicki W.: Asymptotycznie optymalne algorytmy rozpoznawania i identyfikacji w warunkach probabilistycznych. *Prace Naukowe Instytutu Cybernetyki Technicznej Politechniki Wrocławskiej* 1974, nr 18.
- [8] Karczmarek Z.: Metody statystyczne w hydrologii i meteorologii. Wydawnictwa Komunikacji i Łączności, Warszawa 1970.
- [9] Parzen E.: On Estimation of a Probabilist Density Function and Mode. *An. Math. Statist.* 1962, 33, 1065–1076.
- [10] Silverman B.W.: Density Estimation for Statistic and Data Analysis. Monographs on Statistics and Applied Probability. Ed. Chapman and Hall. London, New York 1986.
- [11] Szopa T.: Niezawodność i bezpieczeństwo. Podstawy konstrukcji maszyn, t. I, Wydawnictwa Naukowo-Techniczne, Warszawa 1999.
- [12] Tapia R., Thomson J.: Nonparametric Probability Density Estimation. Baltimore, Maryland Z/Z/8, The Johns Hopkins University Press 1978.

- [13] Wawrzyniak B.: Szkoła Zarządzania. Państwowe Wydawnictwo Ekonomiczne. Warszawa 1987.

Wojciech FELUCH

Probabilistic Assessment of the Reliability of the Network System Infrastructure Using the Kernel Estimation and in Conditions of the Measuring Gap

The article presents a probabilistic method for assessing the reliability of the measurement gap in conditions associated with a lack of data relating to failure in the long term operation of the network infrastructure. Used data bus failure of water supply network. Nonparametric approach was applied in the form of kernel estimation using Parzen estimator.

Keywords: water supply network, the gap measurement, reliability, kernel estimation, Parzen estimator.

SUMMARY