

Algorytm porównania histogramów wielomodalnych opóźnień komunikacyjnych w transmisji bezprzewodowej

Streszczenie. Do opisu parametrów transmisji komunikatów, zwłaszcza w sieciach z losowym dostępem do nośnika, bardzo często stosuje się histogram. W praktyce często zachodzi konieczność porównania dwóch lub więcej histogramów – mówi się wtedy o identyfikacji parametrów rozkładu. W referacie przedstawiono metodę porównywania histogramów w kontekście transmisji bezprzewodowej.

Abstract. To describe the wireless transmission parameters, in particular in networks with random access to the medium, very often are used histograms. In practice, it is often necessary to compare two or more histograms – this is the identification process of parameters of the distribution. The paper presents various methods of comparing histograms in the context of wireless transmission. (**Comparison algorithm of multimodal histograms from wireless transmission**)

Słowa kluczowe: histogram, odległość Euklidesa, ZigBee, WiFi, opóźnienia.

Keywords: histogram, Euclidean distance, ZigBee, WiFi, delays.

doi:10.12915/pe.2014.11.10

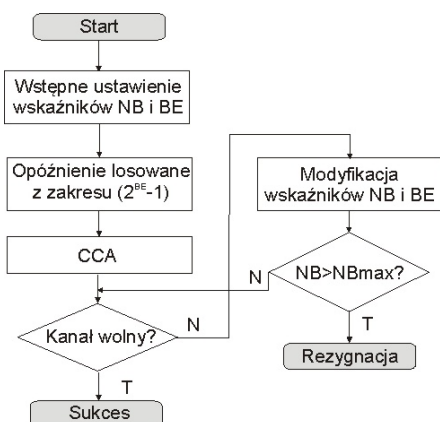
Wstęp

Badając opóźnienia oraz inne parametry transmisyjne w komunikacji bezprzewodowej w sieciach małego zasięgu stwierdzono [1], że do poprawnego interpretowania zjawisk zachodzących w kanale transmisyjnym są niezbędne duże serie danych pomiarowych. Jednym ze sposobów prezentacji dużych serii wyników pomiarowych jest przedstawienie ich w postaci histogramu [1, 2, 3].

W przypadku histogramów jednomodalnych istnieje szereg testów statystycznych oraz zależności, które pozwalają ocenić, czy dany histogram ma rozkład normalny oraz wyznaczyć wartość modalną i odchylenie standardowe rozkładu. Jednak dla histogramów składających się z wielu modów (o rozkładzie normalnym), często połączonych ze sobą, trudno jest stosować standardowe testy. W artykule przedstawiono opracowany algorytm umożliwiający porównywanie dwóch histogramów wielomodalnych z zastosowaniem dedykowanej metryki bazującej na parametrach modelu opóźnień opisanego w [1, 2].

Algorytm CSMA/CA

Cechą charakterystyczną sieci bezprzewodowych jest protokół dostępu do medium komunikacyjnego – CSMA/CA (ang. Carrier Sense Multiple Access with Collision Avoidance) [4, 10]. Ogólnie jego zadaniem jest pewien arbitraż w przypadku, gdy kilka urządzeń sieciowych chce w tym samym czasie nadawać dane.

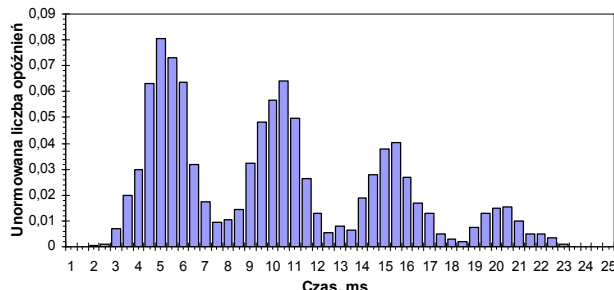


Rys. 1. Uproszczony algorytm CSMA/CA

Sieci bezprzewodowe małego zasięgu używają dwóch typów mechanizmów dostępu do medium – bezszczełinowego oraz z wykorzystaniem tzw. super-ramek [4]. Algorytm dostępu do łącza rozpoczyna się od sprawdzenia, który mechanizm jest w danej sieci stosowany. Szersze informacje na temat opisywanego algorytmu można znaleźć w pracach [4, 10].

W momencie, gdy urządzenie chce transmitować dane musi odczekać pewien losowy odcinek czasu. Jeśli kanał nie jest zajęty, węzeł badając stan medium (tak zwany mechanizm CCA), rozpoczyna transmisję. Jeśli kanał jest zajęty, węzeł czeka losowy okres czasu (tzw. backoff period określony stałą BE) i ponawia próbę dostępu do kanału. Liczba powtórzeń algorytmu CSMA/CA jest zapamiętywana w stałej NB. Gdy wartość stałej przyjmie wartość większą od wartości ustalonej jako maksymalna dane zostaną utracone. Ich ponowna transmisja może się wtedy odbyć tylko na drodze retransmisji programowej. Uproszczony algorytm CSMA/CA przedstawiono na rysunku 1.

W przypadku, gdy jakaś ramka danych nie zostanie odebrana przez odbiornik nadajnik nie otrzyma ramki potwierdzenia. Skutkuje to tym, że nadajnik powtórnie wyśle zagubioną ramką danych. Proces taki nazywa się retransmisją. Liczba możliwych retransmisji jest wielkością zależną w dużej mierze od producenta urządzeń bezprzewodowych, zwykle wynosi 3 lub 5. W związku z takimi algorytmami dostępu do łącza, potwierdzenia i retransmisji histogram opóźnień pakietów w transmisji danych może wyglądać podobnie jak na rysunku 2.



Rys. 2. Histogram opóźnień pakietów danych ilustrujący mechanizm CSMA/CA

Histogram pokazany na rysunku 2 jest nazywany wielomodalnym. W przedstawionym przypadku, otrzymanym w sposób symulacyjny, można wyróżnić: mod pierwszy – ilustrujący pierwszą transmisję danych, mod

drugi – przedstawiający pierwszą retransmisję oraz mody trzeci i czwarty pokazujące kolejne retransmisje. Z rysunku można wnioskować o maksymalnym czasie trwania transmisji pojedynczego pakietu, który wyniósłby, w najgorszym przypadku, ok. 23 ms. Czas ten może być większy, gdy zostanie zwiększona liczba dostępnych retransmisji.

Problematyka identyfikacji

Prowadząc badania opóźnień w transmisji bezprzewodowej uzyskuje się szereg histogramów wielomodalnych [9]. W każdym z nich można otrzymać inną liczbę modów, które mogą być przesunięte w czasie w stosunku do histogramu wzorcowego. Przesunięcie czasowe można uzyskać transmitując dane przez inną liczbę urządzeń bezprzewodowych, ale także wtedy, gdy do transmisji użyte zostanie chociaż jedno urządzenie innego typu (innego producenta lub wykonane w innej wersji). Bardzo często ten sam czas pierwszej transmisji danych oraz kolejnych retransmisji zachowany jest w obrębie urządzeń jednego producenta, a czasem w obrębie tylko jednej serii produkcyjnej. Inną ważną kwestią jest „szerokość” modu, która może być opisywana za pomocą odchylenia standardowego. Dla różnych urządzeń może być ona różna. Często szerokość każdego z modów jest na tyle duża, że mody częściowo wzajemnie się pokrywają.

Tego typu histogramy można porównywać definiując odpowiednią metrykę. Problem wyboru odpowiedniej metryki w przestrzeni cech został szeroko opisany w pracy [6, 8]. Natomiast przykłady obliczania odległości między histogramami są dostępne w pracy [7]. Do najczęściej używanych metryk $d(H_1, H_2)$ przy porównywaniu histogramów H_w i H_b można zaliczyć między innymi:

- metrykę Euklidesową

$$(1) \quad d_E(H_w, H_b) = \sqrt{\sum_{i=1}^N (H_w(i) - H_b(i))^2}$$

- metrykę Manhattan

$$(2) \quad d_M(H_w, H_b) = \sum_{i=1}^N |H_w(i) - H_b(i)|$$

- metrykę Bhattacharyya

$$(3) \quad d_B(H_w, H_b) = -\log \sum_{i=1}^N \sqrt{H_w(i)H_b(i)}$$

- unormowaną korelację wzajemną

$$(4) \quad d_X(H_w, H_b) = \left(\sum_{i=1}^N H_w(i)H_b(i) \right) \div \left(\sum_{i=1}^N H_w^2 \right)$$

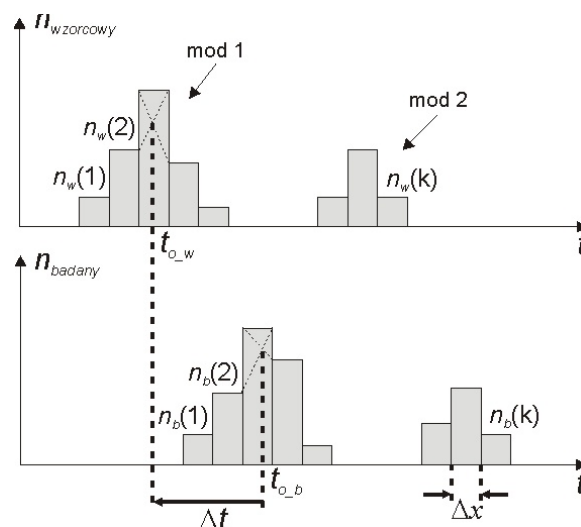
gdzie: H_w – histogram wzorcowy, H_b – histogram badany, i – kolejny przedział klasowy histogramu, N – liczba przedziałów klasowych, d – odpowiednia metryka z indeksem oznaczającym stosowaną metrykę.

Cechą wszystkich przedstawionych metryk jest to, że im uzyskana wartość jest mniejsza tym dwa zbiory elementów (histogramy) są do siebie bardziej podobne.

Algorytm porównywania histogramów

Histogram badany może mieć tyle samo modów co histogram wzorcowy, ale często tak nie jest, co wynika z podstaw działania sieci z losowym dostępem do medium. Może też mieć inne odchylenie standardowe. W dalszej

analizie przyjęto więc pewne założenia. Dla badanych urządzeń bezprzewodowych każdy mod miał charakter rozkładu normalnego, co zostało potwierdzone przeprowadzonymi testami statystycznymi. Stąd wynika założenie wstępne algorytmu, że każdy mod ma rozkład normalny. W celu ułatwienia analizy zakłada się identyczne szerokości, równe Δx , słupków obu histogramów oraz to, że histogramy są unormowane do wartości 1. Opracowana metoda porównania histogramów (częściowo oparta na [5]) bazuje na wykorzystaniu algorytmu porównawczego składającego się z dwóch kroków. Na rysunku 3 przedstawiono dwa histogramy – wzorcowy ($n_{wzorcowy}$) i badany (n_{badany}). Dla każdego z nich oznaczono kolejne słupki histogramów posługując się skróconym zapisem, np.: $n_w(1)$ – pierwszy słupek histogramu wzorcowego, $n_b(k)$ – k -ty słupek histogramu badanego.



Rys. 3. Oznaczenie parametrów potrzebnych do stosowania opracowanego algorytmu

Krok 1.

W histogramie wzorcowym i w histogramie badanym wyznaczana jest w pierwszej kolejności liczba modów. Dla każdej wartości modalnej wyznaczany jest czas opóźnienia t_o , co pokazano na rysunku 3. Następnie dokonuje się przesunięcia czasowego histogramu badanego o taki odcinek czasu Δt , aby wartości modalne pierwszego modu dla histogramu badanego i wzorcowego się pokryły. Czas, o który został przesunięty histogram badany może zostać wykorzystany do zgrubnego porównania histogramów. Można przyjąć, że gdy:

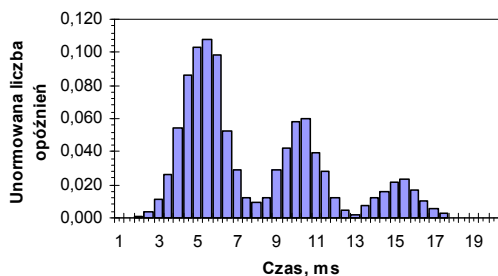
$$(5) \quad \Delta t > \Delta x,$$

wtedy uznaje się że histogramy są różne. Algorytm kończy działanie. Przykładowo, mając dwa histogramy, gdzie szerokość słupków wynosi 0,5 ms, w pierwszym z nich wyznaczona wartość modalna jest równa 5,10 ms a w drugim 5,25 ms. Wtedy wartość Δt wynosi 0,15 ms. Wartość ta jest mniejsza niż szerokość słupka histogramu i zgodnie z kryterium (5) można przejść do drugiego etapu algorytmu. W praktyce można przyjąć inne kryterium. Jeżeli czas przesunięcia jest większy od zadanego czasu progowego (np. 1 ms) wtedy histogramy są różne i prawdopodobnie badany histogram został otrzymany z innych modułów bezprzewodowych. Wartość ta (1 ms), wynika z not katalogowych producentów urządzeń, gdzie zwykle czas ten podawany jest z dokładnością do milisekund.

Krok 2.

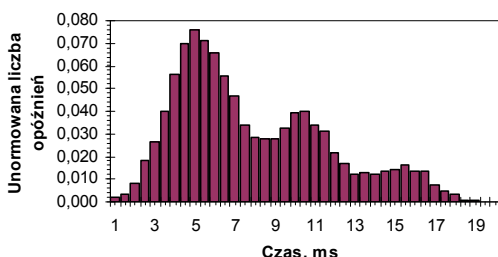
Jeżeli w kroku pierwszym stwierdzono, że dane dwa histogramy należy dalej przetwarzać, to wówczas algorytm przechodzi do ich bardziej szczegółowej analizy. Polega ona na wyznaczeniu współczynnika porównania obu histogramów bazującego na jednej z metryk przedstawionych w rozdziale opisującym problematykę identyfikacji.

Badania obejmowały histogramy uzyskane w sposób pomiarowy i symulacyjny. Histogramy te miały różne przesunięcia czasowe i różne odchylenia standardowe poszczególnych modów. Na rysunkach 4a do 4c przedstawiono trzy przykładowe histogramy. Jeden z nich uznano za wzorcowy natomiast pozostałe są histogramami badanymi.

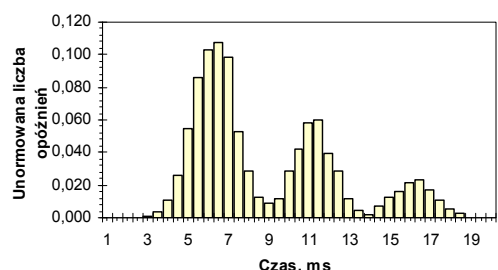


Rys.4a. Histogram wzorcowy

Rysunek 4b przedstawia histogram, w którym poszczególne mody mają takie same średnie opóźnienia jednak inne jest odchylenie standardowe (odchylenie równe 1,5) każdego z modów w porównaniu do histogramu wzorcowego (odchylenie równe 1). Na rysunku 4c przedstawiono histogram uzyskany w sposób symulacyjny, gdzie wszystkie mody są przesunięte w stosunku do histogramu wzorcowego o 1 ms.



Rys.4b. Histogram o odchyleniu standardowym 1,5



Rys.4c. Histogram przesunięty o 1 ms

Do każdej pary histogramów zastosowano wszystkie wspomniane metryki porównawcze. Wyniki badań podsumowano w tabeli 1. Na jej podstawie można stwierdzić, że w prawie każdym przypadku użyte metryki są bardziej wrażliwe na zmiany czasowe przedstawianych histogramów niż na zmiany odchylenia standardowego modów, czyli optyczną zmianę kształtu histogramu. Jedynie w przypadku zastosowania metryki korelacji wzajemnej otrzymuje się wyniki zbliżone dla obu przypadków.

Zgodnie z opisanym algorytmem i wynikami przedstawionymi w tabeli 1 bardziej podobne do siebie są histogramy a i b, pokazane na rysunkach 4a i 4b.

W kolumnie oznaczonej jako „różnica” wyznaczono bezwzględne wartości różnicy współczynników a-b oraz a-c. Wartość ta może być użyta do określenia, która z metryk najwyraźniej pokazuje różnicę pomiędzy histogramami.

Tabela 1. Wyniki porównania histogramów – oznaczenia w tabeli są spójne z rysunkiem 4

	a-b	a-c	różnica $ (a-c)-(a-b) $	jednostka
Krok 1	0	1	1	ms
Krok 2, d_E	0,08	0,16	0,08	-
Krok 2, d_M	0,38	0,69	0,31	-
Krok 2, d_B	0,01	0,04	0,03	-
Krok 2, d_X	0,80	0,79	0,01	-

Podsumowanie

Przedstawiony algorytm porównania histogramów został opracowany tak, aby w prosty sposób uzyskać informację, czy dwa histogramy są do siebie podobne, a co za tym idzie czy transmisja danych miała miejsce dla tych samych warunków pomiarowych, czy też nie. Jest to istotne dla celu identyfikacji parametrów modelu opóźnień transmisyjnych opisanego w [1]. Z tabeli 1 (kolumna „różnica”) wynika, że metryka Manhattan daje najszerszy zakres zmienności. Z tego powodu została wybrana jako najlepsza w prezentowanym przykładzie. Jako graniczną wartość metryki d_M decydującą o tym czy dwa histogramy są podobne przyjęto wartość 0,5. Dla tak przyjętej wartości wynika, że histogramy a i b są podobne, a histogramy a i c są różne. Metryka ta daje dobre rezultaty ponieważ w praktyce ważniejsze jest przesunięcie czasowe niż odchylenie standardowe pojedynczego modu histogramu. Należy tu podkreślić, że dla różnych urządzeń bezprzewodowych może się okazać konieczne przyjęcie innej metryki i innej wartości granicznej.

LITERATURA

- [1] Krupane B.: Modelowanie opóźnień transmisji spowodowanych zaburzeniami w sieciach bezprzewodowych w standardzie IEEE 802.15.4. *Praca doktorska*, Gliwice, 2012r.
- [2] Jakubiec J., Krupane B., Grygiel M.: Probabilistyczne modelowanie opóźnień komunikacyjnych w sieciach komputerowych. *Zeszyty Naukowe Pol. Śl., Elektryka*, Z. 216, Gliwice 2011, s. 79-92.
- [3] Krupane B., Bogacz R.: Badania opóźnień transmisji w wielowęzłowych sieciach ZigBee. *Mechanik*, 7/2013.
- [4] Farahani S.: ZigBee Wireless Networks and Transceivers. *Elsevier Ltd*, USA, 2008.
- [5] Bityukova S., Krasnikov N., Nikitenkoc A., Smirnovaa V.: A method for statistical comparison of histograms, *Cornell University Library, Physics*, 2008.
- [6] Tadeusiewicz R., Flasiński M. Rozpoznawanie obrazów. *Państwowe Wydawnictwo Naukowe*, Warszawa, 1991.
- [7] Stricker M., Orengo M. Similarity of Color Images, *SPIE Storage and Retrieval for Image and Video Databases III*, vol. 2185, 1995, pp. 381-392.
- [8] Krupane B., Bogacz R.: Metoda porównywania histogramów wielomodalnych dla celów analizy transmisji bezprzewodowej. *Materiały X Konferencji Systemy Pomiarowe w Nauce i w Przemysle*, Łagów, 2014.
- [9] Krupane B., Bogacz R.: Investigations of transmission delays in ZigBee networks, *Przegląd Elektrotechniczny*, 2014, nr 1, s.70-75.
- [10] Elahi A., Gschwender A.: ZigBee Wireless Sensor and Control Network, *Prentice Hall*, 2010.

Autorzy: dr inż. Beata Krupane, Politechnika Śląska, Instytut Metrologii, Elektroniki i Automatyki, ul. Akademicka 10, 44-100 Gliwice, E-mail: beata.krupane@polsl.pl; dr inż. Ryszard Bogacz, Politechnika Śląska, Instytut Metrologii, Elektroniki i Automatyki, ul. Akademicka 10, 44-100 Gliwice, E-mail: ryszard.bogacz@polsl.pl.