

# System transmisji obrazu ruchomego przez Internet jako hierarchiczny system złożony

Zbigniew Omiotek

Wyższa Szkoła Zarządzania i Administracji w Zamościu

**Streszczenie:** System transmisji obrazu ruchomego przez Internet jest obecnie postrzegany najczęściej jako system prosty, a jego poszczególne elementy badane są w sposób niezależny od pozostałych. Zdaniem autora wspomniany system jest systemem złożonym, który tworzy struktura składająca się z trzech podstawowych warstw. Są to: źródło, kanał transmisji oraz węzły sieci. Każda z tych warstw traktowana oddzielnie jest również systemem złożonym, a cały system transmisji należy rozpatrywać łącznie, a nie oddzielnie, jak zakłada większość obecnych kierunków badań. Wiąże się to z koniecznością stosowania narzędzi badawczych oraz metod projektowania odpowiednich dla systemów złożonych. W artykule zaprezentowano ideę przetwarzania autonomicznego jako metody projektowania systemów złożonych prowadzącej do poprawy ich wydajności. Przedstawiono także propozycję adaptacyjnego zarządzania jakością usług podczas transmisji wideo. Zakłada ona wykorzystanie zjawisk wyższych rzędów, których do tej pory nie uwzględniano.

## Wstęp

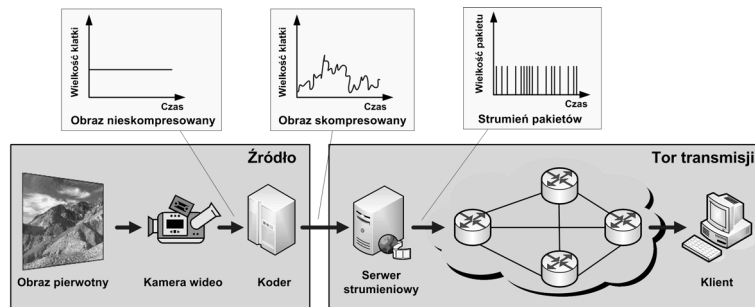
Jakość usług w sieciach IP, a tym samym w Internecie, nabrała w ostatnich latach szczególnego znaczenia. Jest to spowodowane naturalną i ustawiczną tendencją podnoszenia poziomu świadczonych usług we wszystkich dziedzinach, a szczególnie w informatyce. Przyczyną tego jest między innymi gwałtowny rozwój aplikacji multimedialnych. Należy podkreślić, że w ogólnym przypadku jakość usług może być postrzegana jako rezultat dopasowania oczekiwań do ograniczonych możliwości ich zaspokojenia. Świadomość ograniczonych możliwości towarzyszy ludzkości od zarania jej istnienia. Wyzwania te podjęto w sposób naturalny, a obecnie próbuje się je rozwiązywać również w obszarze informatyki, koncentrując się na kontekście technicznym, czyli wrodzonej jakości usług. Sprowadza się to do kluczowego zagadnienia dopasowania realizowanych procesów do zawsze ograniczonych zasobów rzeczywistego systemu wykonawczego.

Naturalną cechą pierwotnego źródła informacji multimedialnej jest występowanie w nim procesów zarówno o charakterze krótko-, jak i długoterminowym. Ponadto, internetowy, wielowarstwowy tor transmisji pakietowej dodatkowo pogłębia procesy o charakterze długoterminowym źródła, jak i sam generuje procesy długoterminowe. Przyczyną tego zjawiska jest fakt, iż zarówno źródło, jak i kanał są systemami złożonymi. Jak wynika z doświadczeń i statystyki nieekstensywnej, procesy długoterminowe są generowane lokalnie, degradując parametry systemu w skali globalnej. W związku z tym globalną wydajność systemu można efektywnie poprawiać poprzez usuwanie wąskich gardeł na poziomie lokalnym, dlatego na obecnym etapie badań najbardziej efektywnym zarządzaniem jakością przekazu obrazu ruchomego przez Internet wydaje się rozwiązanie wykorzystujące predykcję procesów o charakterze długoterminowym w źródle do automatycznego, dynamicznego ustawiania priorytetu transmisji w trybie kroczącym „od grupy do grupy pakietów”. Rozwiązanie to umożliwiłoby bardziej efektywne wykorzystanie zasobów systemu i pozwoliłoby zagwarantować wyższą jakość usług zdecydowanie szerszemu gronu użytkowników niż w dotychczas stosowanych trybach zarządzania.

## 1. Hierarchiczny model systemu transmisji obrazu ruchomego

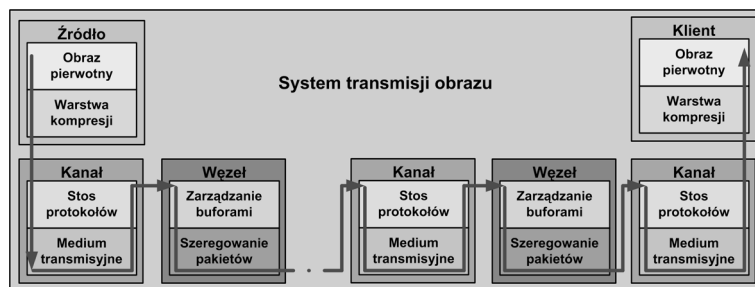
Podstawowe elementy funkcjonalne systemu transmisji obrazu ruchomego przez Internet to źródło sekwencji wizyjnej oraz tor transmisji (rys. 1). Tor transmisji składa się z kanału transmisji oraz węzłów sieci, natomiast

źródło realizuje przechwycenie obrazu pierwotnego (kamera wideo) oraz kompresję tego obrazu (koder). Oczywiście, w niektórych rozwiązaniach oba wymienione wyżej procesy realizowane są w jednym urządzeniu funkcjonalnym. Na wyjściu źródła otrzymujemy więc obraz skompresowany według określonego algorytmu kompresji w sposób programowy lub sprzętowy. W dalszej części pracy pojawi się pojęcie obrazu pierwotnego. Jest to fragment otaczającej nas rzeczywistości, który za pomocą kamery wideo jest w danym momencie czasu przechwytywany w postaci klatki obrazu.



Rys. 1. Elementy systemu transmisji obrazu ruchomego przez Internet

Większość współczesnych badań zmierzających do poprawy wydajności scharakteryzowanego wyżej systemu transmisji koncentruje się na poszczególnych elementach architektury i rozważa je w oderwaniu od reszty systemu. Zdaniem autora, system transmisji obrazu należy traktować jako system złożony. Wówczas infrastrukturę wykonawczą takiego systemu można przedstawić w postaci zwartej struktury hierarchicznej, w której występują trzy podstawowe warstwy: źródło, kanał transmisji oraz węzły sieci (rys. 2).



Rys. 2. Hierarchiczny model systemu strumieniowej transmisji obrazu

Obraz pierwotny po zamianie na postać cyfrową jest poddawany kompresji zgodnie z określonym algorytmem (MPEG-1,2,4, H.263 itd.). Jak wiadomo, zasadniczym celem kompresji jest eliminacja nadmiarowości nieistotnych z punktu widzenia percepcji obrazu przez człowieka. Dzięki zmniejszeniu objętości obraz skompresowany może być przesyłany za pomocą dostępnego kanału transmisji. Wynikiem kompresji jest strumień klatek, który następnie jest dzielony na pakiety określonej wielkości. Pakiety są opatrywane nagłówkami protokołów czasu rzeczywistego (RTP, RTCP, RTSP itp.), a następnie nagłówkami protokołów niższych warstw (UDP, TCP, IP). W ten sposób formowany jest strumień transportowy, który dzięki protokołom jest przesyłany przez kanał transmisji. Wskutek dynamicznych zmian charakterystyk kanału strumień może być narażony na straty pakietów, dlatego ważną rolę odgrywa zarządzanie parametrami kanału, np. przepustowością. Po dotarciu do węzła (routera) strumień jest poddawany oddziaływaniom mechanizmów zarządzania buforami, np. *tail drop*, EPD czy RED. Mechanizmy te są odpowiedzialne za podjęcie decyzji, które pakiety trzeba zniszczyć w okresie przeciążenia, gdy bufor nie mogą pomieścić całego nadchodzącego ruchu. Przed opuszczeniem węzła na strumień oddziałują jeszcze mechanizmy szeregowania pakietów w procesie przyznawania pasma. Są to np. mechanizmy FIFO, WFQ czy szeregowanie według priorytetów przyporządkowanych na stałe.

Każdy element na rys. 2, tj. źródło, kanał transmisji i węzły sieci, ma charakter systemu złożonego, składającego się z kilku warstw. Każdą z tych warstw można z kolei zastąpić kilkoma kolejnymi, czyniącymi prezentowany

model bardziej szczegółowym. W takim modelu każda warstwa wpływa na jakość transmisji obrazu, lecz waga tego wpływu nie jest jednakowa. Im wyżej w strukturze leży dana warstwa, tym większy jest jej wpływ na działanie całego systemu. Taka właściwość jest charakterystyczna dla systemów hierarchicznych. W modelu przedstawionym na rys. 2 najniższa warstwa źródła jest położona bezpośrednio na najwyższej warstwie kanału. W kanale sytuacja jest analogiczna, najniższa warstwa kanału leży na najwyższej warstwie węzła. Zatem można ogólnie przyjąć, że występują trzy systemy, z których każdy jest systemem złożonym. Te trzy systemy razem tworzą również system złożony (pojawia się więc struktura fraktalna), dlatego bazując na analizie opartej na systemach złożonych, można w szerszym zakresie wyeksponować własności tak omawianego systemu transmisji.

### 1.1. Źródło sekwencji wizyjnej jako system złożony

Większość obrazów obserwowanych w naturze (drzewa, twarze, domy, góry, chmury itd.) wykazuje własność samopodobieństwa przestrzennego. Wyraża się ona w tym, iż obraz składa się z kopii (odpowiednio przekształconych) części samego siebie. Poziom samopodobieństwa charakteryzującego obraz można wyznaczyć, wykorzystując wymiar fraktalny obrazu oraz jego związek ze współczynnikiem Hursta. Wymiar fraktalny charakteryzuje stopień złożoności obiektów, wykorzystując ocenę tego, jak szybko wzrastają długość, powierzchnia czy objętość, jeśli pomiar dokonywany jest z coraz większą dokładnością. W przypadku obiektów fraktalnych wykorzystana jest zasada, że dwie wielkości – długość, powierzchnia czy objętość – z jednej strony, a stopień dokładności – z drugiej – nie zmieniają się w sposób dowolny, lecz są związane prawem potęgowym, które pozwala wyznaczyć jedną wartość na podstawie drugiej. Istnieje wiele estymatorów wymiaru fraktalnego, z których najwięcej zastosowań ma wymiar pudełkowy.

W celu obliczenia wymiaru pudełkowego badany obraz należy umieścić na regularnej siatce składającej się z elementów o długości boków równej  $\delta$ , a następnie policzyć, ile elementów siatki (pudełek) pokrywa obraz. Liczba, którą otrzymamy ( $N(\delta)$ ), będzie zależna od rozmiaru elementów siatki. W dalszych iteracjach należy stopniowo zmniejszać wartość  $\delta$  i określać odpowiednie wartości  $N(\delta)$ . Istota określenia wymiaru pudełkowego polega na obserwacji, jak zmienia się  $N(\delta)$  przy zmianie  $\delta$  [Jurgens i inni 2002]. Jak zostało wykazane przez Pentlanda [Pentland 1984], parametr Hursta ( $H$ ) jest bezpośrednio związany z wymiarem fraktalnym  $D_F$  zależnością:

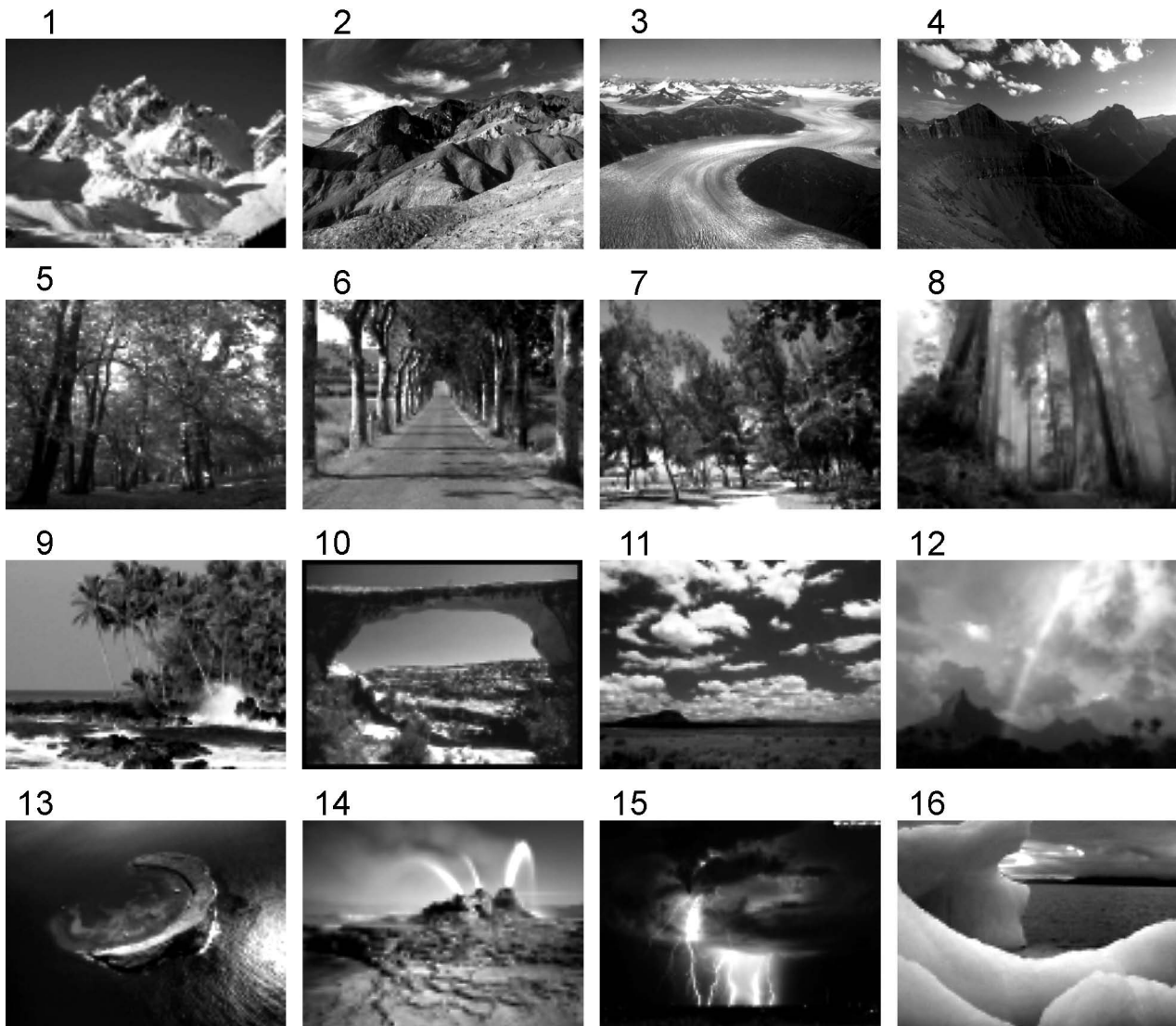
$$[1] \quad H = E + 1 - D_F,$$

gdzie  $E$  jest wymiarem topologicznym. Jeżeli  $0,5 < H \leq 1$ , badany obraz charakteryzuje się samopodobieństwem przestrzennym.

W dalszej części artykułu przedstawiono wybrane wyniki analizy fraktalnej obrazów występujących w naturze (rys. 3). Badane obrazy po przekształceniu do skali szarości poddawane były analizie za pomocą metody pudełkowej. Do analizy wykorzystano oprogramowanie Fractal Analysis System [Omiotek i Grabowski 2006a; Sasaki]. Wyniki analizy fraktalnej pokazały, że pierwotna struktura danych (obraz źródłowy) w większości analizowanych przypadków posiada cechy fraktalne oraz charakteryzuje się samopodobieństwem przestrzennym (współczynnik Hursta znacznie przekracza poziom 0,5). Oznacza to, że wielkości takie jak długość czy powierzchnia – z jednej strony, a stopień dokładności – z drugiej, nie zmieniają się w sposób dowolny, lecz są związane prawem potęgowym. Są to własności fraktalne, które powodują generowanie procesów z zależnościami długoterminowymi. Te procesy, jako naturalna cecha informacji zawartej w obrazie są wnoszone do kodera, a następnie do dalszych stopni toru transmisji.

Tab. 1. Wyniki analizy fraktalnej obrazów przedstawionych na rys. 3

Nr	$D_b$	H	Kategoria	Nr	$D_b$	H	Kategoria
1	2,4956	0,5044	Góry	9	2,4455	0,5545	Horyzont
2	2,4017	0,5983		10	2,3702	0,6298	
3	2,3324	0,6676		11	2,2748	0,7252	
4	2,2692	0,7308		12	2,1625	0,8375	
5	2,6375	0,3625	Drzewa	13	2,4805	0,5195	Krajobraz
6	2,5495	0,4505		14	2,3955	0,6045	
7	2,4998	0,5002		16	2,1231	0,8769	
8	2,3914	0,6086		16	2,0502	0,9498	



Rys. 3. Zdjęcia przedstawiające badane obrazy (rozmiar oryginalny 800×600 pikseli)

W dalszej części artykułu zaprezentowano wyniki badań sygnałów na wyjściu kodera. Do pomiaru wielkości kolejnych klatek tworzących szeregi czasowe (strumień obrazu na wyjściu kodera) zastosowano oprogramowanie VirtualDubMod<sup>1</sup>. Podczas badań wykorzystane zostały różne algorytmy kompresji (stratne i bezstratne).

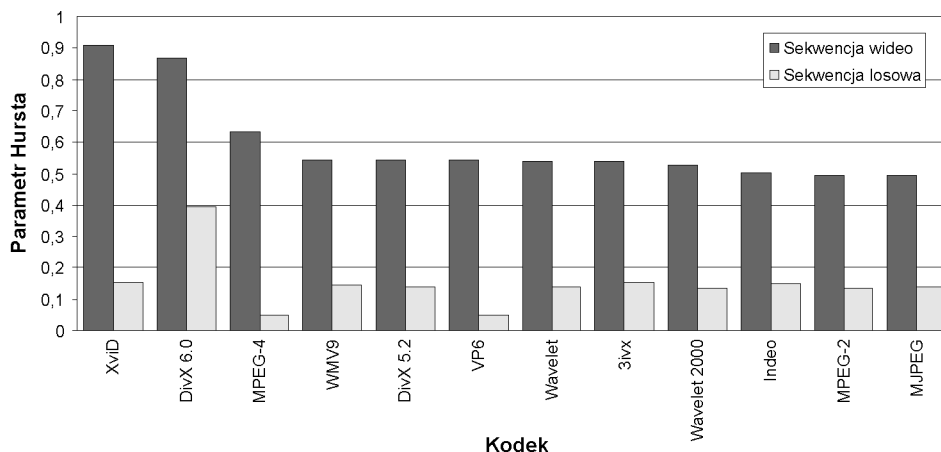
Na początku zastosowano dwa rodzaje nieskompresowanego obrazu źródłowego – sekwencję wideo i sekwencję losową. Następnie badania uzupełniono o kilka rodzajów sekwencji wideo. Charakteryzowały się one różnymi poziomami akcji, zmian scen, zbliżeń itd., co przekładało się na różnorodność wymagań dla ruchu pakietów stawianych sieci oraz systemom końcowym. Wykorzystane klipy umownie przyporządkowano do trzech kategorii: *Przemówienie* (orędzie premiera RP), *Koncert* (nagranie z koncertu Warszawskiej Orkiestry Symfonicznej) oraz *Film akcji* (fragment filmu pt. *Mission Impossible*). Na rys. 4 przedstawiono wybrane klatki kodowanych sekwencji.

<sup>1</sup> <http://www.virtualdub.org>.

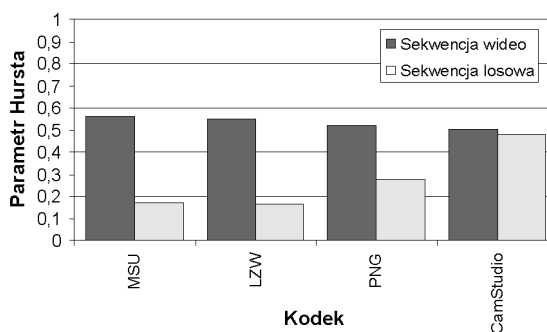


Rys. 4. Przykładowe klatki kodowanych sekwencji wideo: kategoria *Przemówienie* (a), kategoria *Koncert* (b), kategoria *Film akcji* (c)

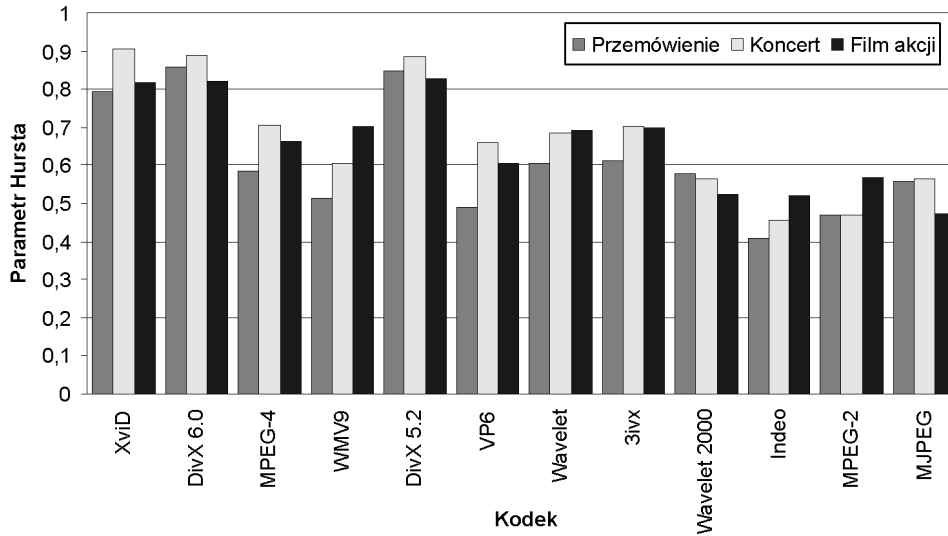
W przypadku klipów wideo (fragmenty filmów) kolejne klatki zawierały cechę samopodobieństwa przestrzennego. Obecność tej cechy dla większości obrazów naturalnych potwierdziły wyniki badań zaprezentowane wcześniej. Z kolei w przypadku obrazu generowanego losowo cecha samopodobieństwa nie występuje. Zmierzone na wyjściu kodera wielkości kolejnych klatek tworzyły szeregi czasowe, które następnie były poddawane analizie statystycznej.



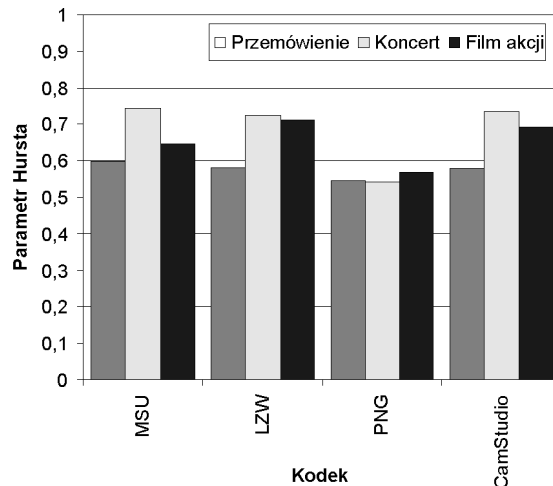
Rys. 5. Porównanie samopodobieństwa strumienia ramek sekwencji wideo i sekwencji losowej dla algorytmów stratnych



Rys. 6. Porównanie samopodobieństwa strumienia ramek sekwencji wideo i sekwencji losowej dla algorytmów bezstratnych



Rys. 7. Porównanie samopodobieństwa strumienia ramek różnych sekwencji wideo dla algorytmów stratnych



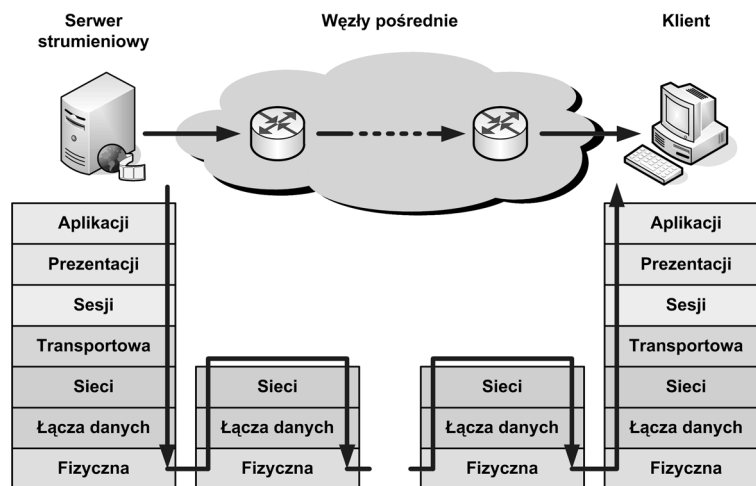
Rys. 8. Porównanie samopodobieństwa strumienia ramek różnych sekwencji wideo dla algorytmów bezstratnych

Wyniki badań zaprezentowane na rys. 7 i 8 dowodzą, iż w przypadku, gdy obraz wejściowy zawiera cechę samopodobieństwa (sekwencje wideo), skompresowany strumień ramek obserwowany na wyjściu kodera również zawiera tę cechę. Taka sytuacja występuje zarówno dla algorytmów stratnych, jak i bezstratnych. Natomiast w przypadku sekwencji losowej cecha samopodobieństwa w skompresowanym strumieniu klatek nie została zaobserwowana (rys. 5 i 6). Wskazuje to na fakt, iż koder bez względu na rodzaj algorytmu kompresji nie jest źródłem samopodobieństwa obserwowanego w ruchu pakietów. Ten etap eksperymentu wykazał niewielki wpływ algorytmów kompresji na poziom samopodobieństwa obserwowanego w strumieniu obrazu. Okazało się, że koder propaguje samopodobieństwo wnoszone przez dane wejściowe do dalszych stopni toru transmisji i nie jest on źródłem powyższej cechy [Omiotek i Grabowski 2006b]. Można obrazowo powiedzieć, że samopodobieństwo przestrzenne jest przenoszone od poziomu pojedynczych pikseli danych pierwotnych do poziomu strumienia klatek na wyjściu kodera. Przyjmuje ono tam postać samopodobieństwa w czasie, czemu towarzyszy zmienna wielkość klatek w strumieniu obrazu.

Dalsze stopnie toru transmisji przenoszą samopodobieństwo obecne w strumieniu klatek na wyjście źródła sygnału i emitują do sieci strumień pakietów o zmiennej częstotliwości. Gwałtowny wzrost natężenia transmisji danych sprawia, że w węzle sieci obserwowane są „wybuchowości”.

## 1.2. Kanał transmisji jako system złożony

Podczas przesyłania danych od serwera strumieniowego do klienta sekwencja wideo na serwerze strumieniowym przechodzi w dół stosu protokołów (zgodnie z modelem ISO/OSI), a więc od warstwy aplikacji do warstwy fizycznej, po czym następuje transmisja przez medium fizyczne. Pokonując kolejne węzły w Internecie, strumień pakietów jest przetwarzany na poziomie pierwszych trzech warstw modeli ISO, czyli warstwy fizycznej, łącza danych i warstwy sieci. Po dotarciu do interfejsu klienta dane przechodzą w górę stosu protokołów – od warstwy fizycznej do warstwy aplikacji – po czym następuje ich prezentacja. Proces ten w sposób obrazowy przedstawiono na rys. 9.



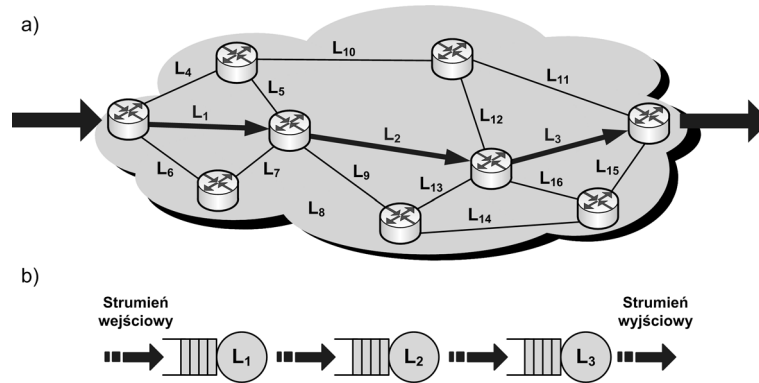
Rys. 9. Model warstwowo kanału transmisji

Zatem kanał komunikacyjny łączący dwie stacje końcowe można sobie wyobrazić jako szeregowe połączenie odcinków medium fizycznego łączącego poszczególne węzły oraz wielu warstw protokołów przetwarzających dane w każdym węzle. Mamy więc połączenie bardzo wielu warstw funkcjonalnych i każda z tych warstw ma wpływ na jakość transmisji strumienia danych w relacji *end to end*. Wskutek niedopasowania struktury wykonawczej (poszczególnych warstw kanału) do realizowanych zadań (przetwarzanych strumieni), na styku tych warstw powstają zaburzenia o cechach długoterminowych. Mają one postać szumów nadmiarowych typu  $1/f$ , które dla małych szerokości pasma kanału prowadzą do degradacji jego pojemności. W związku z tym w celu określenia wydajności kanału należy rozpatrywać łączny wpływ wszystkich elementów składowych opisanej wcześniej struktury. Takie podejście jest możliwe, jeżeli zdefiniowany wcześniej kanał transmisji będzie rozpatrywany w kategoriach systemu złożonego i stosownie do tego zastosowane będą odpowiednie narzędzia badawcze [Strzałka i Grabowski 2005].

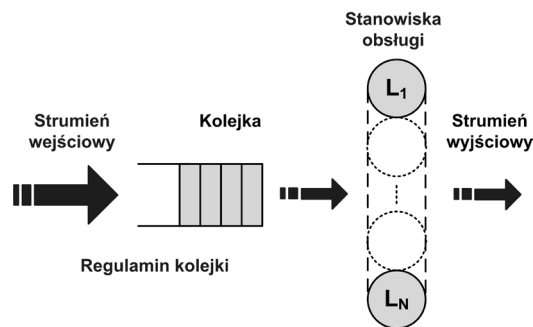
## 1.3. Węzeł sieci jako system złożony

Do analizy rzeczywistych systemów sieciowych stosowane są modele kolejkowe w postaci sieci stanowisk obsługi. W takich modelach stanowiska obsługi to linie transmisyjne łączące węzły sieci, a klienci to przesyłane pakiety, natomiast kolejki ustawiane przed stanowiskami obsługi to kolejki pakietów oczekujące w węzłach (np. w buforach ruterów) na wysłanie w ustalonym kierunku [Czachórski 1999]. Analiza wydajności węzłów sieci sprowadza się zatem do analizy sieciowego systemu kolejkowania (rys. 10).

Pakiety przesyłane w sieci pomiędzy stacjami końcowymi pokonują w istocie złożony system kolejkowy zbudowany z wielu kolejek elementarnych. Podstawowe elementy tworzące taką kolejkę elementarną to wejściowy strumień zgłoszeń, regulamin kolejki oraz stanowiska obsługi (rys. 11).



Rys. 10. Przykładowa trasa rutowania (a) oraz równoważny jej model kolejkowy w postaci sieci stanowisk obsługi (b)



Rys. 11. Elementy składowe systemu kolejowania

Optymalną wydajność całego systemu osiąga się wówczas, gdy wszystkie elementy składowe są nawzajem dopasowane. Klasyczne metody rozwiązywania modeli kolejkowych bazują na analizie procesów stochastycznych w dziedzinie czasu. Traktują one kolejkę jako skupiony system prosty i analizują oddzielnie wpływ poszczególnych elementów na działanie całego systemu. W efekcie takiej praktyki powstała duża liczba modeli, które uwzględniają tylko szczególne przypadki funkcjonowania systemu kolejkowego i nie oferują rozwiązania ogólnego. Takie podejście było słuszne we wczesnej fazie rozwoju teleinformatyki, gdy ruch sieciowy cechował się niewielką zmiennością procesów. Postęp sprawił jednak, że obecnie ruch sieciowy posiada charakterystyki dynamiczne, nieregularne i nie może być już opisywany za pomocą modeli klasycznych traktujących kolejkę jako system prosty.

W rzeczywistości na wydajność systemu kolejkowego wpływają jednocześnie wszystkie jego elementy składowe, zatem wszystkie elementy systemu należy uwzględniać łącznie, a nie oddzielnie, jak zakładają modele proste. Takie podejście prowadzi do analizy kolejki jako systemu złożonego. Do badania takiego systemu można zastosować m.in. analizę widmową, która jest równoważna analizie procesów w dziedzinie czasu. Analiza widmowa oferuje narzędzie badawcze w postaci szumów typu  $1/f$ , które oddają wewnętrzne, dynamiczne procesy systemu. Szумы  $1/f$  pozwalają opisać własności systemu kolejkowego w sposób kompleksowy, z uwzględnieniem jednoczesnego wpływu wszystkich elementów składowych na cały system [Grabowski i Strzałka 2005].

## 2. Przetwarzanie autonomiczne – nowe podejście do projektowania systemów złożonych

Duża złożoność współczesnych rozproszonych systemów komputerowych sprawia, że stają się one trudne do ręcznego sterowania przez administratorów. Sterowanie ręczne jest ponadto czasochłonne, drogie i może być źródłem błędów. Pojawia się więc bariera ograniczająca dalszy rozwój systemów, którą dodatkowo potęguje złożoność procesów projektowania aplikacji, konfiguracji i zarządzania. W przypadku aplikacji, środowisk programowania czy infrastruktur komunikacyjnych występuje również problem integracji, zarządzania i bezpieczeństwa. Pojawia się więc potrzeba badań nad alternatywnym modelem projektowania systemów i aplikacji, który byłby w stanie sprostać aktualnym wyzwaniom w zakresie skali, złożoności i różnorodności.



Wychodząc tym wyzwaniom naprzeciw, firma IBM wystąpiła z inicjatywą nowego podejścia do projektowania systemów i aplikacji określanego jako przetwarzanie autonomiczne [Practical... 2006]. Jest to pojęcie stosowane do opisanego szerokiego zakresu standardów, architektur, produktów, usług i narzędzi, które pozwalają systemom na samozarządzalność. Inspiracją, która legła u podstaw opracowania modelu przetwarzania autonomicznego, był autonomiczny system nerwowy człowieka. System ten pełni w organizmie człowieka funkcję regulatora procesów wegetatywnych toczących się w warunkach wewnętrznych, decydujących o funkcjonowaniu całego ciała, a niekontrolowanych na ogół z poziomu świadomości. Używając terminologii informatycznej, można powiedzieć, że autonomiczny układ nerwowy realizuje ideę przetwarzania rozproszonego, przy czym stopień rozproszenia jest bez porównania wyższy niż w jakimkolwiek obecnie istniejącym systemie komputerowym. Fakt ten czyni układ autonomiczny szczególnie interesującym obszarem obserwacji dla naukowców zajmujących się nowymi architekturami systemów przetwarzania danych. Cel działania systemu autonomicznego, jakim jest regulacja i stabilizacja parametrów fizycznych oraz chemicznych środowiska wewnętrznego organizmu, stawia go na czele listy obiektów interesujących z punktu widzenia automatyków i specjalistów z zakresu techniki sterowania.

Wzorowany na autonomicznym systemie nerwowym model przetwarzania autonomicznego powinien mieć mechanizm, dzięki któremu zmiany jego podstawowych parametrów mogłyby wywołać takie zmiany zachowania systemu, dzięki którym, uwzględniając warunki środowiska, powróciłby on z powrotem do równowagi. Stan stabilnej równowagi jest warunkiem koniecznym dla przetrwania organizmu. W przypadku autonomicznych systemów przetwarzania pojęcie przetrwania rozumiane jest jako zdolność systemu do samodzielnej ochrony przed atakami, powrotu do normalnej pracy po awarii, rekonfiguracji spowodowanej zmianami warunków środowiska oraz ciągłej pracy z wydajnością zbliżoną do optymalnej. Równowaga systemu może być zachwiana zarówno przez środowisko wewnętrzne (np. nadmierne wykorzystanie pamięci czy procesora), jak i zewnętrzne (np. ataki). Strategia przetwarzania autonomicznego udostępnia sieciowym systemom komputerowym możliwość samodzielnego zarządzania bez bezpośredniej interwencji człowieka. Z tego względu jedna z głównych idei leżąca u podstaw funkcjonowania systemów autonomicznych określana jest właśnie jako samozarządzalność. Człowiek odgrywa w tych systemach nową rolę – nie steruje on systemem bezpośrednio, lecz określa jedynie ogólne strategie i cele działania, które służą jako wejściowe dane sterujące dla procesów samozarządzania.

Przykładowa funkcjonalność systemu zdolnego do samodzielnego zarządzania obejmuje instalację oprogramowania, gdy zostanie wykryte, że go brakuje, restart elementu, który się zawiesił, zwiększenie bieżącego obciążenia w przypadku, gdy zostanie wykryty wzrost przepustowości, przełączenie się w tryb *offline* przy wykryciu próby włamania do systemu. Każda z powyższych cech reprezentuje obszar aktywnych badań [Parashar i Hariri 2005]. Pojęcie samozarządzania dotyczy czterech zasadniczych aspektów funkcjonowania systemów, a mianowicie: konfiguracji, optymalizacji, ochrony i przywrócenia normalnej pracy systemu po awarii.

Główna koncepcja, na której ma opierać się działanie systemów autonomicznych, to zastosowanie układów sterowania w trybie *online* wykorzystujących zamknięte pętle sterowania. Ich główne zadanie to monitorowanie określonych zasobów (sprzętowych lub programowych) oraz podejmowanie decyzji i adaptacja systemu w celu autonomicznego utrzymywania parametrów w pożądanym zakresie.

Koncepcja sterowania w trybie *online* wykorzystuje pewne miary wydajności (np. pobór energii, tryb przełączania, itd.), które na bieżąco są optymalizowane. Miara wydajności jest funkcją stanu systemu oraz zmiennych wejściowych (wyjściowych). Najczęściej stosowana jest zasada, zgodnie z którą zmienne są dodawane z różnymi wagami odzwierciedlającymi wpływ danej zmiennej na działanie całego systemu. Moduł prognozowania na podstawie parametrów środowiska pracy, np. obciążenia nadchodzącymi zadaniami, dokonuje estymacji tych parametrów w określonym zakresie czasu (zakresie predykcji). Dane otrzymane na podstawie predykcji są następnie wykorzystane przez matematyczny model systemu do prognozy przyszłego zachowania systemu fizycznego. Prognozowane zachowanie systemu jest z kolei poddane optymalizacji zgodnie z określonymi wymaganiami QoS (*Quality of Service*). Moduł optymalizacji znajduje najlepsze dane sterujące i podaje je na wejście systemu.

### 3. Koncepcja adaptacyjnego zarządzania QoS systemu transmisji obrazu

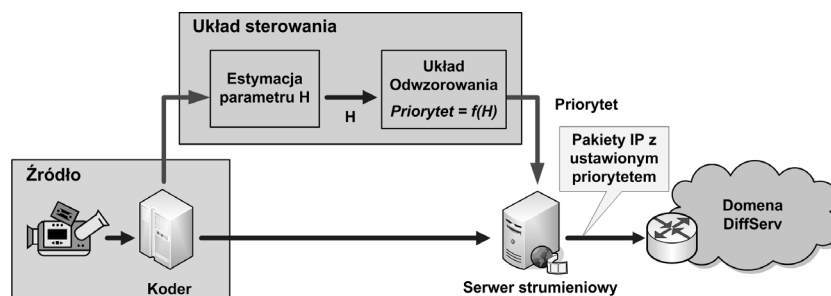
W literaturze są proponowane różne modele prognozowania służące estymacji wydajności systemów komputerowych. W pozycji [Shen i Hellerstein 2000] zaprezentowano model autoregresyjny służący do prognozy trendów w sieciach komputerowych. Połączenie filtra Kalmana z modelem autoregresyjnym służące do wykrywania zmian obciążenia serwera WWW przedstawiono w [Zhang i Hellerstein 2000]. Z kolei w [Vilalta

i inni 2002] zaprezentowano algorytmy predykcji służące do estymacji różnych parametrów wydajności systemu komputerowego uwzględniające zdarzenia nietypowe takie jak naruszenie QoS czy uszkodzenie systemu.

Algorytmy predykcji stosowane w dotychczasowych modelach regulacji w trybie *online* szacują trend na podstawie aktualnych parametrów środowiska. Taka predykcja ma charakter krótkoterminowy i nie oddaje w pełni charakteru procesów zachodzących w rzeczywistych środowiskach sieciowych. Wada dotychczasowych układów sterowania polega na tym, że jeżeli układ fizyczny zostanie dostrojony na podstawie bieżących parametrów środowiska, to czynność ta występuje już „po fakcie”. Zawsze działanie takiego układu będzie opóźnione względem zmian środowiska (w stosunku do tego, co się dzieje obecnie, dostrajany jest następny krok).

W związku z powyższym w algorytmie predykcji jako miarę trendu zmian parametrów środowiska autor pracy proponuje wykorzystać parametr Hursta. Jest to parametr długoterminowy, który nie zmienia się od kroku do kroku w zakresie predykcji. Dzięki temu zmiany parametrów środowiska można prognozować z wyprzedzeniem („do przodu”). Na tej podstawie można adaptacyjnie reagować na samoorganizację kanału i węzła. W proponowanym rozwiązaniu model systemu w układzie sterowania obejmować będzie również zjawiska wyższych rzędów, których do tej pory nie uwzględniano.

Systemem fizycznym, w odniesieniu do którego należy zastosować proponowane rozwiązanie, jest system transmisji obrazu ruchomego przez Internet. Jest to hierarchiczny systemem złożony, w którym na samym szczycie leży warstwa danych pierwotnych (obraz źródłowy) [Grabowski i Omiotek 2005]. Warstwa ta wnosi zaburzenia do ruchu sieciowego, ponieważ są one naturalną cechą informacji zawartej w obrazie. Tych zaburzeń nie można wyeliminować, można natomiast podjąć próby dopasowania struktury wykonawczej (tor transmisji) do przetwarzanego zadania (strumień obrazu). Zdaniem autora należy w tym celu wykorzystać długoterminowy parametr Hursta określający poziom zaburzeń obecnych w strumieniu i na tej podstawie dokonywać adaptacyjnej rekonfiguracji toru transmisji. W celu realizacji tego zadania należy wykorzystać architekturę wspierającą różnicowanie jakości usług (*DiffServ*) i dokonać jej integracji ze wspomnianą wcześniej metodą szacowania zaburzeń ruchu (rys. 12).



Rys. 12. Ustawianie wartości priorytetu pakietu IP na podstawie poziomu zaburzeń wnoszonych przez dane wejściowe

Dokładna charakterystyka proponowanej metody różnicowania QoS w sieci *DiffServ* na podstawie własności długoterminowych strumienia obrazu została zamieszczona w pracy *A Method for QoS Differentiation in DiffServ Network Based on the Long-Term Properties of a Video Stream*<sup>2</sup>, natomiast wstępne wyniki badań symulacyjnych zaprezentowano w [Omiotek 2010].

## Podsumowanie

System transmisji obrazu ruchomego przez Internet jest systemem złożonym, charakteryzującym się strukturą hierarchiczną. Na samym szczycie tej struktury leży warstwa obrazu źródłowego, której naturalną cechą jest występowanie procesów o charakterze długoterminowym. Są one główną przyczyną zaburzeń ruchu i w konsekwencji powodują degradację wydajności całego systemu transmisji. Procesów tych nie można usunąć ze strumienia danych, ponieważ są one własnością informacji zawartej w obrazie źródłowym, w związku z tym w celu poprawy jakości przekazu należy zastosować układ sterowania *online* do automatycznego, dynamicznego ustawiania priorytetu transmisji. Takie rozwiązanie wpisuje się w koncepcję przetwarzania autonomicznego, jednak

<sup>2</sup> Referat zgłoszony na X Międzynarodową Konferencję Informatyka – Badania i Zastosowania (IBIZA 2011).

w odróżnieniu od dotychczasowych rozwiązań układów sterowania *online* do predykcji trendu zmian parametrów środowiska (zaburzeń ruchu) autor proponuje wykorzystać parametr Hursta. Jest to parametr długoterminowy, który nie zmienia się od kroku do kroku w zakresie predykcji. Dzięki temu zmiany parametrów środowiska można prognozować z wyprzedzeniem („do przodu”). Na tej podstawie można adaptacyjnie reagować na samoorganizację kanału i węzła. W proponowanym rozwiązaniu model systemu w układzie sterowania obejmować będzie również zjawiska wyższych rzędów, których do tej pory nie uwzględniano.

## Literatura

- CZACHÓRSKI T. (1999): *Modele kolejkowe w ocenie efektywności pracy sieci i systemów komputerowych*, Wydawnictwo Pracowni Komputerowej Jacka Skalmierskiego, Gliwice.
- GRABOWSKI F., OMIOTEK Z. (2005): *Hierarchiczny model systemu strumieniowej transmisji obrazu ruchomego przez Internet*, [w:] A. KWIECIEN, A. GRZYWAK i A. BANASIK (red.): *Wysokowydajne sieci komputerowe. Zastosowanie i bezpieczeństwo*, Wydawnictwa Komunikacji i Łączności, Warszawa.
- GRABOWSKI F., STRZAŁKA D. (2005): *Wpływ procesów długoterminowych na wydajność kolejki*, [w:] S. WĘGRZYN, B. POCHOPIEŃ, T. CZACHÓRSKI i M. BACH (red.): *Wysokowydajne sieci komputerowe. Nowe technologie*, Wydawnictwa Komunikacji i Łączności, Warszawa.
- JURGENS H., PEITGEN H.-O., SAUPE D. (2002): *Fraktale : granice chaosu*, cz. 1, PWN, Warszawa.
- OMIOTEK Z. *A Method for QoS Differentiation in DiffServ Network Based on the Long-Term Properties of a Video Stream*, Referat zgłoszony na X Międzynarodową Konferencję Informatyka - Badania i Zastosowania (IBIZA 2011), [praca niepublikowana].
- OMIOTEK Z. (2010): *Symulacja sieci DiffServ z różnicowaniem priorytetu pakietów na podstawie parametru Hursta strumienia obrazu*, „Barometr Regionalny. Analizy i prognozy”, nr 3(21), s. 55–79.
- OMIOTEK Z., GRABOWSKI F. (2006a): *Metody badania źródeł informacji*, [w:] S. WĘGRZYN (red.): *Nowe technologie sieci komputerowych: praca zbiorowa*, t. 1, Wydawnictwa Komunikacji i Łączności, Warszawa.
- OMIOTEK Z., GRABOWSKI F. (2006b): *Propagacja samopodobieństwa w torze transmisji obrazu*, [w:] S. WĘGRZYN (red.): *Nowe technologie sieci komputerowych : praca zbiorowa*, t. 1, Wydawnictwa Komunikacji i Łączności, Warszawa.
- PARASHAR M., HARIRI S. (2005): *Autonomic computing: An overview*. „Unconventional Programming Paradigms”, Vol. 3566, s. 257–269.
- PENTLAND A.P. (1984): *Fractal-Based Description of Natural Scenes*. „Ieee Transactions on Pattern Analysis and Machine Intelligence”, Vol. 6, No. 6, s. 661–674.
- Practical Autonomic Computing: Roadmap to Self Managing Technology* (2006), Enterprise Management Associates, [http://www.item.ntnu.no/\\_media/academics/courses/ttm4128/lectures/23-ac\\_practical\\_roadmap\\_whitepaper.pdf](http://www.item.ntnu.no/_media/academics/courses/ttm4128/lectures/23-ac_practical_roadmap_whitepaper.pdf), A White Paper Prepared for IBM.
- SASAKI H. *Fractal analysis system for Windows*, <http://cse.naro.affrc.go.jp/sasaki/fractal/fractal-e.html>.
- SHEN D., HELLERSTEIN J.L. (2000): *Predictive models for proactive network management: application to a production Web server*. Network Operations and Management Symposium, 10–14.04.2000, Honolulu.
- STRZAŁKA D., GRABOWSKI F. (2005): *Wpływ procesów długoterminowych kanału na jego przepustowość*, [w:] S. WĘGRZYN, B. POCHOPIEŃ, T. CZACHÓRSKI i M. BACH (red.): *Wysokowydajne sieci komputerowe. Nowe technologie*, Wydawnictwa Komunikacji i Łączności, Warszawa.
- VILALTA R., APTE C.V., HELLERSTEIN J.L., MA S., WEISS S.M. (2002): *Predictive algorithms in the management of computer systems*, „Ibm Systems Journal”, nr 3(41), s. 41–56.
- ZHANG F., HELLERSTEIN J.L. (2000): *An approach to on-line predictive detection*, „8th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems, Proceedings”, s. 549–556.