

Rozdział 9

Wpływ warstwy fizycznej na wydajność protokołu IEEE 802.11

Bartłomiej ZIELIŃSKI
Politechnika Śląska, Instytut Informatyki
Bartlomiej.Zielinski@polsl.pl

Streszczenie

W rozdziale omówiono wpływ warstwy fizycznej na wydajność protokołu warstwy liniowej w bezprzewodowych sieciach lokalnych zgodnych ze standardem IEEE 802.11. Na podstawie definicji tych warstw wyprowadzono zależności analityczne, które umożliwiają oszacowanie wydajności protokołu warstwy liniowej. Przedstawiono wyniki obliczeń dla poszczególnych warstw fizycznych i porównano ich wpływ na wydajność protokołu.

1. Wprowadzenie

Standard IEEE 802.11 [2] jest obecnie wiodącym rozwiązaniem w zakresie bezprzewodowych sieci lokalnych. Adaptery sieciowe zgodne z tym standardem są dostępne dla wielu typów urządzeń, nie tylko komputerów osobistych. Można je wykorzystać w komputerach przenośnych, palmtopach, a także w cyfrowych aparatach fotograficznych (np. Canon EOS 1D), odtwarzaczach DVD czy cyfrowych tunerach satelitar-nych.

Architektura standardu zawiera definicję warstwy fizycznej i liniowej według modelu odniesienia ISO/OSI. W ciągu około 10 lat istnienia standardu warstwa liniowa nie została znacząco zmodyfikowana, określono natomiast kilka nowych wariantów warstwy fizycznej, wykorzystujących różne sposoby modulacji i kodowania w celu uzyskania różnych prędkości transmisji. Jednakże, zgodnie z założeniami modelu ISO/OSI, warstwa fizyczna musi być przezroczysta dla warstw wyższych. Konieczne jest zatem wprowadzenie określonych mechanizmów sygnalizacji na poziomie warstwy fizycznej. Podejście takie powoduje jednak wprowadzenie dodatkowego narzutu protokołu, co nie pozostaje bez wpływu na wydajność sieci.

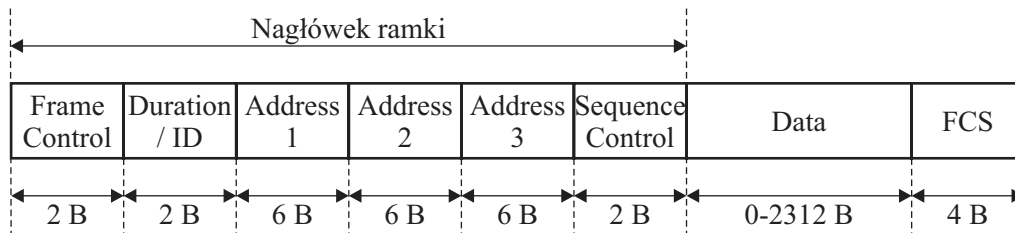
2. Standard IEEE 802.11

Definicja standardu IEEE 802.11 zawiera opis warstwy liniowej oraz kilku wariantów warstwy fizycznej. Sposób działania warstwy liniowej nie zależy od warstwy fizycznej – uwzględnia się tu jedynie pewne jej parametry.

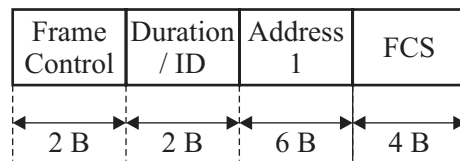
2.1. Warstwa liniowa

Opis warstwy liniowej wyjaśnia działanie protokołu dostępu do łącza, formaty ramek oraz zasady ich wymiany. Ramki dzieli się na trzy typy – sterujące, zarządzające i danych. Każdy typ ramki zawiera pewną liczbę podtypów, określających dokładne przeznaczenie ramki.

Dla potrzeb dalszej analizy potrzebne są tylko ramki Data i Ack (potwierdzenie), używane podczas transmisji danych. Formaty tych ramek pokazane są na rys. 1 i 2.



Rys. 1. Format ramki danych (Data)



Rys. 2. Format ramki potwierdzenia (Ack)

Ramka danych zaczyna się nagłówkiem warstwy liniowej, zawierającym w większości przypadków 24 B. W tej liczbie znajdują się trzy 2-bajtowe pola – sterujące (ang. *Frame Control*), czas trwania (ang. *Duration/ID*) oraz sterowanie kolejnością (ang. *Sequence Control*) – a także trzy 6-bajtowe pola adresowe. Chociaż standard przewiduje użycie czwartego adresu, wykorzystuje się go rzadko, podczas transmisji między dwiema sieciami poprzez tzw. system dystrybucyjny (ang. *Distribution System*). Po nagłówku znajduje się pole danych o pojemności do 2312 B oraz 4-bajtowa suma kontrolna. Ramka potwierdzenia zawiera jedynie pole sterujące, czas trwania, adres i sumę kontrolną, co daje razem 14 B.

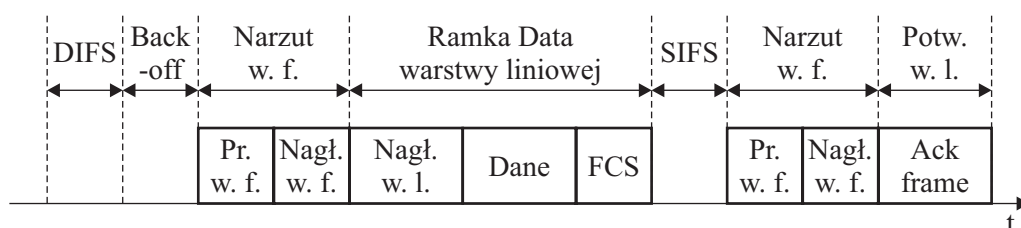
W czasie transmisji danych ramki Data i Ack przesyłane są naprzemiennie, gdyż każda prawidłowo odebrana ramka danych musi być potwierdzona; w przeciwnym razie konieczna jest jej retransmisja. Między ramkami musi upłynąć czas SIFS (ang. *Short Interframe Space*), natomiast przed ramką danych występuje czas DIFS (ang. *Distributed Interframe Space*). Czasy te uwzględniają czas przełączania nadajnika-odbiornika i zależą od warstwy fizycznej. Oczywiście każda ramka warstwy liniowej jest także poprzedzana preambułą i nagłówkiem warstwy fizycznej.

2.2. Warstwa fizyczna

Jak dotąd określono cztery warianty warstwy fizycznej, wyłączając 802.11n, która w chwili pisania rozdziału nie jest jeszcze ukończona:

- „podstawowy” 802.11 z prędkościami transmisji 1 i 2 Mb/s, wykorzystujący rozpraszanie widma metodą przeskokuw częstotliwości (FHSS) lub kluczowania bezpośredniego (DSSS), a także podczerwień;
- 802.11b z prędkościami 5,5 i 11 Mb/s, wykorzystujący modulację HR-DSSS i kodowanie PBCC lub CCK; zgodny wstecznie z wariantem 802.11 DSSS;
- 802.11a z prędkościami 6, 9, 12, 18, 24, 36, 48 i 54 Mb/s, wykorzystujący modulację OFDM; niezgodny z 802.11 ani z 802.11b;
- 802.11g z prędkościami i modulacją jak w 802.11a, ale z możliwością zgodności z 802.11b za cenę zwiększonego narzutu warstwy fizycznej w porównaniu z 802.11a.

Jedną z ważniejszych cech wszystkich warstw fizycznych standardu 802.11 jest możliwość prowadzenia transmisji z różnymi prędkościami. Pozwala to użyć prędkości optymalnej dla danych warunków, co jest szczególnie istotne dla sieci bezprzewodowych ze względu na ich zmienność w czasie. Prędkość transmisji warstwy fizycznej nie może jednak być widoczna dla warstw wyższych, konieczne zatem jest wprowadzenie odpowiedniej sygnalizacji na poziomie warstwy fizycznej. Niestety, zwiększa to narzut protokołu. Dodatkowy narzut powodują także pola synchronizujące, niezbędne dla transmisji radiowej. Biorąc pod uwagę, iż każda ramka musi być poprzedzona preambułą i nagłówkiem warstwy fizycznej, wymiana ramek może przebiegać jak pokazano na rys. 3.



Rys. 3. Proces wymiany ramek

Warto zauważyć, że długość preambuły i nagłówka warstwy fizycznej oraz czasy DIFS i SIFS zależą od wariantu tej warstwy. Tym niemniej, preambuła i nagłówek przesyłane są zwykle z najniższą prędkością transmisji (zwykle 1, 2 lub 6 Mb/s), aby jak najwięcej stacji mogło prawidłowo odebrać zawarte tam informacje. Elementy pochodzące z warstwy liniowej przesyła się natomiast z najwyższą prędkością możliwą do uzyskania w danych warunkach. Dla potrzeb transmisji potwierzeń wykorzystuje tzw. prędkości podstawowe, nie przekraczające prędkości, z jaką odebrano potwierdzane właśnie dane (np. ramka Data – 54 Mb/s, Ack – 24 Mb/s).

W celu zmniejszenia narzutu warstwy fizycznej przy użyciu wyższych prędkości transmisji, wprowadzono tzw. „krótką preambułę”. W tym przypadku sama preambuła jest rzeczywiście dwukrotnie krótsza od „długiej”, a dodatkowo nagłówek przesyłany jest dwukrotnie szybciej. W sumie narzut warstwy fizycznej maleje o połowę.

W tabeli 1 zebrano parametry zależne od warstwy fizycznej, niezbędne dla dalszych rozważań i oszacowania wydajności protokołu.

Tabela 1. Parametry zależne od warstwy fizycznej¹

Warstwa fizyczna	Czas [μ s]				Dodatkowy narzut
	SIFS	Preambuła	Nagłówek	Slot	
802.11 DSSS	10	144	48	20	
802.11 FHSS	28	96	32	50	Kodowanie 32/33
802.11 Ir (1 Mbps)	10	16	41	8	
802.11 Ir (2 Mbps)	10	20	25	8	
802.11b Short	10	72	24	20	
802.11a	16	20	4	9	≥ 22 bity
802.11g Long	10	144	48	9 / 20	18 μ s
802.11g Short	10	72	48	9 / 20	18 μ s

3. Ocena wydajności standardu IEEE 802.11

Jest wiele miar wydajności protokołów. W niniejszym rozdziale wydajność ta jest rozumiana jako iloraz czasu niezbędnego do transmisji samych danych i czasu niezbędnego do transmisji tych danych z użyciem danego protokołu. W ten sposób można oszacować narzut protokołu dla różnych warunków pracy sieci.

W celu wyprowadzenia potrzebnych zależności użyto sposobu podobnego do opisanego w pracy [1]. Posłużył on do porównania wpływu warstwy fizycznej na wydajność protokołu warstwy liniowej. Przyjęto następujące założenia:

- sieć składa się z dwóch stacji, a w czasie transmisji nie występują kolizje ani inne błędy transmisji – nie ma więc retransmisji;
- wymiana danych odbywa się na poziomie warstwy liniowej, zatem nie uwzględnia się narzutu warstw wyższych, jak np. TCP/IP;
- nie używa się mechanizmów zabezpieczających;
- czas przetwarzania ramek można zaniedbać.

Dla wszystkich wariantów warstwy fizycznej wydajność protokołu można wyrazić następująco:

$$\eta = \frac{L/R_{wl}}{T_{DIFS} + T_{SIFS} + T_{BO} + 2 \cdot T_{PLCP} + T_{Data} + T_{Ack}}. \quad (1)$$

We wzorze (1) T_{DIFS} oraz T_{SIFS} określają czas trwania okresów DIFS i SIFS, natomiast T_{PLCP} – czas trwania preambuły i nagłówka warstwy fizycznej. Wartości te zdefiniowane są w specyfikacji poszczególnych warstw fizycznych, a zebrano je w tabeli 1. Z kolei T_{BO} określa czas trwania okresu wycofywania (ang. *backoff period*), który, zgodnie z wyjaśnieniami w pracy [1] można uprościć do postaci:

¹ Dane przedstawione w tabeli zaczerpnięto z pracy [2]

$$T_{BO} = \frac{CW_{min}}{2} \cdot T_{slot} \quad (2)$$

T_{slot} określa czas trwania szczeliny (tabela 1), natomiast CW_{min} (ang. *Contention Window*) określa minimalną liczbę szczelin okresu rywalizacji. Z kolei

$$T_{Data} = \frac{8 \cdot (28[B] + L)}{R_{wl}} \quad (3)$$

oraz

$$T_{Ack} = \frac{8 \cdot 14[B]}{R'_{wl}}, \quad (4)$$

gdzie: L – pojemność pola danych ramki [B], R_{wl} – prędkość transmisji ramki danych [b/s], natomiast R'_{wl} – prędkość transmisji potwierdzenia [b/s].

Podczas obliczania T_{Data} i T_{Ack} należy zwrócić uwagę na dodatkowy narzut wynikający z przyjętej metody modulacji, np. kodowanie 32/33 w warstwie 802.11 FHSS. Podobnie modulacja OFDM wprowadza dodatkowy narzut o wielkości co najmniej 22 B na każdą przesyłaną ramkę.

Obliczenia wykonano dla pojemności pola danych 2304 B (największa dla standardu 802.11), 1500 B (największa dla sieci Ethernet, z którą 802.11 często współpracuje), 256 B (największa dla protokołu AX.25 używanego w sieci Packet Radio) [3] oraz 48 B (zbliżona rozmiarowi komórki ATM oraz najmniejszej ramki sieci Ethernet) [4]. Podczas obliczeń przyjęto, że ramki potwierdzeń przesyłane są z tą samą prędkością co dane w 802.11 i 802.11b. Natomiast dla sieci 802.11a i 802.11g przyjęto, że potwierdzenia przesyła się z prędkością podstawową (6, 12 lub 24 Mb/s), nie przekraczającą prędkości, z jaką odebrano potwierdzane dane. Takie zachowanie sieci potwierdzono testami z wykorzystaniem monitora łącza bezprzewodowego.

3.1. IEEE 802.11

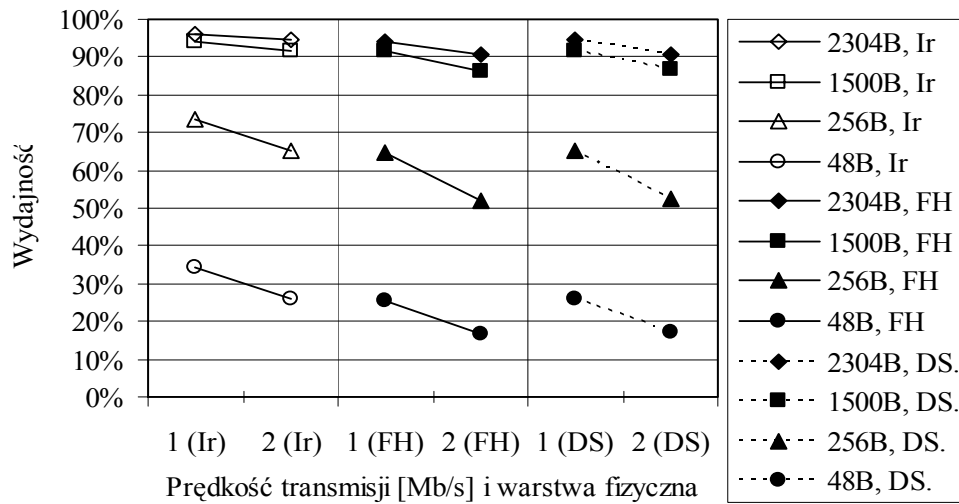
Wyniki otrzymane dla wszystkich wariantów „podstawowego” standardu 802.11 – podczerwieni (Ir), przeskoków częstotliwości (FH) i kluczowania bezpośredniego (DS) – przedstawiono razem na rys. 4. Najbardziej wydajną warstwę tworzy podczerwień, najmniej – DS. W każdej odmianie transmisja z prędkością 2 Mb/s jest mniej efektywna (ale szybsza) niż 1 Mb/s. Różnica między tymi prędkościami jest najmniej widoczna dla podczerwieni, a dla dłuższych ramek – 1500 lub 2304 B – jest pomijalna. W każdej odmianie zmniejszanie rozmiaru pola danych zmniejsza zarówno wydajność protokołu, jak i przepustowość efektywną.

3.2. IEEE 802.11b

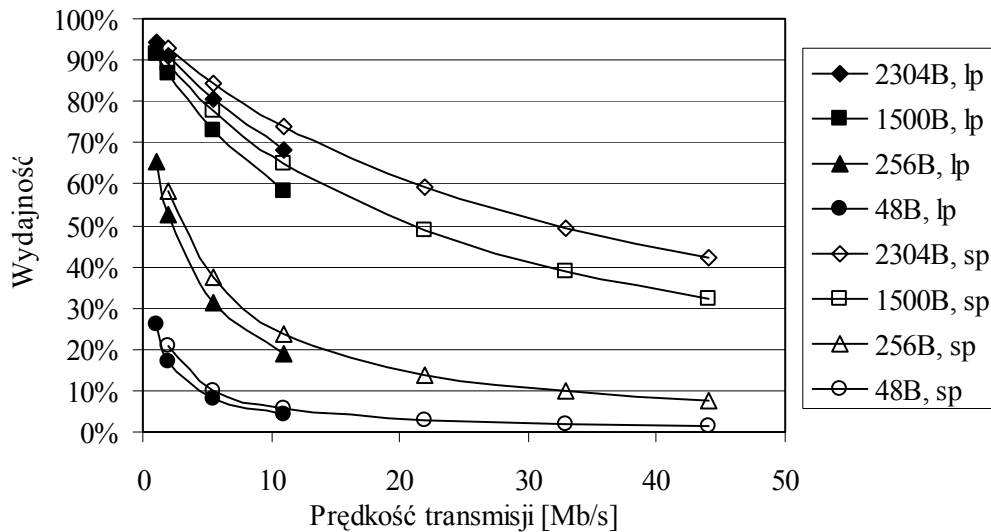
Wyniki otrzymane dla warstwy fizycznej 802.11b, z użyciem długiej (lp) lub krótkiej (sp) preambuły, pokazano na rys. 5. Nietrudno zauważyć, że skrócenie preambuły przynosi pewne korzyści. Przykładowo, wydajność protokołu z najdłuższymi ramkami danych (2304 B) i długą preambułą jest zbliżona do wydajności z ramkami 1500 B i krótką preambułą. Różnica między preambułami jest najbardziej widoczna w przy-

padku ramek zawierających 256 B danych. Można także zauważyć, iż wraz ze wzrostem prędkości transmisji spada wydajność protokołu, lecz rośnie efektywna przepustowość.

Pomimo że standard określa prędkości transmisji nie większe niż 11 Mb/s, przeprowadzono obliczenia także dla prędkości wyższych, mianowicie 22, 33 i 44 Mb/s. Prędkości te zdefiniowane są jako opcjonalne w standardzie 802.11g, który można rozpatrywać jako rozszerzenie 802.11b. Ponadto, są dostępne urządzenia, w których użycie tych prędkości transmisji jest możliwe. Obliczenia dla tych prędkości przeprowadzono przy założeniu występowania wyłącznie krótkiej preambuły.



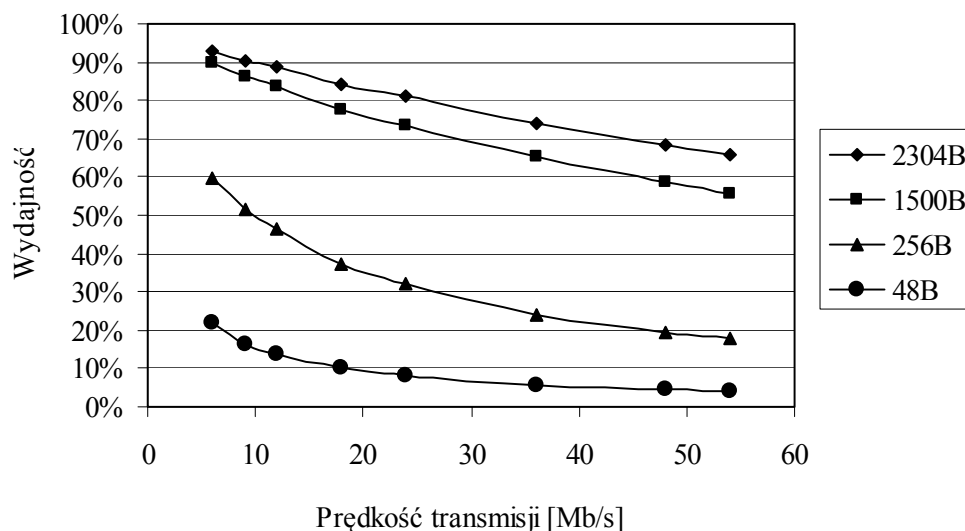
Rys. 4. Wydajność „podstawowego” wariantu standardu IEEE 802.11



Rys. 5. Wydajność protokołu 802.11b dla różnych pojemności pola danych – preambuła długa (lp) i krótka (sp)

3.3. IEEE 802.11a

Wyniki obliczeń dla warstwy fizycznej 802.11a, uzyskane dla różnych długości pola danych, pokazano na rys. 6. Warto zauważyć, że wraz ze wzrostem prędkości transmisji wydajność protokołu szybko maleje – nawet dla ramek o maksymalnej długości – od około 92% dla 6 Mb/s do około 66% dla 54 Mb/s. Spadek ten – podobnie jak dla 802.11b – jest liniowy dla ramek dłuższych i hiperboliczny dla krótszych.

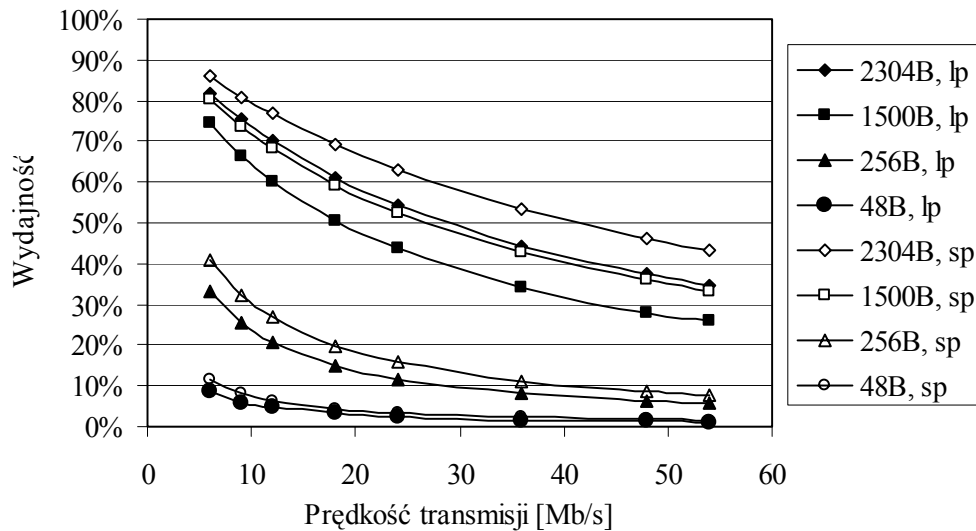


Rys. 6. Wydajność protokołu 802.11a dla różnych pojemności pola danych

3.4. IEEE 802.11g

Wyniki uzyskane dla warstwy fizycznej 802.11g pokazano na rys. 7. Obliczono je dla wariantu DSSS-OFDM z użyciem krótkiej lub długiej preambuły. Wyniki dla wariantu ERP-OFDM byłyby zbliżone do 802.11a, ale tylko wówczas, gdy żadne urządzenia starszego typu (802.11b lub 802.11 DSSS) nie są podłączone do sieci. W tym przypadku bowiem ERP-OFDM zastępuje się przez DSSS-OFDM. Z kolei wyniki dla ERP-CCK i ERP-PBCC są zbliżone do 802.11b (rys. 5).

Nietrudno zauważyć, że w porównaniu z warstwą 802.11a (rys. 6) spadek wydajności następuje znacznie szybciej. Jest to spowodowane użyciem preambuł zapewniających zgodność wsteczną, które są znacznie dłuższe od używanych w modulacji OFDM. Przykładowo, dla prędkości 54 Mb/s wydajność spada poniżej 50% nawet dla najdłuższych ramek. Podobnie do warstwy 802.11b różnice efektywności między ramkami 1500 B i 2304 B nie są duże. Natomiast różnica między krótką i długą preambułą jest najbardziej widoczna dla ramek 1500 B.



Rys. 7. Wydajność protokołu 802.11g dla różnych pojemności pola danych – preambuła długa (lp) i krótka (sp)

3.5. Porównanie warstw fizycznych

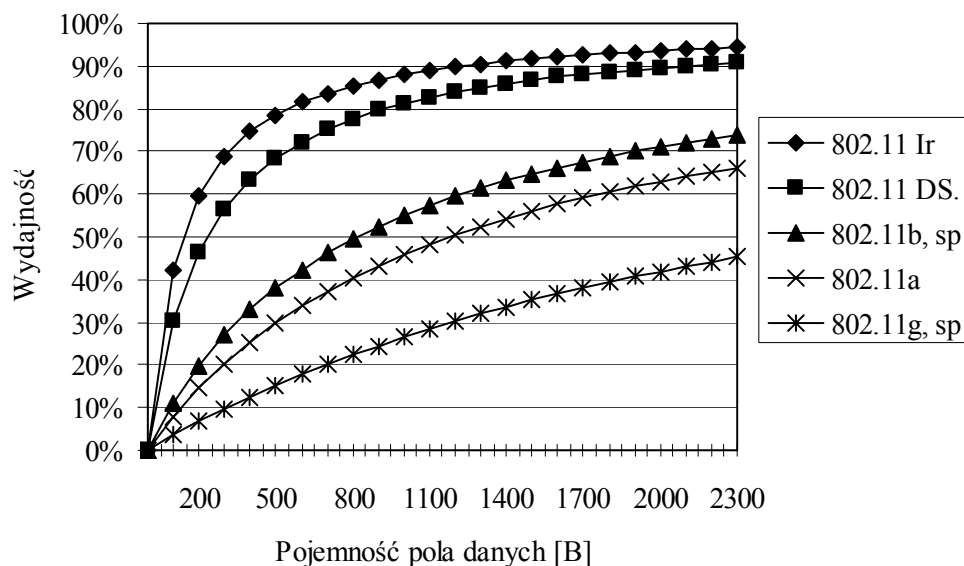
Opierając się na poprzednich wyliczeniach, wybrano kilka odmian warstw fizycznych dla dalszych porównań, których celem jest określenie wpływu pojemności pola danych ramki na wydajność protokołu. Wybrano następujące warstwy:

- 802.11 (podczerven) z prędkością transmisji 2 Mbps,
- 802.11 (DSSS) z prędkością 2 Mbps,
- 802.11b (krótka preambuła) z prędkością 11 Mbps,
- 802.11a z prędkością 54 Mbps,
- 802.11g (krótka preambuła) z prędkością 54 Mbps.

Dzięki takiemu wyborowi można porównać najwyższe prędkości transmisji określone dla danej warstwy fizycznej. Podejście takie jest uzasadnione, ponieważ niższe prędkości używane są rzadko – gdy moc sygnału jest zbyt niska, aby zapewnić prawidłowy odbiór z prędkościami wyższymi. Ponadto, różnice między osiąganymi osiągnięciami poszczególnych warstw przy małych prędkościach transmisji nie są duże – wydajność przekracza wówczas 85%.

Wyniki porównania pokazane są na rys. 8. Nietrudno zauważyć, że wraz ze wzrostem prędkości transmisji spada wydajność protokołu. Jest to spowodowane narzutem warstwy fizycznej wynikającym z użycia preambuły i nagłówka, przesyłanych z najniższą prędkością dla danej warstwy. Istotne jest przy tym, że najważniejsze i najpopularniejsze odmiany warstw fizycznych – 802.11 DSSS, 802.11b i 802.11g (z wyjątkiem ERP-OFDM) mają taki sam format preambuły i nagłówka. Biorąc pod uwagę zasady wymiany ramek (rys. 3) i prędkości transmisji elementów warstwy fizycznej, nietrudno zauważyć, iż, wraz ze zwiększaniem prędkości transmisji, czas transmisji ramki warstwy liniowej maleje, co oczywiście przynosi wzrost efektywnej przepu-

stowości. Jednakże elementy warstwy fizycznej przesyła się nadal z tymi samymi (niskimi) prędkościami, zatem narzut protokołu wynikający z ich użycia rośnie, a jednocześnie wydajność protokołu maleje. Wyjaśnienie to nie dotyczy przesyłów realizowanych wyłącznie z użyciem modulacji OFDM, np. w sieci 802.11a czy 802.11g ERP-OFDM. Tym niemniej, narzut warstwy fizycznej jest tam i tak wystarczająco duży, by wydajność tych protokołów przy najwyższych prędkościach transmisji była niższa niż dla 802.11b.



Rys. 8. Wpływ długości pola danych na wydajność protokołu 802.11 dla różnych warstw fizycznych

4. Podsumowanie

Uzyskane wyniki nie są zaskakujące. Niska efektywność sieci zgodnych ze standardem IEEE 802.11, szczególnie w porównaniu z przewodową siecią Ethernet, jest powszechnie znana. Efektywna przepustowość protokołu TCP w sieci Ethernet o prędkości 100 Mb/s osiąga powyżej 8 MB/s. W sieci 802.11 o prędkości 54 Mb/s – co można uznać za połowę prędkości sieci Ethernet – efektywna przepustowość jest znacznie niższa i osiąga na ogół 2,5-3 MB/s. Jest to zatem trzy-, a nawet czterokrotnie mniej, niż w sieci Ethernet. Uzyskane wyniki analityczne nie są zbyt odległe od tych wartości, a różnice wynikają z tego, że przedstawione obliczenia uwzględniają protokół warstwy fizycznej i liniowej, podczas gdy w rzeczywistej sieci występuje także stos protokołów TCP/IP, który także wnosi określony narzut. Tym niemniej przedstawione rozważania wyjaśniają przyczyny względnie niskiej wydajności bezprzewodowych sieci lokalnych.

Prawdopodobnie jedną z przyczyn niezadowolających osiągnięć standardu 802.11 jest zmienna prędkość transmisji. W każdej odmianie warstwy fizycznej nagłówek PLCP zawiera informację o prędkości transmisji, z jaką przesyłana będzie ramka war-

stwy liniowej. Nagłówek ten musi jednak być poprawnie odebrany i zrozumiany przez możliwie jak najwięcej urządzeń, nie wyłączając tych, w których zaimplementowano jedynie niektóre prędkości transmisji. Ponadto, przesył z najniższą prędkością jest najmniej narażony na błędy transmisji, zapewnia więc łatwą adaptowalność sieci do panujących w jej otoczeniu warunków. Niestety, ceną, jaką trzeba zapłacić za elastyczność sieci, jest jej zmniejszona efektywność. Warto przy tym zauważyć, że preambuły stosowane są nie tylko w sieciach bezprzewodowych, lecz także np. w sieci Ethernet. W sieci tej preambuła jest jednak stosunkowo krótka, a przesyła się ją zaw sze z tą samą prędkością, co pozostałą część ramki.

Innymi czynnikami, wpływającymi negatywnie na wydajność sieci 802.11, są mechanizmy typowe dla sieci bezprzewodowych, jak np. potwierdzenia warstwy liniowej oraz ramki sterujące RTS i CTS. Potwierdzenia podnoszą niezawodność przesyłu, natomiast RTS i CTS używa się dla eliminacji stacji ukrytych w procedurze dostępu do łącza. Żaden z tym mechanizmów nie występuje w sieci Ethernet. Warto zatem zauważyć, iż pakiet TCP odpowiada pojedynczej ramce sieci Ethernet, natomiast w sieci 802.11 – ciągowi ramek. Nawet gdy ramki sterujące RTS i CTS nie są używane, każdy pakiet TCP wymaga przesłania aż dwóch ramek (Data i Ack), poprzedzonych preambułą i nagłówkiem warstwy fizycznej.

Efektywne wykorzystanie wyższych prędkości transmisji, sięgających kilkuset Mb/s, wymaga zatem modyfikacji standardu. Przy obecnych prędkościach transmisji wystarczającym mechanizmem wydaje się potwierdzenie blokowe (ang. *Block Acknowledgement*) [2], pozwalające wspólnie potwierdzać wiele ramek danych. Dla przyszłych prędkości potrzebne są jednak dalej idące poprawki, np. łączenie ramek (ang. *Frame Aggregation*), pozwalające przesłać ciąg ramek danych, poprzedzony tylko jedną preambułą i nagłówkiem warstwy fizycznej [5].

Literatura

1. Qiao D., Choi S., Shin K. G.: Goodput Analysis and Link Adaptation for IEEE 802.11a Wireless LANs. *IEEE Trans. on Mobile Computing*, Vol. 1, No. 4, Oct-Dec. 2002, pp. 278-292.
2. IEEE Standard for Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. IEEE Std 802.11™-2007. IEEE, New York, 2007.
3. Beech W. A., Nielsen D. E., Taylor J.: AX.25 Link Access Protocol for Amateur Packet Radio. Tucson Amateur Packet Radio Corporation, Richardson, TX, 1998.
4. Tanenbaum A. S.: *Computer Networks* (4th ed.). Prentice Hall, Upper Saddle River 2003.
5. Skordoulis D., Ni Q., Chen H.-H., Stephens A. P., Liu C., Jamalipour A.: IEEE 802.11n MAC Frame Aggregation Mechanisms for Next-Generation High-Throughput WLANs. *IEEE Wireless Communications*, Vol. 15, No. 1, Feb. 2008, pp. 40-47.