

ALESSANDRO PAULO DE OLIVEIRA

AVALIAÇÃO DE DESEMPENHO DO TCP EM SISTEMAS VBLAST CODIFICADOS

Dissertação submetida ao Programa de Pós-Graduação em Informática da Pontifícia Universidade Católica do Paraná como requisito parcial para a obtenção do título de Mestre em Informática.

Curitiba PR
Dezembro de 2008

ALESSANDRO PAULO DE OLIVEIRA

AVALIAÇÃO DE DESEMPENHO DO TCP EM SISTEMAS VBLAST CODIFICADOS

Dissertação submetida ao Programa de Pós-Graduação em Informática da Pontifícia Universidade Católica do Paraná como requisito parcial para a obtenção do título de Mestre em Informática.

Área de concentração: *Ciência da Computação*

Orientador: Prof. Dr. Marcelo Eduardo Pellenz

Curitiba PR
Dezembro de 2008

Oliveira, Alessandro Paulo.

Avaliação de Desempenho do TCP em Sistemas VBLAST Codificados. Curitiba, 2008. 63p.

Dissertação - Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Informática.

1. TCP 2. MIMO 3. VBLAST 4. RLP. I. Pontifícia Universidade Católica do Paraná. Centro de Ciências Exatas e de Tecnologia. Programa de Pós-Graduação em Informática II-t.

*Esta folha deve ser substituída pela ata de defesa devidamente assinada,
que será fornecida pela secretaria do programa após a defesa.*

A Deus todo poderoso criador e mantenedor do universo. A minha mãe Neuza Aparecida Marquês e a meu pai Vicente Paulo de Oliveira (in memoriam).

Agradecimentos

A Deus por permitir e ajudar-me a realizar mais um sonho. A minha mãe, Neuza, que sempre ajudou-me em tudo o que precisei. Ao Prof. Dr. Marcelo pela oportunidade disponibilizada e pela boa orientação. Ao Prof. Dr. Richard por disponibilizar materiais e recursos necessários para a realização deste trabalho. A todas as pessoas que de alguma maneira contribuíram para a realização deste trabalho.

Resumo

Os novos padrões para redes locais sem fio e redes metropolitanas sem fio, como por exemplo os padrões IEEE 802.11 n (WiFi) e IEEE 802.16 (Wimax) utilizam as técnicas MIMO de maneira a propiciar serviços multimídia com altas taxas de transmissão. O protocolo de controle de transmissão (TCP) é geralmente implementado na camada de transporte para garantir comunicação confiável dos dados. Por esta razão, a otimização do desempenho do protocolo TCP para sistemas MIMO é atualmente uma importante área de pesquisa. Neste estudo são propostos e avaliados esquemas com diferentes técnicas de modulação. Os esquemas empregam códigos convolucionais com diferentes taxas de codificação e capacidade de correção. O desempenho dos diferentes esquemas VBLAST, em termos de taxa de erro de bit e taxa de erro de bloco, são analisados através de simulações computacionais. Os dados de desempenho da camada física são utilizados para avaliar o desempenho do TCP/RLP através de simulações computacionais. Baseado nas curvas de desempenho do TCP foi proposto um esquema de transmissão adaptativo de maneira a otimizar o seu desempenho para cada faixa de relação sinal ruído. **Palavras-chave:** TCP, MIMO, RLP.

Abstract

The standards to local wireless network and metropolitan wireless network, such as IEEE 802.11 n (WiFi) and IEEE 802.16 (Wimax) use MIMO technique to provide high rate multimedia service transmission. The transmission control protocol (TCP) is generally employed at the transport layer for reliable data communication. Thus, the optimization of TCP performance over MIMO systems is currently an important research area. In this study, are proposal and evaluate schemes with different modulation techniques. The schemes use convolutional codes with different coding rates and correction capacity. The performance of different VBLAST coded schemes, in terms of bit error rate and frame error rate, are analysed through the computer simulations. The datas of physical layer performance are used to evaluate TCP/RLP performance through computer simulation. By analysing TCP performance was proposal aadaptive transmittion scheme in wich TCP performance is optimized to each signal noise ratio range.

Keywords: TCP, MIMO, RLP.

Sumário

Resumo	xi
Abstract	xiii
Lista de Figuras	xvii
Lista de Tabelas	xix
Lista de Abreviações	xx
1 Introdução	1
1.1 Motivação	1
1.2 Proposta	2
1.3 Organização	2
2 Sistemas MIMO	5
2.1 O Canal de Rádio	5
2.2 Modelagem do Canal de Rádio	5
2.3 Técnicas de Diversidade	8
2.4 Diversidade de Transmissão	10
2.5 Esquemas MIMO	12
2.6 Tecnologia BLAST	14
2.6.1 Tecnologia V-BLAST	15
2.7 Conclusão	18
3 O Protocolo TCP	19
3.1 Controle de Transmissão Confiável	20
3.2 Controle de Fluxo	21
3.3 Controle de Congestionamento	22
3.4 Desempenho do TCP em Enlaces sem Fio	23
3.5 Estratégias para Minimizar a Redução do <i>Throughput</i> do TCP em Enlaces sem Fio	24
3.5.1 Mudanças na Implementação do TCP	24
3.5.2 Dividir a Conexão do TCP	26
3.5.3 Retransmissão Local	28
3.6 Conclusão	28

4	Esquema Proposto	29
4.1	Avaliação da Camada Física	29
4.2	Avaliação da Camada de Rede	31
4.3	Conclusão	36
5	Conclusões	39

Lista de Figuras

2.1	Comparação de desempenho da modulação BPSK no canal AWGN e Rayleigh.	7
2.2	Técnica - <i>Selection Combining</i> .	9
2.3	Técnica - <i>Maximal Ratio Combining</i> .	9
2.4	Desempenho do esquema de diversidade MRC.	10
2.5	Esquema do Alamouti.	11
2.6	Codificação Espaço-Temporal.	11
2.7	Simulações do Esquema Alamouti e MRC.	13
2.8	Modelo de Canal MIMO.	14
2.9	Arquitetura D-BLAST.	15
2.10	Arquitetura V-BLAST.	15
3.1	Transmissão com confirmação positiva.	20
3.2	Transmissão com ACK duplicado.	21
3.3	Perda de segmento por <i>timeout</i> .	22
3.4	Controle de fluxo do TCP.	22
3.5	Medida do atraso relativo.	24
3.6	Determinação do estado das filas de acordo com $D_{j,i}^F$.	25
3.7	Divisão de uma conexão TCP em duas conexões.	27
3.8	Arquitetura de retransmissão local.	28
4.1	Arquitetura Proposta.	29
4.2	Resultados das simulações da BER x SNR para o detector ZF.	30
4.3	Resultados das simulações da BER x SNR para o detector MMSE.	31
4.4	Resultados da FER x SNR para o detector ZF.	32
4.5	Resultados da FER x SNR para o detector MMSE.	33
4.6	Cenário de rede implementado no ns2.	33
4.7	Desempenho do TCP para os esquemas VBLAST-ZF.	34
4.8	Desempenho do TCP para os esquemas VBLAST-MMSE.	35
4.9	Desempenho do TCP usando as estratégias otimizadas de transmissão para o detector VBLAST-MMSE.	36
4.10	Desempenho do TCP usando as estratégias otimizadas de transmissão para o detector VBLAST-ZF.	36
4.11	Comparação entre as estratégias otimizadas de transmissão para os detectores VBLAST-ZF e VBLAST-MMSE.	37

Lista de Tabelas

4.1	Esquemas codificados VBLAST simulados	30
4.2	Parâmetros de simulação da camada de enlace	34
4.3	Faixa de operação para os esquemas codificados VBLAST-ZF	35
4.4	Faixa de operação para os esquemas codificados VBLAST-MMSE	35

Lista de Abreviações

BER	<i>Bit Error Rate</i>
SNR	<i>Signal Noise Ratio</i>
SISO	<i>Single Input Single Output</i>
MIMO	<i>Multiple Input Multiple Output</i>
MISO	<i>Multiple Input, Single Output</i>
SIMO	<i>Single Input Multiple Output</i>
MRC	<i>Maximal Ratio Combining</i>
STBC	<i>Space Time Block Code</i>
TX	Transmissor
RX	Receptores
ACK	<i>Acknowledgment</i>
WPAN	<i>Wireless Personal Area Network</i>
WLAN	<i>Wireless Local Area Network</i>
WMAN	<i>Wireless Metropolitan Area Network</i>
AWGN	<i>Additive White Gaussian Noise</i>
DBLAST	<i>Diagonal Bell Layered Space-Time</i>
VBLAST	<i>Vertical Bell Layered Space-Time</i>
MMSE	<i>Minimum Mean Square Error</i>
ZF	<i>zero-forcing</i>
SIC	<i>Successive Interference Cancelling</i>
OSIC	<i>Ordered Successive Interference Cancelling</i>
BPSK	<i>Binary Phase Shift Keying</i>
QPSK	<i>Quadrature Phase Shift Keying</i>
QAM	<i>Quadrature Amplitude Modulation</i>
MAC	Media Access Control
WiMAX	<i>Worldwide Interoperability for Microwave Access</i>
WiFi	<i>Wireless Fidelity</i>
RLP	<i>Radio Link Protocol</i>
TCP	<i>Transmission Control Protocol</i>
FER	<i>Frame Error Rate</i>
FTP	<i>File Transfer Protocol</i>
MSS	<i>Maximum Segment Size</i>
CW	<i>Congestion Window</i>
ARQ	<i>Automatic Repeat Request</i>
RTT	<i>Round Trip Time</i>
NACK	<i>Negative Acknowledgment</i>

Capítulo 1

Introdução

A utilização dos sistemas MIMO *Multiple Input Multiple Output* nas especificações da camada física dos padrões IEEE 802.11 n (WiFi) e IEEE 802.16 (Wimax), tornam o estudo de tais sistemas de grande importância para as comunicações sem fio. Pesquisas a nível de *crosslayer*, onde as camadas superiores de protocolo trocam informações com a camada física e a camada MAC (*Media Access Control*) para melhor explorar os recursos da rede, fazem-se necessárias para completar a análise do impacto que os sistemas MIMO provocam nas redes sem fio. Existem diversos trabalhos recentemente publicados na área de projeto *crosslayer*, que investigam o impacto da camada física nas camadas superiores de protocolo, em especial as interações entre a camada física e o protocolo TCP da camada de transporte [Balakrishnan et al., 1996].

1.1 Motivação

Os sistemas MIMO, quando comparados com os sistemas SISO *Single Input Single Output*, oferecem vantagens fundamentais relacionadas a taxa de transmissão e à confiabilidade do sistema. Os sistemas MIMO podem ser implementados de duas maneiras distintas: multiplexação espacial ou codificação espaço temporal [Toledo and Wang, 2006]. A multiplexação espacial oferece alta eficiência espectral ou altas taxas de transmissão. Neste esquema, as diferentes antenas são utilizadas para transmitir diferentes símbolos de informação codificados, ao mesmo tempo. Por outro lado, a codificação espaço temporal melhora a confiabilidade do enlace, explorando o ganho de diversidade. Estes sistemas geralmente implementam códigos de bloco espaço temporais, denominados STBC (*Space Time Block Code*). Nesta abordagem as diferentes antenas transmitem diferentes símbolos codificados do mesmo símbolo de informação, num dado instante de tempo. A comunicação se torna mais confiável, porém a taxa de transmissão pode ser diminuída, sendo ainda menor que em sistemas SISO, dependendo da taxa de codificação do STBC.

Vários padrões de redes sem fio, como WPAN (*Wireless Personal Area Network*), WLAN (*Wireless Local Area Network*), WMAN (*Wireless Metropolitan Area Network*) e redes celulares devem utilizar sistemas MIMO [D.Gesbert et al., 2003]. As normas internacionais de WMAN, em específico o WiMAX [Yang et al., 2006], e Celular (3GPP/3GPP2) incluem MIMO em suas especificações.

Nos últimos anos, várias pesquisas foram feitas sobre os sistemas MIMO. Porém, somente alguns trabalhos têm focado a análise do desempenho de protocolos da camada de

transporte quando tendo os sistemas MIMO presente na camada física. O protocolo TCP, atualmente, é o protocolo da camada de transporte mais utilizado devido à popularização crescente da internet. Sendo assim, pesquisas sobre o desempenho do TCP, nas mais diversas condições do canal, têm se tornado importante.

Nas redes sem fio com múltiplos saltos e redes em malha as condições do canal são instáveis e imprevisíveis. Este fato é característico de canais sem fio. Diante desta situação o desempenho do protocolo TCP fica prejudicado, pois o mesmo não foi projetado para adaptar-se a estas condições variantes de canal. Nestas condições, seria de suma importância que os sistemas de transmissão sem fio levassem em consideração o comportamento aleatório do canal para que o funcionamento do protocolo na camada de transporte, TCP, não seja penalizado prejudicando o seu bom funcionamento. Para solucionar este problema as pesquisas feitas seguem, basicamente, três diferentes caminhos:

- Modificação no protocolo TCP de maneira que o mesmo fosse capaz de adaptar-se ao comportamento imprevisível do canal. Esta solução foi proposta em [S.Floyd and T.Henderson, 1999].
- Dividir a conexão do TCP em duas partes: conexão que tem seu fluxo em uma rede cabeada e a outra que tem seu fluxo em uma rede sem fio [Bakre and Badrinath, 1995].
- Utilizar mecanismo de retransmissão na camada de enlace ARQ (*Automatic Repeat Request*), pois é a maneira mais comum de esconder o comportamento aleatório do canal do protocolo TCP. Esta solução já está presente nos padrões sem fio tendo como exemplo o RLP (*Radio Link Protocol*) utilizado no UMTS [RLC, 1999].

Contudo, este método de retransmissão local pode introduzir atrasos e variações na taxa de transmissão se um número excessivo de retransmissões ocorrer. Isso faz com que exista uma redução na taxa útil do enlace [Chan and Ramjee, 2002]. Portanto, uma solução baseada apenas em ARQ não é suficiente para melhorar o desempenho do TCP em enlaces sem fio.

1.2 Proposta

Este trabalho tem como objetivo investigar diferentes esquemas VBLAST codificados para diferentes tipos de modulações e propor uma estratégia de transmissão adaptativa para ser usada em redes sem fio com múltiplos saltos e redes em malha. Esta estratégia de transmissão adaptativa associada ao protocolo RLP constituirá uma solução para diminuir o problema do canal sem fio, pois permite melhorar o desempenho fim-a-fim do TCP para aplicações cliente-servidor e peer-to-peer. Os esquemas codificados utilizam códigos convolucionais com diferentes taxas. O objetivo é avaliar qual sistema desempenha melhor em uma faixa de relação sinal-ruído específica. Como resultado derivamos um conjunto de regras para selecionar o melhor esquema VBLAST codificado em função da relação sinal-ruído média entre os nós que estão se comunicando.

1.3 Organização

Os demais capítulos desta dissertação estão organizados da seguinte forma:

- Capítulo 2: Uma análise geral sobre os sistemas MIMO.
- Capítulo 3: Uma análise funcional sobre o TCP e um estudo quando o mesmo está operando em sistemas sem fio.
- Capítulo 4: Apresenta a análise para os diversos sistemas VBLAST codificados, considerando os detectores ZF e MMSE, propondo uma transmissão adaptativa.
- Conclusão

Capítulo 2

Sistemas MIMO

O interesse pelas comunicações sem fio vem crescendo ao longo dos anos. O principal motivo por esse interesse é a mobilidade dos equipamentos que a utilizam. No entanto, pelo menos inicialmente, os sistemas que utilizavam comunicações sem fio tinham um desempenho, no que diz respeito a velocidade de transmissão e confiabilidade, inferior quando comparado com os sistemas que utilizavam as comunicações cabeadas. O canal sem fio degrada mais severamente o sinal do que o canal cabeado devido a um fenômeno chamado de desvanecimento. Em face a este problema é que ao longo dos anos tem sido feitas pesquisas para tornar os sistemas de comunicações sem fio mais rápidos e mais confiáveis. O resultado destas pesquisas apontaram para os sistemas MIMO que podem, pela maneira que são utilizados, oferecer maior capacidade de transmissão e confiabilidade quando comparado com os sistemas de comunicações sem fio tradicionais.

2.1 O Canal de Rádio

Um canal de rádio impõe limitações no desempenho de sistemas de comunicações sem fio. O caminho de transmissão, entre transmissor e receptor, pode variar de uma simples linha direta, linha de visada, para um caminho que pode ser severamente obstruído por prédios, montanhas, e etc. Além disso, objetos movendo-se no canal de rádio e o próprio sistema de recepção movendo-se com relação ao sistema de transmissão, tornam os canais de rádio, diferentemente de canais com fio que são estacionários e previsíveis, extremamente aleatórios e de difícil análise. Portanto, a modelagem de um canal de rádio tem sido, historicamente, uma das mais difíceis partes do projeto de sistemas de rádio, e é tipicamente feita com um modelo estatístico, baseado em medidas feitas especificamente para um desejado sistema de comunicação ou alocação de espectro [Rappaport, 1996].

2.2 Modelagem do Canal de Rádio

Entre os problemas enfrentados por um sinal, ao se propagar em um canal de rádio, pode-se destacar dois: desvanecimento em larga escala e desvanecimento em pequena escala.

O desvanecimento em larga escala representa a atenuação média da potência do sinal, ou a perda de percurso, devido ao movimento do mesmo sobre grandes distâncias (várias centenas ou milhares de metros). Este fenômeno é afetado por notórios contornos de terreno (tais

como montanhas, florestas, prédios, e etc) entre o transmissor e o receptor [Rappaport, 1996] [Sklar, 1997].

O desvanecimento em pequena escala, ou simplesmente desvanecimento, é usado para descrever a rápida flutuação da amplitude do sinal de rádio sobre um curto período de tempo ou distancia de separação entre transmissor e receptor (numa escala da ordem do comprimento de onda do sinal). Neste caso, os efeitos da perda de percurso podem ser ignorados. Esse desvanecimento é causado pela interferência entre duas ou mais réplicas do sinal transmitido que chega no receptor em tempos ligeiramente diferentes. Estas ondas, chamadas ondas multipercurso, combinam-se na antena receptora para dar um sinal resultante que pode variar amplamente em amplitude, fase e atraso.

No desvanecimento em pequena escala, quando entre o transmissor e o receptor não tem linha de visada, a distribuição de Rayleigh é comumente usada para descrever estatisticamente a envoltória do sinal recebido na antena receptora. A função de densidade de probabilidade de Rayleigh é definida como:

$$p(r) = \begin{cases} \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} & 0 \leq r \leq \infty \\ 0 & r < 0 \end{cases} \quad (2.1)$$

onde σ^2 é a variância. Para o caso em que existe uma linha de visada entre o transmissor e o receptor o desvanecimento é representado estatisticamente por uma distribuição de Rice. A função de densidade de probabilidade da distribuição de Rice é definida como:

$$p(r) = \begin{cases} \frac{r}{\sigma^2} e^{-\frac{r^2+A^2}{2\sigma^2}} I_0\left(\frac{Ar}{\sigma^2}\right) & A \geq 0 \quad r \geq 0 \\ 0 & r < 0 \end{cases} \quad (2.2)$$

onde A denota o pico de amplitude do sinal dominante e I_0 é a função de Bessel modificada de primeira espécie de ordem zero [Rappaport, 1996].

A figura 2.1 apresenta um gráfico onde é mostrado a curva da BER (*Bit Error Rate*) em função da relação sinal ruído SNR (*Signal-to- Noise Ratio*) para um sistema de comunicação BPSK no canal AWGN (*Additive White Gaussian Noise*) e no canal Rayleigh. Deste gráfico pode-se observar que o desempenho num canal Rayleigh (canal sem fio) é significativamente pior que no caso do canal AWGN. Por exemplo para uma BER= 10^{-2} é necessário uma SNR de aproximadamente 7 dB maior no canal Rayleigh.

Em se tratando de canais multipercurso, o canal pode ser analisado sob o ponto de vista da frequência ou do tempo.

Sob o ponto de vista da frequência, o parâmetro importante que é definido é a largura de banda de coerência do canal. A largura de banda de coerência do canal é uma medida estatística da faixa de frequências sobre a qual o canal pode ser considerado plano, ou seja, um canal que permite passar todas as componentes espectrais, de um sinal, com ganho aproximadamente igual e fase linear. Em outras palavras, a largura de banda de coerência é a faixa de frequências sobre a qual duas componentes de frequência tem um forte potencial para correlação da amplitude [Rappaport, 1996]. Neste contexto, define-se que o canal é seletivo em frequência se a largura de banda de coerência do canal for menor que a largura de banda do sinal. Dois sinais com separação de frequência maior do que a largura de banda de coerência do canal são afetados de maneira diferente pelo canal, tendo uma correlação baixa entre eles. Em uma situação contrária, sendo a largura de banda de coerência maior que a largura de banda ocupada pelo

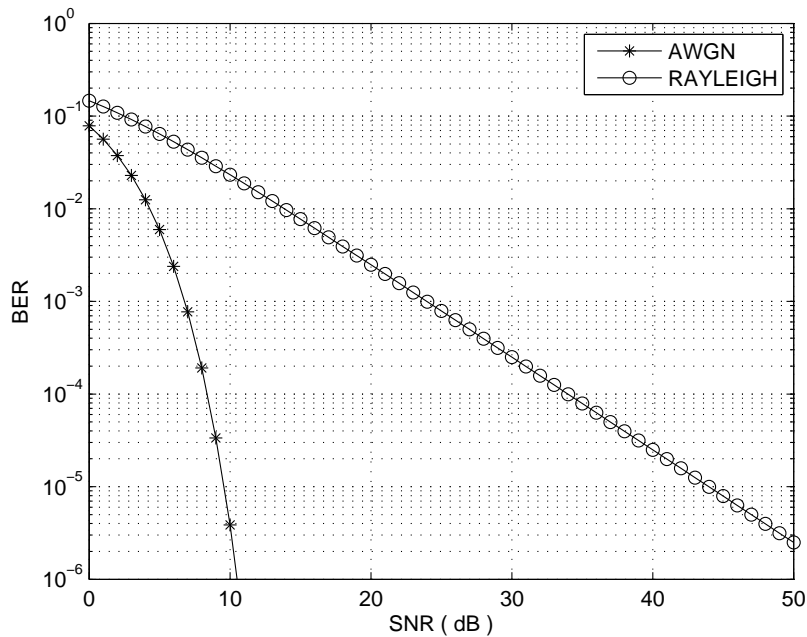


Figura 2.1: Comparação de desempenho da modulação BPSK no canal AWGN e Rayleigh.

sinal transmitido, o canal afetará por igual todas as componentes espectrais do sinal e o desvanecimento é definido como desvanecimento plano. Neste trabalho, considera-se que a largura de banda do sinal é menor que a largura de banda de coerência do canal, sendo assim, todas as componentes do sinal serão afetadas igualmente pelo canal.

Do ponto de vista do tempo, um parâmetro importante é o tempo de coerência. O tempo de coerência é na verdade uma medida estatística da duração do tempo sobre o qual a resposta ao impulso do canal é essencialmente invariante. Em outras palavras, o tempo de coerência é a duração de tempo sobre o qual dois sinais recebidos tem um forte potencial para correlação de amplitude [Rappaport, 1996] [Sklar, 1997]. O canal é dito seletivo no tempo se o tempo de coerência do canal for menor quando comparado a duração do sinal recebido. Caso contrário, o canal não é seletivo no tempo, ou plano em tempo.

Existe ainda a classificação do canal como desvanecimento lento ou desvanecimento rápido. Um canal com desvanecimento rápido é aquele em que o período do símbolo do sinal transmitido é maior que o tempo de coerência. Um canal com desvanecimento lento é aquele em que o período do símbolo do sinal transmitido é menor que o tempo de coerência [Rappaport, 1996] [Sklar, 1997].

Uma outra classificação importante do canal é quando ele é quasi-estático. Neste caso, a resposta ao impulso do canal varia tão lentamente durante a duração de um bloco de comunicação que pode ser considerada invariável, apenas variando de um bloco para outro [Jr et al., 2005]. Isto significa que o canal não mudará durante a transmissão de um bloco de dados, sendo possível variar de um bloco para outro.

2.3 Técnicas de Diversidade

O fenômeno do desvanecimento é uma característica inerente dos canais sem fio. Dada esta realidade física, uma das maneiras de se diminuir o problema é usando a diversidade, que pode ser vista como uma forma de redundância. Se várias réplicas do sinal de informação podem ser transmitidos simultaneamente sobre canais com desvanecimento independentes, então há uma boa probabilidade de que pelo menos um dos sinais recebidos não será severamente degradado pelo desvanecimento do canal. Existem alguns métodos para a diversidade onde pode se destacar três: diversidade em frequência, diversidade em tempo e diversidade espacial.

Na Diversidade em Frequência o mesmo sinal de informação é transmitido por meio de várias portadoras que são espaçadas umas das outras suficientemente para fornecer versões do sinal com desvanecimento independente. Isto pode ser alcançado escolhendo-se um espaçamento entre as frequências das portadoras igual ou maior do que a largura de banda do canal [Haykin and Moher, 2005].

Na Diversidade em Tempo o mesmo sinal de informação é transmitido em diferentes instantes de tempo, com intervalo entre os sucessivos instantes de tempo sendo igual ou maior do que o tempo de coerência do canal. Se o intervalo é menor do que o tempo de coerência do canal, pode-se conseguir ainda alguma diversidade, mas com o custo no desempenho. A diversidade no tempo pode ser comparada ao uso de um código de repetição [Haykin and Moher, 2005].

A Diversidade Espacial é o uso de múltiplas antenas transmissoras ou receptoras, ou ambas, separadas de maneira a garantir a descorrelação (independência) de desvanecimento entre os caminhos percorridos pelo sinal [Haykin and Moher, 2005] [Proakis, 2001]. Este espaçamento entre as antenas adjacentes receptoras deve ser de 10 a 20 comprimentos de onda de rádio. Define-se três tipos de diversidade espacial de acordo com a localização do equipamento com múltiplas antenas nas extremidades do sistema de comunicação sem fio: múltiplas antenas na transmissão e uma única antena de recepção MISO (*Multiple Input, Single Output*), múltiplas antenas tanto na transmissão como na recepção MIMO e múltiplas antenas receptoras e uma transmissora SIMO (*Single Input Multiple Output*).

A tecnologia MIMO resolve a questão fundamental de ter que lidar com duas realidades práticas das comunicações sem fio: um usuário de equipamento com energia de bateria limitada e um canal de RF (Rádio Frequencia) com largura de banda limitada [Haykin and Moher, 2005].

A combinação dos sinais provenientes de canais multipercurso, com desvanecimento descorrelacionado, com algum critério, traz uma melhora no desempenho do equipamento receptor final. As técnicas mais importantes de combinação desses sinais são três:

a) *Selection Combining*: É composto por múltiplas antenas receptoras ligadas a entrada de um circuito lógico. Este circuito lógico seleciona a antena com maior SNR, dentre todas as antenas receptoras, e a conecta na sua saída. A figura 2.2 apresenta este sistema.

b) *Maximal Ratio Combining* (MRC): A técnica de combinação por seleção não apresenta um desempenho satisfatório porque ele considera somente o sinal da antena que tem maior SNR, ignorando os outros sinais disponíveis nas outras antenas. A técnica MRC procura resolver esta limitação substituindo o bloco executado por um circuito lógico por um bloco que é executado por um combinador linear. A figura 2.3 e a equação 2.3 representam esta técnica [Haykin and Moher, 2005][Haykin, 2004].

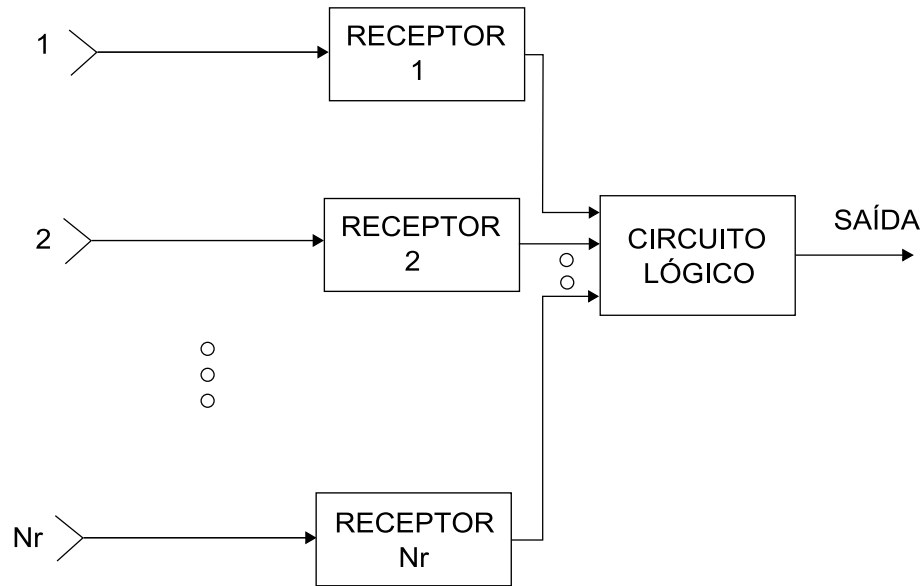


Figura 2.2: Técnica - *Selection Combining*.

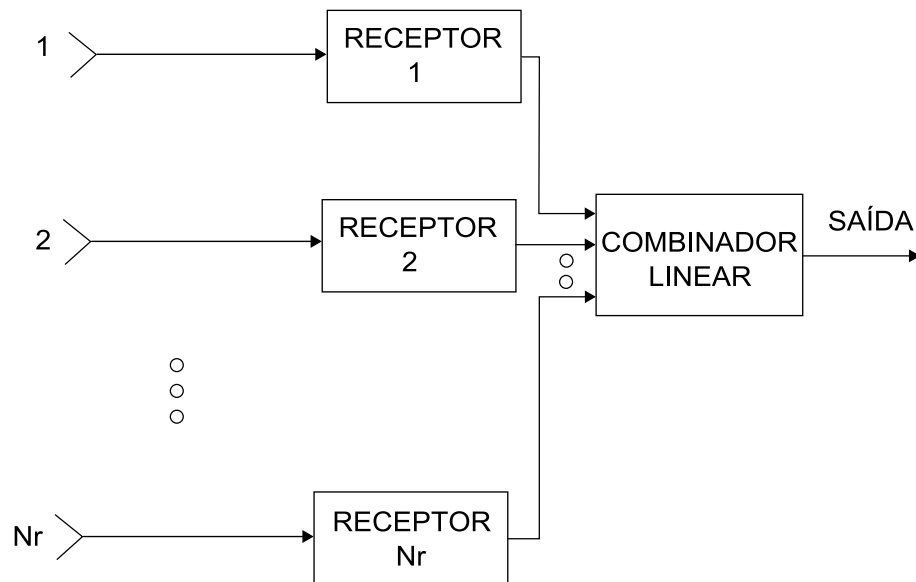


Figura 2.3: Técnica - *Maximal Ratio Combining*.

$$y(t) = s(t) \sum_{k=1}^{Nr} (a_k \alpha_k e^{j\theta_k}) + \sum_{k=1}^{Nr} (a_k w_k(t)), \quad (2.3)$$

onde a_k é o parâmetro de peso complexo que caracteriza o combinador linear. Este parâmetro é alterado constantemente conforme a variação do sinal, em cada antena, provocado pelo desvanecimento. Em outras palavras, este parâmetro representa o próprio ganho do canal cofasado, ($a_k = \alpha_k e^{-j\theta_k}$). O termo $s(t)$ é o sinal da antena transmissora e $w_k(t)$ é o ruído gaussiano branco aditivo.

Esta técnica tem a vantagem de apresentar em sua saída um valor aceitável de relação sinal/ruído mesmo se nenhuma antena individual tenha um sinal aceitável. Quando

o ruído é descorrelatado nos diversos ramos, esta técnica é a que apresenta a melhor redução estatística de desvanecimento do que qualquer combinador linear conhecido [Haykin and Moher, 2005][Haykin, 2004].

c) *Equal Gain Combining*: Em determinados casos, não há necessidade de implementar circuitos que fazem o ajuste dos parâmetros de pesos do MRC. Pode-se reduzir a complexidade do sistema. Nestes casos, a magnitude de a_k é unitária, ou seja, $a_k = e^{-j\theta_k}$. Dessa maneira, ainda existe a possibilidade do sistema produzir em sua saída um valor aceitável a partir de vários valores de entrada inaceitáveis com uma redução no desempenho do sistema mínima quando comparado ao MRC e com um desempenho superior à técnica *Selection Combining* [Haykin and Moher, 2005][Haykin, 2004].

A figura 2.4 apresenta um gráfico que demonstra a melhora no sistema de comunicação sem fio, com modulação BPSK, quando se utiliza diversidade espacial na recepção. Quanto maior for o número de antenas na recepção melhor será o desempenho do sistema.

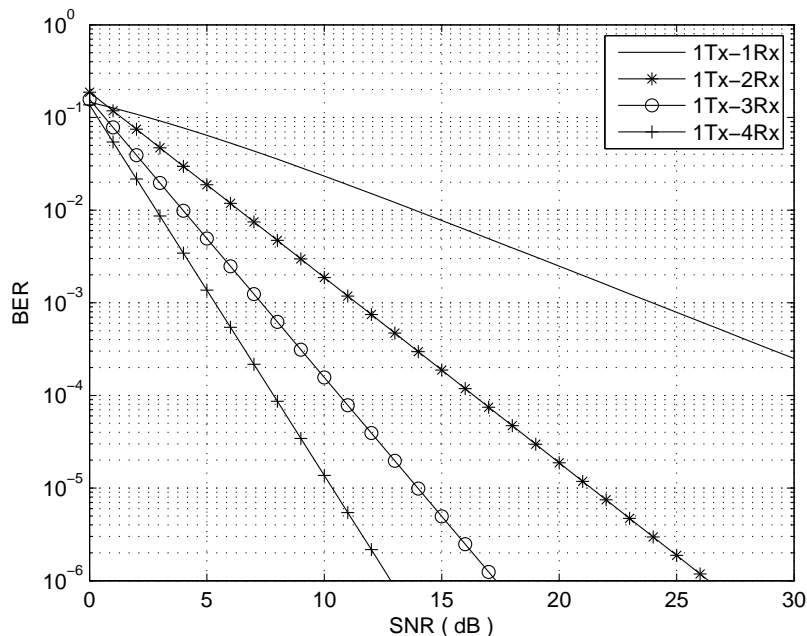


Figura 2.4: Desempenho do esquema de diversidade MRC.

2.4 Diversidade de Transmissão

A diversidade espacial na recepção tem problemas de implementação tais como: custo, tamanho e energia da unidade remota. Múltiplas antenas na recepção e circuitos de seleção ou combinação tornam as unidades remotas mais caras e maiores. Portanto, a técnica de diversidade tem sido aplicada quase que exclusivamente em estações base, pois estas servem centenas a milhares de unidades remotas. Sendo assim, é mais barato adicionar equipamento às estações base do que adicionar equipamento às unidades remotas.

Alamouti [Alamouti, 1998] propôs um simples esquema de diversidade com duas antenas transmissoras e uma receptora (MISO), podendo ser generalizado a N antenas receptoras,

(MIMO), fornecendo uma ordem de diversidade de $2N$ e com pequena complexidade computacional na decodificação dos códigos espaço-temporais. A figura 2.5 apresenta o Esquema do Alamouti. Este esquema possui três etapas:

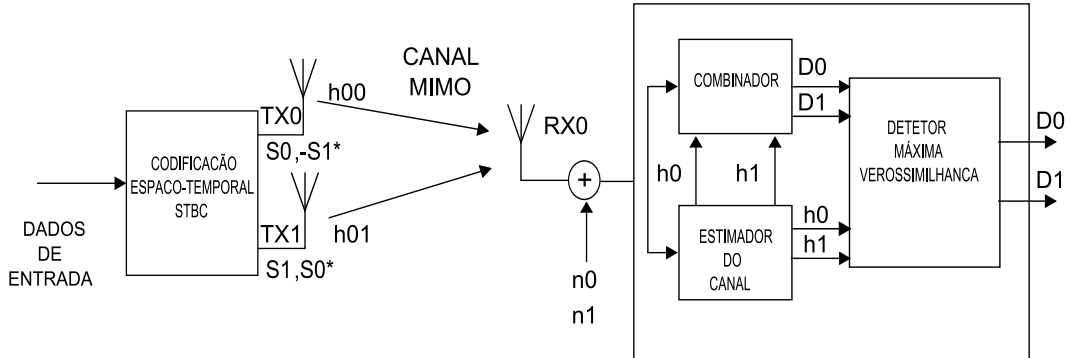


Figura 2.5: Esquema do Alamouti.

a) Codificação e seqüência de transmissão: Para um dado período de símbolo, duas antenas transmitem dois sinais simultaneamente. Da antena zero é transmitido um sinal denotado por s_0 e da antena um s_1 . No próximo período de símbolo, da antena zero é transmitido $-s_1^*$ e da antena um s_0^* onde $(*)$ é a operação conjugado complexo. Esta codificação é feita em espaço e tempo, por isso recebe a denominação de codificação espaço-temporal STBC (*Space Time Block Code*). A figura 2.6 ilustra a codificação espaço-temporal de Alamouti. O canal no

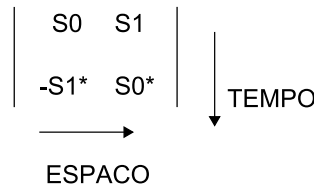


Figura 2.6: Codificação Espaço-Temporal.

tempo t pode ser modelado como uma distorção multiplicativa h_0 para a antena transmissora zero e h_1 para antena um. Assumindo que o canal permanece constante sobre os dois períodos de símbolos consecutivos, os sinais recebidos podem ser expressos como:

$$r_0 = r(t) = h_0 s_0 + h_1 s_1 + n_0 \quad (2.4)$$

$$r_1 = r(t+T) = -h_0 s_1^* + h_1 s_0^* + n_1 \quad (2.5)$$

onde n_0 e n_1 são variáveis aleatórias complexas representando o ruído no receptor e T é o período do símbolo. Estas equações podem ser equivalentemente escritas como [Stamoulis and Al-Dhahir, 2003]:

$$\begin{bmatrix} r_0 \\ r_1^* \end{bmatrix} = \begin{bmatrix} h_{00} & h_{01} \\ -h_{01}^* & h_{00}^* \end{bmatrix} \cdot \begin{bmatrix} s_0 \\ s_1 \end{bmatrix} + \begin{bmatrix} n_0 \\ n_1 \end{bmatrix} \quad (2.6)$$

b) Esquema de combinação: o combinador gera os dois seguintes sinais que são enviados ao detetor de máxima verossimilhança:

$$D_0 = h_{00}^* r_0 + h_{01} r_1^* \quad (2.7)$$

$$D_1 = h_{01}^* r_0 - h_{00} r_1^* \quad (2.8)$$

c) Regra de decisão de máxima verossimilhança: Os sinais das equações 2.7 e 2.8 são enviados a um detector de máxima verossimilhança, o qual estimará os sinais e com probabilidade de erro mínima. Com este novo esquema é possível fornecer diversidade ao sistema com apenas uma antena no receptor e mais de uma antena no transmissor. Tradicionalmente, a diversidade acontecia com várias antenas na recepção e uma na transmissão. O esquema de detecção pode ser generalizado para N antenas receptoras conforme definido na equação 2.9.

$$\begin{bmatrix} r_{0,0} \\ r_{0,1}^* \\ \vdots \\ r_{N-1,0} \\ r_{N-1,1}^* \end{bmatrix} = \begin{bmatrix} h_{00} & h_{01} \\ -h_{01}^* & h_{00}^* \\ \vdots & \vdots \\ h_{N-1,0} & h_{N-1,1} \\ -h_{N-1,1}^* & h_{N-1,0}^* \end{bmatrix} \cdot \begin{bmatrix} s_0 \\ s_1 \end{bmatrix} + \begin{bmatrix} n_{0,0} \\ n_{0,1} \\ \vdots \\ n_{N-1,0} \\ n_{N-1,1} \end{bmatrix} \quad (2.9)$$

onde $r_{i,t}$ é o i -ésimo sinal recebido e $n_{i,t}$ é o ruído na antena receptora i durante o tempo t . Os sinais detectados são:

$$D_0 = \sum_{i=0}^{N-1} (h_{i0}^* r_{i0} + h_{i1} r_{i1}^*) \quad (2.10)$$

$$D_1 = \sum_{i=0}^{N-1} (h_{i1}^* r_{i0} + h_{i0} r_{i1}^*) \quad (2.11)$$

Estes símbolos estimados, D_0 e D_1 , são obtidos através de um detector de máxima verossimilhança. A figura 2.7 apresenta um gráfico comparando o esquema do Alamouti com o MRC. Desta figura observa-se que o desempenho do esquema do Alamouti é 3 dB pior do que o MRC. Isto acontece porque cada antena transmissora irradia metade da potência de uma antena transmissora quando operando no sistema MRC. Este fato assegura que a potência gasta para transmitir no esquema do Alamouti é a mesma que a gasta para transmitir no sistema MRC. No entanto, se cada antena transmissora do esquema Alamouti transmitir com a mesma potência da antena do sistema MRC o desempenho seria o mesmo.

2.5 Esquemas MIMO

Nos sistemas MIMO considera-se que os arranjos de M antenas transmissoras e N receptoras não tem linha de visada, ou seja, adota-se a função de densidade de probabilidade de Rayleigh para representar a envoltória do sinal recebido na antena receptora. Esta suposição baseia-se em que os sinais, nos canais MIMO, encontram uma série de obstáculos que criam uma grande quantidade de componentes multipercurso. Dessa maneira, um número significativo de componentes de onda com amplitudes diferentes e ângulos de chegada uniformemente distribuídos devido ao espalhamento sofrido pelo sinal, chegam a cada um dos receptores. Portanto, o sinal no receptor terá uma envoltória com distribuição de Rayleigh.

O modelo do canal MIMO sem fio que apresenta a relação entrada-saída do sistema é representado pela seguinte equação:

$$\mathbf{r} = \mathbf{H} \cdot \mathbf{s} + \mathbf{n} \quad (2.12)$$

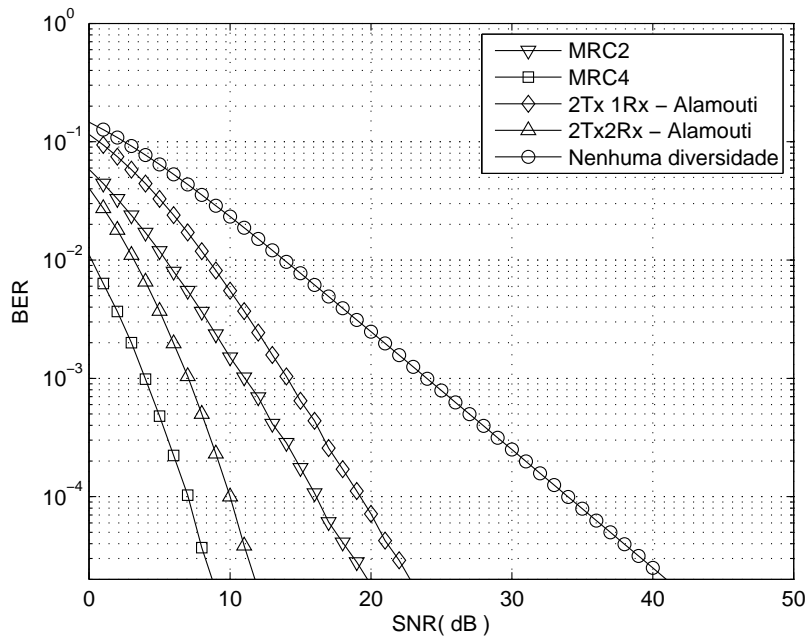


Figura 2.7: Simulações do Esquema Alamouti e MRC.

onde $\mathbf{s} = [s_1, s_2, \dots, s_M]^T$ é o vetor de símbolos transmitidos, $\mathbf{r} = [r_1, r_2, \dots, r_N]^T$ é o vetor de símbolos recebidos, $\mathbf{H}_{(NXM)}$ a matriz que representa o canal, $\mathbf{n} = [n_1, n_2, \dots, n_N]^T$ é o vetor complexo de AWGN (Additive White Gaussian Noise). Este ruído Gaussiano é um processo aleatório com valor médio igual a zero. Esta suposição é uma descrição razoável do tipo de ruído presente em muitos sistemas práticos de comunicação [Haykin and Moher, 2005].

A matriz do canal MIMO, \mathbf{H} , pode ser representada da seguinte maneira:

$$\mathbf{H} = \begin{bmatrix} h_{1,1} & h_{2,1} & \dots & h_{1,M} \\ h_{2,1} & h_{2,2} & \dots & h_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N,1} & h_{N,2} & \dots & h_{N,M} \end{bmatrix} \quad (2.13)$$

Cada um dos elementos h_{ij} representa o coeficiente de ganho de percurso do canal entre o transmissor j e o receptor i . Os elementos h_{ij} são identicamente distribuídos de uma variável aleatória complexa gaussiana e são considerados como sendo amostras independentes. A envoltória de h_{ij} segue uma distribuição Rayleigh. A figura 2.8 apresenta o modelo do canal MIMO.

Em [Foschini and Gans, 1998][Foschini, 1996] foi analisado a capacidade dos sistemas MIMO partindo da conhecida fórmula de capacidade de canal de Shannon:

$$C = B \log_2(1 + SNR) \quad \text{bits/s} \quad (2.14)$$

onde B é a largura de banda do canal e SNR é a relação sinal/ruído. Partindo desta fórmula e inserindo a influência das antenas transmissoras e receptoras do sistema MIMO, Foschini obteve

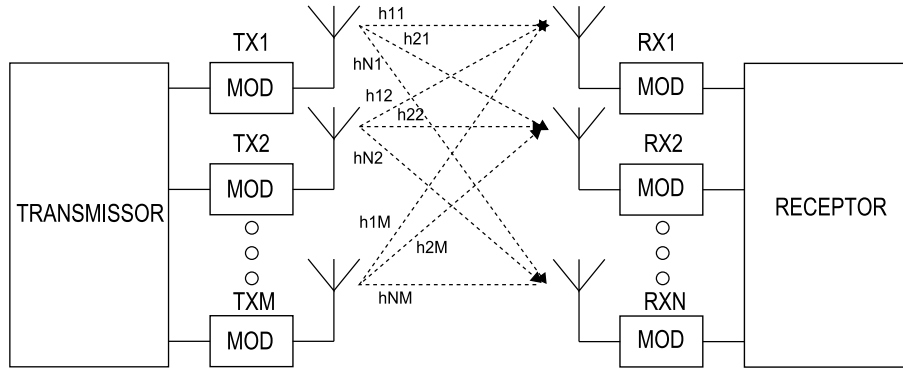


Figura 2.8: Modelo de Canal MIMO.

a seguinte expressão, em função das antenas transmissoras e receptoras, para a capacidade do canal:

$$C = E \left\{ \log_2 \left[\det \left(I_N + \frac{SNR}{M} H^H H \right) \right] \right\} \text{ bits/s/Hz} \quad (2.15)$$

onde E é a esperança sobre a matrix do canal H , \det é o determinante, I_N é a matriz identidade $N \times N$ e H^H significa a transposta conjugada.

Com base na equação 2.15, os sistemas MIMO apresentam três pontos importantes para serem considerados:

1) O fenômeno do desvanecimento é visto como uma fonte ambiental de possível enriquecimento e não como um problema.

2) A diversidade espacial tanto na recepção como na transmissão fornece as bases para um aumento na eficiência espectral ou capacidade do canal. Este aumento de capacidade é alcançado através da transmissão simultânea de símbolos, na mesma banda de frequência, a partir das antenas transmissoras. Num ambiente de espalhamento com desvanecimento Rayleigh independente, os sistemas MIMO podem alcançar taxas que crescem linearmente com o número de antenas transmissoras.

3) Diferentemente das técnicas convencionais, o aumento da capacidade nos sistemas MIMO é obtido aumentando-se a complexidade computacional e mantendo-se os recursos de comunicação primários constantes (largura de banda do canal e potência total de transmissão).

2.6 Tecnologia BLAST

Considerando um sistema MIMO de M antenas transmissoras e N antenas receptoras, cada uma das N antenas receptoras recebe todos os sinais transmitidos a partir de cada uma das M antenas transmissoras. A partir disto, o esquema de recepção deve fornecer um apropriado procedimento de decodificação para detectar os símbolos. A tecnologia BLAST é uma das tecnologias que realiza tal procedimento.

A primeira estrutura que surgiu, como procedimento de decodificação para detectar os símbolos em um sistema MIMO, foi a D-BLAST (Diagonal Bell Layered Space-Time) proposta por Foschini [Foschini, 1996]. A estrutura D-BLAST usava uma codificação para cada camada, ou seja, para cada antena. Para entender esta estrutura, considere, como exemplo, um sistema MIMO com três antenas transmissoras conforme a figura 2.9. Observando-se esta

figura percebe-se que em cada instante de tempo os símbolos são transmitidos repetidamente de maneira a formar uma diagonal espaço temporal.

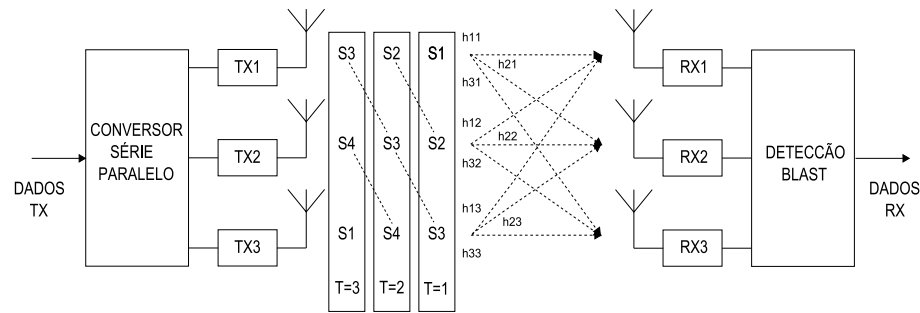


Figura 2.9: Arquitetura D-BLAST.

Entretanto, esta estrutura era muito complexa para implementação [Foschini, 1996][Shiu and Kahn, 1999]. Sendo assim, uma nova estrutura chamada V-BLAST (Vertical-BLAST) foi proposta [Wolniansky et al., 1998][Wolniansky et al., 1999].

2.6.1 Tecnologia V-BLAST

A estrutura V-BLAST é mais simples de ser implementada, no lado da transmissão, do que a D-BLAST, pois não necessita de codificação de bloco no vetor de entrada. O fluxo de dados de entrada é demultiplexado, através de um conversor série-paralelo, em M símbolos para cada período de símbolo, e cada grupo de símbolos é simultaneamente enviado, ao receptor, pelo grupo de M transmissores. Este grupo de símbolos transmitidos simultaneamente formam um vetor de M símbolos. Os transmissores operam na mesma faixa de frequência. Assume-se que todos os símbolos são modulados a partir da mesma constelação de modulação [Wolniansky et al., 1998]. A tecnologia V-BLAST é formada por M antenas transmissoras e N antenas receptoras onde $N > M$ dado que é necessário a inversão da matriz que representa o canal no processo de detecção. A figura 2.10 apresenta a arquitetura V-BLAST.

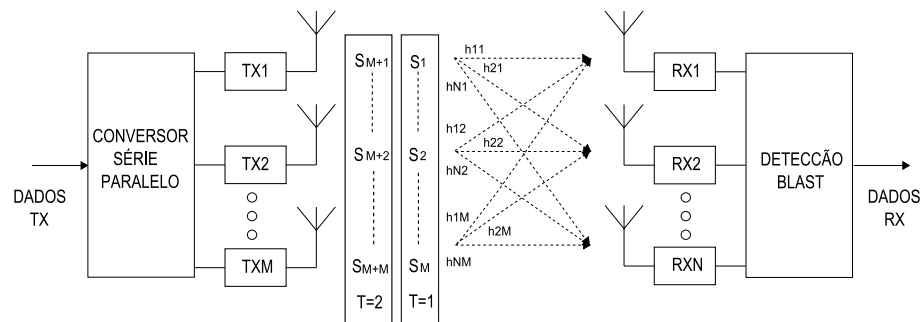


Figura 2.10: Arquitetura V-BLAST.

A equação que representa o sistema MIMO V-BLAST é descrita a seguir:

$$\begin{bmatrix} r_1 \\ \vdots \\ r_N \end{bmatrix} = \begin{bmatrix} h_{11} & \dots & h_{1M} \\ \vdots & \ddots & \vdots \\ h_{N1} & \dots & h_{NM} \end{bmatrix} \cdot \begin{bmatrix} s_1 \\ \vdots \\ s_M \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_N \end{bmatrix} \quad (2.16)$$

Para o caso do receptor RX1, o símbolo recebido r_1 será composto pelos seguintes termos:

$$r_1 = h_{11} \cdot s_1 + h_{12} \cdot s_2 + \dots + h_{1M} \cdot s_M + n_1 \quad (2.17)$$

Assume-se que o receptor conhece o canal através de seqüências de treinamento enviadas pelo transmissor. Cada um dos elementos do vetor $[r_1, r_2, \dots, r_N]^T$ representa o sinal recebido em cada uma das N antenas receptoras. Cada um dos sinais recebidos r_i com $1 \leq i \leq N$, é uma superposição dos sinais enviados pelas antenas transmissoras, onde para o caso da i -ésima antena receptora, cada um dos símbolos das M antenas transmissoras é multiplicado pelo seu coeficiente de ganho de percurso h_{ij} e somados em cada uma das antenas receptoras como se observa na equação 2.16. A cada um dos sinais nas antenas receptoras é somada também a componente do ruído AWGN n_i , conforme a equação 2.16.

O fato de que o sinal presente em cada uma das antenas receptoras tenha contribuição de todas as antenas transmissoras, faz com que exista a necessidade de um mecanismo de detecção que detecte cada símbolo transmitido. Existem dois métodos para realizar a detecção: detecção linear e detecção não-linear.

No método da detecção linear o sinal transmitido da antena i é tratado como o sinal desejado, enquanto que os sinais de outras antenas transmissoras são tratados como interferência. Processamento linear no vetor recebido é então usado para suprimir a interferência das outras antenas. Neste procedimento, também conhecida como anulamento (*nulling*), a detecção dos símbolos é feita pela ponderação linear satisfazendo algum critério de desempenho como o critério do mínimo erro quadrático médio MMSE (*Minimum Mean Square Error*) [Toledo and Wang, 2006] ou forçagem a zero ZF (*zero-forcing*) [Wolniansky et al., 1999]. O critério ZF, que tem como objetivo estimar um símbolo tratando os demais como interferentes, é realizado através de \mathbf{H}^+ que é a matriz pseudo-inversa [Pse, 1998] da matriz do canal \mathbf{H} . \mathbf{H}^+ é definida pela equação 2.18.

$$\mathbf{G} = \mathbf{H}^+ = \{\mathbf{H}^T \cdot \mathbf{H}\}^{-1} \cdot \mathbf{H}^T \quad (2.18)$$

O critério MMSE é realizado através da equação 2.19.

$$\mathbf{G} = (\mathbf{H}^H \cdot \mathbf{H} + \sigma^2 \cdot I_{(M,M)})^+ \cdot \mathbf{H}^H \quad (2.19)$$

O vetor de saída \mathbf{S} do detector MIMO-ZF ou MIMO-MMSE é obtido pela equação 2.20.

$$\mathbf{S} = \mathbf{G} \cdot \mathbf{r} \quad (2.20)$$

O detetor VBLAST ótimo é o detetor de máxima verossimilhança que faz a detecção dos símbolos recebidos com a mínima probabilidade de erro. No entanto, a complexidade computacional deste detetor cresce exponencialmente com o número de antenas transmissoras. Por este motivo, na prática os dois detetores de baixa complexidade, ZF e MMSE, são comumente implementados no receptor. Comparando-se as equações 2.18 e 2.19 conclui-se que algoritmo ZF possui uma implementação mais simples que o MMSE. Porém, a custo de um menor desempenho. Ambos os detetores são assintoticamente similares em altas relações sinal-ruído [Shiu and Kahn, 1999].

Assim como é possível estimar todos os símbolos do vetor \mathbf{S} simultaneamente através da equação 2.20, também é possível realizar o procedimento detectando símbolo a símbolo.

Para que isto aconteça, escolhe-se a linha da matriz \mathbf{G} correspondente ao símbolo do vetor \mathbf{S} que se deseja estimar. Por exemplo, se fosse desejado estimar somente o símbolo \mathbf{S}_4 , escolher-se-ia a quarta linha da matriz \mathbf{G} , \mathbf{G}_4 , e se realizaria a operação da equação 2.21 para este símbolo e esta linha.

$$\mathbf{S}_4 = \mathbf{G}_4 \cdot \mathbf{r} \quad (2.21)$$

Quando os detectores não lineares utilizam a técnica de cancelamento sucessivo de interferência SIC (*Successive Interference Cancelling*), a contribuição dos símbolos já detectados no sinal recebido é sucessivamente removida em etapas. Tomando-se decisões exatas em relação ao símbolo transmitido, o sinal ficará livre de interferência dos símbolos anteriormente detectados. No algoritmo SIC, inicialmente o sinal recebido é detectado através de processamento linear, através de ZF ou MMSE; isto gera um primeiro símbolo estimado \hat{s}_j . A contribuição deste símbolo é subtraída nos sinais recebidos, calculando-se novos sinais de saída livres da contribuição do símbolo \hat{s}_j . A recursão continua até ser encontrado o último símbolo. O algoritmo SIC pode ser crítico se um símbolo é detectado erroneamente, pois sua arquitetura sequencial que depende dos símbolos encontrados previamente, faz que a partir dessa detecção errônea se gerem erros para frente. Este problema pode ser atenuado com uma ordem determinada no processo de detecção. Esta ordem pode ser determinada, tomando-se o primeiro símbolo a ser detectado como aquele com maior SNR; isto é feito porque o símbolo com maior energia no receptor terá menos probabilidade de ser detectado erroneamente, e gerará menos erros nas seguintes detecções, já que elas dependem dos símbolos previamente detectados. Neste caso, o algoritmo SIC é chamado de OSIC (*Ordered Successive Interference Cancelling*).

Algoritmo de Detecção de Símbolos na Tecnologia V-BLAST

A seguir é analisado o algoritmo detalhadamente. Assumindo que os componentes s_j do vetor \mathbf{S} são extraídos na ordem incremental desde s_1 a s_M , o algoritmo ZF-SIC ou MMSE-SIC, da tecnologia V-BLAST, pode ser escrito como um procedimento recursivo como segue [Wolniansky et al., 1998]:

i) Visando aplicar a detecção linear através do critério ZF ou MMSE, calcula-se a matriz \mathbf{G} do canal.

$$\mathbf{G} = \mathbf{H}^+ \quad \text{ZF} \quad (2.22)$$

$$\mathbf{G} = (\mathbf{H}^H \mathbf{H} + \sigma^2 \mathbf{I}_{(M,M)})^+ \mathbf{H}^H \quad \text{MMSE} \quad (2.23)$$

Conforme explicado anteriormente, a matriz \mathbf{G} fornece agora a possibilidade de detectar cada um dos símbolos no vetor recebido \mathbf{r} , o qual é conhecido.

ii) Determine a linha da matriz \mathbf{G} correspondente ao j -ésimo símbolo a ser detectado. Este símbolo no algoritmo OSIC será aquele com maior energia no receptor, para obter uma maior probabilidade de detecção correta. Para detectar o símbolo com maior energia no receptor, procura-se a linha da matriz \mathbf{G} que tenha a menor norma; esta linha com a menor norma significa a coluna da matriz \mathbf{H} com a maior norma, ou seja, a coluna tendo os coeficientes de ganho de percurso com os valores mais altos; cada coluna da matriz \mathbf{H} possui os ganhos de percurso que multiplicam o símbolo enviado por cada antena em particular; por exemplo, a coluna u da matriz \mathbf{H} são os ganhos de percurso associados aos símbolos enviados pelo primeiro transmissor a cada um dos receptores. Assim, a coluna com a maior norma em \mathbf{H} gerará o maior

valor de energia no receptor, multiplicando o símbolo correspondente à coluna pelos ganhos de percurso. Então esta linha da menor norma da matriz \mathbf{G} será denotada pelo vetor \mathbf{w} .

$$\mathbf{w} = \mathbf{G}_j \quad (2.24)$$

iii) Detecte o j -ésimo símbolo, o qual será dado por y_j .

$$y_i = \mathbf{w} \cdot \mathbf{r} \quad (2.25)$$

iv) Estime o símbolo detectado.

$$\hat{a}_j = Q(y_i) \quad (2.26)$$

v) Começa o processo de cancelamento de interferência. O efeito do símbolo \hat{a}_j é então subtraído do sinal recebido.

$$\mathbf{r} = \mathbf{r} - \mathbf{H}_j \hat{a}_j \quad (2.27)$$

Onde \mathbf{H}_j é a j -ésima coluna de \mathbf{H} .

vi) Já que o símbolo \hat{a}_j foi estimado é possível anular a j -ésima coluna de \mathbf{H} correspondente a este símbolo. Sendo assim,

$$\mathbf{H}^- = \mathbf{H} - [0 \ 0 \ \dots \ \mathbf{H}_j \ \dots \ 0 \ 0] \quad (2.28)$$

vii) Repita o processo desde o passo (i) até que todos os fluxos sejam detectados.

2.7 Conclusão

A transmissão de sinais sobre o canal de rádio apresenta maior dificuldade quando comparado com o canal cabeado. Existem técnicas de diversidade que visam minorar os problemas da transmissão no canal sem fio. O canal de rádio, em função de suas características, pode ser modelado, sob o ponto de vista da frequência e do tempo, e classificado.

Este capítulo mostrou, também, que com os sistemas MIMO pode-se obter um sistema de transmissão sem fio mais confiável ou com altas taxas de transmissão sem o aumento da largura de banda e/ou da potência do sinal. Quando se deseja alta capacidade de transmissão, em sistemas de comunicações sem fio, o sistema a se utilizar é o MIMO-VBLAST. O MIMO-VBLAST tem dois algoritmos de detecção mais utilizados: VBLAST-ZF e VBLAST-MMSE. Para o caso em que se é necessário maior confiabilidade de transmissão uma das opções a se utilizar é o sistema MIMO-STBC. Devido as características de alta capacidade de transmissão e confiabilidade os sistemas MIMO estão sendo utilizados nos sistemas de comunicações sem fio. Portanto, é comum que se tenha o protocolo TCP, o mais utilizado da camada de transporte, sendo implementado sobre os sistemas MIMO.

Capítulo 3

O Protocolo TCP

É um excelente protocolo, pertencente à camada de transporte no modelo de referência OSI, pois foi projetado para se adaptar aos mais variados tipos de redes, com diferentes taxas de erros e capacidade de transmissão. Além disso, o TCP pode ser implementado em diversas tecnologias, tais como conexões por linha discada, rede local com par trançado, redes de fibra ótica e redes *wireless*. Atualmente encontra-se várias versões do protocolo TCP (*Transmission Control Protocol*), muitas delas com peculiaridades particulares para trabalhar sob uma tecnologia de rede específica, como por exemplo em redes *wireless*. Por destacar-se em termos de desempenho e adaptação nos mais variados tipos de redes, a versão bastante usada no mundo atualmente é a TCP Reno. Atualmente, é o protocolo mais utilizado pela internet.

O protocolo TCP [Tanenbaum, 2003] é um protocolo orientado a conexão, pois antes que processos de aplicação comecem a se comunicar eles devem trocar segmentos preliminares para estabelecer os parâmetros da transferência de dados em questão. Um serviço orientado a conexão possui três fases:

- Estabelecimento de conexão: é definido apenas um caminho entre a origem e o destino.
- Transferência de dados: com o objetivo de que os dados cheguem ao destino na ordem em que foram enviados, os mesmos são transmitidos em sequência pelo caminho estabelecido.
- Término de conexão: terminar a conexão entre destino e origem quando a conexão não está mais em uso.

O protocolo TCP injeta dados na rede através de segmentos. Estes segmentos são conjuntos de octetos recebidos dos programas aplicativos. Conforme descrito por [Tanenbaum, 2003], um segmento TCP é composto por um cabeçalho fixo de 20 bytes seguido de zero ou mais bytes de dados. Quando uma conexão é estabelecida, o TCP define o tamanho do segmento através de uma negociação entre a unidade transmissora e a unidade receptora. Para ser eficiente, este tamanho não deve ser pequeno para não haver uma má utilização da rede, uma vez que o cabeçalho é grande, e nem grande para que não ocorra uma alta taxa de perda de segmentos.

3.1 Controle de Transmissão Confiável

Para garantir que dados perdidos, durante a transmissão, sejam recuperados, o TCP faz uma troca de números de sequência durante a sequência da conexão. Um hospedeiro começa uma conexão enviando um pacote indicando seu número de sequência inicial com um determinado bit no cabeçalho definido para indicar um pedido de conexão. Na sequência, o outro hospedeiro recebendo o pacote memoriza o número de sequência A e anexa seu próprio número de sequencial inicial B respondendo com uma confirmação A + 1. A confirmação recebida A + 1 indica que o hospedeiro recebeu todos os octetos até A e que está aguardando A + 1 na sequência.

O TCP é responsável por adaptar a taxa de transmissão às condições de banda disponíveis no enlace de comunicação, evitando o congestionamento da rede e criando uma conexão confiável através da retransmissão de pacotes perdidos. A confiabilidade é obtida através do envio de segmentos de confirmação ACK (*Acknowledgment*) transmitidos pelo receptor em resposta aos segmentos de dados já recebidos. Os ACKs indicam ao transmissor a numeração do próximo segmento esperado no receptor. A figura 3.1 ilustra esta idéia.

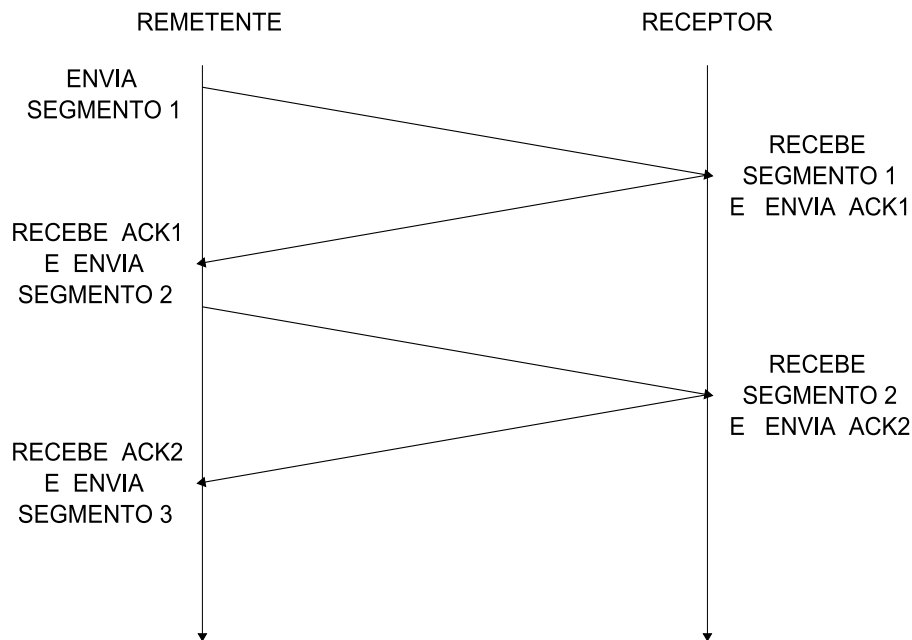


Figura 3.1: Transmissão com confirmação positiva.

Um segmento é considerado perdido quando ocorre uma das duas situações:

- Três ACKs repetidos chegam ao transmissor. Suponha que um transmissor envie os pacotes 1, 2, 3 e 4 e o pacote 1 seja perdido no caminho. Nesta condição os três ACK's que o receptor enviará ao transmissor, como resposta ao envio dos pacotes 2, 3 e 4, será o ACK 1. Sendo assim o transmissor enviará novamente o segmento 1. A figura 3.2 ilustra esta idéia.
- Quando o transmissor envia um segmento ao receptor ele inicia uma contagem de tempo. Caso o ACK correspondente aquele segmento enviado não chegue dentro de um período

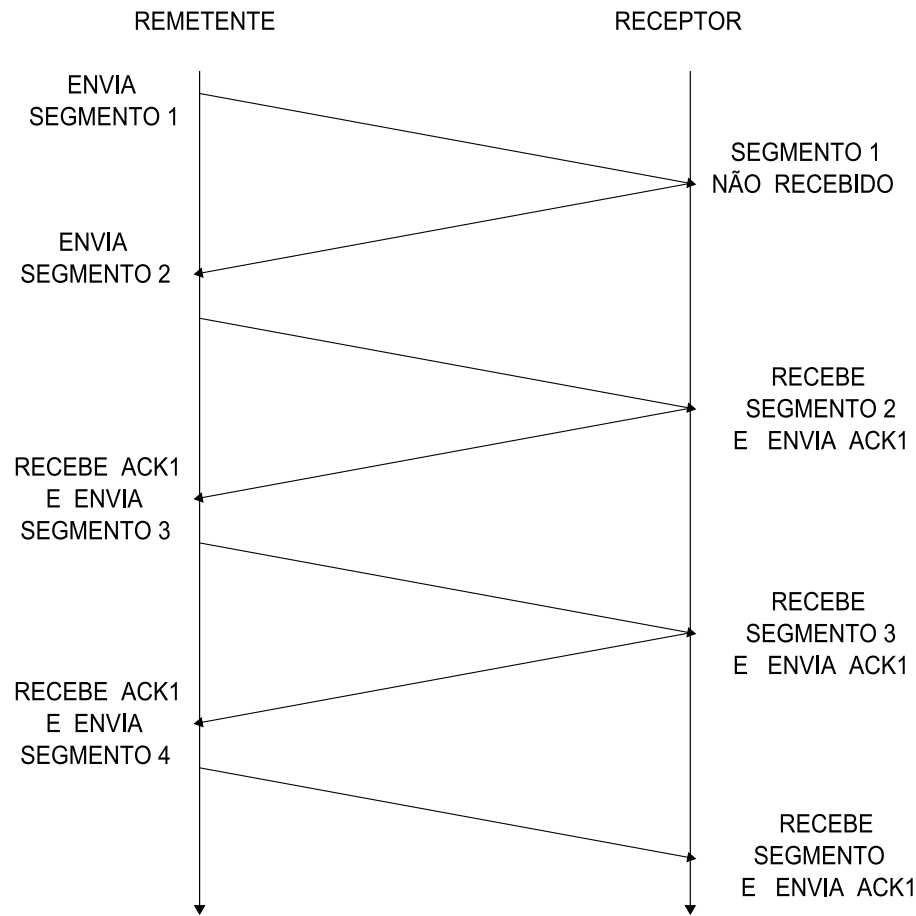


Figura 3.2: Transmissão com ACK duplicado.

de tempo específico (*timeout*), o TCP considera o segmento perdido e retransmite-o novamente [Tanenbaum, 2003]. A figura 3.3 ilustra esta idéia.

3.2 Controle de Fluxo

Para controlar o fluxo de dados o TCP utiliza um sistema de janela deslizante. No momento em que uma conexão é estabelecida, tanto o transmissor como o receptor disponibiliza um buffer para armazenamento de dados que chegam, e os dois enviam, um ao outro, o tamanho deste buffer. Ao passo que os segmentos chegam ao receptor o mesmo envia ACK's que contém informação a respeito do tamanho do buffer restante. A quantidade de espaço disponível nos buffers é chamada de janela. Toda vez que um receptor envia um ACK, a informação da janela está contida neste ACK. Se o aplicativo no lado do receptor consegue ler dados na mesma velocidade com que eles chegam, então o receptor enviará ACK's contendo a informação de janela positiva. No entanto, se o lado transmissor é mais rápido que o receptor, então os dados que estão chegando causarão o enchimento do buffer, forçando que o receptor envie um ACK com a informação de janela zero. Ao receber a informação de janela zero, o transmissor interrompe o envio de segmentos até que o receptor envie novamente o anúncio de janela positiva [Comer, 2007]. A figura 3.4 descreve este processo.

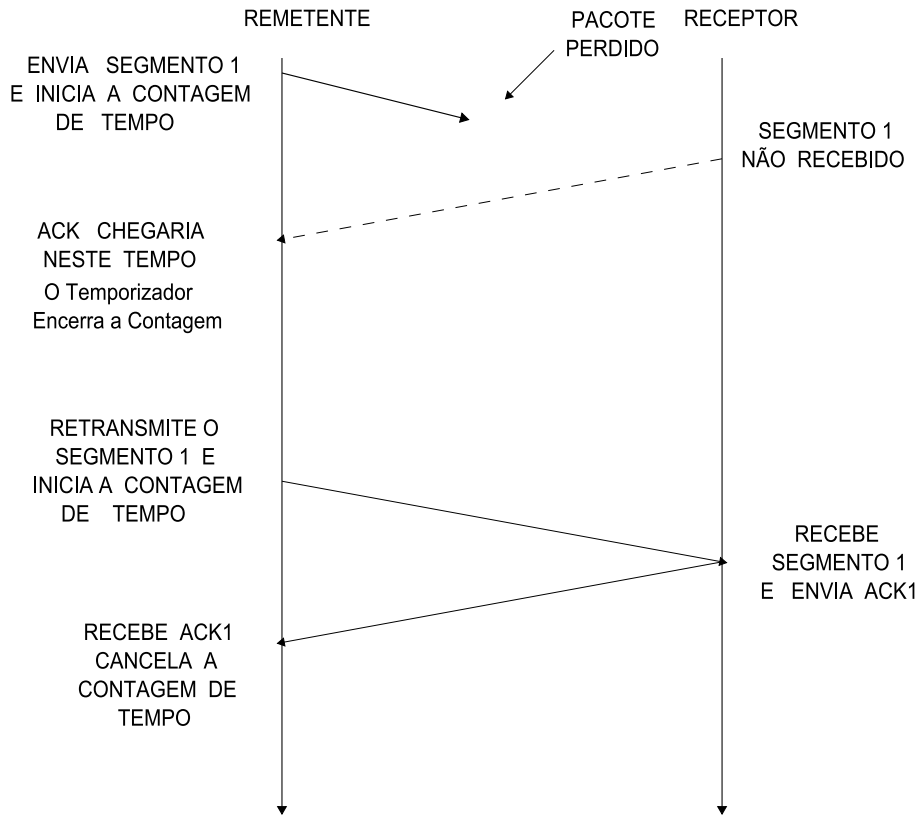


Figura 3.3: Perda de segmento por *timeout*.

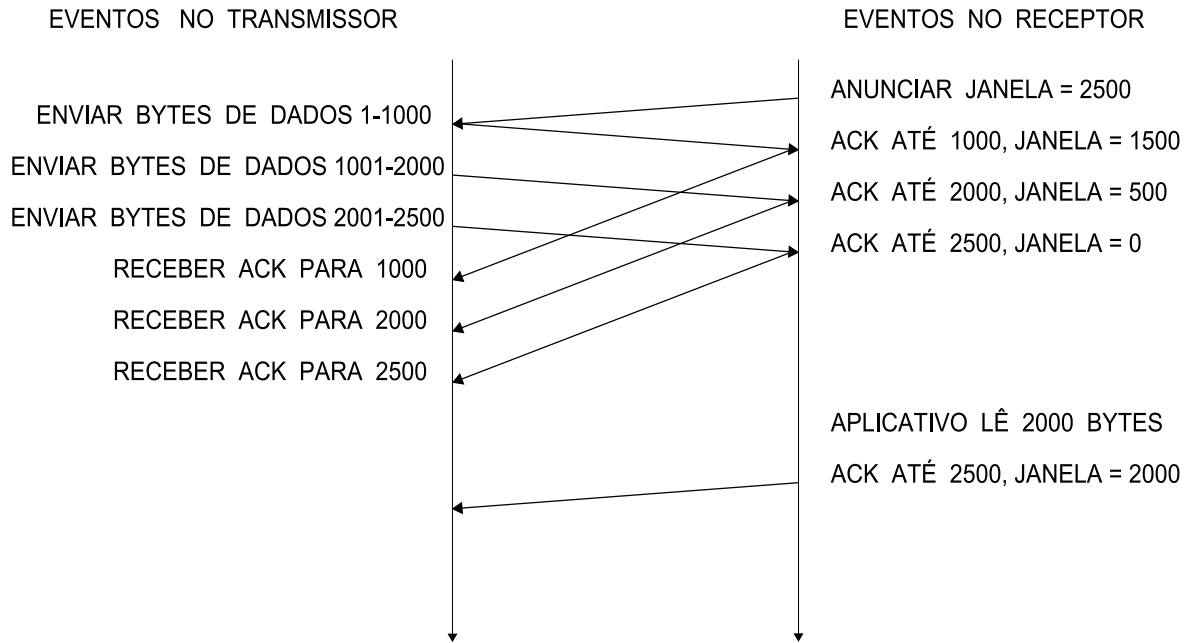


Figura 3.4: Controle de fluxo do TCP.

3.3 Controle de Congestionamento

De maneira a se controlar a taxa de transmissão e evitar o congestionamento da rede, o número de pacotes que podem ser transmitidos sem o recebimento do ACK é limitado por um

parâmetro denominado janela de congestionamento CW (*Congestion Window*). O tamanho da CW é controlado dinamicamente de acordo com o estado da rede [Allman et al., 2002]. Para uma nova conexão, ou quando uma retransmissão é gerada por *timeout*, a CW é ajustada para a fase de inicialização lenta (*slow start*). A cada vez que um ACK é recebido, a CW é incrementada de um segmento. Durante a fase *slow start* a CW aumenta até que um determinado valor limite é atingido ou até que ocorra uma retransmissão por *timeout*. Quando isso acontece, a fase *slow start* termina e uma segunda fase denominada *congestion avoidance* inicia. Nesta fase a CW tem um crescimento linear. Um mecanismo adicional denominado *fast retransmit* e *fast recovery* também é implementado. Neste caso, um segmento perdido é detectado e recuperado antes que ocorra a retransmissão por *timeout*. Portanto a taxa de transmissão não é reduzida tanto quanto ocorreria quando a retransmissão é gerada por *timeout*, porque existe uma detecção e retransmissão antecipada do pacote perdido, aumentando o desempenho do TCP. Existem diversas variantes do protocolo TCP. Na versão Tahoe, sempre que uma perda é detectada, o tamanho da janela de congestionamento reduz para um e a fase *slow start* inicia novamente. Em outras versões bastante utilizadas como o TCP Reno e New Reno, o tamanho da CW se reduz para um apenas se uma perda for detectada por *timeout*. Quando uma perda é detectada pela ocorrência de ACKs repetidos, o tamanho da CW se reduz pela metade. Assim a fase de *slow start* não é iniciada novamente e o protocolo se mantém na fase *congestion avoidance*.

3.4 Desempenho do TCP em Enlaces sem Fio

O protocolo TCP foi originalmente projetado para operação em redes cabeadas. O *backbone* da rede Internet é composto por enlaces de fibra ótica de alta velocidade que conectam os roteadores da rede. Neste nível, a transmissão dos dados pode ser considerada como livre de erros. As perdas de pacote são usualmente geradas por estouro de memória nas filas de roteamento. Uma parte menos confiável da rede é a rede de acesso. A rede de acesso cabeada pode ser implementada usando-se fibra ótica, cabo coaxial ou par trançado. Nestes cenários os mecanismos de controle do TCP descritos na anteriormente são considerados bastante efetivos para garantir um bom desempenho em termos de taxa de transmissão, mesmo sob congestionamentos moderados da rede. O canal definido pelo par trançado não pode ser assumido como livre de erros e apresenta na prática uma taxa de erro de bit considerável induzida pelo ruído, interferência e distorção. Contudo, este canal ainda é bastante previsível se comparado com canais sem fio.

Em contraste, a comunicação sem fio é afetada por problemas adicionais, como o desvanecimento, o que pode resultar em altas taxas de erro de bit. O desempenho do TCP em redes sem fio pode ser severamente reduzido, pois perdas adicionais de pacotes são induzidas pelo canal de rádio. O TCP interpreta estas perdas como sendo causadas por congestionamento na rede. Neste caso o TCP executa os mecanismos de controle de congestionamento, reduzindo o taxa de transmissão útil (*throughput*) da conexão e aumentando o atraso fim-a-fim para transferência dos dados. Em alguns casos, a redução do *throughput* é desnecessária porque o canal de rádio apresenta desvanecimentos intensos, que geram altas taxas de erro de bit e muitas perdas de pacotes, mas que duram pouco tempo. Nesta situação todos os pacotes são perdidos e o protocolo TCP reduz o número de pacotes injetados na rede. Basicamente, existem três estratégias para minimizar a redução do *throughput* do TCP em enlaces de rádio [Toledo and Wang, 2006].

3.5 Estratégias para Minimizar a Redução do *Throughput* do TCP em Enlaces sem Fio

3.5.1 Mudanças na Implementação do TCP

Um exemplo de mudança na implementação do TCP é o TCP Santa Cruz [C.Parsa and Garcia-Luna-Aceves, 1999]. O algoritmo de controle de congestionamento Santa Cruz detecta não somente o estágio inicial de congestionamento, que possibilita um aumento ou uma redução na janela de congestionamento como resposta a sinais de aviso antecipados, mas também identifica se o congestionamento está se desenvolvendo no caminho do transmissor para o receptor ou do receptor para o transmissor. A determinação da direção do congestionamento permite uma ação de isolamento, do *throughput* dos dados, de um eventual congestionamento no caminho do receptor para o transmissor. A medida de um atraso relativo que um pacote experimenta com relação a outro, quando eles propagam através da rede, é a base do algoritmo de controle de congestionamento do TCP Santa Cruz.

Para determinar este atraso relativo, o transmissor do TCP Santa Cruz mantém uma tabela com dois tempos para cada pacote: o tempo de transmissão do pacote de dados na fonte e o tempo de chegada do pacote de dados no receptor informado pelo ACK. A partir destas informações, o transmissor calcula os seguintes intervalos de tempo para quaisquer dois pacotes i e j (onde $j > i$): $S_{j,i}$ que representa o intervalo de tempo entre a transmissão dos pacotes e $R_{j,i}$ que representa o intervalo de tempo de chegada dos pacotes no receptor. A equação 3.1 determina o atraso relativo do caminho dos dados.

$$D_{j,i}^F = R_{j,i} - S_{j,i} \quad (3.1)$$

Onde $D_{j,i}^F$ representa a mudança no atraso do caminho dos dados experimentado pelo pacote j com relação ao pacote i . A figura 3.5 ilustra a medida do atraso.

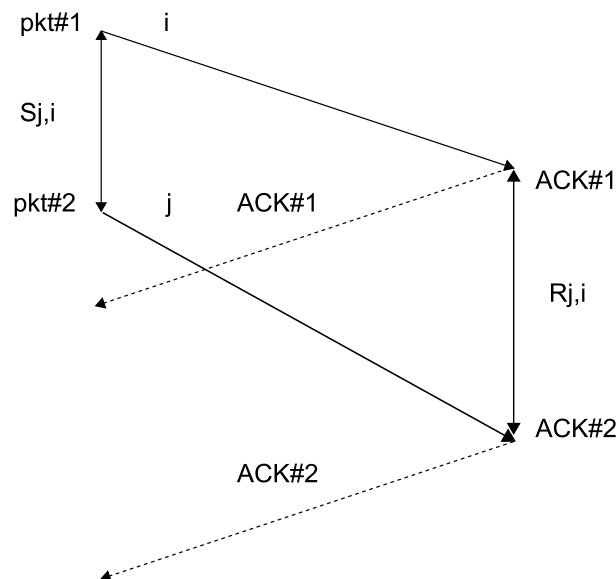


Figura 3.5: Medida do atraso relativo.

Observando-se a equação 3.1 conclui-se que:

1. Quando $D_{j,i}^F=0$ os dois pacotes possuem o mesmo atraso.
2. Quando $D_{j,i}^F<0$ o primeiro pacote esta mais atraso do que o segundo.
3. Quando $D_{j,i}^F>0$ o segundo pacote esta mais atraso do que o primeiro.

Em qualquer tempo durante uma conexão, as filas na rede (e especificamente a fila no gargalo da rede) estam em um dos três estados:

- Aumentando em tamanho.
- Diminuindo em tamanho.
- Mantendo-se constante.

A obtenção do atraso relativo do caminho de dados, $D_{j,i}^F$, permite a determinação da mudança no estado das filas. O diagrama de estado da figura 3.6 mostra como isto acontece. O objetivo

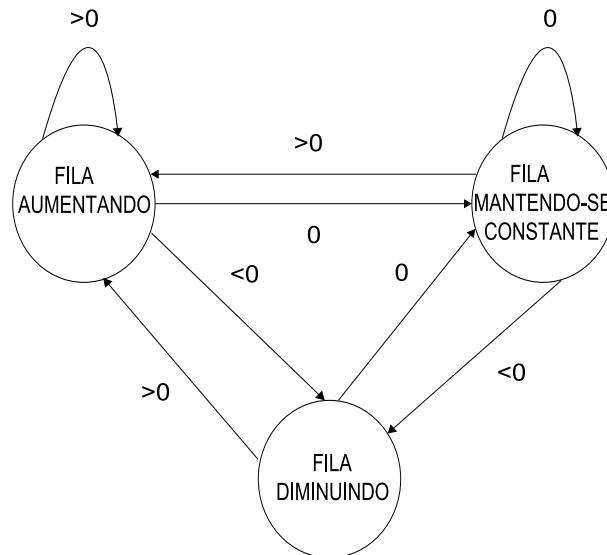


Figura 3.6: Determinação do estado das filas de acordo com $D_{j,i}^F$.

do TCP Santa Cruz é permitir que as filas da rede (especificamente a fila do gargalo) cresça para um desejado tamanho. O algoritmo abaixo descrito realiza tal tarefa.

O valor do atraso relativo quando é positivo significa um aumento nas filas da rede e quando é negativo significa uma redução nas filas da rede. Somando-se as medidas do atraso relativo sobre um periodo de tempo tem-se uma indicação do nível de enfileiramento no gargalo da rede. Sendo assim, o algoritmo de controle de congestionamento do TCP Santa Cruz opera somando os atrasos relativos do começo de uma sessão e atualiza estas medidas em intervalos discretos. Cada intervalo corresponde a uma quantidade de tempo para transmitir uma janela completa de dados e receber seus correspondentes ACKs. Entretanto, a unidade de $D_{j,i}^F$ é em tempo, portanto este valor é convertido para um número equivalente de pacotes que represente este atraso. O algoritmo tenta manter a condição estabelecida pela equação 3.2.

$$n_{t_i} = N_{op} = n_{t_{i-1}} + M_{w_{i-1}} \quad (3.2)$$

Onde n_{t_i} é o número total de pacotes enfileirados no gargalo da rede no tempo t_i ; N_{op} é o ponto de operação (o desejado número de pacotes, por sessão, para ser enfileirado no gargalo da rede); $M_{w_{i-1}}$ é a quantidade adicional de fila introduzida pela janela anterior W_{i-1} . A CW do TCP Santa Cruz é ajustada de maneira que a equação 3.2 é satisfeita dentro de uma faixa de $N_{op} \pm \delta$, onde δ é algumas frações de um pacote. Os ajustes realizados na CW ocorrem somente em intervalos discretos, como por exemplo, o tempo gasto para esvaziar uma janela completa de dados. Sobre este intervalo, $M_{w_{i-1}}$ é calculada e no fim do intervalo é somado a $n_{t_{i-1}}$. Se o resultado estiver dentro da faixa $N_{op} \pm \delta$, a CW é mantida no seu tamanho corrente. Se, entretanto, $n_{t_{i-1}}$ resulta em um valor abaixo de $N_{op} - \delta$, então o sistema não está sendo utilizado o tanto que poderia ser e por isso a CW é aumentada linearmente durante o próximo intervalo de tempo. Se $n_{t_{i-1}}$ resulta em um valor acima de $N_{op} + \delta$, então o sistema acima do ponto de operação desejado e por isso a CW é decrementada linearmente durante o próximo intervalo.

O TCP Santa Cruz também é capaz de fornecer uma melhor estimativa do RTT (*Round Trip Time*) quando comparado com a estimativa do TCP tradicional. Esta melhor estimativa é conseguida através da medição do RTT de cada segmento transmitido em que um ACK é recebido, incluindo as retransmissões. Para alcançar tal desempenho, o TCP Santa Cruz precisa que cada ACK que chegue ao transmissor indique com precisão o pacote que o gerou e o transmissor mantenha um registro de tempo para cada transmitido ou retransmitido pacote. A identificação é feita da seguinte maneira: a primeira transmissão do pacote 1 é especificada como 1.1, a segunda transmissão é especificada por 1.2, e assim por diante. Os ACKs dessas transmissões são identificados da seguinte maneira: por exemplo, um ACK (5.7.2) especifica um ACK acumulado de 5, quinto ACK enviado, que foi gerado pela segunda transmissão de um pacote com número de sequência 7. A atualização do RTT usa o mesmo algoritmo utilizado pelo TCP Tahoe e Reno; entretanto, a computação é realizada para cada ACK recebido, em vez de um por RTT.

Para ajudar na identificação e recuperação de pacotes perdidos, o receptor do TCP Santa Cruz retorna uma janela de ACK para o transmissor com o objetivo de indicar qualquer ausência de pacotes no fluxo sequencial recebido. No caso de múltiplas perdas por janela, a janela de ACK permite o TCP Santa Cruz retransmitir todos os pacotes perdidos sem esperar pelo *timeout* do TCP.

3.5.2 Dividir a Conexão do TCP

No algoritmo de controle de congestionamento do TCP, quase todas as implementações do TCP atuais interpretam que quando o *timer* do TCP expira é devido ao congestionamento na rede e não a pacotes perdidos. Nessas condições, o TCP inicia a fase *slow start* com o objetivo de reduzir a carga da rede e assim diminuir o congestionamento. Entretanto, o canal de rádio possui uma taxa de erro de bits elevada e, por isso, existem perdas de pacotes a todo instante. A melhor maneira de lidar com essas perdas seria enviá-los novamente o mais rápido possível [Tanenbaum, 2003]. Nesta situação, se for diminuída a taxa de transmissão de pacotes o problema ficará ainda pior, pois o *throughput* efetivo do TCP diminuirá e, no entanto, continuará acontecendo a perda de pacotes. Por exemplo, suponha-se que o transmissor envie ao receptor 200 pacotes/segundo e que 10% seja perdido devido a taxa de erro de bit do canal. Então, o *throughput* nesse caso é de 180 pacotes/segundo. Entretanto, se o transmissor diminuir sua taxa de transmissão para 100 pacotes/segundo o *throughput* diminuirá para 90 pacotes/segundo. Em resumo, em uma rede cabeada, quando ocorrer a perda de um pacote, o transmissor

deve diminuir sua taxa de transmissão. Mas quando essa perda de pacote acontece em uma rede sem fio o transmissor deve aumentar ainda mais sua taxa de transmissão. O problema é que se o transmissor não souber em que tipo de rede ele está injetando pacotes, fica difícil tomar a decisão correta.

Frequentemente, o caminho entre o transmissor e o receptor é heterogêneo, ou seja, os pacotes trafegam por canais cabeadas e por canais sem fio até chegarem ao seu destino. Nessas circunstâncias, fica ainda mais difícil saber porque motivo o *timeout* aconteceu, pois é necessário saber onde ocorreu o problema. A solução proposta para este problema foi denominada por TCP indireto e ela consiste em dividir a conexão TCP em duas conexões separadas, como mostra a figura 3.7[Tanenbaum, 2003]. A primeira conexão inicia no transmissor e termina na estação-base. A segunda inicia-se na estação-base e termina no receptor. A função da estação-base é copiar pacotes entre as conexões em ambos sentidos. Nessas condições, sabe-se

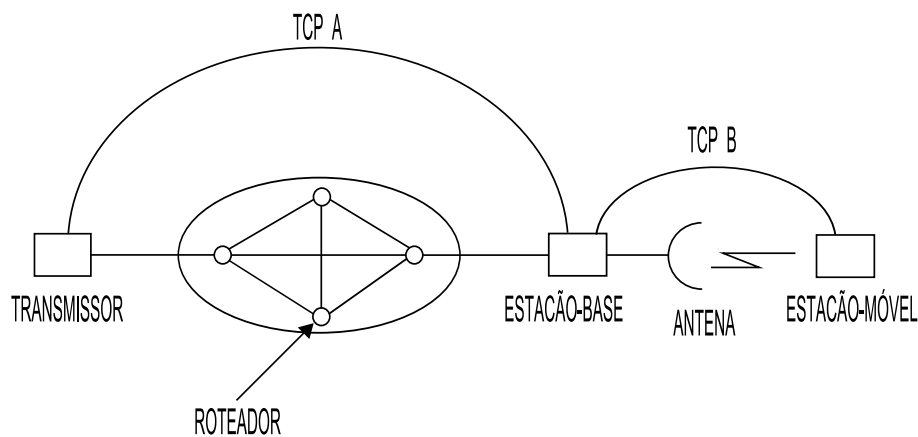


Figura 3.7: Divisão de uma conexão TCP em duas conexões.

exatamente o motivo pelo qual o *timeout* foi gerado e, portanto, o transmissor poderá diminuir sua taxa de transmissão, quando a rede for cabeada, ou aumentar sua taxa de transmissão quando a rede for sem fio. Entretanto, essa proposta não se adapta a semântica do TCP. A estação-base confirma cada segmento do TCP recebido, pois cada divisão da conexão é uma conexão completa. Sendo assim, quando o transmissor receber um ACK, isso não significa que o segmento chegou no receptor, mas sim à estação-base.

Para solucionar esse problema, uma outra proposta é sugerida sem que a semântica do TCP seja rompida[Tanenbaum, 2003]. Nesta proposta são feitas várias modificações pequenas no código da camada de rede da estação-base. Uma dessas mudanças é o acréscimo de um espião que monitora e armazena em um *cache* os segmentos do TCP que saem para a estação-móvel e os ACK's que são emitidos por essa estação. Quando o espião percebe que um segmento TCP partiu para a estação-móvel e não houve um ACK para ele antes da ocorrência do *timeout*, ele simplesmente retransmite esse segmento não informando sua ação à origem. Uma retransmissão também ocorre quando o espião detecta confirmações duplicadas vindas da estação-móvel. Essas confirmações duplicadas, que informam que a estação-móvel perdeu algo, são imediatamente eliminadas para impedir que a origem interprete, erroneamente, como um sinal de congestionamento. No entanto, uma desvantagem desta proposta é que se o canal sem fio provocar muitas perdas de pacotes, isto poderá gerar um *timeout* no transmissor devido a espera por uma ACK e, na sequência, o algoritmo de controle de congestionamento é invocado. Essa mesma proposta contempla, também, uma solução para o problema dos segmentos

pedidos originados na estação-móvel. Ao detectar um intervalo nos números de sequência de entrada, a estação-base requisita uma retransmissão seletiva dos bytes perdidos utilizando uma opção do TCP.

3.5.3 Retransmissão Local

A Retransmissão Local, ARQ (*Automatic Repeat Request*), é a maneira mais comum de esconder perdas, devido ao canal sem fio, do TCP. O mecanismo de retransmissão local acontece em baixo do nível IP no enlace sem fio. A figura 3.8 apresenta este mecanismo.

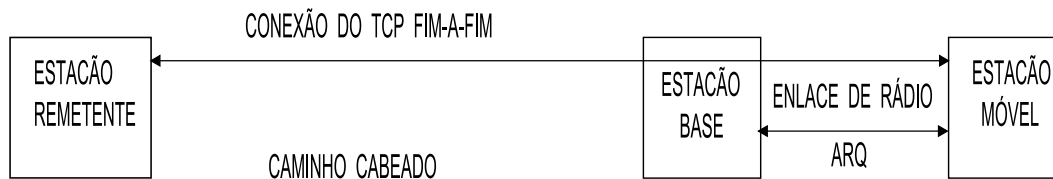


Figura 3.8: Arquitetura de retransmissão local.

O RLP (*Radio Link Protocol*), que utiliza o sistema ARQ, é o protocolo mais comum para melhorar o desempenho do TCP sobre canais sem fio. Este protocolo é transparente para as camadas mais superiores e tem como característica a redução da FER do canal. Muitos padrões de comunicação sem fio de nova geração, como por exemplo o UMTS [RLC, 1999], implementam o RLP. Seu objetivo é evitar que o TCP reinicie na fase de *slow start*. Retransmissões na camada de enlace são possíveis porque o tempo de *timeout* do TCP é um valor suficientemente grande para permitir a retransmissão de alguns quadros perdidos.

O RLP realiza fragmentação e remontagem de pacotes na camada de enlace e utiliza filas de armazenagem no transmissor e receptor. Existe uma temporização para *timeout* para cada quadro perdido no receptor. Este temporizador é usualmente configurado para o tempo de transmissão de quatro quadros RLP. Quando ocorre o *timeout* para um determinado quadro, ou um quadro RLP não consecutivo é recebido, uma perda é identificada. Neste caso, o receptor envia um NACK (*Negative Acknowledgment*) para o transmissor, indicando que um quadro está faltando, que então procede a retransmissão. Este processo é executado para um mesmo quadro até um máximo de n ocorrências de *timeout* (usualmente entre 3 e 10 vezes) ou até que o quadro correto seja recebido. Após as n tentativas, se o quadro RLP ainda estiver faltando, a camada de enlace de rádio descarta todos os quadros precedentes, referentes ao mesmo segmento TCP. Assim, o protocolo da camada de transporte, TCP, inicia o seu próprio processo de recuperação.

3.6 Conclusão

Este capítulo apresentou o protocolo TCP demonstrando como ele funciona de maneira a apresentar uma transmissão confiável. Também foi abordado os mecanismos de controle de fluxo e o de controle de congestionamento. Uma análise do TCP sobre canal sem fio foi realizada demonstrando o motivo pelo qual ele apresenta um baixo desempenho e, também, as soluções existentes e propostas para melhorar o seu desempenho sobre esse canal. O próximo capítulo apresenta a proposta de esquemas, que juntamente com o RLP, visa a melhorar o desempenho do TCP sobre canal sem fio.

Capítulo 4

Esquema Proposto

O objetivo deste trabalho é investigar e propor um esquema de transmissão adaptativo da camada física, que maximiza o *throughput* do TCP dependendo da faixa SNR de operação [Oliveira et al., 2008a][Oliveira et al., 2008b]. Tal abordagem nos permite introduzir uma degradação de desempenho suave na faixa de operação do VBLAST. É importante ressaltar que esta estratégia é diferente do que simplesmente alternar entre esquemas VBLAST e STBC conforme proposto em [Toledo and Wang, 2006], onde foi utilizada uma modulação fixa.

4.1 Avaliação da Camada Física

Foi adotado esquemas VBLAST usando diferentes taxas de codificação e diferentes ordens de modulação. A arquitetura proposta para o sistema VBLAST é mostrada na figura 4.1. Foi assumido um esquema MIMO com $M = N = 4$ antenas. Os quatro subfluxos de dados

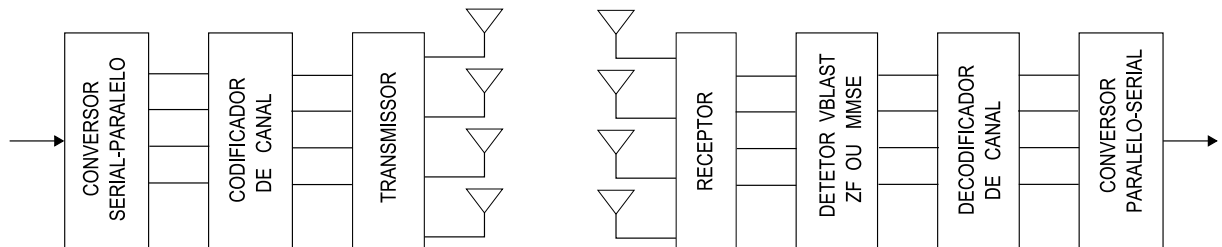


Figura 4.1: Arquitetura Proposta.

são codificados separadamente usando um codificador convolucional de memória igual a dois, ou seja, o codificador convolucional tem dois registros de deslocamento. A decisão usada no Viterbi, algoritmo de decodificação do código convolucional, foi a *hard*. Os parâmetros dos sistemas VBLAST simulados são apresentados na tabela 4.1. A modulação 8-QAM, modulação em amplitude e quadratura, foi usada no lugar da modulação 8-PSK porque possui um melhor desempenho. Este melhor desempenho ocorre porque a distância mínima entre símbolos da modulação 8-QAM é maior do que a distância mínima entre os símbolos da modulação 8-PSK [Proakis, 2001]. A BER para estes esquemas foi obtida através de simulação usando o software MATLAB. Nas simulações considerou-se um canal com desvanecimento plano e independente de símbolo para símbolo transmitido. As figuras 4.2 e 4.3 apresentam os resultados das simula-

ções no MATLAB para os detectores ZF e MMSE, respectivamente, avaliando-se BER x SNR.

Tabela 4.1: Esquemas codificados VBLAST simulados

Esquema	Modulação	Taxa do Código	Gerador do Código	bits/s/Hz
S1	8-QAM	2/3	K=[2 2];g=[3 1 3,1 2 2]	8
S2	8-QAM	1/2	K=3;g=[5 7]	6
S3	8-QAM	1/3	K=3;g=[5 7 7]	4
S4	QPSK	2/3	K=[2 2];g=[3 1 3,1 2 2]	16/3
S5	QPSK	1/2	K=3;g=[5 7]	4
S6	QPSK	1/3	K=3;g=[5 7 7]	8/3
S7	BPSK	2/3	K=[2 2];g=[3 1 3,1 2 2]	8/3
S8	BPSK	1/2	K=3;g=[5 7]	2
S9	BPSK	1/3	K=3;g=[5 7 7]	4/3

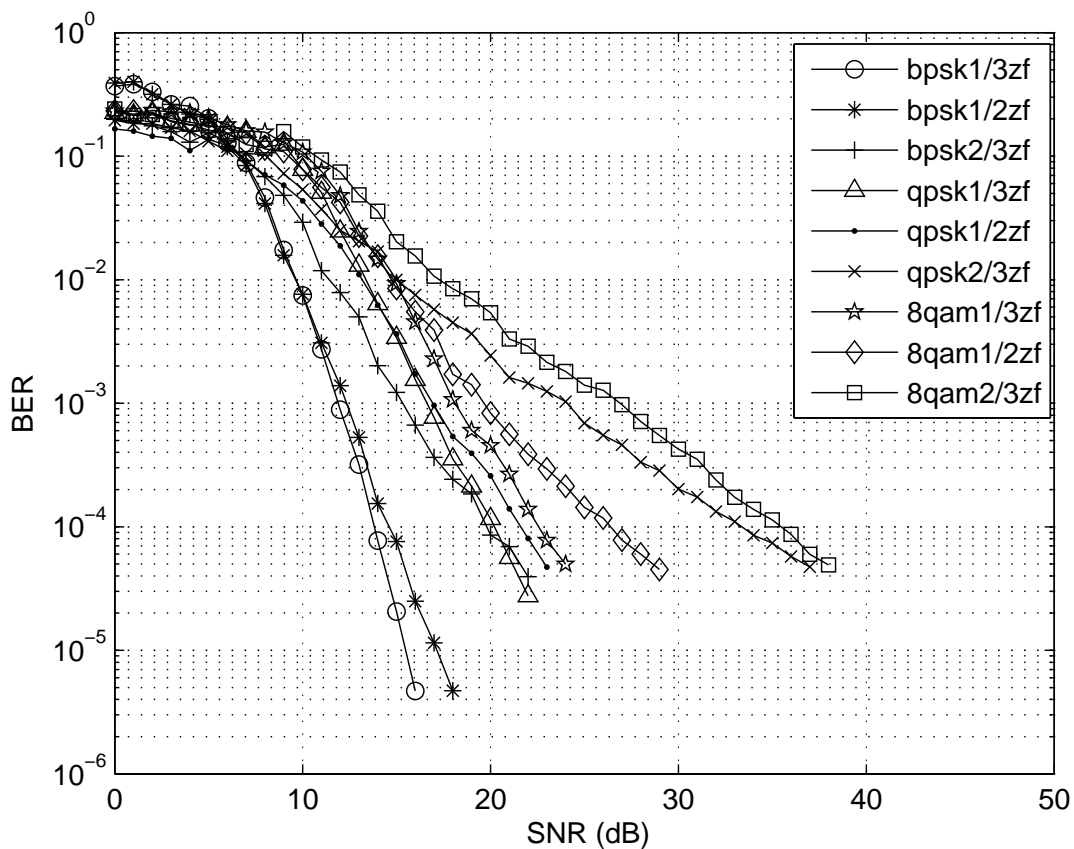


Figura 4.2: Resultados das simulações da BER x SNR para o detector ZF.

Para estes esquemas apresentados na tabela 4.1, determina-se, a partir da equação 4.1, a FER (*Frame Error Rate*). As figuras 4.4 e 4.5 apresentam os gráficos da FER para cada esquema.

$$FER = 1 - ((1 - BER)^T), \quad (4.1)$$

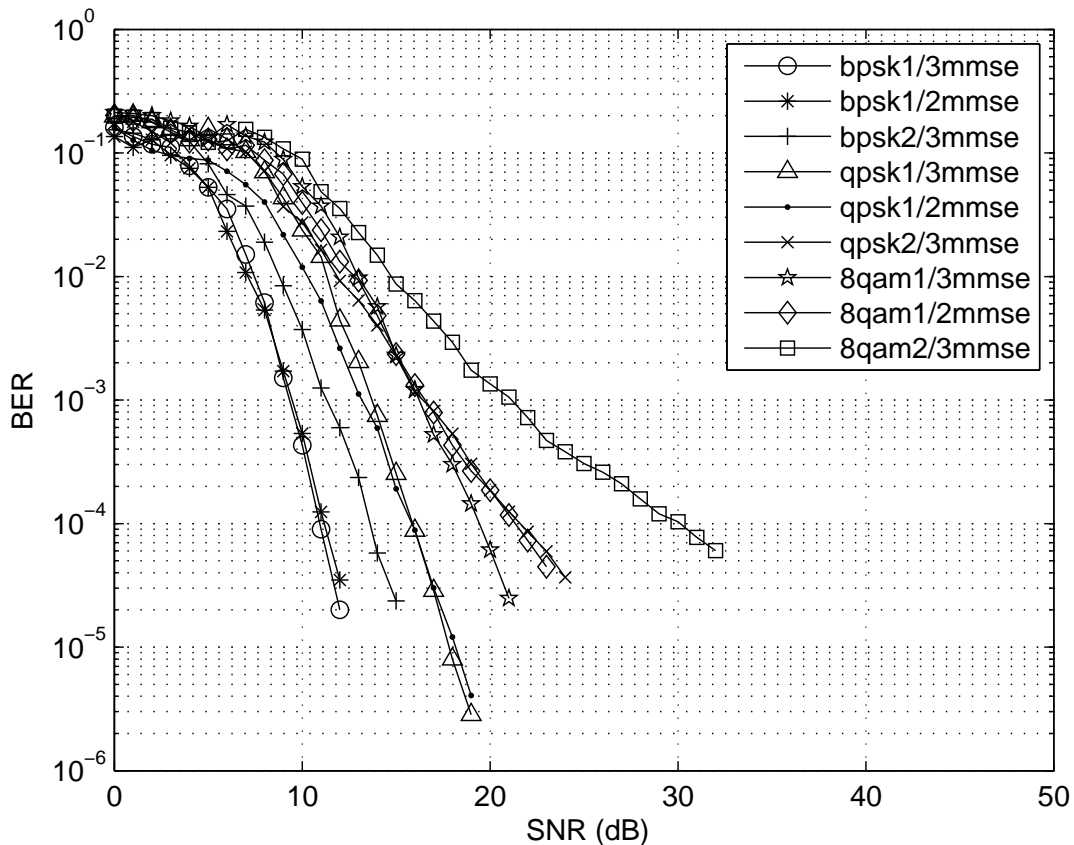


Figura 4.3: Resultados das simulações da BER x SNR para o detector MMSE.

onde T é o tamanho do quadro RLP, em bits, apresentado na tabela 4.2 para cada esquema considerado.

4.2 Avaliação da Camada de Rede

O valor da FER, obtida da equação 4.1, foi usado como parâmetro para modelar o enlace sem fio no cenário de rede apresentado na figura 4.6. O cenário é constituído por seis nós, onde os nós 2 e 3 se comunicam através do enlace sem fio. Os nós 1 e 4 representam os nós intermediários, roteadores como exemplo, que um segmento TCP geralmente precisa passar antes de chegar ao seu destino. A taxa do enlace sem fio varia de acordo com o esquema a ser utilizado conforme a tabela 4.2 e o seu atraso de propagação é de 30 ms. Os demais enlaces são cabeados e possuem uma taxa de 1 Mbps e um atraso de propagação de 10 ms. Um arquivo de dados é transmitido usando o protocolo de aplicação FTP (*File Transfer Protocol*) através do TCP. A pilha de protocolos TCP/IP/LL/PHY, implementado para os nós 0 e 5, é responsável pela execução do FTP e a transferência dos dados do arquivo. A pilha de protocolos PPP/RLP/LL/PHY, implementado para os nós 2 e 3, transmite os dados através do enlace sem fio. O padrão RLP implementado na camada de enlace de rádio é o protocolo 3GPP2 RLP3 [3GP, 2004]. A versão TCP implementada foi o protocolo TCP Reno. O tamanho máximo do segmento de dados MSS (*Maximum Segment Size*) usado foi de 536 bytes. Este é o tamanho

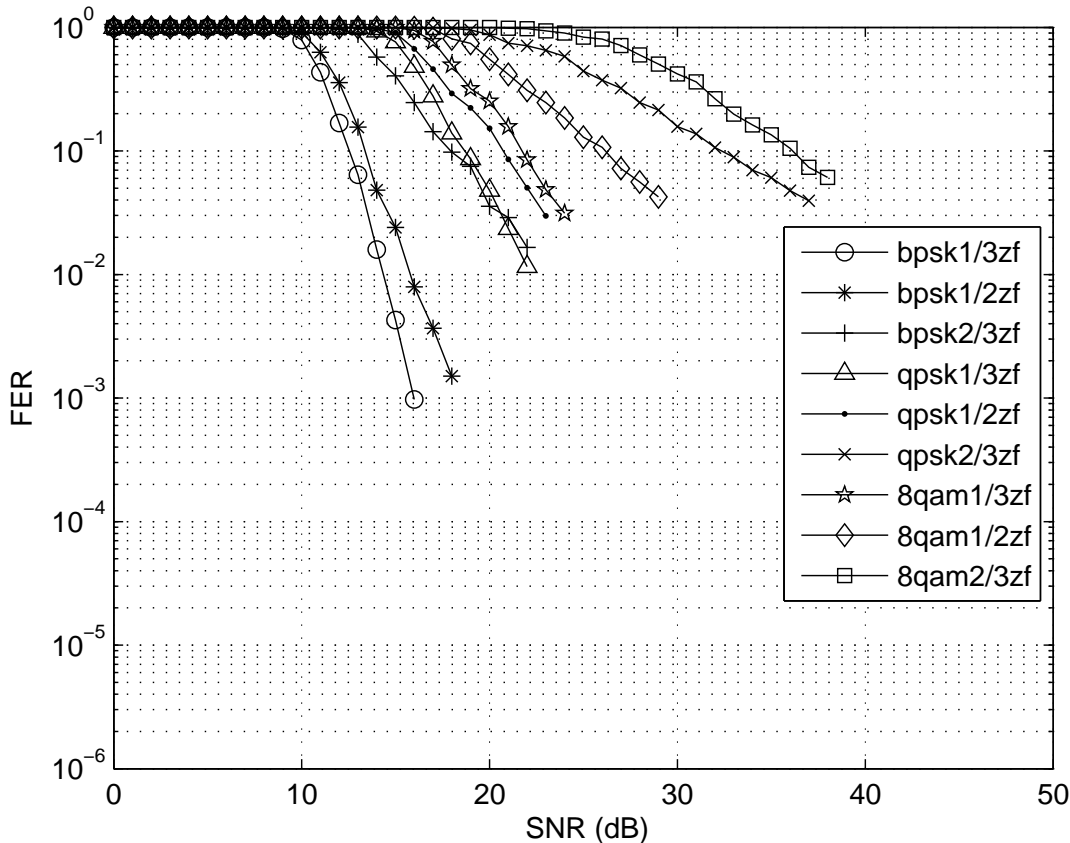


Figura 4.4: Resultados da FER x SNR para o detector ZF.

padrão do MSS usado na camada RLP [3GP, 2004]. O cenário de rede foi simulado no software NS2 [ns2, 2007].

Para se comparar corretamente os esquemas simulados, que possuem diferentes eficiências espectrais, uma compensação na taxa de transmissão do enlace sem fio foi feita. Para o esquema com maior eficiência espectral (8QAM; $r=\frac{2}{3}$), a capacidade do enlace foi especificada como sendo de 64kbps, conforme adotado em [Toledo and Wang, 2006]. Os outros esquemas possuem uma alocação de capacidade de enlace proporcional, conforme mostrado na tabela 4.2. O próprio protocolo RLP determinou o tamanho do seu quadro de acordo com a taxa do enlace. As figuras 4.7 e 4.8 apresentam os resultados das simulações para o *throughput* do TCP em função da SNR para os diferentes esquemas codificados VBLAST-ZF e VBLAST-MMSE. Destas curvas pode-se identificar a estratégia de codificação apropriada que deve ser implementada para cada faixa de SNR. Os esquemas selecionados são apresentados na tabela 4.3 para detector VBLAST-ZF e na tabela 4.4 para detector VBLAST-MMSE. É importante ressaltar que os esquemas QPSK $r = \frac{1}{3}$, QPSK $r = \frac{2}{3}$ e 8-QAM $r = \frac{1}{3}$ não foram selecionados para nenhuma faixa de operação. Analizando as figuras 4.10 e 4.9 pode-se observar que haveria uma perda significativa de desempenho se o método de seleção não fosse utilizado. Considere por exemplo que o sistema opere apenas com o esquema QPSK $r = \frac{1}{2}$. Para a faixa de SNR abaixo de 16 dB, considerando o detector VBLAST-ZF, e para a faixa de SNR abaixo de 13 dB, considerando o detector VBLAST-MMSE, o desempenho do TCP é zero. Para a faixa de SNR acima de 24 dB,

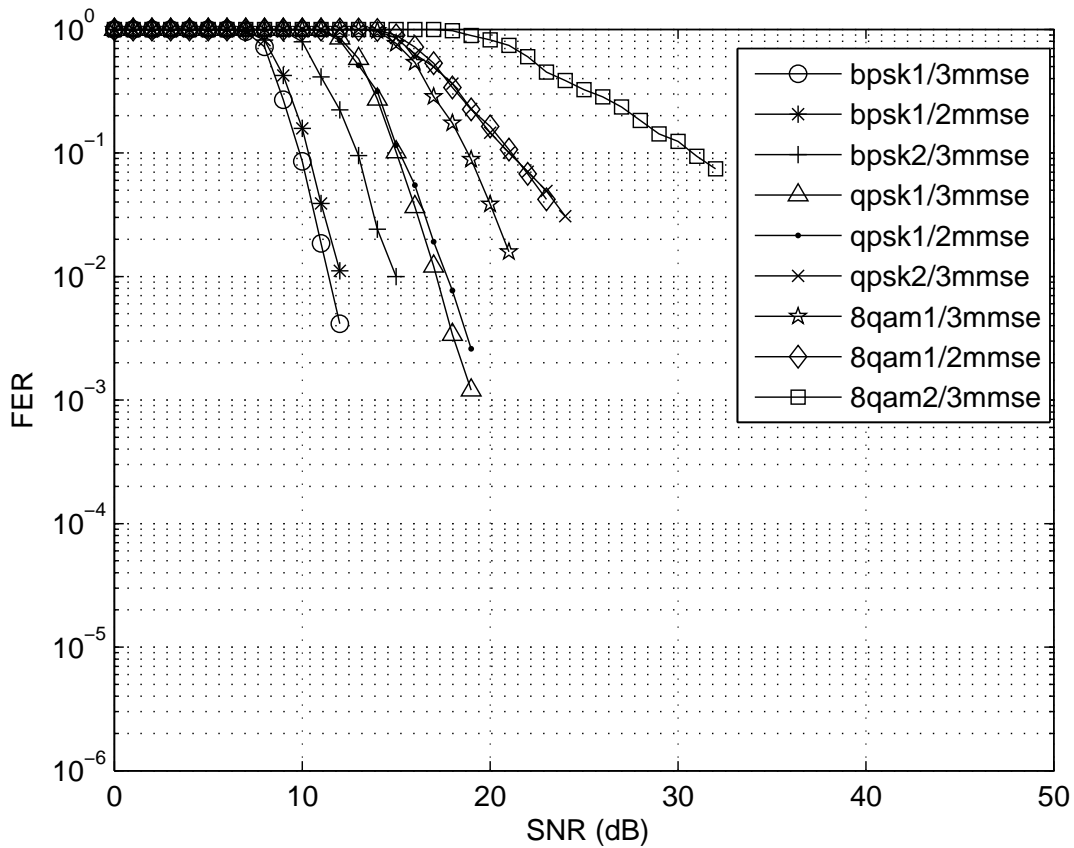


Figura 4.5: Resultados da FER x SNR para o detector MMSE.

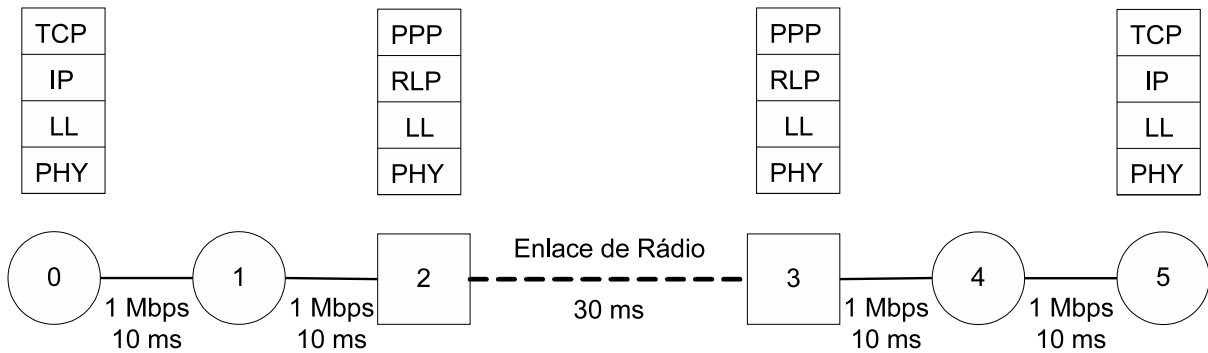


Figura 4.6: Cenário de rede implementado no ns2.

considerando o detector VBLAST-ZF, e para a faixa de SNR acima de 18 dB, considerando o detector VBLAST-MMSE, o throughput do TCP atinge uma taxa de saturação de 30kbps. Fica claro que um desempenho muito melhor é obtido quando o esquema apropriado de modulação e codificação é selecionado para cada faixa de SNR, conforme indicados na tabela 4.3 para detector VBLAST-ZF e na tabela 4.4 para detector VBLAST-MMSE.

A figura 4.10, para detector VBLAST-ZF, e a figura 4.9, para detector VBLAST-MMSE, apresentam a curva de operação para o *throughput* do protocolo TCP operando com os

Tabela 4.2: Parâmetros de simulação da camada de enlace

Esquema	bits/s/Hz	Taxa do Enlace (kbps)	Quadro RLP (bytes)
S1	8	64	160
S2	6	48	120
S3	4	32	80
S4	16/3	128/3	106
S5	4	32	80
S6	8/3	64/3	53
S7	8/3	64/3	53
S8	2	16	40
S9	4/3	32/3	26

esquemas propostos para cada faixa de SNR. A linha vertical pontilhada delimita a faixa ótima de operação de cada esquema apresentado na tabela 4.3 para detector VBLAST-ZF e na tabela 4.4 para detector VBLAST-MMSE. A figura 4.11 apresenta uma comparação entre as curvas de operação para os detectores VBLAST-ZF e VBLAST-MMSE.

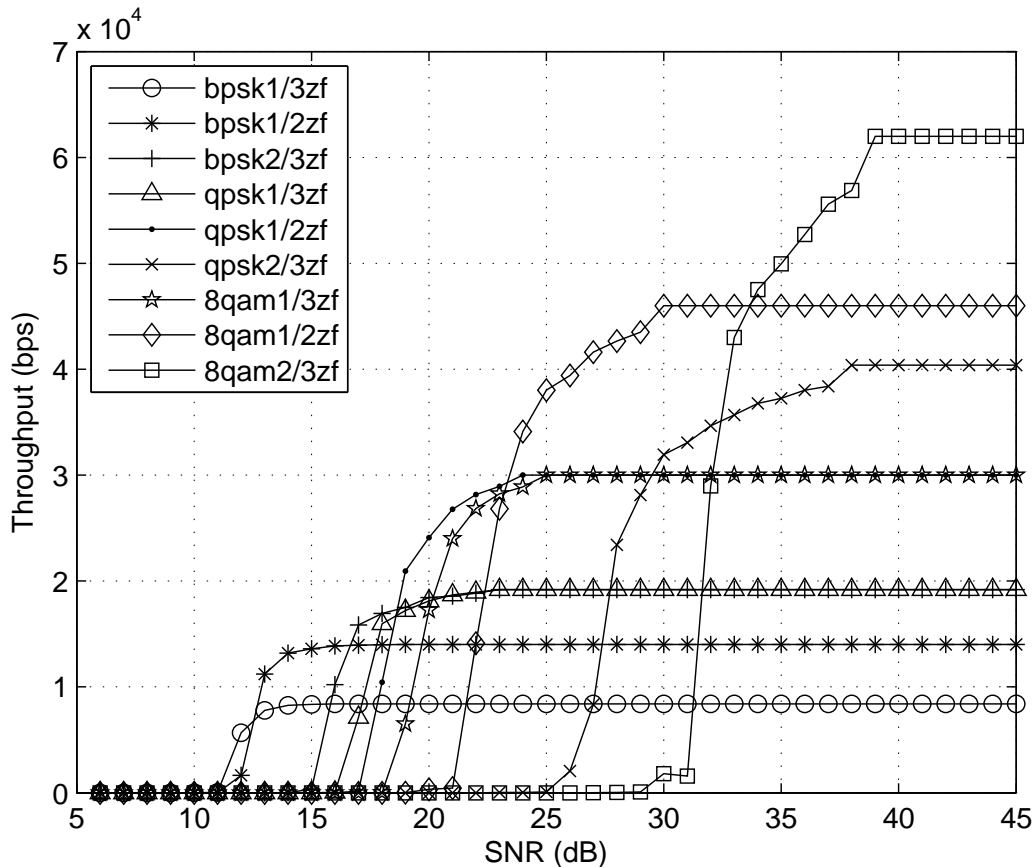


Figura 4.7: Desempenho do TCP para os esquemas VBLAST-ZF.

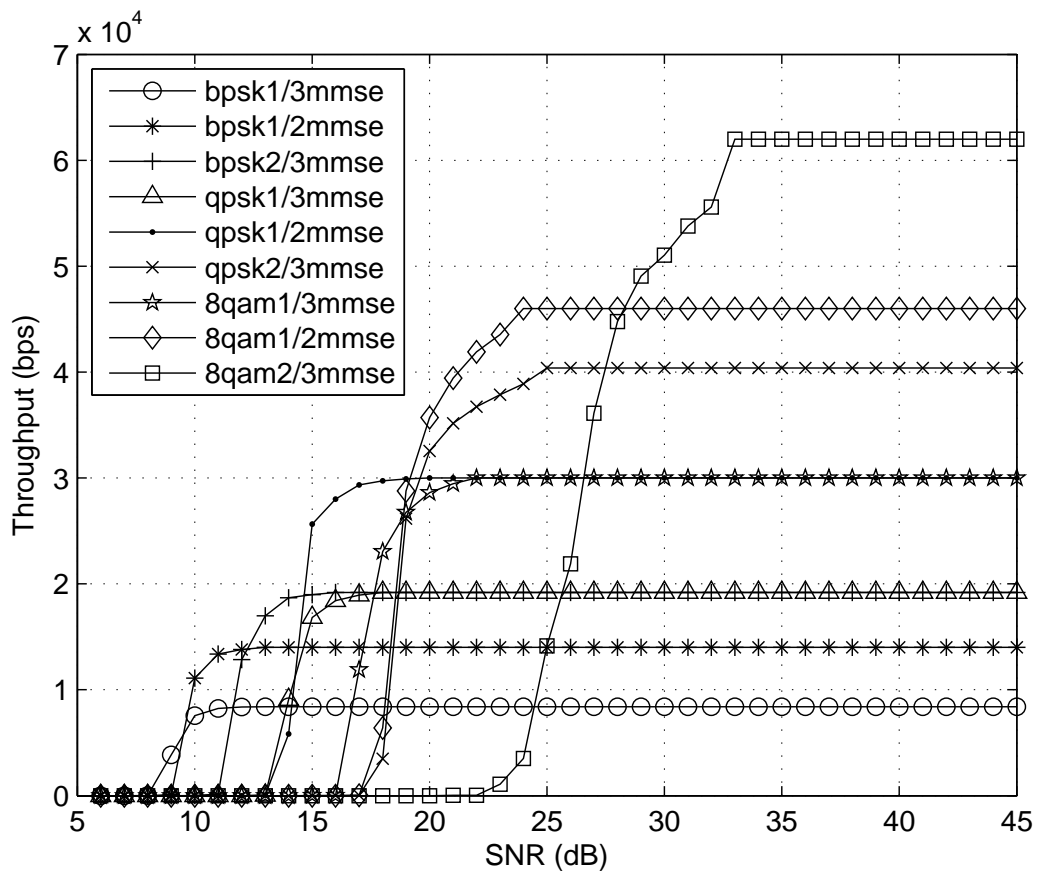


Figura 4.8: Desempenho do TCP para os esquemas VBLAST-MMSE.

Tabela 4.3: Faixa de operação para os esquemas codificados VBLAST-ZF

Esquema Selecionado	Faixa de SNR
BPSK 1/3 - S6	$\text{SNR} \leq 12.5 \text{ dB}$
BPSK 1/2 - S5	$12.5 \text{ dB} \leq \text{SNR} \leq 16.6 \text{ dB}$
BPSK 2/3 - S4	$16.6 \text{ dB} \leq \text{SNR} \leq 18.6 \text{ dB}$
QPSK 1/2 - S3	$18.6 \text{ dB} \leq \text{SNR} \leq 23.3 \text{ dB}$
8QAM 1/2 - S2	$23.3 \text{ dB} \leq \text{SNR} \leq 33.6 \text{ dB}$
8QAM 2/3 - S1	$\text{SNR} \geq 33.6 \text{ dB}$

Tabela 4.4: Faixa de operação para os esquemas codificados VBLAST-MMSE

Esquema Selecionado	Faixa de SNR
BPSK 1/3 - S6	$\text{SNR} \leq 9.5 \text{ dB}$
BPSK 1/2 - S5	$9.5 \text{ dB} \leq \text{SNR} \leq 12.2 \text{ dB}$
BPSK 2/3 - S4	$12.2 \text{ dB} \leq \text{SNR} \leq 14.6 \text{ dB}$
QPSK 1/2 - S3	$14.6 \text{ dB} \leq \text{SNR} \leq 19.1 \text{ dB}$
8QAM 1/2 - S2	$19.1 \text{ dB} \leq \text{SNR} \leq 28.3 \text{ dB}$
8QAM 2/3 - S1	$\text{SNR} \geq 28.3 \text{ dB}$

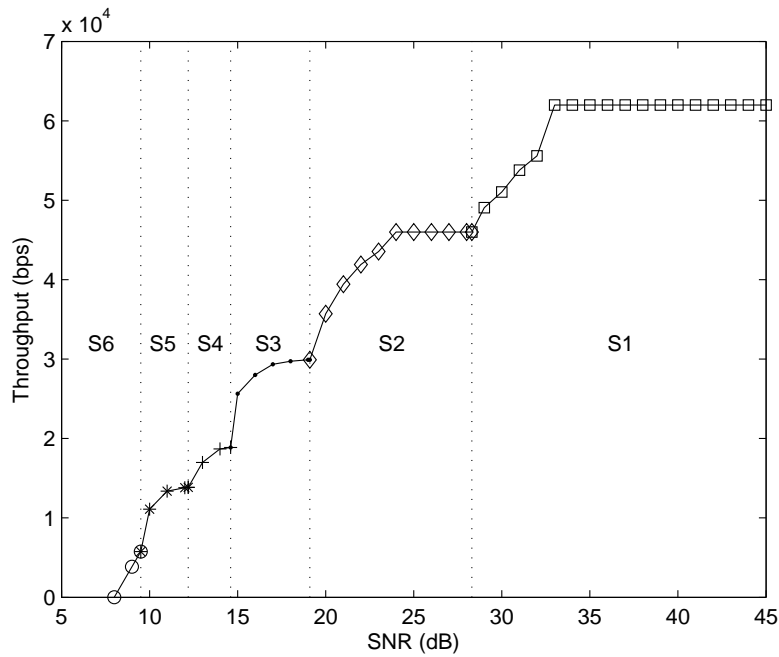


Figura 4.9: Desempenho do TCP usando as estratégias otimizadas de transmissão para o detector VBLAST-MMSE.

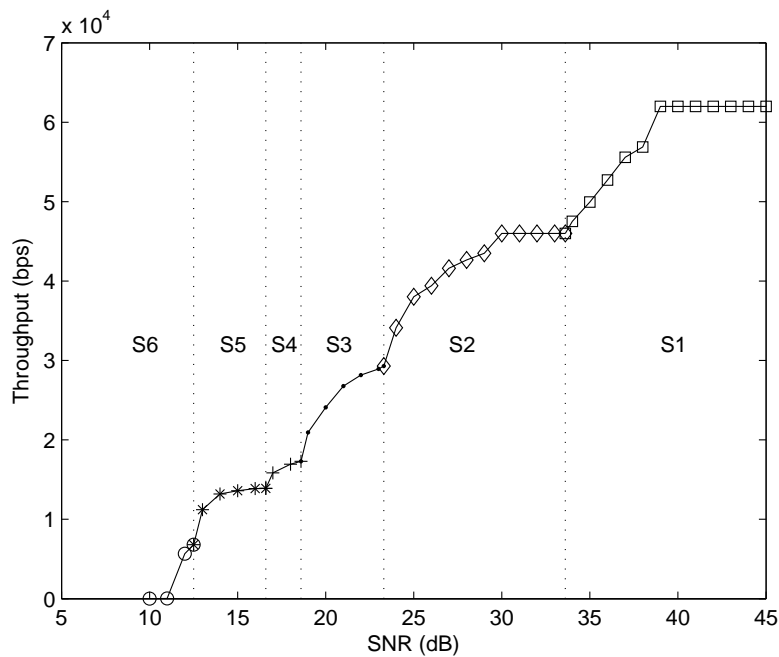


Figura 4.10: Desempenho do TCP usando as estratégias otimizadas de transmissão para o detector VBLAST-ZF.

4.3 Conclusão

O esquema proposto neste trabalho é realizado por um esquema MIMO com $M = N = 4$ codificado por código convolucional com três diferentes taxas e três diferentes ordens

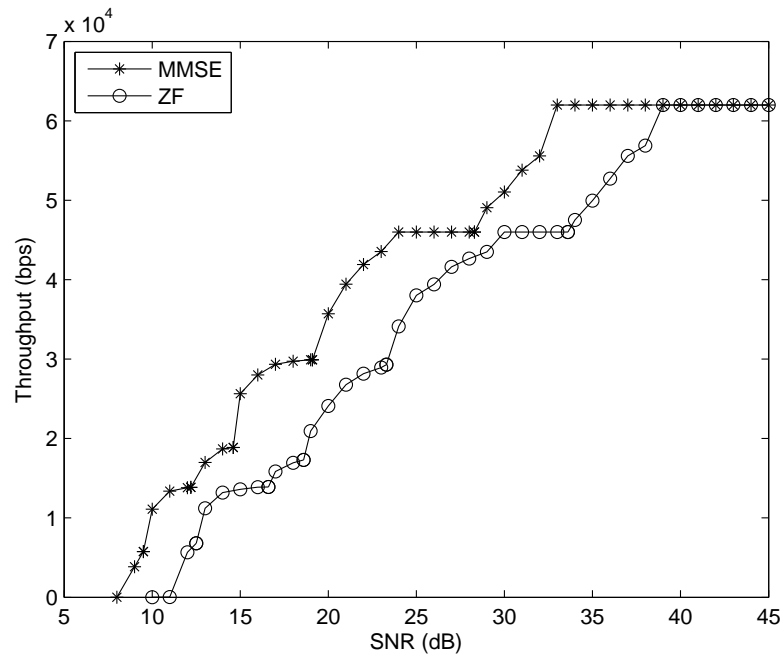


Figura 4.11: Comparação entre as estratégias otimizadas de transmissão para os detectores VBLAST-ZF e VBLAST-MMSE.

de modulação. O esquema proposto é estudado para os dois algoritmos de detecção usados na prática: ZF e MMSE. A combinação das diferentes taxas de codificação e modulação resulta em seis sistemas diferentes. Estes sistemas, com diferentes eficiências espectrais, apresentam *throughput* do TCP com diferentes inícios, na faixa de SNR, e saturação. Este fato permite que se proponha a utilização de um chaveamento entre estes sistemas, de acordo com a faixa de SNR, de maneira que o *throughput* do TCP seja o maior possível. Para a SNR variando de 12.5 a 33.6 dB, considerando o detector VBLAST-ZF, e para a SNR variando de 9.5 a 28.3 dB, considerando o detector VBLAST-MMSE, utiliza-se o chaveamento entre seis sistemas sendo três desnecessário o uso. A utilização de tal esquema proposto otimiza o desempenho do protocolo TCP, quando o mesmo está trabalhando sobre sistemas de comunicação sem fio MIMO-VBLAST e quando a SNR do sistema apresenta condições variantes.

Capítulo 5

Conclusões

Neste trabalho foi investigado o desempenho das camadas TCP/RLP para transmissão usando diferentes esquemas VBLAST codificados. Os esquemas utilizam códigos convolucionais com diferentes taxas e ordens de modulação. Foram simulados nove esquemas diferentes, porém apenas seis foram propostos para efetuar o chaveamento que otimiza o *throughput* do TCP. Isto aconteceu porque pela tabela 4.2 os esquemas S_3 e S_5 e os esquemas S_6 e S_7 possuem a mesma eficiência espectral, porém o esquema S_5 tem melhor desempenho do que o S_3 e o esquema S_7 tem melhor desempenho do que o S_6 em termos de BER x SNR. Para sistemas com a mesma eficiência espectral, o sistema a ser usado é o que possui melhor desempenho em termos de BER x SNR. A não utilização do esquema S_4 ocorreu naturalmente devido a suas características de eficiência espectral e desempenho em termos de BER x SNR serem inferiores aos outros esquemas.

Observando-se gráficos 4.2 e 4.3 percebe-se o melhor desempenho, que já era esperado, do detector MMSE. Este melhor desempenho em termos de BER x SNR proporcionou, também, um melhor desempenho do *throughput* do TCP visto no gráfico 4.11. Baseado nesta conclusão, pode-se dizer que quanto menor a BER, de um esquema implementado na camada física, maior será o *throughput* do TCP. Isto acontece porque quanto menor a BER menor, também, será a perda de segmentos TCP e isto evita retransmissões e, conseqüentemente, que o protocolo reinicie na fase de inicialização lenta (*slow start*), fase que o *throughput* é reduzido.

O resultado principal derivado das simulações apresentadas nas figuras 4.2 e 4.3 é a possibilidade de modificação da taxa de codificação e da ordem da modulação de maneira a se maximizar o *throughput* do TCP, dependendo da faixa de operação de SNR. Esta abordagem permite introduzir uma degradação suave de desempenho na faixa de operação do VBLAST. Alguns aspectos ainda devem ser investigados em trabalhos futuros, como outros esquemas de codificação de canal, o comportamento do TCP para uma outra modelagem de canal e a proposta de um protocolo específico para adaptar dinamicamente a estratégia de transmissão de acordo com as condições do canal.

Referências Bibliográficas

- [Pse, 1998] (1998). *R.Buck,Pseudoinverse*. <http://rkb.home.cern.ch/rkb/AN16pp/node220.html>.
- [RLC, 1999] (1999). *Third Generation Partnership Project, RLC Protocol Specification*. 3GTS 25.322.
- [3GP, 2004] (2004). *Data Service Options for Spread Spectrum Systems: Radio Link Protocol Type 3*. Third Generation Partnership Project 2 (3GPP2).
- [ns2, 2007] (2007). *Network Simulator 2*. <http://www.isi.edu/nsnam/ns>.
- [Alamouti, 1998] Alamouti, S. M. (1998). A simple transmit diversity technique for wireless communications. *IEEE Journal on Select Areas in Communications*, 116:1451–1458.
- [Allman et al., 2002] Allman, M., Floyd, S., and Partridge, C. (2002). Increasing tcp's initial window. Technical Report 3390, Internet Eng. Task Force (IETF) - Request for Comments RFC.
- [Bakre and Badrinath, 1995] Bakre, A. and Badrinath, B. (1995). I-tcp:indirect tcp for mobile hosts. In *Proceedings Int'l Conf. Distributed Computing Systems*.
- [Balakrishnan et al., 1996] Balakrishnan, H., Padmanabham, V. N., Seshan, S., and Katz, R. H. (1996). A comparison of mechanisms for improving tcp performance over wireless links. In *Proc. of ACM SIGCOMM*.
- [Chan and Ramjee, 2002] Chan, M. C. and Ramjee, R. (2002). Tcp/ip performance over 3g wireless links with rate and delay variation. In *Proc. of The ACM Annual International Conference on Mobile Computing and Networking (Mobicom)*.
- [Comer, 2007] Comer, D. E. (2007). *Redes de Computadores e Internet*. Bookman.
- [C.Parsa and Garcia-Luna-Aceves, 1999] C.Parsa and Garcia-Luna-Aceves, J. (1999). Improving tcp congestion control over internets with heterogeneous transmission media. In *Proc. Seventh IEEE Int'l Conf. Network Protocols (ICNP)*.
- [D.Gesbert et al., 2003] D.Gesbert, Shafi, M., Shiu, D. S., Smith, P., and Naguib, A. (2003). From theory to practice: An overview of mimo space-time coded wireless systems. *IEEE Journal on Selected Areas in Communication*, 21(3):281–302.
- [Foschini, 1996] Foschini, G. J. (1996). Layered space-time architecture for wireless communication in a fading environment when using multiple antennas. *Bell Labs Technical Journal*, 1(2):41–59.

- [Foschini and Gans, 1998] Foschini, G. J. and Gans, M. J. (1998). On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communication*, 6(3):314–335.
- [Haykin, 2004] Haykin, S. (2004). *Sistemas de Comunicações Analógicas e Digitais*. Bookman.
- [Haykin and Moher, 2005] Haykin, S. and Moher, M. (2005). *Modern Wireless Communications*. Upper Saddle River, NJ: Pearson Prentice Hall.
- [Jr et al., 2005] Jr, W. C. F., Cavalcante, C. C., and Cavalcanti, F. R. (2005). Estratégias mimo-ofdm para sistemas de comunicações móveis. In *Simpósio Brasileiro de Telecomunicações - SBRT'05*, pages 23–54, Campinas SP.
- [Oliveira et al., 2008a] Oliveira, A. P., PELLEENZ, M. E., SOUZA, R. D., and PENNA, M. C. (2008a). Avaliação de desempenho do tcp em sistemas vblast codificados. In *Anais XXVI Simpósio Brasileiro de Telecomunicações*, pages 1–6, Rio de Janeiro.
- [Oliveira et al., 2008b] Oliveira, A. P., PELLEENZ, M. E., SOUZA, R. D., and PENNA, M. C. (2008b). On the performance of the tcp/rtp layers over vblast coded schemes. In *Proceedings of the IEEE International Symposium on Wireless Communication Systems (ISWCS)*, pages 1–5, Reykjavik.
- [Proakis, 2001] Proakis, J. (2001). *Digital Communications*. McGraw-Hill.
- [Rappaport, 1996] Rappaport, T. S. (1996). *Wireless Communication: Principles & Practice*. Prentice Hall.
- [S.Floyd and T.Henderson, 1999] S.Floyd and T.Henderson (1999). The new reno modification to tcps fast recovery algorithm. Technical Report 2582, Internet Eng. Task Force (IETF) - Request for Comments RFC.
- [Shiu and Kahn, 1999] Shiu, D. and Kahn, J. M. (1999). Scalable layered space-time codes for wireless communications: Performance analysis and design criteria. In *Proc. of The IEEE Wireless Communications and Networking Conference (WCNC)*, pages 159–163.
- [Sklar, 1997] Sklar, B. (1997). Rayleigh fading channels in mobile digital communication systems part i: Characterization. *IEEE Communication Magazine*, 32(4):90–100.
- [Stamoulis and Al-Dhahir, 2003] Stamoulis, A. and Al-Dhahir, N. (2003). Impact of space-time block codes on 802.11a network throughput. *IEEE Transactions on Information Theory*, 49(5):1073–1096.
- [Tanenbaum, 2003] Tanenbaum, A. S. (2003). *Redes de Computadores*. Campus.
- [Toledo and Wang, 2006] Toledo, A. L. and Wang, X. (2006). Tcp performance over wireless mimo channels with arq and packet combining. *IEEE Transactions on Mobile Computing*, 5(3):208–223.
- [Wolniansky et al., 1999] Wolniansky, P. W., Foschini, G. J., Golden, G. D., and Valenzuela, R. A. (1999). Detection algorithm and initial laboratory results using v-blast space-time communication architecture. *Electronics Letters*, 35(1).

- [Wolniansky et al., 1998] Wolniansky, P. W., Foschini, G. J., Golden, G. D., Valenzuela, R. A., Labs, B., Technologies, L., and Holmdel, N. J. (1998). V-blast: An architecture for realizing very high data rates over the rich-scattering wireless channel. Technical report, Bell Labs.
- [Yang et al., 2006] Yang, G., Shen, D., and Li, V. O. K. (2006). Unequal error protection for mimo systems with a hybrid structure. In *Proceedings IEEE International Symposium on Circuits and Systems*, pages 682–685.