

Evidências da Falha do Modelo de *Gilbert-Elliott* em Redes sem Fio

Luis J. Rohling, Carlos A. G. da Silva, Evelio M. G. Fernández e Carlos M. Pedroso

Resumo— Este artigo descreve evidências de falha no modelo de Gilbert-Elliott na modelagem de perda de pacotes em redes WiFi. Para este estudo foi realizada a amostragem de tráfego na rede WiFi do Departamento de Engenharia Elétrica da Universidade Federal do Paraná. Os resultados obtidos indicam a existência de uma dependência temporal entre os surtos de perda, e também um comportamento de cauda pesada no tempo de duração do estado bom, a partir da avaliação realizada com a utilização de ferramentas estatísticas. Ao final é apresentado um modelo alternativo para representar o comportamento observado.

Palavras-Chave— Gilbert-Elliott, Perda de Pacotes, Redes sem fio

Abstract— This article presents evidences of failure of the Gilbert-Elliott packet loss model for WiFi networks. To perform the analysis, real traffic samples were collected in Wifi network of the Department of Electrical Engineering, Federal University of Parana. The results indicate the existence of a time dependence of the loss bursts and heavy tail behavior in the duration of the good state. The samples were evaluated by using statistical tools. It is also presented an alternative model to represent the observed behavior.

Keywords— Gilbert-Elliott, Paket loss, Wireless LAN

I. INTRODUÇÃO

Entre as tecnologias de redes de acesso está o padrão IEEE 802.11 (WiFi), que tem sido a principal opção utilizada nos ambientes onde existe mobilidade de usuários, pois as redes cabeadas não atendem este tipo de necessidade. Estes ambientes compreendem, principalmente, as áreas de grande circulação de pessoas, onde empresas e governo buscam uma solução para disponibilizar o acesso à Internet. Neste cenário, deve-se considerar também o crescente número de usuários conectados através dessa tecnologia, e que deverá sofrer um acréscimo considerável com a nova onda chamada de Internet das Coisas (IoT), que envolve diversas tecnologias, protocolos e aplicações [1]. Assim, é de fundamental importância entender o comportamento destas redes, de forma a prever o desempenho da rede, dimensionando adequadamente os recursos necessários. Um dos parâmetros críticos a ser considerado é a ocorrência de erros na transmissão, que se reflete na perda de pacotes, impactando significativamente no desempenho e na percepção do usuário quanto à qualidade da rede.

Para prever este desempenho da rede são utilizados modelos de perda, sendo que um dos modelos mais empregados nestes estudos é o modelo de Gilbert-Elliott [2] [3]. O objetivo

Luis J. Rohling, Carlos A. G. da Silva, Evelio M. G. Fernández e Carlos M. Pedroso, Departamento de Engenharia Elétrica, Universidade Federal do Paraná, E-mail: lrohling@gmail.com, carlos.gouvea@ufpr.br, eve-lio@eletrica.ufpr.br, pedroso@eletrica.ufpr.br

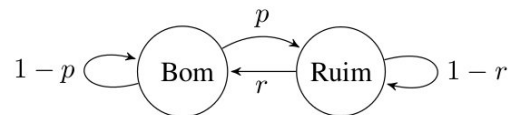


Fig. 1. Modelo de Gilbert-Elliott de 2 estados

deste trabalho é evidenciar que, para a análise das perdas de pacotes em redes sem fio, o modelo de Gilbert-Elliott não é o mais adequado, apresentando-se um modelo alternativo. Para esta finalidade a metodologia empregada foi a realização de diversas medições de perdas de pacotes, em uma rede sem fio real, analisando-se estatisticamente os dados obtidos e elaborado-se um novo modelo. Este modelo foi validado pela comparação dos valores reais obtidos com modelo proposto.

Além desta seção introdutória, este artigo está organizado da seguinte forma: A seção II apresenta alguns conceitos fundamentais dos modelos envolvidos; a descrição dos experimentos e metodologia é indicado na seção III; a seção IV apresenta os resultados, bem como uma proposta para um modelo que melhor descreve as perdas observadas; e na seção V são apresentadas as conclusões finais e trabalhos futuros a serem desenvolvidos.

II. CONCEITOS FUNDAMENTAIS

A. Modelo de Gilbert-Elliott

O modelo apresentado por Gilbert [4] e Elliott [5] é uma das principais abordagens utilizadas atualmente para a modelagem de perdas de pacotes. Este modelo utiliza uma cadeia de Markov de 2 estados, denominados *Bom* e *Ruim*, com probabilidade de incidência de erros pequena e grande, respectivamente [6].

O modelo de Cadeias de Markov é aplicado para fenômenos que não apresentam dependência temporal, ou seja, a transição para os estados *Bom* e *Ruim* depende apenas do estado atual e não dos anteriores. A recomendação G.1050 da ITU-T [7] sugere modelos para avaliar o desempenho de transmissão multimídia sobre o protocolo IP. Esta recomendação utiliza o modelo de Gilbert-Elliott para modelar a perda de pacotes em redes. Também é apresentada uma possível parametrização do modelo para o acesso, redes locais e núcleo da Internet com diversos níveis de qualidade.

No modelo de Gilbert-Elliott a probabilidade de erro depende do estado atual, com uma probabilidade de erro de $1-k$ para o estado *Bom* e $1-h$ para o estado *Ruim*. A Figura 1 ilustra a transição de estados do modelo. A probabilidade para a troca do estado *Bom* para *Ruim* é dado por p e a

probabilidade para a troca do estado *Ruim* para o *Bom* é dada por r . Para determinar-se a probabilidade de estado estacionário para este modelo aplica-se o teorema do limite das cadeias de Markov [8], obtendo-se uma probabilidade de perda dada por $r(1-k)/(p+r)$ para o estado *Bom* e de $p(1-h)/(p+r)$ para o estado *Ruim*. A cadeia de Markov pode ser ampliada para mais estados, o que torna o modelo de Gilbert-Elliott mais complexo, exigindo maiores recursos computacionais para sua parametrização [9]. Além disso, quanto maior a ordem da cadeia de Markov adotada, maior é a precisão do modelo.

O modelo Gilbert-Elliott porém, por ser baseado em cadeias de Markov, somente pode ser aplicado quando não existe dependência temporal entre os surtos de erros. A distribuição de probabilidade do tempo de duração dos estados *Bom* e *Ruim* deveria apresentar então uma distribuição do tipo exponencial, para estar aderente ao modelo de Gilbert-Elliott.

Porém, o método de controle de acesso usado pelo Wifi pode causar um efeito de memória, principalmente em situações de maior congestionamento da rede. Ou seja, a ocorrência de um surto de erros pode aumentar a probabilidade de um surto no futuro próximo. Deste modo, a sequência de estados *Bom* e *Ruim* apresentará uma dependência temporal, não sendo o modelo de Gilbert-Elliott totalmente adequado para a modelagem dos erros em redes sem fio, como demonstrado neste artigo.

B. Trabalhos relacionados

Um estudo realizado sobre o comportamento do tráfego de aplicações multimídia em redes sem fio [10] já apontava para um modelo baseado em distribuição logarítmica para a perda de pacotes neste tipo de tráfego, analisando o comportamento das rajadas de erros em diversos cenários de transmissão. Para cada cenário foi determinado o parâmetro da distribuição, em função dos diferentes tamanhos de pacotes, característicos do tráfego de voz ou de vídeo.

Em outro estudo também é proposto um modelo para perda de pacotes, porém nesse caso utilizando Cadeias de Markov Escondidas (*Hidden Markov Model*, HMM), sendo que melhores resultados foram atingidos com um modelo de 11 estados [11].

A partir dos modelos de Gilbert-Elliott e *PH-Distributed Fault Model* [12] é proposto um modelo formado por um conjunto de estados bons $\{G_1, G_2, G_3, G_4\}$, cada um com uma probabilidade de transição para um único estado ruim, e o estado ruim com probabilidade de transição apenas para o estado G_1 . Resultados simulados indicam a boa aderência do modelo proposto aos dados reais.

C. Distribuições de Cauda Pesada

Variáveis aleatórias caracterizadas por distribuições de cauda pesada apresentam uma maior incidência de valores grandes. Uma das distribuições de cauda pesada mais utilizadas é a distribuição de Pareto tipo II, ou distribuição de Lomax, dada por [13]:

$$P[X \leq x] = 1 - \left(1 + \frac{x}{\beta}\right)^{-\alpha} \quad (1)$$

onde α é o parâmetro de forma e β o parâmetro de localização, cujos valores determinam a forma da curva da distribuição.

Um dos métodos para verificação de que uma distribuição possui cauda pesada é a utilização do gráfico da distribuição complementar em escala logarítmica (*log-log complementary distribution*, *LLCD*). Este gráfico representa o valor da distribuição de probabilidade acumulada complementar $\bar{F}(x) = 1 - F(x)$ com $F(x) = \Pr\{X \leq x\}$, em escala logarítmica, de modo a procurar a invariância $d \log(\bar{F}(x))/d \log(x) = -\alpha$, $x > \theta$. Na prática, escolhe-se um valor de θ a partir do qual o gráfico passa a ser linear, caracterizando uma distribuição de cauda pesada.

TABELA I
CONFIGURAÇÃO DA REDE EM TESTE

Característica	Amostra 1	Amostra 2
Tempo de amostragem	$7,2 \times 10^3$ s	$7,2 \times 10^3$ s
Distância do AP	15 metros	10 metros
Padrão do prédio	Alvenaria	Alvenaria
Visada	Indireta	Direta
Padrão de rede	IEEE 802.11bg	IEEE 802.11bg
Canal	6	1
Pot. do sinal no AP	-52dBm	-57dBm
Relação sinal ruído no AP	34dB	40dB
Pot. do sinal no tx	-31dBm	-33dBm
Pot. do sinal no receptor	-45dBm	-47dBm
Ocupação média da rede	38%	34%
Picos de ocupação	85%	81%

III. DESCRIÇÃO DOS EXPERIMENTOS E METODOLOGIA

Para a realização das medições foram utilizados dois computadores conectados em uma rede sem fio, com diversos outros usuários utilizando esta mesma rede para o acesso à Internet. A rede onde foram efetuados os testes era composta de um AP (*Access Point*) instalado no Departamento de Engenharia Elétrica da UFPR. Os parâmetros da rede foram obtidos por medições realizadas com o testador de redes sem fio portátil AirCheck™ Wi-Fi Tester, da Fluke Networks.

Os demais estudos que buscavam modelar os erros em redes sem fio adotavam redes que continham apenas os dispositivos em teste [14], ou até mesmo uma rede ad-hoc [11]. Porém o objetivo deste estudo foi o de analisar a rede com o perfil de tráfego real, no modelo de infraestrutura, com diversos clientes conectados simultaneamente, de modo que as perdas de pacotes fossem reflexo do tráfego efetivo de uma rede WiFi. A Tabela I mostra as demais condições da rede em teste, incluindo as potências de sinais, relação sinal ruído e ocupação dos canais onde foram realizadas as coletas de dados.

Foi realizada a geração do fluxo de tráfego a partir do computador configurado como cliente, utilizando-se o *software* aberto *iperf* [15]. As mensagens consistiam de pacotes de 125 bytes transmitidas em uma taxa de 1000 pacotes por segundo, com a identificação de tempo (*time stamp*) e numeração de sequência. A taxa total do fluxo gerado foi de 1 Mbps, correspondendo a cerca de 2% da capacidade total da rede. O computador configurado como servidor recebia as mensagens, armazenando os dados dos pacotes recebidos e adicionando o tempo de chegada de cada pacote. Para o ajuste dos relógios do

transmissor e receptor foi utilizado o protocolo NTP (*Network Time Protocol*), de forma que o cliente e o servidor estavam sincronizados a partir de um servidor NTP (*ntp.pop-pr.rnp.br*) localizado na rede local, de forma a garantir que a diferença entre os relógios dos dois computadores tivesse um erro menor do que 1 microssegundo.

Para a primeira amostragem, foi realizada a transmissão de dados durante duas horas, sendo enviados mais de seis milhões de pacotes. Uma segunda amostragem foi realizada, com o mesmo tempo de duração, porém em outro dia e hora. Para identificar e contabilizar a perda de pacotes foi utilizado o número de sequência dos pacotes, contido no cabeçalho IP. Os pacotes não recebidos dentro do tempo máximo admitido pela aplicação foram considerados perdidos.

A. Agrupamento de pacotes (clusterização)

A partir dos dados recebidos foi necessário aplicar um método para realizar o agrupando dos pacotes perdidos, caracterizando os surtos de erros. O método escolhido foi considerar-se o início de um surto de erros uma sequência de três pacotes perdidos, passando então a rede para o estado ruim. Da mesma forma, foi considerado como retorno ao estado bom da rede a ocorrência de uma sequência de três pacotes recebidos sem perdas.

Assim, o critério adotado para a clusterização foi a distância entre os elementos de início e fim da cada transição de estado, entre *Bom* e *Ruim*, cujo valor deve ser maior ou igual a três. Este método é chamado de *complete linkage*, que considera a vizinhança mais distante possível para a formação dos clusters. O critério de validação do processo de clusterização aplicado foi o Critério Interno, com a definição do valor mínimo da distância entre os elementos de cada conjunto, neste caso com valor igual a três.

Aplicando-se este tratamento dos dados, foi obtido um novo arquivo, contendo os valores de tempo da ocorrência de uma mudança de estado e o estado para o qual houve a mudança. A partir deste novo conjunto de informações puderam ser gerados dois arquivos distintos, um deles contendo a duração dos estados *Bom* e outro a duração dos estados *Ruim* da rede. Estes arquivos foram utilizados para a análise estatística dos dados, visando obter-se o modelo de distribuição que efetivamente represente as perdas de pacotes em redes sem fio.

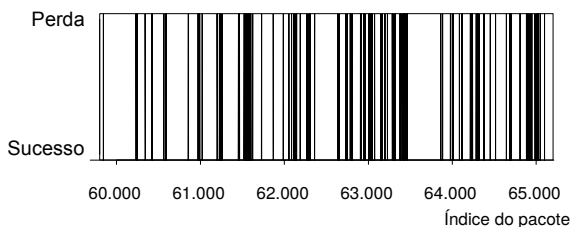


Fig. 2. Sucessão de pacotes transmitido com sucesso e perdidos

IV. RESULTADOS E MODELO PROPOSTO

A Figura 2 ilustra uma sequência de pacotes transmitidos com erro e com sucesso, durante um determinado intervalo de

tempo. Neste gráfico pode-se observar o comportamento das rajadas de tráfego com perdas, caracterizando o fenômeno de dependência temporal, com os surtos de erros apresentando-se claramente correlacionados.

A partir dos dados coletados foi realizada a análise estatística, com a utilização do *software* R [16], estudando-se o tempo de duração do estado *Bom*. Buscando identificar uma distribuição de cauda pesada, foi elaborado o gráfico LLCD, em escala logarítmica. A Figura 3(a) e 3(b) mostra respectivamente o gráfico LLCD para a primeira e segunda amostras. Como referência foi incluído o gráfico da LLCD para uma distribuição exponencial na Figura 3(c). Observa-se nas duas amostras um comportamento de cauda pesada, sendo que a segunda amostra possui um comportamento de cauda pesada mais acentuado que a primeira amostra. Caso o modelo de Gilbert-Elliott pudesse ser aplicado, o gráfico esperado para LLCD em escala logarítmica seria o apresentado na Figura 3(c).

A Figura 4(a) e 4(b) apresenta o gráfico da função de autocorrelação para a sequência representada pelo tempo de duração do estado *Bom*. Observa-se uma forte correlação entre os tempos. Como referência, foi incluído o gráfico da função de autocorrelação para uma distribuição exponencial na Figura 4(c), que seria o esperado caso o sistema pudesse ser modelado com uma cadeia de Markov conforme proposto por Gilbert-Elliott. No entanto, observa-se que o tráfego real o tempo de duração do estado *Bom* é fortemente correlacionado. O mesmo ocorre com o tempo de duração do estado *Ruim*. Isto significa que a ocorrência de um surto de erros aumenta a probabilidade de ocorrência de um surto no futuro próximo. Este comportamento *não* pode ser reproduzido pelo modelo de Gilbert-Elliott.

A. Um possível modelo

A partir dos resultados obtidos, observou-se que o modelo mais adequado é uma variação do modelo de Gilbert-Elliott, com um tempo de permanência em cada estado sendo modelado com uma distribuição de cauda pesada. Este modelo já foi estudado anteriormente no cenário de geração de tráfego em redes.

Para análise dos resultados, foi utilizado o gráfico *Quantile-Quantile* (QQPlot) [17], que é uma ferramenta gráfica para teste de aderência visual, com grande sensibilidade, utilizada para comparar características de duas populações. Nesta técnica, o conjunto de dados é ordenado em ordem de grandeza, sendo os valores que dividem o conjunto em N partes, denominados de *quantis*. Os pontos do gráfico representam os *quantis* de cada uma das amostras, colocados nos eixos x e y . Se as duas amostras possuem a mesma distribuição, os pontos devem estar em torno da linha diagonal em 45° sobre a origem. Foi empregado o *software* estatístico R para construção do gráfico.

Utilizando-se o método proposto por Hosking et al. [18], obtém-se uma estimativa para o parâmetro de forma de 0,59 e de 0,023 para o parâmetro de localização, para o tempo de duração do estado *Bom* da amostra 2. A Figura 5 mostra o QQPlot comparando a distribuição de Pareto com o tempo

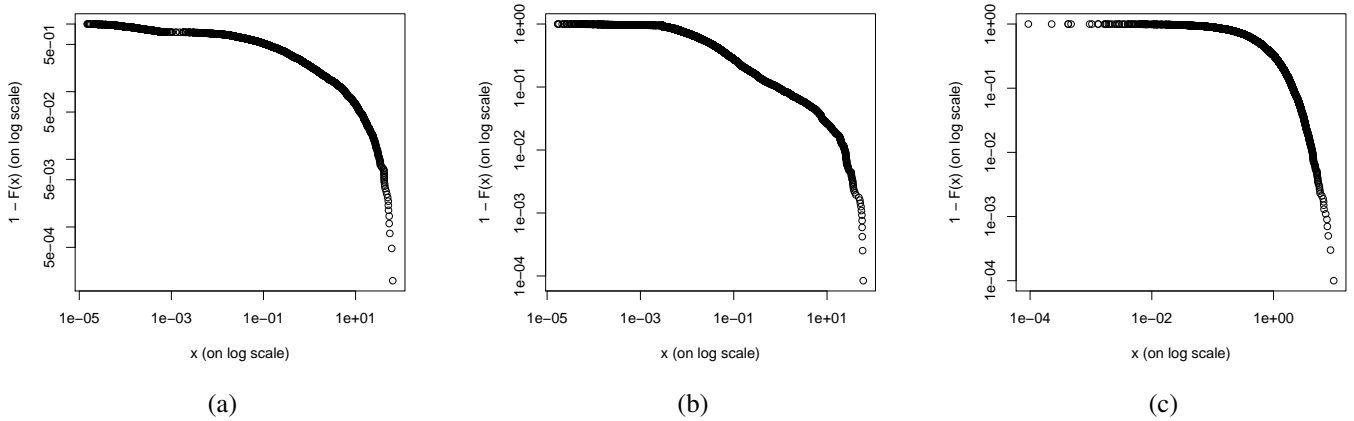


Fig. 3. Gráfico da LLCD para: (a) primeira amostra (b) segunda amostra e (c) distribuição exponencial

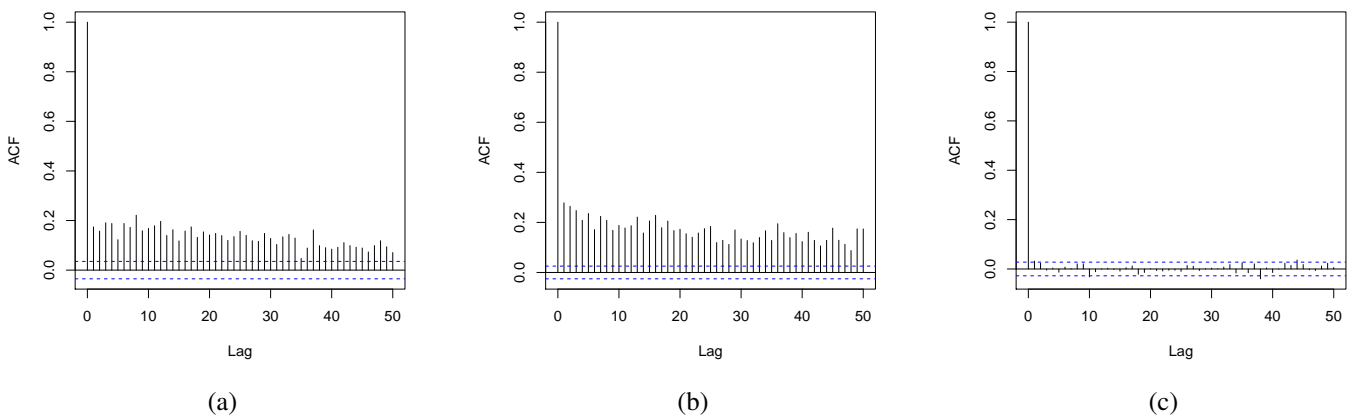


Fig. 4. Gráfico da função de autocorrelação para: (a) primeira amostra (b) segunda amostra e (c) distribuição exponencial

de duração do estado dessa amostra. É possível verificar a boa aderência apresentada, confirmando a hipótese que a distribuição de Pareto pode ser utilizada para modelar o tempo de duração dos estados.

A Figura 6 mostra o gráfico da distribuição de probabilidade acumulada do tempo de duração do estado *Bom* da amostra 2. Para comparação, foi inserida uma linha tracejada representando a distribuição teórica de Pareto, com os parâmetros ajustados conforme descrito anteriormente. Observa-se novamente a boa aderência entre os dados.

Testes semelhantes realizados com o tempo de duração do estado *Bom* da amostra 1 também mostram um comportamento de cauda pesada. Além disso, os mesmos testes realizados com o tempo de duração do estado *Ruim* também apresentam esta característica.

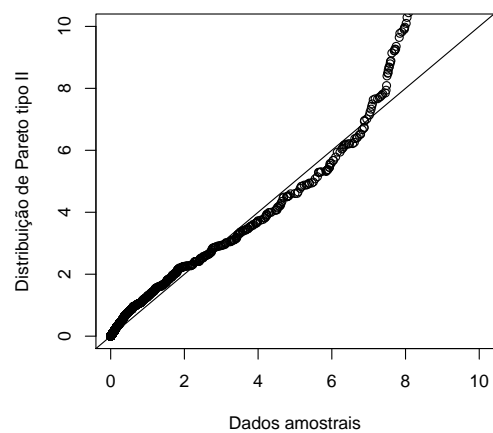


Fig. 5. QQPlot comparando o tempo de duração do estado *Bom* da amostra 2 com a distribuição de Pareto

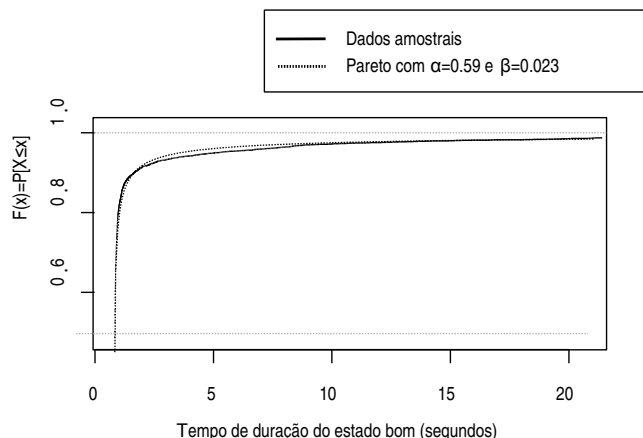


Fig. 6. Distribuição acumulada do tempo de duração do estado *Bom* da amostra 2 com a distribuição de Pareto

V. CONCLUSÕES E TRABALHOS FUTUROS

Neste artigo foram apresentadas evidências da falha do modelo de perda de pacotes de Gilbert-Elliott, quando aplicado a redes Wifi. Foram coletadas amostras de tráfego na rede Wifi do Departamento de Engenharia Elétrica da Universidade Federal do Paraná.

Após a análise estatística dos dados foi possível observar que o tempo de duração dos estados *Bom* e *Ruim* apresentou comportamento de cauda pesada, significando ser este mais adequado do que o modelo de Gilbert-Elliott.

Para aprimorar este estudo, e aperfeiçoar o modelo proposto, os autores estão desenvolvendo novas ações em duas frentes, que são (1) aplicação de algoritmos de clusterização que sejam mais adequados para agrupar os surtos de perda, realizando novas análises estatísticas e (2) realização de medições em outros ambientes de rede, incluindo o padrão 802.11n e redes LTE.

REFERÊNCIAS

- [1] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 2347–2376, Fourthquarter 2015.
- [2] G. Hasslinger and O. Hohlfeld, "The gilbert-elliott model for packet loss in real time services on the internet," in *Measuring, Modelling and Evaluation of Computer and Communication Systems (MMB), 2008 14th GI/ITG Conference*, March 2008, pp. 1–15.
- [3] Y. Qi, P. Cheng, and J. Chen, "Dynamic sensor data scheduling for remote estimation over gilbert-elliott channel," in *2014 IEEE/CIC International Conference on Communications in China (ICCC)*, Oct 2014, pp. 26–30.
- [4] E. N. Gilbert, "Capacity of a burst-noise channel," *The Bell System Technical Journal*, vol. 39, no. 5, pp. 1253–1265, Sept 1960.
- [5] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *The Bell System Technical Journal*, vol. 42, no. 5, pp. 1977–1997, Sept 1963.
- [6] G. Hasslinger and O. Hohlfeld, "Analysis of packet errors in gilbert-elliott channels," in *2011 IEEE 12th International Workshop on Signal Processing Advances in Wireless Communications*, June 2011, pp. 216–220.
- [7] ITU-T, "Network model for evaluating multimedia transmission performance over internet protocol," International Telecommunication Union, Geneva, Recommendation G.1050, Nov. 2007.
- [8] C. Grinstead and J. Snell, *Grinstead and Snell's Introduction to Probability*, ser. Titolo collana. University Press of Florida, 2009. [Online]. Available: <https://books.google.com.br/books>
- [9] X. Yu, J. W. Modestino, and X. Tian, "The accuracy of gilbert models in predicting packet-loss statistics for a single-multiplexer network model," in *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 4, March 2005, pp. 2602–2612 vol. 4.
- [10] L. Carvalho, J. Angeja, and A. Navarro, "A new packet loss model of the ieee 802.11g wireless network for multimedia communications," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 3, pp. 809–814, Aug 2005.
- [11] K. V. Cardoso and J. F. D. Rezende, "Accurate hidden markov modeling of packet losses in indoor 802.11 networks," *IEEE Communications Letters*, vol. 13, no. 6, pp. 417–419, June 2009.
- [12] K. Wolter, P. Reinecke, T. Krauss, D. Happ, and F. Eitel, "Ph-distributed fault models for mobile communication," in *Proceedings of the 2012 Winter Simulation Conference (WSC)*, Dec 2012, pp. 1–12.
- [13] C. I. Mattos, E. P. Ribeiro, E. Fernandez, and C. M. Pedroso, "An unified voip model for workload generation," *Multimedia tools and applications*, vol. 70, no. 3, pp. 2309–2329, 2014.
- [14] S. H. Russ and S. Haghani, "802.11g packet-loss behavior at high sustained bit rates in the home," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 2, pp. 788–791, May 2009.
- [15] J. Dugan, S. Elliott, B. A. Mah, J. Poskanzer, K. Prabhu, M. Ashley, A. Brown, A. Jaii, S. Sahani, B. Simpson, and B. Tierney, "iperf - the network bandwidth measurement tool," 2016, on Line. [Online]. Available: <https://iperf.fr/>
- [16] B. Laboratories, "The R project for statistical computing." 2016. [Online]. Available: <https://www.r-project.org/>
- [17] Z. M. Durovic and B. D. Kovacevic, "QQ-plot approach to robust Kalman filtering," *International Journal of Control*, vol. 61, no. 4, pp. 837–857, 1995.
- [18] J. R. M. Hosking and J. F. Wallis, "Parameter and quantile estimation for the generalized pareto distribution," *Technometrics*, vol. 29, no. 3, pp. 339–349, 1987.