

可纠正插入/删除错误的 Marker-Polar 级联码研究

白羽, 柳元

(天津师范大学电子与通信工程学院, 天津 300387)

摘要: 为进一步提升 Marker-LDPC 级联码纠正插入/删除错误的能力, 本文选用具有更强纠错能力和更低错误平层特性的 Polar 码作为外码, 提出 Marker-Polar 级联方案。外译码器采用列表连续消除译码算法纠正内译码器输出的残留错误, 并结合循环冗余校验和奇偶校验的双重机制提升系统的纠错性能。仿真结果表明, 在误帧率为 0.01, 级联码长为 768 比特的条件下, 所提方案较传统 Marker-LDPC 结构每帧可多纠正 6 个插入/删除错误, 纠错性能有明显的提升。

关键词: 通信工程; 插入/删除; 标记码; Polar 码; 级联码

中图分类号: TN911

Research on Marker-Polar Concatenated Codes for Correcting Insertion/Deletion Errors

BAI Yu, LIU Yuan

(Electronics and communication engineering College, Tianjin Normal University, Tianjin 300387)

Abstract: To enhance the error correction capability of Marker-LDPC concatenated codes against insertion/deletion errors, this paper proposes a Marker-Polar concatenated scheme by employing Polar codes as outer codes, which exhibit superior error correction performance and lower error floor characteristics. The outer decoder adopts the successive cancellation list (SCL) decoding algorithm to correct residual errors from the inner decoder, while incorporating a dual verification mechanism combining cyclic redundancy check (CRC) and parity check (PC) to improve the system's error correction performance. Simulation results demonstrate that at a frame error rate of 0.01 with a total code length of 768 bits, the proposed scheme can correct 6 more insertion/deletion errors per frame compared to conventional Marker-LDPC structures, showing significant performance improvement.

Key words: Communication engineering; Insertions/deletions; Marker code; Polar codes; Concatenated coding scheme;

0 引言

在信息传输的过程中, 接收机的采样时钟不稳定, 会导致插入或删除错误^[1]。这种错误

作者简介: 白羽 (1999-), 女, 主要研究方向: 无线通信系统与理论

通信联系人: 柳元 (1988-), 女, 副教授, 硕导, 主要研究方向: 无线通信系统与理论、同步理论及应用, 高效编码方案的设计及验证. E-mail: liuyuan@tjnu.edu.cn

在存储系统、DNA 测序及无线通信等场景中具有显著的破坏性。特别的，存在插入/删除错误的系统具有记忆性。单个的插入/删除错误会引发突发性的替代错误。针对插入/删除-替代（Insertion/Deletion-Substitution, IDS）信道带来的错误问题，级联码方案通过内外码的协同工作机制可实现有效纠错。文[2]将标记码作为内码方案，在 LDPC 码的编码序列中以固定间隔周期性插入已知标记模式来建立同步参考。该方案采用基于加权莱文史特距离（Weighted Levenshtein Distance, WLD）的符号级前向-后向译码算法，通过动态规划法有效计算译码网格中的分支度量，不仅克服了传统方法受限于内码结构的缺陷，还能更准确纠正接收序列中的插入/删除错误。外译码层负责处理经内译码校正后残余的随机错误[2]。然而，现有方案中外码多采用二进制 LDPC 码，其在插入/删除信道中的纠错效能有限，导致高误码率环境下系统整体性能下降[3-4]。

为进一步提升级联码在 IDS 信道中的纠错能力，本文提出一种 Marker-Polar 级联码方案。该方案采用 Polar 码作为级联码框架的外码，基于 Polar 码在串行消除译码框架下展现出的优异纠错性能[5]，能够有效处理经过内译码校正后残余的随机错误。进一步，选用 Polar 码的列表连续消除（Successive Cancellation List, SCL）译码算法[6]，作为外译码器的译码算法。同时提出一种基于循环冗余校验（Cyclic Redundancy Check, CRC）和奇偶校验（Parity-Check, PC）的辅助译码机制[7]，通过在信息比特中加入 CRC 和 PC 校验比特，能够有效剔除 SCL 译码过程中的错误路径，降低错误概率[7]。仿真结果表明，所提方案在插入/删除信道下的纠错性能优于传统级联码方案，尤其是在外码采用 SCL 译码算法并结合 CRC 和 PC 校验时，较传统方案有明显的提升[8]。该方案为同步敏感场景下的通信系统提供了一种可靠的错误控制解决方案，具有潜在的工程应用价值。

1 级联码的编译码结构

级联码将检测插入/删除错误的码与传统纠错码相结合，以实现插入/删除的检测定位和纠正。级联码的编译码算法包括内译码算法和外译码算法。内译码过程通常用于纠正插入、删除错误，外译码过程用于进一步处理内译码无法纠正的残留替代错误。

1.1 系统模型

如下图 1 中所示的方案使用已知的标记码作为内部编码器，并使用二进制 Polar 码作为外部代码。二进制信息 \underline{b} 首先在尾部加入 PC 奇偶校验比特和 CRC 码，然后进入外部 Polar 码编码，并被映射到二进制代码 \underline{d} (N_L, K_L) 中。然后将已知标记码 \underline{w} 均匀地插入到 \underline{d} 中，产生长度为 N_c 的传输码 \underline{x} 。 \underline{x} 通过随机二进制插入/删除-替换通道[9]发送。发送码字 \underline{x} 经插入/删除-替代信道产生长度为 N_c^* 的接收序列 \underline{y} ，之后基于 WLD 的标记码译码器，纠正接收序列 \underline{y} 中的错误，输出似然比序列 \underline{l} ；利用 \underline{l} 初始化二进制 Polar 译码器，采用列表连续消除译码算法进行译码，并进行 PC 奇偶校验以及 CRC 校验，进一步提高译码的准确性，输出最后的信息序列估计值 $\hat{\underline{b}}$ 。

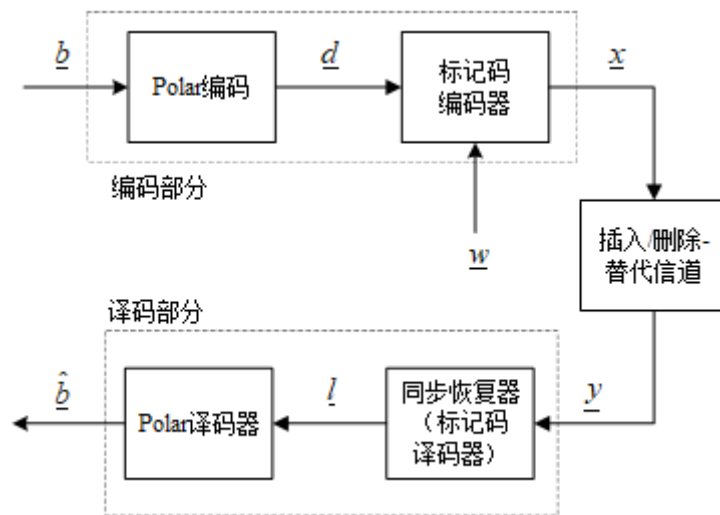


图 1 级联码系统框图

Fig. 1 Schematic of a Concatenated Coding System

IDS 信道的模型如图 2 所示，对于序列中的每一位，可能会发生四种情况。具体地说，将随机位插入到 P_i 序列中，或者用 P_d 删除发送位，或者用 P_s （模 2）添加到位“1”，或者正确发送位，其中参数 P_i 、 P_d 和 P_s 分别表示插入、删除、替换概率。传输概率 $P_t = 1 - P_i - P_d$ 。对于每个区块，信道产生 N_i 插入错误， N_d 删除错误， N_s 替代错误的概率计算公式如下：

$$P(N_i, N_d, N_s) = P_i^{N_i} \cdot P_d^{N_d} (P_t P_s)^{N_s} (P_t (1 - P_s))^{N_c - N_d - N_s} \quad (1)$$

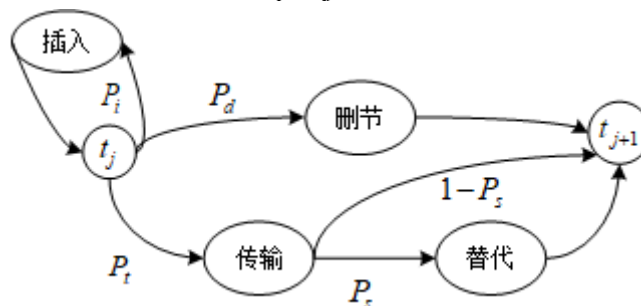


图 2 比特 IDS 信道

Fig. 2 Bit-IDS Channel

1.2 编码方法

外编码器由于其优异的纠错性能，采用了二进制 Polar 编码方法^[5]。首先将 Polar 码编码后的 d 序列划分为 (N_L/m) 个符号，每个符号都有 m 个位。取符号满足的值 $q = 2^m$ 。本文选择伪随机序列作为标记码。内部编码器将标记码的 λ 位分配给 d 的每个子序列，以生成 x 。

1.3 译码方法

本文采用基于 WLD 的前向后向译码方法作为内译码方案^[2]。WLD 通过构建动态规划矩阵计算序列间的编辑距离，利用加权因子（如插入、删除和替换的权重）量化编辑操作成本，从而更准确地反映插入/删除错误对序列的影响。具体而言，通过最小化加权编辑距离来优化序列匹配，基于 WLD 的内译码算法能够为外译码算法提供初步纠正后的软信息。外译码器利用该软信息进一步纠正残留错误。

传统的连续消除（Successive Cancellation, SC）译码算法虽然实现简单，但在插入/删除

信道中性能受限^[5]。为提升译码性能，本文采用 SCL 译码算法^[6]，通过维护一个大小为 L 的路径列表，保留多条候选路径以减少错误传播。此外，本文在 SCL 译码中引入 CRC 校验和 PC 校验^[7]，通过校验机制辅助路径修剪，剔除错误路径，从而进一步提高译码的准确性。

2 仿真结果分析

仿真中采用伪随机序列作为内部标记码，其中 $\lambda=1$ ， $m=2$ ， $I_{\max}=5$ ， $P_i=P_d$ 。表 1 为仿真参数设置。图 3 是不同外码设计对整体纠错性能的影响。

表 1 仿真基本参数设置表
Tab. 1 Simulation Parameter Settings

参数	Marker-LDPC	Marker-Polar	CRC-Marker-Polar	PC-Marker-Polar	PC-CRC-Marker-Polar
码长 N_c (bits)	864	768	768	768	768
外码码长 N_L (bits)	576	512	512	512	512
信息位 K_L (bits)	288	256	256	256	256
PC 校验位数 r_1 (bits)	-	0	0	8	8
CRC 校验位 r_2 (bits)	-	0	24	0	24
译码列表长度 L	-	4	4	4	4
替代概率 P_s	0.001	0.001	0.001	0.001	0.001
码率	0.33	0.33	0.3	0.32	0.29

图 3 展示了不同外码设计方案对级联码性能的改进效果。仿真结果表明，本文提出的 Marker-Polar 级联方案在纠错性能上优于传统 Marker-LDPC 方案。通过引入 PC 校验和 CRC 校验的联合设计，系统误帧率进一步降低。具体来看，在误帧率为 10^{-2} 时：采用 768 比特码长的基础 Marker-Polar 方案每帧可纠正 44 个插入和删除错误；加入 PC 校验后可纠正 46 个错误；加入 CRC 校验后可纠正 50 个错误；CRC-PC 联合校验方案同样可纠正 50 个错误。相比之下，采用 864 比特码长的传统 Marker-LDPC 方案只能纠正 44 个错误。这表明本文提出的方案比传统 Marker-LDPC 方案可多纠正 6 个插入和删除错误。

从校验过程来看，在 SCL 译码过程中，24 位 CRC 校验作为核心路径筛选模块，通过对候选译码路径进行全局校验来确保译码可靠性。当某条路径的 CRC 校验结果与编码端预设值不匹配时，该路径被剔除。仿真结果显示，CRC 校验的引入使系统在插入删除概率为 0.03 时的误帧率从 10^{-2} 量级降低至 10^{-3} 量级，纠错能力有明显的提升，充分发挥其在错误路径筛选中起到的作用。8 位 PC 校验则作为辅助筛选机制，主要针对 CRC 校验可能漏检的连续突发错误。当多条候选路径同时通过 CRC 校验时，PC 校验通过检测各路径的奇偶一致性进行二次筛选。值得注意的是，插入删除概率为 0.025 以及 0.04-0.06 之间时 CRC-PC 校验的方案相比仅使用 CRC 校验的方案展现出更优的性能曲线，这表明 PC 校验能够填补 CRC 在某些信道条件下的检测盲区。CRC-PC 的联合校验使系统在保持原有可靠性的同时，增强了对特定错误的抵抗能力。

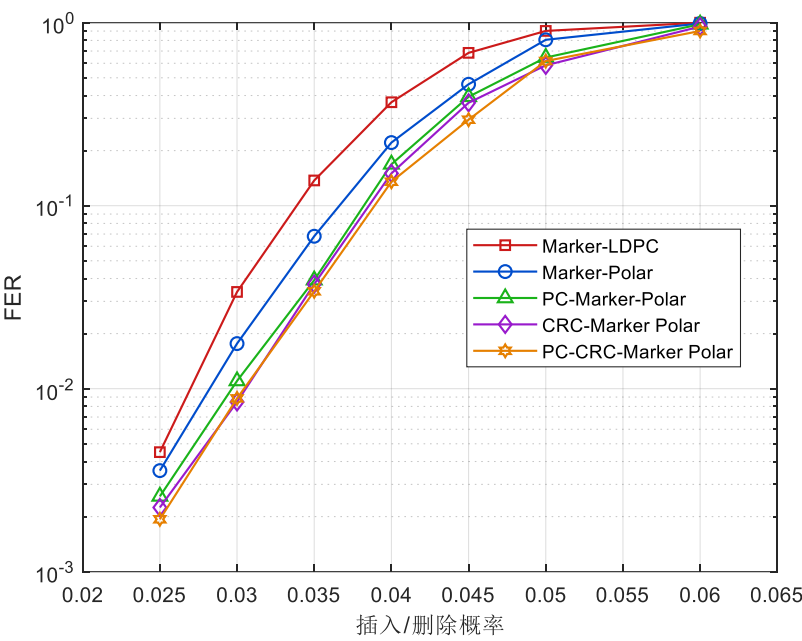
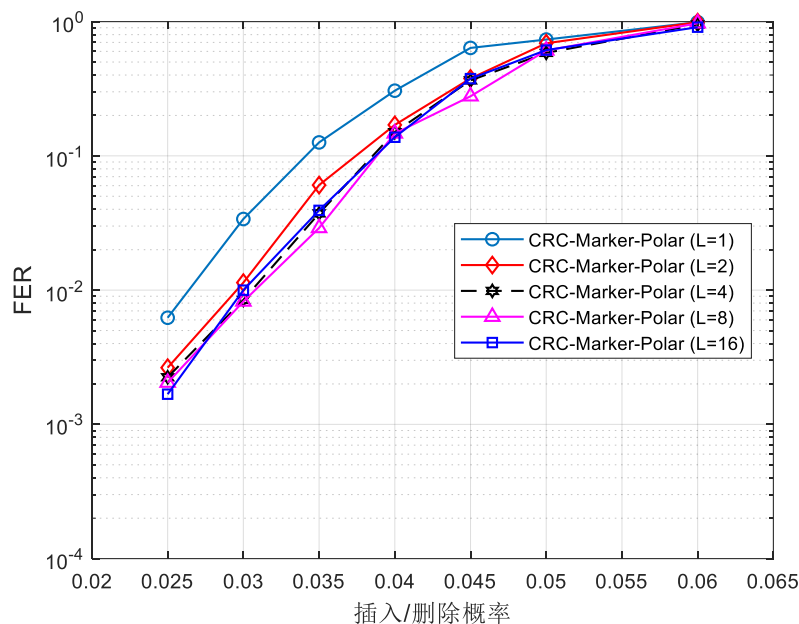


图 3 不同外码设计方案对级联码性能的影响
Fig. 3 Impact of Different Outer Code Designs on Concatenated Code Performance

图 4 研究了外码码长为 512 比特时，CRC-Marker-Polar 级联码在不同列表长度 L 下的纠错性能对比。基于表 1 的数据，只改变 CRC-Marker-Polar 组的列表长度 L ，其他数据保持不变。仿真结果表明，CRC 校验通过剪除错误路径显著降低了无效路径的保留概率，这一机制使得即使采用较小的列表长度例如 $L=4$ 时，系统性能也能接近较大列表长度的情况。具体来说，当 $L=1$ 即 SC 译码时，此时仅保留一条路径，没有路径扩展和选择的过程，误帧率曲线最高，性能最差。而 $L=2$ 相比 $L=1$ 性能明显改善，但由于只能保留 2 条路径，性能提升有限。进一步，增大至 $L=4$ 时性能继续有明显提升，当列表长度增大到 $L=8$ 时，在 0.03 至 0.05 的插入/删除概率区间内，误帧率曲线出现小幅下降，性能略有提升。然而继续增大到 $L=16$ 时，误帧率曲线与 $L=4$ 基本重合，表明此时增大列表长度带来的性能改善已十分有限。综合考虑性能和外译码复杂度如表 2 所示，最终选择 $L=4$ 作为最优方案，其复杂度比 $L=8$ 时更低，性能和 $L=8$ 时比较相差不大。此外，系统采用的 24 位 CRC 校验会引入额外的计算量，每帧需要 24 次多项式模 2 除法运算，但由于这部分开销是固定的，对整体复杂度影响较小。这说明 $L=4$ 在性能和复杂度之间实现了较好的平衡。

表 2 SCL 译码复杂度分析
Tab. 2 Complexity Analysis of SCL Decoding

列表长度 L	1	2	4	8	16
外码时间复杂度	$O(M\log N)$	$O(2M\log N)$	$O(4M\log N)$	$O(8M\log N)$	$O(16M\log N)$
外码空间复杂度	$O(N)$	$O(2N)$	$O(4N)$	$O(8N)$	$O(16N)$

图 4 不同列表长度 L 对级联方案的性能影响Fig. 4 Impact of List Size L on Concatenated Scheme Performance

145

图 5 研究了 Marker-Polar 级联码在保持 CRC 校验提供的稳定路径筛选基础上, 改变不同码长和码率对级联方案纠错性能的影响。基于表 1 的仿真数据, 通过调整 CRC-Marker-Polar 组的码长和码率参数可得, 随着码长增加和码率降低, 级联方案的误帧率呈现明显的下降趋势, 其中在码率为 0.19 时误帧率曲线最低, 纠错性能最优。具体来说, 当总码长达到 3072 比特, 码率为 0.32 时, 在误帧率为 10^{-2} 的条件下, 该方案每帧可同时纠 107 个插入/删除错误和 3 个替代错误。实验数据进一步表明, 较长的码长通过增加冗余提升了纠错能力, 而适当的码率配置在不显著牺牲编码效率的前提下, 实现了纠错性能的最优化。

150

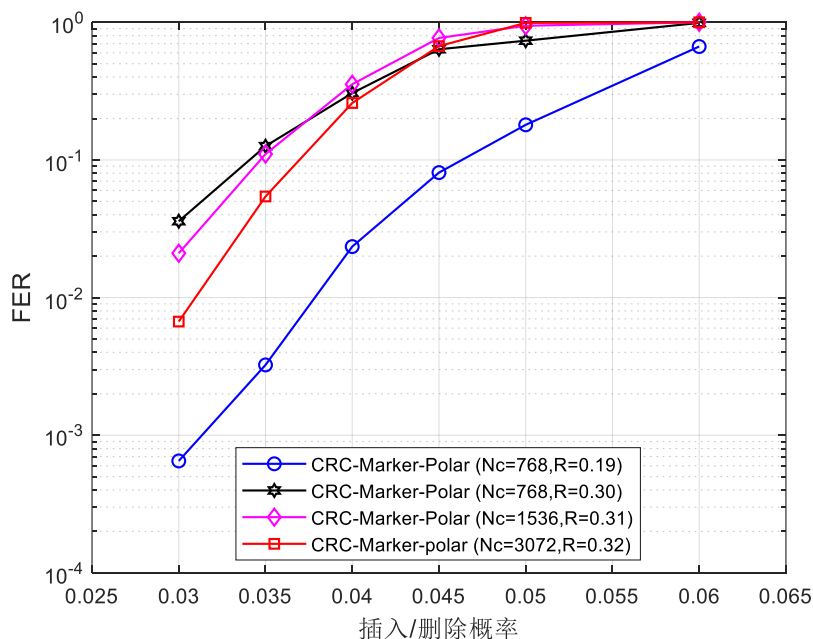


图 5 不同码长与码率对级联方案的性能影响

Fig. 5 Impact of Code Length and Rate on Concatenated Scheme Performance

155

图6研究在PC-CRC-Marker-Polar级联方案的基础上,分别比较PC校验个数为2比特、4比特、8比特、12比特和16比特时对误帧率的影响。本组仿真数据固定替代概率为0.002,总码长为768,CRC校验个数为24位,SCL译码列表长度为4,其中码率随着不同的PC校验个数会有一些改变,但码率相近,对整体影响结果较小。仿真结果表明,PC校验位数的选择直接影响路径修剪效果:当采用2或4比特校验时,系统存在错误路径剔除不足的问题;而12或16比特校验则可能导致过度修剪,增加正确路径被误删的风险。在外码Polar码长为512比特的条件下,8比特PC校验方案表现出最优的平衡性,既能有效筛选错误路径,又可避免过度修剪。这说明在实际通信系统中PC校验位数的选择需要综合考虑码长和信道条件等因素。

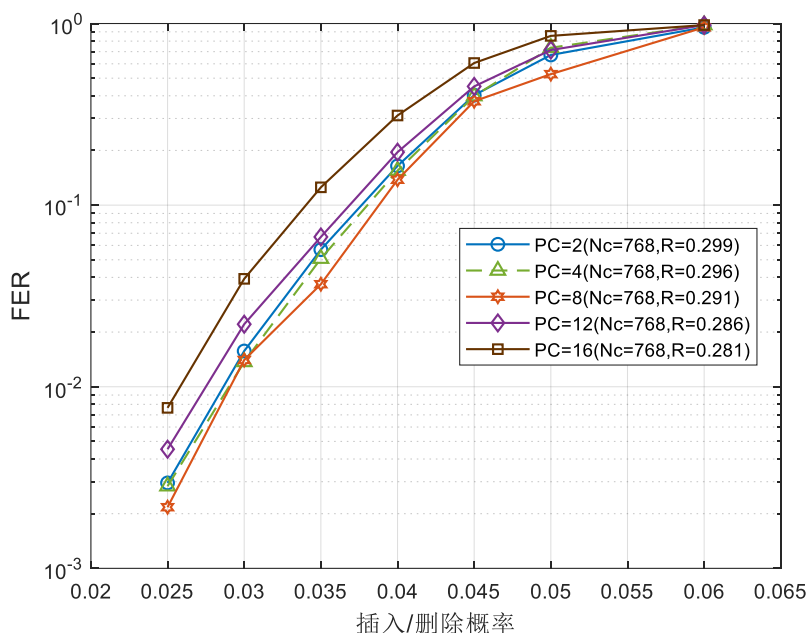


图6 不同PC校验个数对级联方案的性能影响

Fig. 6 Impact of PC Check Count on Concatenated Scheme Performance

3 结论

本文提出一种Marker-Polar级联码方案,用于纠正插入/删除和替代错误。仿真结果表明,外译码采用Polar码SCL译码算法时,其纠错性能提升明显。进一步,通过引入多级校验机制,包括CRC校验和可配置的PC奇偶校验,有效优化了译码路径选择过程,显著降低了系统误帧率。本研究为高可靠通信系统,特别是对同步错误敏感的应用场景,提供了一种高效的纠错编译码方案。未来研究工作可着重于算法复杂度优化、硬件实现验证。

[参考文献] (References)

- [1] Cheng L. Coding techniques for insertion/deletion error correction [D]. University of Johannesburg, 2010.
- [2] Liu Y, Lu Y W, He Y S, et al. Marker codes using the decoding based on weighted Levenshtein distance in the presence of insertions/deletions [J]. IEEE Access, 2020, 8: 206003-206008.
- [3] Gallager R. Low-density parity-check codes [J]. IRE Transactions on information Theory, 1962, 8(1): 21-28.
- [4] Richardson T, Urbanke R. The capacity of low-density parity-check codes under message-passing

decoding[J].IEEE Transactions on Information Theory,2001,47(2):599-618.

[5] Arikan E. Channel polarization:A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels[J].IEEE Transactions on Information Theory,2009,55(7):3051-3073.

185

[6] Tal I, Vardy A. List decoding of polar codes[J].IEEE Transactions on information Theory,2015,61(5):2213-2226.

[7] Niu K, Chen K. CRC-aided decoding of polar codes[J].IEEE Communications Letters,2012,16(10):1668-1671.

[8] Hagenauer J. Rate-compatible punctured convolutional codes(RCPC codes) and their applications[J].IEEE Transactions on Communications,1988,36(4):389-400.

190

[9] Davey M C, Mackay D J C. Reliable communication over channels with insertions,deletions,and substitutions [J].IEEE Transactions on Information Theory,2001,47(2): 687-698.