ON THE COMPARISON OF NEW METHODS OF NOISEPROOF CODING

(Quick report translation at conference 2016 in RSREU, Riazan, Russia)

Doctor of Technical Sciences, prof. Zolotarev V.V., Doctor of Technical Sciences, prof. Ovechkin G.V.
Institute of Space Research of the Russian Academy of Sciences
Ryazan State Radio Engineering University

The efficiency of known error correction methods for block codes in binary Gaussian channels, in channels with erasures, and in q-point symmetric channels is analyzed. Special attention is paid to the methods of decoding polar codes, which are widely discussed in publications, as well as low-density (LDPC) codes, multithreshold decoders (MTD) of self-orthogonal codes and block modification of the decoding algorithm Viterbi for short convolutional codes. Along with the probability of error per block, their computational complexity is estimated. The analysis of a number of publications has shown that the efficiency of basic methods for decoding polar codes with a length of up to several thousand bits is quite far from the theoretical limits and is comparable to the efficiency of MTD, which have significantly less computational complexity. The best list methods of decoding polar codes are effective, but their complexity is proportional to the size of the list, which significantly complicates their use in high-speed communication systems. The complexity of polar code decoders is many times greater when they are used to decode Reed-Solomon codes. At the same time, the available results on MTD for symbolic codes that have a linear implementation complexity depending on the code length show that they can be used to provide higher levels of data transmission and storage reliability.

The rapid development of the decoding technique, its noticeable improvement in various parameters, requires regular analysis of the current capabilities of known and new error correction algorithms. The possibilities of the Viterbi decoding algorithm (VA) and the multithreshold decoder (MTD), which are well known to specialists, in block modification, in comparison with the algorithms for polar codes (PC), are considered below [1, 2, 6].

The currently developed special presentation of the material on the PC, when instead of the bit error probabilities $P_b(e)$, the error probabilities per WER block are estimated, which much less familiar in publications on decoding techniques, it creates a lot of difficulties for readers when comparing different algorithms. Similarly , the performance characteristics of the PC are somewhat conditional, which are often presented without any comments as numerical results. The report provides a comparison of different decoding algorithms in the binary Gaussian channel for block codes, which will allow a relatively realistic assessment of the capabilities of the main most well-studied codes for the WER parameter. Apparently, the most problematic issue at the moment is the complexity of decoding for polar codes. It requires specific refinement, for example, using data for the performance of algorithms for them on standard processors for personal computers. In addition, for a very long time, polar codes have been presented very inconsistently by various authors, as, for example, in [4],: "Polar codes are the first class of codes that achieve channel capacity with a code length tending to infinity, having the complexity of encoding and decoding O(nlogn), where n is the length of the code. However, the correction ability of polar codes with practically significant lengths is significantly worse than, for example, in the case of codes with a low density of parity checks (LDPC)." Note that this recommendation of polar codes makes them seem attractive in terms of complexity, but questions immediately arise about the reasons for low efficiency, even in comparison with low-density codes. Note also that the proposed complexity estimates are usually independent of the noise level in any way channels, and even turning out to be of the same order for fundamentally different work the purpose of the encoding and decoding procedures looks like a very problematic issue . Another feature of the PC is the lack of meaningful results on the use of convolutional codes by these methods. The above contradictory comments on the properties of the PC could be compensated for by using convolutional codes, for which the exponent of reliability is significantly higher at a high noise level than for block codes. This could greatly improve the efficiency of the PC.

Below we compare the capabilities of a number of known codes and algorithms for decoding them in terms of efficiency and, so far, very conditionally, in terms of complexity, which will help specialists better navigate the realities of known and new error correction methods. Of course, more accurate estimates and comparative characteristics of algorithms and codes of the PC class are still in the future, although quite soon we will celebrate a decade of development in this direction. Figure 1 shows the dependencies of the error probability per word (per block) WER for different binary block codes as a function of the noise level of the Gaussian channel (AWGN) E_b/N_0. The S curve is an estimate for the efficiency of the best possible block codes of length n=1024 with a code rate R=1/2. It indicates the potential of codes of such a relatively small length that can, in principle, be achieved. Ratings received by the spherical packing method, as a result of which the real WER codes are always significantly weaker. This is what all the methods whose characteristics are discussed below demonstrate further for R=1/2.

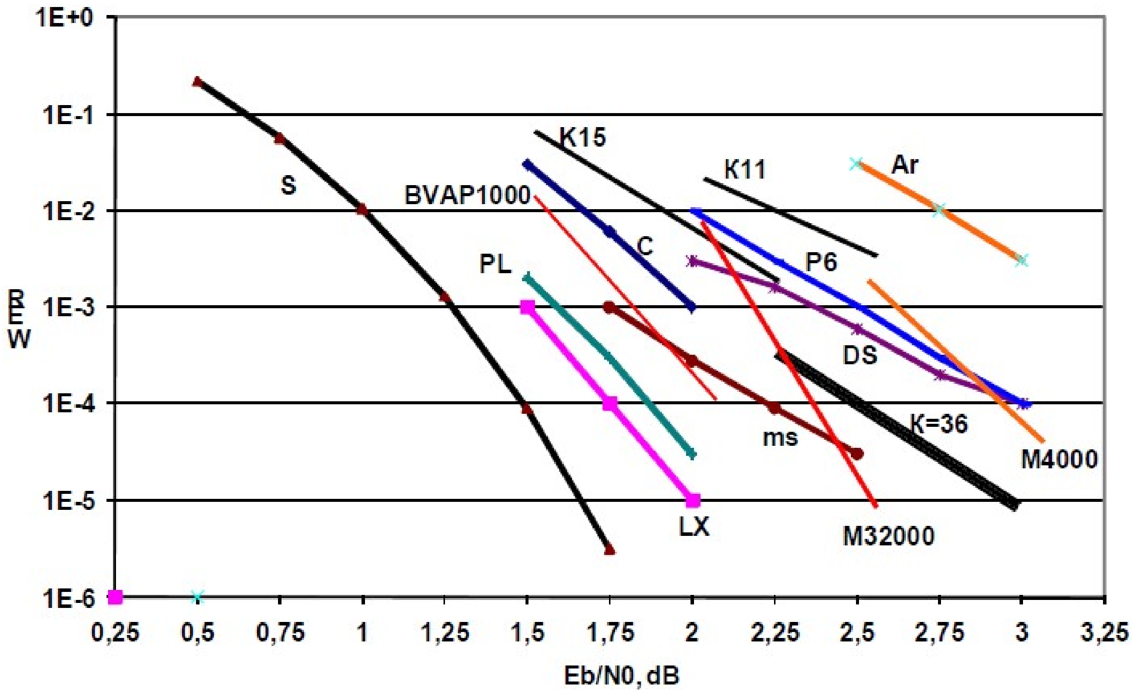

Fig. 1. Characteristics of modern error correction methods

Graphs C, PL, ms, P6, and LX represent the probabilities WER for PC of length n=1024, moreover, for PL, the size of the solution list is L=128, which is quite a lot, and for LX, it is even L=2048, which really means that for the small values of n discussed , the decoding result is very uncertain [6]. The remaining curves refer to length codes n~2048 or close to them. Next to polar codes: Ar chart according to [5] refers to [1], the curve DS obtained in [5] according to [3] for L=32, the lower limit of ms corresponds to several methods of [6] with a list of up to L=32, the boundary relates to the four the methods of [4] with a list of up to L=32, using part of cascading PC up to nine components codes. Four more results from [2] can be bounded from the bottom of the P6 curve.

We will indicate other well-studied decoding methods that can be applied without any special adaptation to the conditions of comparison with the characteristics of polar codes. The results of modeling their operation in a channel with AWGN: K15-the result for a block AB [12] with a generating polynomial of the convolutional code of length K=15 with a block code length of n=400. At the same time, the decoding speed of the block AB is 1000 bits/s on a laptop with a Core-i7 processor, for which the following data will be provided performance characteristics for other algorithms as well. Note that this is a very simple decoder, which was created back in 1990 for the Cassini project (flight to Saturn). Next, K11 is a block AB with a block length of n=200 for K=11. The decoding speed under the same conditions is 15 Kbit/s. This is a very simple AB. Finally, we pay attention to the line K36, which refers to (the only one on the graph!) the probability of error per bit Pb (e) for a convolutional AB with K=36 with viewing only N=64 The BVAP1000 graph corresponds to the estimate for a block AB with an incomplete view of the paths. Its performance is still estimated at about 3000 bps.

Figure 1 shows the possibilities of MTD methods [7-9, 14]. For a code with n=4000 bits, an M4000 graph with a performance of 230 Kbit/s is given, and for a code with n=32000, an M32000 graph is given for a decoder with a performance of 110 Kbit/s. The codes for the MTD were selected from the available databases of binary codes without any special adaptation or special choice and for them the standard ordinary MTD algorithm was implemented.

Let us consider further the relation of the general properties of the decoding algorithms listed above . To begin with, the situation with polar codes (PC) is very similar to the history of the development of methods for sequential decoding of convolutional codes in the 70s of the last century. As is known, the main modifications of algorithms in this direction do not work in the Gaussian channel at a noise level greater than the computational speed of the channel R_1, when, for example, at R=1/2, this ratio is E_b/N_0=2.5 dB, which is 2.3 dB worse in energy than at the capacity of the channel C=1/2 [7]. At the same time, a huge number of articles in those years analyzed the statistical parameters of the distribution of the number of operations of the algorithm, but the real characteristics of reliability, of course, which were of the main interest, were difficult to find anywhere for a long time. Various modifications of the sequential algorithms also failed to eliminate their shortcomings in the future. About the same was the situation with the BCH codes, about which it was known that their capabilities were even more limited. These codes, for example, at R=1/2, could be used mainly in the usual BSC with an error probability in the channel of no more than 0.037, which corresponded to the level of E_b/N_0>5 dB, and their use in the Gaussian channel was a complex problem that could not be solved even when other methods, such as Chase, were used together with them (see [7]).

The development of algorithms for the PC is also not without similar features and omissions. Indeed, the low performance characteristics of the PC almost immediately with the advent of this new direction in coding theory were "improved" in such a way that, similar to the already well-developed methods for Reed-Solomon codes (RS), the solution of the PC decoders began to consider a certain list of possible messages of size L, from which a message is selected as the decoding result, for example, with a running CRC. If the true solution is included in such a list, the decoding is considered successful. Of course, the size of L lists can be tens or even thousands of message variants, which greatly reduces the probability of error for this convenient decision-making method, which is very suitable for theorists. We emphasize that the complexity of such a modified algorithm increases in this case also by L times, which dramatically worsens the already low technological capabilities of the PC (compare the simplest encoding and decoding schemes for quite effective algorithms considered in [7]). With other approaches to PC decoding, their initially low initial performance is improved by other well-known powerful means of increasing the efficiency of encoding and decoding. These include conventional concatenated methods, systems with generalized concatenated codes [10], as well as a whole range of sequential decoding methods adapted to the PC needs. At the same time, it is possible to build a PC with higher values of the minimum code distance d than the original non-concatenated methods, and the use of iterative approaches to their decoding leads to the possibility of working at a slightly higher noise level. At the same time, the high heterogeneity of calculations in the decoders of such PCs significantly complicates the already ideologically quite difficult means of encoding/decoding in complex schemes of such processing (see curve C for the concatenated code [4] with nine (!) component codes!).And the use of iterative methods for PC with the number of iterations of the order of several hundred, and sometimes significantly more than 1000, takes the complexity of their decoding to levels so high that the discussion of the simplicity of the implementation of these algorithms becomes generally irrelevant. Further, we note that the style of polar codes for decoding RS codes is implemented by algorithms, as indicated by some authors [6], with complexity O (Ln^3log(n)). It is possible that it is because of their high complexity that it has so far been possible to find published results in the style of PC only for short RS codes. Compare this situation with the results of [7- 9], where symbolic codes of great length are decoded with linear to the length of code complexity on those the same processors at speeds of hundreds of Bit/s to Mbit/s for the big noise (demo soft can be rewritten at the above resources, along with instructtions for application and then analyzed on their own or with our support).

Another area of application of the PC – for channels with erasures - is also possible, but still, as indicated in [1], not as effective as we would like. In [8] the characteristics of the MTD for such channels are also given with the linear complexity of the recovery algorithms depending on the code length, which can be used for comparison with [1]. The specified MTD works quite effectively even at R~0.96 C, when the probability of unrecoverable characters at R=1/2 in the channel with the probability of erasure p_{ers}~0.48 does not exceed the level of 10^{-6}. Results [8] followed by publications on symbolic codes [9, 11, 14] in fact, they completely close another competitive field in the theory and technology of noiseproof coding, where the absolute advantage is probably for a long time, but it is possible, what is almost forever fixed to the ideology of Optimization Theory and specifically to the methods of MTD in these two most important areas of the development of decoding algorithms. Taking into account the above considerations, we will consider the characteristics of the methods shown in Fig. 1. Recall once again that, while claiming high PC decoding characteristics, the absolute majority of these codes are considered for very small lengths, which was already noted above. But it is quite obvious that for short codes of length less than $n \sim 10^4$ about the efficient operation of algorithms near the channel capacity it can't go for any codes. At the same time, the degree and speed of approximation of the characteristics of any algorithms to the bandwidth of the channel with increasing length for different classes of codes are always very different. Therefore, for small code lengths, the comparison of decoding algorithms requires a careful analysis of specific implementations of the algorithm layouts for the PC. Demo programs for many other code classes can be rewritten from resources [7] and then analyzed under various code parameters.

Let's start with the fact that the Ar graph for the source code really shows low PC capabilities presented by the author of the method. All competing methods clearly have significantly higher characteristics in all variants of the selected implementations VA and MTD, and the latter method decodes its codes at a very high speed. Both methods use codes both larger and shorter than 1024, which shows a good range of capabilities of these methods. Moreover, the algorithms that are actually the simplest in terms of implementation and structure of the VA and MTD codes actually have the same characteristics as the classes methods marked as P6 and DS that use codes with a complex structure, including concatenation .

Estimates show that the application of the simplest well-known cascading methods to the AB and MPD algorithms, which almost do not reduce their speed, will reduce the probability of error per block of these methods by about 2 orders of magnitude. This will further improve the quite acceptable characteristics of the above codes, taken for preliminary comparison with without any special selection. These codes, along with the application of the principles of divergent decoding [13], are successfully developed in simple effective combinations and have broad application prospects. Of course, some unprincipled clarification of the situation with different codes on issues of efficiency and complexity is possible if reliable information on the implementation of algorithms for the PC becomes available. In the meantime, the very long absence of any meaningful concrete results on the implementation of decoders for PCs only confirms the rather common opinion on the unjustified complexity of decoding algorithms for at least a number of types of polar codes.

Back to Figure 1. The PL, LX, ms, and C graphs correspond to various methods that involve complex iterations of PC algorithms, apparently much more complex than those of MPD. In addition, many of them relate to list decoding. It is important to emphasize that the introduction of lists is a forced measure that saves inefficient algorithms. But their use makes it very difficult to apply the decoding technique in real systems connections, when it is not clear how to find the true messages in a bag of possible hundreds of "similar" equally possible data received from the channel. These methods of receiving a "list" should always be analyzed only separately. Conventional decoding methods that are more consistent with real telecommunications systems should not be compared with list systems, since this is a fundamentally different formulation of coding problems, the content of which still needs to be proved or somehow justified.

Note further that the BVAP1000 boundary for block modifications of AB with simplified the implementation shows together with the K36 graph, so far in the form of estimates, the characteristics of another very uniform and simple method for decoding small-length codes. Its correct use, together with the simplest concatenated methods, has already created the conditions for the implementation of successful simple decoding of short messages at a high noise level.

In conclusion, we note that the current state of many years of research on structures and algorithms for polar codes is still in the state of a prolonged initial stage of formation. So far, there is no basis for the claims that an effective and simple method has emerged that can be used in many real-world communication systems. The above-mentioned complexity of algorithms for RS codes of order O ($Ln^3 log(n)$) also does not allow us to hope for an early effective extension of this approach to non-binary codes. The current state of applied coding theory and technology is characterized by good levels of efficiency and an acceptable level of complexity in implementing LDPC codes, as well as the rapid and successful development of new code modifications and decoding algorithms with direct metric control. New results for such codes will allow in the near future to clarify the ratio of their capabilities for Gaussian channels. And the absolute advantage the MPD algorithms in comparison with other methods for non-binary and erasing channels has already been formed and is so significant that, apparently, it will be fixed for these powerful system error correction tools from the arsenal of Optimization Theory of error-correcting coding for a very long time.

Literature

1. Arikan E. Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels. // IEEE Transactions on Information Theory, Vol. 55, No. 7, 2009, - pp. 3051-3073.
2. Seidl M., Huber J. B. An Efficient Length - and Rate-Preserving Concatenation of Polar and Repetition Codes. // International Zurich Seminar on Communications (IZS), February 26 – 28, 2014.
3. Dumer I., Shabunov K. Soft-Decision Decoding of Reed–Muller Codes: Recursive Lists. // IEEE Transactions on Information Theory, Vol. 52, No. 3. - March 2006. - p. 1260-1266.
4. Semenov P. K. Decoding of generalized cascade codes with internal polar codes. // Information and control systems, No. 5 (60) / 2012 .
5. Morozov R. A. Decoding of polar codes using the Dumer-Shabunov algorithm. // Ref.: spisok.math.spbu.ru/2013/txt/papers/s7_1.odt
6. Miloslavskaya V. D. Methods of construction and decoding of polar codes. // PhD thesis. SPSPU, 2014, 206 p.
7. Zolotarev V. V., Ovechkin G. V. Noiseproof coding. Methods and algorithms. Handbook. M., "Hotline – Telecom", 2004, 126 p.
8. Zolotarev V.V., Zubarev Yu.B., Ovechkin G.V., Averin S.V., Ovechkin P.V. 25 years

Optimization Theory of Coding: New Perspectives / / **Plenary report**. Materials of 18th International Scientific and Technical Conference "Problems of Information Transmission and processing in Telecommunications Networks and Systems", 2015, p. 10-17.
9. Resurses: **http://decoders-zolotarev.ru** , **www.mtdbest.ru** , **www.mtdbest.iki.rssi.ru** .
10. Bloch E., Zyablov V. Generalized concatenated codes. Moscow: Svyaz,1976, 240 p.
11. Zolotarev V.V., Chulkov I.V., Ovechkin G.V., Satybaldina D. Zh. Methods of acceleration of algorithms for decoding symbolic codes / / Modern problems of remote sensing of the Earth from the Space., Moscow, IKI RAS, 2014, Vol. 11, no. 2, pp. 138-151.
12. Zolotarev V.V., Ovechkin P.V. Characteristics decoding of block codes for Viterbi algorithm in remote sensing systems // XIII all-Russian open conference "Modern problems in remote sensing of the Earth from space", Moscow, IKI. 2015.
13. Zolotarev V.V., Ovechkin G.V. the Use of divergent coding in channels satellite communication and remote sensing. // XIII All-Russian Open Conference "Modern problems of remote sensing of the Earth from Space", Moscow, IKI RAS. 2015.
14. Zolotarev V.V., Zubarev Yu.B., Ovechkin G.V. Multithreshold decoders and optimization theory of coding. // Edited by Academician of the Russian Academy of Sciences V. K. Levin. M., "Hotline-Telecom", 2012, 238 p.