Gaussian multi-level FM for high-bandwidth satellite communications

Michael Fitch, Keith Briggs, Ivan Boyd (BT Exact) & Fred Stentiford (University College London)
2004 Feb 27

1 Introduction

This contribution describes a modulation scheme suitable for satellite communication, which is based on FM and provides an increased bandwidth of 2 to 3 times that which is currently available commercially with QAM schemes. The system provides best advantage when used in high bandwidth links that occupy the whole transponder. This provides a potential throughput of over 400Mbits/s through a conventional 72MHz transponder where the current state-of-the-art provides 155Mbit/s using 8 PSK systems.

The immediate commercial case for the proposed modulation system is point-to-point high-bandwidth services to large receiving antennas (>5m). Applications include ISP backbone connections, and satellite news-gathering. As the scheme is further optimised, it is believed that it will also become suitable for small dish services such as TV broadcast and provide a viable and better alternative to the traditional QAM schemes used, for example, in DVB.

The modulation scheme is multi-level Gaussian frequency-shift keying (MGFSK). It is a narrow-band FM scheme and as such it has a constant envelope and is therefore suited (a) to low-cost transmitters where the output amplifier can be saturated and (b) to satellite channels where the satellite transponder can be saturated. While MGFSK of itself is not new, it has not been applied to commercial satellite communications before. In designing the satellite MGFSK system, an innovative demodulator has been developed which uses look-up tables that train themselves according to the channel characteristics. The channel is characterized by sending a unit impulse into the channel and measuring the response at the receivers.

Following this introduction, section 2 of the paper provides a very brief view of the market opportunities. A technical overview of the modulation scheme is presented in section 3. Section 4 describes the symbol detector and presents results obtained to date. Finally, section 5 provides conclusions and proposals for further work.

2 Market opportunities

Satellite communications has been a very important part of the global telecommunications backbone for the past three decades. Although the technology is mature, the industry is still expanding, achieving a 15% average annual growth rate over the last 6 years. The industry value chain includes: satellite manufacturing; launch vehicle manufacturing; ground equipment manufacturing; and satellite communications services. Communication services is the fastest growing part of the value chain, trebling in size from 1996 to 2002. Satellite communications now accounts for around 10% of global telecommunications revenues [1].

The proposed modulation scheme would enable more efficient use of satellite transponders for this growing communications demand. Given that the typical market rate for leasing a Ku band transponder is $1.6M to $2M per year, there is significant incentive for service providers to use transponders very efficiently.

The key applications areas where the proposed scheme could be employed are satellite trunking of ISP backbone traffic, satellite news gathering (SNG) and military applications. These are the most suitable applications because large dish antenna receivers can be used and it is feasible to upgrade the systems because of their relatively small installed base.

Satellite operators provide Internet Services Providers (ISPs) with long distance point-to-point (trunk) links to the Internet backbone. Currently, approximately 4% of satellite transponders are used for Internet trunking services [2]. Although optical fibre connections are increasingly being used in the mature communications markets, in other less developed or remote locations, satellite remains the only option, or at least the only cost-effective option. Global demand for satellite Internet trunking is predicted to approximately double from 2005 to 2010 [3] and hence it provides an attractive market for a more efficient modulation scheme.

Satellite news-gathering is becoming increasingly important with the popularity of TV news (and, in particular, ‘rolling news’ channels such as CNN, Sky News and BBC News 24). The demand for the rapid transmission of up-to-date reports from international trouble spots has grown sharply. However, the quality of the broadcast images and sound are often poor because of the limited bandwidth caused by the lack of capacity in some regions, often the very regions where major world conflicts or catastrophes occur. The ability of the proposed modulation scheme to provide 2-3 times increased capacity would significantly enhance the quality of news broadcasts from remote regions.

There has been a huge increase in demand for bandwidth in the military sector, due to new strategies of 'network-centric warfare'. The fundamental principle of these strategies is to be able to conduct military operations with a numerically inferior force, by ensuring that all members of that force have a full and up to date picture of both the enemy and own forces. The result of these strategies is that intelligence assets (mainly aircraft and special forces) are generating very large quantities of sensor data (especially imagery) that needs to be rapidly transmitted to command and control systems and then to weapons systems. Typically satcoms provide 70% of the total bandwidth requirements, with terrestrial wireless and optical fibre networks providing the rest. There is thus a very significant driver from the military for more efficient use of satellite transponder capacity.

3 Technical overview

An MGFSK scheme has been developed specifically for satellite use, to meet the following technical requirements

- High bandwidth efficiency, to achieve low-cost space segment
- High tolerance to equalisation error (amplitude and group-delay ripple) in the channel, since the cost of setting up earth stations to achieve the required equalisation error across the signal bandwidth is very high, becoming worse for wide-band signals.
- High tolerance to non-linearities in the channel, such as one or more amplifier elements that could be running close to or at saturation point

The state-of-the-art in commercially viable bandwidth efficient modulation is represented by 8PSK with error-coding in a Trellis-coded modulation (TCM) scheme, which achieves 155Mbit/s (STM-1) in a 72MHz satellite transponder. There is not much market take-up of satellite links at this speed, primarily because of cost of space segment but also because the links are very difficult to set up due to the tight specifications on linearity and equalisation errors.

If the cost of such links and ease of set-up were to be improved, the proposition of high bandwidth links over satellite would be much more attractive. Whereas 8PSK offers a bandwidth efficiency of just under 3 bits/s/Hz (uncoded), MGFSK has been demonstrated to offer nearly 6 bits/s/Hz. This efficiency is comparable to 64QAM, but the use of 64QAM over satellite is not considered viable because it requires the channel to be highly linear and well equalised.

The chosen parameters for MGFSK for the satellite application are:

- Multiple levels. 16 levels gives 4 bits per symbol, but others values are possible
- Partial response signalling. The symbols are formed by the impulse response of a low-pass Gaussian filter that stretches over adjacent symbols, similar to that in use with GSM. This enables a reduction in occupied bandwidth at the expense of some inter-symbol interference (ISI) and loss of orthogonality. The combination of multiple levels and partial response signalling results in a bandwidth efficiency of almost 6 bits/s/Hz
- Low modulation index. The filtered waveform is frequency modulated onto a carrier using a low modulation index (β) in order to keep the occupied bandwidth in the narrow-band class. Because it uses frequency modulation (FM) instead of the more common PSK, the signal is tolerant to equalisation errors and non-linearities in RF amplifiers

These parameters are designed to maximise the bandwidth efficiency and at the same time maintain a constant envelope. The use of FM maximises tolerance to non-linearities and equalisation error. However, the choice of parameters above comes with the penalty of inferior noise performance when compared to the PSK and QAM schemes. Simulations using Matlab Simulink with a simple symbol detector design indicate that this penalty is about 5dB for a symbol error-rate of 1 in 10^5 when compared with 64QAM which has equivalent bandwidth efficiency. However this penalty is considered tolerable because 64QAM requires that the amplifiers in the channel be backed off by at least 10dB whereas the 16-GFSK signal is not degraded even when there is no back-off.

A block diagram of 16-GFSK is shown in figure 1.

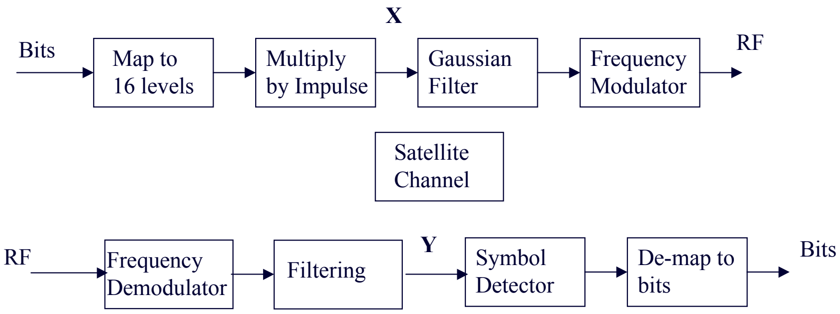

Figure 1 Block diagram of MGFSK system

The modulator section at the top begins by mapping the bits to an alphabet of 16 symbols, which equates to 4 bits per symbol. The 16 symbols are represented by different amplitudes symmetrical about zero, that is, -7.5, -6.5, -5.5 …..-0.5 , +0.5…….+5.5, +6.6, +7.5. These are multiplied by impulses of height 1 to produce impulses of –7.5, -6.5 …….+6.5, +7.5 at point '**X**'.

The response of the Gaussian filter to these impulses will be Gaussian in shape, that is, smooth waveforms with no negative or oscillatory characteristics, which are applied to the frequency modulator. This method of modulation is very similar to the scheme used by the GSM air interface, except that GSM uses only 2 levels (+/-1). The bandwidth of the Gaussian filter is a critical parameter; if it is set small, the occupied bandwidth is reduced, but the intersymbol interference is increased since it sets the width (in time) of the impulse response. With this scheme, the BT product (bandwidth * symbol period) has been set at 0.35 as a reasonable trade-off, although it could be varied as part of an optimisation process and indeed different values are simulated later in the paper. GSM cellular uses BT = 0.3.

The modulation index of the FM modulator is held at $\beta = 1$, which according to Carson's rule means that the occupied bandwidth can be considered as $2f_m$, where f_m is the maximum frequency component of the modulating (baseband) signal which is set by the bandwidth of the Gaussian filter. Under this condition the occupied bandwidth can be calculated to be

BW = Symbol rate * BT * 2 = Symbol rate * 0.7 (Hz)

and the (uncoded) bandwidth efficiency E is E = 4 (bits / symbol) / 0.7 = 5.71 bits /s/ Hz.

A simulated 16-GFSK spectrum is shown in figure 2.

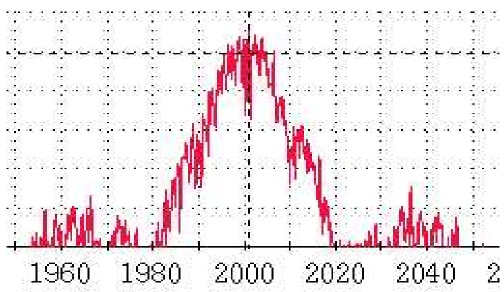

Figure 2 16-GFSK simulated spectrum

In figure 2, the carrier frequency is held low for the purposes of simulation at 2kHz and the symbol rate is 10 baud. The 3dB bandwidth (shown by the dotted horizontal line) of this spectrum is just 7Hz. When scaled to 400Mbit/s, the symbol rate is 100Mbaud and the bandwidth will be 70MHz.

The demodulator consists of an FM demodulator, filtering and a symbol detector as shown in figure 1. The filter has to be designed carefully to band-limit the noise but at the same time preserve a well-behaved impulse response; that is, it should not add significantly to the inter-symbol interference. The actual filtering implementation is performed in two stages at different sampling rates. The waveform at '**Y**' in figure 1 should be as close as possible to the waveform at the output of the Gaussian filter in the modulator.

4 Symbol detector

In order to determine the ultimately achievable performance of the system, various parameters need to be optimised, subject to design and engineering constraints. Any coding/decoding scheme is optimised when:

- parameters governing the coding procedure are adjusted to minimize the error rate for a given decoding procedure; and
- parameters governing the decoding procedure are adjusted to minimize the error rate for a given encoding procedure.

Ideally, one would do a single optimization over both stages. For the second phase (optimal decoder design), we have tried two approaches:

a. Assume all variables are discrete (that is, effectively work at the digital level), which converts the problem into a very large linear programming (LP) problem. This is in principle possible to solve, but we rapidly reach limitations of computer hardware (at about 1M LP variables).
b. Maintain the continuous model.

Work is still in progress on both these optimisation approaches. However, as a proof-of-concept, we have built a simple detector using Matlab Simulink and some optimization heuristics in order to provide a lower bound to the possible performance. This simple detector makes use of look-up templates that store expected waveforms for every possible symbol transition (256 (16*16) templates) and its operation is illustrated in figure 3.

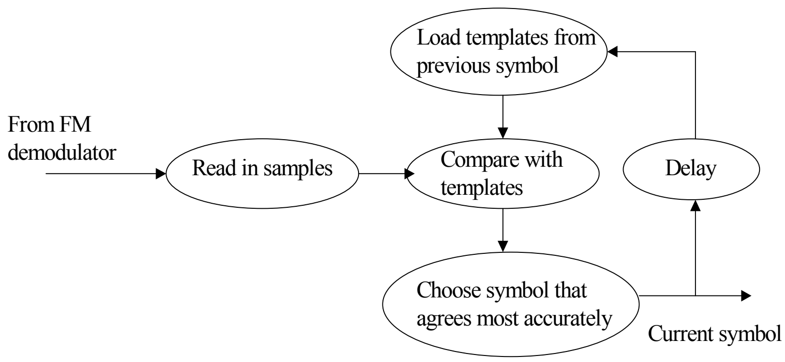

Figure 3. Detector top-level functionality

The entire set of templates can be built from the unit impulse response. Since FM is used over the channel, linearity can be assumed, even though the channel is highly non-linear. The templates can be re-built at regular intervals if needed.

This detector, which has a memory of one symbol, has been built in Matlab Simulink and onto an FPGA hardware prototype. We have fixed the number of samples per symbol at 4, as this is the maximum that can be handled by FPGAs at a symbol rate of 100M baud, due to the current limitations of the silicon. The performance of the simple detector in noise is plotted on figure 4, together with the theoretical plots for QPSK and 64-QAM for easy comparison [4]. It was noticed that the error patterns are very different from those experienced with QAM schemes, in that when the symbol detector makes mistakes due to noise, the detected symbols are wrong only by +/- 1 symbol and the erroneous symbols frequently appear in pairs. These error characteristics, together with confidence information that can be derived from the template comparison error values, can be used to design an efficient forward-error correction (FEC) scheme based on simplified Reed-Solomon codes with erasures. The design of this FEC will be the subject of future work.

As the impulses are spread over more than one symbol, a separate heuristic activity has investigated the performance enhancement that would stem from increasing the memory to two or more symbols and from varying the sample weightings used in the template comparison (an increase in memory to N symbols implies 16^N templates). A repeating pattern of symbols with labels *0, 1, 2, …, 15* is generated with successive groups of symbols in the following 9-symbol format:

000RABCDR

where R is a random symbol and ABCD cycles through all 65536 possible combinations of four symbols. In this way a single data set consists of 9 x 65536 x 4 = 2,359,296 signal values. Each group of four symbols is separated by a buffer of three 0 symbols and two random symbols to minimise the effects of ISI both between groups, and between groups and the buffer. Each data set is synthesised with different levels of noise and ISI that reflects levels likely to be encountered in practice.

Transmitted symbols are received as a stream of equally spaced signal samples x_i, every four of which correspond to one of N different symbols at the receiver. These signal values differ from those transmitted owing to various levels of noise and the effects of intersymbol interference.

A nearest neighbour classifier is used to determine the identity of the received symbol C in each group in which the classification is determined by the label a of the reference point $\boldsymbol{R}_a$ that is closest to the candidate point $\boldsymbol{C}$. A Euclidean metric is used in which the weighted squared distance between $\boldsymbol{C}$ and $\boldsymbol{R}_a$ is given by

$$D(C, R_a) = \sum_{k=1}^{d} \lambda_k (x_k - r_{ak})^2$$

where $\boldsymbol{R}_a = (r_{a1}, r_{a2}, r_{a3}, \ldots, r_{ad}, r_{a(d+1)}, \ldots, r_{at})$
$\boldsymbol{C} = (x_1, x_2, x_3, \ldots, x_d)$
$a = 0, 1, 2, \ldots, 15$
λ_k is the weighting applied to individual sample values
d is the number of signal samples used in the metric, and
t is the total number of samples in the reference.

The classification of $\boldsymbol{C}$ is given by $\underset{a}{Min}^{-1} D(C, R_a)$.

The effects of ISI mean that a single reference $\boldsymbol{R}_a$ is not able to represent satisfactorily all variations of $\boldsymbol{C}$ that belong to class a without introducing errors. Thus a number of reference vectors $\boldsymbol{R}_{aj}$ are used to capture the effects of the intersymbol interference from the preceding combinations of symbols A and B, or B alone and the classification of $\boldsymbol{C}$ becomes

$$\underset{a}{Min}^{-1} \underset{j}{Min}\, D(C, R_{aj})$$ where j = 0, 1, …, 15 for the 16 template case (table 1)

or j = 0, 1, …, 255 for the 256 template case (table 2).

The performance was evaluated using signals with three levels of noise in combination with three levels of ISI. In each case the algorithm was tested using reference templates $\boldsymbol{R}_{aj}$ derived from the noise free signal; 16 templates per label are used to take account of the ISI effects of the immediately preceding symbol B, and 256 templates per label to take account of the two preceding symbols A and B.

	ISI		
Eb/No	**BT=0.35**	**BT=0.3**	**BT=0.25**
λ_k $k = 0,1,2,\ldots$	1, 1, 1, 1	0.7, 1, 1, 1	1, 1, 0.6
No noise	0	0	0
24	0	3	846
19	86	580	5014

Table 1. Number of errors in 65,536 symbols – 16 templates per label.

	ISI		
Eb/No	**BT=0.35**	**BT=0.3**	**BT=0.25**
λ_k $k = 0,1,2,\ldots$	1, 1, 1, 1	0.6, 1, 1, 1	0.9, 1, 1, 1, 0.4, 0.1
No noise	0	0	0
24	0	4	670
19	85	585	4657

Table 2. Number of errors in 65,536 symbols – 256 templates per label.

The results, which show the number of errors in 65,536 symbols, indicate that there is little advantage in using templates extending more than one symbol into the past where the ISI is less than or equal to 0.3. However, for ISI above this, only three sample values are necessary to obtain the best performance in the 16 template case, and up to six yields the best results when 256 templates are used. More data and further analysis is necessary to increase confidence in the lower valued error rates. The values for E_b/N_0 = 19dB are plotted on figure 4 for values of Gaussian filter BT = 0.25, 0.3 and 0.35.

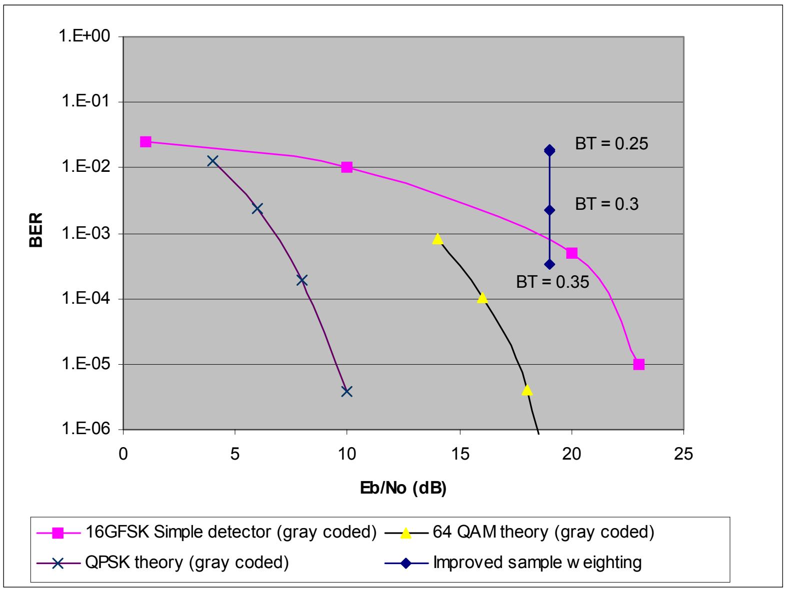

Figure 4 Performance in noise

Figure 4 shows that optimisation of the sample weights improves the noise performance by roughly 1dB at BT = 0.35. If the BT is moved to 0.3, then bandwidth is saved and figure 5 shows that the penalty is 1dB degradation in noise performance.

5 Conclusions and further work

The use of 16GFSK for satellite channels gives a 2 to 3 times improvement in bandwidth efficiency, suited to large dish services and is tolerant to channel impairments such as non-linearity and equalisation errors

A symbol detector for 16GFSK has been demonstrated to have about 5dB worse noise performance than 64-QAM for an equivalent bandwidth efficiency, however this degradation can be compensated for by operating transponders in saturation, that is, at approximately a 10dB higher power level.

Increasing the memory in the detector beyond one symbol does not improve the noise performance for BT > 0.3, however there is some improvement gained through experimentation with sample weighting

Further work is planned to develop a more rigorous theory for an optimum symbol detector, to design a suitable FEC scheme and to produce schemes to recover the symbol and bit timing. This will enable a fully functional production prototype to be completed and rigorously tested.

6 References

[1] Telecommunications Fixed Satellite Services US Equity Research, Credit Lyonnais Securities

[2] Satellite Industry Statistics 2002, Satellite Industry Association

[3] Mark Williamson, *Internet-over-satellite: backbone or spare rib?,* Telecommunications International, November 2002.

[4] K. Feher, Advanced Digital Communications. ISBN 0-13-011198-8. Prentice-Hall 1987, pages 328 and 329.