Eavesdropping and Network Analyzing Using Network Dispersion

Shlomo Marciano[1,2], Shalva Ben-Ezra[1] & Er'el Granot[2]

[1] Finisar Israel, Golda Meir 3, Nes-Ziona, Israel

[2] Department of Electrical Engineering, Ariel University, Ariel, Israel

Correspondence: Shlomo Marciano, Finisar Israel, Golda Meir 3, Nes-Ziona, Israel. E-mail: shlomo.marciano@finisar.com

Abstract

The dispersion effect in an optical network is utilized in several applications. The pattern of the received signal is matched to a pre-prepared data bank. Useful data can be retrieved from this by data matching. It is demonstrated that in a Time-Division-Multiplexing configuration, not only that the data can be retrieved but also the data of the adjacent detectors, i.e., this method can be utilized for optical eavesdropping. Moreover, it is also shown that this method can be utilized as an affordable method to synchronize the data stream and to find the length of the dispersive fiber (or, for a given fiber length, this method can be used to measure the fiber's dispersion coefficient). One of the benefits of this method is that it can operate, while the network is running without deteriorating its performances. The conclusions are based on theoretical and numerical analysis as well as on experimental results.

Keywords: dispersion compensation devices, fiber optics communications, fibers, single-mode, dispersion compensation devices

1. Introduction

Dispersion is one of the main obstacles to optical communications. In fact, as optical fiber communication has developed and problems related to attenuation have been addressed by using low attenuation fibers (like the ubiquitous smf 28) and relatively low cost amplifiers (mostly Erbium doped fiber amplifiers (Desurvire et al., 2002)), the main hindrance in the optical communication line is dispersion (Ramachandran, 2010). When a pulse propagates in a dispersive medium, such as an optical fiber, it is deformed. As a consequence, the pulse, which indicates a specific digital bit, exceeds the allotted time slot, and therefore beyond a certain distance, the signal data cannot be recovered.

There are numerous means to mitigate the dispersion effects, either by hardware elements, like Dispersion Compensating Fibers (Ramachandran, 2010), or by software, such as DSP methods, which are incorporated in coherent detection (Shieh & Djordjevic, 2010). From this perspective, dispersion is a negative phenomenon which should be fixed. However, dispersion may have some very useful characteristics; for example, dispersion is used to mitigate the influence of nonlinear effects.

In this work dispersion addresses several applications.

In the first application, dispersion is used to decode data from neighboring bits in a Time-Division Multiplexing (TDM) network (Prat, 2010). TDM is used in many networks, where there is large gap between the fiber's data capacity and the capacity of the network's end elements. For example, in ordinary optical communication line, the fiber can carry more than 1Tb/s, while a single transmitter or receiver can process only a fraction of that rate (usually several 10Gb/s).

To be more specific, we are referring to interleaved TDM, in a system with Q interleaved channels transmitted at an aggregate bit rate of QB b/s (see Figure 1 for Q=4). Each receiver has an output bit rate of B b/s and gets similar optical signal as other receivers, but only measures every Q'th bit and receiver q measures at a time shift of one bit, at the QB bitrate, relative to the q-1 and $q+1$ receivers ($0 < q < Q$-1).

Most of the long reach networks today are based on Wavelength (or Frequency) Division Multiplexing (WDM or FDM) (Antoniades, Ellinas, & Roudas, 2012), in which every transmitter/receiver is dedicated to a given portion of the fiber's spectrum which was allocated to it. In the TDM approach, the data is encoded in very short pulses,

which are transmitted in a high bit rates, but the data are routed by temporal multiplexers to numerous receivers. Correspondingly, the data from numerous transmitters is routed to the fiber.

The first object of this work was to demonstrate, that in such a TDM network, by measuring the data received by a single receiver, it is possible to decode the data, which was transmitted to its adjacent receivers, i.e., to the one allocated to its neighbor bits channels.

For example, suppose four 10Gb/s receivers and transmitters are multiplexed to a single 40Gb/s channel, and let us further assume that the signal of, say, the 3rd channel is measured with high precision, then the first conclusion of this work is that it is possible to decode the data of the 1st, the 2nd and the 4th channels as well from the 3rd channel measurements alone. That is, all four channels data can be deciphered from only the 3rd channel measurements.

It should be stressed that in ordinary fiber tapping and eavesdropping a special hardware is required, and the tapping must be placed in close proximity to the fiber(see, for example, Iqbal, Fathallah, and Belhadj, 2011; Grishachev, 2012; Hayes, ND)). In the proposed technique, on the other hand, no such hardware is required.

Furthermore, there are numerous methods to eliminate or to mitigate signal deformation caused by dispersion. However, in all cases, dispersion measurement is required. When the communication line is fixed, the total dispersion can be estimated from the length of the fiber and knowledge of the fiber's dispersion coefficient, or can measured when the fiber is first deployed. However, many networks are flexible, and vary in time; therefore it is very difficult to evaluate the dispersion exactly at any given time. Methods to estimate dispersion usually require a pilot tone, which, after being detected at the other end of the fiber, can be used to determine the fiber's length, provided the group velocity dispersion (GVD) parameter is known (see, for example Noh, Kim, Oh, & Pack, 1999; F. Hakimi & H. Hakimi, 2001; Fernando, 2001; Derickson, 1998). These methods are cumbersome, for they require an established protocol between the transmitters and the receivers, and they either reduce the data spectral bandwidth or reduce signal to noise ratio. Moreover, these methods require access to both ends of the fiber, which in many cases is an unattainable requirement. A simple method to estimate the exact GVD parameter, such as the one suggested in this work, (which does not require any additional hardware, nor does it require an access to both ends of the fiber, nor does it require an allocation of spectral bandwidth for the dispersion measurement, and it can measure while the network is operating), can be useful in mitigating dispersion in time-varying networks.

In the second part of this work, a similar method is used in estimating the total dispersion of the medium when the length of the fiber is known. Conversely, in the case where the GVD parameter is known, the length of the medium can be estimated. This method does not require sending special pulses, signals or protocol. Instead, the evaluation is done by comparing the pattern of the arriving signal to a pre-prepared data bank. Matching the pattern allows decoding the data, evaluating the GVD parameter, and signal synchronization – all in one algorithm, and without the need to access both ends of the fiber.

2. Motivation

Dispersion is a linear but non-local effect. Hence, the deformations that affect any single pulse in the data stream, also affect other pulses in the data stream, particularly the pulses closest to it. Thus, measuring the correlations between neighboring pulses, which usually have a negative effect on optical communication, can be harnessed to gather information about the values of nearby bit slots or to evaluate the length of the fiber.

The general idea of the algorithm is based on the assumption that since the dispersion dynamics are well known, and, more specifically, the dynamics of any smooth rectangular pulse is known with high precision, then it is possible to anticipate the signal shape for any pattern of pulses after dispersion.

One implementation is to create a database, which includes the waveform of all possible digital patterns of a given length, after transmission through different dispersive media. Then, by comparing a small section of the entire received signal to the different shapes in the database, and locating the waveform that best fits the measured data, the exact digital pattern that was sent can be recovered. In addition, the waveform includes information not only on the bits which comprise the transmitted pattern, but also information on neighboring patterns. In cases where the neighboring patterns are not necessarily decoded into binary representation by the receiver, such as TDM, information from neighboring patterns nonetheless is present in the received waveform (see Figure 1).

3. Mathematical and Physical Background

When an electromagnetic pulse propagates in a dispersive medium its envelope, in the slowly varying approximation, is governed by the Schrödinger equation (Ramachandran, 2010). In real systems the shape of the resulting pulse is governed by nonlinearities, attenuation, polarization effects and wave equation corrections to the Schrödinger equation. However, in relatively short networks (less than 60km, which is the maximum distance of all PON's) polarization, attenuation and nonlinearities effects can, in the first approximation, be neglected.

Moreover, provided the pulse rise time is substantially shorter than the reciprocal of the optical carrier's frequency, the Schrödinger equation is an excellent approximation to the Maxwell wave equation. Finally, third order dispersion can also be neglected. As Granot, Luz, and Marchewka (2012) explained, even for data rates as high as 50GB/s, the third order dispersion term is negligibly small in comparison to the second one.

Hence, the dynamics of the *envelope* of the electromagnetic field $A(z,t)$ is governed by the following Schrödinger equation:

$$\frac{\partial A(z,t)}{\partial z} = i\frac{\beta}{2}\frac{\partial^2 A(z,t)}{\partial t^2} - \alpha A(z,t) \tag{1}$$

where $t = \tau - z/v$ is the time measured with respect to the fiber's time of flight, τ is the time and v is the group velocity in the fiber, α is the attenuation coefficient and β is the GVD parameter (which for brevity we omit the subscript 2).

At $z = 0$ (the beginning of the fiber) an infinite sequence is sent into the fiber.

The sequence can be either OOK, i.e., for example, $x_n = [\cdots 1,0,0,1,0,1,1,0\cdots]$, a PSK or DPSK, i.e., for example, $x_n = [\cdots 1,-1,-1,1,-1,1,1,-1\cdots]$, a 4QAM, i.e., $x_n = [\cdots(1+i)/\sqrt{2},(1-i)/\sqrt{2},(-1+i)/\sqrt{2},(1-i)/\sqrt{2},\cdots]$, or *any other digital sequence*. Furthermore, assume that the laser is modulated by (smooth) rectangular pulses then the envelope of the electromagnetic field can be written (this formulas were derived in Ref.13 as the optical equivalence of the quantum dispersion analysis that was derived in (Granot & Marchewka, 2011; Granot & Marchewka, 2005; Marchewka, Granot, & Schuss, 2007))

$$A(z=0,t) = A_0 + \sum_{n=-\infty}^{\infty} x_n \, \mathrm{srect}_{\xi T}(t-nT,z,0) \tag{2}$$

where

$$\mathrm{srect}_{\xi T}(t,\Delta,0) \equiv$$
$$\frac{1}{2}\left\{ \mathrm{erfc}\left[\left(t-\frac{\xi T}{2}\right)\frac{\pi\Delta}{\sqrt{2\ln 2}}\right] - \mathrm{erfc}\left[\left(t+\frac{\xi T}{2}\right)\frac{\pi\Delta}{\sqrt{2\ln 2}}\right]\right\} \tag{3}$$

$T = 1/B$ is the bit period (B is the Bitrate), ξ is a measure of the duty cycle (i.e., it determines the normalized pulses width), A_0 is a dc term, which is a manifestation of the modulation extinction ratio, Δ is the reciprocal of the pulses rise time and erfc is the complementary error function (Abramowitz & Stegun, 1965).

It should be noted that Equation (2) and (3) can be utilized for both NRZ and RZ with an arbitrary duty-cycle. $\xi = 1$ corresponds to 100% duty cycle and therefore to NRZ sequence, while $\xi < 1$ corresponds to a duty cycle of $\xi \times 100\%$, which is the general definition of an RZ signal. In most practical RZ cases, $\xi \cong 0.5$. It should be emphasized again that Equation 2 is totally generic for any digital sequence (OOK, PSK, DPSK, 4QAM, 16QAM, etc.).

Furthermore, the signal in Equation (2), is the equivalent of launching a stream of rectangular pulses,

$A(z=0,t) = \sum_{n=-\infty}^{\infty} x_n \, \mathrm{rect}_{\xi T}(t-nT)$, which consists of perfect rectangular pulses $\mathrm{rect}_{\xi T}(t) \equiv \begin{cases} 1 & |t| \le \xi T/2 \\ 0 & |t| > \xi T/2 \end{cases}$, through a

spectrally Gaussian filter $H(f) = \exp\left[-2\ln 2(f/\Delta)^2\right]$.

The Gaussian filter will be responsible for smoothing the pulse edges.

The transfer function of the dispersive fiber can be written

$$G(f) = \exp\left(i\beta L(2\pi f)^2/2\right) \tag{4}$$

where β is the GVD parameter (see Equation 1), and L is the fiber's length. Hereinafter we disregard higher dispersion terms, since, as was explained in Ref.13, the influence of the next term even for high bitrates, is negligible.

In this case the exact solution of the signal at the other end of the fiber can be written directly *without any approximation* (Granot, Luz, & Marchewka, 2012; Granot & Marchewka, 2011)

$$A(z>0,t) = A_0 + \sum_{n=-\infty}^{\infty} x_n \, \mathrm{srect}_{\xi T}(t-nT,\Delta,z) \tag{5}$$

where

$$\text{srect}_{\xi T}(t,\Delta,z) = \frac{\exp(-\alpha z)}{2} \times$$

$$\left\{ \text{erfc}\left[\frac{t-\xi T/2}{\sqrt{i2\beta z + 2\ln(2)/\pi^2\Delta^2}}\right] - \text{erfc}\left[\frac{t+\xi T/2}{\sqrt{i2\beta z + 2\ln(2)/\pi^2\Delta^2}}\right] \right\} \tag{6}$$

and α again represents the attenuation coefficient of the fiber (see Equation 1).

That is, the function srect (6) describes the exact evolution of a smooth rectangular pulse, i.e., a rect pulse after a spectrally Gaussian filter, in dispersive medium.

Now, let us assume that we can measure only what happens within a single bit time-slot, which is what happens in a switchable TDM system, see Figure 1.

Therefore, let the optical signal be sampled at $2M+1$ points within a time slot T, i.e.,

$$-MT/2 \le t_m = m\delta t \le MT/2 \tag{7}$$

Furthermore, let us assume that we are interested only in N nearest neighbors bits, i.e., we assume that the $2M+1$ signal envelope measurements in the zero time slot ($n=0$) can be approximated by

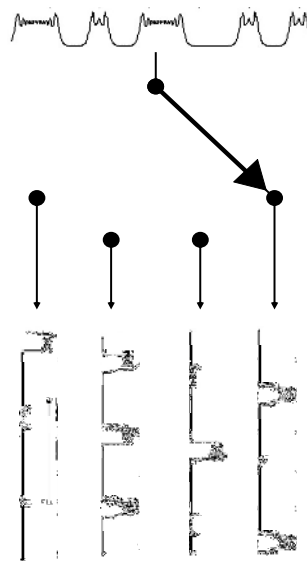$$A(z,t_m) = A_0 + \sum_{n=-N}^{N} x_n \, \text{srect}_{\xi T}(m\delta t - nT, \Delta, z) \tag{8}$$

Figure 1. Switchable TDM schematic. The sequence is splitted into 4 channels, each of which is connected periodically for only 1/4 of the total time.

Hence, the relation between the measured signal $\mathbf{A}$ and the digital data sent $\mathbf{x}$ can be written in a matrix form

$$\mathbf{A} = S\mathbf{x} \tag{9}$$

Where

$$\mathbf{A} \equiv \begin{pmatrix} A(z,t_{-M}) \\ \vdots \\ A(z,t_0) \\ \vdots \\ A(z,t_M) \end{pmatrix}, \quad \mathbf{x} \equiv \begin{pmatrix} x_{-N} \\ \vdots \\ x_0 \\ \vdots \\ x_N \end{pmatrix}, \text{ and } \quad S \equiv \begin{pmatrix} S_{-M,-N} & \cdots & S_{-M,0} & \cdots & S_{-M,N} \\ \vdots & \ddots & & & \vdots \\ S_{0,-N} & & S_{0,0} & & S_{0,N} \\ \vdots & & & \ddots & \vdots \\ S_{M,-N} & \cdots & S_{M,0} & \cdots & S_{M,N} \end{pmatrix}$$

whose matrix components are

$$S_{mn} \equiv \text{srect}_{\xi T}(m\delta t - nT, \Delta, z) \tag{10}$$

4. Eavesdropping in Coherent Detection

In the case of coherent detection (Betti, De Marchis, & Iannone, 1995) the electromagnetic field is measured directly. Hence, the pulse envelope vector **A** is known by direct measurements. Since the matrix S can be approximated by (10) then if the number of unknown bits (2N+1) is equal to the number of measurements in a *single* bit period (2M+1), i.e., when $N=M$ the data can easily be retrieved by inverting the matrix

$$\mathbf{x} = S^{-1}\mathbf{A} \tag{11}$$

However, since the measurements are accompanied by noise it is advantageous to repeat as many measurements as possible. Since the measurements are taken on a live sequence, there are no repetitions, and therefore all the measurements have to be taken in the bit period, however, in principle, the measurements number can be considerably larger than the number of unknown bits. In this case, i.e., when $M>N$, the matrix cannot be inverted, however, the data can be retrieved by using least square error algorithm (Cherkassky & Mulier, 1998), i.e.,

$$\mathbf{x} = \left(S^T S\right)^{-1} S^T \mathbf{A} \tag{12}$$

Since one realization of this technique is to be activated in real time, then it is advantageous to calculate the matrix $M \equiv \left(S^T S\right)^{-1} S^T$ beforehand. In real time only a matrix product with relatively small computational burden is required.

The problem with this method is that it is based on coherent detection, which is currently not used in short-range optical networks, although there is a good chance that it will be in the near future (Jung, Cho, Takushima, & Chung, 2010; Lavery et al., 2012; Cvijetic et al., 2011).

5. Eavesdropping in Non-Coherent Detection

In the non-coherent case, the intensity, rather than the field, is measured, i.e., instead of measuring the *field*, the intensity is measured at 2M+1 different measuring points:

$$I_m = \left|A(z,t_m)\right|^2 = \left|\sum_{n=-N}^{N} x_n \operatorname{srect}_{\xi T}(m\delta t - nT, \Delta, z)\right|^2 =$$
$$\sum_{n=-N}^{N} \sum_{q=-N}^{N} x_q x_n \operatorname{srect}_{\xi T}(m\delta t - qT, \Delta, z)\operatorname{srect}_{\xi T}(m\delta t - nT, \Delta, z)^* \tag{13}$$

Or in a matrix form:

$$I_m = \left|\mathbf{S}_m \mathbf{x}\right|^2 = \left(\mathbf{S}_m \mathbf{x}\right)^* \left(\mathbf{S}_m \mathbf{x}\right) = \left(\mathbf{x}^T \mathbf{S}_m^T\right)\left(\mathbf{S}_m \mathbf{x}\right) = \mathbf{x}^T \left(\mathbf{S}_m^T \mathbf{S}_m\right)\mathbf{x} \tag{14}$$

Hence, for every vector **x** an error function can be defined:

$$E(\mathbf{x}) \equiv \sum_m \left[I_m - \mathbf{x}^T \left(\mathbf{S}_m^T \mathbf{S}_m\right)\mathbf{x}\right]^2 \tag{15}$$

The minimum of this function $E_{\min}$, for all possible bit sequences **x**, is the best estimate of **x**, which should correspond to the sequence with the highest probability $\mathbf{x}_{best}$. Ideally, the exact solution of (14) is the correct **x**, for which $E_{\min} = 0$, which does not occur in real scenarios.

When the number of unknown bits is relatively small, there is no need to guess the combination, but instead all the possible sequences of **x** of a given length can be examined.

In order to find this solution, a database is created for all possible combinations of **x**. In most examples used in this work N=3, i.e., the database included all possible combinations of 7 bits, i.e., 128 combinations. For every the error (15) is calculated. When the minimum error is reached, then the sequence $\mathbf{x}_{best}$ is considered to be the algorithm solution (see Figure 2). It should be emphasized that the term $\mathbf{x}^T \left(\mathbf{S}_m^T \mathbf{S}_m\right)\mathbf{x}$ is the database, and therefore does not have to be calculated in real time.

6. Experimental validation of the optical signal mathematical model

In the first section we demonstrate the ability of the function (5) to express the dynamics of a real 20Gb/s NRZ sequence in a long dispersive medium. The theoretical function prediction is plotted with the experimental results for a single mode fiber with $\beta = 20\,ps^2/km$, and $\alpha = 0.2\,dB/km$ (smf28).

An optical communication laser was utilized with a carrier wavelength 1550nm, modulated with an indium phosphide modulator.

In Figures 3-5, measurements were taken with a fast real time oscilloscope (80Gsamples/s) where every bit consists of only 4 sampling points. It should be stressed that a real time sampling scope was utilized and not a Digital Communications Analyzer (DCA) to demonstrate real time operation without exploiting the PRBS periodicity which does not occur in real network data.
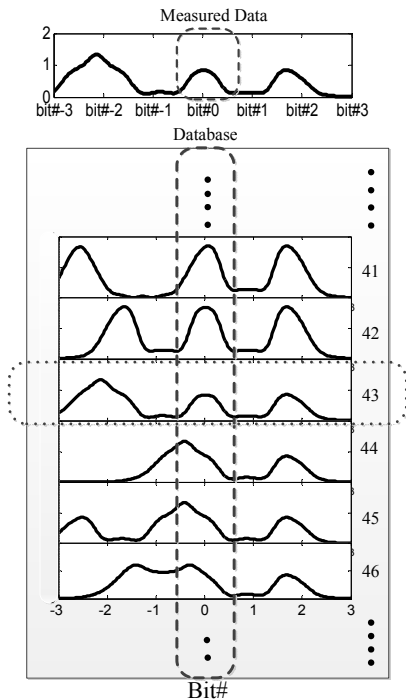

Figure 2. Schematic presentation of the comparison between the measured signal (upper panel) and the database (lower panel). In this case, combination #43, which corresponds to specific bit sequence, was found to yield the lowest error
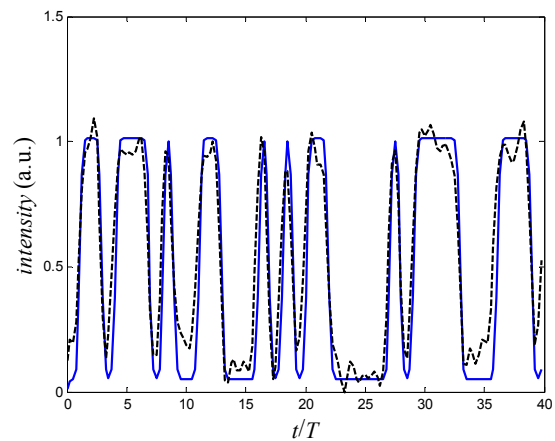
Figure 3. A comparison between measured signal at 20Gb/s, 2^7PRBS (dashed curve) with high bit-rate scope (80Gs/s) and the theoretical analytical signal (solid curve) of Equation 5 at z=0km
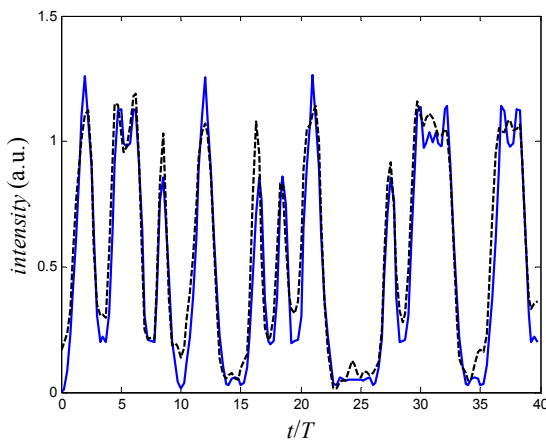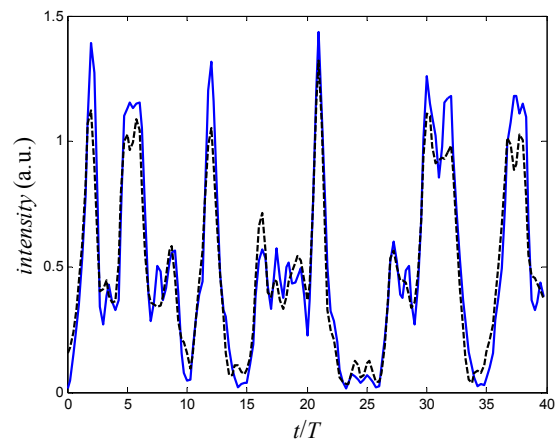

Figure 4. same data as in Fig(2) for z=20km

Figure 5. same data as Fig(2) for z=40km

As can be seen from Figures 3-5, despite the noisy measurements, which are a consequence of the fast sampling rate (note that averaging is not an option here) the theoretical *analytical* prediction is excellent, and it is even improved for 20-40km since the dispersion deformation smooths out the effects of noise.

7. Simulation validation of the decoding algorithm

In the second part we demonstrate the ability of the algorithm (neighbors eavesdropping) to decipher the data of the adjacent unknown bits. After seeing that the mathematical description (5) describes fairly well the signals in the communication line, we use the mathematical formula to evaluate the algorithm's capabilities in the presence of noise.

First, a database, which consists of all the 128 combinations of 7 consecutive bits was generated for several fiber's lengths (10km, 20km, 30km, 40km, 50km, 60km, and 70km).

Then, a PRBS digital sequence was generated and substituted in Equation 5 to generate an optical signal. Optical noise was added to the sequence with different Signal to Noise Ratios (SNR's).

In the deciphering process, the bit at the center of each sequence in the database is compared to a specific (but random) bit in the noisy and dispersed sequence for evaluating the relative error of each sequence.

Finally, the adjacent neighbors of the chosen bit are reconstructed from the sequence in the data base with the smallest error.

In Figures 6-9, the ability to decode the neighbors' bits for different distances and different SNR's is presented.

The bit number stands for relative distance between the central bit, i.e., the one where the measurements took place, and the adjacent ones, i.e., bit#0 is the measured bit, bit#1 is an average of its two nearest neighbors etc.
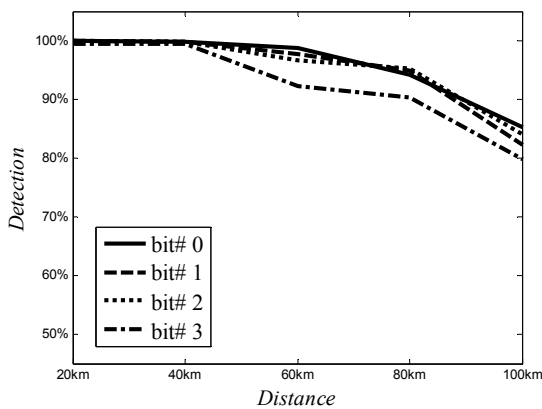

Figure 6. Detection statistics for 2048 bits and SNR=23.5 for different distances
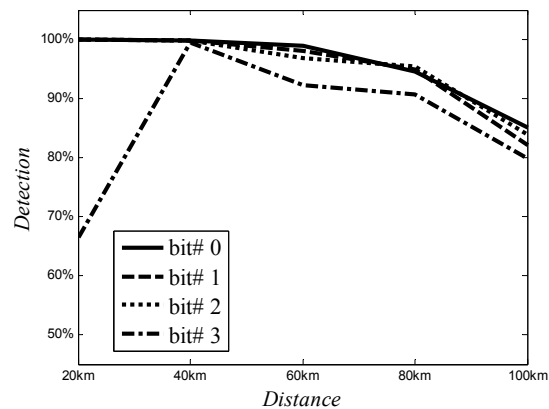
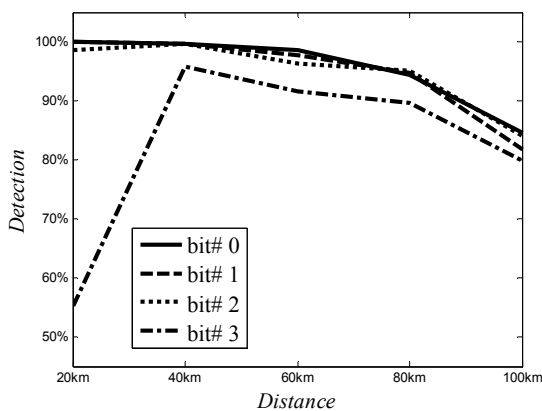Figure 7. Same as Figure 6 but for SNR=13.5
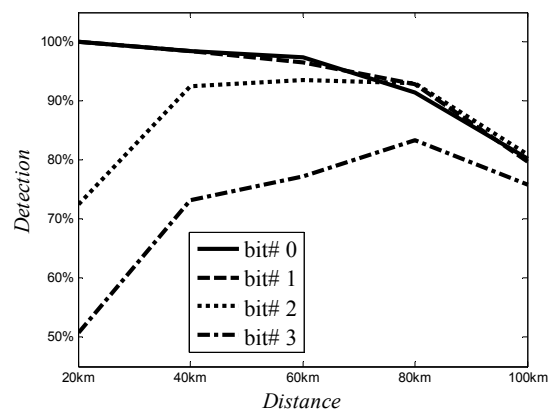
Figure 8. Same as Figure 6 but for SNR=8.8

Figure 9. Same as Figure 6 but for SNR=3.5

As can be seen, this algorithm works reasonably well. Clearly, when the dispersion deformations are smaller than the noise level, it is more difficult to decipher the adjacent data. This is the reason that after 40km (in the low SNR case) it is easier to reconstruct the data than after 20km. Eventually when the deformations are too high (i.e., for

long distances), the boundary effects of the finite combination in the data base are felt, and the reconstruction abilities are reduced again. That is the reason that it is more difficult to reconstruct the data beyond 60 km than after 40 km for the relatively far bits (except for the last SNR=3.5 case, where the optimum point is moved toward the 70 km).

An important conclusion is that the technique is relatively insensitive to the precise shapes of the initial signals in the database. Beyond 20Km the dispersion becomes the major distortion effect.

It should be stressed that for practical purposes the first adjacent (bit#1) is sufficient to prove that eavesdropping is possible. From Figures 6-9 it is shown that decoding of these bits (#1) has proven to be excellent for the first 20km scenario – considerably less than 10^{-3}. In cases where the SNR is not too low (higher than 3.5) the BER increases above 1% only for distances larger than 60km, which is sufficient for PON.

8. Experimental validation of the model and algorithm

To validate the algorithm experimentally, the algorithm was applied on an experimental signal, which was deformed by dispersion. A 2^7 PRBS optical signal at 10Gb/s was generated with carrier wavelength of 1550nm and launched into a 90km of smf28 optical fiber.

In Figure 10 the optical intensity of the deformed signal is plotted.

It should be noted, that this experiment is equivalent for sending a 20Gb/s sequence into 22.5 km of fiber. In Figures 11-13 the detection probability as a function of the bit slot is presented for three different scenarios.

In Figure 10 the measurements were taken only in the time-slot of bit#0, and the adjacent bits# ± 1, ± 2 and ± 3 were reconstructed from the algorithm based on the data base.

As can be seen, the reconstruction of bits ± 2 and ± 3 yielded relatively poor detection probability, however, the detection probability of the nearest (i.e., ± 1) neighbors was higher than 97%.

In Figure 12, bit#0 was reconstructed based on the measurements in both bits +1 and -1. In this configuration 100% detection probability was generated.

Similarly, in Figure 13 the same idea is generalized in decoding the odd bits (+1,-1,+3 and -3) based on the measurements in the even bits (0,+2 and -2). These results show, that this kind of neighbors eavesdropping works with 100% precision within the experiment limitations of BER<10^{-3}.

Clearly, this eavesdropping method can work provided the TDM remains in the optical domain. Conversion to the electrical domain will regenerate the initial digital signal and will eliminate the necessary deformation.

In these experiments the database contained only 128 bits combinations times 4 samples per bit. Hence, only 512 products are required to evaluate the correct combination, which is a task that can easily be done in real time.
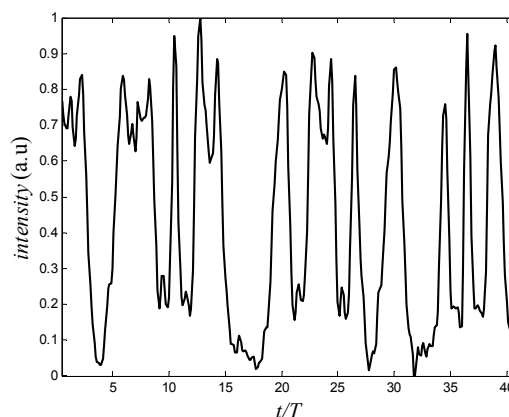

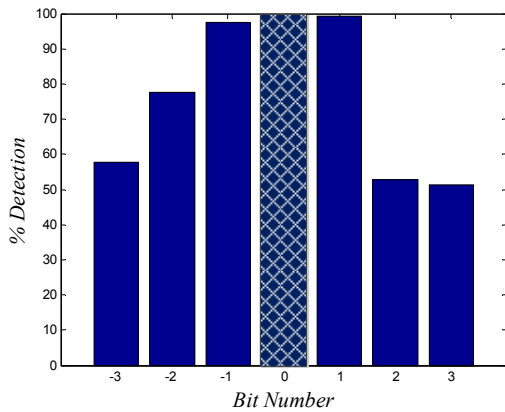Figure 10. The signal intensity of a 2^7 PRBS at 10Gb/s after 90 km of fiber

Figure 11. Detection probability of different neighbors bits, based only on measurements in the bit #0 (marked with crosses) time slot
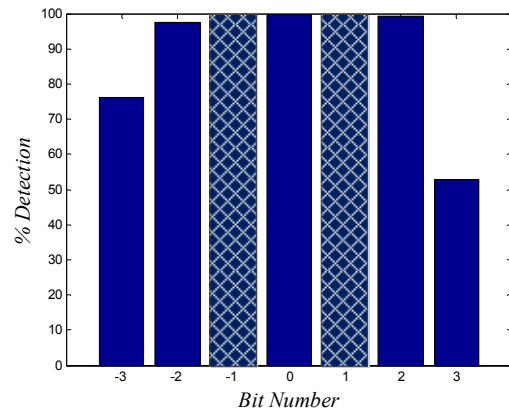
Figure 12. Detection probability of different neighbors bits, based on measurements in the bits +1 and -1 (marked with crosses) time slots

In fact, according to these charts one can identify unknown bits in between two known ones with absolute certainty.
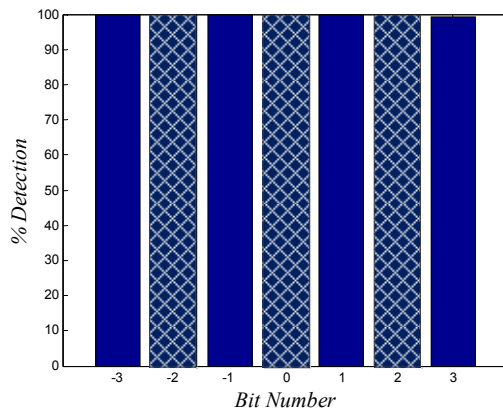

Figure 13. Similarly to Figure 11, the odd bits (+1,-1,+3 and -3) are decoded with 100% accuracy based on the measurements in the even bits (0,+2 and -2).

9. Applying the algorithm to measure the length of a fiber

In fact, there is no reason in assuming a priori knowledge of the fiber's length. Knowing the fiber length in advance saves a lot of time in decoding the neighboring channels. However, if the fiber length is unknown, one can extend the database for different fiber lengths. We can use the same algorithm, but instead of limiting the search to different kinds of bit combinations, we can include other unknowns such as the fiber's length (or the GVD parameter). Obviously, this method can be utilized in evaluating the fiber's length in any network *without the need of sending control signals*. Modern networks are highly dynamic. Frequently, they evolve faster than the network documentation and in many cases the network even changes during operation. Hence, the network manager requires a tool to evaluate the fiber length, or equivalently, the fiber's overall dispersion, between specific points. The most ubiquitous methods consist of control signals (such as specific tones), which are monitored at the other end of the channel. The main problem, of course, with these methods is that they utilize part of the spectral bandwidth, and thus hinder the channel capacity. In addition, these methods require access to both ends of the fiber, which in many cases this is an unattainable requirements.

The method presented here solves these problems completely, since the data signal at the end of the channel is compared to the pre-calculated database (instead of to a control signal), and therefore, this method can work

continuously and simultaneously with the normal operation of the channel, where its full capacity is used. Moreover, it does not require access to both ends of the fiber.

As was said above, even the offset can be determined by the algorithm. Clearly, the offset is already determined in a TDM system; however, this algorithm can work even in a system, where there is no a-priori knowledge of the data sequence transmitted and therefore the offset is undetermined. This feature is important in cases, where there is no knowledge on the characteristics of the fiber line, no preliminary compensation is used, and the data BER is as high as 0.5.

The new database is therefore considerably larger, it consists the intensity $I_p^b(z)$ as a function of the combination (b, for all the binary combinations of $2N+1$ bits), the bin number (p, i.e., the measurement number in a single bit) as in the previous case, but also as a function of the fiber's length (i.e., this is a 3D database), i.e.,

$$I_p^b(z) = \left| A(z, t_p) \right|^2 = \left| \sum_{n=-N}^{N} x_n^b \, \mathrm{srect}_{\xi T}(p\delta t - nT, \Delta, z) \right|^2 =$$

$$\sum_{n=-N}^{N} \sum_{q=-N}^{N} x_q^b x_n^b \, \mathrm{srect}_{\xi T}(p\delta t - qT, \Delta, z) \mathrm{srect}_{\xi T}(p\delta t - nT, \Delta, z)^* \tag{16}$$

For $p = 0, \pm 1, \pm 2, \ldots \pm M$, where, for example, $M = \mathrm{round}(T / \delta t)$.

Since b is the number of the specific binary combinations, then $0 \le b \le 2^{2N+1} - 1$.

Now let us assume that we similarly measure the incoming signal at $2M+1$ different times, but there is a small offset $t_r \equiv r\delta t$ between the two (smaller than a bit period) i.e.,

$$I_{p,r}^{in} = \left| A(z, t_r + t_p) \right|^2 \tag{17}$$

Then an error function can be determined

$$E_r^b(z) = \sum_{p=-M}^{M} \left(I_{p,r}^{in} - I_p^b(z) \right)^2 \tag{18}$$

which is a function of the offset r, the combination number b and the fiber's length z. The parameters, for which $E_r^b(z)$ is at minimum' determines the correct time offset r, the length of the fiber z, and even the specific combination b.

A schematic presentation of the process is illustrated in Figure 14. The minimum of the error function determines the fiber's length. In Figure 13 the process, by which comparison between measured and the database takes place, is schematically presented. In all cases, the center bit (where the data is measured) is compared to the central bit in every combination in the entire database.

In Figure 15 the simulation results are presented, where the minimum error $E(z) = \min_{b,r} E_r^b(z)$ (i.e., for the best combination number b, and best offset r) is presented as a function of the fiber's length. As can be seen for the figure, the fiber's length can thus be estimated from the algorithm.

 In Figure 16 we demonstrate the minimum error seeking algorithm for a real experimental setup. In this setup the data stream was sent at 10Gb/s into approximately 90km of optical fiber. Since the data was sampled at 80GS/s, every bit was oversampled at 8 different places. As can be seen from four different experimental results the fiber's length can be correctly estimated. After applying the algorithm for a large number of different places in the data stream the fiber length was measured to be $z = 86 \pm 4km$.

It should be noted that the main calculation time consumption is the generation of the database, which is needed to be created only once. After the database generation is complete, the comparison itself takes less than a couple of seconds on an ordinary pc with a matlab platform.

The database contained (128 bit options) x (4 samples per bit) x (~100 fiber's lengths) elements, i.e., only 51,200 products are needed to evaluate the length of the fiber, which can easily be done in less than a second. Since the fiber length does not have to be measured in real time, this time scale is practically instantaneous.

The object of this investigation was to demonstrate the feasibility of the idea, and therefore, there was no attempt to improve the process or to make it more efficient. Nevertheless, the deciphering, and comparison times are sufficiently low for most practical purposes, and they can be reduced considerably with efficient programming.

It should be stressed that despite the fact that this technology was applied in this work to OOK, and despite the fact that in the non-coherent method only intensity is measured, this technology (even the non-coherent one) is not restricted to the OOK protocol, and, in principle, it should work on the BPSK and QPSK ones as well. The reason

for that is that the dispersion operates on the field, and therefore differences in the initial field's phase will contribute differently to the intensity profile at the end of the fiber. Any phase sequence will have eventually a different footprint on the intensity.

However, this technology will not work on Polarization Division Multiplexing, since the fiber does not maintain polarization, and cannot be detected in a simple intensity detector.
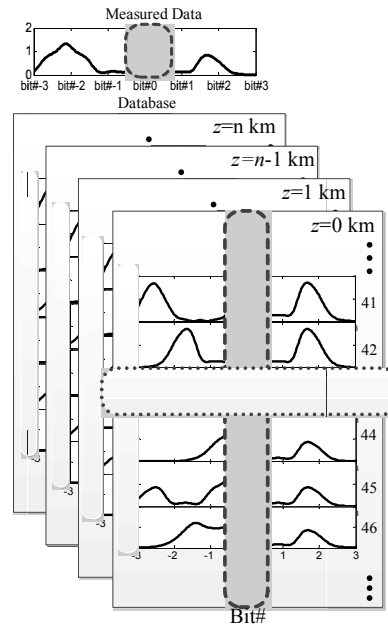

Figure 14. Schematic view of the comparison between the measured signal (upper panel) to the database (lower panel) for each distance
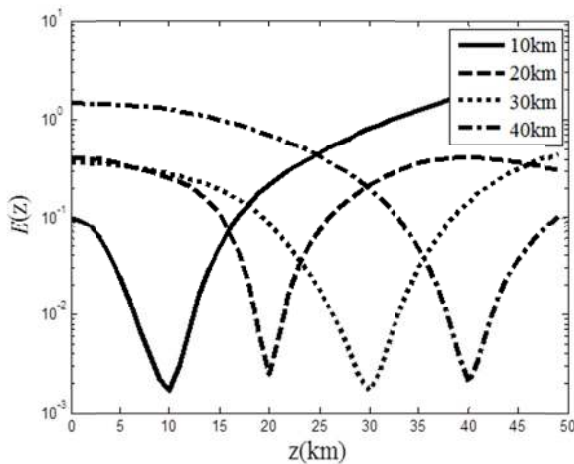

Figure 15. The error as a function of the fiber's length, for different simulation scenarios. In the legend the real fiber's length is presented
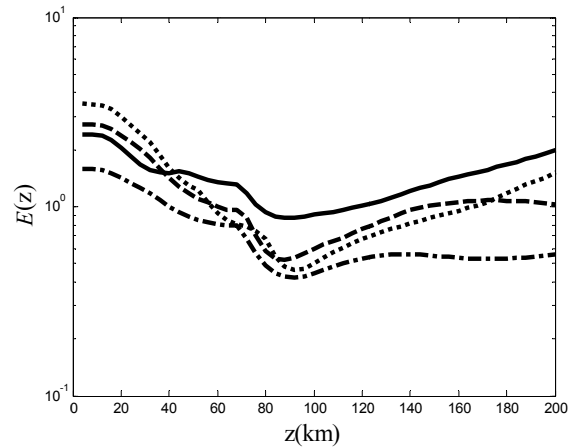

Figure 16. The error as a function of the fiber's length for the experimental ($z\sim90$km) case for four experimental repetitions

10. Summary

The dispersion in optical communication fibers was utilized as a tool to decode data of nearby TDM channels. It was shown that in a TDM scenario, where only a specific time channel is measured, then after traveling through many kilometers of optical fiber, not only is it possible to determine the bits of the specific measured channel, but also to determine the data sent to its adjacent time channels. This capability was demonstrated with a noisy

simulation as well as in an experimental case. Hence, unlike other tapping technologies, the proposed technique does not required cumbersome tapping equipment, which makes it an affordable technique.

Moreover, it was shown that a similar algorithm can be used to measure the length of an optical communication channel without the need for a control (or pilot) signal generation and detection. In fact, by comparing the received data of an ordinary OOK signal sequence to a database comprising different combinations for different fiber's length, the exact length of the fiber can be retrieved. This method can be utilized as a low cost dispersion measuring technique, which unlike other methods can be operated in a running network without deteriorating its performances.

Acknowledgement

References

Abramowitz, M., & Stegun, A. (1965). *Handbook of Mathematical Functions*. New-York: Dover publication.

Antoniades, N., Ellinas, G., & Roudas, I. (2012). *WDM Systems and Networks: Modeling, Simulation, Design and Engineering (Optical Networks)*. Springer Science. http://dx.doi.org/10.1007/978-1-4614-1093-5

Betti, S., De Marchis, G., & Iannone, E. (1995). *Coherent Optical Communications Systems*. Wiley Series in Microwave and Optical Engineering.

Cherkassky, V., & Mulier, F. (1998). Learning from Data: Concepts, Theory and Methods. New-York: John Wiley & Sons.

Cvijetic, N., Huang, M. F., Ip, E., Shao, Y., Huang, Y. K., Cvijetic, M., & Wang, T. (2011). 1.92 Tb/s coherent DWDM-OFDMA-PON with no high-speed ONU-side electronics over 100km SSMF and 1: 64 passive split. *Optics express*, *19*(24), 24540-24545. http://dx.doi.org/10.1364/OE.19.024540

Derickson, D. (1998). *Fiber Optic Test and Measurement*. Prentice Hall, Upper Saddle River, NJ.

Desurvire, E., Bayart, D. O. M. I. N. I. Q. U. E., Desthieux, B. E. R. T. R. A. N. D., & Bigo, S. É. B. A. S. T. I. E. N. (2002). *Erbium-doped fiber amplifiers* (pp. 207-306). John Wiley.

Fernando, J. (2001). *Two methods measure chromatic dispersion*. EDN Network. Retrieved from http://www.edn.com/design/test-and-measurement/4386390/Two-methods-measure-chromatic-dispersion

Granot, E., & Marchewka, A. (2005). Generic Short-Time Propagation of Sharp-Boundaries Wave Packets. *Europhysics Letters, 72*, 341-347. http://dx.doi.org/10.1209/epl/i2005-10264-2

Granot, E., & Marchewka, A. (2011). Emergence of currents as a transient quantum effect in nonequilibrium systems. *Physical Review A, 84*, 032110-032115. http://dx.doi.org/10.1103/PhysRevA.84.032110

Granot, E., Luz, E., & Marchewka, A. (2012). Generic pattern formation of sharp-boundaries pulses propagation in dispersive media. *J. Opt. Soc. Am. B, 29*, 763-768. http://dx.doi.org/10.1364/JOSAB.29.000763

Grishachev, V. V. (2012). Detecting Threats of Acoustic Information Leakage through Fiber Optic Communications. *Journal of Information Security, 3*, 149-155. http://dx.doi.org/10.4236/jis.2012.32017

Hakimi, F., & Hakimi H. (2001). Measurement of optical fiber dispersion and dispersion slope using a pair of short optical pulses and Fourier transform property of dispersive medium". *Opt. Eng. 40*(6), 1053-1056. http://dx.doi.org/10.1117/1.1369144

Hayes, J. (ND). *How To Tap Fiber Optic Cables*" The Fiber Optic Association, Inc. Retrieved from http://www.thefoa.org/tech/ref/appln/tap-fiber.html

Iqbal, M. Z., Fathallah, H., & Belhadj, N. (2011, December). Optical fiber tapping: Methods and precautions. In *High Capacity Optical Networks and Enabling Technologies (HONET), 2011* (pp. 164-168). IEEE.http://dx.doi.org/10.1109/HONET.2011.6149809

Jung, S. P., Cho, K. Y., Takushima, Y., & Chung, Y. C. (2010). *Recent Progresses in Coherent WDM PON Technologies*. Tu.A1.3, ICTON 2010.

Lavery, D., Maher, R., Millar, D., Thomsen, B. C., Bayvel, P., & Savory, S. (2012, March). Demonstration of 10~ Gbit/s colorless coherent PON incorporating tunable DS-DBR lasers and low-complexity parallel DSP. In *Optical Fiber Communication Conference* (pp. PDP5B-10). Optical Society of America.

Marchewka, A., Granot, E., & Schuss, Z. (2007). Short time propagation of a singular wave function: some surprising results. *Optics and Spectroscopy, 103*, 330-335. http://dx.doi.org/10.1134/S0030400X07080255

Noh, Y. O., Kim, D. Y., Oh, S. K., & Pack, U. C. (1999, August). Dispersion measurement of a short length optical fiber using Fourier transform spectroscopy. In *Lasers and Electro-Optics, 1999. CLEO/Pacific Rim'99. The Pacific Rim Conference on* (Vol. 3, pp. 599-600). IEEE. http://dx.doi.org/10.1109/CLEOPR.1999.817738

Prat, J. (2010). *Next-Generation FTTH Passive Optical Networks: Research Towards Unlimited Bandwidth Access*. Springer Science.

Ramachandran, S. (2010). *Fiber Based Dispersion Compensation (Optical and Fiber Communications Reports)*. Springer-Science.

Shieh, W., & Djordjevic, I. (2010). *OFDM for Optical Communications*. Elsevier Inc.