Journal of Universal Computer Science, vol. 23, no. 5 (2017), 454-478
submitted: 19/12/16, accepted: 17/5/17, appeared: 28/5/17 © J.UCS

Steganography in VLC Systems

Grzegorz Blinowski
(Institute of Computer Science, Warsaw University of Technology, Warszawa, Poland
g.blinowski@ii.pw.edu.pl)

Krzysztof Szczypiorski
(Institute of Telecommunications, Warsaw University of Technology, Warszawa, Poland
ksz@tele.pw.edu.pl)

Abstract: One of the areas in which wireless networks based on visible light communication (VLC) is considered superior to traditional radio-based communication is security. The common slogan summarizing VLC security features is WYSIWYS - "What You See Is What You Send". However, especially in the case of infrastructure downlink communication, security must be accounted for also in case of VLC. In this paper, we analyze the possibilities of implementing a steganographic channel in VLC systems on the physical (PHY) and medium access control (MAC) layers. We introduce new steganographic techniques for VLC: hiding data in dimming patterns and color visibility dimming (CVD) frames, and adapt methods previously proposed for radio networks, such as using unused header bits and pseudo-corrupted frames to hide data. We use the IEEE 802.15.7 standard as a reference model for protocol structure, but the steganographic techniques that we describe can also be used in networks based on different standards.

Keywords: VLC, visible light communication networks, steganography, physical layer security
Categories: C.2.1, C.2.2

1 Introduction

Visible light communication (VLC) is a wireless optical communication technology employing the visible light spectrum from 380 to 780 nm. The decreasing cost and hence rapid adaptation of LED-based light make VLC a promising communication technique and a significant alternative to radio-based wireless communication. Wi-Fi, Bluetooth and similar "traditional" radio-based communication systems that use bandwidth below 6 GHz suffer from limited channel capacity and transmission rate due to the limited radio spectrum available. At the same time, user requests for data transmission throughput and availability continue to increase. With a potential 300 THz bandwidth, VLC data transmission networks provide an attractive alternative to traditional wireless techniques. Multi-gigabit-per-second data rates are available over short distances, with arrays of LEDs in multiple-input multiple-output (MIMO) configurations. With VLC technology, gigabit-per-second data rates can be provided in conjunction with lighting with only simple LEDs and photodetectors (PDs) - compared to expensive radio solutions that require high power consumption for transmitting, sampling, and processing RF signals.

VLC has been proposed both for indoor and outdoor applications (see [Hranilovic, 13] and [Tsiatmas, 14]). Indoor VLC applications include office communication [Rahaim, 11], multimedia conferencing [Chen, 14], peer-to-peer data exchange, multimedia broadcasting (home-audio and video streams) [Langer, 08; O'Brien, 08; O'Brien, 09], and positioning [Yoshino, 08; Ren, 14]. The majority of commercial VLC systems currently available provide data broadcasting services and include solutions for museums, shopping centers, exhibition centers, airports and train stations as well as accessibility for disabled persons. General two-way communication solutions based on VLC are currently the subject of intensive study, especially as a component of hybrid communication systems [Ayyash, 16]. VLC systems provide a safe alternative to electromagnetic interference from radio frequency communications in hazardous environments such as mines and petrochemical plants, and in applications where traditional WLAN communication may interfere with specialized equipment, for example in hospitals and in in-flight entertainment systems in aircraft passenger cabins (where an additional benefit is the reduced weight of cabling and the potential for integration with passengers' own mobile devices) – see [Ghimire, 12].

One of the features in which VLC techniques are considered superior to traditional radio-based communication is security. The directivity and high obstacle impermeability of optical signals are considered to guarantee a secure way of transmitting data within an indoor environment, making the data difficult to intercept from outside. The common slogan summarizing VLC security features is "What You See Is What You Send" (WYSIWYS) [Conti, 08]. However, as recent history has taught us, a common mistake in the development of novel communication techniques is to neglect or downplay security issues. Such was the case with the internet protocol suite (both on the network and application layer), fiber-optic based networks, and more recently radio-based wireless networks. Currently the VLC industry seems to be on the same path again: the indubitable "pro-security" physical characteristics of VLC have steered the developers' focus away from security issues. Yet the shared nature of the medium allows wireless VLC networks to be easily monitored and broadcast on. Attackers can not only gain access to the communication, but also launch various attacks (e.g. jamming or denial of service). Ensuring the security of VLC up to now has been mainly tackled with respect to the physical layer from the information theory point of view [Classen, 16]. To our best knowledge there has been very little work on other aspects of VLC security - in [Classen, 15] transmission "snooping" was addressed, while in [Blinowski, 16a] the risk of various forms of attacks on VLC networks was classified.

In this paper we address a specific security issue of VLC – the transmission of hidden data, i.e. the issue of steganography on the physical (PHY) and media access control (MAC) layers of VLC. As far as VLC standards are concerned, we will refer to the IEEE 802.15.7 standard [IEEE, 11]; however, our discussion should also be relevant to other proposed VLC techniques not covered by the current IEEE norm.

According to [Zielinska, 14] "steganography […], is the art of embedding secret messages (steganograms) in a certain carrier, possibly to communicate them in a covert manner." Network steganography is the family of methods that uses telecommunication protocols as a carrier for hidden data. These methods [Mazurczyk, 16] utilize modifications of packets to perform covert communication by modifying the structure of the packet (e.g the payload or protocol-specific fields) or modifying

time relations among packets (like changing the sequence of the packets or inter-packet delays). Wireless technologies are an excellent environment for applying network steganography as they have a relatively huge bandwidth that is ready to use, especially at the PHY layer.

The purpose of this paper is to propose the widest possible spectrum of steganographic methods that can be used for VLC communication – ranging from the modulation to the MAC level. For each introduced method, we will discuss the potential for bandwidth and undetectability. We do not address the issue of practical implementation in this paper as our focus is on covering different steganographic methods: some adapted from wireless radio-based communication, and some introduced by us originally for VLC. However, we would like to note that we have developed a steganographic system based on VLC which will be presented in a separate paper (see section 3.3 for more details).

The structure of this paper is as follows: in section 2 we describe the basics of the PHY layer of VLC technology – modulation as well as dimming and anti-flicker mechanisms – which we subsequently use in our proposed steganographic methods. In section 3 we discuss the possibilities of implementing hidden data transmission by changing modulation parameters. In sections 4 and 5 we discuss (respectively) how PHY and MAC frames, as defined by the IEEE 802.15.7 standard, can be used for steganographic purposes. In section 6 we discuss and summarize the undetectability of proposed scenarios. Section 7 summarizes the paper and outlines the areas of future research.

2 Modulation, dimming and anti-flicker mechanisms in VLC

2.1 VLC system components and modulation techniques

A VLC PHY layer consists of two major elements: the transmitter and the receiver. Their properties are as follows:

Transmitter – Two types of white-light LEDs are used in solid-state lighting: 1) red-green-blue (RGB) emitters; 2) blue LEDs with yellow-light emitting phosphor layers ("single-chip" LEDs). The VLC transmitter may use both types, but the second type is more widespread in illumination due to its energy efficiency and lower complexity. Different types and form factors of LEDs are employed in various environments: high power LEDs or LED arrays are the choice for typical indoor illumination purposes, while low-power devices are used in mobile appliances.

The slow response of yellow phosphor-based white light LEDs limits their spectral component bandwidth to 2MHz and hence their transfer rate to approximately 10 Mbps. When the yellow component is filtered out at the receiver and only the blue component is detected, a bandwidth of 8 MHz throughput in the range of 30-40 Mbps may be attained. By combining a simple pre- and post-equalization, 75 Mbps of throughput can be achieved. Data throughput of up to 100-230 Mbps has been demonstrated in a single-emitter–single-receiver scenario and On-Off Keying (OOK) [Langer, 11]. Higher data rates of about 1 Gbps are also attainable with more advanced modulation techniques such as Discrete Multitone (DMT) and Orthogonal

Frequency Division Multiplexing (OFDM). Similar data rates have also been attained with arrays of separately driven light sources [Azhar, 13].

The receiver collects and then concentrates the incoming light onto a photo-detecting element. Both imaging and non-imaging receivers are used. Photocurrent generated in the photo-detector is amplified and fed to the D/A circuitry. Currently in devices such as smartphones, tablets, etc., low cost photodiodes or typical optical sensors are used as photodetectors for the VLC channel. With current technology achieving sufficient photo-detector sensitivity, bandwidth is not limited in the receiver (the transmitter and channel loss and dispersion are the major elements limiting the bandwidth).

The critical difference between VLC and radio-based communication is that in VLC, data cannot be encoded in the phase of the light signal. The information has to be encoded in the varying intensity of the emitted light. Demodulation depends on direct detection at the receiver, which is why Intensity Modulated/Direct Detection (IM/DD) modulation techniques are used in VLC. Modulation in VLC must also take into account the requirements of dimming and flicker mitigation. Various modulation schemes have been proposed for VLC systems, including

- OOK – the data bits 1 and 0 are transmitted by turning the LED on and off respectively. In the "0" state, the LED is not completely turned off but rather the light intensity is reduced. The advantages of OOK include its simplicity and ease of implementation.
- Pulse Width Modulation (PWM) – the widths of the pulses are adjusted based on the desired level of light dimming while the pulses themselves carry the modulated signal in the form of a square wave.
- Pulse Position Modulation (PPM) – the position of the pulse in a series of pre-defined time-slots identifies the transmitted symbol.
- Variable Pulse Position Modulation (VPPM), proposed specifically for VLC systems, is a combination of PPM and PWM. As in PPM, VPPM optical symbols are distinguished by pulse position, and pulse width may be changed to achieve different dimming levels (see Figure 1).
- OFDM – the communication channel is divided into multiple orthogonal subcarriers and data is sent in parallel substreams modulated over the subcarriers. Standard "radio-based" OFDM techniques need to be adapted for application in IM/DD techniques because OFDM generates complex-valued bipolar signals which need to be converted into real values.
- Frequency Shift Keying (FSK) – the instantaneous frequency of a constant amplitude carrier signal is changed between two (for BFSK) or more (for MFSK) values by the baseband digital message signal. This type of modulation is typically used in simple LED-ID systems [Blinowski, 16b].
- Color Shift Keying (CSK) requires a system with RGB emitters. CSK is similar to FSK in that the bit patterns are encoded as color (wavelength) combinations. For an illustration of the principle of CSK, see Figure 2.

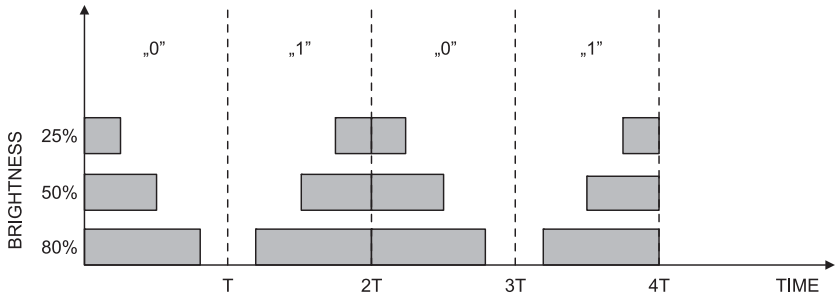


Figure 1: The principle of VPPM: pulse position indicates whether "0" or "1" is transmitted, pulse width is used for dimming control.

Current VLC systems based on white LEDs use either OOK or VPPM, and the IEEE 802.15.7 standard is based on these two modulation schemes. For RGB emitters, the same standard is based on CSK.

CSK is specifically designed for VLC. In CSK, the combined intensity of the three bands of an RGB LED is held constant while their relative intensities vary. CSK relies on the color space chromaticity diagram as defined by CIE, 1931 [CIE, 31]. The chromaticity diagram maps all colors perceivable by the human eye into two chromaticity parameters – x and y. The entire human visible wavelength is divided into seven bands (Figure 2). The CSK signal is generated using three color light sources out of the seven available color bands – depending on the light emitters used. The principles of CSK are as follows:

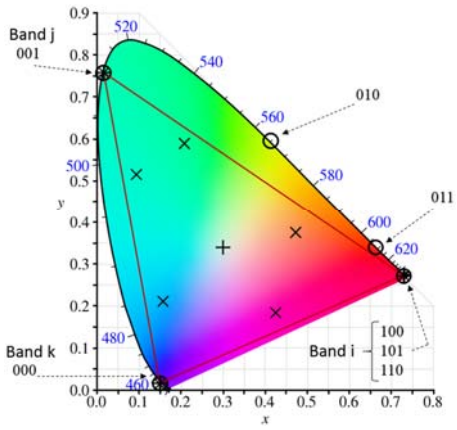
1. The three vertices of the CSK constellation triangle are decided by the center wavelength of the three color bands on *x*,*y* color coordinates.
2. Data bits are mapped to chromaticity (*x*,*y*) values. Depending on the modulation scheme used (4CSK, 8CSK, 16CSK) 4, 8 or 16 symbols are mapped to points inside the constellation triangle. Each symbol is assigned a bit sequence (for example in 4CSK: 00, 01, 10, and 11). Symbols are equally distributed in the triangle so that the combined light emitted when transmitting different symbols is perceived by the human eye to be white. In the transmitter, a scrambler is used to create a pseudo-random data stream and prevent data-pattern-dependent color shifts.
3. The intensities of RGB LEDs driven by transmitter D/A circuitry must be determined. The individual intensities of the three LEDs (P_i, P_j and P_k) for each symbol is calculated by solving the following equations:

$$x_s = P_i x_i + P_j x_j + P_k x_k \tag{1}$$

$$y_s = P_i y_i + P_j y_j + P_k y_k \tag{2}$$

$$P_i + P_j + P_k = 1 \tag{3}$$

where x_s and y_s are chromaticity values of the symbol and (x_i,y_i), (x_j,y_j), (x_k,y_k) are constants determined by the chromaticity values of the central wavelength of the LEDs used.



Constellation triangle coordinates	4CSK symbol coordinates	8CSK symbol coordinates
(0.734,0.265)	[0 0]: (0.011,0.733)	[0 0 0]:(0.064,0.491)
(0.011,0.733)	[0 1]:(0.305,0.335)	[0 0 1]:(0.188,0.237)
(0.169,0.007)	[1 0]:(0.169,0.007)	[0 1 0]:(0.470,0.366)
	[1 1]:(0.734,0.265)	[0 1 1]:(0.452,0.136)
		[1 0 0]:(0.011,0.733)
		[1 0 1] (0 169,0.007)
		[1 1 0]:(0.252,0.577)
		[1 1 1]:(0.734, 0.265)

Figure 2: CIE color space chromaticity diagram. The outer boundary is the spectral (or monochromatic) locus, with wavelengths (blue digits) shown in nanometers. The circles mark seven color bands defined for CSK. An example constellation triangle for codes (110, 001, 000) is drawn with solid red lines. Crosses show the locations of symbols used in 4CSK, while circles and "X-s" mark the locations of 8CSK symbols. (Source of the CIE diagram: https://en.wikipedia.org/wiki/CIE_1931_colour_space #/media/File:CIE1931xy_CIERGB.svg CC)

2.2 Overview of dimming mechanisms available in VLC

In terms of VLC infrastructure, the transmitter must provide the basic function of visibility, i.e. an adequate level of illumination for its surroundings, at all times, even during "idle periods," when no data is being transmitted. It must also be able to provide light dimming. In VLC, light dimming is the variation of the perceived brightness of the light source according to the user's requirements. Although it is necessary for VLC infrastructure to provide visibility and dimming, it is inconvenient from the data communication point of view, as it potentially reduces the obtainable

transfer rate and makes system implementation more troublesome. On the other hand, mechanisms which are used to achieve proper visibility and light dimming may also be used to transfer hidden data (see [Section 3]).

2.2.1 Duty cycle and signal intensity dimming

Light dimming in VLC systems can be provided by two general mechanisms:

1. Changing the signal's duty cycle - an arbitrary level of illumination is obtained by modifying the number and width of modulated pulses.
2. Changing the signal's intensity – the transmitted power of luminaires is varied according to dimming requirements.

The combination of both of the above methods may also be used. The general principle of achieving dimming by altering the duty cycle is shown in Figure 3.

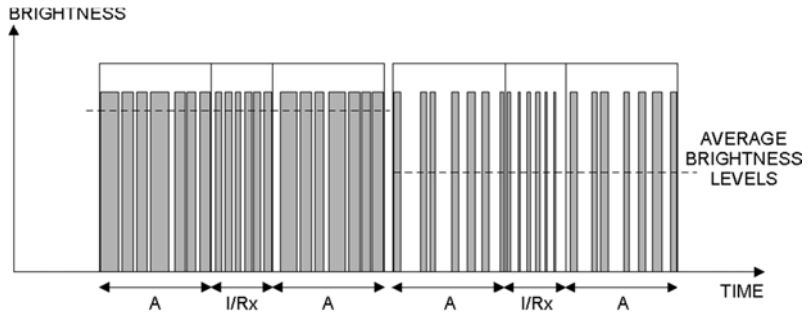


Figure 3: Dimming patterns and data duty cycle adapted for high and low brightness. "A" indicates active transmitter periods and "I/Rx" indicates idle periods or receive mode – during this time an idle pattern is transmitted to ensure proper lighting.

2.2.2 Idle patterns and compensation time dimming

An idle pattern must be inserted between data frames for light dimming and to avoid flicker. The duty cycle of the idle pattern can be varied to provide brightness variation. An idle pattern can be transmitted either in-band, using the same optical clock speed that is used to transmit data, or out-of-band. An in-band idle pattern does not require any change in the clock and is "seen" by the receiver. An out-of-band idle pattern is typically sent at a much lower optical clock rate (but high enough to avoid flicker), and can even maintain visibility via a DC bias only. An out-of-band pattern does not lie in the receiver's modulation-domain bandpass, hence is not seen by the receiver.

Compensation time dimming is implemented either by turning the light source "ON" or "OFF," with no modulation, for a specified period of time, or by transmitting a dimming pattern with a specific duty cycle. Compensation time periods can be inserted into either the idle pattern or the data frame to decrease or increase the average brightness of a light source (refer to the discussion in the next sections and see Figures 4 and 5).

With regard to using the above described dimming mechanisms as a potential base for steganographic transmission, one important note must be made: lowering the signal's intensity reduces the transmitter range while the bit rate remains constant, and the insertion of compensation time results in a lower bit rate as the light dims, which implies a reduced bit rate with constant range.

2.3 Dimming in the IEEE 802.15.7 standard

2.3.1 Intra-frame and modulation based dimming

The IEEE 802.15.7 standard allows three PHY types and three modulation schemes – OOK, VPPM and CSK. These are summarized in Table 1. Each PHY type contains mechanisms for light signal modulation, run length limited (RLL) coding, and channel coding for forward error correction (FEC). RLL coding is used to avoid long repeated sequences of 0s and 1s that may cause flicker and clock recovery problems. Moreover, depending on the modulation scheme, different dimming and flicker elimination mechanisms are available.

Modulation scheme	RLL code	Optical clock rate range	Data rate range	Comments
PHY I				
OOK	Manchester	200-400 kHz	11.67 – 100 kbps	PHY I is intended for outdoor usage with low data rate applications under potentially high ambient interference levels.
VPPM	4B6B	400 kHz	35.56 – 266 kbps	
PHY II				
VPPM	4B6B	3.75 – 7.5 MHz	1.25 – 5 Mbps	PHY II is intended for indoor usage with moderate data rate applications.
OOK	8B10B	15 – 120 MHz	6 – 196 Mbps	
PHY III				
4-CSK 8-CSK	n/a	12 MHz	12 – 18 Mbps	PHY III is intended for applications using CSK that have multiple light sources and detectors.
4-CSK 8-CSK 16-CSK	n/a	24 MHz	24 – 96 Mbps	

Table 1: IEEE 802.15.7 PHY Types: modulation, coding and data rate

Three dimming mechanisms are provided for by the IEEE 802.15.7 standard: visibility patterns, compensation time and pulse width adjustment for VPPM modulation.

Figure 4 shows the principle of OOK dimming: because DC-balanced RLL codes (Manchester or 8B10B) are used for data transmission with this type of modulation, the average brightness for PHY frames is always close to 50% of the pulse brightness. To achieve the desired brightness, the standard allows: (1) 0s and 1s levels of intensity to be modified or (2) the average duty cycle of the signal to be changed with the insertion of idle patterns and compensation symbols. Idle patterns (marked as "IDL" on the figure) are inserted between frames. The PHY frame may also be split into subframes, each proceeded by a resync field and with compensation symbols added between subframes. In effect, an average brightness of N% is maintained.

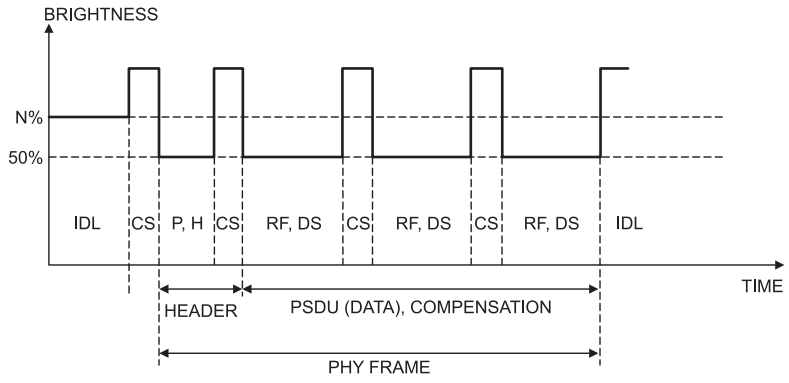


Figure 4: An example of OOK dimming for brightness control. A single PHY frame is shown with idle patterns transmitted before and after the frame. IDL – idle patterns, CS – compensation symbols, P – preamble, H – header, RF – resync field, DS - data

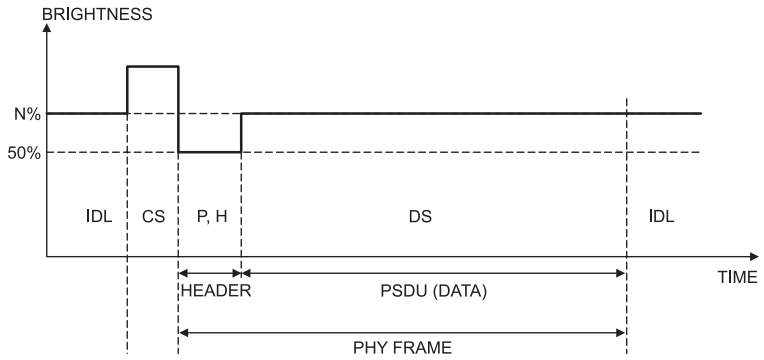


Figure 5: An example of VPPM dimming for brightness control. A single PHY frame is shown with idle patterns transmitted before and after the frame. IDL – idle patterns, CS – compensation symbols, P – preamble, H – header, DS – data.

With VPPM, dimming in the IEEE 802.15.7 standard is provided for by changing the pulse width. There is no need to insert in-frame compensation symbols with this modulation scheme. Since the preamble and header are always transmitted with a

50% duty cycle, compensation symbols are typically required to be inserted before the frame. Both data and idle patterns are transmitted with VPPM, hence the required brightness is automatically assured.

With CSK, dimming support is easy – simple amplitude dimming is used where the driving current of the LEDs is varied to change the brightness of the resulting white light.

2.3.2 Dimming with color visibility dimming frames

As stated in section 2.2.2, an adequate level of dimming must also be provided when no data is transmitted, i.e. during MAC idle and receive periods. To achieve this, the IEEE 802.15.7 standard provides support for special color visibility dimming (CVD) frames that do not contain data but send only in-line idle patterns. A CVD frame is used for color, visibility and dimming support. It may visually provide information such as communication status and channel quality to the user via various colors and/or light intensity. The CVD frame may also be sent during idle or receive modes of operation for continuous visibility and dimming support. During CVD frame transmission, the device is still emitting light but not communicating, and it is thus able to fulfil its lighting function. The payload of the frame consists of visibility patterns of appropriate intensity and color.

For OOK modulation, the IEEE 802.15.7 standard defines a set of eleven basic low resolution bit patterns to be used for in-band dimming (it should be noted that the usage of these predefined patterns is not a requirement). These low resolution patterns are used to develop high resolution visibility patterns by averaging them across time to generate the required high resolution pattern. The dimming algorithm proposed by the standard defines how many times each of the two pre-defined adjacent visibility patterns should be transmitted. The algorithm always uses one or two 11 bit patterns from the predefined set. When two patterns are interleaved, the adjacent patterns on both sides of the precision requirement are used (for example, a pattern with five 1s for 50% dimming and a pattern with six 1s for 60% dimming are used for a dimming level of 55%). The principle of the algorithm is shown in Figure 6 and table 2, which illustrate pattern choice and pattern count respectively.

The IEEE 802.15.7 standard does not define the sequence of dimming patterns. Since RLL codes are used for data transmission, it must be ensured that there is no conflict between the visibility pattern and the valid RLL code. This requirement reduces the number of bits of data that may be transmitted in one visibility pattern code word, which has important consequences for steganographic transmission and will be addressed later in section 3.4.1.

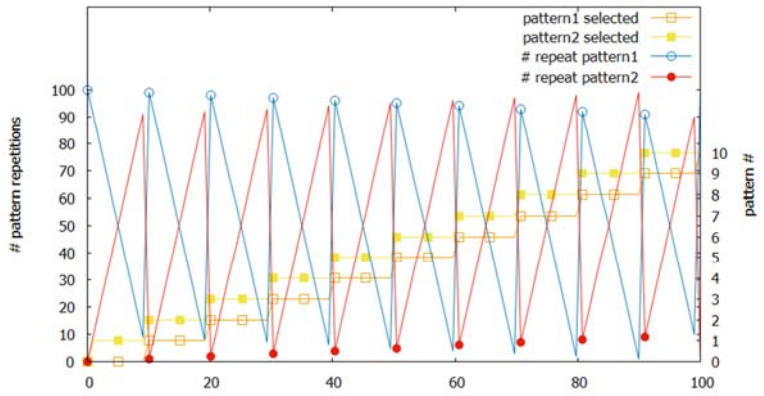


Figure 6: Illustration of dimming pattern choice and repetition count selection algorithm as a function of dimming.

light level	pattern
100%	11111 11111
90%	11110 11111
80%	11110 11110
70%	11101 11100
60%	11001 11100
50%	10001 11100
40%	00001 11100
30%	00001 11000
20%	00001 10000
10%	00001 00000
0%	00000 00000

Table 2: Bit pattern choice for required dimming level, 10% steps as defined in IEEE 802.15.7 standard.

3 Steganography on the physical level

3.1 Steganography via modifying signal intensity

The steganographic scenario that we propose on the PHY level involves varying the peak transmit power to convey the data in the secret channel. As the light changes could be noticed (e.g. as flickering), the overt signal duty cycle should also be modified to assure that the average transmitted power remains constant. This can be done via standard mechanisms provided for dimming: idle pattern modification, compensation time adjustment, etc.

3.1.1 Steganography with OOK

For OOK, the transmit power for "1" remains constant. However, we can introduce additional information to the transmitted signal by varying the maximum signal intensity (see Figure 7). "On" and "Off" levels at normal transmit power are used in standard transmission mode. If steganographic data is introduced into the stream, intermediate signal levels are used to transmit "hidden" symbols. In this example, one additional bit of information may be transmitted with 4-ASK modulation (4 level amplitude shift keying) and two bits with 6-ASK. "Stegano-aware" receivers will recognize and decode additional signal levels, while standard receivers will simply decode additional symbols as 0s or 1s.

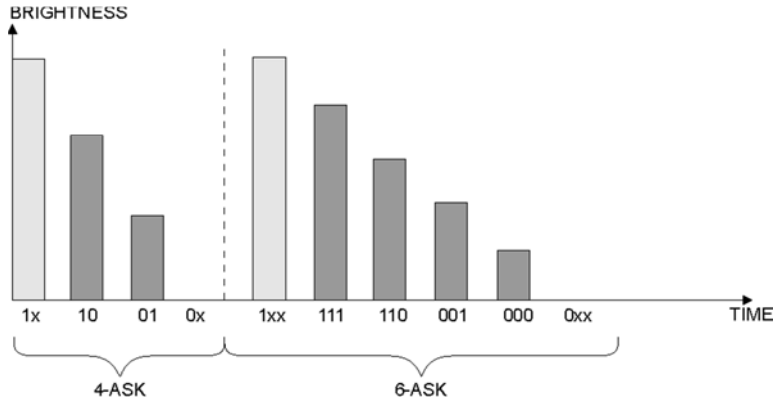


Figure 7: Introducing steganographic symbols into OOK by changing the modulation type to 4-ASK and 6-ASK.

This solution comes with a price: the introduction of higher order n-ASK modulation requires a higher signal to noise ratio (SNR) in the channel to achieve the desired bit error rate (BER). The BER to SNR relation for n-ASK modulation (assuming white Gaussian noise) is given as (4):

$$BER = \frac{M-1}{M \cdot \log_2 M} \cdot Q\left(\frac{1}{M-1} \cdot \sqrt{SNR \cdot \log_2 M}\right) \qquad (4)$$

where M is the number of signal levels and Q is defined by the error function: $Q(x) = \frac{1}{2}\mathrm{erfc}\left({}^{x}/_{\sqrt{2}}\right)$

Sample plots of BER with relation to SNR for different modulation levels are shown in Figure 8. To conclude, in a typical scenario, the maximum physical range of steganographic transmission will degrade faster than the range of a standard OOK transmission.

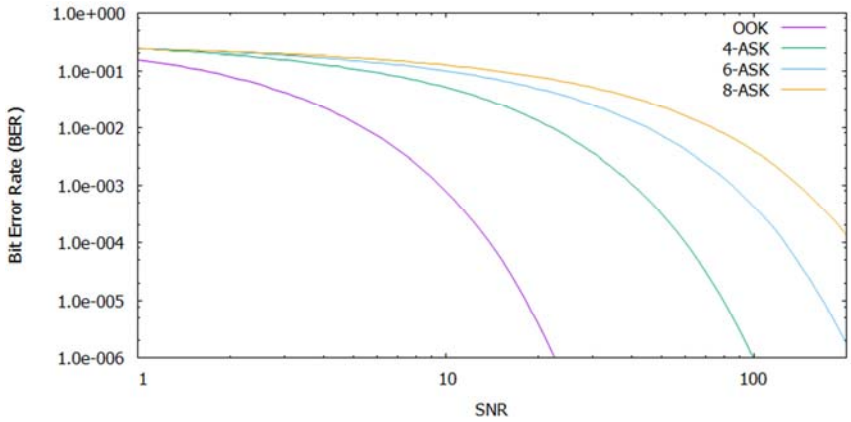


Figure 8: Spectral properties of BER performance of OOK/ASK systems.

3.1.2 Steganography with VPPM

With VPPM, we can use the technique described above for OOK, and we can also introduce additional information to the transmitted signal by encoding it in the pulse length (see Figure 9). In this case, the dimming functions become depreciated and an additional mechanism such as that described above can be used to compensate for the required light intensity level and potential flicker.

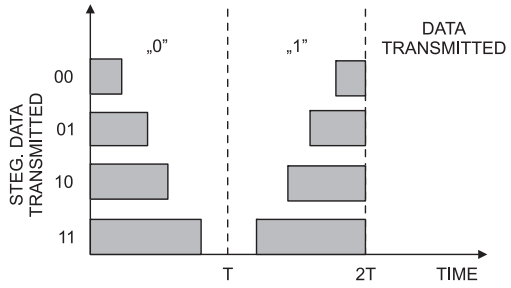


Figure 9: Steganographic data encoded as pulse width in VPPM modulation.

As was the case with OOK/ASK modulation, encoding data in VPPM pulse width also comes with a price: as the signal's duty cycle approaches 0% (or 100%), the requirement for a high SNR grows. Sample BER plots for different duty cycle values are shown in Figure 10. The discussion of an analytical formula for BER in this case is beyond the scope of this paper – we refer the reader instead to [Yoo, 13].

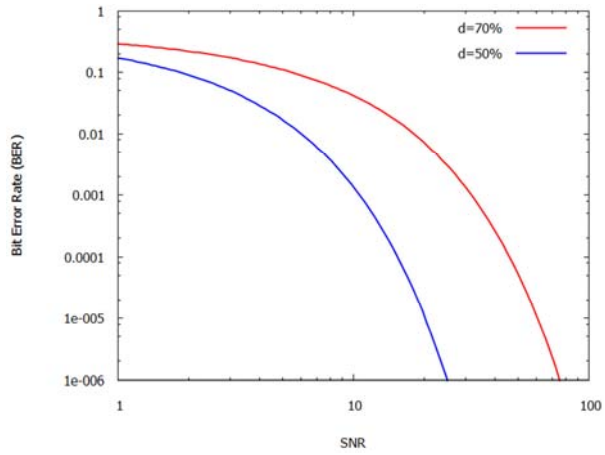


Figure 10: Spectral properties of BER performance of VPPM system depending on duty cycle parameter d. (Line of Sight channel)

3.2 Steganography with CSK

In CSK, a TriLED transmitter is used and the data is mapped to 2, 3 or 4 bit symbols which are transmitted as "color points" on the CIE color plane. Before data transmission takes place, CSK constellation must be established. This is done by estimating the channel propagation matrix [Monteiro, 12]. With this type of modulation, additional information may be introduced to the transmitted signal by using extra symbols from outside the chosen constellation. For example, we may use 4CSK for the overt channel while transmitting hidden data with 8CSK. CSK does not use compensation time or idle pattern dimming and LED transmitter power is linearly varied by D/A circuitry. However, CSK may be combined with other modulation types, for example OOK, ASK, PPM or VPPM to convey hidden data. The downside of such an approach is that it requires serious modifications on the transmitter and receiver side to handle such a combined modulation scheme. Some promising results of such a "hybrid" CSK-PPM modulation in a VLC system can be found in [Pergoloni, 15].

3.3 Steganography with other modulation techniques

Depending on the modulation technique used in the VLC system, other steganographic methods can also be used. In particular, in PPM and the Direct Sequence Code Division Multiple Access (DS-CDMA) approach, steganographic data can be added to the code spreading sequence. A VLC system achieving a 100 Mbps transmission rate and up to 1 Mbps of hidden data transmission is currently being investigated by us and will be described in a separate paper.

3.4 Steganography via dimming and flicker elimination mechanisms

3.4.1 Encoding data in an idle pattern

Dimming mechanisms can also be used as a basis for a covert channel. In OOK, the covert channel can be implemented by tweaking the idle pattern, the compensation time or both. As described in section 2.3.2, the IEEE 802.15.7 standard allows an idle pattern to be inserted between the data frames for light dimming. The standard defines a set of patterns and an algorithm of pattern selection for "high resolution dimming". However, a set of different patterns may be used as long as two conditions are met: (1) an adequate dimming level is maintained; (2) the symbols used in patterns do not conflict with RLL coding.

The number of possible bit patterns in a code word of length N satisfying the required dimming level d can derived as follows. The number of patterns with n bits set to "1" is:

$$dpat(N,n) = \frac{N!}{n!\,(N-n)!} \tag{5}$$

From (5), by substituting parameter d (dimming measure: $0 \leq d \leq 1$) where $d = {}^{n}/_{N}$ we obtain (square brackets indicate nearest integer, i.e. rounding function):

$$dpat(N,d) = \frac{N!}{([dN])!\,([N(1-d)])!} \tag{6}$$

Calculating $\log_2 dpat(N,d)$ gives us the number of bits in a codeword available for steganographic transmission. Figure 11 illustrates this number for 8 and 10 bit coding as a function of the dimming level.

To meet the second condition, we must eliminate RLL codes from allowed patterns. For example, for 8B10B coding, 124 patterns with 4 or 6 bit sets and 192 patterns with 5 bit sets must be eliminated. The number of bits available for coding is shown in Figure 11 as squared line.

The method of hiding data described below has its limitations: (1) it might introduce flicker; (2) clock recovery may not be trivial for certain sequences of code words; (3) data capacity approaches zero for very low and high levels of dimming. To avoid (1) and (2), some additional limitations on the code space must be imposed or additional synchronizing symbols must be introduced (which in turn will limit the steganographic transmission rate). Regarding (3), it should be noted that 0% and 100% dimming does not necessarily imply transmission of all 1s or 0s in CVD frames, nevertheless the data rate (both overt and covert) in this case is seriously limited.

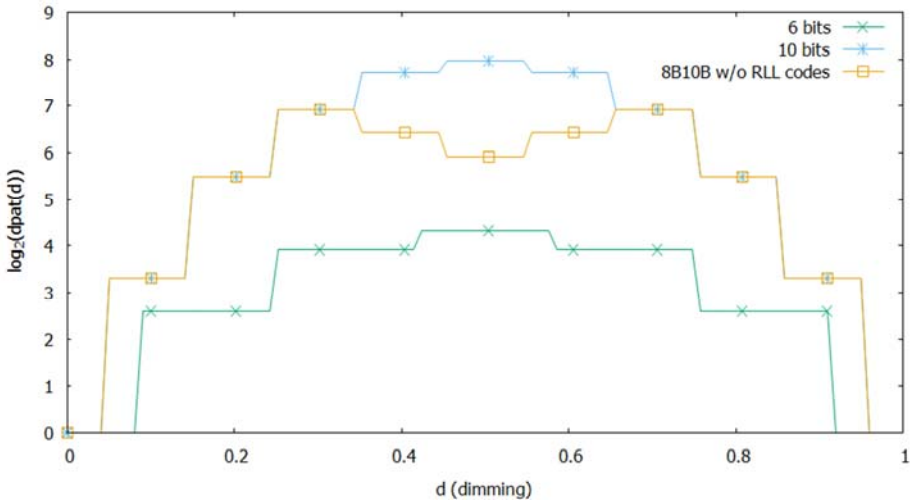


Figure 11: Number of available bit patterns as a function of dimming level (d); for 6 and 10 bit codes, third plot: for 10 bit codes adjusted for the elimination of RLL 8B10B symbols.

3.4.2 Encoding data in a sequence of idle patterns

The idle pattern mechanism provides the additional possibility of hiding covert data with the use of the sequence of patterns. Information may be encoded not in the pattern itself, but in the sequence in which the patterns are transmitted. This solution may be used both with predefined and arbitrary patterns. The dimming pattern selection algorithm has already been described in section 2.3.2. For steganographic coding, we can use a fixed pattern as a "0" or "1". We denote:

- N - the pattern sequence length (N is a power of 10 according to the desired dimming precision order (expressed as a logarithm value) p: $N = 10^{1-p}$
- r_1 , r_2 – the number of repetitions of the first and second pattern (r_2=N-r_1)

The number of bits that can be encoded in the pattern sequence is:

$$nbitsteg = \log_2 \frac{N!}{(N - r_1)!\,(r_1)!} \tag{7}$$

Hence, the maximum number of bits that we can transmit with a pattern sequence of length 10 is 7.97, and with a pattern sequence of length 100 we can transmit up to 96.3 bits. The number of bits that can be encoded in this way is highly dependent on the dimming precision and the ratio of r_1/r_2. The IEEE 802.15.7 standard proposes using 0.1% precision, hence N=100. When the dimming level is set to multiples of 10%, only one pattern is used and no data can be transmitted. We can circumvent this problem by enforcing at least one repetition of the first or second pattern. The difference of 0.1% between the actual and desired light level will not be perceivable by users (due to nonlinear sensitivity of the human eye, we can allow for a larger

divergence between the actual and desired dimming level depending on the intensity of the perceived light. This we leave as a topic for separate work).

4 Hiding data in PHY header fields

In this section, we show the possibility of hiding data in the PHY layer of IEEE 802.15.7 standard VLC. The structure of the VLC physical layer convergence protocol data unit (PPDU), as defined in [IEEE, 11], is shown in Figure 12. The PPDU consists of:

- Preamble, which is used to obtain optical clock synchronization
- PHY header, which is a fixed length of 4 octets
- Header Control Sequence (HCS)
- Optional fields, whose presence is indicated beforehand in the header
- Physical layer convergence protocol service data unit (PSDU) – transmitted data

The details of the header are shown in Figure 12. From our point of view, two fields are of interest:

- Reserved fields – 5 bits at the end of the header
- Channel number – 3 bits used to transmit the "band plan"

In a manner similar to the one proposed for ZigBee [Baronti, 07] and other PHY level protocols, the reserved fields of the header may be used for steganographic transmission. According to the IEEE 802.15.7 standard, “reserved bits” should be set to zero on transmission and ignored on reception. A secret message can be divided into 5 bit chunks, distributed into "reserved" bits of outgoing packets, and the receiver should aggregate the chunks to recover the secret message.

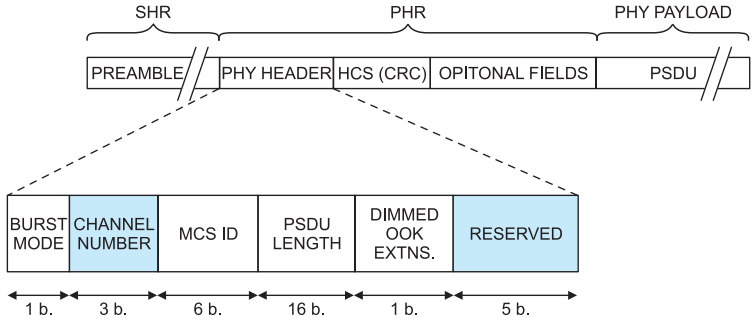


Figure 12: General structure of the IEEE 802.15.7 PHY frame.

According to the standard, the VLC receiver should support reception of the entire visible light spectrum with any type of optical light source. This requirement guarantees association without knowledge of receiver capabilities as well as the

support of unidirectional broadcasting. The "channel number" field may be used by the receiver for optimizing its performance. However, with current technology, i.e. wideband white LEDs, it can be assumed that spectrum optimization via channel number will rarely be used, hence the extra 3 bits may also be used for steganographic transmission.

Finally we should also note that some of the steganographic methods originally proposed for the PHY level of radio-based WLAN networks are also applicable to VLC systems, namely:

1) Extra symbols in OFDM and manipulating the safety interval between OFDM symbols
2) Manipulating Inter Frame Spacing (IFS) time
3) Introducing data into the padding bits used in the Forward Error Correction (FEC) encoder which is based on Reed-Salomon (RS) coding.
4) Using symbol padding in OFDM

5 VLC Steganography on the MAC level

5.1 Overview of the IEEE 802.15.7 MAC level

The MAC level of IEEE 802.15.7 standard VLC bears some resemblance to radio-based WLAN protocols such as Wi-Fi (802.11) and Bluetooth (IEEE 802.15.4). Each MAC frame consists of the following basic components (as shown in Figure 13):

- A MAC header (MHR), which comprises frame control, sequence number, address information, and security-related information
- A MAC service data unit (MSDU) of variable length, which contains information specific to the frame type. Acknowledgment frames do not contain a payload.
- An MAC Footer (MFR) field, which contains a frame control sequence (FCS).

The frame control field is 2 octets in length and contains information that defines the frame type, addressing fields, and other control flags. Four types of MAC frames are used by the standard, as specified by the "Frame Type" subfield of the FCS: beacon, data, acknowledgment and command. In further discussion, we will distinguish the frame type if necessary.

The frame control field itself contains 4 unused (reserved) bits which can be used to transmit hidden data.

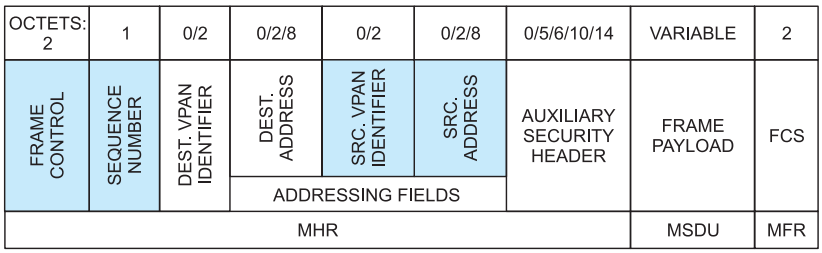


Figure 13: General structure of the IEEE 802.15.7 MAC frame.

5.2 Sequence number field

The sequence number field is 1 octet in length and specifies the sequence identifier for the frame. For a beacon frame, the sequence number field specifies a beacon-sequence number (BSN). Each coordinator should store its current BSN value in the MAC physical-layer personal-area-network information base (PIB) attribute – macBSN, and initialize it to a random value. Each time a beacon frame is generated, the MAC sublayer copies the value of macBSN into the sequence number field of the MHR of the outgoing frame and then increases it by one. For data, an acknowledgment, or a MAC command frame, the sequence number field specifies a data-sequence number (DSN) that is used to match an acknowledgment frame to the data or MAC command frame. In a manner similar to BSN generation, each device stores its current DSN value in the MAC PIB attribute – macDSN, and initializes it to a random value. Each time data or a MAC command frame is generated, the MAC sublayer copies the value of macDSN into the sequence number field of the MHR of the outgoing frame and then increases it by one. (The algorithm for choosing the random initiating number is out of the scope of the standard.) If we set a non-random value to macBSN (for a coordinator) or macDSN (for a VLC device), we can "hide" steganographic data inside it. Up to 8 bits may be transmitted in this manner in a single packet.

5.3 The addressing field

The addressing field is comprised of the destination address field and/or the source address field, depending on the settings in the frame control field. According to destination and source addressing modes specified in the FCS, a frame may contain a short (2 octet), long (8 octet) or zero-length source or destination address. Short addresses are used to identify VPANs and long addresses – individual devices. The source addressing field may be used to transmit hidden data if a non-existent source address is used. With this technique, data of up to 64 bits can be transmitted in a single packet. This technique can be difficult to detect if the number of nodes in the network is not known, especially in big networks where nodes appear and disappear frequently over time.

5.4 Beacon frame field

A beacon MSDU contains some additional fields which may be used to transmit hidden data. The structure of a beacon frame is shown in Figure 14. The "Superframe Specification" field defines the parameters of the periodically transmitted superframe – a 1 bit reserved field may be used to hide data in a manner identical to that described in section 5.1. Guaranteed time slot (GTS) variable field includes: "GTS Specification" and "GTS direction" (responsible for the management of guaranteed time slot parameters) which contain respectively 4 and 1 unused bits. Finally, the "Pending Address" field which provides notification of pending frames contains 2 reserved and unused bits. In effect a total of 8 bits in the beacon MSDU may be used to transmit hidden data.

OCTETS: 2	1	4/10	0/5/6/10/14	2	VARIABLE	VARIABLE	0/1	VARIABLE	2
FRAME CONTROL	SEQUENCE NUMBER	ADDRESSING FIELDS	AUXILIARY SECURITY HEADER	SUPERFRAME SPEC	GTS FIELDS	PENDING ADDRESS FIELDS	OTHER	BEACON PAYLOAD	FCS
MHR				MSDU					MFR

Figure 14: Structure of the IEEE 802.15.7 MAC beacon frame.

5.5 Acknowledgment frame field

In addition to the frame control fields described above, the acknowledgment frame field contains an optional B-ACK payload MSDU with a 1 bit reserved field which may be used for hidden transmission.

5.6 The MAC checksum

As originally proposed in [Szczypiorski, 03] for Wi-Fi protocols, VLC steganography may also be based on the transmission of frames with intentionally corrupted checksums. The basic idea behind VLC HICCUPS is as follows: when a regular (ie. unaware of this steganographic method) client client detects an error in a broadcast frame by checking the packet's FCS checksum, it simply drops the corrupted frame. A HICCUPS-enabled client extracts data from the allegedly corrupt frames, which are in fact used for hidden transmission. The key issue with undetectability is to keep BER at a level close to the original one – otherwise the covert channel will be very easy to compromise. Another option is to hide information as an FCS checksum and construct the corresponding payload to conform this value. This approach offers limited bandwidth, but has no influence on error rate.

5.7 Summary - MAC level information hiding

On the MAC level it is possible to hide a substantial amount of data in the IEEE 802.15.7 standard MAC frame. Table 3 summarizes the number of steganographic bits in each MAC frame type.

Frame type	Frame control	Sequence	Source address	Other	Max. bits
Data	4	8	64		76
Beacon	4	8	64/16	8	84
Acknowledgment	4	-	-	1	5
Control	4	8	64	-	76

Table 3: Potential number of steganographic bits for different IEEE 802.15.7 frame types.

6 Detectability of the proposed steganographic methods

According to [Fridrich 98], network steganography communication can be characterized by three basic and interdependent features: bandwidth, undetectability and robustness. Their interdependence can be illustrated by representing them as vertices of a triangle. For example, the higher the required bandwidth, the more difficult it is to achieve high undetectability and robustness.

It should be emphasized that the most desirable characteristic of network steganography communication is usually undetectability [Moskowitz 02]. In this chapter, we will present a short summary of the undetectability features of each of the presented steganographic methods. The classification that we use is that of "a moving observer", originally introduced for wireless radio networks and presented in [Szczypiorski 15]. For a full explanation of this classification method, we refer the reader to this cited article. A short summary of the method follows:

Three levels of undetectability are “good”, “bad”, and “ugly”, defined as

- “Good” – the observer is unable to detect hidden communication at the source of the steganograms.
- “Bad” – the observer is able to detect hidden communication at the source of the steganograms, but is unable to detect this communication when he/she is moved away from the source
- “Ugly” – the observer is able to detect hidden communication anywhere in the network, even at the steganographic receiver.

The summary of the undetectability of the proposed steganographic methods is presented in Table 5. It is worth mentioning that with dimming and flicker elimination mechanisms, the actual undetectability is dependent on the implementation of this mechanism in a given system, and hence in general, the undetectability will vary between "good" and "bad".

Level and method	**Mark**		
	Good	**Bad**	**Ugly**
Physical / modulation			
Steganography with OOK, VPPM, CSK		x	
OFDM extra symbols, padding			x
DS-CDMA - data in code spreading sequence		x	
Dimming / flicker elimination			
Encoding data in an idle pattern or in a sequence of idle patterns	x	x	
Hiding data in protocol headers			
PHY header fields, MAC header fields			x
extra bits in FEC			x
Manipulating Inter Frame Spacing (IFS) time	x		

Table 4: Summary of undetectability level for the proposed steganographic methods

7 Summary

Paradoxically, hidden data in VLC should not be visible, but rather as invisible as possible to protect covert channels. We presented two brand new methods for steganography: dimming patterns and CVD frames, which both depend on VLC features. Two other methods for steganography were adapted from methods already used for wireless and wired networks: unused header bits and pseudo-corrupted frames. There is no doubt that the issue of steganography in VLC is interesting, but there are still questions about the real life application of steganography in limited coverage networks like VLC systems. Because we treat steganography as a potential security flaw, we strongly believe that looking for covert channels even in local area networks and personal area networks is a necessity and must be part of security analysis.

To our best knowledge, this is the first work on the subject of steganography in VLC. We have shown methods of transmitting hidden data on the PHY level by manipulating modulation parameters of typical schemes used in VLC systems, namely OOK, VPPM and CSK. Steganographic transmission via modulation "tweaking" may be potentially applied to any VLC system. We have also referred to PHY and MAC levels of the IEEE 802.15.7 standard to demonstrate how data may be hidden in protocol frames. In our opinion, VLC systems are especially prone to hidden transmission for the following reason: on the PHY level, the IM/DD technique opens the door to piggybacking additional information directly onto the transmitted data stream. In the IEEE 802.15.7 standard, the PHY level complexity, especially in

terms of the various dimming and anti-flicker functions, and the complex MAC level functionality allow for numerous schemes for introducing hidden data.

This work presents different possibilities of introducing steganography in VLC, this leads to a question regarding the actual feasibility of implementation. Recently the authors have developed a proof of concept VLC steganographic system – "LuxSteg", which is described in a separate paper currently under review [Blinowski 17].

References

[Ayyash, 16] Ayyash, M., et al.: "Coexistence of WiFi and LiFi toward 5G: concepts, opportunities, and challenges", In *IEEE Communications Magazine*, *54*(2), 64-71, 2016

[Azhar, 13] Azhar, A.H., Tran, T., O'Brien, D.: "A Gigabit/s Indoor Wireless Transmission Using MIMO-OFDM Visible-Light Communications", In IEEE Photonics Technology Letters, vol. 25, No. 2, 2013

[Baronti, 07] Baronti, P., Pillai, P., Chook, V. W., Chessa, S., Gotta, A., Hu, Y. F.: "Wireless sensor networks: A survey on the state of the art and the 802.15. 4 and ZigBee standards," Computer communications, 30(7), pp. 1655-1695, 2007

[Blinowski, 16a] Blinowski, G.: "Practical aspects of physical and MAC layer security in visible light communication systems,". International Journal of Electronics and Telecommunications, 2016, 62(1), pp. 7-13.

[Blinowski, 17] Blinowski, G., Januszewski, P., Stępniak, G., Szczypiorski, K.: "LuxSteg: First practical implementation of steganography in VLC", International Journal of Computing Communication and Control, ISSN 18419836, (under review), 2017.

[Blinowski, 16b] Blinowski, G., Kmieciak, A.: "Modelling and evaluation of a multi-tag LED-ID platform", Proc. Computer Science and Information Systems (FedCSIS), 2016 Federated Conference on. IEEE, DOI:10.15439/2016F89, pp. 1049 – 1056, 2016

[Chen, 14] Chen, L.B., et al. "Development of a dual-mode visible light communications wireless digital conference system," In Consumer Electronics (ISCE 2014), The 18th IEEE International Symposium on, 2014, pp. 1-2.

[CIE, 31] CIE, "Commission Internationale de lEclairage Proc." Cambridge University Press, 1931.

[Classen, 15] Classen, J., Chen, J., Steinmetzer, D., Hollick, M., Knightly, E.: "The Spy Next Door: Eavesdropping on High Throughput Visible Light Communications" In Proceedings of the 2nd ACM MobiCom Workshop on Visible Light Communication Systems, ser. VLCS (Vol. 15), 2015

[Classen, 16] Classen, J., Steinmetzer, D., Hollick, M.: "Opportunities and pitfalls in securing visible light communication on the physical layer," In Proceedings of the 3rd Workshop on Visible Light Communication Systems (pp. 19-24). ACM, 2016

[Conti, 08] Conti, J.P.: "What you see is what you send," Engineering & Technology, 2008, pp. 66-67.

[Fridrich 98] J. Fridrich, "Applications of data hiding in digital images (tutorial)," In Proc. IEEE International Workshop on Intelligent Signal Processing and Communication Systems (ISPACS'98), 1998.

[Ghimire, 12] Ghimire B., Haas, H.: "Self-organising interference coordination in optical wireless networks," EURASIP Journal on Wireless Communications and Networking, 2012(1), pp. 1-15.

[Hranilovic, 13] Hranilovic, S., Lampe L., Hosur, S.: "Visible light communications: the road to standardization and commercialization," In IEEE Communications Magazine, vol. 51, Iss. 12, ISSN: 0163-6804, 2013, pp. 24-54.

[IEEE, 11] "IEEE standard for local and metropolitan area networks–part 15.7: Short-range wireless optical communication using visible light", IEEE Std 802.15.7-2011, https://standards.ieee.org/findstds/standard/802.15.7-2011.html

[Langer, 08] Langer, K.D., et al.: "Optical Wireless Communications for Broadband Access in Home Area Networks," In Proc. International Conference on Transparent Optical Networks, ICTON, 2008, pp. 149 - 154, DOI: 10.1109/ICTON.2008.4598756

[Langer, 11] Langer, K.D., et al.: "Exploring the potentials of optical-wireless communication using white LEDs," in Proc. 13th Int. Conf. Transp. Opt. Netw., Jun. 2011, pp. 1–5.

[Mazurczyk, 16] Mazurczyk, W., Wendzel, S., Zander, S., Houmansadr, A., Szczypiorski, K.: "Information Hiding in Communication Networks: Fundamentals, Mechanisms, Applications, and Countermeasures", Wiley-IEEE Press; 1 edition, February 2016.

[Monteiro, 12] Monteiro, E., Hranilovic, S.: "Constellation design for colour-shift keying using interior point methods." *2012 IEEE Globecom Workshops*. IEEE, 2012.

[Moskowitz 02] Moskowitz, I.S., Chang, L. and Newman, R. E., "Capacity is the wrong paradigm," in Proc. New Security Paradigms Workshop (NSPW 2002), 2002, pp. 114–126.

[O'Brien, 08] O'Brien, D.C., et al: "Home access networks using optical wireless transmission," In Proc. Personal, Indoor and Mobile Radio Communications, 2008, IEEE 19th International Symposium on, pp. 1-5, 2008

[O'Brien, 09] O'Brien, D.C. , et al: "Gigabit Optical Wireless for a Home Access Network," in Proc. IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications, 2009, pp. 1-5.

[Pergoloni, 15] Pergoloni, S. et al.: "Merging colour shift keying and complementary pulse position modulation for visible light illumination and communication." *Journal of Lightwave Technology* 33.1 (2015): 192-200.

[Rahaim, 11] Rahaim, M. B., Vegni A.M., Little, T. D.: "A hybrid radio frequency and broadcast visible light communication system," in Proc. IEEE Global Communications Conference (GLOBECOM) Workshops, 2011, pp. 792–796.

[Ren, 14] Ren, Z. X., Zhang, H. M., Wei L., Guan, Y.: "A High Precision Indoor Positioning System Based on VLC and Smart Handheld," in Applied Mechanics and Materials, Vol. 571, 2014, pp. 183-186.

[Szczypiorski, 03] Szczypiorski, K.: "HICCUPS: Hidden communication system for corrupted networks," International Multi-Conference on Advanced Computer Systems, pp. 31-40, 2003

[Szczypiorski, 15] Szczypiorski, K., Janicki, A., and Wendzel, S.: "The Good, The Bad And The Ugly: Evaluation of Wi-Fi Steganography," In: Journal of Communications (JCM), ISSN: 1796-2021 (Online), ISSN: 2374-4367 (Print), Vol. 10, No. 10, October 2015, pp. 747-752, DOI: 10.12720/jcm.10.10.747-752

[Tsiatmas, 14] Tsiatmas, A., et al.: "An illumination perspective on visible light communications," In Communications Magazine, IEEE, 52.7, 2014, pp. 64-71.

[Yoo, 13] Yoo J. H., Jung, S. Y.: "Modeling and analysis of variable PPM for visible light communications," EURASIP Journal on Wireless Communications and Networking, 2013(1)

[Yoshino, 08] Yoshino, M., Haruyama S., Nakagawa, M.: "High-accuracy positioning system using visible LED lights and image sensor," Radio and Wireless Symposium, IEEE, vol., no., 2008, pp.439-442.

[Zielinska, 14] Zielinska., E. Mazurczyk, W., Szczypiorski, K., "Trends in steganography, " Communications of the ACM, 57(3), 2014, pp. 86-95.