Delft University of Technology
Master's Thesis in Embedded Systems

Leveraging VLC for energy disaggregation in Smart Buildings

Johnny Verhoeff

Leveraging VLC for energy disaggregation in Smart Buildings

Master's Thesis in Embedded Systems

Embedded Software Section
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands

Johnny Verhoeff
j.s.c.j.verhoeff@student.tudelft.nl

26th March 2016

Author
Johnny Verhoeff (j.s.c.j.verhoeff@student.tudelft.nl)
Title
Leveraging VLC for energy disaggregation in Smart Buildings
MSc presentation
26th March 2016

Graduation Committee

TODO GRADUATION COMMITTEE	Delft University of Technology
TODO GRADUATION COMMITTEE	Delft University of Technology

Abstract

TODO ABSTRACT

Preface

TODO MOTIVATION FOR RESEARCH TOPIC

TODO ACKNOWLEDGEMENTS

TODO AUTHOR

Delft, The Netherlands
26th March 2016

Contents

Chapter 1

Introduction

TODO INTRODUCTION

TODO ORGANISATIONAL DESCRIPTION OF THESIS

Chapter 2

CHAPTER TITLE

INTRODUCTION TEXT TO THIS CHAPTER IN WHICH ALL SECTIONS ARE DESCRIBED ROUGHLY (1 SENTENCE EACH).

This chapter describes the ... In Section 2.1, examples are given of how to use tables and figures in MSc theses.

2.1 SECTION TITLE

Every caption of a table (or figure) should start with a capital letter, and should end with a period. References to tables are given with a capital letter for table, as in "(see Table 2.1)" or "in Table 2.1, ...".

left aligned	centred	right aligned
12	34	56

Table 2.1: Complete sentence describing the tabular data.

References to figures are given with a capital letter for figure, as in "(see Figure 2.1)" or "in Figure 2.1, ...".

[1] [3]

Figure 2.1: Complete sentence describing the figure thoroughly.

2.2 CDMA

This section will explain what CDMA is, alternatives and why it is needed.

When transmitting data from a transmitter to a receiver over a channel, the entire channel is being used for this purpose. If one wants to have multiple transmitters transmitting over one channel, there is a problem. The transmitters interfere with each other, this is called multiple access interference (MAI). There are several ways to get around this problem:

- TDM: Time Division Multiplexing.
  Each transmitter gets assigned a time slot, in which it and only it is allowed to transmit, hereby going around the MAI problem.

- FDM: Frequency Division Multiplexing.
  Each transmitters gets assigned a frequency band. Each transmitter is allowed to transmit the whole time, but only at the assigned frequencies.

- CDM: Code Division Multiplexing.
  Each transmitter gets assigned a code word. The data first needs to be encoded using the code word and then the transmitter can send his message. Each transmitter can transmit all the time using the entire frequency band. These codes will determine how many transmitters can actually transmit with correct decoding results at the receiver end.

The distributed network of the VLC LEDs is inherently uncoordinated, since all the LEDs are basically only lights. They have no receiver of any kind. They can only implicitly transmit data by turning the load or the LED on or off. Because the LEDs cannot receive data, they cannot be assigned a time slot and therefor we cannot use a TDM scheme.

An FDM scheme is also not applicable. This is because the LEDs do not have an explicit hardware transmitter that can modulate a signal. Instead the transmitting is implicitly done via turning the LED on and off. So only binary values of the current draw are sent as signals.

This is where the CDM approach comes into play. This scheme allows the multiple LEDs to transmit at the same time. But the type of code used here is of importance. The code type determines the MAI and what the receiver is able to decode.

2.2.1 Performance metrics of a code

To determine which code is the best for this problem some measures are needed to be able to compare the codes.

One such a measure is called the correlation. Correlation is a measure for determining how much sequence X is similar to sequence Y and can be found in Equation 2.1. With L being the length of the code and τ the time-shift. When sequence X and Y are the same sequence, we speak of the autocorrelation. When they are two different sequences, we speak of the cross-correlation.

$$R(\tau)_{xy} = \sum_{i=0}^{L-1} x(i) \times y(i+\tau) \text{ with } \tau = 0, 1, 2, \ldots, L \tag{2.1}$$

The properties of an ideal code set should be, that the autocorrelation for each code in the set should be 0 for each time-shift $\tau \neq 0$, at $\tau = 0$ the autocorrelation should be L. The ideal cross-correlation properties should 0 for every time-shift τ, so that no code interferes with any other code, hereby causing no MAI.

Other metrics to be considered are the length of the code and how many codes there are in that code set. The code length is of importance because each bit a user will transmit must be encoded. So the the message that will be transmitted via the channel will be the length of the code times the length of the data. If there are only a few codes in a code set then only a few number of users can transmit which does not scale well.

2.2.2 Walsh-Hadamard Sequences

Walsh-Hadamard sequences are sequences which are created using a Hadamard matrix. Hadamard matrices are $n \times n$ matrices which are recursively generated. Starting with a 1×1 matrix: $H_1 = \begin{bmatrix} 1 \end{bmatrix}$, then $H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. Or in general: $H_{2n} = \begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix}$ [2]. The matrix can also be filled with binary values: zero and one. In that case the general matrix will be: $H_{2n} = \begin{bmatrix} H_n & H_n \\ H_n & \overline{H_n} \end{bmatrix}$

The Hadamard matrix has the property that every row in the matrix is orthogonal to every other row. Hadamard matrices exist for every power of 2, so the code length is also a power of 2. So for $\tau = 0$, the cross-correlation is 0, but when $\tau \neq 0$ not all the rows are orthogonal anymore. [6] proved that an Hadamard matrix of size 2^P could be divided into $P + 1$ subsets of rows, where one code could be selected giving $P + 1$ orthogonal codes for each time-shift τ. These codes are called Cyclically Orthogonal Walsh Hadamard Codes (COWHC).

All rows of the matrix have the property that the autocorrelation at $\tau = 0$ is equal to L. But when $\tau \neq 0$, undesirable behavior occurs as can be seen in Figure 2.2. The autocorrelation function has several high peaks where only one is desired. This means that if a transmitter sends an encoded message with this code and the receiver does not know when in time the start of the message is, the receiver would get false positives for data.

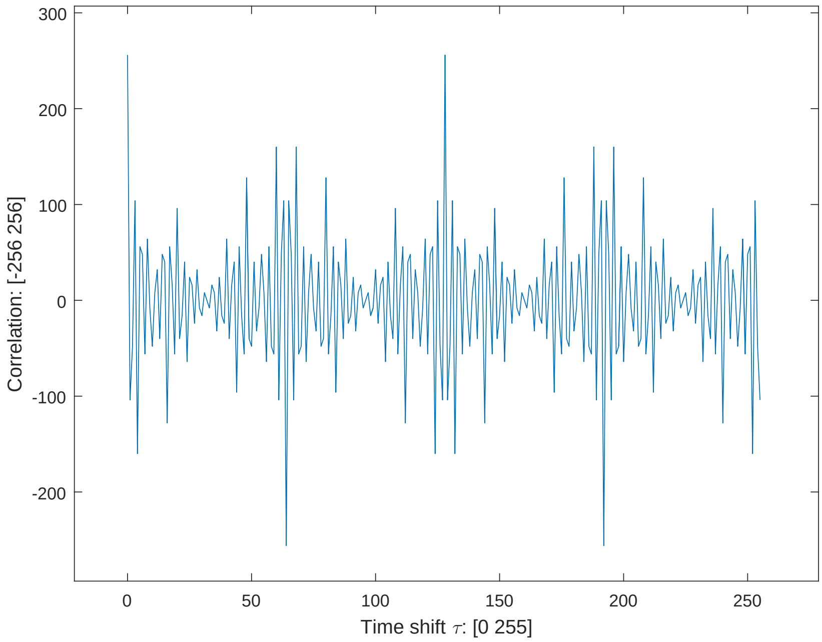

Figure 2.2: Autocorrelation of Hadamard code with index 120 of length 256.

So only a small subset of the codes have 0 cross-correlation for every time-shift τ and the autocorrelation is far from what the ideal code set should have.

2.2.3 PN Sequences

PN sequences are sequences where the numbers looks like they are randomly generated but they are easily generated in software or hardware. The sequences have the following noise-like properties [4]:

- Balanced
  Any PN sequence of length $L = 2^n - 1$ contains exactly 2^{n-1} ones and $2^{n-1} - 1$ zeros.

- Runs
  A run is a subset of the sequence where all the consecutive numbers are the same. In any PN sequence, 1/2 of the runs have length 1, 1/4 have length 2, 1/8 have length 3 and so on.

- Autocorrelation
  The autocorrelation function of a PN sequence will take on two values as can be seen in Equation 2.2 and Figure 2.3.

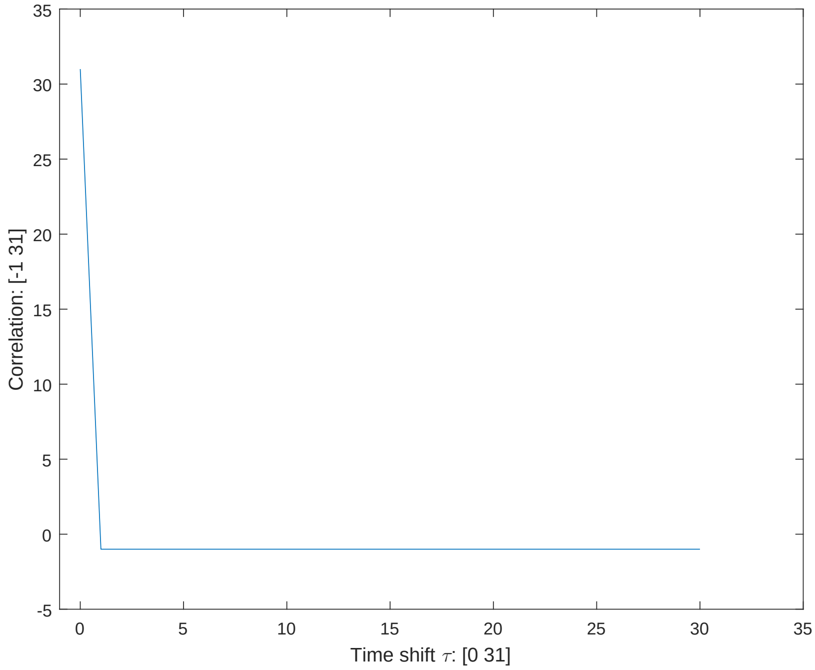

Figure 2.3: Autocorrelation of PN sequence of length 31.

$$R(\tau) = \begin{cases} L & \text{if } \tau = 0 \\ -1 & \text{if } \tau \neq 0 \end{cases} \tag{2.2}$$

PN sequences are generated using a linear feedback shift register (LFSR) [5]. Figure 2.4 shows an n length LFSR with some XOR gates attached to it. The LFSR is defined entirely by the feedback function, also called a characteristic polynomial. It determines the length and the type of sequence generated. The polynomial looks like Equation 2.3. The LFSR in Figure 2.4 contains n shift registers and is initiated with a starting seed. This seed can be any vector apart from the all zero vector. The reason for this is because the XOR function with two zeros as input will output a zero, making the sequence that this LFSR outputs a sequence of all zeros. The output of the shift registers are multiplied with the coefficients of the characteristic polynomial $(C_{n-1}, Cn-2, \dots, C_1, C_0)$. The result output is then fed back into the first shift register, all the bits in the rest of the registers also shift one position and one output bit is created. With n number of registers and the zero start vector excluded it takes $2^n - 1$ steps to output the entire sequence before it starts to repeat itself.

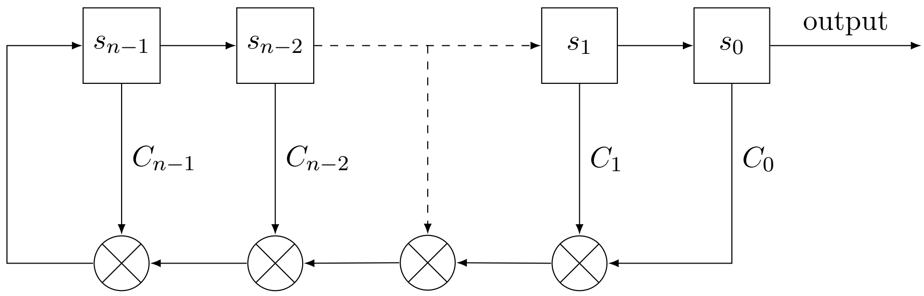

Figure 2.4: Linear feedback shifter register of length n, with XOR gates.

$$p(x) = x^n + C_{n-1}x^{n-1} + C_{n-2}x^{n-2} + \ldots + C_2x^2 + C_1x + C_0 \tag{2.3}$$

Chapter 3

Conclusions and Future Work

3.1 Conclusions

TODO CONCLUSIONS

3.2 Future Work

TODO FUTURE WORK

Bibliography

[1] Test author. Tes title. *c*, e(f):g, h d. i.

[2] E. H. Dinan and B. Jabbari. Spreading codes for direct sequence cdma and wideband cdma cellular networks. *IEEE Communications Magazine*, 36(9):48–54, Sep 1998.

[3] IOJASDOIJASDOIJAS. *b*, volume e of *g*. g, re, e edition, e h. e.

[4] Abhijit Mitra. On pseudo-random and orthogonal binary spreading sequences. *Int. J. Information Techn*, 4(2):137–144, 2008.

[5] L.-T. Wang and E. J. McClusky. Linear feedback shift register design using cyclic codes. *IEEE Trans. Comput.*, 37(10):1302–1306, Oct. 1988.

[6] Yongxiang Xia, Pu Li, Xiuming Shan, and Yong Ren. Cyclically orthogonal subsets of walsh functions. In *Communication Systems, 2002. ICCS 2002. The 8th International Conference on*, volume 1, pages 107–111 vol.1, Nov 2002.