**Interim Report**

**EE5500**

Name: Wojciech Lesnianski

Student number: 1644612

Electronic and Computer Engineering

School of Engineering and Design



Dr. Ali Mousavi

Thursday, January 11, 2018

School of Engineering & Design

Electronic & Computer Engineering

MSc Distributed Computing Systems Engineering

Brunel University West London

Simulation and Performance Analysis

of a Distributed Position Correction

Scheme for Unmanned Aerial Vehicles

Dionysios Satikidis

Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Declaration:** I have read and I understand the MSc dissertation

guidelines on plagiarism and cheating, and I certify that this

submission fully complies with these guidelines.

Abstract

Acknowledgements

Me for taking teh time to write all this stuff

List of Abbreviations

|  |  |
| --- | --- |
| VIN | **V**ehicle **I**dentification **N**umber |
| SF | **S**auberes **F**ahren |
| C# | C-Sharp |
| UMTS | **U**niversal **M**obile **T**elecommunications **S**ystem |
| XML | **E**xtensible **M**arkup **L**anguage |
| Wifi | **Wi**reless **Fi**delity |
| OBD | **O**n-**B**oard-**D**iagnose |
| OSM | **O**pen **S**treet **M**ap |
| GUI | Grafische Benutzeroberfläche (**G**raphical **U**ser **I**nterface) |
| WPF | **W**indows **P**resentation **F**oundation |
| PNG | **P**ortable **N**etwork **G**raphics |

List of Therms

|  |  |
| --- | --- |
| VIN / Vehicle Identification Number | **Vehicle Identification Number**,  ist die international genormte, 17-stellige Nummer, mit der ein Kraftfahrzeugeindeutig identifizierbar ist. |
| Fahrzeugtyp | Fahrzeugtyp wird in diesem Projekt gleichgesetzt mit von der Informationen, welche aus der VIN gewonnen werden können. Dazu gehören der Hersteller, das Modell und die Motorisierung. |
| OSM / Open Street Map | Open Street Map ist ein frei nutzbares Projekt, welches Geodaten bereitstellt. |

Content

[1 Introduction 1](#_Toc503095427)

[1.1 Context of the Project 2](#_Toc503095428)

[1.2 Problem Description 5](#_Toc503095429)

[1.3 Aims and Objectives 6](#_Toc503095430)

[2 Methodology and Project Organisation 8](#_Toc503095431)

[2.1 Software Development Process 8](#_Toc503095432)

[2.2 Tools and Architectures 8](#_Toc503095433)

[2.3 Strengths and Risks 8](#_Toc503095434)

[2.4 Project Management 8](#_Toc503095435)

[3 Literature Review 9](#_Toc503095436)

[3.1 Water-Plants 9](#_Toc503095437)

[3.2 Data Harmonization 10](#_Toc503095438)

[3.3 Comparable Industries and Projects 10](#_Toc503095439)

[4 Design and Implementation 11](#_Toc503095440)

[4.1 Alalalalala 11](#_Toc503095441)

[5 Experimental Results and Analysis 12](#_Toc503095442)

[5.1 Wasmachicheigentlichhier 12](#_Toc503095443)

[5.2 Was ist das 12](#_Toc503095444)

[6 Conclusions and Further Work 13](#_Toc503095445)

[6.1 Bla 13](#_Toc503095446)

[6.2 Blub 13](#_Toc503095447)

[6.3 Blib 13](#_Toc503095448)

1. Introduction

Access to clean water is the most basic and fundamental type of the human infrastructure. The quality of life highly depends on the accessibility to clean water. We require water not only for drinking, but also for cooking, and washing. Additionally, various professions and commercial establishments, like farms or restaurants, could not exist without certain quality and quantity of water. The quantity of clean water in most cases, depends on collecting water and sewage from rivers and lakes, cleaning it in dedicated water-plants and thus bringing it to a specific quality standard, and then distributing it.

A software groundwork for acquisition, analysis and modelling of historical and real-time data of water-plants could become the first step to provide an infrastructure capable of “keeping track of the water” – its amount, quality and source, as well as making forecasts and statistics easier. This thesis will focus on the acquisition, harmonization and provision of water-related historical and real-time data.



(Source: [http://www.water.org.uk/about-water-uk/wastewater 18.09.2017](http://www.water.org.uk/about-water-uk/wastewater%2018.09.2017) last accessed: 28.09.2017)

Even though for a common user, the most significant outcome is the one of step 6 – Final treatment[[1]](#footnote-1), which also indicates the quality of water available for public usage, the incoming and outgoing water of the other steps provide data for other kinds of interesting analyses, especially due to the fact, that each step deals with a specific kind of water quality related problem, meaning that all “possible to gather” data should also be gathered, harmonized and stored by t system, for further investigation.

* 1. Context of the Project

To understand how important waste water treatment is, one has to understand the dimensions of its effect. On average, a UK water-customer pays 1 pound a day to be able to enjoy high quality water. This money is much needed and is used for the treatment of around 16 billion litres of wastewater, gathered in around 345 thousand kilometres of culvert and cleaned in about 9000 wastewater plants – every day, as well as developing the water infrastructure in the country. [3]

Every country has its own way of dealing with the regulations of the water treatment process. (<https://www.eea.europa.eu/data-and-maps/indicators/urban-waste-water-treatment/urban-waste-water-treatment-assessment-4>): Taking into consideration the 4 Steps:

* Collecting wastewater
* Primary treatment[[2]](#footnote-2)
* Secondary treatment
* Tertiary treatment

the trend in Europe over the past decades was to connect more and more population to waste water treatment plants. In northern countries 80% have been reached already in 1995, with over 70% of the water receiving tertiary treatment. In central Europe (this includes the UK) on average over 95% of the population enjoy treated water. The trend also shows increasing amount of tertiary treating among all countries:



What UK is doing great, is the fact that 100% of the population is connected to at least one of the waste water treatment plants. Also, all of the collected water receives at least a secondary treatment (Data from 2015). What UK still lacks, is tertiary treatment which is close to 100%, like in other countries including Germany, Netherlands, Denmark and Austria. Again, the trend shows improvement over the past years, but it seems rather slow in the progress. The tertiary treatment usually significantly reduces nitrogen and phosphorus pollution and might not always be required, but is still recommended by the Urban Waste Water Treatment Directive.

On a big scale, the UK can compare its achievements to other countries to find out what causes their good and bad results. The comparison on a small scale includes comparing the single water-plants within the UK to see which treatments steps are lacking, where the water quality is better, the treatment more effective and efficient and why.

The UK-wide water supply regulations are set by the government and regulate the water treatment process of every water provider whose area is wholly or partially in the United Kingdom. The list of indicator parameters is long and contains minimum, maximum values and ranges within which values are allowed to lie. Only if all regulations apply, the water may be called drinking water. With all the regulations and monitoring organizations the quality of UKs water might seem assured – yet the process of doing so is very troublesome and laborious. Twelve big companies, responsible for water and sewerage, cover most of UKs water supply. Additionally, there are some water-only companies providing water for some of the remaining regions. [4] [5]



(Source: <http://www.ofwat.gov.uk/households/your-water-company/map/> last accessed: 28.09.2017)

**[a]**

**Water transfer and interconnection**

Depending on the location, the population and the climate, some regions of the UK have less available clean water than others. Those regions usually lie in the south and east. This is why a system for water moving was needed and build already in the 17th century (New River to transfer water from Hertfordshire to London). This transfer and interconnection system was since then optimized and expanded to provide water to regions in need, despite other water companies being “responsible” for this region[[3]](#footnote-3).

Water can be transferred treated or untreated using canals, pipes, aqueducts or rivers. Treated water is typically transported over buried pipes. Transferring the water is costly, especially when lowland reservoirs have to supply upland reservoirs and the water needs to be pumped instead of having the gravity doing most of the work. This is why water should not be transferred if not explicitly necessary. Even though water transfer – or “water trade” when referring to water transfer between companies, is costly, it is still in most cases cheaper in money and energy than water desalination, and thus promoted by the government and regulators in the UK.

* 1. Problem Description

The water quality is regulated UK-wide, yet the way the different companies ensure their quality and monitor their water treatment process is not unified. This makes comparison of data between companies and water-plants, as well as getting a global picture difficult. Reacting to lack of quality water in specific regions, or forecasting such a scenario, while still monitoring which of the remaining regions has enough “spare” quality water to help out the company in need would be a lot easier with a common information base. It would simplify the monitoring of local area changes caused by changes in the water and wastewater treatment regulations. To assure better forecasts or more meaningful reports, other information bases, like weather information might be taken into account – but using those external systems are not a topic in this part of the (data-gathering) system.

The advantages of a big dataset from various sources are obvious – especially in a case where the geographical location of sources also matter. Co-operating, comparing, planning, monitoring and analysing is a lot easier when all the data is stored at seemingly one place in a unified format.

Several questions need to be investigated upon before the approach of gathering the water information from various water-plants can be attempted:

* Which data is available for the existing water-plants
  + Interesting categories:
    - Water-related data
    - Productivity data (such as reliability of technical equipment)
    - Energy data (such as the energy consumption)
    - Environmental data (such as carbon footprint)
  + How can the data be accessed
  + How can the data be legally used
  + What needs to be done to access the data
* Which data is mandatory to create a useful platform
  + Who are the stakeholders
  + Do all / most water-plants provide this data
* What is the best way to handle this data in terms of
  + Storage
  + Security
  + Processing
* What is the best way of harmonizing the data in terms of
  + Flexibility (adding new sources / reacting to changes in source schema)
  + Performance
  + Usability (Target-Schema)
* Which other sources of data (aside of the water-plants) could be used to enrich the dataset for
  + Historical data
  + Actual data
  + Non-water-data which could be useful anyway (i.e. for forecasts or to put more context on the stored data)
  1. Aims and Objectives

The purpose of this dissertation is to investigate and design knowledge and data engineering infrastructure for big amounts of water and wastewater treatment process specific data. This includes answering the questions asked in chapter 1.2 – Problem Description, as well as designing a solution with a proof of concept for the data harmonization. The below picture shows the basic idea of the systems interaction with the outside world. The scope of this project lies within the Large Network Performance Collider, which represents the harmonization and storage layer.



(Source: [8] Page 3)

The most essential objective of this dissertation is the investigation and answering of the questions as well as designing a software solution.

The creation of a fitting data-schema based upon most interesting and best available data is of high priority.

Recognizing deviation between input data schema and the predefined system schema. Additional aim is to make adjustments in the received data, or suggest “fixes” for the incoming schema to fit with the systems schema.

Next aim is to validate the stored data in terms of datatype and value range.

**Important system design objectives**:

The software design should take into consideration a high amount of different data sources for all the different water-plants and eventually other sources. Thus parallel processing of the data and concurrent storing should be considered.

Access security of the data has to be considered, as different roles might have different rights, as well as some stored data might not be legally given out to the public.

1. Methodology and Project Organisation
   1. Software Development Process
   2. Tools and Architectures
   3. Strengths and Risks
   4. Project Management
2. Literature Review

Designing the software and creating a fitting data-schema, as well as making decisions about the project infrastructure requires a deeper understanding of the subject. For this purpose it should be examined in detail with all its components.

The first survey will examine water-plants related topics. It will explain how a water-plant work and what the most important steps of the water cleaning process are. Additionally it will investigate on the similarities and differences between single water-plants in terms of water-based data – which data is stored, how it is stored, which data can be accessed, how and by whom.

The second survey will investigate on the topic data harmonization. It will explore the common methods of harmonization and the theory behind them.

The third survey will investigate on similar projects and industries, which also harmonized data coming from different sources for a common use. The advantages and disadvantages of each approach will be taken into consideration to see, which approaches – or parts of approaches, are fitting for this kind of a project.

HIER ÜBER DIE WEITEREN SURVEYS SCHREIBEN WENN MENR ALS 3.1 3.2 3.3 BESCHLOSSEN WURDE!!! :D:D:D: ASDF WTF LOL OMG

* 1. Water-Plants

**[A]**

Water-plants (wastewater-plants or sewage-plants) are used to clean water – primary household, -/ but also industries and businesses sewage – for further use. This is accomplished by speeding up the natural process by which water is purified. Today it is mostly done in three steps which will be explained later.

Water was privatised in the UK in 1989, since then the quality of water and its availability increased, but the administration of each water-plant differed more than before.

Functionality

The picture below shows a full life-cycle of water: 

(Source: https://www.water.org.uk/about-water-uk/wastewater)

**[B]**

As mentioned before, there are several steps within which the water (already filtered of grit and large solids) gets treated during the cleaning process. Those steps are:

1. **Primary Treating**

Physical and chemical settlement of suspended solid waste which didn’t get removed before, as well as reduction of its biochemical oxygen demand.

This step should reduce:

* Biochemical oxygen demand by 20%-30%
* Suspended solids by 60%

1. **Secondary Treating**

This step involves biological treatment to break down and reduce residual organic matter. This treatment step must comply with the standards of the Directive.

1. **Tertiary Treating (Final Treatment)**

This treatment step depends upon the location. It can involve disinfection by violet light, nutrient removal or the removal of specific toxic substances.

Statistics expand

There are around 9000 water-plants in the UK which collect around 16 billion liters of waste water each day.

Difficulties

The problematic in the context of this project lies mainly in the data acquisition. First of all, the data in the UK is not only not available to the public, but also not available to any authorities as online provided real time data.

Sensors which analyze the water and provide information are not 100% reliable. They may provide wrong data, data-holes or provide data in an inconsistent frequency. Those sensors may differ from water-plant to water-plant and be placed on different places within a plant, which makes comparison between different water-plants difficult. The sensors might not provide all the data which is needed to do a representative comparison between different water-plants. Also the data formats may not only differ in form but also in type. While the one might use an XML schema, the other one might use JSON. Treatment steps may differ from water-plant to water-plant, meaning that even if a sensor is placed on the same position of a primary treatment in two water-plants, the data may still differ a lot.

There is not much real data to develop the system. Real, water-plant-created data, is very useful when designing a harmonizing system, because it gives an idea about the different formats and possible deviations within the schemas. Additionally, having multiple sources gives statistical insight on which data is ‘usually’ available/tracked, and which is rather rare.

* 1. Data Harmonization

**[b] [D]**

Data harmonization’s aim is to create a single source of information based on multiple sources. The general problem during a harmonization process is, that the different base-sources differ the form in which their information is provided. To present a clean, harmonized information source, the data needs to be cleared of inaccurate and misleading entries. This means, that harmonization alone might not be enough for every process. To make sure the created information set is useful, it needs to undergo some additional processing.

Why Data Harmonization?

As already mentioned, harmonization is needed when dealing with different sources of information, but trying to get a common knowledge base. Sometimes the same type of information is stored in different types of data-formats. An example for this is the DateTime Format.

The DateTime might be stored in formats like:

* DD/MM/YY
* MM/DD/YY
* DD-MM-YYYY
* YYMMDD
* DD MONTH YYYY
* …

All the formats store the exact same information, but to extract this information further processing is needed. In some cases it might be enough to look at the data to find out which of the formats it is using, but in other cases (DD/MM/YY and MM/DD/YY) it is mandatory to know the format before extracting the needed information.

Data Simplification

This step eliminates unnecessary and redundant data. It is usually happening before the harmonization takes place. This has the advantage, that in high amounts of data, not all the data needs to undergo further processing. The disadvantage to have it as first step is, that some data might be removed that is actually needed, but not recognized as such and would be only after the harmonization step. A good approach is to have a simplification take place once before and once after the harmonization took place. The first simplification in this approach aims at clearly unneeded data and leave out data classified as ‘unsure’ to handle it in the second iteration after the harmonization is done.

Data sources are likely to have additional data, which is only relevant for the specific source itself and not for inter-source-analyses. This data needs to be identified and ‘cleaned’ from the source dataset.

Example with two different sources A and B:

Target Schema:

|  |
| --- |
| DataA |
| DataB |
| DataC |

Source A:

|  |  |  |
| --- | --- | --- |
| DataA | Simplification | DataA |
| DataB | DataB |
| DataD | ~~DataD~~ |

Source B:

|  |  |  |
| --- | --- | --- |
| DataA | Simplification | DataA |
| DataB | DataB |

Simplification needs to know the target schema to function. It doesn’t add values which are missing when comparing source to target. It only removes the unnecessary.

Data Standardization

Data standardization is about processing of the dataset into a standard form through standard bodies. It is not said that a harmonized dataset matches the standard forms outside of the system. It might be harmonized but still useless to the outside world. This is why a standardization step might be needed to do further processing of the data, change its form or format, before publishing it.

* 1. Comparable Industries and Projects

1. Design and Implementation
   1. Developing the Schema

The development of a common schema is one of the most important parts of this dissertation. It has to take into consideration, that the data-sources may vary a lot in their format, type of data and frequency, but also be specific enough to provide meaningful data for further analyses. Some values needed to determine a key-factor may not be available on a data source, or only as an interpolation. Some values may not be as current as other values, or update equally frequent.

Key-Factors

Key-Factors may also be seen as high level abstractions of the gathered raw data. On the most molecular level we have the raw water-plant data itself. The aim is to provide the key-factors in nearly real-time, calculated with the help of the raw-data gathered from the water-plants. The 4 key-factors used in this project are:

* **Productivity**
* **Energy**
* **Environment**
* **Quality**
  + - 1. Productivity

Technical Equipment (reliability, performance)

* + - 1. Energy

Based on **[C]**

* + - 1. Environment
      2. Quality

1. Experimental Results and Analysis
   1. Wasmachicheigentlichhier
   2. Was ist das
2. Conclusions and Further Work
   1. Bla
   2. Blub
   3. Blib

Literaturverzeichnis

|  |  |
| --- | --- |
| **Bücher:** | |
| [BUN09] | Hans Joachim Bungartz, et al. „Modellbildung und Simulation“, Springer, 2009 |
| [GOL12] | Joachim Goll „Methoden des Software Engineering“, Springer Vieweg, 2012; ISBN: 978-3-8348-2433-2 |
| [MEY14] | Bertrand Meyer, „Agile! – The Good, the Hype and the Ugly“, Springer, 2014; ISBN: 978-3-319-05154-3 |
| [LOR00] | Edward Lorenz „The Chaos Avant-grande: Memories oft he Early Days of Chaos Theory“, Kapitel 7 „The Butterfly Effect“, 2000; ISBN: 978-981-02-4404-0 |
| **Internetquellen:** | |
| [TUD] | <http://ls4-www.cs.tu-dortmund.de> (Letzter Zugriff: 04.01.16) |
| [AGI1] | <http://agilemanifesto.org/iso/de/principles.html> (Letzter Zugriff: 04.01.16) |
| [AGI2] | <http://agilemanifesto.org/iso/de> (Letzter Zugriff: 04.01.16) |
| [TUD1] | <http://ls4-www.cs.tu-dortmund.de/download/LehreMaterialien/MAO2011/Kap7.pdf> (Letzter Zugriff: 04.01.16) |
| [SWI] | http://sumo.dlr.de/wiki (Letzter Zugriff: 04.01.16) |
| [SWI1] | <http://sumo.dlr.de/wiki/SUMO_User_Documentation> (Letzter Zugriff: 04.01.16) |
| [SWI2] | <http://sumo.dlr.de/wiki/Theory/Traffic_Simulations> (Letzter Zugriff: 04.01.16) |
| [SWI3] | <http://sumo.dlr.de/wiki/Networks/Building_Networks_from_own_XML-descriptions> (Letzter Zugriff: 04.01.16) |
| [SWI4] | <http://sumo.dlr.de/wiki/Definition_of_Vehicles,_Vehicle_Types,_and_Routes> (Letzter Zugriff 04.01.16) |
| [SWI5] | <http://www.sumo.dlr.de/wiki/Simulation/Output> (Letzter Zugriff: 04.01.16) |
| [CUN04] | <http://www.artima.com/intv/simplestP.html> – Ein Gespräch zwischen Bill Venners und Ward Cunningham; „A Conversation with Ward Cunningham, Part V“ (Letzter Zugriff: 04.01.16) |
| [ZAV] | www2.research.att.com/~pamela/faq.html – Quelle angegeben in [MEY14]. (Letzter Zugriff: nicht erreichbar) |
| **Weitere Arbeiten und Projekte:** | |
| [C2X] | Matthias Schulze, Timo Kosch, Ilse Kulp, Thomas Benz, Andrea Tomatis, Ilja Radusch, Gerhard Noecker, Luisa Andreone, Tanja Kessel, Carola Klessen; „Pre-Drive C2X“ – Deliverable D0.3 Final Report; 30.06. 2010 |
| [EMMA] | Sven Lahde, Michael Doering, Wolf-Bastian Pöttner, Moritz Rosin, Gerrit Lammert und Lars Wolf; „Environmental Monitoring in/of Metropolitan Areas“ |
| [GAW] | Christian Gawron, „Simulation-Based Traffic Assignment – Computing User Equilibria in Large Street Networks“, 1998; <http://www.sumo.dlr.de/pdf/GawronDiss.pdf> (Letzter Zugriff: 04.01.16) |

Anhang A

**Prinzipien [MEY14 – Kapitel 4]**

*(Übersetzung vom Englischen ins Deutsche)*

In der agilen Entwicklung bildet eine Anzahl grundlegender Prinzipien, als methodologische Regeln, die Basis für die spezifischen Praktiken und Artefakte. Diese Prinzipien ermöglichen einen Gesamtüberblick wie Software entwickelt werden sollte. Wir werden uns nun mit den Prinzipien,

1. The final treatment not always is the 3rd treatment. If there are only two treatment steps the secondary treatment becomes the final treatment [↑](#footnote-ref-1)
2. The 3 treatment stems will be explained later in detail [↑](#footnote-ref-2)
3. „Intra-Company transfer“ [↑](#footnote-ref-3)