CZECH TECHNICAL UNIVERSITY IN PRAGUE

Faculty of Electrical Engineering

BACHELOR THESIS

Jan Blaha

Inferring Temporal Models of People Presence from Environment Structure

Department of Cybernetics

Thesis supervisor: **George Broughton, MSc.**

May, 2020

BACHELOR‘S THESIS ASSIGNMENT

I. Personal and study details

Student's name: **Blaha Jan** Personal ID number: **478158**
Faculty / Institute: **Faculty of Electrical Engineering**
Department / Institute: **Department of Cybernetics**
Study program: **Open Informatics**
Branch of study: **Computer and Information Science**

II. Bachelor’s thesis details

Bachelor’s thesis title in English:

Inferring Temporal Models of People Presence from Environment Structrure

Bachelor’s thesis title in Czech:

Odhadování parametrů prediktivních modelů přítomnosti lidí na struktuře prostředí

Guidelines:

1) Research spatio-temporal models of human presence in mobile robotics.
2) Research computer vision methods for detection and localisation of common objects in human-populated environments.
3) Investigate the possibility to infer models of human presence or activity based on the distribution of the objects at a given location.
4) Propose a set of scenarios to evaluate the impact of the inferred prior models on the accurracy of human presence predictions.
5) Select of gather datasets relevant to the proposed scenario.
6) Design and implement a tool capable to automatically evaluate the performance of the individual methods.
7) Perform the experimental evaluation and discuss the feasibility of using objects to forecast people presence over time.

Bibliography / sources:

[1] Krajník, et al: Fremen: Frequency map enhancement for long-term mobile robot autonomy in changing environments. IEEE Transactions on Robotics, 2018.
[2] Krajnik, Vintr et al.: Warped hypertime representations for long-term autonomy of mobile robots. IEEE Robotics and Automation Letters, 2019.
[3] Coppola, Claudio, et al. "Learning temporal context for activity recognition." Proceedings of the Twenty-second European Conference on Artificial Intelligence. IOS Press, 2016.
[4] David Hall, Feras Dayoub, John Skinner, Haoyang Zhang, Dimity Miller, Peter Corke, Gustavo Carneiro, Anelia Angelova, Niko Sünderhauf, 2018. Probabilistic Object Detection: Definition and Evaluation. Available from arXiv.
[5] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).
[6] Kunze, Lars, et al. "Combining top-down spatial reasoning and bottom-up object class recognition for scene understanding." 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2014.
[7] Weiming Zhi et al.: OCTNet: Trajectory Generation in New Environments from Past Experiences. arXiv preprint arXiv:1909.11337.

Name and workplace of bachelor’s thesis supervisor:

George Broughton, MSc., Department of Computer Science, FEE

Name and workplace of second bachelor’s thesis supervisor or consultant:

Date of bachelor’s thesis assignment: **10.01.2020** Deadline for bachelor thesis submission: **22.05.2020**

Assignment valid until: **30.09.2021**

George Broughton, MSc. Supervisor’s signature	doc. Ing. Tomáš Svoboda, Ph.D. Head of department’s signature	prof. Mgr. Petr Páta, Ph.D. Dean’s signature

III. Assignment receipt

The student acknowledges that the bachelor’s thesis is an individual work. The student must produce his thesis without the assistance of others, with the exception of provided consultations. Within the bachelor’s thesis, the author must state the names of consultants and include a list of references.

.

Date of assignment receipt

Student’s signature

Author statement

I declare that the presented work was developed independently and that I have listed all sources of information used within it in accordance with the methodical instructions for observing the ethical principles in the preparation of university theses.

Prague, date ....

Signature

Ackowledgment

I would like to express my gratitude to my supervisor George Broughton and to the scientific guru of the Chronorobotics laboratory, Tomáš Krajník. I am thankful for their endless support with words of motivation in hard times and help in following the ways of science with all that it takes.

Abstract

The goal of this thesis is to present the FreMEn contra COVID project from a technical perspective and experimentally evaluate its impact. To allow for deployment in new areas, because of low amounts of data the possibility of transferring chronorobotic temporal models in the application is tested. Temporal transfer presents a way to deal with extremely small amounts of data, common to robotics, that pose a significant problem to quick deployment of robotic systems dependent on those. Because of the social need caused by the world-wide pandemic of 2020—for which the FreMEn contra COVID project was founded—the temporal transfer has been evaluated in the context of boosting the performance of a system meant to aid individuals to implement social distancing measures FreMEn Advisor app. This app gives recommendations to the time of visits to public locations, where many people concentrate so that the user can avoid crowded times. With the temporal transfer, also the impact of the FreMEn Advisor is tested to the risk people experience while doing necessary tasks in public places, like shops, with different possible sources of predictions of human behaviour in given places. The results show that the risk is significantly lower for users following the recommendations and that transferred models present a promising way to provide recommendations for places not covered by data for exact modelling.

Abstrakt

Cílem této práce je prezentovat projekt FreMEn contra COVID, jeho technickou stránku, a experimentálně ohodnotit jeho přínos. Aby bylo možné systém nasadit v nových oblastech i přes malé množství dat, možnost přenosu chronorobotických temporálních modelů, při modelovaní lidského davového chování, je taktéž testována. Přenos temporálních modelů je způsob, jak se vypořádat s extrémně malými množství dat, běžnými pro robotiku, která jsou velkým problémem při potřebě rychlého nasazení robotického systému, pokud je jeho funkcionalita na temporálních modelech závislá. Z důvodu společenské potřeby, způsobené světovou pandemií v roce 2020—kvůli které project FreMEn contra COVID vznikl—přenos temporálních modelů byl vyhodnocen pro zlepšení aplikace Nebojsa (*angl.* FreMEn Advisor) s cílem pomoci lidem zavést do jejich života principy sociálního odstupu. Tato aplikace doporučuje čas k návštěvě veřejných míst, kde jsou vysoké koncentrace lidí tak, aby se mohl uživatel vyhnout vytíženým časům. Společně s přenosem temporálních modelů i efekt systému Nebojsa na riziko, kterému se lidé vystavují při nutných pochůzkách, je testován včetně různých možných zdrojů predikcí zaplněnosti daných míst. Výsledky ukazují, že riziko je významně nižší pro uživatele, kteří se řídí doporučeními a že přenesené modely jsou slibným způsobem, jak službu zajistit i v místech, kde nejsou data pro přesné modelování k dispozici.

Contents

List of Figures

1 Introduction

Since ancient times, humankind has invested much effort in making the life of human easier and in making stronger, more sophisticated tools. Tools, enabling him to build object and perform tasks at higher speed, with better precision or simply out of reach of capabilities of the human body.

At one point in history, a new task for researchers and engineers emerged, that is freeing human of the need to operate the tools and machinery in the full extent. This gave birth to the field of robotics, a field designed to bring human better tools—robotic systems—that need as little supervision as possible. For many years the development in the field of robotics has brought human such robots.

As the robots became more sophisticated, and their purpose has got extended from one specific task for a wide range of capabilities, the main task became to allow for the autonomy of such robotic systems. First, autonomous robotic systems operating entirely on their own were performing given tasks in controlled environments of industrial sites.

Because the domain of robotic systems were not only industrial applications, even from other domains arose a demand for full autonomy. At the time of writing, full autonomy in natural environments is by far not solved question of research. On all levels of the autonomous robotic system (ARS)—being it low-level hardware or its control software, higher-level software, artificial intelligence and machine learning tools or algorithms for multi-robotic systems— much research is being done.

To deal with problems of long-term autonomy, that are tied to human-populated and natural environments, a field of chronorobotics, discussed in this thesis, was founded.

In the robotic research—including the field of chronorobotics—many methods and results have been developed. Such methods usually manifest properties that come from many specifics of the domain of robotic applications. This by no means implies, that their only possible use is in robotics. While there are many strongly tied, for example to hardware operation or sensory fusion and alike, many are much more general and can take advantage of their traits to solve problems in other domains.

This thesis ambitiously takes on two tasks. One lies in showing a successful transfer—in the sense of a technology transfer—of robotic methods for temporal modelling into a software product with human users and also preliminary testing of its impacts. The other lies in testing possibilities of boosting the learning time an ARS needs in a new environment for creating temporal models of quality necessary for deployment, using data already collected in different environment—a temporal transfer or transfer of temporal models.

Initially, only the second was supposed to be the main task of this thesis, but due to an unexpected crisis of the world-wide pandemic of 2020, all efforts in society world-wide have been redirected to help as many people as possible. Chronorobotics laboratory at FEE CTU in Prague[1], of which author is a member, initialised a collective effort, called

---

[1] Chronorobotics laboratory of Artificial intelligence centre at Department of computer science

FreMEn contra COVID, of deployment of its existing methods in a software product, helping people with implementing social-distancing measures into their lives. Since any such system needs a lot of data to cover large, populated areas, the possibility of temporal transfer would allow for much faster deployment in a new location.

Because the initial topic of the thesis coincided with the social need at the time of writing, the author decided after consultation with his supervisor to change the target domain from an ARS deployed temporal modelling methods, to the FreMEn contra COVID project (the Project), because he was in charge of the development of the software products for the Project. This does not neglect the initial proposal, but instead of understanding the distribution of common objects as the layout of individual objects inside for example a human dwelling, the author adopted understanding it as the purpose of different locations determined by types of objects commonly found at such locations. Therefore computer vision methods are not discussed and researched, as this topic no longer is relevant for the rest of the work.

Only one method for context-based temporal model transfer is tested. The interpretations of context can be numerous, and the main target of this thesis was to boost one specific application domain using high-level contextual information about the purpose of locations. For that, the focus was on the possibility of such transfer, its justification and proposing of one simple, quickly deployable method, rather than a comparison of a large number of possible approaches and evaluation of small differences in their performance.

Because of the strong motivation for this work in the existing social need, in the experimental section, only temporal modelling method deployed by the Project is considered and compared with different sources of temporal information commonly accessible online.

1.1 Organisation

The thesis is organised into five Sections, first being this Introduction. Second Section is titled “Related work” and groups theoretical research, the field of chronorobotics is discussed in detail, the Project and temporal model transfer as well.

Third Section “Project” describes in detail the software products developed as part of the Project. Mainly the overall architectural view is presented, as that was designed for the Project by the author himself, but essential algorithms are described as well. The temporal modelling algorithm is a work of the Chronorobotics laboratory. The author of this thesis claims authorship of the presented algorithm for recommendations.

Fourth Section “Experiments” describes the conducted experiment. First, the goals are stated, then data it was based upon is listed. The section concludes with a presentation of the results and their discussion.

The whole work is concluded with fifth Section “Conclusion”.

2 Related work

This section further describes temporal modelling and with it the newest results of the field of chronorobotics. Also, it presents the general context of work on the FreMEn contra COVID project.

2.1 Chronorobotics

First, a general view of spatial representations is presented, which are in standard robotics methods static, lacking the ability to represent dynamical phenomena. The gravity of this insufficiency when robots are deployed for extended periods of time motivated researchers to develop a new set of methods, that are able also to capture periodical temporal context. This resulted in a new field of study called chronorobotics, that focuses on "investigation of scientific methods allowing robots to adapt to and learn from the perpetual changes occurring in natural and human-populated environments"[2]. The complete motivation is analysed in part 2.1.2.

Results of the field of chronorobotics are then presented in part 2.1.3.

2.1.1 Spatial representations

One of the main tasks of any robotic system deployed in a previously unseen environment is to secure its reliable navigation accounting for the uncertainty of the sensorial observations as well as of the environmental properties. There are multiple ways to solve this problem, dividable into two sets—navigation in a map and reactive navigation—that differ on whether they employ localisation in some kind of a spatial model or not.

The first approach lies in the creation of a spatial representation—model of an environment—usually in the form of a map, which allows the robot to move reliably, plan trajectories, understand the purpose of individual sub-locations, such as rooms in a building. Of course, there are scenarios where such a map can be provided in advance, but those are usually not very interesting for the general long-term autonomy and life-long learning. Because of that, many methods of probabilistic mapping were developed, that allow the robot to create the map from noisy sensory data, while at the same time localising in it—collectively named SLAM (for Simultaneous Localisation And Mapping)—thus effectively performing assigned tasks in a location where the map is not known in advance.

The second approach builds on the analogy with human behaviour. In many situations, the navigation performed by people does not rely on a precise knowledge of the environment but instead on the ability to determine the general direction of the target and the sequence of steps needed to achieve it or at least approach it to some extent. It has been shown, that for some robotic tasks, like the teach-and-repeat scenario the precise localisation is unnecessary for state-of-the-art system.[3]

For most of the common tasks robot will eventually need some kind of spatial model, to be able to reason about its actions, their outcomes, usually with the use of previously collected sensorial data, processed or not, that either need to be fitted into the frame of the existing spatial model or it can represent one if the data also contains a spatial component.

2.1.2 Spatio-temporal environmental representations

The problem arises when the system is deployed for a longer period, where environments natural and common to people tend to change their properties with time. Deterministic robotic systems need to account for such changes, and even the spatial representations they employ must be adapted for this. This presents a significant problem as most of the standard methods for creating such maps assume the only source of uncertainty to be the noise from used sensors. Such an assumption is no longer valid once the models are supposed to understand the environmental property, that is the subject to the observation, to be governed by a probabilistic process of its own, originating in the nature of the given environment. A robot has to be able to model such dynamics, to understand its surroundings and perform better in assigned tasks. [4]

Most of the spatial representations build on the idea of discretisation of space[5], being it on the low-level cells of occupancy grids or the individual locations of a topological map. The simplest way to expand the capabilities of spatially discrete representation with the information about temporal context is to employ temporal-only model for every individual part of the considered space. For temporal modelling, important characteristics of such extensions of standard types of spatial representations are, that the variables which need to be modelled are usually binary. Whether is it the visibility of a fiducial marker in a markers-based map or the traversability of an edge in a topological map, these are usually represented by binary values—true/false or 0/1. So in terms of probability theory, these are modelled using Bernoulli distribution. This is of course, not a general property, but it is prevalent in spatial representations and presents a narrow use-case, on which the development of efficient methods was possible.

2.1.3 Temporal modelling

Temporal models are generalised representations of phenomena with a temporal context. Such a model needs to provide at least the following capabilities to be useful in robotic applications: update with a new observation and prediction of the state of modelled phenomena in the future or retrospectively in the past. Without the updating capability, the robot would not be able to adapt to a new environment or learn in general. The prediction ability is then the basis to the application of the gathered knowledge to performed tasks—once the robot needs to reason about its surroundings it needs to be able to access the model's encoded information efficiently.

This subsection presents several possible temporal models. The following list is not meant to be exhaustive, as the topic of time-series analysis has been a subject of active research since the 17th century [6] the field presents many results, the vast majority outside of the scope of this thesis. What strongly differentiates robotic applications from methods developed, for example, for time-series analysis in economics comes from the nature of the modelled data. The dominating characteristics are

1. small amounts of data — data in robotics are expensive in terms of the time it takes to collect them because that usually means the robot would not perform other assigned tasks, and in many situations, most of the data cannot be observed, for example, if two events happen at the same time far apart from each other,
2. random distribution in time (and space) — if the robot has got other tasks than data gathering, it is going to receive observations along its trajectory, which can be a random process if the goal changes in time (a very common scenario in service robotics),
3. uncertainty of observation — as discussed above the data in robotics are very noisy, which in combination with their small volumes requires more caution.

Non-trivial result of previous research also states that in human-populated environments the dominating processes of medium-term (i.e. weeks) temporal evolution to be periodical.[7] Most of the changes in such environments come from people, like opening or closing doors—effectively changing the map of traversability in a building—or natural processes, like the day-night cycles in lightning conditions. When a change occurs, and the process no longer exhibits strong periodical evolution, then in most cases, this is due to the change in the generating process itself, where for example installation of automatic light intensity triggered blinds might change the way movement of the sun exhibits itself on lightning conditions inside of the given building.

All the methods are described as modelling Bernoulli distribution of a variable over some time interval. I.e. they can be characterised as a function f of parameter t for time, returning parameter p to Bernoulli distribution, denoting the modelled variable to take the value 1 with probability p and 0 with probability $q = 1 - p$. The generalisation to a more complicated scenario, like categorical distribution, is possible, but only for the specific method used in experiments complete description will be provided.

Mean Listed here only for completeness, mean presents an important baseline method for evaluation of the performance of temporal models, due to its optimality on the training dataset with respect to the popular RSME metric. It neglects any dynamics to the temporal evolution and statically estimates always one value—the mean of all observations. As it is very simple and also does not need much storage space, it is also the first choice for many situations where one does not expect any reasonable periodic evolution over time.

Histogram One of the simplest methods for estimating probability distributions build upon histograms. The foundational idea being that with enough draws from a probability distribution, we can estimate density in a particular value of the domain by measuring how many—proportionally to their total number—draws fall into some neighbourhood of the given value.

Modelling any temporal phenomena with histogram is not very useful as new observations come chronologically and with time new observations are more distant from the previous once in the time domain then those before—they are not repeating, and the domain is only growing. If the observations are considered as ordered by the time of measurement, then for any division of the range $[first\,observation,\ last\,observation]$ into bins, the resulting histogram only interpolates the values, not giving any more general insight. With the assumption that the process is periodical with a period T, following a trivial transformation of the time component of the observed data (t, v) can be used:

$$(t, v) \mapsto (t \pmod T), v), \tag{1}$$

where t is some linear representation of the time of observation and v observed value. Such transformation then maps all the data onto one time-interval of length T, which allows for sound statistical reasoning about the observed process for any time t, based on observations.

The way histograms may apply for modelling of the Bernoulli distribution is that two models are created "Zeros" and "Ones"—of the same properties like the number of bins, their position and their sizes—one accumulating the number of observations for individual bins, where the observed value was 0 and the other where the value was 1, with the assumption of periodicity T. The estimation of the parameter p of sought distribution at time t is then

$$p = \frac{Ones_T(t)}{Ones_T(t) + Zeros_T(t)}, \tag{2}$$

where $model_T(t)$ denotes the number of observations in the *model* for time $t \pmod T$.

Regarding the employment of this method apart from disadvantages that come from histogram itself—like the need to specify bins, at minimum their size—suffers from the need to set the period T explicitly. This method then can explain the effects of processes, that have their period smaller then T and T is dividable by it. Research shows, that when the main subject to modelling is people, then a reasonable choice is either a period of a day or a week, which is able to describe most of the common habits people display.

From the perspective of performance, the main problem of this method is the need for a lot of data. All the bins need to be covered by the data enough to make a meaningful estimate of the ratio between individual values of the modelled variable.

FreMEn This method tries to model generating process as a simple periodic function. It is based on spectral analysis. Specifically, it uses a modified version of the Fourier transform described in [8], that does not require data to be sampled in the time domain with a constant frequency and also allows for incremental computation, which is very convenient as robotic systems usually need to work with limited resources.

The output of this method is a list of triples $(\gamma_i, \omega_i, \psi_i)$, where the γ_i is the amplitude, ω_i is the frequency and ψ_i is the phase shift of one of the components of the resulting

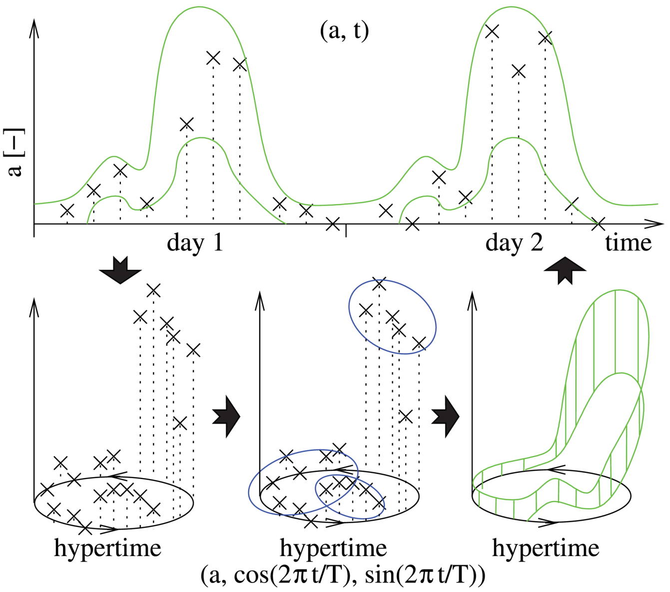


Figure 1: Visualisation of the hypertime method and its individual steps—the hypertime space expansion (A), clustering or other statistical method (B,C) and back projection (D). Courtesy of [1]

modelling function. Using this list the resulting modelling function can be written as

$$f(t) = \gamma_0 + \sum_{i=0}^{n} \gamma_i cos(\omega_i t + \psi_i), \tag{3}$$

where n is one of two parameterters to this method. The parameter n describes how many of the most significant components to include in the predictive function. This is usually estimated using cross-validation during training, so it is not necessary to specify it beforehand and can be dynamically changed during deployment, but if wanted it is possible to provide the method with some default value.

In contrast with histogram, the number of periodic processes that FreMEn can model is determined only by the parameter n. What is necessary to specify in advance is the maximal considered periodicity, which is needed for the spectral analysis.

FreMEn has been deployed in many scenarios, for localisation, navigation and more. It has been integrated into the navigational stack of the Robotic Operating System (ROS)[9], which allowed the robot to access maps temporarily local to its operation in an abstracted way, so no change to the user program was necessary. Experiments comparing the performance of standard maps to their temporal extension in the individual tasks have shown it boosting both indoor and outdoor localisation and navigation [1]

Hypertime The simplicity of implementing FreMEn into existing discrete spatial representations comes with one problem. Many of these representations are very dense because the operating robot needs a high degree of resolution. Very common is to partition space in occupancy grid representations into cubes of 5 cm long edges. The number of such cells then grows to the power of three with most of the cells being empty because they represent air. In the standard occupancy grids every cell is represented by one number, whereas with extension by FreMEn, every cell is represented by the chosen number of triplets, as discussed above. Such representation grows in size significantly with most of the space being wasted on constantly empty cells.[10]

These problems have led to the development of a new modelling method, that is capable of respecting time and space in a continuous manner.[1, 11] Due to the physical nature of the world around, the raw data comes in as many individual data points that exhibit extreme spatial correlation. Because of this, this approach is even able to work with much more sparse spatial data, then FreMEn operating over individual occupancy cells.

The first step of this method uses FreMEn to identify periodicities $T_1, \ldots, T_n$ in the training data. Based on these a transformation called "hypertime space expansion" is applied

$$(t, a) \mapsto (cos(\frac{2\pi t}{T_1}), sin(\frac{2\pi t}{T_1}), \ldots, a), \tag{4}$$

where (t, a) is one data point, tuple of time t (it's linear representation) and observed value $a \in A$, and $T_i, i = 1, \ldots, n$ are periodicities identified by FreMEn. This transformation projects all the data into a higher dimensional vector space, a cartesian product of a Clifford hypertorus and a vector space A from which values of the observed variable are drawn.

Over these extended values a clustering can be performed, which then gives an estimate of density in the combined space, which can be then projected back onto a standard, linear time axis providing estimates to the value of the observed spatial (and temporal) phenomena based on time. An intuitive visualisation is depicted in Figure 1.

This method has a wide range of applications. From the perspective of the data, it can model theoretically any kind of time-dependent periodic process. It has been applied to robotic binary time-series modelling [12], pedestrian density, i.e. the observed values were twodimensional positions [10], and even pedestrian density with the direction of the movement included [13]. The method performed well in all of these scenarios, at least as well as standalone FreMEn, while maintaining much smaller (up to 5 orders of magnitude) storage requirements when spatial data were modelled.

From perspective of performance in application, the Hypertime method has been applied to anomaly detection in Bernoulli distribution on a robotic dataset, in comparison to the state-of-the-art method, and showed to converge fast (up to 10 days before any other method) to desirable result [12]. Different experiments with long-term robotic datasets for localisation and navigation were also performed, showing it outperform FreMEn and other used methods [1].

2.2 FreMEn contra COVID

During the December of 2019, in Chinese province Hubei an outbreak of a new type of coronavirus, named SARS-CoV-2, causing respiratory disease COVID-19 sometimes causing pneumonia with a possibility of acute respiratory failure, began. During the first months of 2020, it turned into a worldwide pandemic, impacting most of the Earth's population in some way, be it a direct consequence of the virus presence or measures to mitigate its fast-spreading. [14]

The efforts to get epidemiological models fitted for the disease started soon, for example [15], and their results were, that the virus has a high potential for spread. Because of its spread through air with droplets of saliva, apart from more intensive hygiene, list of recommendations to prevention from World health organisation (WHO) also contains social distancing as a way of limiting the potential of exposure to the virus.[16] Social distancing, also known as physical distancing is a set of measures in which people limit their visits to public places, meeting friends and relatives and keep physical distance when dealing with other people. If the number of contacts people have is small, then the potential number of people one infectious person can meet is much lower than under normal circumstances.

2.2.1 Other approaches and technical solutions

Of course, many initiatives arose to the challenge of a world crisis, advocating different solutions and principles to fight the spread of the virus. This section tries to provide a brief overview of those solutions and ideas, with a commentary of shortcomings and disadvantages. First technology-based solutions, targeting individual people are presented and then more general instruments of governmental bodies.

Contact tracing The most prevalent idea for a technological solution lies in contact tracing. Contact tracing is a method for identification of potentially infected people, by accessing the history of human contacts of people with a confirmed positive diagnosis. Many implementations have emerged, based on either a GPS location tracking—for example solution by Seznam.cz[17]—or a Bluetooth technology (BT)—for example eRouška[18]—that make use of personal mobile devices of users.

While BT based schemes have seemed to be prefered over GPS tracking because of privacy concerns tied to location tracking, because they only store lists of encountered users, the problems generally lies in need for always-on service. The limitations of mobile platforms lie in the way operating system (OS) handles application, that can be usually preempted "at its will". The two most dominant providers of OSs for mobile phones—companies Apple Inc. and Google LLC—have teamed up and came up with a vendor-interoperable BT-based software development kit (SDK), moving the responsibility from the user app to the system level of the OS. Unfortunately, while two GPS recorded locations present an opportunity for interpolation of the trajectory when BT scanning is not active for a short period of time, the data is not recoverable.

Apart from all the concerns about personal privacy and limitations of technology, the main disadvantage is that such a system needs a massive penetration in the society to work, as its efficiency grows with the second power of the penetration. It also requires the users, to have their device on, with the running application and with them.

Reservation systems Of all public places, very frequented ones are shops or for example offices. Places like those usually have one simple purpose, and their operation is dependent on some staff. That allows them to deploy some reservation system, which visitors use to block some time interval for themselves.

The clear disadvantage of this is that such a system needs much support from the side of the administration; therefore, it is not possible to use, for example, for parks. Moreover, such a system will strongly depend on the standard functioning of the services at the location, so no universal system is possible, and the user would need to use different reservation systems for all locations they visit.

Real-time business reporting Due to the spread of mobile technologies, it is possible to develop and operate a system, capable of reporting in real-time business in various locations. It can be based, for example on location services of OS of mobile devices, or mobile network operators.

The technological drawback is that to be covered reliably; a location has to be frequented a lot by users of relevant technologies to be noticed by the system. However, the main problem of the idea lies in a limited list of cases, where a live data are useful—a normal person needs to create its schedule ahead of time, so real-time is of little use.

Sharing position of infected people By general public most demanded tool is one, that would be able to show in real-time where is a high risk of infection, because of the presence of an infected person. However, before proper diagnosis, there is no way of knowing who is infected and who is not and once diagnosed, patients are usually required to stay at home. Second, when dealing with infectious diseases, of great challenges is the prevention of stigmatisation [19]. One of the sources for stigma is public labelling of infected [20]; these can then face ostracism or even physical violence. Third, there exists a hypothesis, that this can lead to stimulation of spread because people that would be in an area of high risk would be tempted to getaway.

Stratification – spatial ("Geofencing") and temporal When dealing with an exponentially spreading virus, the primary standard goal is to limit the number of human contacts. If the society would be divided into groups of people, that would never interact, then the virus could not transfer from one to the other. This division can be done in multiple ways.

First out of two tools of governments is spatial stratification of the population. This is restricting where people can move to some defined areas. An example can be a city lockdown issued in the city of Wuhan, even though in this specific case, the target goal was more in limiting the transmission of the virus from an epicentre.

Stratification of society does not need to be only spatial. One of the employed measures during the pandemic of 2020 consisted of devoting specific times of day in for senior shopping[21], as seniors were indicated as facing higher risks when infected. [22] In those times no person below some age was allowed entering into shops.

Even in terms of the whole society, there are schemes, that would divide the whole society—for example, the proposed regulation, where people could enter shops only according to the last one or two digits of their credit card.

Even though these schemes are generally considered to be quite effective, their strongly restrictive nature makes them hard to authorise on large scale.

2.2.2 Project motivation and targets

Even when implementing social distancing or under restrictive orders requiring people not to leave their home, that some governments implemented, people still need to visit grocery stores to get supplies or pharmacies to get necessary medications. When restricting orders are kept for a long time people also need to leave their homes to take a walk and get some fresh air to protect their mental health and to support their immune system, both of which is important for fighting any kind of disease.

The main idea of the FreMEn contra COVID project was to give public a tool for better decision making as to when to perform such necessary activities in public places, while maintaining distance and choosing times, when these places are less crowded, combining the results of medical science and chronorobotics. By applying temporal modelling to people densities in various locations, the system should be able to predict the future states, thus allowing the user to choose a time of visit with low potential of exposure, because this potential drops with the number of people present.

2.2.3 Robotic research perspective

The research in modelling people presence over time even in the spatial domain has shown, that application of methods from chronorobotics can help many robotic applications. The basic scenario of modelling people presence over time and space in some location was explored in [1, 13]. Temporal modelling has been applied to the search-and-rescue scenario, where a service robot was supposed to find its owner as fast as possible because of an emergency situation. [23] The use of the knowledge of temporal behaviour acquired during the previous operation was shown to shorten the time needed to find a person in its apartment. [24] Another use-case was boosting recognition of human indoor activity, by modelling the prior probability of its happening as a temporal phenomenon. [25]

Previous results conclude the suitability of presented tools for modelling of statistical properties of human behaviour. Thus when working with people and natural processes, methods of chronorobotics can be used even outside of the robotic domain, while maintaining all the properties, that come from the specific needs of robotic applications.

What this project brings new is the concept of collective information gathering as being a process separated from the agent—the user—using it in its operation. It creates a system where individual agents can split the time needed for exploring the environment and can reason about a place they have never even seen.

2.3 Temporal transfer

Transfer learning is a general task in machine learning, where a domain-specific knowledge gained during solving of one problem is applied to solving a different problem, mostly from the same application domain.

In the field of chronorobotics and modelling human behaviour, research has been conducted, for example, in the prediction of complex pedestrian trajectories [26] from an environmental structure. It is based on generalising the representation of trajectories and how they depend on environmental structure. Then when faced with the unseen environment, the method is able to generate likely trajectories by analysis its structure.

On presented temporal and spatio-temporal models, research in this direction is starting with this thesis.

3 FreMEn contra COVID project

The main objective of FreMEn contra COVID project is to give people necessary information allowing them to act rationally while implementing social distancing measures. In the same way, people are used to consulting weather forecast before leaving home for long time and reasoning, whether to go out in the rain or not, they should have similar kind of information source regarding the expected business of a shop, or alike they are going to visit, as these places present a high risk of exposure.

One important distinction of this project from all the other big projects that arose during the world pandemic of COVID-19 is the effort to present an alternative to general surveillance and breach of privacy that comes from people tracking. It has been argued, that from the experience of Asian countries like South Korea from SARS epidemic of 2003, quick tracking of people using modern technologies and tracing of positively tested is the only effective way to fight the spread of the virus. This project turns its attention to finding a way that would be compatible with the principles of western democracies, preserve individual freedom and pose little to no threat to the privacy of its users.

3.0.1 Sources of data - analysis

This Section enumerates possible and considered sources of data for the recommending system. This list is not chronological, all the possibilities were considered simultaneously, the specific reasoning to each one is presented, finishing with the one, that was adopted.

The first idea to this project was to get data about the crowd creation from holders of big data like telephone companies. This turned out as a not feasible solution, because of regulation Moreover it would also work only in larger cities and not very reliably, because the accuracy in determining the location of mobile devices using trilateration is insufficient. On small villages the number of base transceiver stations (BTSs) is usually quite limited, for example, accounting for one station per village and one larger for the whole area. Two stations are not enough for a precise trilateration, so this idea was abandoned.

Another, much more precise information would be the number of credit card transfers from a particular shop. This information is of course much harder to get as this would have to be cleaned from all the details of the payment, credit card number and other sensitive information and banks are in general conservative and one of the most regulated domains. The main disadvantage of this source of data is that it covers only trade-focused locations, but completely neglects public places like parks or offices.

Last rejected approach to data gathering was the effort to engage people in charge of the individual locations. It would create some overhead in reaching out to them, but for places under active administration (any place with always present personnel), this could bring very precise data with a full-time coverage. The problem of this approach is that for most shops, the actual number of customers is sensitive information, they do not want to give up because of competition.

The final decision was made that the data will be gathered by the users themselves in a crowdsourced way. Any user could cover places in their neighbourhood with data and help others in their vicinity in the spirit of protecting the more vulnerable. Based on this data, anyone visiting a particular place can act according to the predicted situation. This approach also covers any kind of public place with no distinction, because the active user can submit their observations from anywhere.

3.1 Architecture

This Section describes the target architecture of the software product of the Project. First, a list of requirements and their analysis is presented. All decisions that were made in the design process are based on these. Then a top-level of the created architecture view is presented and commented. Following this, the design of individual components is discussed.

3.1.1 Initial requirements

The list of requirements can be divided into two sets, based on the viewpoint of their origin. The first set of requirements comes from the functionality, that should be presented to the user, the way they should interact with the system, the usability and privacy preservation. The second set was defined by the long-term goals of the Project, its personal resources and its philosophical motivation.

User view First are listed requirements related to data collection and second those related to the use of the output of the system by users.

1. The user should be able to submit information about the number of people and their density in their favourite location.

   This information must contain position, time, number of people and the level of business at the time. For a better analysis of the data type of location should also be submitted, so that we can correctly distinguish between observations close together, but of a completely different generating process.

2. There should be no way to connect the submitted data with an individual user.

   If there was such a way, then we could get a list of observations of one user, effectively getting their historical trajectory and a list of places they visit.

3. The presentation of information to the user should be clean and straightforward, so even the users less skilled or unfamiliar with new technologies could use it.

4. The user should be able to find a place they want to visit on a map.

   To keep the simplicity of usage and with the assumption that users will use the system repeatedly, there should be an option to keep favourite places in one place.

5. The user should be able to look up information relevant to choose time to visit the place that interests him.

   With simplicity in mind, the user should first receive a recommendation as to when to visit the target place. The full predicted graph should then be optional but easily accessible information, as the given recommendation cannot always match with other constraints of users personal schedule.

Requirements arising from the project vision

1. During the design and implementation, it must be proceeded with extreme caution as the system deals with people and can potentially put some at risk.

   When dealing with actual humans, all measures must be taken to prevent any harm that might arise from the use of the system. There is a higher chance for people with some previous medical condition, that are at higher risk of complications from exposure to the virus, so take higher precaution, than people healthy.

   One of the implications of this is that the system needs to beware of giving all its users the same recommendation—for example, the time of minimal potential for crowds. If the users all followed it, the effect could be inverted—a crowd could form.

2. Preserving the privacy of users is a top priority.

   Apart from complying with the standard regulation like the European General Data Protection Regulation (GDPR), the Project must ensure that no personal data are being used or manipulated in the system if not necessary. This is due to motivation of the Project, to provide an alternative to mass surveillance, that can preserve the privacy of its users while at the same time help to deal with the 2020 coronavirus outbreak.

3. Despite the purpose of the software being specific, targeting epidemiological measures, the underlying ideas of modelling spatio-temporal processes on a large scale are very general.

   Many of the principles are very general, so the resulting software could be easily reused in other domains with all its benefits if designed properly.

4. The backend of the system has to be reliable and scalable.

   The Project has no target goal as to the number of users, but from its nature, more users will follow the recommendations, greater the impact it will have on the epidemic situation. Because of the reach of the world pandemic, the goal is applicable and relevant in almost all countries around the globe, so the potential number of users is very large.

   During the life of the Project multiple potential partners appeared, that considered the incorporation of the backend systems into their existing products, with as much as two millions of users. This means that the backend has to be able to adapt to a quick growth of the number of requests it has to be able to serve.

5. The whole system must comply with all of local regulation.

   With the ambition to provide discussed service all around the globe, the Project must be able to comply with the regulation in all countries its users are. One of such regulations is that many countries consider the data the system operates with as sensitive on a large scale, as it describes the habits and behaviour of its population. Therefore the data must remain in the country it originates from.

3.1.2 Top level architecture

There are two points of view of this architecture. One captures the way data flows in the whole system, starting by its creation and ending by the presentation of the modelling results and recommendations to the user. The second one focuses on the software side, describing how individual parts were designed to be implemented and all the supporting parts for a legally sound and scalable system. It is further divided into two layers—one dealing with the regionality of data and the other with the data for one region.

Data flow The way data flows in the whole system is depicted in Figure 2, where the individual stages are labelled 1 to 5. For better organisation, every stage is described independently.

1. Creation of the data.

   First, the data needs to be created, which happens in the device of a user. An interface in the form of a mobile application is provided to the user by the Project, called “FreMEn Explorer”, which consist of a very simple form, where the user can enter the number of people they see, their subjective density and a type of place they is at. The app then appends the current GPS location and time, which completes the list of information the system needs to work.

   This app is implemented for Android and iOS devices natively, so using Java and Swift technologies.

   The data, of course, does not have to originate from the provided app as long as it fits with the rules for its generation. As the Project has potential partners, the data can come from those as well, as they can ensure the process correctness, while implementing this frontend into their existing product.

2. Collection and long-term storage.

   All data after the creation has to be concentrated in one place for further treatment. That is the collection backend to which the sources upload the data upon its creation.

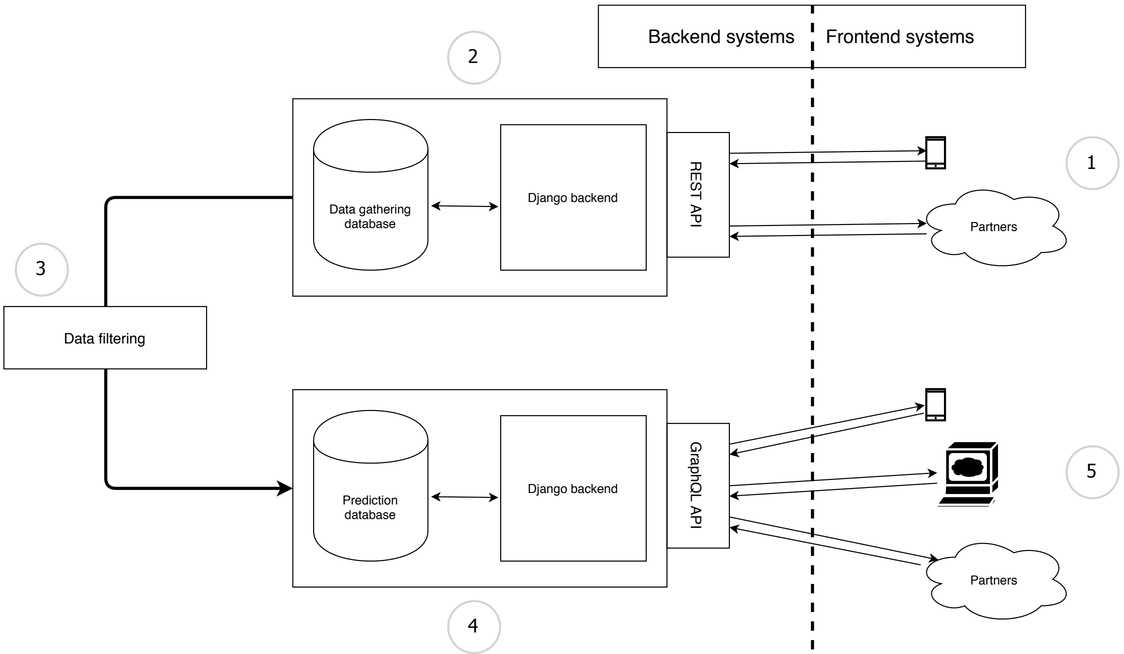


Figure 2: Diagram of how the data flows through the whole system. The data is created (1), collected (2), filtered for assuring quality (3), used for predictive models and computing recommendations (4) and finally presented to the user (5).

This backend is implemented in Python using Django framework[27] for web development. Aside from the Project having developers familiar with the technology, the development in Python is generally fast and Django allows to leverage a large set of tools already prepared for developing such applications in contrast to for example the minimalistic framework Flask, so it was a natural choice.

This backend system exposes a REST API implementing endpoints for submitting individual observations and upon success returns back a number of submissions in the given area (500m radius) in the last 24 hours, so that the user gets feedback on the status of data collection for the place they are submitting.

Collected data is stored in a persistent long-term storage unit. Specifically, an open-source Postgresql[28] relational database was chosen with the PostGIS[29] extension, that adds the functionality for working with geospatial data, including spatial queries, which are necessary for the creation of the models. A relational database over a NoSQL one was chosen for very simple interfacing through Djangos ORM, that supports all of PostGIS features and even allows managing the database definitions as a Python code.

3. Filtering of collected data.

   Next, the data must undergo a process for ensuring its quality. [2] For this, some

---

[2] In the time of writing, this part of the system has not yet been implemented or even specified to the

additional meta-information is stored with the data in the collection database. It is a hash of IP address from which the submission request came, a version of the application and a randomly generated string that gets regenerated on every start of the mobile application.

The hash of IP address gets deleted after two weeks, so it cannot be later accessed, but in case of a large scripted attack on the database, this can help to detect it and remove all flawed information coming from a single source. But the exact source cannot be determined.

The version of the application can affect legal issues arising from the user agreeing to some terms of service or alike. Because of that, it is important that the information connecting data with the terms under which they were provided, is persisted in the database. The further treatment of the data should be decided in the filtering stage.

The randomly generated string is a measure very similar to the hash of the IP address. Where it differs is, that there is some pattern to how this is generated, so even if the malicious user would change their IP address, which is simple, they should not have access to such a key. If they get one, it is probably by eavesdropping on some authorised traffic, where they can get only a limited number of those, allowing for filtering out all data submitted with the same key.

The last part of the filtering process is based purely on the data itself. It is a set of hand-crafted rules that specify what data is too unlikely to appear, so it should be discarded. An example of such rule can be, that two data points of the exact same GPS coordinates (or recurring set of points with coordinates similar to the resolution of several meters) are very unlikely to occur in a short period of time.

4. Modelling and recommendations.

   When the data has been filtered out, it is moved into temporal storage, a database of the same type and properties as the collecting one. From this database, the data is served to all of the modelling and recommending computations. Apart from the raw data, it contains cached results of temporal models.

   The backend system for the analyses of the data in the prediction database(s) is written using the same set of technologies as the data gathering backend system, i.e. Django framework.

   The user can have two requests for this endpoint, the service being exposed through a GraphQL API. As it is supposed to provide services to frontends presenting the recommendations, it has to provide a set of locations in the requested area. The second type of request is the recommended time and graph of estimated business based on the requested time, place type and its location. GraphQL was chosen in the later stage of the Project as a better alternative to a REST-type API for spec-

---

full extent. Therefore only the ideas and support information prepared for this stage are presented. The implementation consists of a scheduled migration of data from collection to prediction database.

ification and communication between the development teams, with a performance benefit of easily combining multiple requests into one.

Specific employed algorithms are further described in their own Sections 3.3,3.4.

5. The presentation of results to the end-user.

   As presenting the results to the user is a task requiring much less control, in contrast to data gathering, the Project has developed apart from mobile Android app "FreMEn Advisor" also a web application.[3] Version for Apple devices was planned, but not finished, because of the time pressure and the amount of work.

   The list of platforms, where these results could be presented is very large, including the existing solutions of several potential partners of the Project. The API is prepared for having a version of the presentation layer done by a different group of people.

World level At the topmost level, the system is supposed to be able to cover any desirable area. Due to the principle of spatial locality, where the for the computation of a temporal model for one place only the data in its proximity are needed, a solution to this problem lies only in data management, so does not prevent from scaling the computations over an arbitrarily large area. The constraint by the local regulation that prohibits the data from crossing any border is the main motivation for the highest level of architecture, but also does not introduce any changes to the computations, as the data only is relevant in the place where it was created; therefore the solution lies in data management.

The solution has two parts. First part needs to cover the data gathering, where the data upon creation must not leave the country of origin, so the decision of the target collecting service must happen on the creating device, before submission. The second part handles the fact, that the data for the creation of models is fragmented across the globe, but the user can theoretically request the service for any place, however distant from their current location, so the request then has to be dispatched accordingly to the regional instance, that holds the necessary data.

The layer implementing this functionality on top of the basic one is depicted in Figure 3. The main idea is to split responsibilities and data between some regional instances, maintain a centralised table of these and its distribution between all parties. All happens in 4 kinds of requests:

1. A centralised server is upon request serving current version of a Central routing table (CRT), consisting of location-names and their respective named servers, responsible for the given region. Every device operating a mobile app must ensure its copy of CRT is not older than a week if it is being used.

---

[3] In the time of writing, the web application has not yet been released to the public, because of minor technical issues.

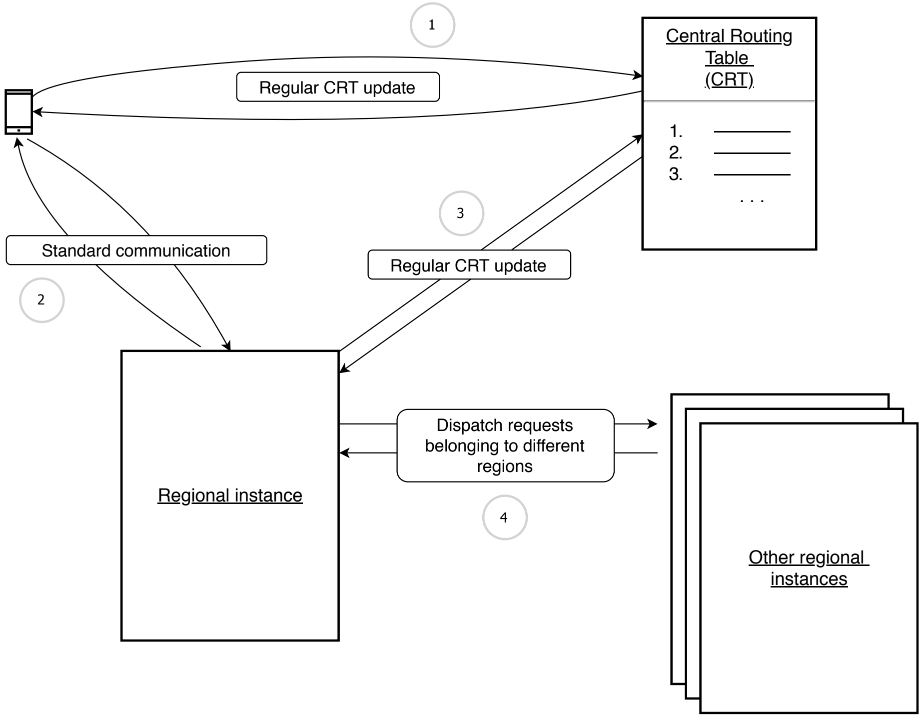


Figure 3: Diagram of communication in the "regionality layer". User device regularly request a copy of Central Routing Table (CRT) (1) from responsible server. Standard communication (2) happens according to local CRT. Regional instances also request CRT regularly, but with higher frequency (3) and if they receive user request not for their region, they dispatch it according to their copy of CRT.

2. In a standard communication, the device requests the name of its current timezone from its operating system in the long format, including a location, like "Europe/Prague". Based on the location, the device chooses to communicate with the instance responsible for the given location. This way, in the data collection process the data is sent directly to the responsible server, and in case of the recommendation application, the request is also targeted directly to the regional instance with a high chance of it being geographically close.
3. The individual instances themselves must keep a local copy of the CRT, in case they receive a request for a recommendation concerning a place not in their region. In such a case, they need to know where this request should be rerouted. Every regional instance should update its copy on a daily basis because unlike for the mobile devices, it is not much overhead for an always running server. This way, introducing a new regional instance with expansion takes only one day from entering of the record into CRT to propagate the information to all currently running instances.

4. When a regional instance receives a request outside of its region, then this request is forwarded to the responsible instance. The result is served back to the user device by the first instance in a proxy-like fashion.

Regional level On the level of regional instance, we can divide the solution into two almost independent parts according to the data flow and the respective frontend applications. One takes care of data gathering and the other of recommending times of visit. As the target architecture is supposed to be scalable and robust, it makes use of the principles of the microservices architectures, where the service is divided into independent components, that focus on one task only and provide some API to the rest of the system. Such an approach brings advantages in the organisation of code and development work, but most importantly if designed well brings naturally ease in scaling the solution. If combined with containerisation technology like Docker[30], then even deployment is much easier. This can be further enhanced by an orchestration tool, for example Kubernetes, which is a tool for simplification of automated deployment of a big number of individual containers. Both these tools were included in the draft of the architecture for the Project[4].

First, one instance of the part handling the data collection is described; it is depicted in Figure 4, individual numbered parts are described separately.

1. The entry point for the whole system is a “Request handler” microservice. It reads the request, determines, whether it should be handled by this instance, if no, it is dispatched according to CRT (3). The local copy of CRT is updated every day, according to the “Regionality layer” specification in Section 3.1.2 from (4).

   If the request is local, it is handled, and the data is written into the data collection database (2).

2. Second part is the database with collected data. Based on the amount of request per time unit or the amount of data needed to store, this can be implemented either as one database server or a more complicated database cluster containing multiple servers.

3. Number (3) refers to the other regional instances, as per “Regionality layer”.

4. Number (4) refers to the server in charge of the Central routing table, as per “Regionality layer”.

5. Number (5) refers to the databases holding data in the recommendation backend service.

The second part—the recommendation backend service—is more complicated. The diagram of the architecture is shown in the Figure 5, individual numbered parts are described separately

---

[4] Docker for containerisation was actively used since the beginning with docker-compose for one-machine orchestration. At the time of writing, Kubernetes was not yet implemented, as the need for scalability across multiple machines was not as urgent.

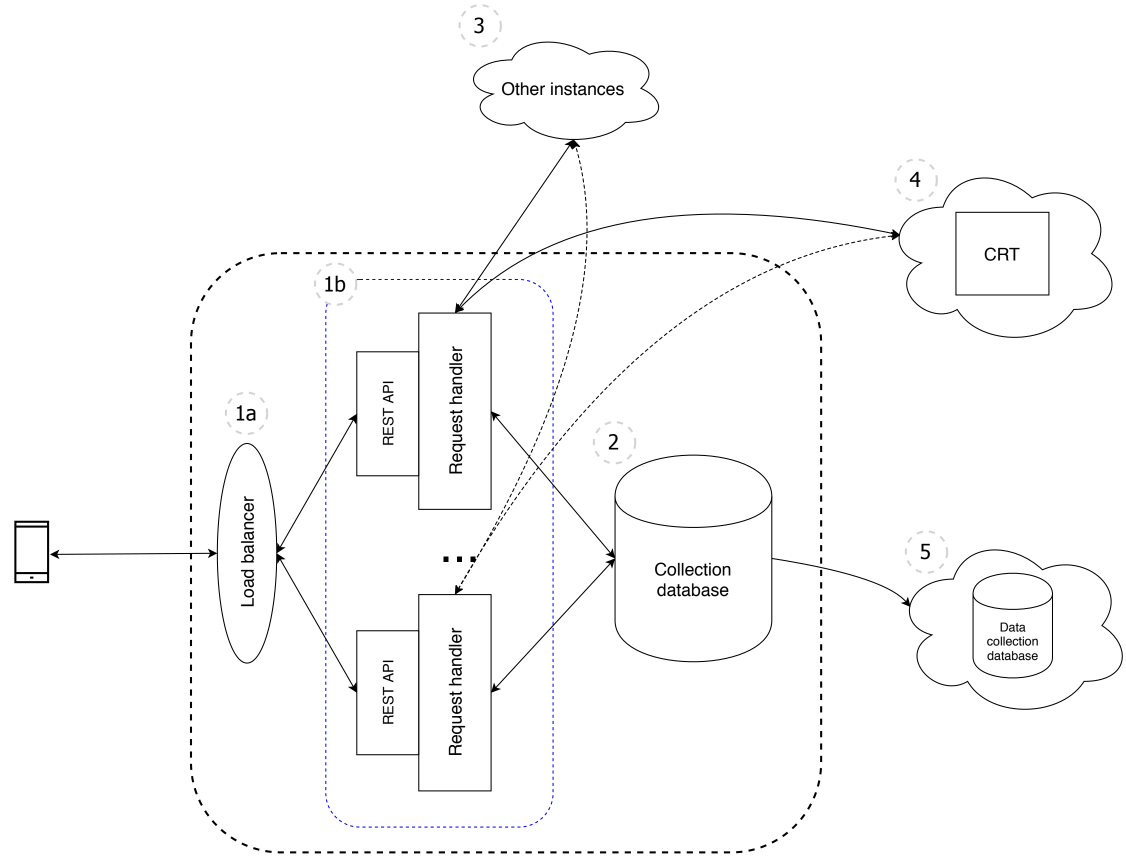


Figure 4: Diagram of the data collection backend. Shown are microservices implementing logic (1) and storing data (2), further other parties are depicted (3,4,5).

1. The first microservice is the entry point to the whole system. In the same fashion as with the entry point of the data collection, it first determines, whether the request should not be handled by another instance, in which case it gets dispatched according to the local copy of the CRT, which is updated every day from the server in charge of the global copy (6).

   If the request is local, then based on its type—the request can either be asking for the list of places in a certain area, or it can be asking for a recommendation for a specific place—it is passed to the respective microservice.

2. The microservice (2) is responsible for the computation of recommendations and managing the data needed for it. It consists of the logic and database part.

   Because of the spatial locality principle (as per Section 3.1.2), when the amount of data gets too big so that it negatively affects the performance of database queries, the data can be further split into spatially smaller sectors. The scaling to higher amounts of data and users can then be smarter, than standard horizontal scaling, where the use of a spatially-aware load balancer (2a) can ensure, that the requests

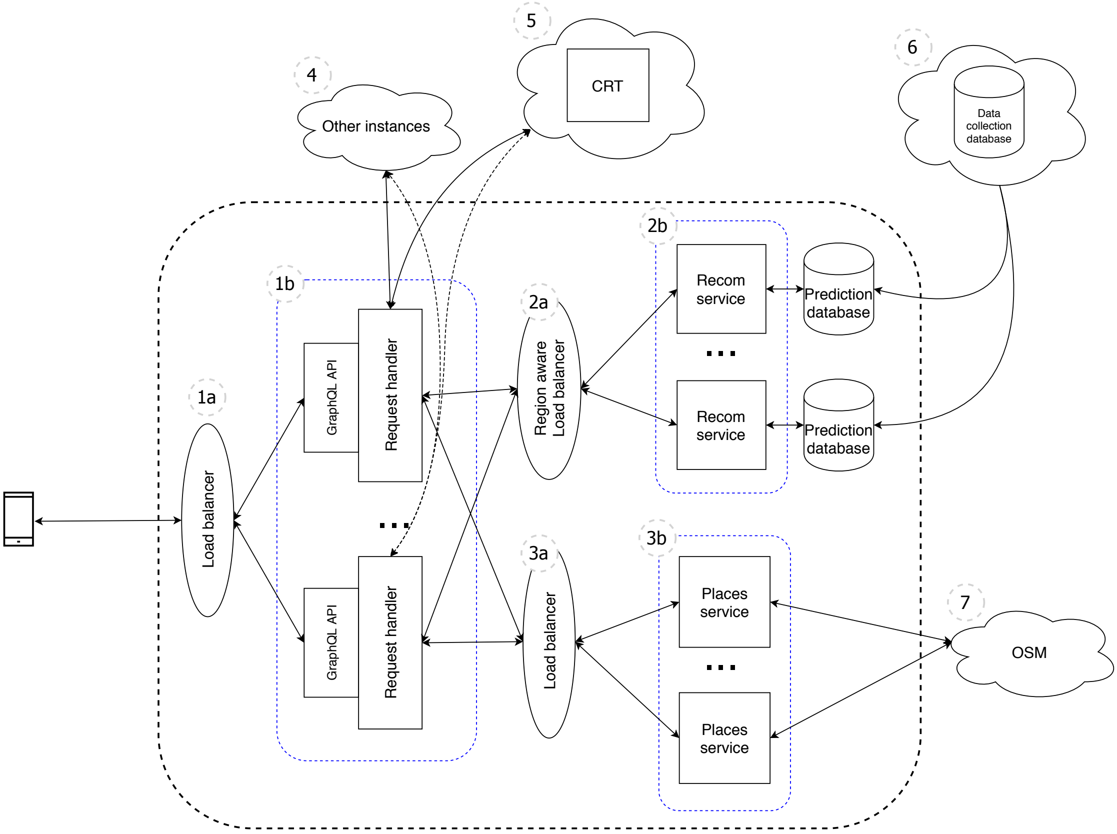


Figure 5: Diagram of the Recommendation backend. Shown are microservices acting as entry points(1), ones responsible for recommendation logic and data storage (2) and ones responsible for serving lists of places (3). Other parties are shown under (4,5,6,7).

are distributed according to the sector they belong to as well as equally between multiple instances of the microservice that holds the data for this sector.

3. The third microservice implements the logic for requests related to places. It processes the request and forwards it to an OpenStreetMap[5] (OSM) provider (7)[6].

   If the performance of this outsourcing turns out to be insufficient, it is possible to deploy an OSM hosting specifically for the Project due to the open nature of OSM.

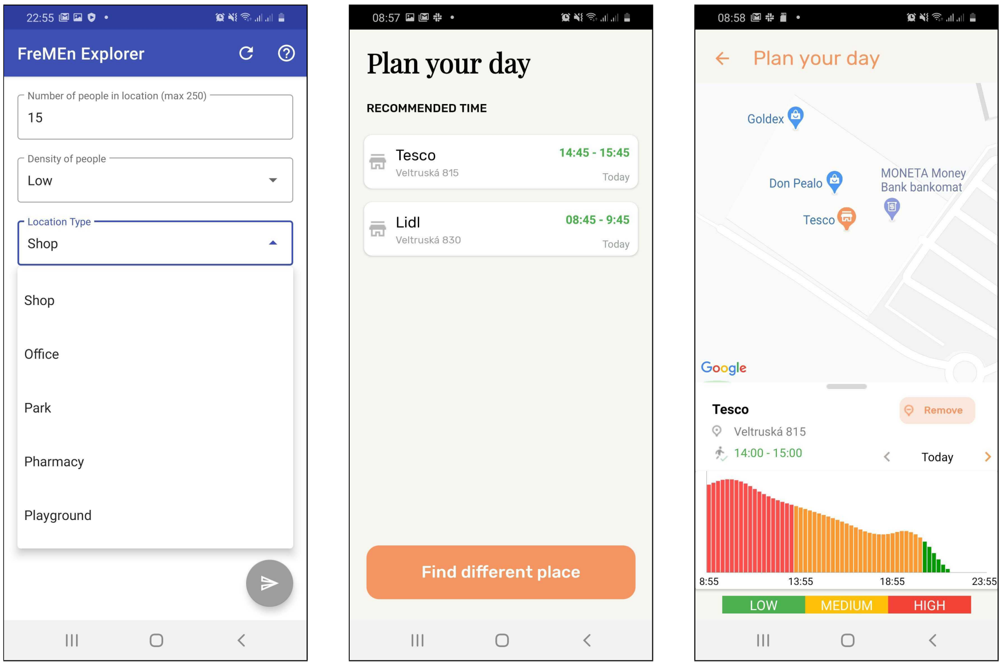
4. Number (4) refers to the other regional instances, as per “Regionality layer”.

5. Number (5) refers to the server in charge of the Central routing table, as per “Regionality layer”.

---

[5] The name is a trademark of the OpenStreetMap Foundation and is used with their permission. Nor the Project nor author is endorsed by or affiliated with the OpenStreetMap Foundation.

[6] At the time of writing this has not yet been implemented, and as a temporary solution, Google Places API (part of the Google Maps^{TM} API) was used.



Number of people in location (max 250): 15

Density of people: Low

Location Type: Shop

RECOMMENDED TIME

Tesco, Veltruská 815, 14:45 - 15:45, Today

Lidl, Veltruská 830, 08:45 - 9:45, Today

Tesco, Veltruská 815, 14:00 - 15:00, Today

(a) FreMEn Explorer

(b) FreMEn Advisor - list of favourites

(c) FreMEn Advisor - detail of favourite with graph for tommorow

Figure 6: Screenshots from two applications developed as a part of the Project. One on the left is from FreMEn Explorer and two on the right are from FreMEn Advisor.

6. Number (6) refers to the databases holding data in the data collection backend service.

7. Number (7) refers to an OpenStreetMap provider.

3.1.3 FreMEn Explorer

The mobile app for data gathering was designed with the ambition to develop a very simple one-purpose tool as soon as possible, so the collection of data could begin early and by the time the development of the frontend for presenting recommendations some initial database could already be in place.

It contains an onboarding activity, a dialogue for collecting users consent to the further academic use of collected data and a form for submitting observations. The form activity is shown in the Figure 6a. When the user presses the submit button, then the content of the form is submitted to the backend service for data collection together with a GPS location and current time, including timezone information.

3.1.4 FreMEn Advisor / Nebojsa

Mobile application for showing recommendations is more complicated, that the one for data collection. The home activity, as shown in Figure 6b, is a list of favourite places—their name, address and an icon showing the type of place. The user can add new places to this list and remove them from it. For every place, a recommended time to visit it is displayed next to its name.

When clicking on the box of the place, the user can view its detail, as shown in Figure 6c. They can see where it is on the map, they can remove the place from the list, and they can also see the predicted graph based upon which the recommendation was given.

Adding of place happens through a search on a map in a which the user is presented with a list of places in proximity of their searched query (they only are allowed to choose from a valid set of locations presented to him as autosuggestions while typing).

3.2 Problem formulation

Before discusing the logic behind the recommending system, it is necessary to formulate the problem, which will be solved. The natural language description is, that given the collected dataset, the system should respond to requests, where the user can ask when to visit some target place and wants to see the prediction of business for the future.

First, the location L is defined as a triple (x, y, c), where x, y are coordinates, $c \in C$ is a type of the place—its category—and $C = \{\text{“park”}, \text{“shop”}, \text{“pharmacy”}\}$.

Next, the data point or observation O is defined as a quadruple (L, t_o, b, n), where L is the observed place, t_o the time of the observation, $b \in \{1, \ldots 5\}$ is the level of business and n the number of observed people.

The main problem is then a triple $(t, P, \mathcal{D})$, where t is the time of the request, P is the target place, and $\mathcal{D}$ is the collected database—a set of observations $\mathcal{O}$ from which the recommendation should be computed, with the possibility to perform SQL-like queries on top of it. To which the result is a tuple $(t_r, \mathbf{p})$, where t_r is the recommended time of visit and $\mathbf{p}$ is a function of time, representing the predictions.

3.3 Temporal modelling

This Section describes the temporal modelling algorithm used for prediction of business by time in a particular location, that is used for creating a recommendation. It is expected that this might undergo drastic changes as more data come into the system, and a scientific method and testing can be employed for its design. In the time of designing this algorithm, only small amounts of data were present in the system not of great quality.

3.3.1 Choice of temporal model

The first challenge comes with the decision of the type of a temporal model. A natural choice for periodic temporal process modelling using only very sparse data is the FreMEn model, as described in Section 2.1.3. The problem with this approach is due to the underlying spectral analysis, that even though the amount of data necessary is small, it still needs 20-50 observations according to their quality, that are spread over about two times the maximal periodicity it will be able to capture. If there are less observations, then even FreMEn model would not be able to deliver reasonable results.

There are some assumptions, which can be taken, that help to choose a temporal model useful until enough data are present in the system to employ more sophisticated technology. First one is that we can expect the behaviour of people to follow daily periodicity, day-night changing is one of the key natural processes affecting all the living organisms. In the habits of people, also the weekly periodicity plays an important role, but this can be added once the data is dense enough with the use of FreMEn. Second is that the generation of the data also follows a temporal pattern. If the users are expected to enter observations while performing standard activities, it can also be expected, that they will do it at the same time they are used to perform these activities. Therefore the observations will be clustered on the time axis at times, the business of the place is higher.

With these assumptions, the histogram model with one-day periodicity was chosen as this is with further adjustments a good starting strategy.

3.3.2 Modifications made to match the use-case

The histogram model was modified and its parameters were chosen to better match with the specific data, their sparsity and also low amount. These consist of

1. two-hour bins, so that enough data falls into a bin;
2. applying an idea inspired by the Parzen window method, implementing some temporal continuity in processing the data, which does not include the data point only in the respective bin of the histogram, but also to bins next to it with half of the weight;
3. reverting the effect of choosing two-hour sized bins resulting in a very coarse histogram is done by second-order spline interpolation over the whole histogram interval.

For the generalisation of the histogram method to categorical distribution, another trick was used, leveraging the natural ordering of predicted values. So instead of actually modelling the parameters of a categorical distribution, the incremental mean method was used for estimation of a value for one bin.

3.4 Recommendation algorithm

The main algorithm that presents the modelling results to the user by giving him advice as to when visit a certain place is described in this Section. It solves the problem, as defined in Section 3.2.

Algorithm 1: Computation of the time recommended for visit

```
Input:  $t$ – time when recommendation is requested,
        ($lat$, $lon$, $cat$) – a place (coordinates and category),
        $\mathcal{D}$ – collected dataset—a database with tuples (position, place category,
        time, business value, number of people)
Output: $\mathbf{p}$ – graph of predicted business starting at $t$ sampled by 15 minutes,
        $t_r$ – recommended time

if $cat$ = "$park$" then
    $select\_distance \leftarrow 500m$
else
    $select\_distance \leftarrow 50m$

$s \leftarrow$ SELECT time, business value FROM $\mathcal{D}$ WHERE
        distance(position, ($lat$, $lon$)) $<$ $select\_distance$ AND place category $==$ $cat$

$f \leftarrow$ create_temporal_model($s$)
// create a prediction in form of an array p and sample temporal
model in 15 minute intervals
$\mathbf{p}$, where $p[i] \leftarrow f(t + 15i), i = 0, \ldots, 48 * 4$
// select a subarray p' of p, so that it represents time in interval
max(8 am, $t$)-7 pm
$\mathbf{p'} = p[j : k]$, where $j, k$ are indexes into $\mathbf{p}$, so that $j$ represents the lowest time
        $t_j$ where $t_j > t \wedge t_j > 8$ am, and $k$ represents the largest time $t_k$ where
        $t_k < 7$ pm.
// create an array h representing hourly averages of p'
$\mathbf{h}$, where $\mathbf{h}[i] \leftarrow \frac{1}{4} \sum_{j=0}^{3} p'[i + j]$.
// create a set $\mathcal{I}$ representing indexes (hours) with below-or-average
business
$\mathcal{I} \leftarrow \{i | i = 0, \ldots, |\mathbf{h}| \wedge \mathbf{h}[i] \leq \frac{1}{|\mathbf{h}|} \sum_{j=0}^{|\mathbf{h}|} h[j]\}$
// fix the ordering
$\mathcal{I} \leftarrow$ array($\mathcal{I}$)
// create an array $\mathcal{I}'$ of values of particular times
$\mathcal{I}'$, where $\mathcal{I}'[i] = 1/h[i]^3, i \in \mathcal{I}$
// the final recommendation - select time according to probability
$r \leftarrow$ select one element from $\mathcal{I}'$, where selection of $\mathcal{I}[i]$ happens with probability
$\mathcal{I}'[i]$
$t_r \leftarrow$ convert index $r$ back to time representation
```

Inputs to this algorithm are

1. local time,
2. the target place, represented by its GPS location and its type.
3. collected dataset

Detailed description If no cached results of temporal modelling are available, then the algorithm then selects all data points that are of the place-type and in the vicinity of the target place. The vicinity is defined to be circle area in 50 meters radius for small and localised places, and 500 meters radius for parks as these are usually large and of various shapes in comparison with shops and alike. From these data points, a temporal model f is created, which is the basis for creating the recommendation.

All the following computation happens with respect to the input time t, including timezones. The created temporal model is used for predicting the business of the target place in next 48 hours (plus 6 hours if the result should be cached), the sampling being done in 15 minutes long intervals, creating an array $\mathbf{p}$, where

$$\mathbf{p}[i] = f(t + 15i),\ i = 0, \ldots, 48 * 4, \tag{5}$$

assuming the time t to be in some linear representation and the unit of time of the temporal model to be minutes. The recommending algorithm only uses the prediction for the current day, but the user is supposed to get information even for the next day. Only prediction for today and in specified interval 8:00 - 19:00 is selected for further processing so that the recommendations are reasonable and not in the middle of the night, this will be denoted $\mathbf{p}'$.

Hourly average is computed using a uniform filter, i.e. create an array $\mathbf{h}$ defined as

$$\mathbf{h}[i] = \frac{1}{4} \sum_{j=0}^{3} p'[i + j]. \tag{6}$$

A set $\mathbf{h}'$ of considered times (indexes) is created, by taking all hours in $\mathbf{h}$, where expected business is not above average of these, so

$$i \in \mathbf{h}' \text{ if } \mathbf{h}[i] <= \frac{1}{|\mathbf{h}|} \sum_{j=0}^{|\mathbf{h}|} h[j]. \tag{7}$$

For every considered time i its value v_i is computed by taking an inverse of the third power of expected business

$$v_i = \frac{1}{h[i]^3}. \tag{8}$$

The choice of the third power is arbitrary, as any would prefer times with lower expected business, but it can be viewed as a parameter effecting how much the algorithm will penalise times of higher business. Initially, a second power was used, which during user-testing turned out to be weak penalisation, so the number was changed accordingly.

The last stage is a stochastic selection, where one of the considered times is selected with the probability proportional to its value.

Output The selected time with a graph of predictions for today and tomorrow is returned to the requesting user.

4 Experiments

The main task of the experimental evaluation is to test what is the effect of the transition, transfer or inference of the temporal model on the benefits users get from using the recommending service, and the benefits themselves. The performance is measured through a comparison of the theoretical impact of different strategies, and different sources of data on the risk people undertake while performing necessary visits to public places.

4.1 Sources of data

For experiments, three sources of data were used. Two were solely for the training and one of the three for training and evaluation. The used datasets represent the temporal evolution of business in several locations.

Graph The first data source was using the backend services of the Project. Therefore, the data was created using the crowdsourcing of observations and methods of chronorobotic temporal modelling. For specific locations, a request was made against the recommendation service, and only the predictions were kept.

Google Popular times As part of their maps service Google Maps$^{\mathrm{TM}}$, Google LLC for several years already publishes a graph for every day of the week called Popular times, which shows average business by an hour in a graph, for specific locations. It is a long-term aggregate from the location history of users of Google services and products, like the Android operating system.

Google "Live business reporting" During the period of formation of the Project and a world-wide pandemic, Googles Popular times were in the Czech republic extended with live data. The live data updates in about 15 minutes intervals and apart from historical aggregations show the actual business in the target place. This data source was used for evaluation, as only it gives full-time coverage for proper evaluation. The data for experiments were gathered using web scraping as a part of the efforts of the Project.

The list of locations was chosen, so that enough data was present in the recommendation database and at the same time, Google gives data about its business as a part of its Popular times with the live data included. Two location categories were considered—shops and parks.

It is the following list of shops:

1. Billa, Tilleho nám. 793/1, 152 00 Praha 5-Barrandov (50.0346471N, 14.3781577E)
2. Globus hypermarket, Kostelecká 822/75, 196 00 Praha-Čakovice (50.1504213N, 14.5054118E)

3. Kaufland, Blanenská, 664 34 Kuřim (49.308819N, 16.534834E)
4. Tesco, Veltruská 815, 278 01 Kralupy nad Vltavou (50.2530286N, 14.3178333E)

And the following list of parks:

1. Riegrovy sady, Vinohrady, 120 00 Praha 2 (50.0802843N, 14.4415605E)
2. Malešický park, Akademická 688/1, 108 00 Praha 10 (50.0850143N, 14.499155E)
3. Vítkov, Žižkov, 130 00 Praha 3 (50.0888836N, 14.45436E)

4.2 Design of experiment

The experiment is built in a manner of comparison in an application. There are several possible strategies a person can adopt when visiting a public place—whether are they based on some previous knowledge or not. In the following text, these will be referred to as different kind of agents. The agents with previous information can be based upon any of the three data sources discussed in the previous section. However, more importantly, the ones based on data can be further divided by whether they use information tied exactly to the specific place, or not.

4.2.1 Strategies

By a strategy, a probabilistic distribution is meant, by which some agent chooses the time when it will visit the target place. Some strategies use data for decision-making; these are referred to as data-aware; others are data-agnostic. In Algorithm 1 this distribution is referred to as $\mathbf{p}'$, which is a sampling of temporal model restricted to a specific interval.

The resulting distribution is always tied to some defined time interval because a unified view must be adopted in terms of opening hours, which can vary between different shops and while for parks is hardly applicable, for the majority of the population, it does not make sense to visit parks in the middle of the night.

Uniform The uniform agent behaves completely randomly with uniform distribution. It is a data-agnostic method, acting as a baseline for the other agents.

Standard The standard agent adopts a behaviour standard to the population, that visits the given place. I.e., this agent will follow the temporal model of business of the given location directly, simulating an agent not aware of when it is preferable to perform the visit.

Recom The recom agent is an agent with some previous knowledge of the location, that tries to minimise the risk of an encounter with high business in the target place using the recommendation algorithm (Algorithm 1).

RecomT The recomT agent is an agent, that implements the same strategy as a recom agent, but does not use the knowledge specific to the target location. Instead, it constructs the prior knowledge by using the type of the target location. It selects all different locations of the same type and makes its strategy for every single one of them and then follows the average of those. This represents the type of generalisation that people do on a daily basis and lies in connecting multiple places according to their semantical similarity and extracting the common patterns.

4.2.2 Agents

By an agent, a certain behaviour is represented. The kind of agent specifies its adopted strategy.

The data-aware agents take as input the distribution of business in time and adjust their behaviour by it and the defined interval of possible visits. The source of the data can be any of the three datasets listed in the previous section, so the label for a specific agent is constructed as *strategy-datasource*, with data source being “graph” for the Project, “poptimes” for Googles Popular times or “live”. In the case of “live” data source the agent gets information about the day before the evaluated one.

As different sources of data come with different level of granularity, linear interpolation is used inside the interval given by the dataset and extrapolation by the closest edge value is used outside, to obtain a temporal evolution covering the whole day.

4.2.3 Experiment details

In the experiment, for every day d and location l in the evaluation dataset, every considered agent creates a distribution over presented times and according to this distribution a weighted average of business values at the given times in the evaluation dataset is computed.

For every pair (d, l) a list of times $\mathbf{t}'_{(d,l)}$ and business values $\mathbf{b}'_{(d,l)}$ is taken from the evaluation dataset. The list of times is cropped to a chosen interval 8 am – 8 pm and the list of business values is restricted accordingly. This is supposed to match with all opening hours and also represent subjective time interval where most of the people prefer to leave their home.

Then, because the quality of data sampling is not great, the $\mathbf{t}'_{(d,l)}$ and $\mathbf{b}'_{(d,l)}$ are extended to the whole interval 8 am – 8 pm using linear interpolation and extrapolation with edge value, this creates a function of time g. This creates lists of times with step of one hour $\mathbf{t} =$ 8 am, 9 am,. . . , 8 pm and g is sampled at times in $\mathbf{t}$, creating $\mathbf{b}_{(d,l)} = g(\mathbf{t})$.

Following list of agents $\mathcal{A}$ is created: a uniform agent, a standard agent, a recom agent and a recomT agent, for every type of dataset, i.e. “graph”, “poptimes” and “live”. The overall principle is depicted in Figure 7.

Every agent $a \in \mathcal{A}$—where a can interchangeably be referring to the agent itself, or its name—is given list $\mathbf{t}$ and according to its strategy it returns a list of probabilities

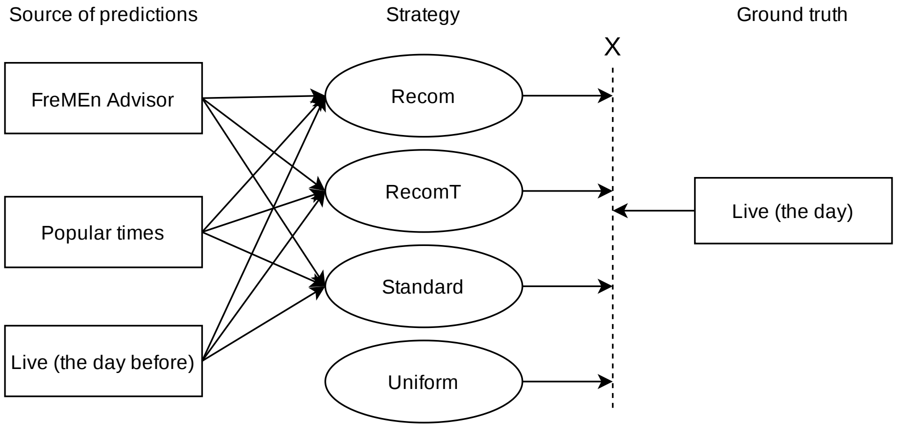


Figure 7: Sheme of the experiment for one location. Nine data-aware agents and one agent implementing uniform stategy are created. Their plan for the next day is compared to ground truth data from “Live” datasource.

$\mathbf{p}^a_{(d,l)}$ for individual times. Since it is obligated to visit the place at least once, this is then normalised, so a distribution $\mathbf{p}^a_{(d,l)}$ over $\mathbf{t}$ is obtained.

Instead of a probabilistic computation of the expected value, this distribution is used as weights for computation of weighted average $\hat{b}^a_{(d,l)}$ of the actual business in the ground truth. As $\mathbf{p}$ is a distribution, the computation can be simplified into

$$\hat{b}^a_{(d,l)} = \langle \mathbf{p}^a_{(d,l)} | \mathbf{b}^a_{(d,l)} \rangle, \tag{9}$$

using the fact, that distribution sums up to one. This average represents the amount of risk the agent would experience when visiting target location l in day d.

4.2.4 Evaluation method

The results from the experiment, as described in the previous section are in form of a set of values $\hat{b}^a_{(d,l)}$ for every $a \in \mathcal{A}$ and (d, l) in evaluation dataset. For a meaningful visualisation and analysis of the results, these are organised into series $\mathbf{r}$. For an agent a, first a fixed ordering I of pairs (d, l) is adopted, then

$$\mathbf{r}^a = (\hat{b}^a_{(d,l)_i})_{i \in I}. \tag{10}$$

This is a series of results achieved by the agent a for all locations and days.

To responsibly compare the performance of two methods, it is necessary to understand, that not the distributions of resulting values are to be compared, but the values themselves. As these are tight to individual situations, pairs arise in the two distributions from the same situation. These can be subtracted, i.e. in this case a series $\mathbf{r} = \mathbf{r}^a - \mathbf{r}^b$ is created for two agents $a, b \in \mathcal{A}$. Then, it is possible to use statistical methods to test whether distribution r has zero mean value if the assumption of r having normal distribution is accepted. This method is called pair t-test.

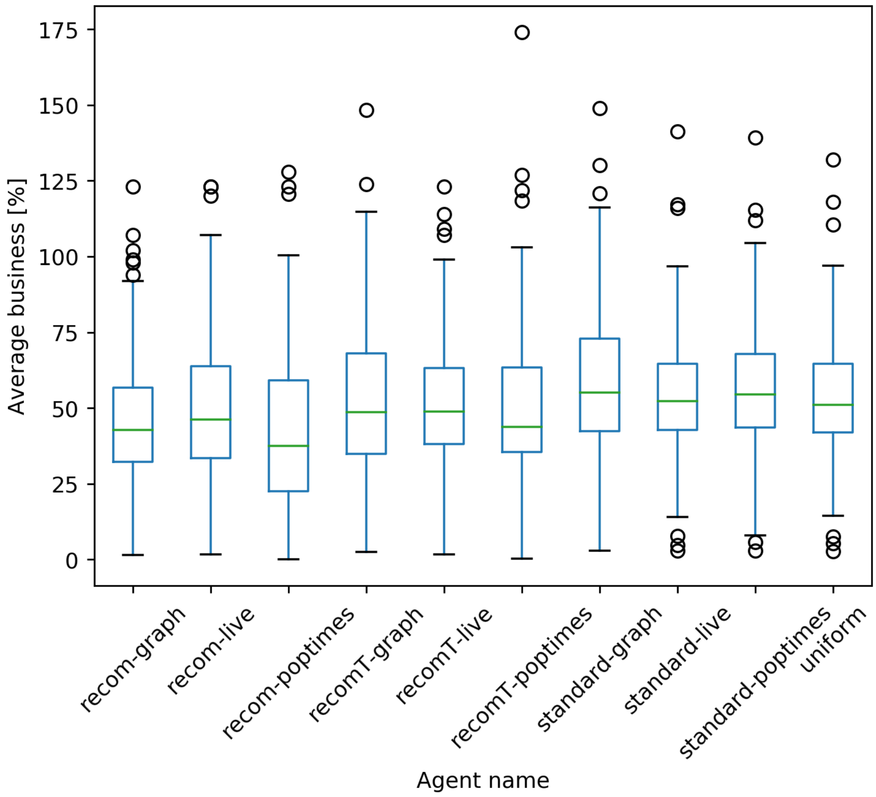


Figure 8: Boxplot of results. For every considered agent achieved daily averages of business in target location are plotted.

4.3 Results

Results of the experiment are mainly to analyse the benefit of a strategy a person can adopt and compare the result between individual strategies and data this strategy can be based upon.

The resulting series are visualised in Figure 8 in the form of a box plot, to show the distribution of resulting values. This presents the general idea of how every strategy performed and how this depends on the source of data.

To compare individual agents more rigorously, a pair t-test between all pairs of agents has been conducted with $\alpha = 0.05$, and the results are shown in the diagram in Figure 9. This diagram has an arrow from agent a to b if and only if a performs statistically significantly better than b. Since the amount of agents is large and the relation depicted with arrows is transitive, only the longest path between two agents has been plotted.

Contribution to social distancing The first observation that can be made from the Figures 8 and 9 is that recom strategy indeed results in a lower amount of met business.

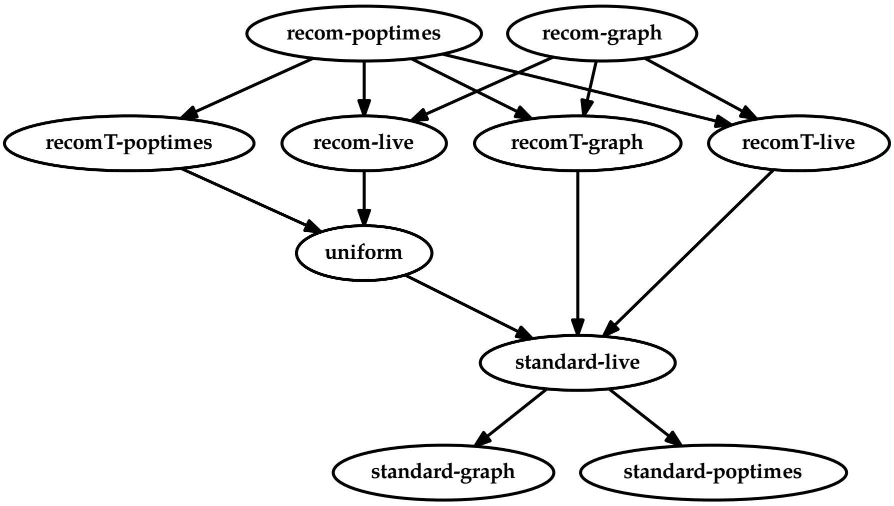


Figure 9: Plot of statistically significant dominance inbetween different agents. An arrow goes from agent A to agent B, if a pair t-test at $\alpha = 0.05$ rejects the hypothesis, that B performs as good or better that A. In other words, an arrow from A to B indicates that people behaving by A will be exposed to statistically significant lower exposure to the virus.

This proves that tools aiding social distancing are possible and that the approach taken by the Project brings benefit to the user.

On average, the recom strategy achieved result $\sim 20\%$ lower, than the standard one, when comparing results from individual days and locations, in the metric plotted in Figure 8. This quantification unfortunatelly does not represent the actual risk of the infection, because the exact models of transmision of the virus are not known. Employed metric only allows for comparison of presented methods, as the connection between the business and probability of infection is expected to be monotonic and dramatic, according to literature[31].

Sources of predictions The results show, that when comparing sources of data between different strategies, then “graph” and “poptimes” capture the best temporal evolution of target location and their performance is generally comparable. When used with the recom strategy, the results are lower; when used with standard strategy, these are higher. That is due to the source of prediction being more precise when standard is meeting more people and recom less.

The fact, that “graph” and “poptimes” are comparable also shows, that even with an extremely low amount of data, it is possible to match with a big-data based model from a major company.

When used as a basis for the temporal transfer—recomT strategy—the “graph” data source is not performing as well. As the variance is highest among other recomT agents, it is probably due to the “graph” generally being based on lowest amounts of data, which seems to need a more sophisticated method to combine individual models, than following a mean distribution.

Temporal transfer When using data source based on more data, the recomT strategy performs quite well, even comparably to the recom strategy using data specifically from the target location. Although recom strategy still dominates recomT. This shows that the temporal transfer is a possible and reasonable way of dealing with limited amounts of data often present in robotic applications.

5 Conclusion

This thesis had two interconnected goals. One was to present a successful technology transfer in the FreMEn contra COVID project, where results of robotic research are used for aiding people to adopt social distancing measures during a world-wide pandemic and preliminary testing of its potential impact on users risk of exposure. The second was to investigate temporal transfer, a possibility of bootstrapping the temporal modelling methods using previous experience from a different environment, to fasten their deployment and deal with low amounts of data, common to robotic applications.

Overview of the Project was presented in the context of the situation it was created. The technical solution providing recommendation on time of visiting public places is described in detail, including motivation for individual decisions. Requirements for the solution are presented and their role in the resulting architecture of the whole system. The whole system is characterized from the architectural level, and key algorithms are discussed.

The experiment was designed and conducted so that the impact of the recommendation algorithm is measured and also the transfer of temporal models is tested as a way of dealing with a low amount of data. The behaviour of following the output of the recommendation algorithm was compared with the standard behaviour of people and a completely random one. At the same time, several sources of predictions were tested—one relying on crowdsourced gathering, one on public models from Google based on historical big-data and one also from Google with real-time reporting of business in a given location.

The result show, that following the recommendations lowers the amount of business in target location. It has been shown, that according to available datasets the sources of predictions, based on historical data outperform short history of real-time reporting.

The temporal transfer of models describing a public place has been shown to give worse results than targeted models, as expected. In terms of the desired boosting, an important result is that they still perform better, than a standard behaviour, therefore they can effectively substitute in cases, where no data is available.

5.1 Future work

There are many questions to pursue in future research. These can be divided into improvements to the presented methods, improvements to the experiments presented in this thesis and new research questions.

Improvements to the presented methods would be to propose, investigate and evaluate more complicated methods for transferring temporal models. There are many kinds of context to be considered for allowing the transfer. For example, the ratio between the business and an absolute number of present people could bring insight into the size of the target location, which intuitively plays a role in how people behave there. Even the way knowledge of context is exploited should be further investigated.

The presented experiments are based on a relatively small sample of locations, expanding this list, and the length of the evaluation dataset would definitely make the result stronger. An evaluation dataset based on physical observation of the number of people would also bring more confidence. Also, different methods for temporal modelling should be evaluated, with regard to previous research in the field of chronorobotics it is expected, that FreMEn method should perform better, that modified histogram-based algorithm deployed by the Project.

As the next question to experimentally answer, the author proposes how the amount of penetration of users of FreMEn Advisor in the population affects the business at given locations. As more visitors of location would be using it, the business throughout the day should approach a uniform distribution, which would be ideal, as it is not possible to spread the demand of visitors in time more. The effect of penetration could be tested by comparing a predicted distribution not only to the actual data but to a convex combination of the actual data and the predicted distribution. The ration would then represent the amount of penetration.

References

[1] Tomáš Krajník, Tomáš Vintr, Sergi Molina, Jaime Pulido Fentanes, Grzegorz Cielniak, Oscar Martinez Mozos, George Broughton, and Tom Duckett. Warped Hypertime Representations for Long-Term Autonomy of Mobile Robots. *IEEE Robotics and Automation Letters*, 4(4):3310–3317, October 2019.

[2] Tomáš Krajník, Tomas Vintr, George Broughton, Filip Majer, Tomáš Rouček, Jiří Ulrich, Jan Blaha, Veronika Pěčonková, and Martin Rektoris. CHRONOROBOTICS: Representing the structure of time for service robots. November 2019.

[3] Tomáš Krajník, Filip Majer, Lucie Halodová, and Tomáš Vintr. Navigation without localisation: reliable teach and repeat based on the convergence theorem. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1657–1664, October 2018. ISSN: 2153-0866.

[4] Tomasz Piotr Kucner, Achim J Lilienthal, Martin Magnusson, Luigi Palmieri, and Chittaranjan Srinivas Swaminathan. *Probabilistic mapping of spatial motion patterns for mobile robots.* Springer, 2020.

[5] Sebastian Thrun et al. Robotic mapping: A survey. *Exploring artificial intelligence in the new millennium*, 1(1-35):1, 2002.

[6] Judy L Klein and Dave Klein. *Statistical visions in time: a history of time series analysis, 1662-1938.* Cambridge University Press, 1997.

[7] Tomas Krajnik, Jaime Pulido Fentanes, Grzegorz Cielniak, Christian Dondrup, and Tom Duckett. Spectral analysis for long-term robotic mapping. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3706–3711, May 2014. ISSN: 1050-4729.

[8] Tomáš Krajník, Jaime P. Fentanes, João M. Santos, and Tom Duckett. FreMEn: Frequency Map Enhancement for Long-Term Mobile Robot Autonomy in Changing Environments. *IEEE Transactions on Robotics*, 33(4):964–977, August 2017.

[9] Tomáš Krajník, Jaime Pulido Fentanes, Marc Hanheide, and Tom Duckett. Persistent localization and life-long mapping in changing environments using the Frequency Map Enhancement. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4558–4563, October 2016. ISSN: 2153-0866.

[10] Tomáš Vintr, Zhi Yan, Tom Duckett, and Tomáš Krajník. Spatio-temporal representation for long-term anticipation of human presence in service robotics. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 2620–2626, May 2019. ISSN: 2577-087X.

[11] Tomas Vintr, Sergi Molina Mellado, Grzegorz Cielniak, Tom Duckett, and Tomas Krajnik. Spatiotemporal models for motion planning in human populated environments. Žilina, Slovakia, September 2017. Czech Technical University in Prague, Faculty of Electrical Engineering.

[12] Tomáš Vintr, Kerem Eyisoy, Vanda Vintrová, Zhi Yan, Yassine Ruichek, and Tomáš Krajník. Spatiotemporal Models of Human Activity for Robotic Patrolling. In Jan Mazal, editor, *Modelling and Simulation for Autonomous Systems*, pages 54–64, Cham, 2019. Springer International Publishing.

[13] Tomáš Vintr, Sergi Molina, Ransalu Senanayake, George Broughton, Zhi Yan, Jiří Ulrich, Tomasz Piotr Kucner, Chittaranjan Srinivas Swaminathan, Filip Majer, Mária Stachová, Achim J. Lilienthal, and Tomáš Krajník. Time-varying Pedestrian Flow Models for Service Robots. In *2019 European Conference on Mobile Robots (ECMR)*, pages 1–7, September 2019. ISSN: null.

[14] WHO. Coronavirus disease 2019 (COVID-19), Situation report 71. Technical report, World Health Organization, Geneva, 31 March 2020.

[15] Adam J Kucharski, Timothy W Russell, Charlie Diamond, Yang Liu, John Edmunds, Sebastian Funk, Rosalind M Eggo, Fiona Sun, Mark Jit, James D Munday, Nicholas Davies, Amy Gimma, Kevin van Zandvoort, Hamish Gibbs, Joel Hellewell, Christopher I Jarvis, Sam Clifford, Billy J Quilty, Nikos I Bosse, Sam Abbott, Petra Klepac, and Stefan Flasche. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases*, 20(5):553–558, May 2020.

[16] WHO. Coronavirus disease (COVID-19) advice for the public. `https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public`. Accessed: 2020-05-19.

[17] Seznam.cz. Zastavme společně šíření koronaviru. `https://www.seznam.cz/zastav-covid/`. Accessed: 2020-05-18.

[18] Ministry of Health of the Czech Republic. eRouška. `https://erouska.cz/`. Accessed: 2020-05-18.

[19] WHO. Coronavirus disease 2019 (COVID-19), Situation report 35. Technical report, World Health Organization, Geneva, 24 February 2020.

[20] Rachel A. Smith and David Hughes. Infectious Disease Stigmas: Maladaptive in Modern Society. *Communication studies*, 65(2):132–138, April 2014.

[21] Ministry of Health of the Czech Republic. Mimořádné opatření - nová nákupní doba pro seniory. `http://www.mzcr.cz/Legislativa/dokumenty/mimoradne-opatreni-nova-nakupni-doba-pro-seniory_18785_4135_11.html`. Accessed: 2020-05-20.

[22] NHS. People at higher risk from coronavirus. `https://www.nhs.uk/conditions/coronavirus-covid-19/people-at-higher-risk-from-coronavirus/`. Accessed: 2020-05-19.

[23] Tomáš Krajník, Miroslav Kulich, Lenka Mudrová, Rares Ambrus, and Tom Duckett. Where's waldo at time t ? using spatio-temporal models for mobile robot search. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2140–2146, May 2015. ISSN: 1050-4729.

[24] João Machado Santos, Tomáš Krajník, and Tom Duckett. Spatio-temporal exploration strategies for long-term autonomy of mobile robots. *Robotics and Autonomous Systems*, 88:116–126, February 2017.

[25] Claudio Coppola, Tomáš Krajník, Tom Duckett, and Nicola Bellotto. Learning temporal context for activity recognition. In *Proceedings of the Twenty-second European Conference on Artificial Intelligence*, ECAI'16, pages 107–115, The Hague, The Netherlands, August 2016. IOS Press.

[26] Weiming Zhi, Tin Lai, Lionel Ott, Gilad Francis, and Fabio Ramos. OCTNet: Trajectory Generation in New Environments from Past Experiences. *arXiv:1909.11337 [cs]*, September 2019. arXiv: 1909.11337.

[27] Django Software Foundation. Django [software]. `https://djangoproject.com`, 2019.

[28] The PostgreSQL Global Development Group. Postresql [software]. `https://www.postgresql.org`.

[29] PostGIS Project Steering Committee. Postgis [software]. `https://postgis.net`.

[30] Docker, Inc. Docker [software]. `https://www.docker.com`.

[31] Chacha M Issarow, Nicola Mulder, and Robin Wood. Modelling the risk of airborne infectious disease using exhaled air. *Journal of theoretical biology*, 372:100–106, 2015.