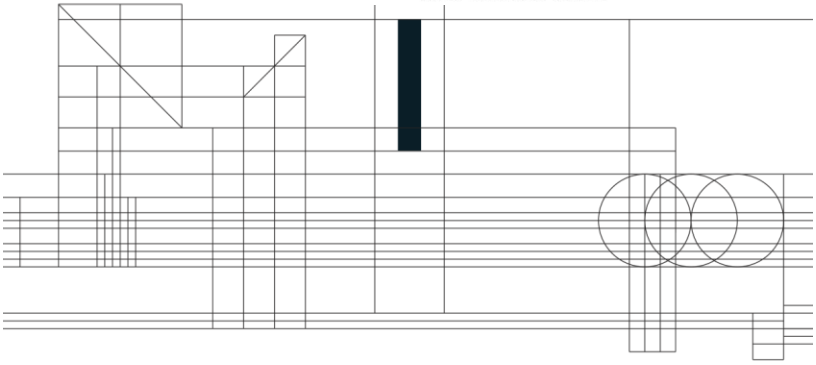
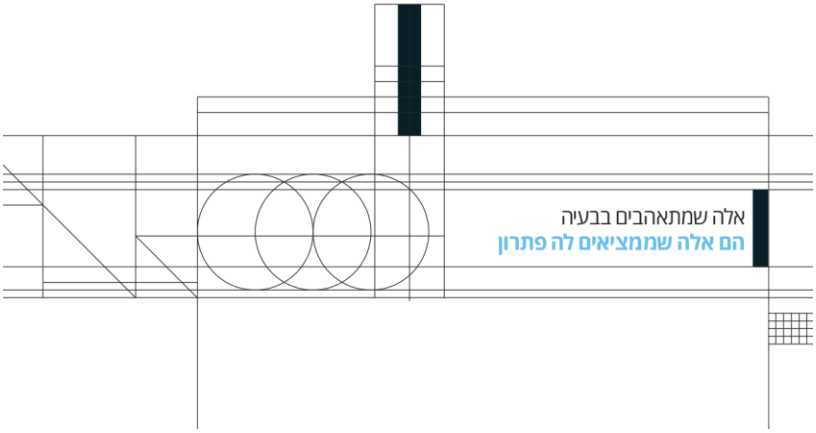
המחלקה להנדסת תעשייה
הפקולטה להנדסה ע"ש איבי ואלדר פליישמן
אוניברסיטת תל אביב



יום עיון מחקרי

10.1.2023

חוברת תקצירים



9:00 - 9:10 – התכנסות ודברי פתיחה
9:10 - 10:00 - מושב א׳
בן חן - בהנחיית ד״ר נטע רבין Spatiotemporal Time-Series Analysis Using Kernel-Based Methods
מירב מופז - בהנחיית פרופ' ארז שמואלי Real-Time Sensing of War's Effects on Wellbeing with Smartphones and Smartwatches
מתן יחזקאל - בהנחיית פרופ' דן ימין Evaluating the Safety Profile of COVID-19 Boosters: Integrating Physician, Patient and Wearable-Level Data
10:00 - 10:15 - הפסקה
10:15 - 11:05 - מושב ב׳
ישי שפירא - בהנחיית ד״ר מור כספי A Column Generation Approach for Public Transit Enhanced Robotic Delivery Services
אור נעים - בהנחיית פרופ׳ עירד בן-גל Malicious Website Identification Using Design Attribute Learning
יובל אשר - בהנחיית ד״ר נעם קניגשטיין Learning to Explain: Explaining Vision Transformers by Explainer-Explainee Models
11:05 - 11:20 - הפסקה
11:20 - 12:10 - מושב ג׳
שחר אייזן - בהנחיית פרופ' טל רביב ופרופ' מיכל צור Dynamic Parcel Routing in Hyper-Connected Networks
יהודה ניסנבוים - בהנחיית ד"ר עמיחי פיינסקי Cross-Validated Tree-Based Models for Multi-Target Learning
מיכאל חבקין - בהנחיית פרופ' ערן טוך Human Aspects of Differential Privacy in Data Markets
12:10 - 12:45 – הפסקה
12:45 - 13:00 - הסברים כלליים על המסלול הישיר לתואר שני מיועד למועמדים למסלול ישיר וכן לסטודנטים לתואר שני שעוד לא בחרו נושא לעבודת גמר
13:00 - 14:00 - מפגש עם סטודנטים ותיקים, מעבדה 424 קומה 4 בניין וולפסון

Spatiotemporal Time-Series Analysis Using Kernel-Based Methods

Ben Hen, Dr. Neta Rabin

Spatiotemporal time-series analysis is a growing area of research. It is applied when data is collected across space and time and describes a phenomenon in a set of particular locations over a predefined time slot. Spatiotemporal data mining research includes different types of tasks, such as forecasting and prediction, clustering, visualization, and its applications span across various domains.

These types of problems pose several challenges, including how to represent the data while capturing the relationships in the temporal and spatial domains and how to efficiently process and analyze the represented data.

Kernel-based techniques play a central role in many algorithms and have become a common way for describing the local and global relationships of data samples. In addition, kernels-based methods may be incorporated into regression methods, thus, enabling to perform forecasting tasks

In this work, we propose a unique forecasting model using Spatiotemporal Auto Adaptive Laplacian Pyramids (SALP), a kernel-based, multi-scale regression method, which learns the spatial connection between nodes, and accordingly predicts future values.

We analyze and compare this algorithm with well-known machine learning-based methods on two real-world different datasets. The results demonstrate the strength and flexibility of the model for addressing diverse types of spatiotemporal prediction tasks.

Real-Time Sensing of War's Effects on Wellbeing with Smartphones and Smartwatches

Merav Mofaz, Prof. Erez Shmueli

Modern wars have a catastrophic effect on the wellbeing of civilians. However, the nature of this effect remains unclear, with most insights gleaned from subjective, retrospective studies. We prospectively monitored 954 Israelis (>40 years) from two weeks before the missile-riddled May 2021 Israel-Gaza armed conflict until four weeks after the ceasefire using smartwatches and a dedicated mobile application with daily questionnaires on wellbeing.

We identified considerable changes in all the examined wellbeing indicators during missile attacks and throughout the war, including spikes in heart rate levels, excessive screen-on time, and a reduction in sleep duration and quality. These changes, however, faded shortly after the war, with all affected measures returning to baseline in nearly all the participants. Greater changes were found in individuals living closer to the battlefield. Furthermore, among those who exhibited greater changes in their reported stress level, we observed a higher proportion of women and individuals from lower socioeconomic backgrounds.

To the best of our knowledge, this is the first study to use a combination of objective and subjective measures in real-time, before, during, and after a war. It sets the know-how for simple, effective, and automatic monitoring. Given the continuing global impact of wars and the urgent need to identify and aid individuals in need in real-time, we believe that our findings are fundamental from both a scientific and public health perspective.

Co-authors: Matan Yechezkel, Prof. Haim Einat, Prof. Noga Kronfeld-Schor, Prof. Dan Yamin, and Prof. Erez Shmueli

Evaluating the Safety Profile of COVID-19 Boosters: Integrating Physician, Patient and Wearable-Level Data

Matan Yechezkel, Prof. Dan Yamin

The effectiveness of the second BNT162b2 (Pfizer–BioNTech) mRNA COVID-19 booster vaccine dose (ie, fourth inoculation) is well established, but its safety has yet to be fully understood. The absence of sufficient vaccine safety information is one of the key contributors to vaccine hesitancy. In this study, we aimed to evaluate the safety profile of the second BNT162b2 mRNA COVID-19 booster vaccine using data from a retrospective cohort and a prospective cohort.

To evaluate the safety profile of the second booster vaccine, we analysed its short-term effects and compared them to those of the first booster by using data from, first, a retrospective cohort of 250 000 random members of the second-largest health-care organisation in Israel (Maccabi Healthcare Services) and, second, a prospective cohort (the PerMed study) of 4698 participants from all across Israel. Individuals who were aged 18 years or older who received the second BNT162b2 mRNA COVID-19 vaccine booster during the vaccination campaign, from Dec 30, 2021, to July 22, 2022, were eligible for inclusion in the retrospective cohort analysis. To be included in the PerMed study, participants needed to be 18 years or older, members of Maccabi Healthcare Services at the time of enrolment, using their own smartphone, and be able to give informed consent by themselves. Participants from the prospective cohort received smartwatches, downloaded a dedicated mobile application, and granted access to their medical records. The smartwatches continuously monitored several physiological measures, including heart rate. For analysis of the prospective cohort data, we used the Kruskal-Wallis test to compare heart rate levels observed before and after vaccination. The mobile application collected daily self-reported questionnaires on local and systemic reactions. Medical records of the retrospective cohort were accessed to examine the occurrence of 25 potential adverse events, and we evaluated the risk differences between 42 days in the periods before and after vaccination in a pairwise method using non-parametric percentile bootstrap.

Our retrospective cohort analysis included 94 169 participants who received the first booster and 17 814 who received the second booster. Comparing the 42 days before and after vaccination, the second booster was not associated with any of the 25 adverse events investigated, including myocardial infarction (risk difference, 2·25 events per 10 000

individuals [95% CI –3·93 to 8·98]) and Bell's Palsy (–1·68 events [–5·61 to 2·25]). None of the individuals was diagnosed with myocarditis or pericarditis following vaccination with the second booster. The prospective cohort included 1785 participants who received the first booster and 699 who received the second booster. We found no significant differences after inoculation with the first booster compared with the second booster (heart rate: day 2 [p=0·3], day 6 [p=0·89]; extent of self-reported reactions [p=0·06]). We found a significant increase in mean heart rate relative to that observed during the week before vaccination (baseline) levels during the first 3 days following the second booster (p<0·0001), peaking on day 2 (mean difference of 1·61 bpm [1·07 to 2·16] compared with baseline). Mean heart rate values returned to baseline levels by day 6 (–0·055 bpm [–0·56 to 0·45] compared with baseline).

In conclusion, both our retrospective and prospective analyses support the safety of the second booster, with our findings reflecting physicians' diagnoses, patients' objective physiological measures, and patients' subjective reactions. We believe this study provides safety assurances to the global population who are eligible to receive an additional COVID-19 booster inoculation. These assurances can help increase the number of high-risk individuals who opt to receive this booster vaccine and thereby prevent severe outcomes associated with COVID-19.

Co-authors: Merav Mofaz, MSc, Amichai Painsky, PhD, Tal Patalon, MD, Sivan Gazit, MD, Erez Shmueli, PhD, and Dan Yamin, PhD.

A Column Generation Approach for Public Transit Enhanced Robotic Delivery Services

Yishai Shapira, Dr. Mor Kaspi

Delivery services of small-size packages are becoming more and more popular in parallel to the rise of on-demand economy services, in particular, e-commerce and fast-food deliveries. In order to meet the growing need for home delivery within urban areas, new innovative delivery concepts have been developed, involving cargo bikes, robots, automated lockers, and more. Autonomous transportation may have a key role in the development of delivery services, and many new technology initiatives that rely on vehicle autonomy capabilities have emerged in recent years. One prominent concept is based on Autonomous Mobile Robots (AMRs). In such systems, small to medium-sized wheel-based robots travel on sidewalks at pedestrian speed, providing point-to-point deliveries. Typically, the robots are powered by small batteries. This feature and the relatively slow traveling speed limit the service range to approximately two to three kilometers.

In the AMR service at the focus of this study, a fleet of robots is distributed among several depot stations, and a set of requests is given, where each request is characterized by a pair of pick-up and delivery locations. To extend the reach of the service, robots are allowed to perform parts of their journey on board public transit vehicles. Specifically, a set of fixed public transit lines that operate in the service area are given. Each fixed line is characterized by a sequence of stations, the traveling times between them, and the frequency of the service. When a robot is assigned to a task, it needs to travel from its origin depot station to a pickup point to load the package, travel to the associated delivery point, and then return to one of the depot stations for recharging. Each leg of this journey can be performed directly by the robot (using battery power) or be performed partially on board a public transit vehicle, which travels at higher speeds and does not require any discharging of the robot's battery.

We consider the static version of the operational problem, i.e., the case where information regarding the requests is known long enough in advance for operational decision-making. In our setting, all given requests must be served, either by the AMRs or by an alternative service (outsourcing) at a greater cost. For each request the problem is to decide upon the means it will be served. Specifically, for requests served by the AMRs, we need to decide

which robot will be assigned to each request, the robots' routes, and the depots to which they will return. This needs to be determined while respecting the maximal number of robots that can be carried simultaneously by a public transit service and the capacities of the robot depots. The goal of the decision problem is to minimize the total operational costs which are composed of AMR's operational costs and costs associated with the alternative service

We develop two Mixed Integer Programming formulations of the problem, namely, an arc-based formulation and a route-based formulation. The former explicitly represents each robot's potential leg and decides upon the legs to be traveled. The latter considers complete feasible routes to select the best route (and robot) for each request. While the route-based formulation has a more compact structure, the number of routes that may be considered grows exponentially with the number of requests, thus this formulation can be solved directly by fully enumerating all potential routes only for very small instances of the problem. To handle medium to large instances we develop a column generation approach. We define an initial set of potentially good routes for every robot-request pair and then formulate the underlying sub-problem of finding new solutions to improving routes as a resource-constrained shortest path problem. Based on the assumption that each robot serves at most one request during the planning horizon, we develop a five-stage dynamic programming algorithm to solve this problem. Furthermore, as the robot-request sub-problems are independent, we apply parallel computing to speed up each column generation iteration.

We have generated hundreds of problem instances, varying in the number of requests, public transit nodes, depot nodes, and robots. Initial results show that the arc-based formulation instances with up to 15 requests can be solved to optimality within a few minutes while the route-based column generation approach can handle instances with 150 requests within five minutes.

Malicious Website Identification Using Design Attribute Learning

Or Naim, Prof. Irad Ben-Gal

Malicious websites pose a challenging cybersecurity threat. Traditional tools for detecting malicious websites rely heavily on industry-specific domain knowledge, are maintained by large-scale research operations, and result in a never-ending attacker-defender dynamic.

Malicious websites need to balance two opposing requirements to successfully function: escaping malware detection tools while attracting visitors. This fundamental conflict can be leveraged to create a robust and sustainable detection approach based on the extraction, analysis and learning of design attributes for malicious website identification.

In this study, we propose a next-generation algorithm for extended design attribute learning that learns and analyzes web page structures, contents, appearances and reputations to detect malicious websites. Large-scale experiments that were conducted on more than 50,000 websites suggests that the proposed algorithm effectively detects more than 83% of all malicious websites while maintaining a low false-positive rate of 2%. In addition, the proposed method can incorporate user feedback and flag new suspicious websites and thus can be effective against zero-day attacks.

Learning to Explain: Explaining Vision Transformers by Explainer-Explainee Models

Yuval Asher, Dr. Noam Koenigstein

We present 'learning to explain' (LTX) - a novel method for producing explanations by learning explanation masks. The LTX framework introduces an explainer-explainee model, in which the explainer learns to explain and justify the explainee's predictions. We demonstrate LTX's ability to produce explanations for ViT models, where it significantly outperforms state-of-the-art methods on multiple explanations and segmentation tests.

Dynamic Parcel Routing in Hyper-Connected Networks

Shachar Eizen, Prof. Tal Raviv and Prof. Michal Tzur

The small parcel delivery industry has grown significantly over the last few years. From an economic perspective, logistic operations in urban areas, notably last-mile delivery, tend to be the most expensive segment of the logistic process. In this study, we adopt a novel framework to economize the delivery of small parcels using the Physical Internet concept. We consider a hyper-connected network of service points, where each point can be served by more than one vehicle route, and parcels may be transferred from their origin to their destination in several legs. We present a method to route and schedule the parcels while considering the service times of the vehicles at the SPs. As a first step in analyzing this intricate problem, our model assumes that the routes of the vehicles are fixed and determined in advance. In particular, we present a dynamic routing policy for the parcels in a given network and a search algorithm to accomplish the vehicle scheduling. A preliminary experiment demonstrated the merits of this approach.

Cross-Validated Tree-Based Models for Multi-Target Learning

Yehuda Nissenbaum , Dr. Amichai Painsky

Multi-target learning (MTL) is a popular machine learning technique which considers simultaneous prediction of multiple targets. The most common MTL schemes utilize traditional linear models and more contemporary deep neural networks. Interestingly, tree-based MTL methods received limited attention over the years. In this work, we introduce a novel tree-based MTL scheme which exploits the correlation between the targets to obtain improved prediction accuracy. Our suggested scheme applies cross-validated splitting criteria to identify correlated targets at every node of the tree. This allows us to benefit from the correlation among the targets while avoiding overfitting. We demonstrate the performance of our proposed scheme in a variety of synthetic and real-world experiments, showing a significant improvement over alternative methods.

Human Aspects of Differential Privacy in Data Markets

Michael Khavkin, Prof. Eran Toch

Differential privacy has been proposed as a rigorous privacy guarantee for computation mechanisms.

However, it is still unclear on how data collectors can properly and intuitively configure the value of the privacy budget parameter ε for differential privacy, such that the privacy of involved individuals is protected.

In this work, we seek to investigate the trade-offs between differential privacy valuation, scenario properties and preferred differential privacy level of individuals participating in a data trade.

Using a choice-based conjoint analysis (N = 139), we mimic the decision-making process of individuals under different data sharing scenarios.

We found that, as hypothesized, individuals required lower payments from a data collector for sharing their data, as stronger perturbation was applied as part of a differentially private data analysis.

Furthermore, respondents selected scenarios with lower ε values (requiring higher privacy levels) for indefinitely-retained data with the purpose of profit generation, than for temporarily-retained data with a non-commercial purpose.

Our findings may help data processors to better tune the differential privacy budget for their data analysis based on both individual privacy valuation and contextual properties.