

אופטימיזציה, אלגוריתמיקה ויישומים מדעיים-הנדסיים

4/3/2021

מתמטיקה

ד"ר מרק ברמן

Title: A graph theoretic solution to a medical student assignment problem.

Abstract: Students of ophthalmology at a medical school in Cape Town, South Africa, need to be assigned to patients with various eye disorders. Pairs of students carry out case studies on these patients, where each student works in several pairs. The problem is to define the pairs in such a way that 1) every pair is assigned precisely one patient and 2) no student is assigned two (or more) patients with the same disease. I will present a solution to this problem based on classical graph theory dating back some 100 years. I will also describe an algorithm I have written for lecturers at the medical school in order to implement the solution.

פרופ"ח חגי כתריאל

כותרת: בעיות אופטימיזציה בהקצאת חיסונים

תקציר: הקצאה של כמות מוגבלת של חיסונים כנגד מחלה מדבקת לאוכלוסייה מצריכה החלטות מורכבות בנוגע לחלוקת החיסון בין קבוצות שונות באוכלוסייה – למשל קבוצות גיל, קבוצות עם רמת חשיפה שונה (כמו עובדי המערכת הרפואית או עובדים חיוניים אחרים), קבוצות עם רמת סיכון שונה, ותושבי אזורים גיאוגרפיים שונים. שאלת ההקצאה מובילה לבעיות מאתגרות באופטימיזציה מתמטית, שבה המשתנים הם כמויות החיסון המוקצות לקבוצות אוכלוסייה שונות, ופונקציית המטרה שאותה רוצים למזער יכולה להיות כמות התחלואה או התמותה. את חישוב פונקציית המטרה יש לבצע בעזרת מודלים מתמטיים להתפשטות המגיפה. מאחר וחיסון של בני קבוצה אחת מעניק הגנה מפני הדבקה גם לבני קבוצות האחרות, הפתרון האופטימלי אינו בהכרח לחסן את אלה שהם בעלי הסיכון הגבוה ביותר מהמחלה. כאשר מדובר בחיסון שזמינותו תעלה בהדרגה במהלך הזמן, כפי שצפוי לקרות עם חיסוני הקורונה, מדובר בבעיית הקצאה דינמית מורכבת.

בהרצאתי אסקור מספר בעיות אופטימיזציה הנוגעות להקצאת חיסונים ואציג תוצאות ידועות ושאלות פתוחות, תוך התייחסות למחקרים הנערכים ברחבי העולם לגבי חיסון לקורונה.

Title: Deep Learning for Hate Speech Detection

Abstract: Hate speech detection is a critical problem in social media, being often accused for enabling the spread of hatred and igniting violence. Hate speech detection requires overwhelming computing resources for online monitoring as well as thousands of human experts for daily screening of suspected posts or tweets. Recently, Deep Learning (DL)-based solutions have been proposed for hate speech detection, using modest-sized datasets of few thousands of sequences. While these methods perform well on the specific datasets, their ability to generalize to new hate speech sequences is limited. Being a data-driven approach, it is known that DL surpasses other methods whenever scale-up in trainset size and diversity is achieved. Therefore, we first present a dataset of 1 million hate and non-hate sequences, produced by a deep generative model. We further utilize the generated data to train a well-studied DL detector, demonstrating significant performance improvements across five hate speech datasets.

Title: Online and approximation algorithms, with applications to optical networks.

Abstract: I will survey results on some optimization problems, that stem from the model of optical networks. Stated in graph-theoretic terms, these problems can be defined as coloring of paths, where the optimization criteria are derived from constraints originated in optical networks. Many of these problems can be stated in terms of the theory of scheduling. For the online problem of minimizing the number of Add-Drop-Multiplexers (ADMs) we present an optimal 1.75-competitive algorithm (that is, an algorithm that uses at most 75% ADMs more than the optimal one). For the offline problem of minimizing total busy time in parallel scheduling we present a 4-approximation algorithm (that is, an algorithm that produces a scheduling that is at most 4 times the optimal one).

Title: Textual Material Analyzing using Signal Processing Methods

Abstract:

A phenomenon of the so-called “mediated culture” is a very outstanding aspect in the modern age. Mass media and social networks mirror the people’s intentions and comprehending specific themes via the massive documents amounts distributed in the virtual space. We discuss the novel attitude to represent successively created textual material in the digital signal fashion. Such a methodology applied to various information sources makes it possible to propose a new standpoint to treat many practical and theoretical tasks connected to disclosing significant social events, plagiarism detection, data fabrication, and authorship determination. In the present talk, we present and discuss new techniques developed in this framework for text to linguistic signal transformations and their combinations with Deep Learning CNN, LSTM, and Digital Signal Processing together with several practical examples. This methodology may be of interest from practical and theoretical perspectives in areas such as industrial engineering, applied mathematics, and signal processing.

1. Z. Volkovich, A Short-Patterning of the Texts Attributed to Al Ghazali: A “Twitter Look” at the Problem, *Mathematics* 2020, 8(11), 1937; <https://doi.org/10.3390/math8111937> (registering DOI)
2. O. Bernikova, O. Granichin, D. Lemberg, O. Redkin, Z. Volkovich, Z. Entropy-Based Approach for the Detection of Changes in Arabic Newspapers’ Content. *Entropy*, 22, 441. 2020
3. K. Amelin, O. Granichin, N. Kizhaeva, Z. Volkovich, Patterning of writing style evolution by means of dynamic similarity, *Pattern Recognition*, 77, 45-64, 2018
4. Z. Volkovich and A. Avros, Detection of computer-generated papers using One-Class SVM and cluster approaches, Conference: 14th International Conference on Machine Learning and Data Mining MLDM 2018, pp 42-55
5. N. Kizhaeva, Z. Volkovich, O. N. Granichin, O. Granichina and V. Kiyaeva, Spectral profiling of writing process, Conference: 2017 IEEE Conference on Control Technology and Applications (CCTA), August 27-30, 2017, Coast, Hawaii, USA, pp. 2063-2068.

6. D. Shalymov, O. Granichin, L. Klebanov, and Zeev Volkovich, Literary writing style recognition via a minimal spanning tree-based approach, Expert System with Application 61, 1, 145–153, November 2016.
7. Z. Volkovich, O Granichin, O. Redkin and O. Bernikova, Modeling and visualization of media in Arabic, Journal of Informetrics, 10, 2, 439–453, May 2016.
8. D. Lemberg, A. Soffer, and Z. Volkovich, New Approach for Plagiarism Detection, International Journal of Applied Mathematics 29, 3, 365-371, June 2016

ד"ר מירי וייס-כהן

Title: Shape Optimization Definiteness using NURBS Curves and Genetic Algorithm

Abstract: A Shape Optimization method performed on multi material Topology Optimization results is presented. The approach is based on two major phases, first, parametrization the material boundaries and second, optimizing those boundaries to definite shapes. The parameterization process involves identifying the boundaries of different materials and representing the boundaries by NURBS curves. This approach conveys the ability to define a new parametrized internal boundary profile, which is much more flexible. Once a parametrized model is defined, the optimization is performed using the parameters of the boundary curves as the optimization variables.

A Genetic Algorithm (GA) is used to optimize the best variables (control points and weights) for achieving the best coordinates of the boundary curves, sufficing the Fitness function criteria. The method was tested on a C-Core magnet, and results are presented.

*This is a joint work with João Batista, and Frederico G. Guimarães

תעשייה וניהול

ד"ר אילנה בן-דוד

Title: PET Appointment Scheduling: a Dynamic Empirical Based Solution and Tool for Maximizing Resource Utilization.

Abstract: Positron Emission Tomography (PET) scan is an example of advance and expansive diagnosis tool (to identify, among other, cancer and Alzheimer diseases) that has an increase demand in the past two decades. Since scan equipment and its operation are rather expensive, there is a need to develop a way to utilize them more effectively. In PET scan, a radiopharmaceutical (radioactive

substance) is injected to the patient prior to the examination, making the timing of the scan to be rigid. This restriction makes patient appointment scheduling challenging.

Previous work suggest that the scan time can be predicted if some patient related details are known in advance (e.g., patient's BMI, height, and body part to be scanned), while another provides Simulation-Based Sequential (SBS) algorithm for appointment scheduling, when scan distribution is known in advance and characterizes the whole population. We suggest utilizing these two works and creating a methodology and a tool that balances between system utilization and quality of service.

The aim of this work will be to develop an empirical based methodology to classify patients and generate multi-type distributions of service duration. In this way, instead of using a general distribution for all patients, we can match a more adapted distribution to each patient and fix better appointments (using data-mining). The next step should be to build a generic and dynamic tool to form a simple and intuitive user interface. The tool should extract real patient service data, generate the different distributions according to patients' classification and use the data to determine the best appointment scheduling policy for a known sequence of patients to be scheduled on a specific day. Finally, since this type of scan equipment and its operation are rather expensive, one of the purpose of the developed tool would be to provide a way to increase resource availability while maintaining high utilization and pre-determined service levels. One way to increase the resource availability would be to dynamically decide, according to a given sequence of patients, the optimal number of patients to schedule in the same working day in order to balance between regular and extra working hours and increase the resource utilization. The decision should be a function of the expected end of day and the optimal balance between regular hours to overtime.

*This is a joint work with Y. Marmor and B. Shnits

גברת אילה אלקובי

כותרת: מזעור זמן מחזור ליח' מודפסת בתלת מימד

תקציר:

Additive Manufacturing (AM) או ייצור בהוספה, הינה תפישה מהפכנית שבניגוד לתפישה שהייתה רווחת עד כה – ייצור בהסרה או Subtractive Manufacturing. הטכנולוגיה שמממשת את התפישה היא הדפסה תלת מימד (3DP). זהו תהליך שבמהלכו מייצרים אובייקט תלת מימדי באמצעות הוספת שכבה של החומר מעל גבי שכבה קודמת. הודות ליתרונות

המלווים בהדפסת תלת מימד הכוללים ייצור על בסיס התאמה אישית, ייצור אובייקטים בעלי גאומטריה מורכבת, הקטנת בזבז חומרים, הקטנת עלות שינוע, גמישות בייצור ועלויות שפחתו באופן דרסטי בשנים האחרונות, שיטת ייצור זו הפכה לנחלתם של ענפי ייצור מגוונים. שתי הדרכים הנפוצות ביותר להדפסה בתלת מימד הן (SLS) Selective Laser Sintering ו-Fused Deposition Modeling (FDM), הראשונה כוללת פיזור אבקת מתכת על משטח ההדפסה, בגובה מסויים והתכתה במקומות הרצויים ע"י לייזר ואילו בשנייה נעשה שימוש בחומר תרמופלסטי מותך שמוזרק ע"י זרבובית ומונח על גבי משטח ההדפסה, בגובה מסויים, במקומות הרצויים. במחקר נתמקד בגישת SLS. הפוקוס במחקר עד כה היה על פיתוח חומרים ופיתוח טכנולוגיות הדפסה. ישנו מחסור במחקרים ששופכים אור על איך לייעל את התהליך שמתחיל בתכן המוצר (Product Design), עובר בתכן התהליך של ההדפסה (Printing) (Process) וכלה בתהליך הגימור של המוצר (Post-processing). המחקר שלנו מתמקד באופטימיזציה של זמן המחזור ליח' מודפסת בתהליך הנ"ל. לצורך כך, זוהו פרמטרים קריטיים לאורך שלושת השלבים, פותחו פונקציות לכימות המרכיבים השונים שמשפיעים על זמן המחזור. פרמטרים של אוריינטציית המוצר (הזווית שבה מונח על משטח ההדפסה), גובה שכבת הדפסה, מהירות סריקה ועוד זוהו כמשפיעים על זמן המחזור שמהווה את פונקציית המטרה. פונקציות מטרה אחרות יכולות לכלול חוזק, חספוס, עמידה בטולרנסים ועוד. לרוב, ישנו טרייד אוף בין סוגי המדדים הנ"ל, ז"א קומבינציית פרמטרים שמביאה לאופטימום מדד תפעולי כלשהוא, עלולה במקביל לפגוע באופן משמעותי במדדים המכניים ולהפך. המידול יתבסס על נתונים אמיתיים שנאספו מארגון המתמחה בהדפסות תלת מימד. בשלב הראשון, נמדל את הבעיה עבור מקרה פרטי שבמסגרתו נדרש לייצר מוצר אחד בלבד במנת ייצור אחת. לאחר המידול הראשוני הנ"ל, נרחיב את המקרה הפרטי לתיאור מתמטי כללי של מספר מנות ייצור שכל אחת מהן מורכבת מכמה מוצרים, כך שבבעיה זו נדרש גם לקבל החלטה באשר להקצאת המוצרים שישוּבצו בכל מנת הדפסה כדי לצמצם במידת האפשר את זמן המחזור מבלי לפגוע באופן משמעותי בטיב פני השטח של המוצרים המודפסים. תהליך המידול כולל תכנות מתמטי, שימוש בעצי סיווג (CART) לזיהוי פרמטרים קבועים בעלי השפעה על מדדי התכן והמכניקה ורשתות עצביות. תהליך האופטימיזציה יכול שיטות מוכרות לטיפול בבעיה מרובת מטרות שהן גישת חזית היעילות ותכנות לפי מטרות כאשר הפוקוס שלנו יהיה על פיתוח היוריסטיקה לפתרון שמנצלת את המבנה של הבעיה.

*עבודה זו מבוססת על פרוייקט גמר בהנחיה משותפת של חוסין נסראלדין ואביב גיבלי

ד"ר ולנטין ונצק

Title: The Number Needed to Treat Adjusted for Explanatory Variables in Regression and Survival Analysis: Theory and Application

Abstract: The number needed to treat (NNT) is an efficacy index that is commonly used in randomized clinical trials. The NNT refers to the average number of patients that are needed to treat in order to attain one successful outcome due to treatment. We introduce a systematic framework to model and estimate the adjusted and the marginal NNT in the presence of explanatory variables, in various models with dichotomous and nondichotomous outcomes. The adjusted NNT is

illustrated in a series of five examples; ANOVA, linear and logistic regression, Kaplan-Meier, and Cox regression. In addition, we establish and prove mathematically the exact relationship between the adjusted and the marginal NNT in the presence of explanatory variables. We introduce four different methods to calculate the asymptotically-correct confidence intervals for both indices. Finally, we conduct a simulation study to provide a numeric demonstration of the aforementioned theoretical results and the five examples. Numerical analysis showed that the parametric estimators of the NNT with bootstrap-based confidence intervals outperformed other examined combinations in linear and generalized linear models, while in survival analysis the analytical type of confidence intervals are the recommended ones. An R-package has been developed and made available online to calculate the adjusted and the marginal NNTs with their corresponding confidence intervals.

*This is a joint work with Y. Goldberg , S. Z. Levine