ORMS TOMORROW

informs Student Magazine

Included in this issue...

STUDENT COMPETITION WINNERS

CONTRIBUTED ARTICLES

@ORMS_Tomorrow

facebook/ORMSTomorrow

linkedin/groups/3788040

Cover image sourced from WIRED.com.

Fall 2019

ORMS TOMORROW informs Student Magazine

Fall 2019

Dear Reader,

We are pleased to present you with the Fall 2019 edition of OR/MS Tomorrow. This issue has been made possible by the hard work of our entire staff.

We are excited to announce the results from our first ever student writing competition conducted earlier this year. The three winning articles, spanning topics from AI in healthcare to OR in gerrymandering, are published in this issue. The judging committee, consisting of our writing staff and editorial board members, evaluated the submissions based on criteria such as argument, topic and form.

Breaking tradition from our theme-based publications, this issue brings to you contributed articles ranging from biodiversity preservation to market design. We also include an informational piece on CMU INFORMS Student Chapter's student-run competition, YinzOR, which has steadily grown over the last several years in participation and impact. We wish them continued success in the years to come.

Additionally, to engage more broadly with the OR/MS student community, our staff has made regular contributions to a new student column at OR/MS Today. Contributed pieces include navigating the OR/MS academic job market - one from a candidate's perspective (here) and one from a hiring committee's perspective (here) – and a column on the industry job market (here). We also compiled an article on student volunteering (here) with perspectives from six contributors on how volunteering impacted their professional and personal lives.

We thank the OR/MS community for its continued engagement with our team. We hope you enjoy this issue's content, and we always welcome your thoughts regarding OR/MS Tomorrow via email at `orms_tomorrow@mail.informs.org`.

Note: We are hiring new team members across the board. More information can be found here.

Best Regards,

Hossein Badri and Rahul Swamy,
Co-Lead Editors,
OR/MS Tomorrow

Operations Research v. Gerrymandering

Samuel Gutekunst
FIRST PRIZE

The U.S. legal system is facing a flood of court cases that argue that certain political districts have been drawn to provide partisan advantage. Such a phenomenon is known as *partisan gerrymandering*, but demonstrating partisan gerrymandering has proved difficult. In the words of Justice Ginsburg, "the court...has not found a manageable, reliable measure of fairness for determining whether a partisan gerrymander violates the Constitution" (Liptak, 2017). Operations research is being applied to meet this challenge.

Figure 1, modified from Ingraham (2015), illustrates gerrymandering and its impact. Fifty people are to be split into five equally-sized, contiguous districts. Each person belongs to either party S (smiley-faces) or party A (angry-faces), and Figure 1 shows two districting plans with extremal outcomes. Districts where S wins are shaded. In the left plan, party S wins 100% of the districts despite having only 60% of the votes. Conversely, in the right plan party A wins a majority of the districts despite having a minority of the votes.

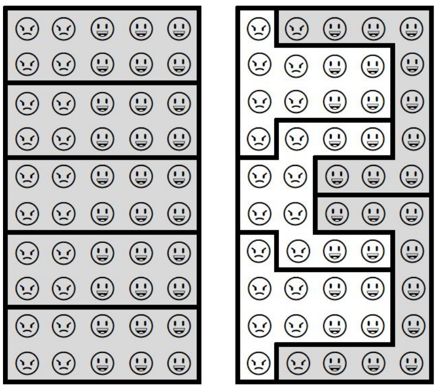


Figure 1

Notice that, in the right plan, party A is able to systematically waste S votes. In the three non-shaded districts that A wins, the four losing S votes are wasted: they could be used to win other districts. In the two shaded districts, party S only needs five votes to win (or, at least, tie); the four extra S votes in each are again wasted. This example begs the question: can we use analytics to measure if district lines are fair?

Stephanopoulos and McGhee recently proposed the *efficiency gap* as a litmus test for partisan gerrymandering (Stephanopoulos and McGhee, 2015). Their proposed formula accounts for wasted votes exactly as above: In districts where a party loses, every vote cast for that party is wasted. In districts where a party wins, every vote cast for that party beyond the 50% threshold for winning is wasted. In every single district of the S-favored plan of Figure 1, for example, party S wastes one vote while party A wastes four. Let W_A and W_S respectively denote the number of wasted votes for A and S across an entire districting plan, and let T denote the total number of votes cast. The efficiency gap then compares the proportion of votes wasted by each party:

$$EG = (W_A - W_S)/T.$$

An efficiency gap near zero indicates a plan where the votes of neither party are being systematically wasted; Stephanopoulos and McGhee propose that an efficiency gap of 0.08 or more indicates that party A is being gerrymandered against.[1] In the left plan of Figure 1, the efficiency gap is $EG = (20-5)/50 = 0.3$ and indicates that party A is being gerrymandered against. In the right plan, in contrast, the efficiency gap is $EG = (5-20)/50 = -0.3$ and indicates that party S is being gerrymandered against.

It is a rare phenomenon, however, that a single number can capture all of the context necessary for making a decision. For example, since the 90's Massachusetts has had either nine or ten seats in the U.S. House of Representatives. Despite a sizable proportion of Republican voters (roughly 30%-40%), no Massachusetts Republican has won a seat in the House since 1994. Such an extreme outcome naturally raises suspicions of gerrymandering, but the issue Republican voters face in Massachusetts is instead based on their geographic distribution throughout the state. As a toy example, consider Figure 2 which has 50 towns to be split into five districts of ten towns each. Every town has 2/3 support for S; any district composed of ten towns will have 2/3 support for party S; despite 1/3-map-wide support for A, party A will not win a single seat!

The political distribution in Massachusetts is less contrived, but the situation is analogous: the locales with majority Republican support are not numerous or clustered. Under election data from several recent state-wide elections, Duchin et al. (2018) uses analytics to argue that there is not a possible districting plan that would give the Republicans a single seat in the U.S. House. Moreover, there is no possible districting plan that produces an efficiency gap of less than 11%. The question

[1] This rule of thumb is specifically for state legislative districting plans where a large efficiency gap is likely to persist through multiple election cycles. Symmetrically, a gap less than -0.08 indicates A is gerrymandering Stephanopoulos and McGhee (2015).

thus becomes: how can the efficiency gap account for local geography and politics so that it can be used in complex, real-world settings?

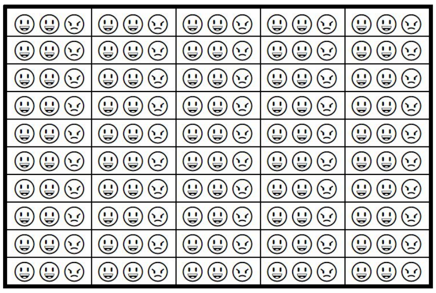


Figure 2

One tool to put metrics like the efficiency gap into context is *Markov chain Monte Carlo (MCMC) simulation*. Here MCMC simulation starts from a proposed districting plan and generates a large collection of random plans (an "*ensemble*"). Generally, these procedures start with a proposed plan and iteratively make small, random adjustments (e.g. moving a small number of voters from one district to another). Such adjustments are constrained to only happen when they preserve traditional districting principles (e.g. keeping districts contiguous).

As part of a recent case in the Pennsylvania Supreme Court, Duchin used MCMC to generate an ensemble with billions of possible districting plans similar to a suspicious plan and computed the efficiency gap of each (Duchin, 2018); the resulting histogram is shown in Figure 3 (modified from Duchin (2018)) where the red line indicates the efficiency gap of the suspicious plan Duchin was evaluating.

Figure 3 provides strong evidence that the suspicious plan was drawn with partisan intent: it seems highly unlikely that a plan designed without partisan intent would have such an extreme, large efficiency gap.

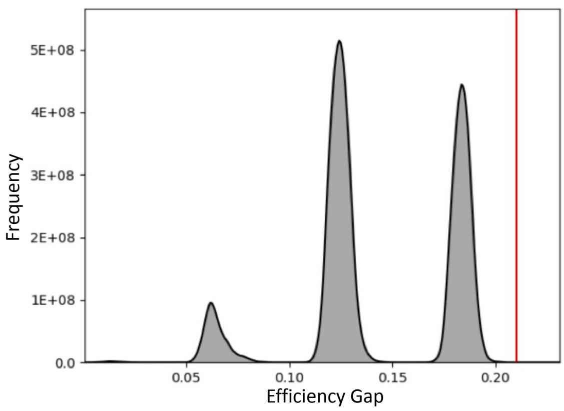


Figure 3

The efficiency gap is an exciting and powerful tool for identifying partisan gerrymandering. It is most useful when it can be evaluated in the proper political and geographical context; sophisticated analytics tools from the operations research community do just that. Many open questions remain, from developing and strengthening metrics for identifying gerrymandering to developing algorithms that draw fair districts. As progress on these questions continues, operations research will play a pivotal role in evaluating and disseminating these tools for use in the real world.

Samuel Gutekunst is a fourth year PhD Student in the School of Operations Research and Information Engineering at Cornell University.

References

1. Duchin, M. (2018). Outlier analysis for Pennsylvania congressional redistricting.
2. Duchin, M., Gladkova, T., Henninger-Voss, E., Klingensmith, B., Newman, H., and Wheelen, H. (2018). Locating the representational baseline: Republicans in Massachusetts. arXiv:1810.09051.
3. Ingraham, C. (2015). This is the best explanation of gerrymandering you will ever see. The Washington Post. Image originally attributed to Stephen Nass.
4. Liptak, A. (2017). On Justice Ginsburg's summer docket: Blunt talk on big cases. The New York Times.
5. Stephanopoulos, N. O. and McGhee, E. M. (2015). Partisan gerrymandering and the effienciency gap. University of Chicago Law Review, 82:831.

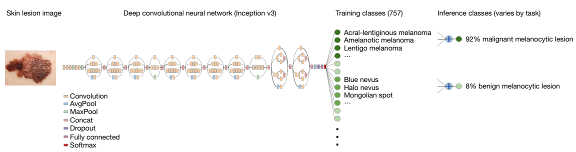
Artificial Intelligence: The New Hype in Healthcare

Jingmei Yang
SECOND PRIZE

THE future of the healthcare industry has never been as bright as today. The application of Artificial Intelligence (AI) has made remarkable progress in its impact across a range of medical applications including drug discovery, remote patient monitoring, medical diagnostics, risk management, virtual assistants and hospital management. Improvements in accuracy and efficiency are made possible by innovations in deep neural networks and high-end computational resources in combination with increasing availability of medical data. In this report, we summarize some significant AI innovations in healthcare followed by a discussion on future challenges and opportunities.

AI is achieving the expert-level disease diagnostics. AI is achieving expert-level prediction and diagnosing of diseases based on image recognition using deep neural networks.

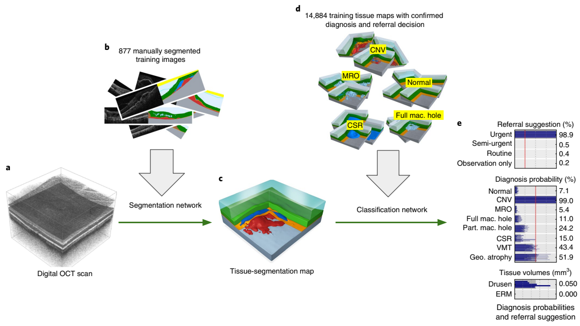
AI can help dermatologists diagnose skin cancer. Roughly 5.4 million incidences of skin cancer are reported annually in the United States. Early detection of skin cancer allows medical practitioners and patients to take proactive action in treatment (Rogers et al., 2012). In 2017, a deep convolution network was built for automated dermatology by Esteva et al. (2017) at Stanford. Trained on 129,450 images comprised of 2,032 diseases, the system achieved accuracy levels on par with dermatologists. Powered with this framework, one possible application is to mobile devices which can potentially extend the reach of specialists, widening the scope of primary care practice and offering a low-cost approach to diagnostic care.



Convolution network for skin diseases detection, Source: Esteva et al. (2017), picture courtesy of the authors.

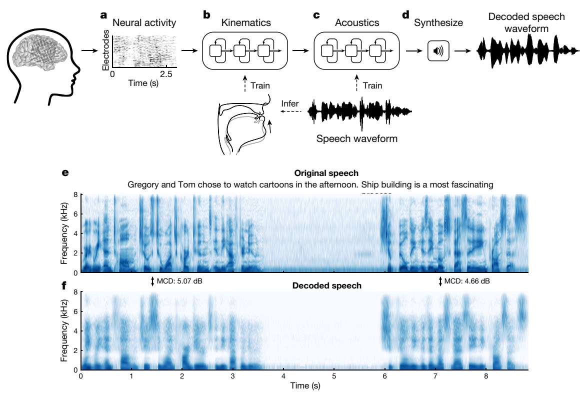
AI supports the diagnosis of cardiovascular diseases. Cardiovascular risk can be revealed through analysis of retinal fundus images, a non-invasive way to visualize blood vessels. Companies such as Google have dedicated resources to developing models which can extract and quantify risk markers in retinal images. Risk factors such as age, gender, smoking status, and systolic blood pressure have been used in well established cardiovascular risk calculators such as SCORE (Framingham and Systemic Coronary Risk Evaluation), however, were previously not extracted from retinal images. In the Nature paper by Poplin et al. (2018), they demonstrated how to identify the presence of these risk factors in the retina. This neural network model is capable of predicting cardiovascular risk directly through retinal images and quantifies the risk factors to a degree of precision not achieved before Poplin et al. (2018). With the aid from such an AI system, cardiovascular risk can be obtained immediately from non-invasive retinal images.

AI can be used for retinal disease diagnosis. A research team in DeepMind developed an innovative framework that investigates eye scans from routine clinical practice. In a paper published in Nature, De Fauw et al. (2018) demonstrated that an AI system is capable of automatically identifying retinal diseases in only a minute. Additionally, the system can classify patients based on their severity and redistribute medical resource to the patients most in need of urgent care. This prioritization attempts to reduce the long delays between scan and treatment resulting from the complexity of interpreting the Optical Coherence Tomography images and the reducing numbers of qualified interpreters. This framework can make referral recommendations for over 50 sight-threatening retinal diseases at a level comparable to clinical experts and has great potential for preventing patients with diabetic retinal disease from sight loss.



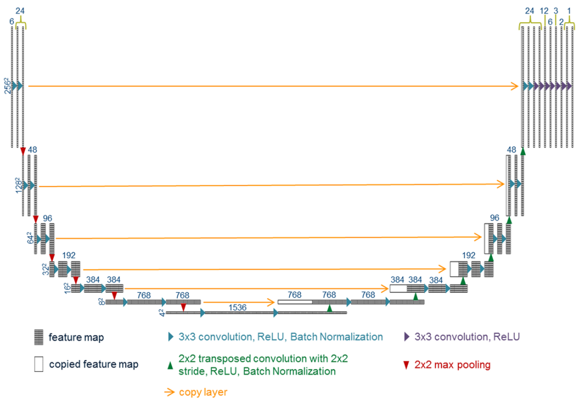
Retinal disease diagnosis and referral suggestions framework. Source: De Fauw et al. (2018), picture courtesy of the authors.

AI can support speech synthesis. Edward Chang, a neurosurgeon, collaborated with a research team at the University of California - San Francisco successfully map brain signals into auditable expression using AI (Anumanchipalli et al., 2019). Recurrent neural networks were trained on brain signals collected from five epilepsy subjects to predict articulatory movements in the tongue, lips, jaw, and larynx. Next, the networks mapped the estimated movements onto synthesized phonic speech. This pioneering technology has a direct application in speech-decoding devices for people with speech impairments resulting from stroke, traumatic brain injury, or neurological diseases like multiple sclerosis and Parkinson's disease.



The neural decoding process with AI, picture taken from Anumanchipalli et al., (2019).

AI in Radiology. Recently there has been a massive amount of publications using deep learning to process medical imaging such as image denoising, segmentation, and super-resolution. Of particular interest is a research team from the University of Texas at Southwestern Medical Center (UTSW) lead by Dr. Steve Jiang. Due to increasing treatment modalities in radiation therapy, treatment planning is complicated and time-consuming for dosimetrists. With an effort to cut down on the planning time while maintaining quality, the team from UTSW built a convolutional neural network to predict the radiation therapy dose for prostate cancer patients (Nguyen et al., 2019). By mapping the patient's contours into local and global features, the model is empowered to predict a dose distribution with impressive accuracy. If equipped with this dose prediction model in clinical practice, physicians can use the prediction as a preliminary plan and cooperate with dosimetrists for further tailoring, making the planning workflow smooth and efficient.



U-net architecture with additional CNN layers used for dose prediction. Source: Nguyen et al. (2019), picture courtesy of the authors.

Future challenges. Despite the promising applications, it is acknowledged that AI has unique limitations when applied to healthcare such as clinical interpretability, data heterogeneity, and patient privacy.

Deep learning is often treated as a black box; its features and parameters are challenging to understand and interpret in a healthcare setting. Not being able to explain the internal mechanics of a model is a barrier for the broad adoption of AI since clinical practitioners place trust heavily on interpretability. As such, researchers have worked on developing interpretable AI systems. Hopefully, this work will make AI systems easier to understand and adopt in practice. Data heterogeneity is another challenge. As shown in the study led by Zech et al. (2018), image data from different hospitals, vendors of imaging modalities and scanners or reconstruction conditions has significant influence on the model performance. Acquiring a training dataset from diverse settings, nonidentical populations, or multiple institutions is beneficial to overcome this problem. Patients' privacy is also a concern. Unlike other domains, the healthcare industry handles a lot of sensitive patients' information. How to balance the usage of all the data and control the infringement of privacy of patients requires care and effort when developing models.

External validation is necessary for AI to prove its promise. All AI-based models need to be validated with clinical trials to test its practical value and performance in a real-world setting. If clinical performance is validated and interpretability of the models are enhanced, AI has the potential to positively impact clinical practice with better performance and increased efficiency.

Jingmei Yang is a PhD Student in Industrial Engineering at the University of Texas at Arlington.

References

1. Anumanchipalli, G.K., Chartier, J., and E.F. Chang (2018). Speech synthesis from neural decoding of spoken sentences. Nature, 568(7753):493-498.
2. De Fauw, J., Ledsam, J.R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., Askham, H., Glorot, X., O'Donoghue, B., Visentin, D., et al. (2018). Clinically applicable deep learning for diagnosis and referral in retinal disease. Nature Medicine, 24(9):1342-1350.
3. Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., and S. Thrun (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639):115-118
4. Nguyen, D., Long, T., Jia, X. Lu, W., Gu, W., Iqbal, Z., and Steve Jiang (2019). A feasibility study for predicting optimal radiation therapy dose distributions of prostate cancer patients from patient anatomy using deep learning. Scientific reports, 9(1):1076.
5. Poplin, R., Varadarajan, A.V., Blumer, K., Liu, Y., McConnell, M.V., Corrado, G.S., Peng, L., and D.R. Webster (2018). Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. Nature Biomedical Engineering, 2(3):158-164.
6. Rogers, H.W., Weinstock, M.A., Feldman, S.R., and B.M. Coldiron (2015). Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the U.S. population, 2012. Jama Dermatology, 151(10):1081-1086.

The Longest Induced Path Problem: How Far Can Information Travel?

Jessica Leung
THIRD PRIZE

IMAGINE you posted something on social media, it may be a food review for that gorgeous restaurant you went to the other day, some brilliant ideas worth sharing or even just a cool meme that you wish everyone can see. Some of your friends shared your post and their friends shared it further...have you ever wondered how far your post could travel and who ended up seeing it? To answer this question, we wish to find the longest induced path in the social network.

The longest induced path problem is to find the largest subset of nodes in a graph that gives a simple path without cycles (Johnson and Garey, 1979; Di Giacomo et al., 2016; Esperet et al., 2017). To do so, we have to consider all possible acyclic paths formed by all possible subsets of nodes in the graph. Considering its combinatorial nature, the longest induced path problem is known to be NP-hard (Johnson and Garey, 1979). Given the inherent computational complexity of the problem, the literature offers limited insight on exact solution approaches in general graphs but instead focuses on subclasses of the problem that are polynomially solvable and on tightening the bounds of the induced path length (Gavril, 2002; Courcelle et al., 2000; Esperet et al., 2017). Fortunately, leveraging the technological and algorithmic advancement in recent decades, we can now formulate the longest induced path problem as an integer programming problem and solve it with an exact solution for any general graph using standard off-the-shelf solvers (e.g. Gurobi).

With the abundance of data available to us, we can now study the longest induced path problem and its implications in many different contexts. For instance, in our motivating example, finding the longest possible path of information transmission in a social network allows us to identify the seed of the information cascade and the final entity which sees this piece of information in the network. Other applications include studying the communication property of complex systems to enhance our understanding of the interacting elements within the system, identifying gene regulation networks in embryonic development, and diagnosing fault in multiprocessor networks. Thus, in light of the numerous practical applications, the longest induced path problem is an interesting class of network optimisation problems to study.

How exactly do we find the longest induced path? Matsypura et al. (2019) provided three conceptually different approaches that yield an exact solution to the longest induced path problem in a general graph setting; developed using interesting facts about the longest induced path.

The first approach is underpinned by the simple observation that the longest induced path in a graph is the subgraph with the maximum diameter. We consider every possible connected subset of nodes in the original graph as a subgraph. In each subgraph, we compute the shortest paths between each pair of nodes where the greatest length of these paths is the diameter of the subgraph. This can be formulated as an integer programming problem that searches for a connected subgraph with the largest diameter which simultaneously consists of the smallest number of nodes among all possible subgraphs. Naturally, the subgraph that gives the maximum diameter is essentially the longest induced path.

The second approach stems from the feature that the graph average distance is simply the average length of the shortest path between any two nodes in a graph. For any connected graph with n nodes, the graph average distance is known to be at most $(n+1)/3$ and is equal to $(n+1)/3$ when the graph is a path (Doyle and Graver, 1977). Therefore, we can search for the largest connected subgraph where its average distance is at least $(n+1)/3$.

The third approach leverages the intuition that an induced path can also be viewed as a walk in a graph with no short-cuts. Imagine we take a walk in the graph, starting with one of the nodes. Then we choose the next node to visit from one of the neighbours of the current node. We traverse through other nodes one at a time until we run out of time or when no further moves are possible. During our journey, we ensure that our route is an induced path by never visiting a single node twice and never taking shortcuts. In such a way, we can formulate an integer programming problem looking for the longest walk.

What's next? Equipped with the tools to find the longest induced path in any general graph, we are offered an alternative perspective to view problems in real world complex systems. We saw that the longest induced path in a social media network let us know how far information travels. From a marketing point of view, this not only indicates the reach of an advertisement but also identifies a group of potential customers that quite literally *share* the same interest. With the additional information on customer segmentation and market reach, businesses can now tailor their digital marketing strategies to different interest groups. A promising direction for future research is to consider different types of induced paths and to develop more computational efficient methods. This opens up new opportunities to gain a better understanding of the communication properties of underlying networks in many other real-life applications.

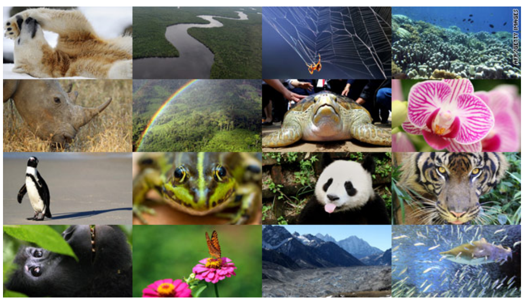
Jessica Leung is a PhD candidate in the Discipline of Business Analytics at the University of Sydney Business School.

References

1. Courcelle, B., J. A. Makowsky, and U. Rotics (2000). Linear time solvable optimization problems on graphs of bounded clique-width. Theory of Computing Systems 33(2), 125–150.
2. Di Giacomo, E., G. Liotta, and T. Mchedlidze (2016). Lower and upper bounds for long induced paths in 3-connected planar graphs. Theoretical Computer Science 636, 47–55.
3. Doyle, J. and J. E. Graver (1977). Mean distance in a graph. Discrete Mathematics 17 (2), 147–154.
4. Esperet, L., L. Lemoine, and F. Maffray (2017). Long induced paths in graphs. European Journal of Combinatorics 62, 1–14.
5. Gavril, F. (2002). Algorithms for maximum weight induced paths. Information Processing Letters 81(4), 203–208.
6. Johnson, D. S. and M. R. Garey (1979). Computers and intractability: A guide to the theory of NP-completeness. WH Freeman.
7. Matsypura, D., A. Veremyev, O. A. Prokopyev, and E. L. Pasiliao (2019). On exact solu- tion approaches for the longest induced path problem. European Journal of Operational Research.

Using Mathematical Optimization to Preserve Biodiversity

Zulqarnain Haider
EDITORIAL BOARD MEMBER



Biodiversity. Reprinted from CNN (2010).

MATHEMATICAL optimization is defined as the "science of better" and is used every day by academics and practitioners to solve a myriad of challenging and interesting problems. Optimization models form the core of algorithms that humans use to move intercontinental cargos, transport billions of human beings, deliver better health outcomes for patients, coordinate emergency relief during disasters, design financial portfolios, and achieve bottom-line savings for corporations. Can magical powers of mathematical optimization also be used to conserve millions of animal and bird species and in the process save our planet and its biodiversity? The answer is a resounding yes. Conservation planning is a burgeoning, although still underserved, field of study that concerns itself with the issues related to maintaining and increasing biodiversity. The unprecedented population growth in the last century coupled with rapid industrialization, and urbanization has strained our planet's resources. Agriculture and other economically beneficial land use alternatives have caused rampant deforestation resulting in the alteration and loss of the habitats for many species (Polasky et al., 2008). According to the Red List of Threatened Species maintained by the International Union for Conservation of Nature, about 28,000 animal and plant species out of more than 105,700 listed are threatened with extinction (IUCN, 2019). Preserving biodiversity is crucial to human societies and the future of planet Earth. Hence, its slow erosion constitutes a threat as consequential as that posed by climate change (Billionnet, 2013). In this article, I briefly describe some of the key problems and issues in the area of conservation planning and how mathematical optimization can help decision-makers in the modeling and implementation of decisions and strategies to protect biodiversity.

Chief among conservation planning problems is the selection and design of natural "reserves": areas set aside for the preservation of natural values – including recreation and ecosystem services (e.g., supply of timber) – or for the protection of biodiversity (Margules Pressey, 2000). The reserves must be selected to fully represent and protect a variety of species over the long term by supporting viable population levels and eliminating threats both natural (coming from other species) and man-made (coming from commercial and development activities). The decisions related to location, size, and design of reserves must incorporate a wide variety of managerial considerations, competing objectives, and physical, economic and political constraints. For example, reserves must be located so they coincide with natural land features, like watersheds. They can also be designed to meet criteria for size, shape, connectivity, compactness, and species complementarity. This problem lends itself well to optimization methods including nonlinear programming, multiobjective optimization, and combinatorial optimization, which have been widely used to solve reserve selection and reserve design problems of increasing complexity. Taking into account multiple species, their survival and growth models, the type and extent of threats they face, and the economic consequences of any management action makes this problem even more interesting and challenging.

Another issue causing damage to biodiversity, especially in developed countries, is land fragmentation, or the division of species' habitats into smaller areas that are not connected to each other. Habitat fragmentation can be caused by new commercial development, housing, roads, highways, or railway lines, and can reduce a species' population, mobility, and genetic diversity. Fragmentation can be measured by multiple indicators. Two of the most popular indicators are the mean nearest neighbor distance (MNND) and the mean proximity index (MPI), which estimate the relative isolation of the parcels (Billionet, 2013). Minimizing these indicator values can mitigate the fragmentation in a given landscape. Another method to offset the effects of fragmented landscapes is to connect them through corridors, that is strips of land connecting larger, isolated parcels through which the species can move to migrate, reproduce, and escape. This extra mobility can help protect the species by supporting larger metapopulations (i.e., spatially separated populations). Various models for reducing land fragmentation have been proposed and can also be solved using optimization models and algorithms.

Another well-studied problem concerns the elimination and control of invasive species. Invasive animal and plant species can cause significant damage to biodiversity through predation, competition for resources, genetic disturbance, and epidemics (Billionet, 2013). Due to a lack of human and financial resources, eliminating an invasive species requires careful evaluation of alternative managerial interventions for optimal deployment of those limited resources. For any control action to bear fruit, the spread dynamics of an invasion need to be carefully considered. The scale, speeds, and vectors of an invasion are highly dependent on the species being considered and the specific physical context of each invasion. To make matters more complicated, the data about invasions and their spatial-temporal spread are sparse and difficult to procure. Thus, the control decisions rely on imperfect information and must be robust to various uncertainties to be truly applicable in real settings.

Some other problems related to conservation planning include long-term land use decisions, adverse effects caused by landscape fragmentation, rational use of forest resources, vegetation management, preservation of species' genetic diversity, wildfire control (Billionet, 2013), optimal deployment of resources to stem illegal poaching and smuggling activity, studying a reserve's resilience to climate change, and planning for long-term risks of climate change to a region's biodiversity (Eaton et al., 2019). Further research into using optimization techniques to solve many different forms of the aforementioned problems is duly warranted and deserves the immediate attention of serious operations research practitioners.

Zulqarnain Haider is a PhD candidate in the Industrial and Management Systems Engineering Department at University of South Florida.

References

1. Billionnet, A. (2013). Mathematical optimization ideas for biodiversity conservation. European Journal of Operational Research, 231(3), 514-534.
2. Eaton, M. J., Yurek, S., Haider, Z., Martin, J., Johnson, F. A., Udell, B. J., ... Kwon, C. (2019). Spatial conservation planning under uncertainty: adapting to climate change risks using modern portfolio theory. Ecological Applications, doi: 10.1002/eap.1962.
3. IUCN [International Union for Conservation of Nature] (2019). Background History. IUCN Red List of Threatened Species. https://www.iucnredlist.org/about/background-history.
4. Margules, C. R., Pressey, R. L. (2000). Systematic conservation planning. Nature, 405(6783), 243-253.
5. Polasky, S., Nelson, E., Camm, J., Csuti, B., Fackler, P., Lonsdorf, E., ... Haight, R. (2008). Where to put things? Spatial land management to sustain biodiversity and economic returns. Biological Conservation, 141(6), 1505-1524.

The Fast and the Furious: Market Design and Operations Research

Sepehr Ramyar
STAFF EDITOR

HAVE you ever lined up in front of a store for several hours on a Black Friday? Have you ever wondered why professional Wall Street (high-frequency) traders spend billions of dollars on IT infrastructure just to get their bids processed before everyone else. The reason for both these situations is fundamentally no different. In situations like these, there is value in being the first one in line. This is the same reason that prompted the 'Sooners' in the early 19th century to jump the gun in the race for registering land claims in Oklahoma. But what can you do when you have so many eager participants and limited resources? Or how do you optimize the performance of a system in presence of different constraints? Beginning to sound like operations research jargon?

It turns out that operations research concepts make up the machinery behind many of our everyday market transactions. From calling Uber on your smartphone, to the ads displayed on the webpages you visit every day, or even in decisions for high school admissions in many urban areas.

In all such marketplaces, the market designer aims to achieve the best performance as fast as possible. The definition of performance, however, depends on the context. For instance, in the case of online ads, the best performance means choosing the most valuable ads (i.e., those with the highest willingness-to-pay) while for high school admissions, the mutual satisfaction of students and schools with the allocations probably makes the most sense. But how can this be achieved?

Oklahoma Land Rush. Sourced from www.soonersports.com.

The primary objective of a market designer is to process the orders of participants as fast and safely as possible. Speed is important because it enables participation of more customers in the market. This is why Amazon and Airbnb have been able to obtain a larger pool of customers than traditional retailers and hotels: they provide a faster means of browsing through items and choosing the one you like most. Safety is also important. It means the customer (or the seller) would not have to worry about the other party 'gaming' the system. Here's where operations research comes in.

Operations research can help us design rules for marketplaces to make them safe for everyone to participate. In the online ad market, as mentioned earlier, the desired outcome is to have each ad slot allocated to the highest bidder. But does that mean choosing the highest bidder and charging them what they bid? This is in fact the mechanism that was in place in the early days of online ads and led to 'bidding wars' where everyone would bid their least affordable price and minimally increasing them with competing bids. This was not safe for the market. First, it strained the market operator's system as competitors would continuously place bids in very short intervals. Second, it did not guarantee whether the winner would be paying equivalent of their true value for the item. The solution came from operations research: an auction where the highest bidder wins but pays the second-highest bid also known as 'Second-Price' auction. The second price auction guarantees 'truth-telling' in the sense that no participant would have an incentive to misreport their value for the item by overbidding or underbidding. The introduction of generalized second price auctions significantly improved the performance in online ad markets that are considered the cash cows of giant tech companies like Google. Later, around 2008, these auctions were optimized with 'reserve' prices and further boosted online ad revenues (Wurman et al., 2001; Roughgarden, 2016).

The idea of second-best choice, however, is not always the solution. In fact, for school admissions, a source of instability of the market has chronically been students and families misreporting their list of preferred schools in hopes of maximizing their chances of getting into the second-best school; because they figure listing their favorite school (which is too hard to get into) on top of the list would lower their chances of getting into the second-best school, so they list their second-best as their first choice. This leads to undesired market outcomes with some schools ending up with empty seats and some others withholding their capacity and assigning their seats based on other criteria. The solution, again, comes from game theory and operations research community: the well-known deferred acceptance (DA) algorithm. In this algorithm, students and schools both list their choices from most to least preferred and then submit it to a 'clearing house'. The market clearing house would then match the most preferred student (by the school) to the most preferred school (by the student) given the capacity of the school. The result is astonishing: no assigned student would want to go to another school (with available capacity) and no school with full capacity would want to exchange one of its students (because that student would prefer to stay in the same school). Here, too, the mechanism is 'incentive-compatible' (or truthful) in the sense that students would have no incentive to misreport their list of preferred school. This algorithm was originally proposed by Lloyd Shapley and David Gale for a hypothetical 'marriage market' in 1962 and has since been improved and built on for many applications (Roughgarden, 2016).

Almost any market these days is run on algorithms that facilitate participation and improve efficiency of the marketplace. From kidney exchanges (Priority Pairwise Kidney Exchange algorithm) to stock exchange (Double Auction), these algorithms help operate markets in a fast and secure way that scales to the level of computation and performance required for today's marketplaces with millions of participants and billions of transactions every day; and operations research is at the heart of what makes all these markets work.

Note: This article has largely been inspired by the book “Who Gets What - and Why” by Alvin Roth, a Nobel laureate in Economics that got his Ph.D. in ‘operations research’. This is a great read for anyone who is interested in learning about market design and matching markets.

Sepehr Ramyar is a PhD student in Technology and Information Management at the University of California, Santa Cruz.

References

1. Roughgarden, T. Twenty lectures on algorithmic game theory. Cambridge University Press, 2016.
2. Wurman, P.R., Wellman, M.P., and W.E. Walsh. A parametrization of the auction design space. Games and economic behavior 35.1-2 (2001): 304-338.

YinzOR: A Journey for the CMU INFORMS Student Chapter

Sagnik Das

YinzOR is a single-track conference that brings together students studying operations research and related fields to facilitate interaction and collaboration with peers. YinzOR came into existence in 2017, when Aleksandr Kazachkov, founding President of the Carnegie Mellon University (CMU) INFORMS Student Chapter, envisioned a “conference of the students, by the students and for the students”. Thiago Serra, the President of our Chapter in 2017, coined the name YinzOR, with “Yinz” being the local Pittsburghese version of you-ones or yous-ones (2nd person plural pronoun), which derives from the Scots language, sometimes called Lowland Scots, spoken in southern Scotland until the late 18th century; and “OR” stands for operations research. More trivia about “yinz” is in our homepage of YinzOR 2019, courtesy of Professor John Hooker. Christian Tjandraatmadja designed the awesome logo for YinzOR, featuring a bridge that symbolizes Pittsburgh, also affectionately known as “The City of Bridges.”

True to the original YinzOR vision, we have hosted high quality conferences centered around PhD students for the past two years. The featured speakers at YinzOR are professionals who have recently graduated from their academic program and joined academia or industry. For invited talks, our goal is to feature PhD students from diverse OR/MS backgrounds. A nomination and voting process is used by the organizing committee to choose and invite outstanding young researchers from various institutions.

The inaugural YinzOR was held in August 2017 at the Tepper School of Business, Carnegie Mellon University (CMU). YinzOR 2017 was chaired by Aleksandr Kazachkov, who was assisted by a team of 11 members. We had the pleasure of hosting as our featured speakers Noam Brown, PhD student from CMU Computer Science Department (now at Facebook AI Research); Hamsa Bastani, Postdoc at IBM Research (now Assistant Professor at Wharton) and Francisco Trespalacios from Exxon Mobil. More than 60 students from CMU, University of Pittsburgh (UPitt), and Lehigh University participated and enjoyed valuable networking opportunities with fellow students. More details about the 2017 conference can be found on our conference website.

YinzOR 2017 Group Photo

Last year marked the second time that CMU INFORMS Student Chapter organized YinzOR. YinzOR 2018 was chaired by Neda Mirzaeian, who was assisted by a committee of 15 members. The conference, sponsored by EQT Corporation and Tepper School of Business, consisted of three featured talks, eight regular-track talks, a 12-person poster competition, as well as a few interactive coffee breaks and a happy hour. During YinzOR 2018, we had the privilege of hosting more than 70 attendees. It was an honor to host invited student speakers from Wharton, University of Michigan, and Lehigh University, in addition to local invited student speakers from CMU and UPitt. The featured speakers were Can Zhang, Assistant Professor at Fuqua School of Business at Duke University, Miles Lubin from Google Research, and Markus Drouven from EQT Corporation.

A well-received addition to YinzOR 2018 was the poster competition open to all PhD students in related fields. The panel of judges for the poster competition consisted of four CMU faculty members: Gerard Cornuejols, John Hooker, Fatma Kilinc-Karzan, and R. Ravi. In a tough competition, Po-Wei Wang won first prize. YinzOR 2018 ended with the most important event of all: Happy Hour. It was held in the biggest balcony of the Tepper Quad, where the participants enjoyed a nice view of the beautiful city of Pittsburgh, and discussed ORMS (and life!). More details about YinzOR 2018 can be found on our conference website and from this detailed blog written by Neda Mirzaeian.

YinzOR 2018 Group Photo

Happy hour during YinzOR 2018

I am personally very excited to share my experiences with the preparation for YinzOR 2019, which took place on August 23rd and 24th at the new Tepper Building, Tepper Quad. I am part of the YinzOR 2019 Organizing Committee chaired by Violet Chen and Ozgun Elci. Our featured speakers this year are Thiago Serra, who recently joined Bucknell University as an Assistant Professor, and Joann de Zegher and Daniel Freund, who both recently joined MIT Sloan as Assistant Professors. The invited speakers are from CMU Tepper, Chicago Booth, Columbia, Cornell, Georgia Tech and MIT Operations Research Center. For our poster competition this year, we have submissions from students from CMU, Georgia Tech, Virginia Tech, Purdue, Johns Hopkins, Lehigh, MIT, Cornell, Polytechnique Montreal, and University of Illinois. This year, we are also introducing another fun event: Flash Talks, where participants will have to explain their work without using "taboo words." The taboo words will be selected by the event coordinator based on the most frequently used words in the abstract.

Our preparations for YinzOR 2019 started on a high note as we were able to secure substantial sponsorship (around four times higher than last year) with FedEx and McKinsey Company as our Silver Sponsors and Simio, Tepper School of Business, and the OR Department at Tepper as our Bronze Sponsors. The external sponsorship was made possible by the support from our faculty members Sridhar Tayur and Willem-Jan van Hoeve. This additional funding played a role in attracting participation from more major universities on the East Coast. Like previous years, we are giving full travel support and accommodation to all featured speakers. For invited speakers, in addition to partial travel support, this year we are providing accommodation. Furthermore, we are providing accommodation to all the poster competition and Flash Talk participants who reside outside of Pittsburgh.

As YinzOR is expanding, we are actively seeking to outsource as much logistic work as possible to make planning the event smoother. In the past years, most of the logistics of the conference were handled by the organizing committee. This year, we will have catering services take care of the coffee breaks. Our chapter's vision is to completely outsource all logistics in future editions of YinzOR, which will enable the organizing committee to focus mainly on improving the quality and diversity of invited talks and poster and Flash Talk submissions, and market the conference to a larger base.

Please visit our conference website to learn more about YinzOR and our student chapter website to learn more about the CMU INFORMS Student Chapter!

Sagnik Das is a PhD Student in Carnegie Mellon University's Tepper School of Business and Vice President of the CMU INFORMS Student Chapter.