SIAG/OPT Views-and-News

A Forum for the SIAM Activity Group on Optimization

Volume 20 Number 1 December 2009

Contents

The Public Side of Optimization

What do school bus schedules and facebook in common? Both are aspects of our social life that can be modeled as optimization problems, and are the topic of the two exciting articles in this issue.

The first article by Armin Fügenschuh shows how a straightforward looking modeling challenge to optimize the school bus schedules can lead to public anger, and even a response from ambitious politicians. Anyone who ever thought that mathematics is apolitical is strongly encouraged to enjoy Armin's article!

The second article by Cody Bredendick and Michael Ferris considers a facebook app to construct a friend wheel: a graph that provides a visual representation of of the relationships between the friends of any given person.

Scheduling School Starting Times and Public Buses

Armin Fügenschuh
Konrad-Zuse-Zentrum für Informationstechnik (ZIB)
Takustraße 7, D-14195 Berlin, Germany
(`fuegenschuh@zib.de`).

1. A real-world problem

Lisa uses public bus transport services to get to her school. At seven o'clock on school days Lisa waits at a bus stop near her home. From the next term on she can stay in bed a little longer, because her school will start at 8:30 instead of 7:50. But she isn't really happy about it: "Then I'll also return home much later. On most days I'll be at home at 2 p.m. So I'll have to quit my sport's club that starts at half past three." Lisa's parents are also unhappy. The primary school of Lisa's younger brother is going to start already at 7:30. Only at the high school of Lisa's older sister the starting time will remain at 8 o'clock. "We can forget about our common lunch. The micro-wave oven will be our family's best friend", fears Lisa's mother.

The situation described in the introduction is typical for many German counties. Traditionally schools start with their first lesson around 8 o'clock. In rural, sparsely populated areas the majority of pupils come to school by bus. They are the largest individual group of customers. For the transport of pupils there are many regulations that give restrictions on the maximum travel time and waiting time at the school. The bus service is therefore strongly oriented towards their special needs. This requirement in combination with all schools starting more or less at the same time leads to a morning peak in the number of deployed buses. During the short period between 7 and 8 o'clock, before the start of the schools, all buses are used. During the rest of the day they stand unused in the depot. But traffic peaks are peaks in cost. These expenses are covered by the county authorities which are by law responsible for the transport of pupils. Since public money is a scarce resource these days, a new idea for potential savings was needed. Of course it is always possible to achieve savings by reducing the quality-of-service, but operating the system with fewer and hence more crowded buses or with longer travel and waiting times is politically undesired.

According to Bussieck, Winter, and Zimmermann [5], the planning of a public transport system is done in several steps. First, one has to plan the **network**: from where to where and at what times do customers want to travel? This results in origin-destination-matrices. Second, the **lines** are planned: where do the buses drive, and where do customers have to transfer (change the line) to reach their destination. Third, the **trips** are assigned to the lines: how often and at what time is a line served? These three steps are on the strategic level, which is visible to the customers, and which defines the quality-of-service. The next steps are on the operative level, usually invisible for the customer. The **schedules** of the buses determine which bus is serving which trip in which order. Finally **drivers** are assigned to schedules, where breaks have the be planned and other working rule limitations have to be taken into consideration. Usually these steps are carried out sequentially, meaning that the output of one step serves as an input for the next step. There exists a limited way for feedbacks in the opposite direction, if the decisions made in an earlier planning stage lead to undesired results on a later stage. Recently integrated approaches have been proposed, for example simultaneous bus and duty scheduling [4, 17].

The central idea of our planning approach "IKOSANA"[1] is the simultaneous planning of school starting times and the schedules of the buses without reducing the quality-of-service. The state's school law allows school starting times to vary between 7:30 and 8:30 a.m. As long as all schools start at 8 o'clock, one bus can only serve one school. If some of the schools start earlier and other change to a later start, then the starting times of the trips have to be altered accordingly, and the bus has enough time to serve a second, late-starting school on a second tour. In this way one bus could be saved (c.f. Figure 1). Note that only the starting times of the trips are changed, but not the routes: the lines remain the same as before, the same number of pupils will be in the bus, and they will change the trip at the same transfer stops, if necessary. Hence the quality-of-service stays the same. In fact, the saved buses can be re-invested to serve either more trips, or to give money to schools as an incentive to alter their starting times.

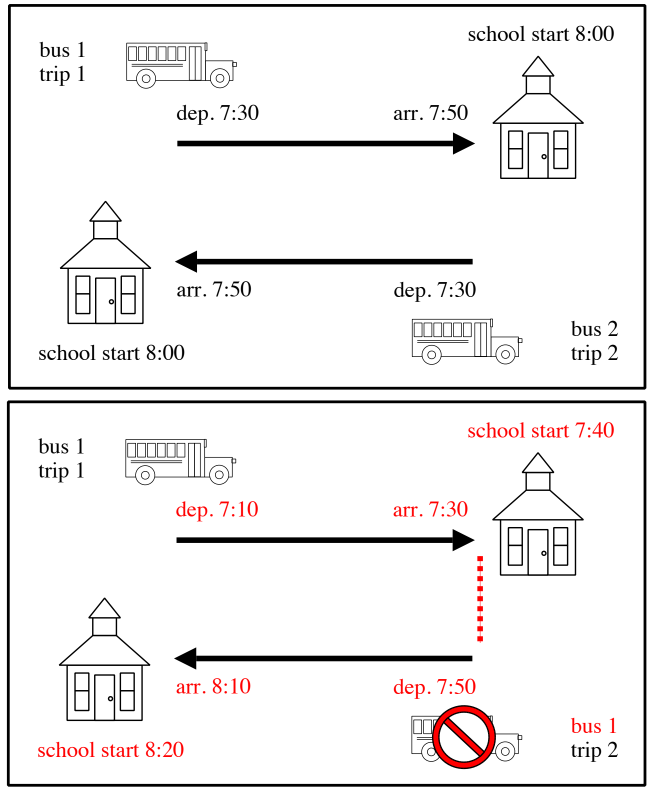

Figure 1: The central idea.

So much for the idea, but how to roll out the planning in a county with up to 150 schools and 300 to 600 trips in the morning rush hours? How to decide which schools will start earlier and which later? In 2002 Alexander Martin, professor for discrete optimization at Darmstadt University of Technology and his PhD student (the author) met traffic consultant Peter Stöveken at the Heureka conference is Karlsruhe, where Stöveken gave a talk on this issue. After the talk we asked him about the how-to, and we found out that the planning was carried out manually. The only help came from a computer

[1] In German: Integrierte Koordinierung von Schulanfangszeiten und des Nahverkehrsangebots.

graphics program, where the bus trips where drawn on a scanned map of the county in order to get at least some visual impression of what is going on. It took about one week to find a new plan, and if the plan has to be changed, another week of re-planning is necessary. Interestingly, Stöveken's consultant company was paid proportional to the savings they could achieve in the end. They accompany the whole conversion process, which lasts up to one year per county, and during this time, re-planning becomes necessary because restrictions emerge in the course of this process. It is no wonder that Stöveken was very interested to obtain an automated planning tool, and we convinced him that the problem behind IKOSANA is in fact a mathematical optimization problem. A collaboration with Stöveken's company was established. Three years later in 2005 we presented our joint results at the next Heureka conference [11], and the author successfully defended his PhD thesis on this subject [7].

2. From practice to theory...

In the sequel we give a formulation of the planning problem behind IKOSANA as a mixed-integer linear program. The basic building blocks of the buses' schedules are the trips. A trip is a given sequence of bus stops with known travel times between the stops. The set of trips is denoted by $\mathcal{T}$. Each bus starts and ends at the depot, and serves one or more trips. Deadheading between trips is allowed, that is, if two consecutive trips in a schedule end and start at different stops, then the bus drives without passengers from the last stop of the previous trip to the first stop of the next trip. The set of potential deadhead trips is denoted by $\mathcal{D} \subset \mathcal{T} \times \mathcal{T}$. The schedules are represented by the following binary variables:

$$\forall\, i \in \mathcal{T} : v_i, w_i \in \{0,1\}, \tag{1a}$$

$$\forall\, (i,j) \in \mathcal{D} : x_{i,j} \in \{0,1\}. \tag{1b}$$

If $v_i = 1$ in a solution then trip i is the first in the schedule, i.e., a new bus starts at the depot to serve i. If $w_i = 1$ then trip i is the last in the schedule, that is, after having served i the bus drives back to the depot. If $x_{i,j} = 1$ then trip j follows directly after trip i in some schedule. Each trip has a unique precessor (or the depot):

$$\forall\, i \in \mathcal{T} : v_i + \sum_{j:(j,i)\in\mathcal{D}} x_{j,i} = 1, \tag{2}$$

and a unique successor (or the depot)

$$\forall\, i \in \mathcal{T} : \sum_{j:(i,j)\in\mathcal{D}} x_{i,j} + w_i = 1. \tag{3}$$

For the trips we introduce variables with integrality constraints

$$\forall\, i \in \mathcal{T} : \alpha_i \in \mathbb{Z} \tag{4}$$

for their starting times (i.e., departure of the bus at the first bus stop of the trip). The starting times of the trips is given as "minutes after midnight". For example, 6:42 as trip starting time translates to $\alpha_t = 402$. Trip starting times can be altered within some given bounds $\underline{\alpha}_i, \overline{\alpha}_i$ which are typically time intervals around the current starting times (for example, plus or minus 30 minutes):

$$\forall\, i \in \mathcal{T} : \underline{\alpha}_i \le \alpha_i \le \overline{\alpha}_i. \tag{5}$$

The time for serving the entire trip i from the first to the last stop is a given parameter δ_i^{trip}, and the time for a deadhead trip between trips i and j is given by $\delta_{i,j}^{\text{dhd}}$. If two trips are combined within a schedule then their starting times have to be synchronized by the following constraints:

$$\begin{aligned} &\forall\, (i,j) \in \mathcal{D} : \\ &\quad \alpha_i + \delta_i^{\text{trip}} + \delta_{i,j}^{\text{dhd}} \le \alpha_j + M \cdot (1 - x_{i,j}), \end{aligned} \tag{6}$$

where M is a sufficiently large constant. The starting times of the trips cannot be changed independent from each other. If currently pupils (or other customers) need to transfer from "feeder" trip i to "collector" trip j then this transfer must still be possible after changing the starting times of these trips. To formally express this condition we introduce some more notation: Let $\delta_{i,j}^{\text{feed}}$ be the time for the feeder trip from the first bus stop of trip i to the transfer bus stop, let $\delta_{i,j}^{\text{coll}}$ be the driving time for the collector trip from the first bus stop of trip j to the transfer bus stop. Given are lower and upper bounds on the waiting times for the passengers at the transfer bus stop while waiting for the collector bus: $\underline{\omega}_{i,j}^{\text{trnsfr}}$ and $\overline{\omega}_{i,j}^{\text{trnsfr}}$. Denote by $\mathcal{C} \subset \mathcal{T} \times \mathcal{T}$ the pairs of trips where a transfer relation exists. Then the synchronization of the two trips is assured by these constraints:

$$\forall\,(i,j) \in \mathcal{C}: \qquad \alpha_i + \delta_{i,j}^{\text{feed}} + \underline{\omega}_{i,j}^{\text{trnsfr}} \leq \alpha_j + \delta_{i,j}^{\text{coll}}, \tag{7a}$$

$$\forall\,(i,j) \in \mathcal{C}: \qquad \alpha_i + \delta_{i,j}^{\text{feed}} + \overline{\omega}_{i,j}^{\text{trnsfr}} \geq \alpha_j + \delta_{i,j}^{\text{coll}}. \tag{7b}$$

For each school s we introduce an integer variable

$$\forall\, s \in \mathcal{S} : \tau_s \in \mathbb{Z} \tag{8}$$

for its planned starting time. Here $\mathcal{S}$ denotes the set of all schools. The school starting time can be altered within some given bounds:

$$\forall\, s \in \mathcal{S} : \underline{\tau}_s \leq 5\tau_s \leq \overline{\tau}_s. \tag{9}$$

The value "5" in this equation indicates that the school starting time must be aligned to 5 minutes. That means, a school can start at 7:40 or 7:45, but not at 7:42. Thus to encode a starting time interval such as 7:30 - 8:30 we have to count the number of 5-minute-timeslots after midnight and set $\underline{\tau}_s := 90$ and $\overline{\tau}_s := 102$. Note that the new school starting time is then given by $5\tau_s$ (and not τ_s). The set $\mathcal{P} \subset \mathcal{T} \times \mathcal{S}$ consists of those pairs (i, s), where trip i transports pupils to school s. The pupils get off the bus at the stop that is closest to their school. The travel time for the bus from the first stop of the trip to this school stop is denoted by $\delta_{i,s}^{\text{school}}$. When pupils leave the bus they need a certain time to reach the classrooms. This time is given by $\underline{\omega}_{i,s}^{\text{school}}$. On the other hand pupils should not wait too long at their school, hence an upper waiting time limit $\overline{\omega}_{i,s}^{\text{school}}$ is also specified. Then the following constraints have to hold:

$$\forall\,(i,s) \in \mathcal{P}: \qquad \alpha_i + \delta_{i,s}^{\text{school}} + \underline{\omega}_{i,s}^{\text{school}} \leq 5\tau_{i,s}, \tag{10a}$$

$$\forall\,(i,s) \in \mathcal{P}: \qquad \alpha_i + \delta_{i,s}^{\text{school}} + \overline{\omega}_{i,s}^{\text{school}} \geq 5\tau_{i,s}. \tag{10b}$$

The most important goal is to save as many buses as possible. On a subordinate level the deadhead trips in total should be as short as possible. An objective function reflecting these two goals is

$$\sum_{i \in \mathcal{T}} C v_i + \sum_{(i,j) \in \mathcal{D}} \delta_{i,j}^{\text{dhd}} x_{i,j}, \tag{11}$$

where C ($\gg \delta_{i,j}^{\text{dhd}}$) stands for the costs per bus. Summing it up, the bi-criteria formulation of IKOSANA is the following MILP:

$$\min\ (11),\ \text{s.t.}\,(1)-(10). \tag{12}$$

In [10] we extend this bicriteria objective to a multicriteria formulation of IKOSANA, that takes the (minimization of) waiting times of the pupils at transfer stops and at their school explicitly into account.

When analyzing (12) several interesting substructures can be identified. First, imagine that there are no schools and no transfers, that is, $\mathcal{P} = \mathcal{C} = \emptyset$. Then (12) reduces to

$$\min\ (11),\ \text{s.t.}\,(1)-(6), \tag{13}$$

which is known as the vehicle scheduling problem with time windows (VSP-TW) in the literature. If the starting times are given and cannot be changed then the remaining problem

$$\min\ (11),\ \text{s.t.}\,(1)-(3), \tag{14}$$

is called VSP [3]. Vice versa IKOSANA (12) can be considered as VSP-TW with additional *coupling* aspects among the time windows, for which we suggested the term VSP-CTW in [8] ([6] suggest to call them *sliding* time windows). Finally, an interesting subproblem occurs when all binary variables are given. Then the remaining problem is to determine feasible starting times for the trips and the schools:

$$\text{s.t.}\,(4)-(10). \tag{15}$$

This system is called IP2, because in each constraint at most two variables occur. Even IP2 are (theoretically) difficult to solve, and several articles are devoted to their study [12, 13, 15, 16]. The most recent result is a pseudopolynomial time algorithm that finds a feasible solution in $O(mU)$, where m is the number of constraints and U is the maximum range between the lower and the upper bound on each of the variables [1].

We note that other researchers have addressed the problem of changing school starting times in order to find schedules with fewer buses [2, 3, 6]. In their work the (new) starting times are computed beforehand, and then the schedules are constructed accordingly. We are, however, not aware of an integrated model such as (12).

For a solution of (12) we implemented several heuristic methods (such as greedy and parameterized greedy construction heuristics with local search enhancement) that are able to find feasible solutions in short time and methods that derive a dual (lower) bound on the objective function value (such as branch-and-cut and column generation). For the details we refer to the literature [7, 9].

3. ...and back into practice

At a certain point within the development of the IKOSANA solution algorithms, it was time to implement these solutions in practice, and to test how good they are under real-world conditions. The first lesson to be learned for a scholar is this: We might develop certificates of optimality as much as we want, but no one in the outside world really cares for them. It simply isn't a strong selling argument.

A typical consulting project has the following steps: data collection from the bus companies, data conversion into a certain XML schema, error correction, and finally solving for the first time. To reach this point requires about three to four months of work, mainly idle time while waiting for the companies to provide data and to correct errors in the input data. After that the solution is handed back to the bus companies who will most likely detect further errors or missing information that render a solution infeasible. In one or more further loops all these mistakes will be corrected, and by a further run of IKOSANA a new solution is computed, that again has to be checked by the companies. This easily takes another three months. Six months after the start of the project we typically have a solution that is feasible from the bus companies' point-of-view. But still it is assumed that all schools involved change their starting times according to our plan.

Schools and parents are informed via public town-hall meetings. At one of these events a journalist was present, taking notes for her newspaper [14]. In the solution that was presented two high schools had to change their start from 8:00 to 8:20, but no one, at least no one present at that particular meeting, was happy with this. To give an impression of the bad mood at that event, we give some quotes from parents and teacher: "The concept is brilliant, but we don't want it." - "The families have to pay a high price for it." - What can an honest scientist reply to such killer arguments? Pointing out proven global optimality is not an option. And it goes on like this. When the kids of a family return at different times, no common lunch will be possible anymore. Or, as a parent said, "Then the microwave oven will be families best friend." Or parents say: "The discrimination of pupils in rural areas - for example in long distances for visits to a museum - must not be increased." A typical argument of teachers is: "As a consequence of this concept, pupils will be taught in the 6th [last] lesson in their biological drop in performance." The dean of a school said: "In no company someone from the outside can meddle in, except the schools. We are the experts for pedagogical affairs. Savings are necessary, of course. But for this, our children are the wrong targets." From that quote one might get the impression that this school dean tries to protect his children from harm. But it could also be the case that he is in a different political party than the current head of county authorities ("Landrat" in German), and now, facing elections, he wants to become the next Landrat.

From these quotes you get an impression how political and hence difficult the task of changing school starting times is. Having IKOSANA as an automatic planning tool helps, though. If some school is not willing to adjust their starting times accordingly, a new solution that respects their wish can be computed quickly. Than either this new solution is not much worse than the previous one (an alternative optimal solution). More likely, imposing new bounds will push the solution closer to the status-quo, and thus increase the number of buses again. Now, this increase can be linked directly to the unwillingness of some stakeholder to cooperate, and in this way can be used as a strong argument against them.

Despite all political difficulties the author has not lost his interest in consulting bus companies and county authorities in order to coordinate their school starting times and public transportation system. Since the end of his PhD in 2005 he successfully consulted four counties in Germany where savings of 5-10 buses were achieved, which corresponds to a cost reduction of 8-15% (or 200-400,000 Euros). From there let's do an extrapolation: Germany is devided into about 330 counties. It is safe to assume that about 90% of counties have uncoordinated school starting times. If on average 5 buses can be saved in each county, and each bus has fix costs of 40,000 Euro per year, then these figures add up to 60 million Euros, which are saved not only once, but year after year. *What are we waiting for?*

REFERENCES

[1] R. Bar-Yehuda and D. Rawitz. Efficient algorithms for integer programs with two variables per constraint. *Algorithmica*, 29(4):595 – 609, 2001.

[2] L. Bodin and L. Berman. Routing and scheduling of school buses by computer. *Transportation Science*, 13(2):113 – 129, 1979.

[3] L. D. Bodin, B. L. Golden, A. A. Assad, and M. O. Ball. Routing and scheduling of vehicles and crews. *Computers and Operations Research*, 10(2):63 – 211, 1983.

[4] R. Borndörfer, A. Löbel, and S. Weider. A bundle method for integrated multi-depot vehicle and duty scheduling in public transit. In *Proceedings of the Conference on Computer-aided Systems in Public Transport (CASPT 2005)*, pages 3 – 24, 2008.

[5] M. R. Bussieck, T. Winter, and U. T. Zimmermann. Discrete optimization in public rail transport. *Mathematical Programming*, 79:415 – 444, 1997.

[6] J. A. Ferland and L. Fortin. Vehicles scheduling with sliding time windows. *European Journal of Operational Research*, 38(2):213 – 226, 1989.

[7] A. Fügenschuh. *The Integrated Optimization of School Starting Times and Public Transport.* PhD thesis, Technische Universität Darmstadt, 2005.

[8] A. Fügenschuh. The vehicle routing problem with coupled time windows. *Central European Journal of Operations Research*, 14(2), 2006.

[9] A. Fügenschuh. Solving a school bus scheduling problem with integer programming. *European Journal of Operations Research*, 193(3):867 – 884, 2009.

[10] A. Fügenschuh and A. Martin. A multicriteria approach for optimizing bus schedules and school starting times. *Annals of Operations Research*, 147(1):199 – 216, 2006.

[11] A. Fügenschuh and P. Stöveken. Integrierte Optimierung der Schulanfangszeit und des Nahverkehrs-Angebots. In *Tagungsband der Heureka'05 – Optimierung in Verkehr und Transport*, pages 265 – 278, Köln, 2005. Forschungsgesellschaft für Straßen- und Verkehrswesen.

[12] D. S. Hochbaum, N. Megiddo, J. Naor, and A. Tamir. Tight bounds and 2-approximation algorithms for integer programs with two variables per inequality. *Mathematical Programming*, 62:69 – 83, 1993.

[13] D. S. Hochbaum and J. Naor. Simple and fast algorithms for linear and integer programs with two variables per inequality. *SIAM Journal on Computing*, 23(6):1179 – 1192, 1994.

[14] U. Holtgrewe. Lottes Eltern, Lehrer und Parteien gegen RVM-Pläne. Neue Osnabrücker Zeitung, March 13 2003.

[15] J. C. Lagarias. The computational complexity of simultaneous diophantine approximation problems. *SIAM Journal on Computing*, 14:196 – 209, 1985.

[16] V. R. Pratt. Two easy theories whose combination is hard. Technical report, Massachusetts Institute of Technology. Cambridge, Massachusetts, 1977.

[17] I. Steinzen, V. Gintner, L. Suhl, and N. Kliewer. A time-space network approach for the integrated vehicle and crew scheduling problem with multiple depots. *Transportation Science*, 2009. To appear.

Facebook Friend Wheels and Quadratic Assignment Problems

Cody Bredendick and Michael C. Ferris[1]
Computer Sciences Department,
University of Wisconsin, Madison,
Wisconsin.
(ferris@cs.wisc.edu, bredendi@cs.wisc.edu).

Abstract

A friend wheel is an image used on Facebook to provide a visual representation of the relationships between the friends any one person may have. We outline a tool, based on the solution of quadratic assignment problems, that generates such a wheel. The tool uses the Facebook API to gather data, the GRASP heuristic to solve the model, and GraphViz to generate the visualization.

1. Introduction

As is the case at many universities, we teach a senior level undergraduate course in optimization tools and modeling at the University of Wisconsin. The course covers a wide range of optimization topics, but concentrates on the types of models that can be easily formulated within a modeling system such as GAMS or AMPL, and attempts to provide the students with a set of techniques that can be readily adapted to a range of different application settings. As is typical for such courses, the content involves a large number of examples and modeling exercises, coupled with an overview of the perceived and observed strengths and weaknesses of different modeling formats. In order to stress the difficulties associated with data collection and processing, the course requires an individual final project that is defined by the student, utilizing the skills they have been developing during the semester and experimenting with the issues of data acquisition, refinement and usage. To ensure the project is well defined and appropriate, a single page proposal is required from the students earlier in the semester that forms the basis of an oral interview to modify the scope or content of the proposal. The final projects include a (stricly enforced) 4 page summary, but students are allowed to provide additional electronic material (including models and data) that may or may not be exercised by the instructor during evaluation. We believe this process works extremely well, and the students (and instructor) learn lots about modeling and optimization in practice. It can also lead to significant research opportunities as such projects can lead to new avenues of exploration initiated by the industry and bright ideas of interested and motivated students. Specifically, Ferris' work on radiation therapy was initiated by a project in an offering of this class carried out by David Shepard [8].

This short note describes the results of one of these projects, carried out by Cody Bredendick while he was enrolled in this class in 2009, to visualize "friend data" within Facebook. The results of other projects from this one class were also impressive, and included an investigation by Joonhoon Kim of some ideas for biological pathway models, utilizing a bilevel program to implement "Opt knock" [2]:

max	bioengineering objective		
	(through gene knockouts)		
s.t.	max cellular objective (over fluxes)		
	s.t.	fixed substrate uptake	
		network stoichiometry	
		blocked reactions (from outer problem)	
	number of knockouts $\leq$ limit		

The goal of the project was to extend a model that accounts for "gene" deletions and regulatory interactions, and solve the resulting problem for different biochemical productions to identify diverse metabolic engineering strategies. The specific model included about 30,000 constraints and 20,000 variables (10,000 binary).

A second project by Brandon Smith concerned video stabilization, which was motivated by the observation that videos shot with a hand-held conventional video camera often appear remarkably shaky. Such camera shake is one of the biggest components of the large quality difference between home movies and professional videos, and Smith observed that while several commercial software packages were available for digitally removing or reducing camera shake as a post processing step, the project would provide an opaque, open source version of such software. Underlying this project was a nonlinear least

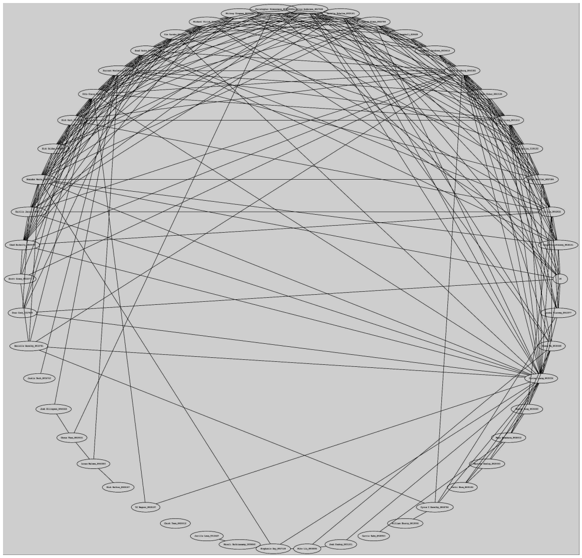

Figure 1: A friend wheel

squares problem:

$$\min_{\{H_f\}_{f=1}^{F}} \sum_{f=2}^{F-1} \sum_{j\in\phi_f} \left\| u_{f,j} - \frac{1}{2}(u_{f-1,j_{prev}} + u_{f+1,j_{next}}) \right\|^2 + \Phi(\{H_f\}_{f=1}^{F})$$

where

$$\Phi(\{H_f\}_{f=1}^{F}) = \sum_{f=2}^{F-1} \lambda_1 \theta_f^2 + \lambda_2 t_{x,f}^2 + \lambda_3 t_{y,f}^2 + \lambda_4 s_{x,f}^2 XS + \lambda_5 s_{y,f}^2$$

H is an unknown transform, parameterized by θ, t and s. Smith was able to solve this problem and demonstrated the results using actual video samples.

2. The Friend Wheel Project

By now, Facebook needs no real introduction. It is one of the largest and most visited websites in existence and has practically defined the modern definition of "social networking". Within this domain, many new features are constantly being developed. Facebook released a developers' API that allows apps to be developed by anyone. A friend wheel is one such app. A specific example is given as Figure 1. The typical size of such wheels involves from 50 to 1000 friends.

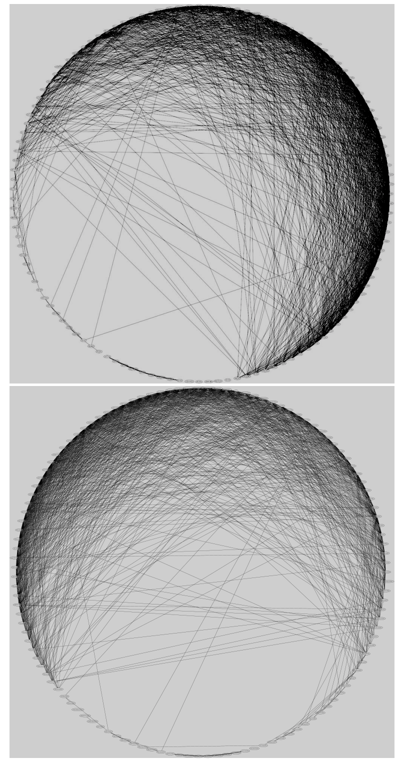

Figure 2: Two representations of a friend wheel using different orderings

A friend wheel is an image used to provide a visual representation of the relationships between the friends any given person may have. It is constructed by gathering each of the target individual's friends and placing them at uniform spacing on the circumference of a circle. Having done that, line segments are drawn between each point (each friend) if the people that those two connecting points represent are (announced) friends with each other. It is clear that ordering determines the resulting image, and that with a random ordering, one can expect a very dense mesh of lines uniformly distributed across the interior of the circle. Such friend wheels existed before this project was conceived: the inspiration of the particular project was to order the friends around the circumference of the circle in order to reduce amount of ink used in displaying the wheel.

In order to complete this project, three main steps were needed. Firstly, the required data needed to be gathered. This was accomplished by building a small *web service* to retrieve friend data using the Facebook API. The data is then manipulated into an appropriate format for the optimization solver using php scripts within Facebook.

The second step is to determine the ordering of friends around the circumference of the circle. We describe that below in Section 2.1.

The final step is to manipulate the optimization output for an image package to allow visualization of the image. The approach used here involved the (publically available) *graphviz* package, with two passes, one to fix the locations on the circle and draw the connecting arcs, and the second to generate the image in a nice format (for example png, jpg, pdf). Specific details on this are available from the authors on request.

2.1 Formulation and Solution via Quadratic Assignment

In order to minimize the ink that is needed to create a facebook wheel, we model the problem as a quadratic assignment problem (QAP). The QAP was first proposed in 1957 by Koopmans and Beckmann [3] as a mathematical model for facility location. Specifically, given n facilities $\{f_1, \ldots, f_n\}$, and n locations $\{l_1, \ldots, l_n\}$, the problem is to determine to which location each facility must be assigned.

Formally, if we let $W = (w_{i,j}) \in \mathbb{R}_+^{n \times n}$ where $w_{i,j}$ represents the flow between f_i and f_j, let $D = (d_{i,j}) \in \mathbb{R}_+^{n \times n}$ where $d_{i,j}$ represents the distance between l_i and l_j, and let $p : \{1, \ldots, n\} \mapsto \{1, \ldots, n\}$ be an assignment of facility f_i to location $l_{p(i)}$ whose cost is

$$c(p) = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} d_{p(i),p(j)}$$

then the problem of interest is:

$$\text{QAP} : \min c(p) \text{ subject to } p \in \Pi_n.$$

Here we use the notation Π_n to denote the set of all permutations of $\{1, \ldots, n\}$. It is known that QAP is strongly NP-hard [7]. A survey is found in [6], and more recent results are detailed in [4] and the references contained therein.

To formulate our minimizing ink problem as a QAP we let n represent the number of friends of a given individual, assign each friend a location on the circumference of the circle and define $w_{i,j} = 1$ if i is a friend of j, and 0 otherwise. The distance from location r on the circle circumference to location s (Euclidean or otherwise) is denoted by $d_{r,s}$. Thus, the only required input to define this problem is n and the links $w_{i,j}$ that represent that i is a friend of j. In this way, the data acquisition is very compact.

Given the instruction from the class, the first attempt was to model the problem in GAMS using binary variables $p_{i,r} = 1$ to indicate that facility f_i is located at l_r. The problem is a mixed integer quadratic program

$$\begin{array}{ll} \min & \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} \sum_{r} \sum_{s} d_{r,s} * p_{i,r} * p_{j,s} \\ \text{subject to} & \sum_{i} p_{i,r} = 1, \forall r \\ & \sum_{r} p_{i,r} = 1, \forall i \end{array}$$

Unfortunately, as is well known, even state-of-the-art commercial codes are unable to solve these formulations: the relaxed problem is not necessarily convex, and even tricks such as fixing one location to remove symmetry are ineffective. Similar results are found if additionally binary variables $q_{i,r,j,s} = (p_{i,r} \text{ and } p_{j,s})$ are defined using

$$\begin{array}{l} q_{i,r,j,s} \leq p_{i,r} \\ q_{i,r,j,s} \leq p_{j,s} \\ q_{i,r,j,s} \geq p_{i,r} + p_{j,s} - 1 \end{array}$$

to make the problem a mixed integer linear program. This is expected since it is well known that exact solution is very hard, even for some small ($n = 40$) problems [1]. More recent work, some using semidefinite programming is also referenced in [4]. While there are sophisticated approaches that may outperform the simple models we implemented, it is clear that the time frame and size of our models precluded exact solution.

Consequently, there are many heuristic solution stategies suggested for these problems. The survey [4] gives a more complete picture. For this project, we utilized a greedy randomized adaptive search procedure (GRASP) as given in [5]. A critical reason for this choice was the availability of the source code for download. While more details are available in the cited reference, the main idea is to generate starting solutions using a greedy heuristic, apply local search until some termination criteria satisfied, and optionally apply "path-relinking" procedures (a heuristic to combine to good solutions) to improve the overall quality of solutions generated.

Some very limited results are provided below to show the use of this heuristic compared to some exact methods.

	25		50		125	
solver -options	time	soln	time	soln	time	soln
dicopt	1.3	195	27.0	1752		
sbb	4.6	189	98.9	1742		
grasp	0.2	186	1.3	1660	16.7	25107
-npr					7.0	25143
-npr -n 2					14.0	25143
-npr -i 64					14.0	25143
-npr -i 10					2.2	25143
-npr -i 5					1.1	25398
-npr -s 250					7.1	25144

It is clear that the grasp code is much more effective than the exact codes shown here, and that specific options to that code facilitate quick solution of reasonable sized models. However, for models of larger size (including the example given in Figure 3, it is clear that there is need for application of other heuristic methods and these kinds of instances undoubtedly provide a rich class of test models for such approaches.

2.2 Details of the procedure

To try these approaches, the following steps are needed:

- Login to your facebook account
- Navigate to `apps.facebook.com/cscbfriend/`
- This page calls index.php on a remote "execution" websever
- Initiates a non-blocking call to start the data gathering process
- php preparedata.php [userID] [sessionID]
  - Communicates with facebook API
  - Generate a list of friends, and then a list of connections (using the Facebook Query Language)
  - Loops over data in order to generate names.txt, arcs.txt, qinput.txt

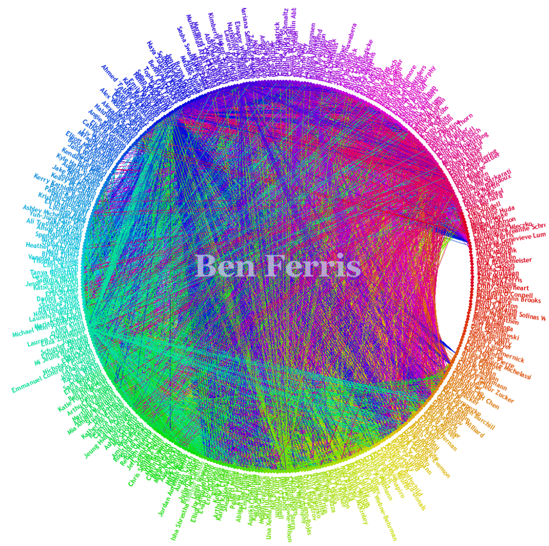

Figure 3: The friend wheel of Ben Ferris

  - Executes a shell script to run "grasp" and "graphviz"

Unfortunately, a timer program was needed to allow the page more than the default time to load. Details on this again are available from the authors.

The approach is one way to generate such friend wheels. As is typical with "apps" there are a number of other competitive versions of these wheels. One such application generated the output in Figure 3. While the output shown was slow to generate, and the solution represented there is not as good as the QAP solution that methods such as the grasp code provide, there is significant other functionality within the displayed output that make it more appealing to facebook users. Specifically, the output is colored and has dynamic facilities within it to allow a user to manipulate the wheel and investigate features of the representation and the social groupings that are clustered along the circumference. Future work aims to collaborate with the developer of this application to switch out the solver engine to use a more sophisticated optimization tool under the hood.

3. Conclusions

Optimization is a powerful tool whose applications are becoming more and more widespread due to increased complexity of systems and enormous growth in data collection and provision. This note detailed a specific application of one optimization formulation that can be used as a tool for visualization of specific data and is available at: `apps.facebook.com/cscbfriend/`

It is clear that optimization is critical in this application, but the underlying process of building such models teaches students that development time is frequently dominated by data handling issues. This note also provides an interesting source of new test programs for QAP, where the overall running time of the application is still dominated by the time to perform optimization.

Finally, it is clear that modern optimizers must have a range of skills including the ability to do rigorous modeling in multiple formats, the ability and flexibility to merge model solutions to perform hard underlying tasks, and computational skills to extract and manipulate data and display the results of optimization in clear and concise ways. It remains an exciting time to be an optimizer!

REFERENCES

[1] K. Anstreicher, N. Brixius, J.-P. Goux, and J. Linderoth. Solving large quadratic assignment problems on computational grids. *Mathematical Programming*, 91:563–588, 2002.

[2] A. P. Burgard, P. Pharkya, and C. D. Maranas. Optknock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnology and Bioengineering*, 84(6):647–657, 2003.

[3] T. C. Koopmans and M. J. Beckmann. Assignment problems and the location of economic activities. *Econometrica*, 25:53–76, 1957.

[4] E. M. Loiola, N. M. M. Abreu, P. O. Boaventura-Netto, P. Hanh, and T. Querido. A survey for the quadratic assignment problem. *European Journal of Operational Research*, 176:657–690, 2007.

[5] C.A. Oliveira, P.M. Pardalos, and M.G.C. Resende. GRASP with path-relinking for the QAP. In T. Ibaraki and Y. Yoshitomi, editors, *Proceedings of the Fifth Metaheuristics International Conference (MIC2003)*, pages 57–1 – 57–6, Kyoto, Japan, 2003.

[6] P. M. Pardalos, F. Rendl, and H. Wolkowicz. The quadratic assignment problem : A survey of recent developments. In P. M. Pardalos and H. Wolkowicz, editors, *Quadratic Assignment and Related Problems*, volume 16 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 1–42, Rhode Island, 1994. AMS.

[7] S. Sahni and T. Gonzales. P-complete approximation problems. *Journal of the Association for Computing Machinery*, 23:555–565, 1976.

[8] D. M. Shepard, M. C. Ferris, G. Olivera, and T. R. Mackie. Optimizing the delivery of radiation to cancer patients. *SIAM Review*, 41:721–744, 1999.

Bulletin

1. Erice Optimization Workshop

Ettore Majorana Centre for Scientific Culture
International School of Mathematics
"G. Stampacchia", Erice, Italy

52nd Workshop on Nonlinear Optimization, Variational Inequalities, and Equilibrium Problems

July 2 - 10, 2010

The Workshop aims to review and discuss recent advances in the development of analytical and computational tools for Nonlinear Optimization, Variational Inequalities and Equilibrium Problems, and to provide a forum for fruitful interactions in strictly related fields of research.

Topics include constrained and unconstrained nonlinear optimization, global optimization, derivative-free methods, nonsmooth optimization, nonlinear complementarity problems, variational inequalities, equilibrium problems, game theory, bilevel optimization, neural networks and support vector machines training, applications in engineering, economics, biology and other sciences.

The Workshop will include keynote lectures (1 hour) and contributed lectures (30 min.). Members of the international scientific community are invited to contribute a lecture describing their current research and applications. Acceptance will be decided by the Advisory Committee of the School.

Invited lecturers who have confirmed the participation are: Ernesto G. Birgin, Francisco Facchinei, Christodoulos A. Floudas, David Gao,Diethard Klatte, Eva K. Lee, Marco Locatelli, Jacqueline Morgan, Evgeni A.Nurminski, Jong-Shi Pang, Mike J. D. Powell, Franz Rendl, Nikolaos V. Sahinidis, Katya Scheinberg, Marco Sciandrone, Valeria Simoncini, Henry Wolkowicz, Ya-xiang Yuan.

A special issue of Computational Optimization and Applications will be dedicated to the Workshop, including a selection of invited and contributed lectures. Further information can be found at the URL: `http://www.dis.uniroma1.it/~erice2010` or requested to the e-mail address: `erice2010@dis.uniroma1.it`

The Scientific and Organizing Committee:

- Gianni Di Pillo, SAPIENZA – Universita' di Roma, Italy
- Franco Giannessi, Universita' di Pisa, Italy
- Massimo Roma, SAPIENZA – Universita' di Roma, Italy

2. New Book: Moments, Positive Polynomials and Their Applications by Jean B. Lasserre

Many important problems in global optimization, algebra, probability and statistics, applied mathematics, control theory, financial mathematics, inverse problems, etc. can be modeled as a particular instance of the Generalized Moment Problem (GMP). This book introduces, in a unified manner, a new general methodology to solve the GMP when its data are polynomials and basic semi-algebraic sets. This methodology combines semidefinite programming with recent results from real algebraic geometry to provide a hierarchy of semidefinite relaxations converging to the desired optimal value. Applied on appropriate cones, standard duality in convex optimization nicely expresses the duality between moments and positive polynomials. In the second part of this volume, the methodology is particularized and described in detail for various applications, including global optimization, probability, optimal control, mathematical finance, multivariate integration, games, etc., and examples are provided for each particular application.

Intended audience: graduate students and researchers in Applied Mathematics, Probability, Engineering, Management Science and Operations Research. The book is published by Imperial College Press, 2009. ISBN-10: 1848164459, ISBN-13: 978-1848164451.

3. SIAM OP11 Conference

Tenth SIAM Conference on Optimization
May 16 – 19, 2011
Darmstadtium Conference Center
Darmstadt, Germany
`http://www.siam.org/meetings/op11`

The Tenth SIAM Conference on Optimization will feature the latest research in theory, algorithms, and applications in optimization problems. In particular, it will emphasize convex relaxations, large-scale problems, optimization under uncertainty and will feature important applications in medicine, biology, networks, manufacturing, finance, aeronautics, control, operations research, and other areas of science and engineering. The conference brings together mathematicians, operations researchers, computer scientists, engineers, and software developers; thus it provides an excellent opportunity for sharing ideas and problems among specialists and users of optimization in academia, government, and industry.

Funding Agency

SIAM and the Conference Organizing Committee wish to extend their thanks and appreciation to Technische Universität Darmstadt for its support of this conference.

Themes

- Convex relaxations, compressed sensing and sparse optimization
- Polynomial optimization
- PDE-constrained and simulation-based optimization
- Stochastic/robust optimization (Optimization under uncertainty)
- Mixed integer nonlinear programming
- Large-scale nonlinear programming
- Discrete optimization
- Derivative-free optimization
- Applications in biology and medicine
- Mathematical Engineering

Organizing Committee

- Erling Andersen, MOSEK ApS, Copenhagen, Denmark
- Kristin P. Bennett, Rensselaer Polytechnic Institute
- Michael C. Ferris (Co-chair), University of Wisconsin
- Nicholas I. M. Gould, Rutherford Appleton Laboratory, United Kingdom
- Alexander Martin, Technische Universität Darmstadt, Germany
- Nikolaos Sahinidis, Carnegie Mellon University
- Alexander Shapiro, Georgia Tech
- Kim-Chuan Toh, National University of Singapore, Singapore
- Stefan Ulbrich, Technische Universität Darmstadt, Germany
- Stephen Vavasis (Co-chair), University of Waterloo, Canada

4. SIAG/OPT Prize 2011

We are soliciting nominations for the SIAM Activity Group on Optimization (SIAG/OPT) Prize. The prize established in 1992, and is awarded to the author(s) of the most outstanding paper, as determined by the prize committee, on a topic in optimization published in English in a peer-reviewed journal. The award is to be made at the 2011 SIAM Conference on Optimization 15th - 19th May in Darmstadt, Germany.

Prize Committee:

- Yinyu Ye, Stanford University (Chair)
- Shabbir Ahmed, Georgia Tech
- Philip Gill, University of California - San Diego
- Etienne de Klerk, Tilburg University, The Netherlands
- Jean Philippe Richard, the University of Florida

Eligibility

Candidate papers must be published in English in a peer-reviewed journal bearing a publication date within the award period. They must contain significant research contributions to the field of optimization, as commonly defined in the mathematical literature, with direct or potential applications. *Nominations should be send to Yinyu Ye by November 15, 2010.* For more information, see http://www.siam.org/prizes/sponsored/siagopt.php. The chair of the SIAG, the vice chair of the SIAG, and the members of the nominating committee are not eligible to receive the prize.

The previous recipients of the SIAM Activity Group on Optimization Prize are:

1996 Dimitris J. Bertsimas and Michel X. Goemans

1999 Michel X. Goemans and David P. Williamson

2002 Kurt Anstreicher, Nathan Brixius, Jean-Pierre Goux, and Jeff Linderoth

2005 Raphael Hauser

2008 Alexandre d'Aspremont, Laurent El Ghaoui, Michael I. Jordan, and Gert R.G. Lanckriet

Yinyu Ye
Management Science and Engineering
Terman Engineering Center 316
380 Panama Way
Stanford University
Stanford, California 94305-4026, USA
yinyu-ye@stanford.edu

Chairman's Column

I am writing this column with two days left before Christmas, an impending snow storm approaching Madison, and a stack of exams to complete grading. I note that I made similar comments in my column this time last year! However, by the time you read this, all of these issues will be long gone, and instead you can focus on the interesting papers in this volume, and reflect on and plan for the exciting events that are announced in this newsletter.

Along those lines, the dates and location for the next SIAG Optimization meeting are May 16–19, 2011 in Darmstadt, Germany. This location was chosen based on feedback obtained during the Boston meeting and the excellent proposal developed by Stefan Ulbrich and Alexander Martin and their colleagues. More details are given in the announcement in this newsletter. The organizing committee is busy preparing a list of plenary speakers and tutorials. Please provide your comments and suggestions to them to help ensure the meeting is a success.

The committee for the SIAG Optimization prize that will be announced at that meeting has been formed and will be chaired by Yinyu Ye. The call for nominations is provided in this newsletter. While it takes some effort to nominate papers for such prizes, I hope that you will make the task of the committee as difficult as possible by ensuring that all the exceptional papers in our field are considered for this prize. I believe that feedback in the form of prizes can help not only the prizewinners themselves, but also allow our community to highlight some of the great pieces of work that are going on in our field.

I believe that both the SIAM Annual Meeting in Denver and the International Symposium on Mathematical Programming in Chicago were great successes this year and I thank you for your participation. I am trusting that if you are reading this you have already renewed your membership of the SIAG - if not please make this a top priority! I welcome your comments on how the SIAG can more effectively help you, and suggestions on how to build stronger ties to other groups within and beyond the optimization field.

I read this morning that a proof of "the fundamental lemma" was announced as one of Time's top 10 scientific discoveries of 2009. It would be interesting to have your thoughts on what would constitute a similar result for us to strive for in the future and enable our field to trumpet the corresponding new opportunities facilitated by such advances. Let's look forward to this in 2010 and beyond!

Michael C. Ferris, SIAG/OPT Chair
Computer Sciences Department
University of Wisconsin-Madison
1210 West Dayton Street,
Madison, WI 53706
USA
`ferris@cs.wisc.edu`
`http://pages.cs.wisc.edu/~ferris/`

Comments from the Editor

It seems as though editing the SIAG/OPT Views-and-News has become a holiday tradition. However, this year I am a thousand miles further West, and the snowstorm that threatens Michael has already passed through Colorado.

A newsletter such as SIAG/OPT Views-and-News is only as successful as the articles that are submitted. I am fortunate to have received two interesting and well-written papers for this issue, but as always we need more papers. Please consider submitting a short paper (maybe summarizing a paper that you submit to SIOPT or Mathematical Programming), or suggest new hot topics for the next issue. Suggestions for new issues, comments, and papers are always welcome!

Sven Leyffer, Editor
Mathematics and Computer Science Division
Argonne National Laboratory
Argonne, IL 60439, USA
`leyffer@mcs.anl.gov`
`http://www.mcs.anl.gov/~leyffer/`