How to Become Outstanding in USA Mathematical Contest of Modeling

Zeyuan Allen Zhu
Fundamental Science Class, Dept. of Physics
Tsinghua University
zhuzeyuan@hotmail.com

- Our Outstanding Paper at MCM 2009
- What Makes Our Paper Outstanding?
- My Advices to MCM Attendees

Problem Description

- Designing a Traffic Circle
  - Many cities and communities have traffic circles—from large ones with many lanes in the circle (such as at the Arc de Triomphe in Paris and the Victory Monument in Bangkok) to small ones with one or two lanes in the circle. Some of these traffic circles position a stop sign or a yield sign on every incoming road that gives priority to traffic already in the circle; some position a yield sign in the circle at each incoming road to give priority to incoming traffic; and some position a traffic light on each incoming road (with no right turn allowed on a red light). Other designs may also be possible.

Problem Description

- Designing a Traffic Circle
  - Many cities and communities have traffic circles—from large ones with many lanes in the circle (such as at the Arc de Triomphe in Paris and the Victory Monument in Bangkok) to small ones with one or two lanes in the circle. Some of these traffic circles position a stop sign or a yield sign on every incoming road that gives priority to traffic already in the circle; some position a yield sign in the circle at each incoming road to give priority to incoming traffic; and some position a traffic light on each incoming road (with no right turn allowed on a red light). Other designs may also be possible.

Problem Description

- Designing a Traffic Circle
  - The goal of this problem is to use a model to determine how best to control traffic flow in, around, and out of a circle. State clearly the objective(s) you use in your model for making the optimal choice as well as the factors that affect this choice. Include a Technical Summary of not more than two double-spaced pages that explains to a Traffic Engineer how to use your model to help choose the appropriate flow-control method for any specific traffic circle. That is, summarize the conditions under which each type of traffic-control method should be used. When traffic lights are recommended, explain a method for determining how many seconds each light should remain green (which may vary according to the time of day and other factors). Illustrate how your model works with specific examples.

Happy Birthday to Somebody

Problem Description

- Designing a Traffic Circle
  - The goal of this problem is to use a model to determine how best to control traffic flow in, around, and out of a circle. State clearly the objective(s) you use in your model for making the optimal choice as well as the factors that affect this choice. Include a Technical Summary of not more than two double-spaced pages that explains to a Traffic Engineer how to use your model to help choose the appropriate flow-control method for any specific traffic circle. That is, summarize the conditions under which each type of traffic-control method should be used. When traffic lights are recommended, explain a method for determining how many seconds each light should remain green (which may vary according to the time of day and other factors). Illustrate how your model works with specific examples.

Remark

- Yield/Stop sign
  - A yield sign asks drivers to slow down and give the right of way to vehicles in the different direction. A stop sign asks drivers to come to a full stop before merging into the flow.
- Traffic light:
  - The function of traffic lights remains controversial. Comparing with yield /stop signs, traffic lights slow down the vehicle movement. At the same time, however, even at a remote motorway traffic circle with few pedestrians, a malfunction of traffic light will probably lead to an accident.

Arc de Triomphe

Three Steps to Make the Traffic Circle Go Round

Zeyuan Allen Zhu.	Dept. of Phys.
Tianyi Mao.	Dept. of Math
Yichen Huang.	Dept. of Phys.
Supervisor: Jun Ye	Dept. of Math

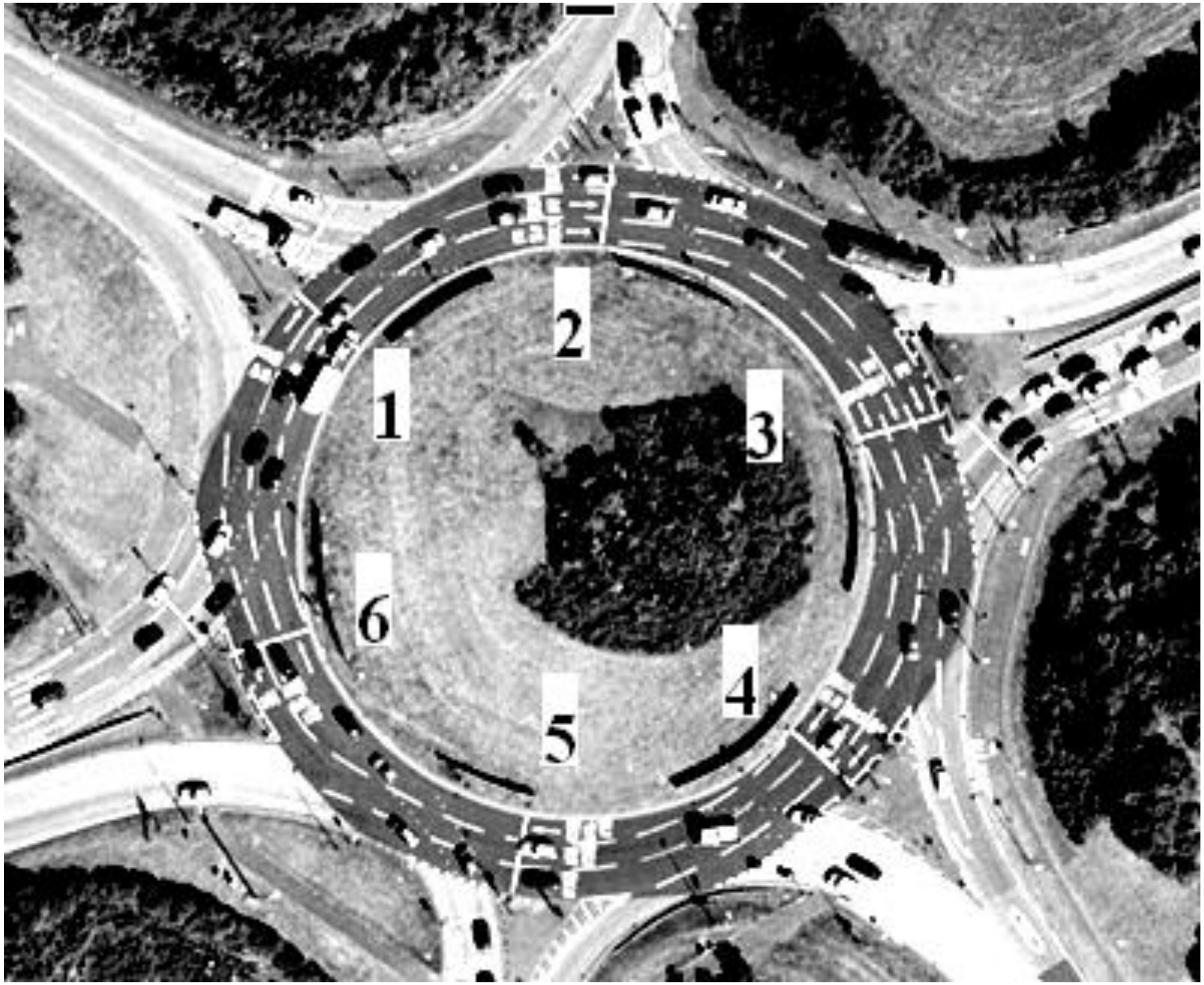
Introduction

- Two different models to simulate the flow.
  - The macro-model: modified Markov process,
  - The micro-model: modified cellular automata.
  - Consistency when applied to a real scenario in Scotland.
- The definition of a "good" method?
  - We choose five main objectives and combine them with an overall measure called the combined expense.
- Optimization
  - Genetic algorithm is employed
  - The green light period
  - The orientation signs
  - Unexpected affairs? like accidents or car breakdowns.
- Technical summary

Assumption

- The geometric design of the traffic circle cannot be changed.
- The traffic circle is a standard one (at grade) with all lanes on the ground, that is, no grade separation structure.
- The incoming vehicle flow in a period we study (e.g. 24 hours) is known.
- People drive on the left (since the example analyzed later is from the UK)
- Since we are considering the traffic flow, the pedestrians are ignored. In fact, the model can be easily modified to take this factor into account.
- Motorcycles move freely even in a traffic jam and are not taken into consideration.

Sheriffhall Roundabout



Circle: 3 lanes
Entrance 2/5: two lanes
Entrance 1/3/4/6: three lanes
3<->6 large flow!

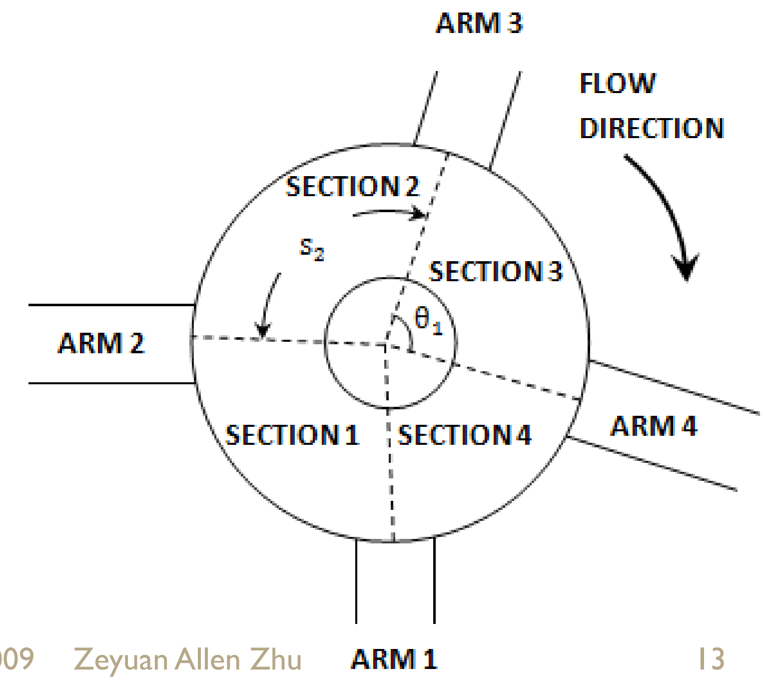
Introduction

- Two different models to simulate the flow.
  - The macro-model: modified Markov process,
  - The micro-model: modified cellular automata.
  - Consistency when applied to a real scenario in Scotland.
- The definition of a "good" method?
  - We choose five main objectives and combine them with an overall measure called the combined expense.
- Optimization
  - Genetic algorithm is employed
  - The green light period
  - The orientation signs
  - Unexpected affairs? like accidents or car breakdowns.
- Technical summary

Simulation Model

- Model I - The Macroscopic Simulation
  - First consider one-lane roads only.

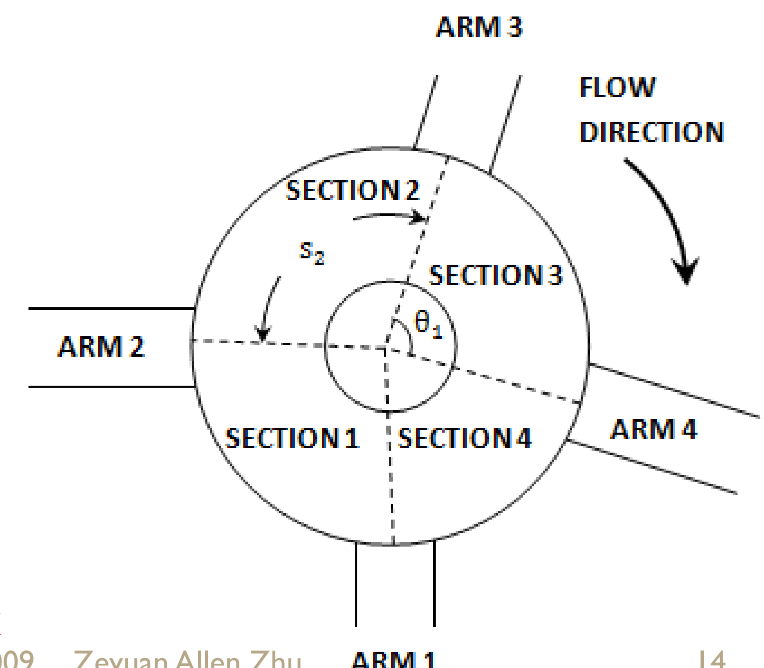
- # of vehicles contained in one section at time t: $\mathrm{num}_\mathrm{i}^\mathrm{t}$
- # of vehicles waiting to enter through one arm at time t: $\mathrm{arm}_\mathrm{i}^\mathrm{t}$
- Max # of vehicles to enter the traffic section through arm i at time t: $\mathrm{cap}_\mathrm{i}^\mathrm{t}$



Simulation Model

- Model I - The Macroscopic Simulation
  - First consider one-lane roads only.

- # of vehicles contained in one section at time t: $\mathrm{num}_\mathrm{i}^\mathrm{t}$
- # of vehicles waiting to enter through one arm at time t: $\mathrm{arm}_\mathrm{i}^\mathrm{t}$
- Max # of vehicles to enter the traffic section through arm i at time t: $\mathrm{cap}_\mathrm{i}^\mathrm{t}$



Steady distribution can be achieved through the normal Markov method?

NO!

Dimension of the transition probability matrix

Simulation Model

- Model I - The Macroscopic Simulation
  - First consider one-lane roads only.

Use the expectation $\overline{\mathrm{num}}_i^t$ and $\overline{\mathrm{arm}}_i^t$ instead of the distribution to denote a state

Traffic light?
One of the two streams is allowed

Stop/yield sign?
Disobey rate α_{stop} or α_{yield}

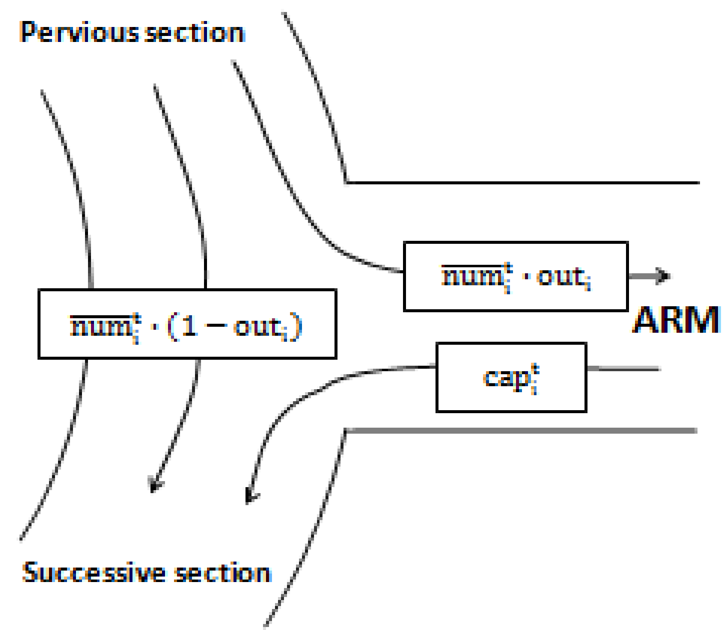


Figure 3: Flows at a junction

Simulation Model

- Model I - The Macroscopic Simulation
  - Multiple-lane?
  - The outflow for a given lane will be distributed into successive lanes according to their popularity

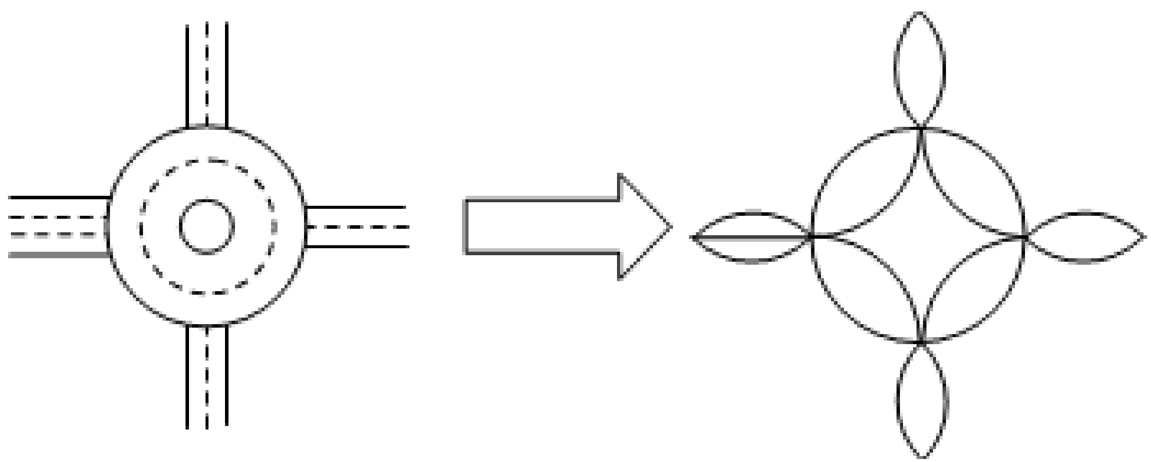


Figure 4: a two-lane circle has been divided into lanes.
Each arc in the right figure denotes a single lane.

Simulation Model

- Model II – The Microscopic Simulation
  - Sequential Cellular Automata?
  - Vehicles can stay at real argument in polar coordinates but with discrete radius values.

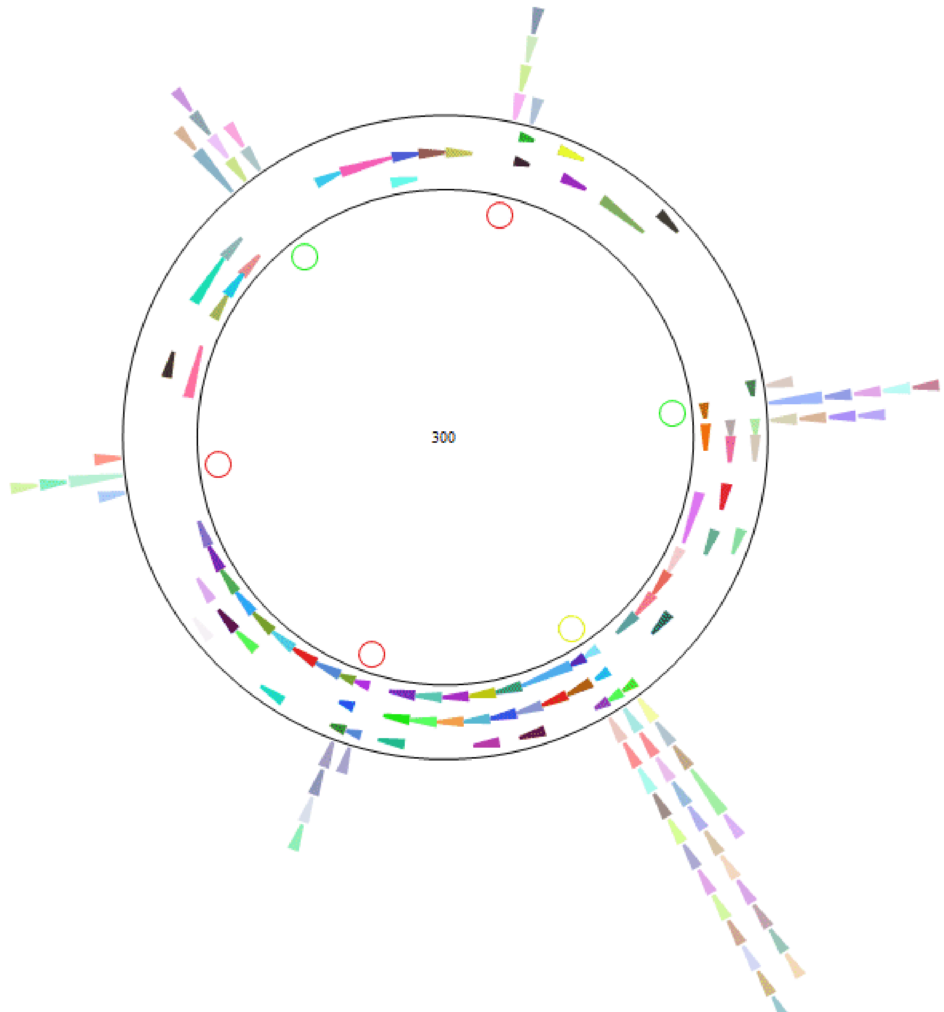
Simulation Model

- Model II – The Microscopic Simulation
  - Traffic coming in
    Given matrix $\left(a_{i,j}\right)_{n\times n}$. Use a Poisson distribution with mean $\frac{a_{i,j}}{T}$ to describe the incoming vehicles from arm i to arm j.
  - Lane choosing and changing
    The hidden principle is: the more sections the vehicle is to pass before its exit, the inner lane the driver wish to take, both in the arm and the circle.
  - Vehicle speed
    Maximum velocity v_{max} and maximum acceleration a_{max} for vehicles, and record the velocity individually.
    - Facing a red light or other vehicles, its speed decreases to zero.
    - When a vehicle changes lanes, it decelerates.
    - Otherwise, it will attempt to accelerate.

Simulation Model

- Model II – The Microscopic Simulation
  - The function of a yield sign.
    When a vehicle faces a yield sign, it checks whether it is empty enough for it to enter the junction. If not, it will wait till it is empty enough, but with a disobey rate α_{yield} - ignore the sign and try to scramble. This reaction affects the accident rate.
  - The function of a stop sign.
    When a vehicle faces a stop sign, it should stops instantaneously. As a next step, it deals with similar procedure as a yield sign. The only difference is that it will accelerate from a zero speed. The disobey rate is α_{stop}
  - The effect of traffic lights – just like normal.

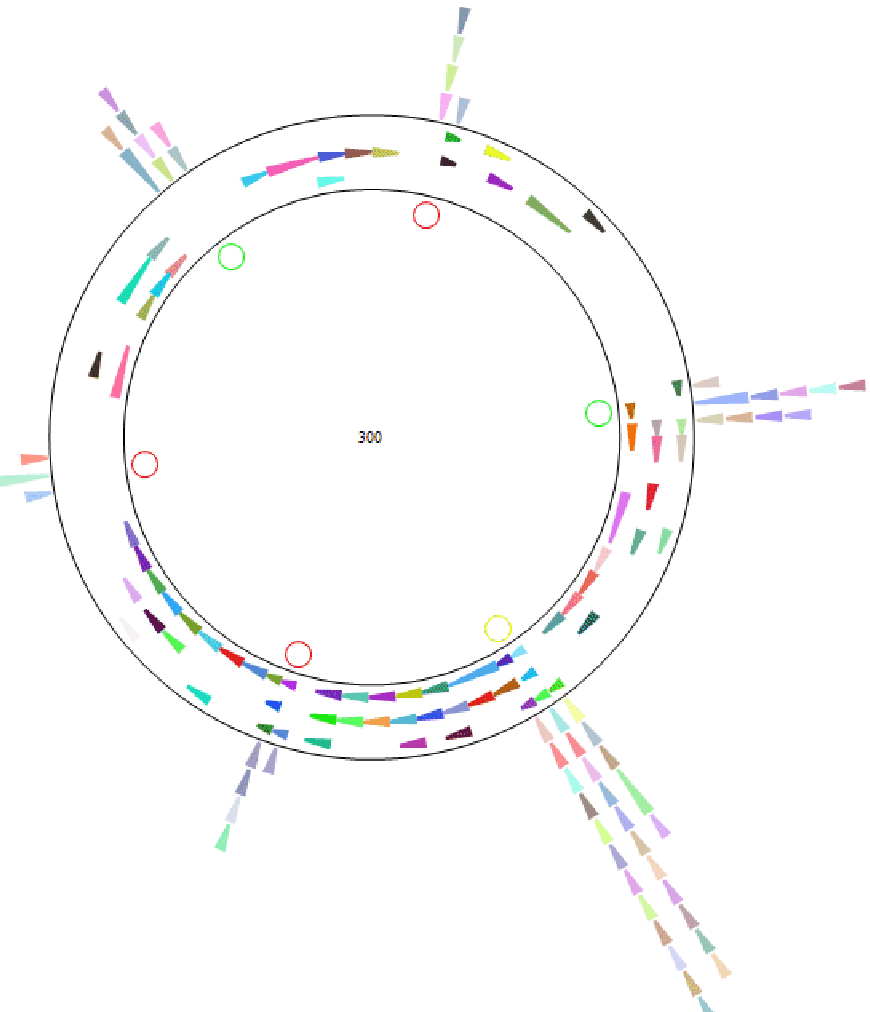
Simulation Model



Compare Two Models

Results?
Sensitivity?
Complexity?

Parameter	variation	model I	model II
v_{max}	+10%	-2.6%	-8.5%
	-10%	10.5%	11.1%
l_0	+1	-19.6%	-16.4
	-1	121%	65.2%
r_{out}	+10%	-7.3%	-3.9%
	-10%	1.1%	3.3%
Traffic flow	+10%	10.6%	7.0%
	-10%	-3.0%	-6.7%



Introduction

- Two different models to simulate the flow.
  - The macro-model: modified Markov process,
  - The micro-model: modified cellular automata.
  - Consistency when applied to a real scenario in Scotland.
- The definition of a "good" method?
  - We choose five main objectives and combine them with an overall measure called the combined expense.
- Optimization
  - Genetic algorithm is employed
  - The green light period
  - The orientation signs
  - Unexpected affairs? like accidents or car breakdowns.
- Technical summary

The Multi-Objective Function

- Saturated flow capacity
  - When traffic flow becomes tremendous, vehicles will accumulate on arms as time goes by. The threshold flux to avoid such an accumulation is called the saturated flow capacity under a certain control method.
- Average delay
  - A vehicle may suffer various kinds of delay in congestion, acceleration and more. The difference between the average time during which a vehicle passes the traffic circle and it passes a void one is called the average delay.
- Equity degree
  - A multi-arm traffic circle may distribute the incoming flow inequably, which annoys drivers more frequently. The relative derivation of average delay is called the equity degree.
- Accident expectation
  - Too much emphasize on speed may mean potential risk of accident. The consideration of this threat calls for the description of accident per hour, which is named the accident expectation.
- Device cost
  - The total expense on the traffic signs and lights is called the device cost.

The Multi-Objective Function

- Non-trivial combination!
- To minimize *the combined expense (CE)* in \$.
  - $CE = Corrected\ Delay\ Expense - Capacity\ Bonus + Accident\ Loss + Device\ Cost$

The Multi-Objective Function

- Non-trivial combination!
- To minimize *the combined expense (CE)* in $.
  - $CE = Corrected\ Delay\ Expense - Capacity\ Bonus + Accident\ Loss + Device\ Cost$
- Device Cost – Easy!
  - Prices of traffic control devices are easy to be accessed on the Internet [8][9].
  - A traffic light is expected to spend $0.23/hour according to the viable data [10][11]

The Multi-Objective Function

- Non-trivial combination!
- To minimize *the combined expense (CE)* in $.
  - $CE = Corrected\ Delay\ Expense - Capacity\ Bonus + Accident\ Loss + Device\ Cost$
- Accident Loss?
  - Accident expense losses are often reported in news.
  - We take the average data from an annual report of local traffic office as the average loss per accident [12]:
  - $Accident\ Loss = \$630 \times Flux$

The Multi-Objective Function

- Non-trivial combination!
- To minimize *the combined expense (CE)* in $.
  - $CE = Corrected\ Delay\ Expense - Capacity\ Bonus + Accident\ Loss + Device\ Cost$
- Delay Expense?
  - From the data of US Department of Transportation and fuel price [13], about $1.2 is lost on a delay of one hour (per vehicle):
  - $Delay\ Expense = \$1.2 \times Flux \times Average\ Delay\ Time$

The Multi-Objective Function

- Non-trivial combination!
- To minimize *the combined expense (CE)* in $.
  - $CE = Corrected\ Delay\ Expense - Capacity\ Bonus + Accident\ Loss + Device\ Cost$
- Delay Expense?
  - From the data of US Department of Transportation and fuel price [13], about $1.2 is lost on a delay of one hour (per vehicle):
  - $Delay\ Expense = \$1.2 \times Flux \times Average\ Delay\ Time$
  - Equity?

The Multi-Objective Function

- Non-trivial combination!
- To minimize *the combined expense (CE)* in $.
  - $CE = \text{Corrected Delay Expense} - \text{Capacity Bonus} + \text{Accident Loss} + \text{Device Cost}$
- Corrected Delay Expense!
  - Equity degree (ED) is a tricky component in the determination. The most annoying situation is to keep two "main arms" open to traffic by sacrificing all other arms, whose equity degree is estimated to be the function of the number of arms n
  - $\text{Reference Equity Degree}\ (RED) = \sqrt{\frac{n(n-2)}{2(n-1)}}$

The Multi-Objective Function

- Non-trivial combination!
- To minimize *the combined expense (CE)* in $.
  - $CE = \text{Corrected Delay Expense} - \text{Capacity Bonus} + \text{Accident Loss} + \text{Device Cost}$
- Corrected Delay Expense!
  - $\text{Corrected Delay Expense} =$

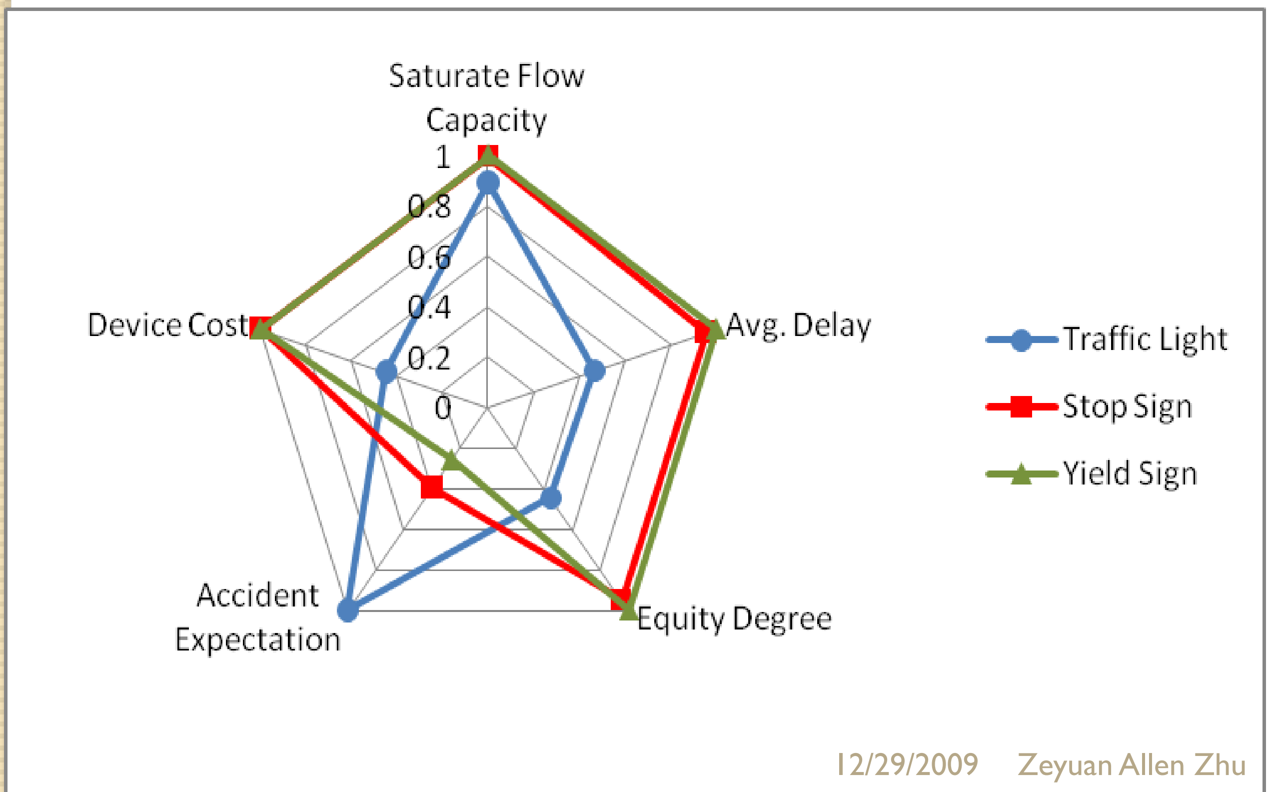
$$\text{Delay Expense} \times \left(1 + \frac{ED}{RED}\right)$$

The Multi-Objective Function

- Non-trivial combination!
- To minimize *the combined expense (CE)* in $.
  - $CE = Corrected\ Delay\ Expense - Capacity\ Bonus + Accident\ Loss + Device\ Cost$
- Capacity Bonus
  - The unused part of the capacity assures for any extra incoming, whose value is estimated as following:
  - $Capacity\ Bonus = 5\% \times \$1.2 \times (Saturate\ Capacity\ - Flux) \times Average\ Delay\ Time,$
    - in which 5% is the possibility of an unexpected vehicle coming.

The Multi-Objective Function

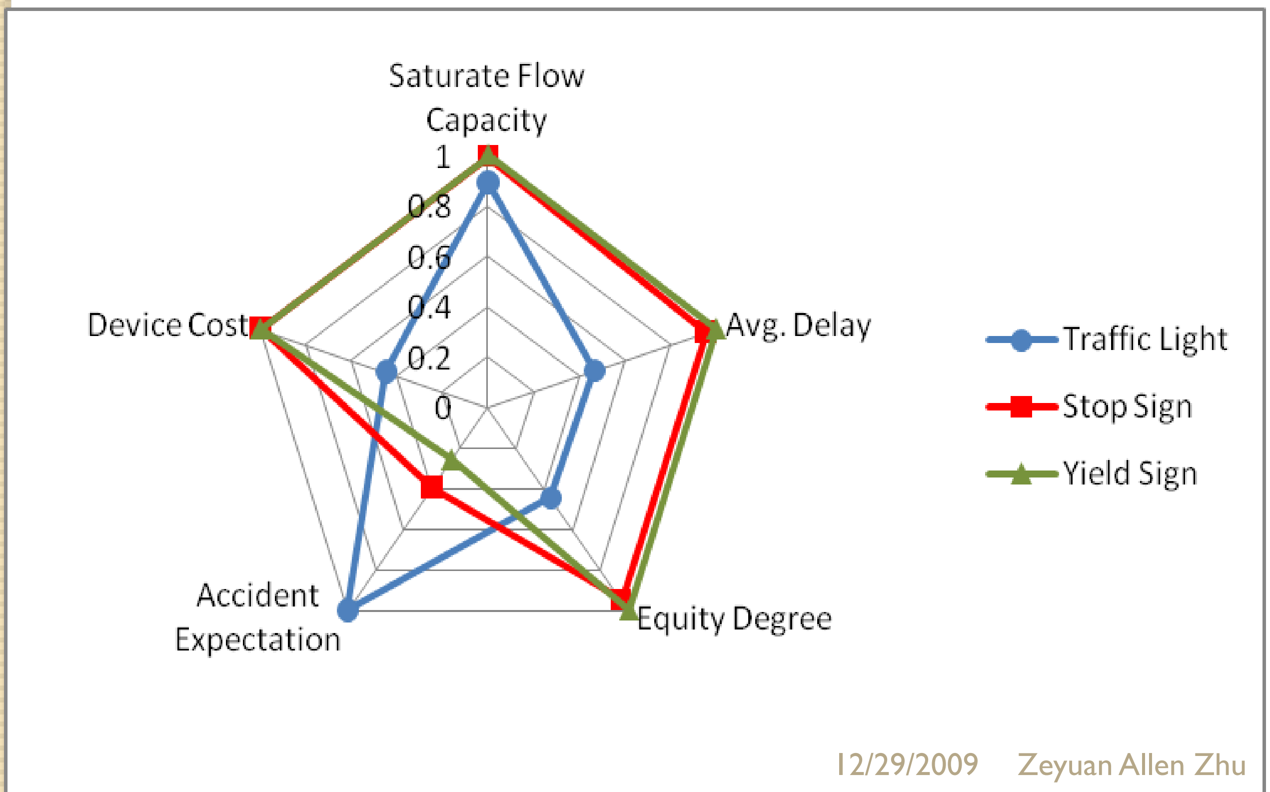
- Application



The Multi-Objective Function

- Application

Control Method	The Combine Expense (US$/hour)
Traffic light	66.76
Stop sign	103.29
Yield sign	116.61



Introduction

- Two different models to simulate the flow.
  - The macro-model: modified Markov process,
  - The micro-model: modified cellular automata.
  - Consistency when applied to a real scenario in Scotland.
- The definition of a "good" method?
  - We choose five main objectives and combine them with an overall measure called the combined expense.
- Optimization
  - Genetic algorithm is employed
  - The green light period
  - The orientation signs
  - Unexpected affairs? like accidents or car breakdowns.
- Technical summary

Optimization Model

- Multi-Objective → Hard to optimize.
  - Quasi-optimal ones, Genetic Algorithm

Optimization Model

- Multi-Objective → Hard to optimize.
  - Quasi-optimal ones, Genetic Algorithm

Breeding
- Combine the traffic control methods of two different configurations.

Mutation
- Randomly mutate the traffic control in a single junction.

Evolution
- Locally adjust the traffic controls in all junctions, and seek for better solution.

Optimization Model

- Step I: Basic Device Placement & Timestamp

Arm	G	R
1	0	14
2	7	14
3	38	18
4	3	19
5	3	13
6	38	14

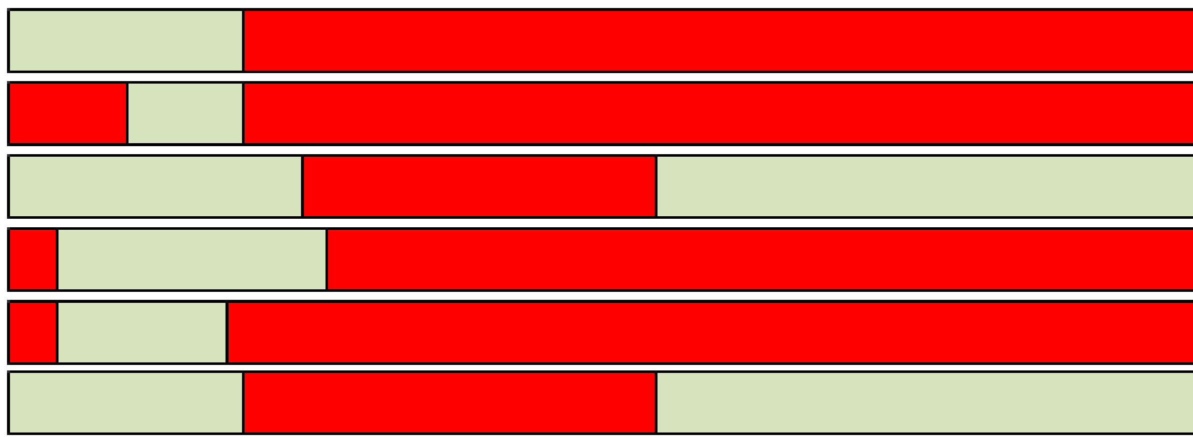


Figure 7: The traffic light timestamps in 6 junctions (light green / and scarlet)

Period = 68s (calculated in assumption). **Original flow info is used.**

Objective	Value
Saturated Flow Capacity	6904 vehicles / hour
Average Delay	42.763 seconds / vehicle · hour = 62.04$ / hour
Equity Degree	0.3187
Accident Expectation	4.63$ / hour
Device Cost	1.38$ / hour
Combined Expenses	78.98$ / hour

Optimization Model

- Step I: Basic Device Placement & Timestamp

Arm	G	R
1	0	12
2	10	18
3	49	15
4	38	57
5	41	52
6	10	45

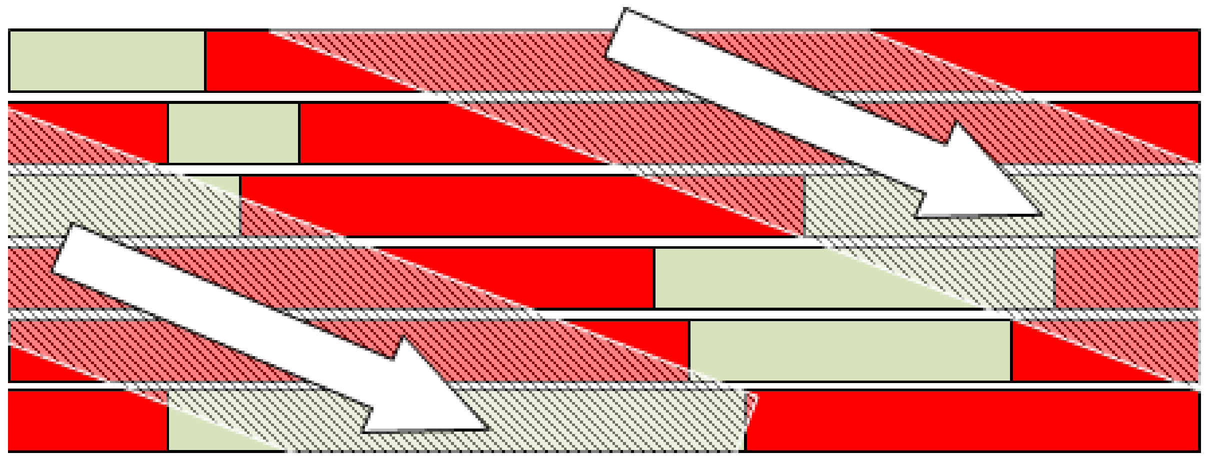


Figure 8: The traffic light timestamps in 6 junctions (light green / and scarlet)

Period = 68s (calculated in assumption). **Original flow × 1.8 is used.**

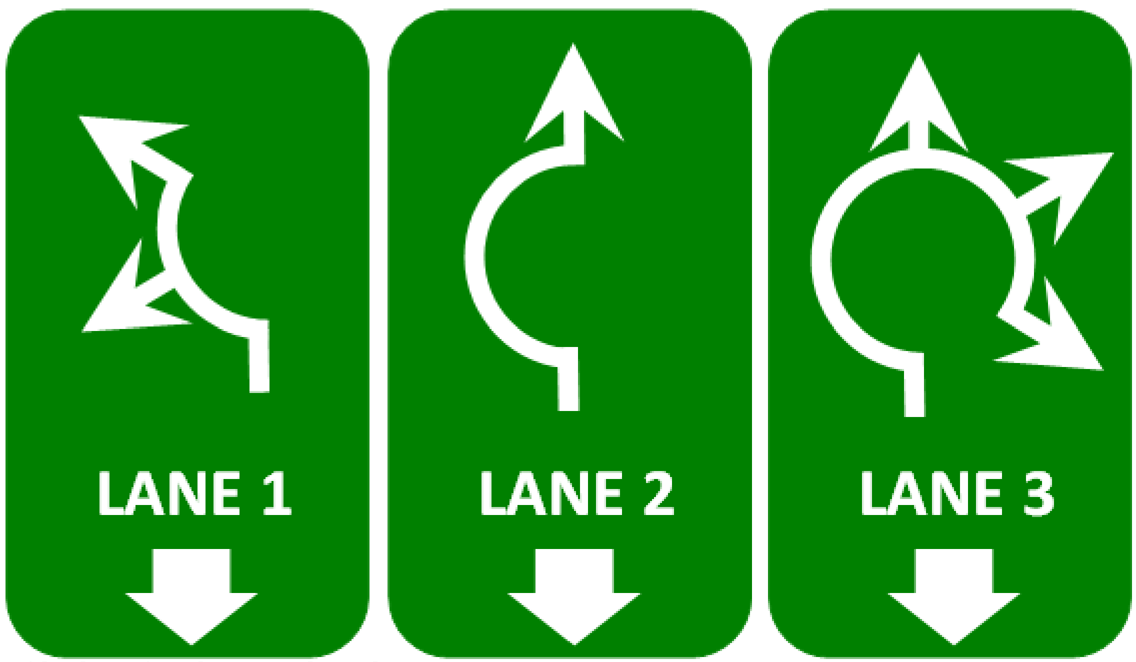
Objective	Value
Saturated Flow Capacity	8354 vehicles / hour
Average Delay	81.278 seconds / vehicle · hour = 117.91$ / hour
Equity Degree	0.3042
Accident Expectation	5.41$ / hour
Device Cost	1.38$ / hour
Combined Expenses	147.14$ / hour

Optimization Model

- Step II: Orientation Sign Placement
  - A hidden rule [5] is: the vehicle nearer to its exit should stay left (Remark: we are driving on the left!) Now we are going to refine this rule.

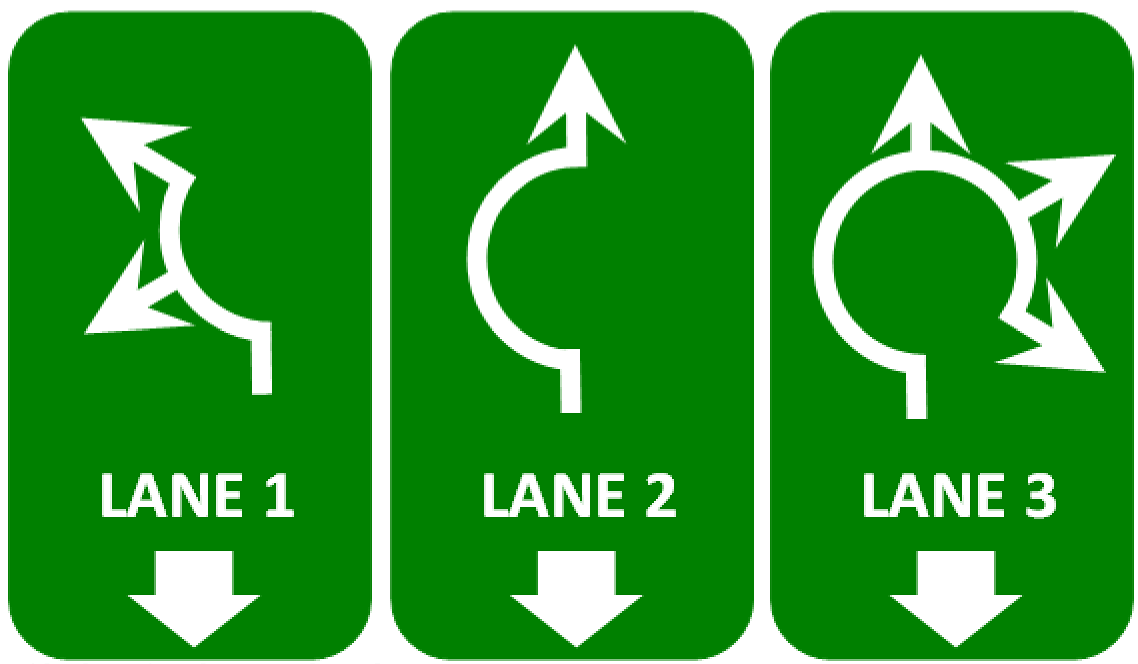
Optimization Model

- Step II: Orientation Sign Placement
  - A hidden rule [5] is: the vehicle nearer to its exit should stay left (Remark: we are driving on the left!) Now we are going to refine this rule.



Optimization Model

- Step II: Orientation Sign Placement
  - Vehicle is at Junct. a with dest. b is suggested stay in the range of $\mathrm{lower}_{\mathrm{a}}^{\mathrm{b}} \leq \mathrm{x} \leq \mathrm{upper}_{\mathrm{a}}^{\mathrm{b}}$.

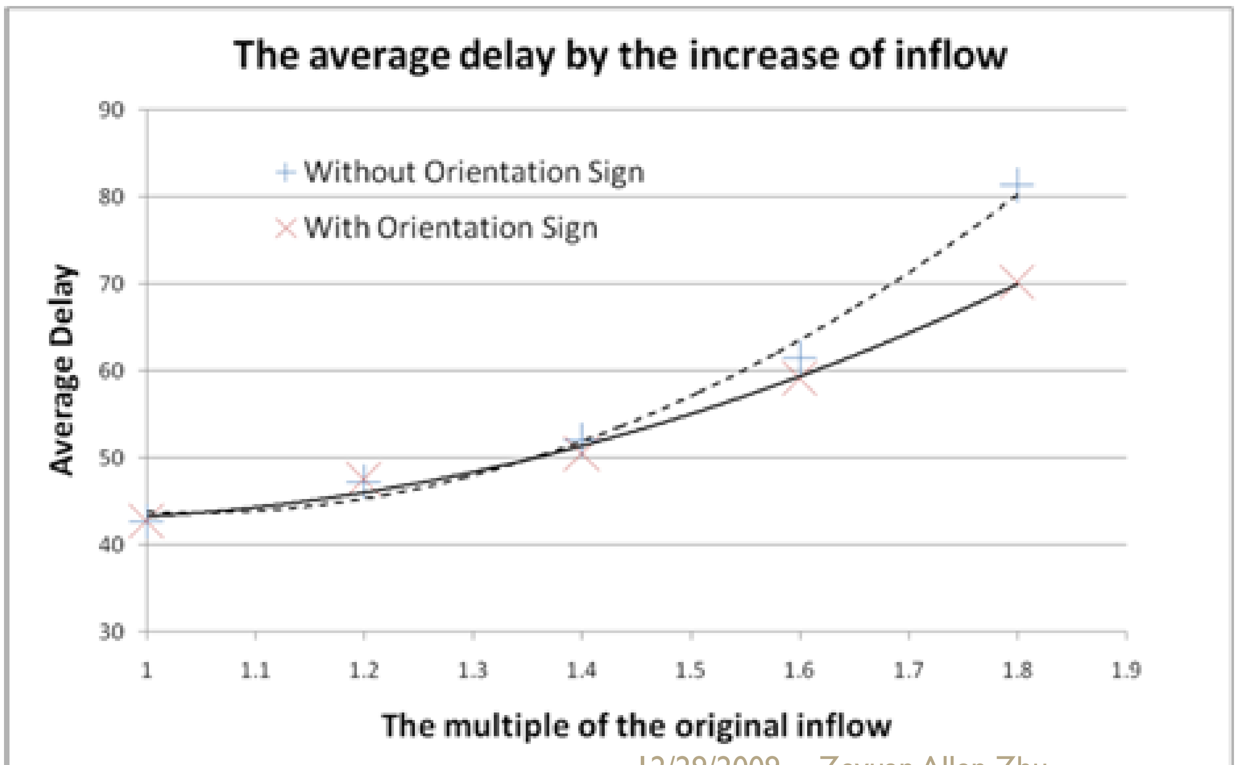


Optimization Model

- Step II: Orientation Sign Placement
  - Vehicle is at Junct. a with dest. b is suggested stay in the range of $\mathrm{lower}_{\mathrm{a}}^{\mathrm{b}} \leq \mathrm{x} \leq \mathrm{upper}_{\mathrm{a}}^{\mathrm{b}}$.
  - Do the Genetic Algorithm again.

Optimization Model

- Step II: Orientation Sign Placement
  - Watch the magic effect!
  - 8354 → 8812 in Saturated Flow Capacity



Optimization Model

- Step III: Time Variance & Self Adaptivity
  - Select the original configuration as our seed, and carry on the Generic Algorithm and gain a similar but better solution

Arm	G	R
1	0	12
2	10	18
3	49	15
4	38	57
5	41	52
6	10	45

Arm	G	R
1	0	14
2	11	18
3	51	15
4	21	46
5	41	53
6	12	45

Figure 11: The Self Adaptivity as the inflow drops from 1.8 to 1.4 × original inflow in an hour

Optimization Model

- Step III: Time Variance & Self Adaptivity
  - As the night falls, traffic demands become considerably small. Now the traffic lights are suggested to be replaced by Yield/Stop Signs.
  - This cannot happen in the usual way. However, a special status of traffic light can be switched on – *flashing yellow*.

Optimization Model

- Verification!
  - First: The Circle is In Work!
  - In 1.8x flow

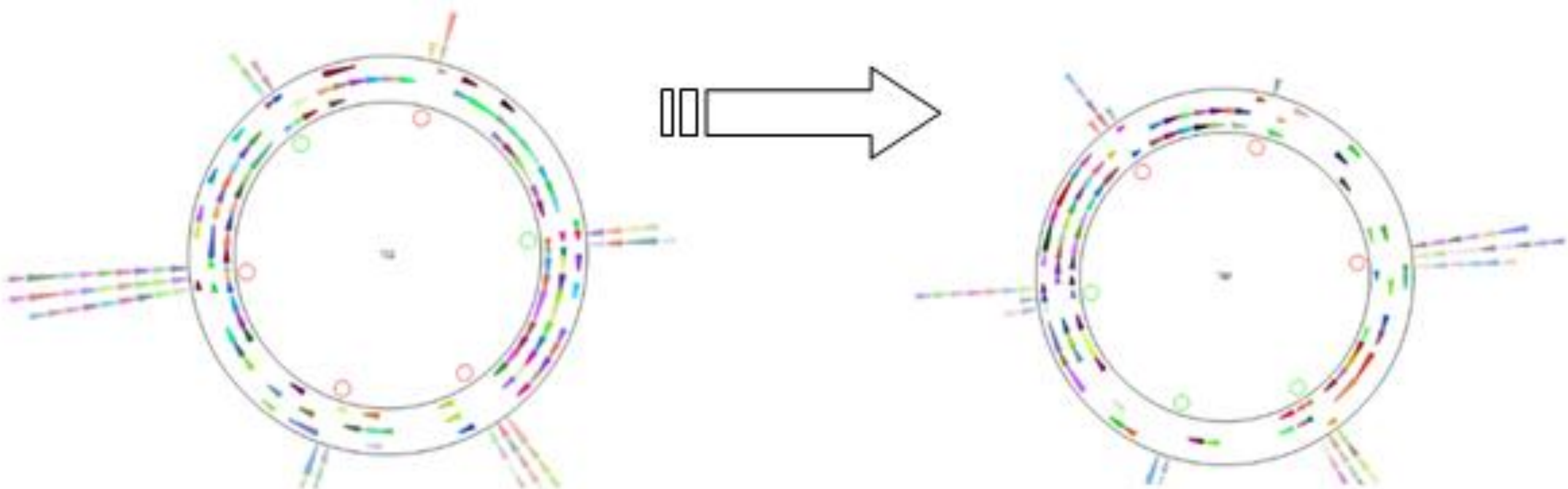


Figure 12: 39 seconds later, most of the vehicles waiting at Junction 6 moves in

Optimization Model

- Verification!
  - Second: Accuracy
  - Use Dummy Circles. Why always traffic lights?
  - Testing on a dummy suburban circle with 4 arms with relatively lower traffic demand

Optimization Model

- Verification!
  - Second: Accuracy
    - 1) the mixture of stop signs and traffic lights;
    - 2) the in circle signs are encountered.

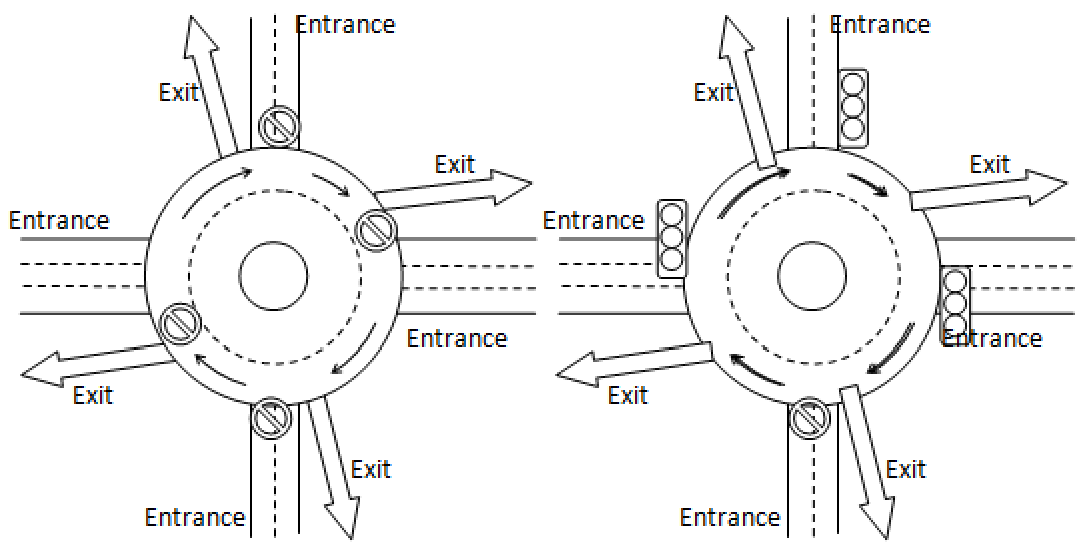


Figure 13: Two intuitive configurations generated by our model.
The left one has two in-circle stop signs and guarantees the left-right fast pass;
The right one has a mixture of traffic lights and stop signs.

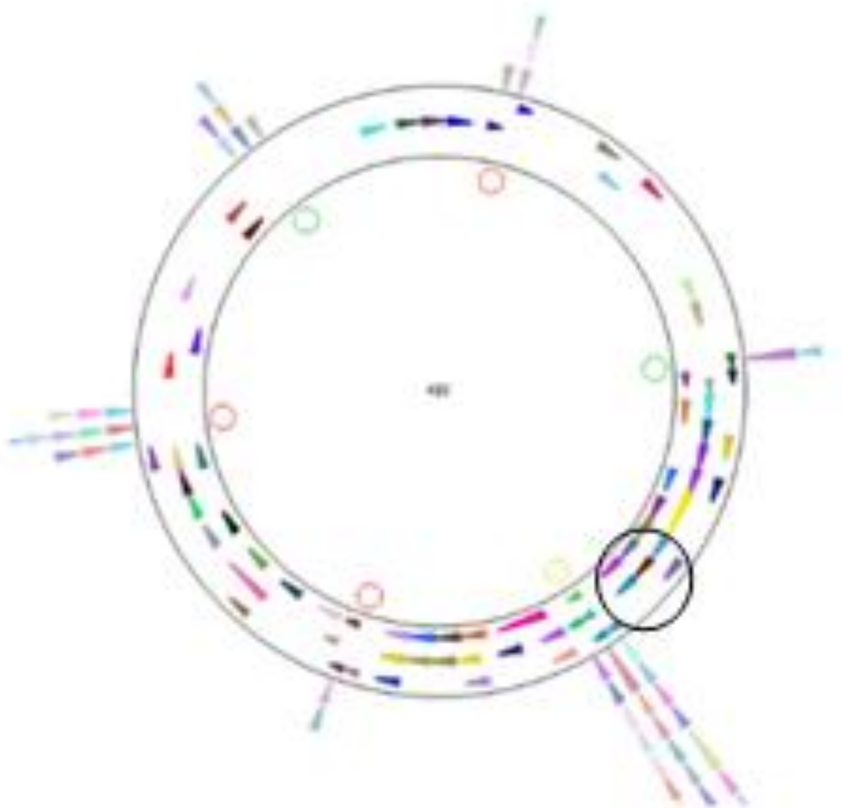
Optimization Model

- Verification!
  - Third: Sensitivity
    - Run the program 50 times.

The multiple of income flow	Average Delay	Standard Deviation
1.0 times	42.76 seconds / vehicle · hour	0.95 seconds / vehicle · hour
1.2 times	47.22 seconds / vehicle · hour	1.56seconds / vehicle · hour
1.4 times	51.99 seconds / vehicle · hour	2.54 seconds / vehicle · hour
1.6 times	61.54 seconds / vehicle · hour	3.81 seconds / vehicle · hour
1.8 times	81.28 seconds / vehicle · hour	8.30 seconds / vehicle · hour

Optimization Model

- Verification!
  - Fourth: Emergency case



Introduction

- Two different models to simulate the flow.
  - The macro-model: modified Markov process,
  - The micro-model: modified cellular automata.
  - Consistency when applied to a real scenario in Scotland.
- The definition of a “good” method?
  - We choose five main objectives and combine them with an overall measure called the combined expense.
- Optimization
  - Genetic algorithm is employed
  - The green light period
  - The orientation signs
  - Unexpected affairs? like accidents or car breakdowns.
- Technical summary

Technical Summary

- Non-interesting part? No!
- What data do we need?
  - Geometry design
  - # of lanes in any road
  - Traffic Flow
- Metropolitan traffic circle?
  - Traffic lights
- Suburban areas?
  - The traffic flows on arms of two vertical directions show great difference
  - a stop sign should stand at the junction of one low-stream arm with traffic lights at the others

What Makes Our Paper Outstanding?

Zeyuan Allen Zhu

What Makes It Above-Meritorious?

- Designing a Traffic Circle
  - Many cities and communities have traffic circles—from large ones with many lanes in the circle (such as at the Arc de Triomphe in Paris and the Victory Monument in Bangkok) to small ones with one or two lanes in the circle. Some of these traffic circles position a stop sign or a yield sign on every incoming road that gives priority to traffic already in the circle; some position a yield sign in the circle at each incoming road to give priority to incoming traffic; and some position a traffic light on each incoming road (with no right turn allowed on a red light). Other designs may also be possible.

What Makes It Above-Meritorious?

- Designing a Traffic Circle
  - The goal of this problem is to use a model to determine how best to control traffic flow in, around, and out of a circle. State clearly the objective(s) you use in your model for making the optimal choice as well as the factors that affect this choice. Include a Technical Summary of not more than two double-spaced pages that explains to a Traffic Engineer how to use your model to help choose the appropriate flow-control method for any specific traffic circle. That is, summarize the conditions under which each type of traffic-control method should be used. When traffic lights are recommended, explain a method for determining how many seconds each light should remain green (which may vary according to the time of day and other factors). Illustrate how your model works with specific examples.

What Makes It Above-Meritorious?

- Answered the question
- Double-Model
  - Discrete – Continuous?
  - Macroscopic – Microscopic?
- Reasonable Optimization
- Reasonable Verification
  - Sensitivity test?
  - Comparison?
  - Other claim to support your solution?

What Makes It Outstanding!

- Outstanding is hard
  - MCM 2009 – 8xUSA / 1xCHN
- Innovation
  - Orientation Sign
- Strengthened Model
  - Multi-Objective
  - Self-Adaptivity
  - Good Demonstration
    - Picture
    - Text

What Makes It More Outstanding?

- Other Innovations:
  - To illustrate exactly what happened when an ambulance enters the traffic circle.
  - Outside Circle – Just like Arc de Triomphe
- Rewording…
- Disadvantages…

My Advices to MCM Attendees

Zeyuan Allen Zhu

Enthusiasm!

- Of Prime Importance!
- Math not so good? Coding not so good?
  - Limited influence only.
- The sole rule I built our team.
- If you are not aiming at an Outstanding prize, skip the next page ☺

Enthusiasm!

- To testify your enthusiasm: would you like to
  - Fight for 3/4 consecutive days, including a sleepless last-night. Skipping all required courses?
  - Play a dummy but simulated competition with your teammates for 3/4 consecutive days?
  - Stay at Tsinghua, giving up your opportunity to stay with your family in The New Year Week?

How to form a team?

1+ coder (1 is dangerous, 2 recommended)

I coded 200KB C# during MCM 2009.

Yichen coded 500+lines in Pascal.

IOI / NOI / ACM experience

1+ math major

Lack of mathematical deduction? Try ICM

2+ writer (2 is dangerous, 3 recommended)

TOEFL writing section 26+

Math & Applied Math

What we used?

- Markov
- Cellular automata
- Genetic algorithm.

What we have used in 7 other MCM competitions?

- Simplex
- Curve fitting
- Stimulated annealing
- Principal component analysis
- Stepwise regression
- Analytic hierarchy process, etc.

What software did we use?

- SPSS (rare)
- Visual Studio (always)
- Mathematica (Visualization)
- Matlab

Before the Contest

- Train your survey skills
  - Google everything!
  - ACM / IEEE paper download?
  - Need to pay?
- Play a dummy contest, to understand whether your team can work as a team
  - What if quarreled?
  - To keep everything scheduled
  - Self-sacrifice.

During The MCM 2009

Day 1

- Brainstorm on the problem, any innovations?
- Search for tools (failed)
- Model II ready

Day 2

- Model II coding
- Model I ready
- Objective working
- Searching for data (difficult)

Day 3

- Model II coding (dangerous)
- Model I coding (dangerous)
- Started paper-writing
- Objective ready

Day 4

- Model II coding (very dangerous)
- Orientation Sign idea ready
- Everybody writing.

During the Contest

- Don't stick to specific model parameters
  - write down the idea instead your detailed model.
  - write everything clear and easy to understand.
- Don't copy other's work – MCM 2008
  - Quote everything you used
- Code from the first day!
  - Except you're an IOI/NOI gold medalist.
- Start paper-writing from the second day!
  - Except …

Advices for the Chief Coder

- Make sure you can code everything in time
  - Make a plan
    - What will be implemented?
    - What should be skipped?
    - What if the result is bad, any substitute algorithm?
  - Make a promise to your teammates
    - I can at least have this done before 06h00.
    - I will expect this to be generated out, so please help me write …

Advices for the Chief Writer

- Make sure you can write everything in time
  - Make a plan
    - What will be written?
    - What should be skipped?
    - What if the result is bad, any substitute paragraph?
    - Write down a memo!
  - Make the paper!
    - Define the Terminology!
    - Eliminate the Self-contradiction.
    - Prepare the References.

Advices for Paper Writing

- (not necessarily correct, but strongly recommended)
- LaTex as a first prior! All team members know LaTex… Not easily satisfiable…
- If choosing Microsoft Word (likely):
  - Use Word 2007 formula!
  - Try to imitate LaTex's style.
  - Try to be scientific.

Advices for Paper Writing

- Word 2003 using Equation 3.0.
- This is actually good!

Following the introduction to DDF in Section 3, we will present a novel boosting algorithm, MPBoost. Specifically, we apply the GentleBoost approach [10] to define loss function and design optimization methods.

For convenience, we combine the preference sets over all queries: $S = \bigcup_q S_q'$ and define the index set I over S : $I = \{(i, j) \mid ((x_i, x_j), \mathrm{dist}(r_i, r_j)) \in S\}$. Note that for sake of simplicity, we will sometimes omit the query subscript in the following discussion, i.e. whenever (x_i, x_j) is involved, we assume that the documents x_i and x_j belong to the same query q , and $r_{qi} \neq r_{qj}$. Similarly, $dist(r_i, r_j)$ means $dist(r_{qi}, r_{qj})$.

Next, to leverage DDF in the MPBoost algorithm, we require that the condition

$$| dist(r_i, r_j) | \approx 1 \qquad (9)$$

is satisfied, which can be attained via carefully setting the parameters within the applied directed distance functions, like α , λ and β in LDD, LOGDD and LOGITDD. This condition will be utilized in the following analysis.

Now, the loss function MPBoost employs is as follows:

$$J(F) = \sum_I e^{-dist(r_i, r_j)(F(x_i) - F(x_j))} \qquad (10)$$

where the strong hypothesis $F(x)$ is a score function based on an

Advices for Paper Writing

- Written by LaTex

We begin our approach from objectives most related to money. The prices of traffic control devices are easy to be accessed on the Internet [8][9]. Along the expense of maintenance and operation, the average cost per hour for each kind of device is calculated. Since traffic lights consume much electricity, we ignore the money spent on other types of devices. A traffic light is expected to spend \$0.23/hour according to the viable data [10][11].

Accident expense losses are often reported in news. We take the average data from an annual report of local traffic office as the average loss per accident [12]:

$$\textit{Accident Loss}=\$630 \times \textit{Flux}$$

The average delay time must be accompanied with a value of delay to get involved. From the data of US Department of Transportation and fuel price [13], about \$1.2 is lost on a delay of one hour (per vehicle):

$$\textit{Delay Expense}=\$1.2 \times \textit{Flux} \times \textit{Averge Delay Time}$$

The saturate capacity shows the endurance of the control design as well as its ability to face sudden challenges. The unused part of this capacity assures for any extra incoming, whose value is estimated as following:

$$\textit{Capacity Bonus} = 5\% \times \$1.2 \times (\textit{Saturate Capacity} - \textit{Flux}) \times \textit{Average Delay Time}$$

in which 5% is the possibility of an unexpected vehicle coming.

Equity degree (ED) is a tricky component in the determination. Current traffic systems often fail to consider this factor. The most annoying situation is to keep two "main arms" open to traffic by sacrificing all other arms, whose equity degree is estimated to be the function of the number of arms n:

Advices for Paper Writing

- Written by Word 2007 with in-build equation

I. INTRODUCTION

In the regularized learning theory, in order to minimize the sum of the *regularization* part and the *loss* part, most of the research works are interested in the *generalization objective* rather than the *empirical objective* [12] [1]. The generalization objective, also known as the *stochastic objective*, is given with respect to a linear predictor $\boldsymbol{w} \in S$, where $S \subset \mathbb{R}^n$ is the domain of $\boldsymbol{w}$:

$$\begin{aligned} F_\sigma(\boldsymbol{w}) &= \sigma \cdot r(\boldsymbol{w}) + l(\boldsymbol{w}) \\ &= \sigma \cdot r(\boldsymbol{w}) + \mathbb{E}_{\boldsymbol{\theta} \sim Dist}[l(\langle \boldsymbol{w}, \boldsymbol{\theta} \rangle; \boldsymbol{\theta})] \end{aligned} \quad (1)$$

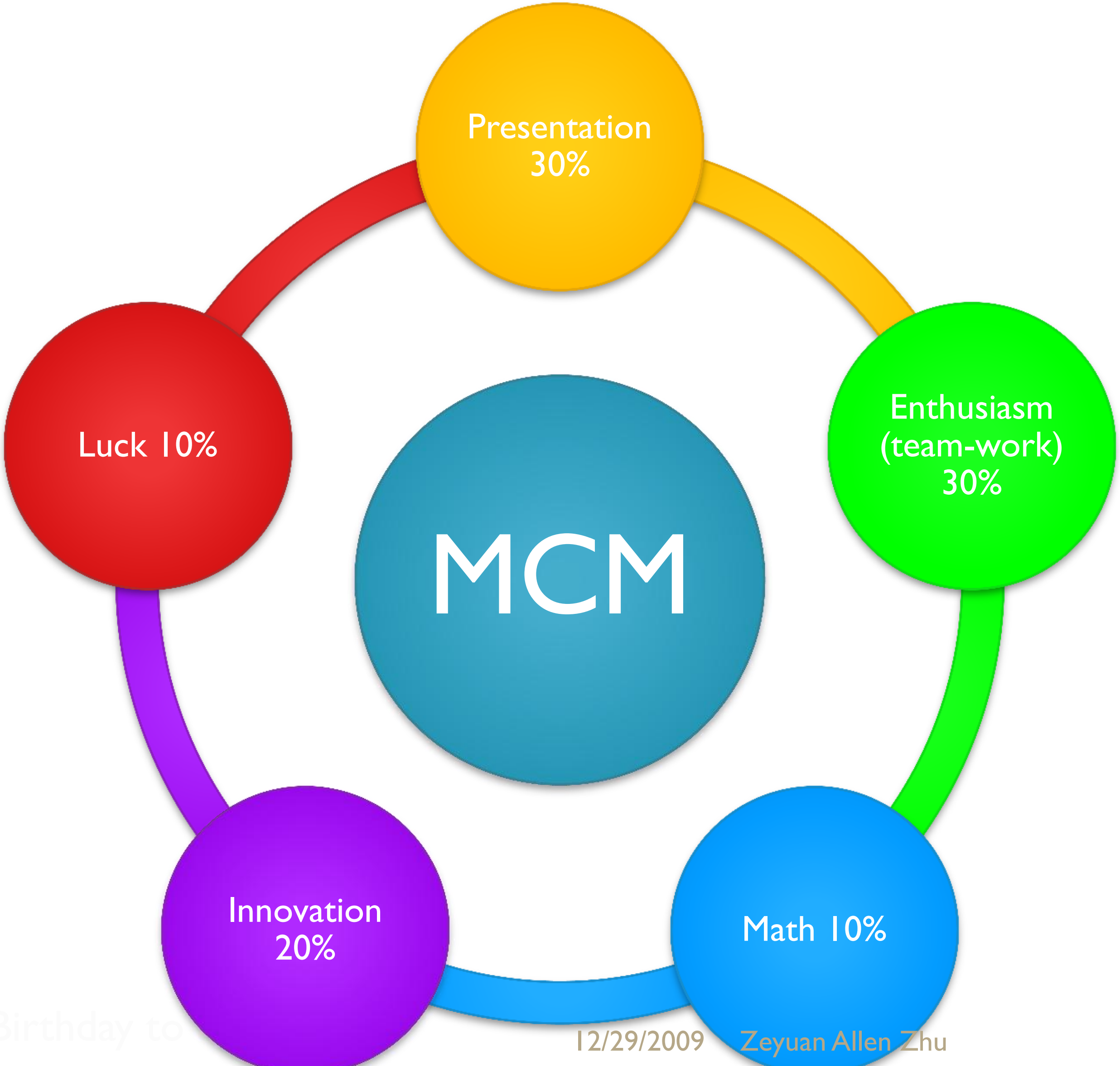
where $r(\boldsymbol{w})$ is the regularizer with a positive weight σ, and $l(\langle \boldsymbol{w}, \boldsymbol{\theta} \rangle; \boldsymbol{\theta})$ is a mapping that calculates the cost or regret by the linear predicting value $\langle \boldsymbol{w}, \boldsymbol{\theta} \rangle$. The expectation is based on a random selection of the sample $\boldsymbol{\theta}$ over the entire sample distribution $Dist$.

Note that the form $\boldsymbol{\theta}$ is used in order to ensure the generality. As an example, $\boldsymbol{\theta}$ can be in the form of $(\boldsymbol{x}, y)$ where $\boldsymbol{x}$ is a vector of features and y is the class identity, adapting (1) to classifications. The loss function l can be for example the SVM hinge loss

Advices for Paper Writing

- The whole paper must be uniquely formatted
  - Same paragraph style for the same *outline level*.
  - Difference between Eng and Chn
    - Watch how other English papers do!
  - Format is the most important at first glance.
- Generate a content page
- Write a good summary page
- # of figures + # of tables >> # of pages
  - Distribute averagely.

Conclusion


Presentation
30%
Enthusiasm
(team-work)
30%
Math 10%
Innovation
20%
Luck 10%
MCM

Bon courage!

"Lift your arms to keep the highest morale"

-- Tianyi Mao

Dedicated to our greatest friendship.

Thanks!   Q&A

Zeyuan Allen Zhu

Fundamental Science Class, Dept. of Physics

Tsinghua University

zhuzeyuan@hotmail.com