For office use only
T1 ________________
T2 ________________
T3 ________________
T4 ________________

For office use only
F1 ________________
F2 ________________
F3 ________________
F4 ________________

2016

19th Annual High School Mathematical Contest in Modeling (HiMCM) Summary Sheet

Team Control Number: 6256

Problem Chosen: A

Summary

Organizing a triathlon event requires taking multiple factors into account. We need to minimize the road closure time as well as alleviating the congestion condition during the competition. In the meantime, we need to satisfy the participants as much as possible and enhance enjoyment of taking part. To address these tasks, we build three main models:

In Model A, we borrow the concept of traffic flow and traffic density to measure the total congestion of the competition of different waves with their respective interval time innovatively, but we have found problems that might introduce inaccuracies into our model.

On the basis of our basic Model, we further improve our model by quantifying congestion of the competition using the total number of surpassing that takes place during the triathlon. Surpassing is the situation where one athlete would get ahead of the other athlete. Based on the international regulations that we have found to ensure the safety of the athletes, we come out with our final arrangement schedule for different divisions and the optimum interval time between waves. This arrangement will lead us to balance the road closure time and congestion.

In Model C, we further explore our model of the best ratio between the swimming, cycling and running distance. This gives us the lowest congestion situation while satisfying the road closure time of 5.5 hours.

Lastly, some sensitivity analysis are done which prove the stability of our model and some key results and arrangements are listed.

Tryathlon

Team # 6256

November 21st, 2016

Contents

1 Introduction

Triathlon, a challenging multiple-stage competition involving three continuous and sequential endurance disciplines, was invented in the 1970s and made its Olympic debut at 2000 Sydney Games. As an alternative to the rigors of track training, triathlon has gained popularity among millions of professional athletes and amateurs. In a standard triathlon which we need to organize in our task, athletes would run over a course of 1500m swim, 40km cycle and 10km of run.

For the purpose of attracting more athletes at professional and premier levels as well as maintaining their enjoyment of participating in the games since we are working with the Super Tread Race Company and the Mayor, we are required to:

1 minimize the congestion take place in the competition

2 minimize the total road closure time

To meet the preceding requirements, we develop models providing our arrangement of the games regarding following aspects:

1 putting all athletes into several different divisions

2 arranging the starting time schedule for different divisions

Then we would consider adjusting the race distance of swimming, cycling and running and explore the advantages that would be gained in terms of congestion and road closure time.

2 General Assumptions

1 We neglect weather influence exerted on road closure time and the extent of competition congestion.

It is hard to quantify and predict the weather condition, and the triathlon Committee will determine whether to hold the game based on official criteria[1] if the weather worsens. Thus, we could neglect the impact of weather condition fluctuation on our final result.

2 We assume that circulation of the race track would not add any extra congestion.

Since the circulation use of the race track are highly dependent of the location that holds the race and are usually designed to avoid congestion, we could neglect the congestion caused by the circulation of the race track.

[1] Figures from *http://triathlon.com.hk/en/races/local-events-weather-policy*

3 Model A: Basic Model

3.1 Model Overview

In this Model, we gave out a basic and conventional solution of measuring road closure time and congestion condition. By utilizing linear programming, we limit the closure time within 5.5 hours. Apart from the time limitation, we establish another model to evaluate congestion through borrowing the concept of traffic flow and road density to evaluate the congestion condition of the competition.

3.2 Symbol Table

Table 1

Symbol Table for **Model A: Basic Model**

Symbol	Definition
	Codes
a_{ij}	the j^{th} triathlete in i^{th} wave
m_i	the number of triathletes in i^{th} wave
M_i	the set of all the triathletes in i^{th} wave
t_{Sij} t_{Cij} t_{Rij}	the time spends by j^{th} triathlete in i^{th} wave on swimming, cycling and running
t_{1ij} t_{2ij}	the time spends by j^{th} triathlete in i^{th} wave in the transition area from swimming to cycling, or from cycling to running
t_{ij}	the time spends by j^{th} triathlete in i^{th} wave during the whole triathlon
r_i	the interval time of the start times between i^{th} wave and $(i+1)^{th}$ wave
l	a specific distance at a time
L	the distance athletes of group i take to cover at a certain speed
t	a certain time in a course

P	the time that athletes of wave i take to cover a certain distance at a specific speed
n_i	the number of athletes in i^{th} wave
R_i	a time delay after the starting gun goes off
v	the speed that athletes of wave i take to cover the distance L
	Constants
n	the number of participants
k	the number of divisions
N	the set of divisions
	Variables
T	the time of closing the roads
Q_i	the traffic flow function of the number of athletes passing through a certain distance per unit time
ρ	the traffic density function of the number of athletes passing through a certain distance

3.3 Road closure time

From our definition, the relationship between variables could be concisely expressed as follows:

$$M_i = \left\{ a_{ij} \,\middle|\, i, j \in Z^+ \quad i \le k \quad j \le m_i \right\},$$

$$\text{For} \left| M_i \right| = m_i \quad m_i < n;$$

$$N = \left\{ M_i \,\middle|\, i \in Z^+ \quad i \le k \right\},$$

$$\text{For} \left| N \right| = k \quad k \le n.$$

The final time for an individual is calculated by combining its swimming time, transition time from swimming to cycling, cycling time, transition time from cycling to running and running time, as expressed:

$$t_{ij} = t_{Sij} + t_{Cij} + t_{Rij} + t_{1ij} + t_{2ij},$$

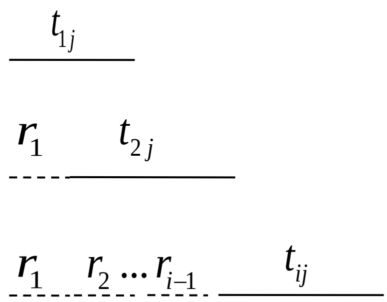

Figure 1 timeline

Figure 1 shows a timeline of the triathlon competition. Clearly, the total road closure time should equal to the longest time for a certain athlete to finish his or her competition and the sum of all interval time before the wave that he or she is in. This could be expressed as following:

$$T = (\sum_{1}^{i-1} r_u + t_{ij})_{\max}$$

According to the Mayor's request, we should keep the road closure time within 5.5 hours:

$$T \le 5.5h\ .$$

3.4 Traffic Flow

Congestion is generally referred to the state where the triathletes are being blocked during the competition by previous contestants. Therefore, we could borrow the concepts of traffic flow and road density from the field of transportation to describe how "crowded" is the entire triathlon event.

The traffic flow Q_i is the number of athletes in wave i passing through a certain distance per unit time (athletes per second), which could be expressed as follows:

$$Q_i(l,t) = n_i f_i^P(t|l)$$

For n_i is the number of the athletes in wave i and $f_i^P(t|l)$ is the density function of the needed time for all the athletes to go through this particular distance.

Therefore, the total traffic flow for this triathlete could be calculated by summing up the traffic flow for all the groups, which is shown as below:

$$Q(l,t)=\sum_{i}Q_i(1,t-R_i),$$

For R_i denotes a time delay between the wave i and the $(i+1)$ group

Traffic density is the number of athletes passing through a per unit distance at a particular moment. Through basic Physics knowledge, we know that:

$$Q=v\rho$$

Accordingly, we express the traffic density of wave i and the total traffic density as:

$$\rho_i(l,t)=\frac{t}{l}Q_i(l,t);$$

$$\rho(l,t)=\sum_{i}\rho_i(l,t-R_k),$$

Then we could evaluate the congestion or the crowding problems in this triathlon event based on the total traffic density. However, there are still limitations to this model. Obviously, in order to utilize this model, we have to assume that the speed for each athlete remains the same throughout the competition. From the data set that is provided for us, we know that this is not the case in reality. If we choose to calculate the speed of the athletes for different disciplines separately, it raises problems when deciding on the delaying time of different waves of cycling and running.

Therefore, the predictions and evaluations of the congestion situation in this model might not be accurate. In order to solve this problem, we decide to improve our model by considering congestion from another perspective.

4 Model B: Improved Model

4.1 Model Overview

Based on Model A which evinces several problems that we cannot solve directly, we further improve our model by proposing an innovative idea of quantifying congestion using the number of surpassing occurring in the entire competition. Then we gave out our arrangement schedule on the basis of several regulations of International Triathlon Union while aiming for minimizing the road closure time and congestion condition of the triathlon simultaneously.

4.2 Symbol Table

Table 2

Symbol Table for **Model B**

Symbol	Definition
	Codes
t	a certain time in the course
	Constants
$t_{s'}$ $t_{c'}$	the time when swimming and cycling ends for the last athlete
t_c t_R	the time when cycling and running begins for the first athlete
X	the number of overtaking
	Variables
$S_{ij}(t)$ $C_{ij}(t)$ $R_{ij}(t)$	the distance travelled by j^{th} triathlete in i^{th} wave on swimming, cycling and running at the moment of t

4.3 Model Construction

Apart from road closure time which we have already given out a detailed solution, the measurement of congestion should also be taken into consideration. In a triathlon competition, congestion occur mainly because there are people in front of one particular athlete which would slow down his or her speed. Thusly, we could use the total number of people that all the athletes surpassed throughout the course to evaluate the congestion extent of the competition. Since surpasses might occur during the transition period which do not really count as congestions in the competition, we take out these special cases by separating the entire triathlon into three stages, swimming, cycling and running.

We define $S_{ij}(t)$ as a function that matches a particular athlete from the moment of t to the distance that he or she has travelled in the competition. And we define a random variable, S'_{ijmnt} would take the value of 1 when a certain athlete has surpassed another through the swimming process since in this case the athlete that he or she has surpassed has caused congestion to him or her, otherwise, it would take the value of 0. This random variable could be expressed as follows:

When $0 < t < t_{S'}$,

$$S'_{ijmnt} = \begin{cases} 1 & S_{ij(t-1)} \le S_{mn(t-1)} \quad S_{ij(t)} > S_{mn(t)} \\ 0 & S_{ij(t-1)} \ge S_{mn(t-1)} \quad S_{ij(t)} < S_{mn(t)} \end{cases}$$

Similarly, for the functions $C_{ij}(t), R_{ij}(t)$ they demonstrate the distance that a certain athlete travels during cycling and running at the moment of t, the congestion of the competition through the cycling and running could be evaluated and expressed as follows:

When $t_C < t < t_{C'}$,

$$C'_{ijmnt} = \begin{cases} 1 & C_{ij(t-1)} \le C_{mn(t-1)} \quad C_{ij(t)} > C_{mn(t)} \\ 0 & C_{ij(t-1)} \ge C_{mn(t-1)} \quad C_{ij(t)} < C_{mn(t)} \end{cases}.$$

When $t_{C'} < t < t_R$,

$$R'_{ijmnt} = \begin{cases} 1 & R_{ij(t-1)} \le R_{mn(t-1)} \quad R_{ij(t)} > R_{mn(t)} \\ 0 & R_{ij(t-1)} \ge R_{mn(t-1)} \quad R_{ij(t)} < R_{mn(t)} \end{cases}.$$

To find the total congestion condition of the competition for all the athletes, we sum up the number of surpasses occuring in three different stages.

$$X = \sum S'_{ijmnt} + \sum C'_{ijmnt} + \sum R'_{ijmnt} \quad .$$

4.4 A Geometric Interpretation

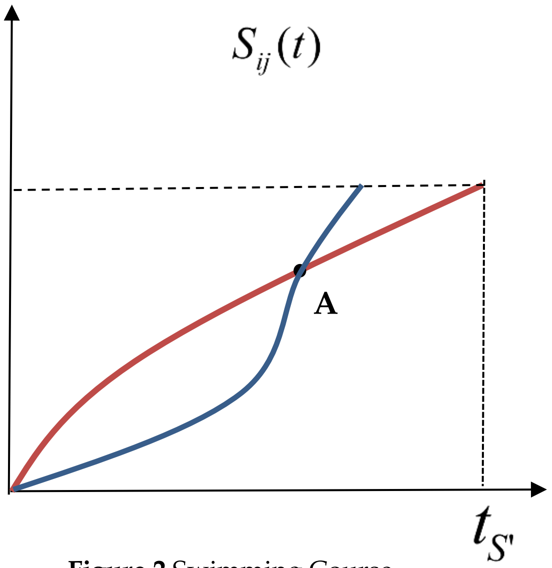

Figure 2 Swimming Course

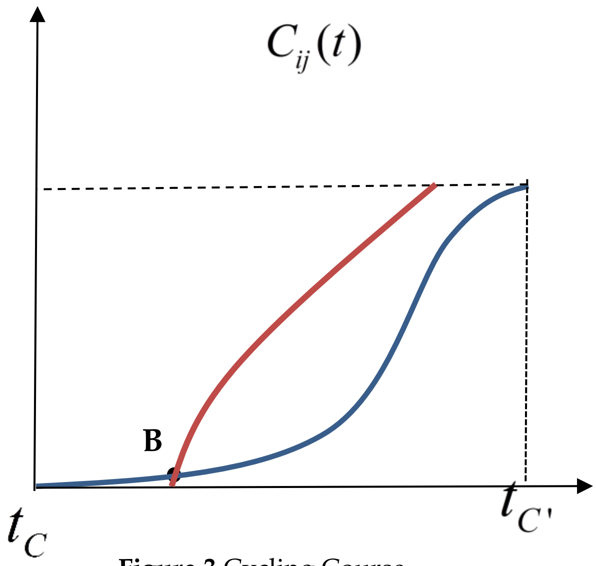

Figure 3 Cycling Course

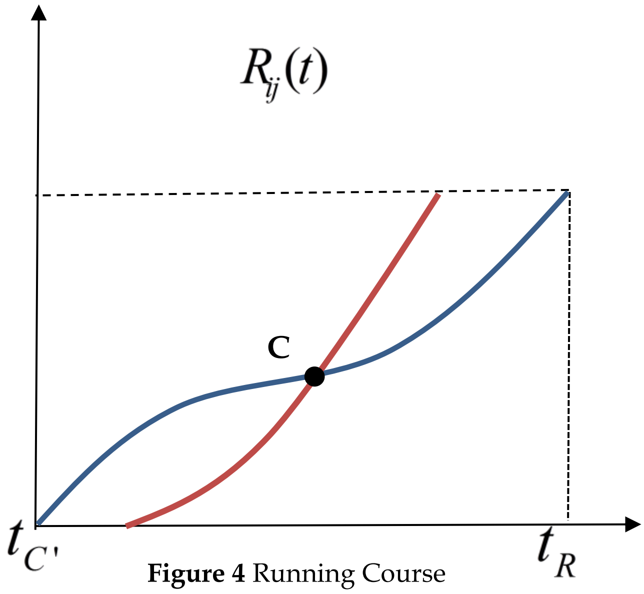

Figure 4 Running Course

The figures shown above gives out a geometric interpretation of our model during different courses and stages:

1 The blue line and the red line show the distance-time relation of two randomly selected athletes.

2 Congestion takes place when a particular athlete surpasses another which is represented by the intersection points like A, B, and C.

3 The transition periods are cut off from the graphs, since, obviously, the surpassing occur during these periods would not result in congestion during the competition.

4 The total number of intersection points shown all of the graphs would be the total number of surpassing happening in the entire competition, which could express the congestion condition.

4.5 Arrangements and Results

According to ITU (International Triathlon Union) Competition Rules, in order to ensure the safety of the athletes, there are limits on the quantity of athletes starting at the same time (in the same wave).

1 For a standard triathlon, the maximum number of athletes in an elite group would be 75 and the maximum number of athletes in age group would be 200.

2 Athletes from the same age group can be divided into two or more wave starts provided these waves would not include athletes from a different Age Group.

3 The interval time between different starting waves should be no less than five minutes

From past triathlon events, the open group athletes would be divided according to their gender and age (usually five ages form a group). The dataset provides more than 3000 data, while in the competition that we organize, only around 2000 participants are being expected. Thus, we reduce the number of participants proportionally for different groups.

For the purpose of attracting more professional athletes and enhancing the enjoyment of every participants, we divide them into different divisions according to preceding requirements. The total number of athletes in corresponding groups are shown in the table below:

Table 3

Number of athletes for different divisions (groups)

Divisions	Estimated Number of People
M Pro	5
M Premier	46
F Pro	4
F Premier	56
M Age 40~44	115
M Age 40~44	115
M Age 30~34	113
M Age 30~34	113
M Age 25~29	152
M Age 45~49	178
M Age 35~39	122
M Age 35~39	122
M Age 50~54	138
M Age 18~24	22
M Age 55~59	97
M Age 60~64	46
F Age 18~34	200
M Age >=65	16
F Age 35~44	157
F Age 45~54	106
F Age >=55	31
ATH	14
CLY	35

It is conventionally acknowledged that the competition should start with the professional and premier groups since they are usually the most competitive and should not be disturbed in the competition by slower athletes. For open Clydesdale and open Athena groups who might not finish the competition, it would be reasonable to place them at the end. By utilizing computer programs, we came out with our arrangement schedule for age groups and all divisions that would lead us to minimum road closure time and congestion condition, expressed in the table below:

Table 4

Arrangement Schedule

Wave	Divisions	Estimated Number of People	Estimated Total Number of People	Interval Time from the Last Wave (min)
Wave 1	M Pro	5	51	0
	M Premier	46		
Wave 2	F Pro	4	60	6
	F Premier	56		
Wave 3	M Age 40~44	115	115	7
Wave 4	M Age 40~44	115	115	6
Wave 5	M Age 30~34	113	113	6
Wave 6	M Age 30~34	113	113	6
Wave 7	M Age 25~29	152	152	6
Wave 8	M Age 45~49	178	178	6
Wave 9	M Age 35~39	122	122	7
Wave 10	M Age 35~39	122	122	6
Wave 11	M Age 50~54	138	160	7
	M Age 18~24	22		
Wave 12	M Age 55~59	97	143	7
	M Age 60~64	46		
Wave 13	F Age 18~34	200	200	7
Wave 14	M Age >=65	16	173	6
	F Age 35~44	157		
Wave 15	F Age 45~54	106	186	7
	F Age >=55	31		
	ATH	14		
	CLY	35		

Conclusions:

If we adopt this arrangement which is the best strategy, then total road closure time would be 329.52 min while the congestion condition would be a total of 820652.

5 Model C: Comparisons Between Strategies

5.1 Model Overview

In order to provide an optimized strategy considering minimizing the time of closing local roads, we adjust the distances of our triathlon according to the international triathlon classification. For the same notion in model A, we work out each participant's new time spend on each section and express the closing time by linear programming.

5.2 Symbol Table

Table 5 Symbol table for **Comparison Model**

Symbol	Definition
	Codes
S' C' R'	new distance of swimming, cycling and running
$v_{Sij}^{'}$ $v_{Cij}^{'}$ $v_{Rij}^{'}$	average speed of j^{th} triathlete in i^{th} wave during swimming, cycling and running after changing the distance
$t_{Sij}^{'}$ $t_{Cij}^{'}$ $t_{Rij}^{'}$	time spend by j^{th} triathlete in i^{th} wave during swimming, cycling and running after changing the distance

5.3 Model Construction

According to the data in Table 6, the triathlon is divided into four kinds based on its distance.

Table 6 **Types** of Triathlons

	Swim(K)	Cycle(K)	Run(K)	Total(K)
Olympic	1.5	40	10	51.5
ITU-Long Distance	3.8	180	42.195	225.195
Long Distance	3	120	25	148
Sprint	0.75	20	5	25.75

Since we are going to organize a traditional open Olympic triathlon, which consists of a 1500m swim, a 40K bike ride and a 10K run,

$$S + C + R = 51.5,$$

For S , C and R represent the original distance of triathlon.

Denoting average time spend by j^{th} triathlete in i^{th} wave during swimming, cycling and running before changing the distance by t_{Sij} , t_{Cij} and t_{Rij} , which is

$$v_{Sij} = \frac{1.5}{t_{Sij}} \quad v_{Cij} = \frac{40}{t_{Cij}} \quad v_{Rij} = \frac{10}{t_{Rij}}.$$

Then, we express the new speed v'_{Sij} , v'_{Cij} and v'_{Rij} as by substituting the above equations, which is

$$t'_{Sij} = \frac{S'}{v_{Sij}} = \frac{S' t_{Sij}}{1.5}, \qquad t'_{Cij} = \frac{C'}{v_{Cij}} = \frac{C' t_{Cij}}{40}, \qquad t'_{Rij} = \frac{R'}{v_{R'ij}} = \frac{R' t_{Rij}}{10},$$

$$\text{And}\, t'_{ij} = t'_{Sij} + t'_{Cij} + t'_{Rij} + t_{1ij} + t_{2ij},$$

For t'_{ij} represents the total time spend by j^{th} triathlete in i^{th} wave during the whole triathlon, and we assume that the time participants spend in transition areas will always remain the same.

Based on our Model A, the time of closing the local roads is

$$T = \left(\sum_{1}^{i-1} r_u + t'_{ij}\right)_{\max} \quad T \le 5.5$$

Taking into account of the international triathlon rules that constrain the minimum distance of each sport, we come out with the optimal ratio between these three sports which is Swimming 3800m, Cycling 20000m, and running 27700m. In this arrangement, the congestion would be 539857 and the road closure time would be 329.6 minutes.

6 Sensitivity Analysis

6.1 Velocity V.S. Road Closure Time and Congestion

Since the datasets provide us the past record of the athletes' performances. There might be some fluctuations in their actual performance. Therefore, we conduct this sensitivity analysis between the velocity of the athletes and our congestion condition as well as road closure time. The results show a 10% change in velocity would only lead to a change in congestion of around 1% and a change in road closure time of 4 %.

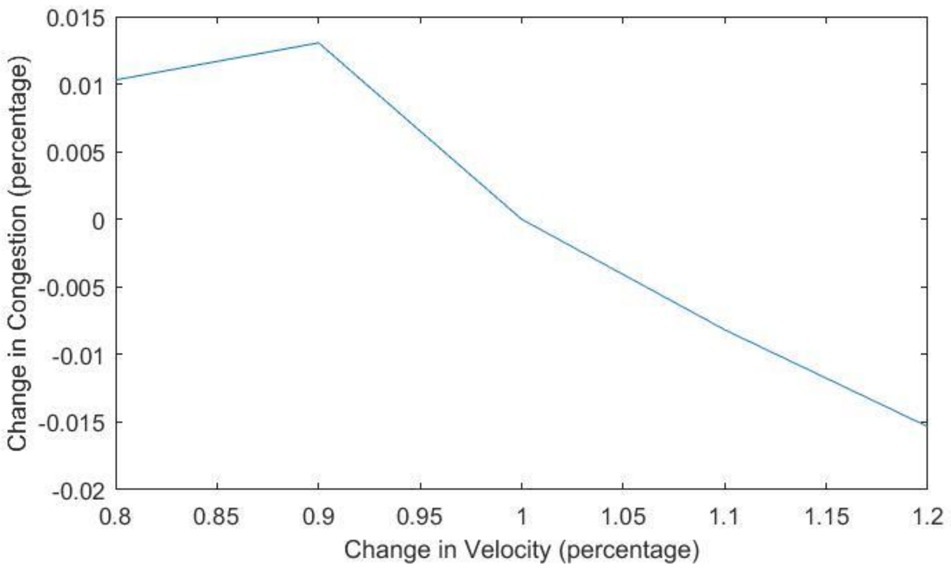

Figure 5 Change in Velocity V.S. Change in Congestion

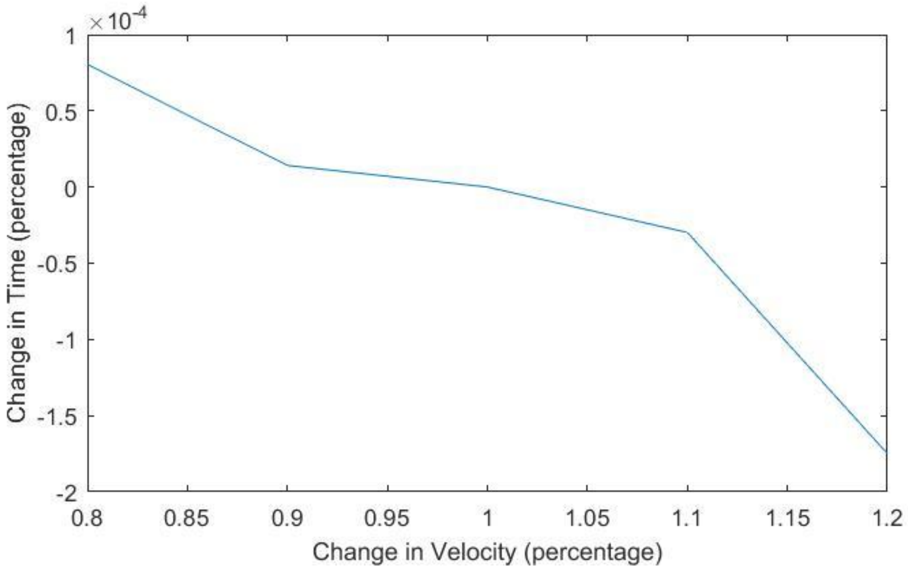

Figure 6 Change in Velocity V.S. Change in Time

6.2 Athletes Number V.S. Road Closure Time and Congestion

Although we expect around 2000 athletes in the competition, there might be some fluctuations in reality when organizing the event. Therefore, we conduct this sensitivity analysis between the number of the participants and the road closure time and congestion condition. The results show that with 2200 participants, the road closure time would change by 0.1% while the congestion would change by around 20%.

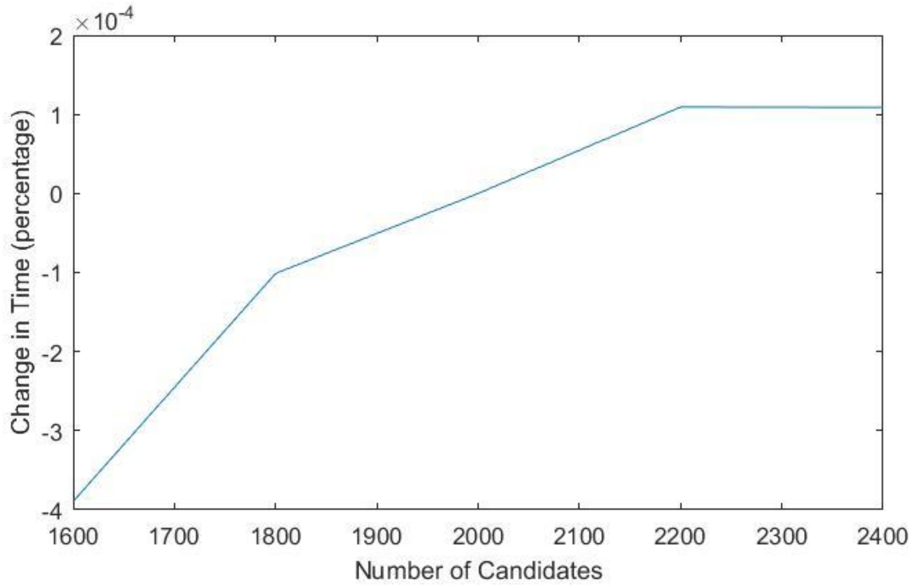

Figure 7 Number of Candidates V.S. Change in Road Closure Time

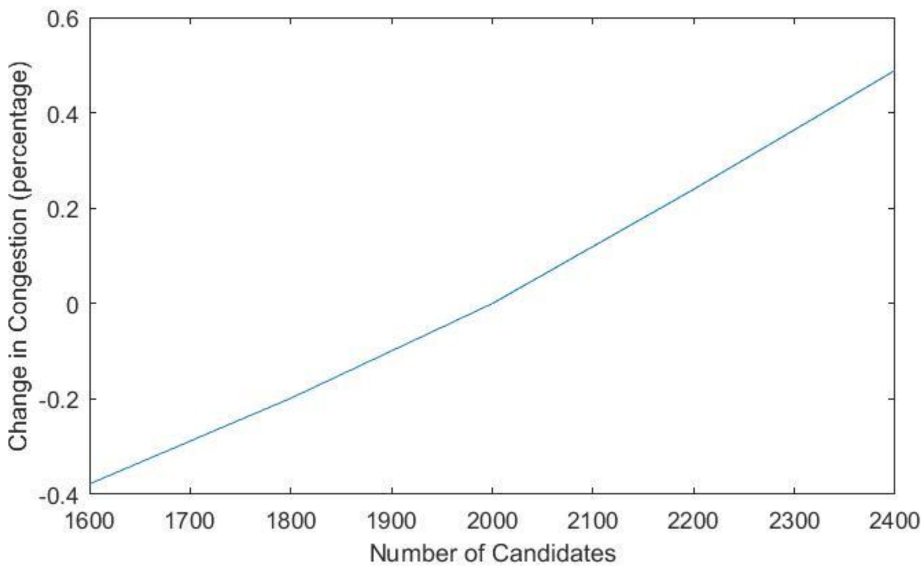

Figure 8 Number of Candidates V.S. Change in Congestion Condition

7 Strengths and Weaknesses

7.1 Strengths

- **Stability**

  From the results of our sensitivity analysis, most of the results show that our model is stable and would not show significant changes due to change in velocity or change in the number of participants.

- **Innovation**

  In our model, we take an innovative approach to measure congestion, which is shown in Model B, using the number of surpasses that takes place to evaluate the congestion condition of the competition. We also uses the concept of traffic flow to model the congestion, which is a novel idea.

- **Reliability**

  We looked up regulations of the ITU on the organizing of the event (the maximum number of wave starts, the minimum interval time) in order to ensure our model and arrangement reasonable and reliable.

7.2 Weaknesses

- **Limited**

  In order to fit as many athletes into the game as possible, we have to reject athletes with total performance over 4 hours from the game which is a total of 64 athletes which is less than 5% of the total number of participants.

8 A letter to the Mayor

Dear Mayor:

As an alternative to the rigors of track training, *triathlon* has gained popularity among millions of professional athletes and amateurs. In order to support the local organization, our team aims to provide best strategies to organize a traditional open Olympic triathlon which contains a course of 1500m swim, 40K bike ride and a 10K run for about 2000 participants.

Since our strategy needs to be feasible and enjoyable that will promote the Super Tread Race Company's brand and sales this year (so that they will support our race again next year), we do not set any required qualification time for participants but do provide awards to the top finishers in each division

To find a balance between the number of divisions and the length of total closure time on local roads, we build up three models by minimizing the congestions and closure time.

The first model, Basic Model, is the basis of all. We first limits the roads' closure time which is less than 5.5 hours in order to prevent the heavy losses by closing local roads.

The second model, Improved Model, is able to evaluate congestions without giving the information of athletes' speeds by utilizing geometry analysis on individuals. We constrain the maximum number of athletes in certain groups based on the information given by ITU (International Triathlon Union) in a standard triathlon, and assume that each age group could be divided into several waves. By selecting the data from 3000 to 2000 participants proportionally, we come out with the divisions shown in the table on the second page.

Then in the third model, we further explore how the change of distance will reduce the congestions within the limitation of 5.5 hours roads' closure time and find out that there exists a proportion in total distance of three events. Swimming should be 3800m, Cycling should be 20000m and Running should be 27700m.

Hope our work can contribute to organize the triathlon this year.

Best regards,

Mathematical Modeling Team #6256

Arrangement Schedule

Wave	**Divisions**	**Estimated Number of People**	**Estimated Total Number of People**	**Interval Time from the Last Wave (min)**
Wave 1	M Pro	5	51	0
	M Premier	46		
Wave 2	F Pro	4	60	6
	F Premier	56		
Wave 3	M Age 40~44	115	115	7
Wave 4	M Age 40~44	115	115	6
Wave 5	M Age 30~34	113	113	6
Wave 6	M Age 30~34	113	113	6
Wave 7	M Age 25~29	152	152	6
Wave 8	M Age 45~49	178	178	6
Wave 9	M Age 35~39	122	122	7
Wave 10	M Age 35~39	122	122	6
Wave 11	M Age 50~54	138	160	7
	M Age 18~24	22		
Wave 12	M Age 55~59	97	143	7
	M Age 60~64	46		
Wave 13	F Age 18~34	200	200	7
Wave 14	M Age >=65	16	173	6
	F Age 35~44	157		
Wave 15	F Age 45~54	106	186	7
	F Age >=55	31		
	ATH	14		
	CLY	35		

References

[1] IRONMAN, the "Fast Wave" – first start wave of IRONMAN Austria 2015 opens for 400 fast swimmers [EB/OL] http://ap.ironman.com/triathlon/events/emea/ironman/austria/athletes/starting-waves.aspx#axzz4Pmq8ef00

[2]Treiber M. Crowd Flow Modeling of Athletes in Mass Sports Events: A Macroscopic Approach[J]. Physics, 2014:21-29.

[3] ITU Competition Rules, International Triathlon Union[S], 17-12-2014

[4]Qiyuan Jiang, Mathematical Model[M], Beijing: Higher Education Press, 1993

[5] John van Rijn ,ROAD CAPACITIES, INDEVELOPMENT, Edition 2004.

[6]Statistical prediction of the mathematical modeling technology[M].Jiangsu science and Technology Press, 2014

[7]Start Lists, ITU World Triathlon Series[EB/OL]. http://wts.triathlon.org/start_lists

[8] Kerner B S. Traffic Congestion, Modeling Approaches to[M]. Springer New York, 2009.

[9] China Triathlon Union, Administrative Measures for Amateur Athletes, 2015