HIGH SCHOOL MATHEMATICAL CONTEST IN MODELING

OUTSTANDING PAPERS

HiMCM

The contest offers students the opportunity to compete in a team setting using applied mathematics in the solving of real-world problems.

November 2014

Additional support provided by the National Council of Teachers of Mathematics (NCTM),
the Mathematical Association of America (MAA),
and the Institute for Operations Research and Management Sciences (INFORMS).

Editor's Comments

This is our seventeenth HiMCM special issue. Since space does not permit printing all nine outstanding papers, this special section includes abridged versions of two papers and summaries from the other six. We wish to emphasize that the papers were not written with publication in mind. Given the 36 hours that teams have to work on the problems and prepare their papers, it is remarkable how much they accomplished and how well written many of the papers are. The unabridged papers from all National and Regional Outstanding teams are on the 2014 HiMCM CD-ROM, which is available from COMAP. ❏

Contest Director's Article

William P. Fox

Department of Defense Analysis
Naval Postgraduate School
Monterey, CA 93943
wpfox@nps.edu

We should all celebrate since the High School Mathematical Contest in Modeling (HiMCM) completed its seventeenth contest year. It is and continues to be a fantastic endeavor for students, advisors, schools, and judges. We have had schools ask for speakers to come in and discuss the modeling process so that their teams can improve and compete. We hope this trend continues. The mathematical modeling ability of students and their advisors is very evident in the professional submissions and work being done. The contest is still moving ahead, growing with a positive first derivative, and consistent with our positive experiences from previous HiMCM contests. We hope that this contest growth continues. Figure 1 is a plot of the growth over time. The trend over the last few years has been an exponential increase.

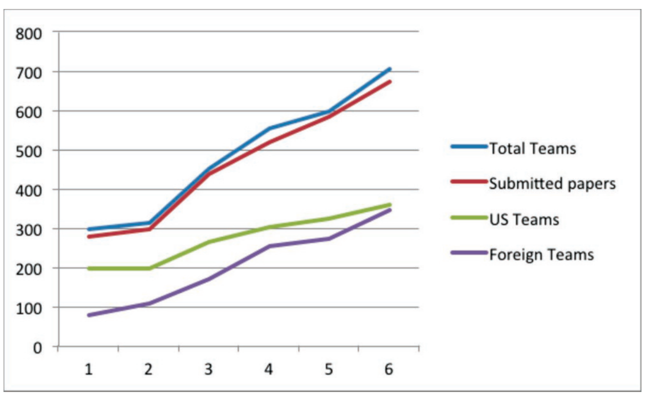

Figure 1: Number of HiMCM teams vs. contest year from 2009-2014

This year the contest had 705 teams and 671 total submissions. This represents an increase of about 17.22% over last year. We had 360 U.S. teams and 345 foreign teams, representing 10.4% and 26.8% growth, respectively. In the United States, these teams represented 79 schools and 24 states. China represented about 93% of the foreign entries. This year we had 2,689 students compete in the contest. This represents an increase of 17.37%. Of the 2,689 students, 1,001 or about 37.2% were female students. Since the beginning we have had 19,822 total participants, of which 7,231 or 36.47% have been female. We are proud of the number of female students that HiMCM attracts. We feel this is remarkable, and we hope that all competing students, male and female, continue on to some STEM education.

The teams accomplished the vision of our founders by providing *unique* and *creative mathematical* solutions to complex open-ended real-world problems. This year the students had a choice of two problems both of which represent "real-world issues." (See the Judges' Commen- tary that follows for the problem statements.)

Commendation: All students and advisors are congratulated for their varied and creative mathematical efforts. Of the 705 registered teams, 671 submitted solutions, which is a 95.17% completion rate. All teams are encouraged to submit their solutions because it is in doing the contest that learning and awards are achieved. These 671 were broken down as follows: 361 doing problem A and 310 doing Problem B. The thirty-six continuous hours to work on the problem provided for quality papers; teams are commended for the overall quality of their work. Judges frequently comment about the amount of material students put together within the 36 hours. It is amazing.

Teams again proved to the judges that they had "fun" with their chosen problems, demonstrating research initiative and creativity in their solutions. This year's effort was again a success!

Judging: We ran three regional judging sites in December 2014. The regional sites were:

- Naval Postgraduate School in Monterey, CA
- Francis Marion University in Florence, SC
- Carroll College in Helena, MN.

Each site judged papers for problems A and B. The papers judged at each regional site may or may not have been from their respective region. Papers were judged as Finalist, Meritorious, Honorable Mention, and Successful Participant. All finalist papers from the Regional competitions were sent to the National Judging in San Antonio, Texas. This year's national judging consisted of ten judges from both academia (high school and college) and industry. They judged the papers and chose the "*best of the best*" as National Outstanding. We usually discuss between 8-12 papers in the final round so all these papers were awarded either "National Finalist" or "National Outstanding". This year there were 19 papers in the final discussions. Nine were deemed National Outstanding and ten National Finalist. These were the "best of the best." The National Judges commended the regional judges for their efforts and found the results were very consistent. We feel that this regional structure provides a good structure for the future as the contest grows.

Seventeenth Annual HiMCM Statistics

Problem	Outstanding	%	National Finalist	%	Finalist	%	Meritorious	%	Honorable Mention	%	Successful Participant	%	Unsuccessful	%
A	5	1%	6	1%	22	1%	70	21%	120	35%	138	40%	2	1%
B	4	1%	4	1%	20	1%	65	23%	106	36%	106	37%	3	1%
Total	9	1%	10	1%	42	1%	135	22%	226	36%	244	38%	5	1%

Judging Results

Problem	National Outstanding	National Finalist	Finalist	Meritorious	Honorable Mention	Successful Participant	Unsuccessful	Total
A	5	6	22	68	114	139	0	383
%	1%	1%	6%	19%	32%	39%	0%	
B	4	4	20	66	106	105	2	198
%	1%	1%	7%	21%	35%	34%	>1%	
Total	9	10	42	134	220	244	2	581
%	1%	2%	6%	20%	33%	37%	>1%	

National Outstanding Teams

Chesterfield County Math/Science High School, VA

Evanston Township High School, IL

High School Affiliated to Fudan University, China

Hong Kong International School

Illinois Mathematics and Science Academy

North Carolina School of Science and Mathematics (2)

Palo Alto High School, CA

Round Rock High School, TX

National Finalist Teams

American Heritage School Plantation, FL

High School Affiliated to Fudan University, China

Hong Kong International School

Illinois Mathematics and Science Academy

Palo Alto High School, CA

Shanghai High School International Division, China

Shanghai Foreign Language School, China

Shenzhen College Of International Education, China

The Stony Brook School, NY

Washtenaw International High School, MI

Common Core State Standards: The director and the judges asked that we add this paragraph. Many of us have read the Common Core Standards and clearly realize the mapping of this contest to the Common Core mathematics standards. This contest provides a vehicle for using mathematics to build models to represent and to understand real-world behavior in a quantitative way. It enables student teams to look for patterns and think logically about mathematics and its role in our lives. Perhaps in a future *Consortium* article we will dissect a problem (paper) and map the standards into it.

General Judging Comments: The Judges' Commentary that follows provides specific comments on the solutions to each problem. As contest director and head judge, I would like to speak generally about solutions from a judge's point of view. Papers need to follow the contest rules posted on the contest site. There is no excuse for not following the rules. Papers need to be coherent, concise, clear, and well written. It is not the number of pages but the ability to complete all contests requirement in a succinct and clear fashion that matters. Students should use spelling and grammar checkers before submitting a paper. Foreign papers should ensure that all symbols in tables and graphs are in English. Papers should use at least 12-point font. Students need to restate the problem in their own words so that the judges can determine the focus of the paper. Papers that explain the development of the model, assumptions with justifications, and its solutions and then support the findings mathematically usually do well. Modeling assumptions need to be listed and justified, but only those that come to bear on the solution (this can be part of simplifying the model). Laundry lists of assumptions that are never referred to in the context of model development are not considered relevant and deter from a paper's quality. The mathematical model needs to be clearly developed, and all variables that are used need to be well defined. Teams that merely present a model to use without any development do not generally do well. Teams that use existing models should reference them during the exposition and not just include a reference/bibliography in the back of the paper. This is true for all graphs and tables taken from the literature. Give proper credit as you go. Thinking outside of the "box" is also considered important by judges. This varies from problem to problem but usually includes model extensions or sensitivity analysis of the solution to the team's inputs. Students need to attempt to validate their model even if by numerical example or intuition. A clear conclusion and answers to specific scenario questions are all key components. The strengths and weakness section is where the team can reflect on their solution and comment on the model's strengths and weaknesses. Attention to detail and proofreading the paper prior to final submission are also important since the judges look for clarity and style. Citations are also very important within the paper as well as either a reference or bibliography page at the end. We encourage citations within the paper in sections that deal directly with data and figures, graphs, or tables. Most papers did not use references within the paper, yet we saw many tables or graphs that obviously came from websites. We have noticed an increase in the use of Wikipedia. Teams need to realize that, although useful, the information might not be accurate. Teams need to acknowledge this fact.

Facts from the 17th Annual Contest

- A wide range of schools/teams competed, including teams from Turkey, Singapore, Jordan, Finland, Hong Kong, and China.
- The 705 registered teams from U.S. and International institutions represent a 17.9% increase in participation.
- There were 2689 student participants, 1686 (63%) male and unspecified and 1001 (37%) female.
- Schools from twenty-four states participated.

The Future: The contest, which attempts to give the under-represented an opportunity to compete and achieve success in mathematics, appears well on its way in meeting this important goal.

We continue to strive to improve the contest, and we want the contest to grow. Any school/team can enter, as there are no restrictions on the number of schools or the numbers of teams from a school. A regional judging structure is established based on the number of teams.

These are exciting times for our high school students. Mathematics continues to be more than learning skills and operations. Mathematics is a language that involves our daily lives. Applying the mathematical principles that one learns is key to future success. The ability to recognize problems, formulate a mathematical model, use technology, and communicate and reflect on one's work is key to success. Students gain confidence by tackling ill-defined problems and working together to generate a solution. Applying mathematics is a team sport!

Advisors need only be motivators and facilitators. They should encourage students to be creative and imaginative. It is not the technique used but the process that discovers how assumptions drive the techniques that are fundamental. Let students practice to be problem solvers. Let me encourage all high school mathematics faculty to get involved, encourage your students, make mathematics relevant, and open the doors to success.

Mathematical modeling is an art and a science. Teach your students through modeling to think critically, communicate effectively, and be confident, competent problem solvers for this new century.

International flavor of the contest: The award format has changed slightly as the contest continues to grow internationally.

Previous Designation and New Designation:

Successful Participant is still Successful Participant.

Honorable Mention is still Honorable Mention.

Meritorious is still Meritorious.

Regional Outstanding Winner is now designated a Finalist.

The papers in the final discussion are now designated National Finalists.

National Outstanding Winner is now designated an Outstanding Winner.

Contest Dates: Mark your calendars early: the next HiMCM will be held in November 2015. Registrations are due in October 2015. Teams will have a consecutive 36-hour block within the contest window to complete the problem and electronically submit a solution. Teams can register via the Internet at **www.comap.com.**

MathModels.org: It is highly recommended that participants in this contest as well as prospective participants take a look at the modeling web site, **www.mathmodels .org**, which has a wealth of information and resources.

The International Mathematical Modeling Challenge, IM^{2}C - An Announcement and an Invitation

We are pleased to announce the establishment of a new international secondary school mathematical modeling competition to be launched in 2015. Please see the following website for details. **www.immchallenge.org**

Rationale: The purpose of the IM^{2}C is to promote the teaching of mathematical modeling and applications at all educational levels for all students. It is based on the firm belief that students and teachers need to experience the power of mathematics to help better understand, analyze and solve real-world problems outside of mathematics itself – and to do so in realistic contexts. The Challenge is being launched in the spirit of promoting educational change.

For many years there has been an increased recognition of the importance of mathematical modeling from universities, government, and industry. Modeling courses have proliferated in undergraduate and graduate departments of mathematical sciences worldwide. Several university modeling competitions are growing and flourishing. Yet at the school level there are only a few such competitions with many fewer students, even amid signs of the growing recognition of modeling's centrality.

One important way to influence secondary school culture, and teaching and learning practices, is to institute a high-level, prestigious new secondary school contest – one that will have both national and international recognition. We have therefore founded the International Mathematical Modeling Challenge (IM^{2}C).

This will be a true team competition, held over a number of days with students able to use any inanimate resources. A major emphasis is for students to experience working with mathematics in a way that mirrors the way the world works with mathematics. Real mathematical problems are messy. Real problems don't come after chapters in a mathematics text so that you know what techniques to use. Real problems require a mix of different kinds of mathematics for their analysis and solution. And real problems take time and teamwork. The IM^{2}C will provide students with a deeper experience both of how mathematics can explain our world and what working with mathematics looks like.

When fully operational the Challenge, inspired by other major international contests, will consist of two rounds of competition. Once the national teams have been chosen, in the first round, they will work on a common problem and submit their solutions to a judging panel. Then there will be a second round hosted each year by a different country, in which the national teams present their solutions in person and engage in additional modeling experiences together.

Plans for 2015: In 2015 our plans call for a contest of only the first phase. We are inviting countries to choose two teams of up to four students with one teacher advisor. Teams will receive a common problem in mid-April. The contest will last until mid-May. During that time teams can choose five (5) consecutive days to work together on the problem. All solutions must be sent in by the faculty advisor, who must certify that the students followed the rules of the contest.

Papers will be judged in early June by an international expert committee and winners announced by mid-June. Papers will be designated as Outstanding, Meritorious, and Honorable Mention, with appropriate plaques and certificates given in the name of students and their schools.

Judges' Commentary, 14th Annual HiMCM Problem A: Unloading Commuter Trains

Trains arrive often at a central Station, the nexus for many commuter trains from suburbs of larger cities on a "commuter" line. Most trains are long (perhaps 10 or more cars long). The distance a passenger has to walk to exit the train area is quite long. Each train car has only two exits, one near each end so that the cars can carry as many people as possible. Each train car has a center aisle and there are two seats on one side and three seats on the other for each row of seats.

To exit a typical station of interest, passengers must exit the car, and then make their way to a stairway to get to the next level to exit the station. Usually these trains are crowded so there is a "fan" of passengers from the train trying to get up the stairway. The stairway could accommodate two columns of people exiting to the top of the stairs.

Most commuter train platforms have two tracks adjacent to the platform. In the worst case, if two fully occupied trains arrived at the same time, it might take a long time for all the passengers to get up to the main level of the station.

Build a mathematical model to estimate the amount of time for a passenger to reach the street level of the station to exit the complex. Assume there are n cars to a train, each car has length d. The length of the platform is p, and the number of stairs in each staircase is q.

Use your model to specifically optimize (minimize) the time traveled to reach street level to exit a station for the following:

Requirement 1. One fully occupied train's passengers to exit the train, and ascend the stairs to reach the street access level of the station

Requirement 2. Two fully occupied trains' passengers (all passengers exit onto a common platform) to exit the trains, and ascend the stairs to reach the street access level of the station.

Requirement 3. If you could redesign the location of the stairways along the platform, where should these stairways be placed to minimize the time for one or two trains' passengers to exit the station?

Requirement 4. How does the time to street level vary with the number s of stairways that one builds?

Requirement 5. How does the time vary if the stairways can accommodate k people, k an integer greater than one?

In addition to the HiMCM format, prepare a short non-technical article to the director of transportation explaining why they should adopt your model to improve exiting a station.

Judge's Comments on Problem A

William P. Fox, Naval Postgraduate School
Problem Author: Joseph Malkevitch

This problem could be addressed through various mathematical models. It was envisioned that students might use linear or integer programming for stair locations and quantity because of the way the problem was posed. We found student approaches both rich and robust. Many teams called what they did optimization but neither used a true optimization technique nor clearly defined their method or decision process as to why it was "optimal". Many teams simulated the process using some Monte Carlo Method. Those teams and future teams are reminded that judges do not read the computer codes attached—teams are required to supply either a flow chart or algorithm for the code they develop and include that in the body of the solution. Judges felt that sensitivity analysis would be a useful discriminator.

Problem B: The Next Plague?

In 2014, the world saw the infectious Ebola virus spreading in western Africa. Throughout human history, epidemics have come and gone with some infecting and/or killing thousands and lasting for years and others taking less of a human toll. Some believe these events are just nature's way of controlling the growth of a species while others think they could be a conspiracy or deliberate act to cause harm. This problem will most likely come down to how to expend (or not expend) scarce resources (doctors, containment facilities, money, research, serums, etc…) to deal with a crisis.

Situation: A routine humanitarian mission on an island in Indonesia reported a small village where almost half of its 300 inhabitants are showing similar symptoms. In the past week, 15 of the "infected" have died. This village is known to trade with nearby villages and other islands. Your modeling team works for a major center of disease control in the capital of your country (or if you prefer, for the International World Health Organization).

Requirement 1: Develop a mathematical model(s) that performs the following functions as well as how/when to best allocate these scarce resources and…

- Determines and classifies the type and severity of the spread of the disease
- Determines if an epidemic is contained or not
- Triggers appropriate measures (when to treat, when to transport victims, when to restrict movement, when to let a disease run its course, etc…) to contain a disease Note: While you may want to start with the well-known "SIR" family of models for parts of this problem, consider others, modifications to the SIR, multiple models, or creating your own.

Requirement 2: Based on the information given, your model, and the assumptions your team has made, what initial recommendations does your team have for your country's center for disease control? (Give 3-5 recommendations with justifications)

Additional Situational Information: A multi-national research team just returned to your country's capital after spending 7 days gathering information in the infected village.

Requirement 3: You can ask them up to 3 questions to improve your model. What would you ask and why?

Additional Situational Information: The multi-national research team concluded that the disease:

- Appears to spread through contact with bodily fluids of an infected person
- The elderly and children are more likely to die if infected
- A nearby island is starting to show similar signs of infection
- One of the researchers that returned to your capital appears infected

Requirement 4: How does the additional information above change/modify your model?

Requirement 5: Write a one-page synopsis of your findings for your local non-technical news outlet.

Judge's Comments for Problem B

William P. Fox, Naval Postgraduate School
Problem Author: Jack Picciuto, Industry

Most teams used a basic SIR model or one of its related differential equation or difference equation models (SIR, SEIR, etc) to answer the initial question. Teams searched the Internet for anything on the subject and it appeared many teams lifted and used equations and graphs directly from the Internet without direct attribution.

General Comments: The judges were surprised with the amount of time and effort put into the modeling. Some papers were unnecessarily long, over 70 pages, without appendices or computer codes. Researching the Internet and putting everything one finds into a solution paper is not necessary. Teams should be brief in their explanations of models that were considered and then not used. The lengths of the papers are sometimes artificial because every possible model form has been included in the discussion although rarely used. Teams are encouraged to include only what they used to answer the posed problem.

Graphs were also an issue. Legends and axes should be labeled and all graphs should make sense. Papers were judged in final hard copy, not electronically. Thus, references to color in a graph were not helpful if the printed copy supplied by the team was not in color. Such teams should restrict their graphs to various gray-scales and perhaps differing markings (line, dashes, dots, etc.)

The executive summaries for the most part are still poorly written, although occasionally they appear to be getting a little better. This has been an ongoing issue since the contest began. Faculty advisors should spend some time with their teams and advise them to write a good summary. Many summaries appear to be written before the teams start and only state how they will solve the problem. Summaries need to be written last and should contain the results of the model as well as a brief explanation of the problem and why their results should be used. The executive summary should entice the reader, in our case the judge, to read the paper. In the real world if the summary is not strong and with "good" results, management may never read the paper.

These comments are also applicable to the non-technical papers or "memos." We do not want equations and methods listed here, but the facts as to why your model is applicable with results and why your model's results are important to the reader.

Again few teams, if any, did sensitivity analysis or error analysis with their models. With the potential *randomness* in each of the problems, this becomes a crucial element. For example, in the Ebola problem a number of infected was reported. How accurate is that number? Could more be affected and not reported? Most likely those numbers are not exact, so sensitivity analysis should be conducted over a range of vales that make sense, and the paper should determine the impact of the solution over that range. This concept is what differentiates this modeling contest from a mathematics contest. In a mathematics contest we have a rate and solve a differential equation, which is what most team did in this problem. In a modeling contest the rate is but a starting point and teams must determine the impact on the solution as all the rates in the problem change.

This issue is true on the commuter train problem as well. Assuming constant exiting speed is a good initial approach, but varying speeds and working with crowds is good sensitivity analysis.

Assumptions with Justifications: All assumptions should have some justification of their importance in the modeling process. Assumptions should not be listed that do not impact on the model used or developed. We also found many teams that listed initial assumptions and then, as they developed model after model, included new assumptions as they went. It would be better to list all assumptions together even if some assumptions were used much later in the paper. For example, a team could state assumption #7 and note that it is used in model #4.

Variables and Units: Teams must define their variables and provide units for each of them.

Computer Generated Solutions: Many papers used extensive computer code. Computer code used to implement mathematical expressions can be a good modeling tool. However, judges expect to see an algorithm or flow chart from which the code was developed. Successful teams provided some explanation or guide to their algorithm(s)—a step-by-step procedure for the judges to follow. Code is not read for papers that reach the final rounds unless the code is accompanied by an algorithm that is clearly and logically explained. A complete algorithm, if code is used, is expected in the body of the paper with the code in an appendix.

Graphs: Judges found many graphs that were not labeled or explained. Many graphs did not appear to convey information used by the teams. All graphs need a verbal explanation of what the team expects the reader (judge) to gain (or see) from the graph. Legends, labels, and points of interest need to be clearly visible and understandable, even if hand written. Graphs taken from other sources should be referenced and annotated.

Summaries: These are still, for the most part, the weakest parts of papers. These should be written after the solution is found. They should contain results and not details. They should include the "bottom line" and the key ideas used in obtaining the solution. They should include the particular questions addressed and their answers. Teams should consider a brief three paragraph approach: a *restatement of the problem* in their own words, a short description of *their method and solution* to the problem (without giving any mathematical expressions), and the *conclusions* providing the numerical answers in context. A good reference is to read it and ask yourselves the question "*does this letter entice me to read the paper?*"

Restatement of the Problem: Problem restatements are important for teams to move from the general case to the specific case. They allow teams to refine their thinking to give their model uniqueness and a creative touch.

Assumptions/Justifications: Teams should list only those assumptions that are vital to the building and simplifying of their mathematical model. Assumptions should not be a reiteration of facts given in the problem statement. Assumptions are variables (issues) acting or not acting on the problem. Every assumption should have a justification. We do not want to see "smoke screens" in the hopes that some items listed are what judges want to see. Variables chosen need to be listed with notation and be well defined.

Model: Teams need to show a clear link between the assumptions they listed and the building of their model or models. Too often models and/or equations appeared without any model building effort. Equations taken from other sources should be referenced. It is required of the team to show how the model was built and why it is the model chosen. Teams should not throw out several model forms hoping to WOW the judges, as this does not work. We prefer to see sound modeling based on good reasoning.

In particular for Problem B, many students immediately wrote the SIR model without any discussion as to why it is appropriate. Often the SIR model and even background material were lifted from sources and not referenced.

Anything not created by a team should be referenced where it is used.

Model Testing: Model testing is not the same as testing arithmetic. Teams need to compare results or attempt to verify (even with common sense) their results. Teams that use a computer simulation must provide a clear step-by-step algorithm. Lots of trials and related analysis are required when using a simulation. Sensitivity analysis should be done in order to see how sensitive the simulation is to the model's key parameters. Teams that relate their models to real data are to be complimented.

For example, consider the Ebola problem where we were given information that 15 out of the 300 villagers were diagnosed with the disease. Is the true rate 15/300? What are the chances that more villagers are sick but unaware or have not sought treatment? Thus, analysis should be performed on rates greater than 15/300 to measure that impact. This is sensitivity analysis. Both problems were open for good sensitivity analysis (also deemed "what if" analysis) but the judges found few papers doing any sensitivity analysis.

Conclusions: This section deals with more than just results. Conclusions might also include speculations, extensions, and generalizations. This is where all scenario specific questions should be answered. Teams should ask themselves what other questions would be interesting if they had more time and then tell the judges about their ideas.

Strengths and Weaknesses: Teams should be open and honest here. What could the team have done better?

References: Teams may use references to assist in their modeling, but they must also ***reference the source*** of their assistance. Teams are reminded that only ***inanimate resources*** may be used. Teams cannot call upon real estate agents, bankers, hotel managers, or any other real person to obtain information related to the problem. References should be cited where used and not just listed in the back of the paper. Teams should also have a reference list or bibliography in the back of the paper. It is requested that teams use some in-line reference to graphs, figures, and direct explanation of materials taken from other sources.

Adherence to Rules: Teams are reminded that detailed rules and regulations are posted on the COMAP website. Teams are reminded that they may use only ***inanimate sources*** to obtain information and that the ***36-hour time limit is a consecutive 36 hours.***

Problem A Summary: North Carolina School of Science and Mathematics

Team Members: Vedant Arora, Abhi Kulgod, Howard Li, Katherine Wang
Advisor: Daniel Teague

Train stations have always been hubs for commuters, but a hub becomes a nightmare once congestion prevents people from leaving the station and getting to work. In certain train stations, the placement and size of staircases is often a limiting factor that prevents the continuous flow of movement from the train platform to the street-level exit. Our models seek to minimize the time it takes to leave the train, move across the platform, and go up the stairs based on the given variables: platform length (p), number of stairs (q), train car length (d), and the number of train cars (n). Each train is at least 10 cars long, with each car having three seats on one side and two on the other. The current staircase can only hold two columns of people.

In approaching this problem, we focused on the overall time it takes to leave the station. This time was broken into three components: the time taken to unload a train and walk to the stairs, the time spent in the queue waiting to walk up the stairs, and the time needed to ascend the stairs. Our goal was to minimize the time it takes to clear the train station, meaning the time it takes the last passenger to reach street level.

We created both a mathematical and a computational model to accomplish this goal. Our mathematical model used flow rate differential equations and Euler's method estimates in an Excel spreadsheet to determine how long it takes to get through the stairway-waiting queue. These time values incorporate the time to unload a train, walk to the stairs, and walk up the stairs. The computational model was a NetLogo program that modeled the time taken to completely empty a train station. For almost all of our calculations, the mathematically and computationally calculated times to empty the station were similar or equal.

Using these tools, we then modified the values of our variables to adjust the timeframe. Modifying the variables represents changing the conditions in the station platform. As we initially expected, the width of the stairwell is the most significant factor that impacts the time; allowing at least 6 people to walk up the stairs at one time halves the amount of time it takes to clear the train station. The placement of stairs, number of staircases, and even the arrangement of train cars do not have as great an influence.

Problem A Summary: North Carolina School of Science and Mathematics

Team Members: Franklin Chen, Ebube Chuba, Maggie Knostman, Shreyas Kolavennu
Advisor: Daniel Teague

Our problem deals with configuring the platform of a commuter train station in a way that minimizes the amount of time that it takes for passengers to exit the train and climb the stairs to street level. We used a worst-case scenario of a full train of ten cars arriving at the platform and unloading all of its passengers, who then go directly to the staircase to exit the station. We optimized our configuration by looking at minimizing the average time spent getting to street level, and then minimizing the time spent waiting in line to get to the staircase.

To solve this problem, we divided exiting the platform into three sections of time: time to get off the train and to the staircase, time spent waiting at the stairs, and time spent climbing the stairs. We first found walking speed to be normally distributed. Then, using both a numerical and analytical approach, found the distribution of arrival times of individuals to the staircase(s). The numerical approach used a program to procedurally generate scenarios and average them for an accurate result. The analytical approach found the cumulative distribution function associated with aggregate inverse scaled normal distributions. Once this data or distribution had been acquired, by taking data on how long it takes the average person to walk up a set of stairs, it was then possible to find both the average time to get to a staircase and the average time spent waiting. The time to ascend the stairs was unrelated to this distribution and found to be fixed. Because we assumed that spending time walking to the staircase was preferable to waiting en masse in front of the stairs, we found that positioning staircases at the ends of the platform was more optimal than having them towards the middle.

Both models recommend three configurations that each work best in a different way. If it's a priority to minimize the overall time, a configuration with three staircases positioned 1/6, 1/2, and 5/6 of the way along the platform is best. However, if it's a priority to minimize the crowding, two staircases at either end of the platform is optimal. Finally, if it is desirable to have only one staircase, it is best to position it on one end of the platform.

Problem A Summary: Evanston Township High School

Team Members: Sean Finn-Samuels, Zane Kashner, Graham Straus, Nate Umbanhowar
Advisor: Mark Vondracek

In today's global economy, efficiency counts. In economic systems, where large volumes of information are exchanged at lightning speed, the smallest inefficiencies can have drastic consequences for the productivity and success of the system as a whole. In logistics, corporations engineer their products' design, production and distribution with these small inefficiencies in mind, as an infinitesimal additional cost for one unit can blossom into the difference between profit and bankruptcy when the product is manufactured in large quantities. Similarly, on Wall Street, stockbrokers harness the power of supercomputers to make trades as quickly as possible and obtain an edge. Just as cost affects industry, a minimal temporal disadvantage for a trader can create a devastating financial loss. Now, more than ever, systems of this kind must be engineered with considerations in mind to eliminate the crippling effects of inefficiency.

Our group was tasked with optimizing a system of just this sort: commuter rail. Each day, corporations rely on commuter trains to bring their employees to the workplace, making commuter rail a crucial form of transportation. Specifically, we focused on the exit of the train station as an area for improvement, for this is traditionally a bottleneck for the speed at which commuters can proceed to work.

To model this situation, our group developed a computer simulation of the train station in Java. We created objects for people and staircases and made a simulator that, based on parameters we entered, told us the average time that it takes one person to make if from their train seat to street level. Using this data, we were able to isolate variables such as the number of train cars, car length, number of steps, staircase positioning and staircase carrying capacity and see their effects on average time. Noteworthy findings from our model include the fact that "seat to street" time is nearly a $1/\sqrt{k}$ relationship with k, the number of columns of individuals on a given staircase as well as linear with a number of the aforementioned variables. We manipulated the positioning of a given number of staircases and found that if there are staircases near at the ends of a train, people may exit the station the fastest. Using assumptions about the train station and train itself we aligned our model with industry standards so we could finalize and backup our model with real data. The M8 train manufactured by Kawasaki nearly perfectly aligned with the problem statement, and Grand Central Station in New York City gave us a similar transportation hub to observe and compare. We believe that our model can be used by industry professionals to redesign their train stations to make commuter's commutes faster and more enjoyable, improving the profitability and efficiency of industry in their city as well.

Problem A Summary: High School Affiliated to Fudan University

Team Members: Motong Chen, Shiyu Ji, Mingyuan Wu, Yue Ying
Advisor: Jun Wan

During our investigation of this situation, our team aimed to build an optimization model to analyze and minimize the unloading time of trains. To achieve this goal, we put forward two models, the mathematical model based on queuing theory and the simulation model based on cellular automaton model. Two models proved each other very well.

We first built our mathematical model, in which the exiting process is divided into four time periods: alighting, walking on platforms, queuing, and walking on staircases. We used the knowledge of former research and $d/d/s$ queuing model in order to calculate the time of the four periods. Using this model, we determined that when there are two staircases, each stair of which contains 2 passengers, the total unloading time for one train is around 261 seconds, while that for two trains is approximately 498 seconds.

In addition, we adopted a simulation approach based on cellular automaton model. Different from the mathematical model, the simulation model simulated several periods of unloading processes as a complete process, and considered the effects of congestion on the moving speed and moving rule of passengers. In addition, the model could visualize the simulation process. The simulation model gave approximately the same result as the mathematical model: the unloading time for one train is 251 seconds, while that for two train is 498 seconds. The results convinced us that both of our models are reliable and accurate.

To understand the effect of different factors, we changed variables such as the number of trains arriving, the location and number of staircases, and the capacity of each staircase to see their impact on unloading time. Running

our simulation model, we found that: the most effective way to minimize the unloading time is to increase the number of staircases; the capacity of the staircase is also a very important factor; the location of staircases does not affect the unloading time very much.

Lastly, we wrote a letter to the director of transportation to recommend that our model be used to help design platforms' layout and schedule the transit system. Although the context of our problem is simple, our simulation model can be applied to the design of any complex transfer stations by slightly adjusting some parameters.

Problem B Summary: Round Rock High School

Team Members: Aravind Ashok, Daniel Munoz, Jenny Xu, Samuel Zhang
Advisor: Janet Sanders

As the world population grows and its distribution changes over time, mathematically modeling the spread of disease is an increasingly important means of developing rapid and effective response measures in diverse areas of the world. Quantitative knowledge of factors related to the spread of disease can be used to influence public policy to determine the best allocation of scarce medical resources to best combat disease. Given this need to model the spread of an unknown disease in a small village on an island in Indonesia, we have created a series of models that, when combined, can both predict the progression of said disease over time and recommend the necessary preventative measures.

Our first model uses a typical SIR model to provide an initial, simplified prediction of how the disease might spread over time absent any preventative measures. From this foundation, we then modified the SIR model into the SIRD model, which accounts specifically for deaths over time and contains a modification that allows for modeling the spread of the disease given some relative abundance of vaccines that have an average protection rate (i.e. the chance that a vaccine has of making an individual immune to said disease). Our third model is a decision tree that categorizes the severity of the disease based on quantitative thresholds of individual and community risk. The model then recommends treatment measures based on the category said disease falls under.

In a real world situation, all three of the previously discussed models would be used in conjunction to ascertain an appropriate response to the disease and adjust it over time. The three models all start with the same initial input of known disease characteristics, and result in a categorization of severity from the Decision Tree Model and an estimation of the progression of the disease from the SIR and the SIRD model. Appropriate treatment measures would likely be determined by adjusting the prescriptions of the Decision Tree using conclusions that can be drawn from the other two models. The interconnectivity between the models means that the effectiveness of the implemented treatment measure can be evaluated by re-running the SIRD model with data from the situation where treatment measures are in effect. If the effects of the disease are shown as being significantly mitigated, the effectiveness of the treatment has been confirmed; otherwise, a re-categorization of the disease and a change in the way resources are allocated may be appropriate. Therefore, the interconnectivity of the models serves as a way to confirm the effectiveness of our response.

Aside from the limitations of modeling with limited concrete data and assuming an essentially isolated population of interest, we have nonetheless generated some important insights from our models. Our models support the effectiveness of vaccination measures, especially on small populations, as each susceptible individual has a greater ability to become vaccinated, and classifies the disease as a "Risk Group 4", the most severe classification, as it crosses our quantitative thresholds of high individual and community risk. This given risk group recommends the allocation of medical resources that prioritizes the reduction of community risk over individual risk.

Problem B Summary: Hong Kong International School

Team Members: Martin Chan, Jiou Choi, Charlotte Chui, Hee Won Chung
Advisor: Edgar Fong

In order to determine how severe the spread of the disease is, we began by classifying the disease based on its mode of transmission. This was done by researching the respective reproductive rates (R_0) of representative diseases for each major mode of transmission and matching the range of values found with the R_0 of the unknown disease based on the data given in the problem. Using this method, we determined the disease to be transmitted primarily through air, with a R_0 of 1.79, meaning that on average an infected person will pass it on to 1.79 other people. Using this value, we also determined that the disease is not contained since the number of infected individuals is still increasing since R_0 is greater than 1. If R_0 is less than 1, then the number of cases will slowly decrease and the disease will slowly disappear from the system because infected individuals are not generating enough new cases.

To model the spread of the disease through the population, we modified the SIR model to include a new factor E, which represents the number of individuals in the 'latency phase', which means that they are exposed and infected by the disease but are not showing symptoms yet. This model is known as the SEIR model. Since most diseases have an incubation period before symptoms are manifested, it is necessary to consider E when modeling the spread of the disease.

To determine the appropriate measures to be taken on the island, we first defined 6 different stages of a disease to match appropriate reactions and resource allocations. These stages are closely related to the variables in our SEIR model. I_1 represents the number of individuals that are not showing symptoms (S and E); I_6 represents the number of individuals that have died as a result of the disease (R); I_2, I_3, I_4, and I_5 represent the four stages of visible illness that have varying degrees of severity (I). These 6 variables were multiplied with respective constants that were determined by the amount of resources necessary for individuals in that stage, then added together in a sum W that roughly represents, based on a scale, the amount of and types of resources an individual in certain stage of the disease will require.

However, these resources are not present in infinite amounts on the island, so we created another model to determine which countries could potentially provide such resources to the island. It is likely that Indonesia will provide a certain amount of aid to its own island, however it is also unlikely that it'll be able to provide all the resources needed to contain the disease. As a result, 12 other countries were selected due to their increased likelihood to provide aid for the island based on proximity to Indonesia and the general amount of supplies the country has in relation to Indonesia. The shorter the distance, the more convenient transportation of resources will be to the island, so the distance between the country and the island is defined as d. The general abundance of supplies is quantified by the ratio of GDP of the country to the GDP of Indonesia and is defined as RGDP.

Problem B Summary: Illinois Mathematics and Science Academy

Team Members: Jason Chen, Kushagra Gupta, Paul Nebres, Pranav Sivakumar
Advisor: Phadmakar Patankar

Faced with the challenge of analyzing an unknown disease and determining appropriate actions, our team developed a model to determine the behavior of the disease and simulate its spread in the village. To add accuracy to our model, given relatively little information, we ran the model countless times with different parameters in an attempt to generalize the disease's behavior.

We first looked toward the SIR model as the foundation of our simulation. We split the people of the village into types: Susceptible, Infected, Recovered, as well as Dead. We also accounted for the potential latent period of infection. Next we looked at the Rumor Spreading Model for insight. We had assumed the disease was spread through human interaction rather than the use of vectors or other means. Therefore, we compared the spreading of a rumor to the spreading of a disease, as both require human interaction and involve spreaders and receivers. We manipulated the formula of rumor spreading to fit our model of the spreading of infection, and determined the probability of a particular infected person infecting a susceptible person.

From these two models, we coded using Java our own model that simulated the spread of infection. We used the given values of infected and susceptible to start with and allowed the members of the village to interact each day. However, since we did not know many values crucial to determining the behavior of spread, we set them as modifiable variables, seeing what would happen under different assumed conditions. We ran this program until the number of infected people reached 0; in other words, the disease is exterminated.

We analyzed the data we received to classify the type and severity of the disease, and whether or not the epidemic was contained. We concluded that this disease followed the behavior of a point source outbreak, and the epidemic was quickly contained. From further analysis of our data, we also devised recommendations to prevent the spread of this disease: namely, preventing contact with infected individuals as well as preventing contact between this village and any other village.

When more information was learned about the particular disease, we were able to create a more accurate model. We were also able to take into account different situations- -including the spread of the disease to other municipalities. However, we have faith that the nature of the disease and the consideration of our recommendations will prevent the continued spread of this dangerous disease.

Problem A Paper

Team Members: Eric Foster, Kathryn Li, Allison Zhang, Andrew Lee
Advisor: Radu Toma

1. Introduction and General Interpretation
Congestion in train stations is a major issue, so optimizing time to travel from seat to exit is crucial. In this problem, a station has a pair of tracks on each side of a platform. We assume that there is a 2-column staircase at the end of a platform. As passengers approach a stairway, a queue forms. People exit a train by walking toward the middle of a platform and then turning to walk toward stairs, but they are held up by the queue. Our model accounts for the fact that people congregate at the base of stairs and then inch forward at a rate determined by queue width and stair capacity.

We first address time optimization when one and two trains unload. We discuss changing car length and number of cars to minimize exit time. Then, we find an optimal staircase location. Finally, we determine how number of stairways and number of people accommodated by a stairway affect exit time.

General Assumptions

1. People use the exit closest to their seat. Justification: There is no way to predict how many people exit through each door without this. Also, it's unlikely someone would go out of the way to use an exit that lengthens travel time.

2. Nothing is to be optimized other than exit time. Justification: We were asked to minimize time and nothing else.

3. Except for Requirements 3 and 4, there is one stairway ($s = 1$). Justification: the problem says "the stairway," which implies one.

4. d is car length in rows (e.g., a car of length $2d$ has 2 rows of seats). Justification: this simplifies calculations on the relationship between number of passengers and car length.

5. Passengers are physically mobile, so mobility impediments are not considered. Justification: people with mobility limitations cannot use a commuter train effectively because it would likely require them to walk from the station to their workplace.

6. Train exits are on the side with two seats per row. Justification: a train company wants to maximize profit, and so prefers displacing two seats to three.

7. A person occupies two steps of a staircase at a given time. Justification: climbing stairs requires a foot on the current step and another on the next.

8. People who sit across from an exit do not move parallel to the tracks to exit. Justification: People want to exit quickly, and these do not need to travel down the car.

9. Passengers walk forward from an exit, turn at the middle of a platform, and walk straight to stairs. Justification: As passengers exit, they feel "pushed" toward the middle of the platform. Although some walk past the middle and others walk less than half the width, the numbers doing each are equal.

10. The two trains are identical. Justification: If not, we would need to optimize a function with four variables. Also, the problem gives car length and number.

Numerical Assumptions

11. People move at 145 ft/min inside cars. As they move to an exit through an aisle, they walk at 250 ft/min. They walk across a platform at 250 ft/min. Justification: Transit companies assume such values.

12. Seat width is 17 inches and distance between rows is 3.67 feet. Justification: (Sources: Haughney and Svensson)

13. Platform width is 29.5 feet. Justification: (Source: California)

14. There are 58 stairs in a staircase. Justification: The depth of a typical train tunnel is 394 inches and stair height in the U. S. is 6.75 inches.

15. A train aisle can hold two people. Justification: most trains are designed this way.

Variables

Given:

n = number of cars to a train; as given by the problem, $n \geq 10$

d = length of the car. (By Assumption 4, $d = 1$ means 1 row)

p = length of platform (feet)

q = number of stairs in each staircase

s = number of stairways

k = number of people who can fit side to side on the stairway

Established:

t = time to travel from car door to bottom stairway

l = distance traveled walking from car seat to street level

l_q = total distance traveled by all passengers within a queue

l_p = total distance traveled by all passenger on a platform, but not in a queue

t_q = sum of all times passengers spend traveling in a queue

t_q = sum of all times passengers spend traveling on a platform, but not in a queue

B = people added to a queue for one train

B_t = people added to a queue for two trains

V = queue growth rate

Vt = queue growth rate for two trains

v_q = velocity of people traveling within a queue

v_p = velocity of people traveling on a platform, but not in a queue

v = general speed

x = number of people per car

2. Single Train Model

Approach: We used $l = vt$ to model time to reach street level. We looked at average distance as a function of n and d. We split distance into three parts: from seat to car door, from car to stair bottom, and from stair bottom to street level.

Part 1: We assigned seat positions, as shown in red in Figure 1. Since seat width is 17 inches, those in position 1 move 17 inches to enter an aisle, those in position 2 move 34, and those in position 3 move 51.

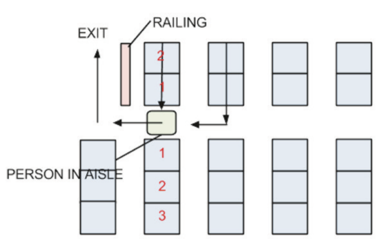

Figure 1: Seat position numbering

We must subtract one position 1 seat and one position 2 seat for each half of each car because there is an exit door in their place.

Position 1 has $d - 1$ people per half car, so time to move into an aisle is $17(d - 1)/1740$. For position 2 this is $34(d - 1)/1740$. For position 3, we use $d/2$ because only one side of a train has position 3 seats (but we do not subtract 1 for the exit door, see Assumption 6). So the people in that position take a total time of $51(d/2)/1740$.

Thus, total time for all to enter aisles is

$$\frac{17(d-1)+34(d-1)+51(d/2)}{1740}, \text{ or } \frac{76.5d-51}{1740} \text{ minutes (1)}$$

Distance between rows is 3.67 feet, so total distance all passengers travel in an aisle to a position where they can travel perpendicular to a train is:

$$3.67\sum_{k=1}^{d/2}(k-1) = 3.67\left(\frac{(d/2)(d/2+1)}{2} - d/2\right) \quad (2)$$

Each person travels $d/2$ times the row they sit in, effectively cutting car length by half. Also, those closer to an exit travel less than those in a car's middle, which is accounted for by the summation of $(k - 1)$, which represents people who have reached aisles.

Since people walk in aisles at 250 ft/min, we get a total time of $3.67\left(\frac{d^2-7d}{2000}\right)$ minutes.

Lastly, total time for all in a car to move from aisle to exit door (the length of three seats: one for getting to aisle from seat plus two seats for walking by a railing) is:

$$\frac{(5d-4)51}{250\times 12} = \frac{255d-204}{3000} \text{ minutes. (3)}$$

Total time for all in a car to exit a train onto a platform is the sum of quantities (1), (2), and (3). Thus, total time for all people to exit a train is

$$2n\left(\frac{76.5d-51}{1740} + 3.67\left(\frac{d^2-7d}{2000}\right) + \frac{22d-204}{3000}\right) \text{ minutes.}$$

Average time to exit a car is negligible compared to time to travel down a platform and up stairs. For instance, when $n = 20$ and $d = 20$, it takes 220.80 seconds to travel through a car, but time on a platform and in a queue is 7647.73 seconds.

To find the number entering a queue per minute, we divided total number, $n(5d - 4)$, by total minutes to exit a train because the number entering a queue per minute equals the number exiting a train per minute. Thus, the queue gets

$$\frac{367(3.67d)^3}{80000} + \frac{4136497(3.67d)^2}{17400000} - \frac{1898321(3.67d)}{4350000} + \frac{868589}{7250}$$

people per minute, which we called B_t.

Net queue growth is our equation above minus the number entering a staircase per minute, which is 120 since a person climbs at 2 stairs/sec and takes up 2 stairs at a time. Since we have 2 columns, 2 people go up the stairs per second, or 120 per minute. The net growth, after simplification and conversion to ft/min is

$\frac{367(3.67d)^3}{80000}+\frac{4136497(3.67d)^2}{17400000}-\frac{1898321(3.67d)}{4350000}+\frac{868589}{7250}$ people/min, which we call V.

Part 2: We found total distance traveled by all passengers in terms of car length (d, in rows), and number of cars (n). Then we found an expression for velocity. We then solved for average time to reach a stairway. Figure 2 shows our scenario with $n = 7$ and $d = 6$, which is explained below.

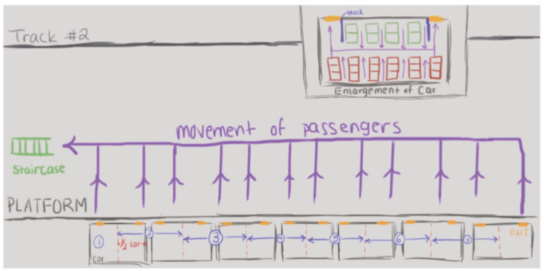

Figure 2: Our scenario with *n* = 7, *d* =6.

Total distance traveled by all people in all cars:

$$\sum_{k=1}^{n}\left((k-1)xd\right)+(1/2)xnd+(1/2)nxw\ ,$$

where x is people per car, d is car length, n is number of cars (blue numbers in Figure 2), and w is platform width. This is found by assuming that a passenger exits through the nearest door (Assumption 1) and goes forward $w/2$ ft. after leaving a train and then turns left or right and walks toward stairs (Assumption 9). Total distance traveled by all passengers from a car to the middle of a platform is thus $\frac{nw(5d-4)}{2}$ (distance 1).

As in Figure 2, we treated adjacent exits of adjacent cars as one exit for $5d - 4$ people since separation between cars is negligible. Thus, if we let the first single exit be #1, the second and third exits be #2, the fourth and fifth be #3, and so on and the last pair of exits be #n (n = number of cars), as shown in Figure 2, then total horizontal distance walked across platforms is $\sum_{k=1}^{n}\left((k-1)d(5d-4)\right)$ (distance 2) At exit #1, passengers walk 0 ft. horizontally. At #2, passengers walk d rows to stairs (see Figure 2). At #3, they walk $2d$ rows, and so on to exit #n from which they walk $(n - 1)d$ rows.

At the last exit (the $n + 1$th), passengers walk nd rows to a stairway (see Figure 2). So the total distance walked is: $\frac{nd(5d-4)}{2}$ (distance 3).

Therefore, our grand total distance (l_{total}) is the sum of distances 1, 2, and 3, which, after simplification is:

$$\frac{n}{2}(5d-4)\left(d(n+2)+27.5\right).$$

Let l_q be total distance traveled through a queue and l_p be total distance traveled on a platform. Let t_p and t_q be total time traveling on a platform and in a queue, respectively. Let V be queue growth rate in ft^2/min. Let v_p and v_q be velocities on a platform and in a queue, respectively.

We know that the total distance travelled by all is: $l_{total} = l_q + l_p$ for l_q.

We also know that $l_q = v_q \times t_q$. v_q is about 0.2712 ft/sec since 2 people leave a queue each second, opening up 8 ft^2 of space. Thus, total queue length decreases by 0.2712 ft/sec.

For each person, queue length when arriving at a queue is the rate of queue change (V) times time spent walking on a platform. A queue does not form until the first group reaches the stairs, and the time it takes them to do so is the time it takes them to walk halfway across a platform: $(w/2)/v_p = (29.5/2 \text{ ft})/(250 \text{ ft/min}) \times (1 \text{ min}/60 \text{ sec}) = 7.08$ sec. So the actual time that we should multiply V by is $t_p - 7.08n(5d - 4)$, where t_p is total time people spend in line because $l_q = l_1 + l_2 + \ldots + l_k = V(t_1 - 7.08 + t_2 - 7.08 + \ldots + t_k - 7.08) = V(t_p - 7.08n(5d - 4))$.
Thus, $l_q = V(t_p - 7.08n(5d - 4))$.

We also know that total distance walking on a platform not in queue (l_p) is $v_p \times t_p$; $l_p = 25/6 \times t_p$.

Solving this system for t_q and t_p:

$$V(t_p - 7.08n(5d-4)) = l_q = v_q \times t_q$$
$$t_q = V(t_p - 7.08n(5d-4))/v_q$$
$$v_q = (8 \text{ ft}^2/\text{s})/29.5\text{ft} \approx 0.2712 \text{ ft/s}$$

We found t_q in terms of t_p, V, and n and d:

$$t_q = \frac{V\left(t_p - 7.08n(5d-4)\right)}{2712}.$$

$$l_{total} - l_q = l_p = v_p \times t_p$$
$$l_{total} - V(t_p - 7.08n(5d-4)) = v_p \times t_p$$
$$v_p = 25/6 \text{ ft/s}$$

We found t_p in terms of l_{total}, V, n, and d:

$$t_p = \frac{l_{total} + V\left(7.08n(5d-4)\right)}{V + 25/6}$$

Now, we have total time ($t_q + t_p$) in terms of n and d, which is what we want.

After substituting t_p into the equation for t_q, adding the highlighted equations above, dividing by number of passengers and simplifying, we have: (*see equation 1 below*), which is average time walking from car exit to stairs on one fully loaded train with d rows per car and n cars.

Part 3: We found average stair-climbing speed to be 2 steps/second. We also set that a person takes up 2 stairs at a time. An average station staircase has 58 steps. Thus, time each person spends climbing stairs is 29 seconds.

Requirement 1: Optimize time traveled to reach street level for passengers on one fully occupied train.

Interpretation: We interpreted this as asking us to relate the time an average person takes to exit a station to n, number of cars, and d, number of rows per car and to find n and d to minimize time or to give a general principle for selecting n and/or d.

Approach: We created a model, A, that is a function of all variables determinant of time to exit a station.

We realized that the only time that matters is time for a passenger to walk on a platform and arrive at stairs, because the time for a passenger to exit a car is constant.

The method for deriving this model is in the Approach section. After using Excel to make a surface plot, we found time is optimized when n and d are minimized.

Our model does not help decide how to minimize wait time as it is unrealistic to have a train with 1 car and 1 row of seats. In our model, however, is also the growth rate of the queue at the bottom of the stairs (V—see Part 1).

V depends only on d because a passenger walks at constant speed when there is no queue (Assumption 9). Thus, the number arriving at a queue is the same for any n since people in farther cars take longer to get to a queue. We found that $V(d) \geq 0$ for $d > 19.7747$. Thus, a queue begins accumulating when the number of rows per car ≥ 20. Therefore, to avoid queueing/fanning, we should have cars with no more than 19 rows. Since a queue does not form for $d \leq 19$, we recommend 19 rows per car.

Requirement 2: Optimize time to reach street level for two fully occupied trains.

Interpretation: This part asks us to use our method in Requirement 1 for two trains (Assumption 10), resulting in twice the number of people as in Requirement 1.

Approach: To optimize time to reach street level, we used the equation in Requirement 1 except we doubled queue grow rate. To look at B_t as queue growth rate, we subtracted queue depletion rate, 120 ppl/min, from B_t, and converted d to feet. Also, the total number of passengers doubles with two trains: $2n(5d - 4)$. Total distance travelled to stairs (assuming there is no queue) doubles as well.

When we plug these changes into the equation for one train and simplify, we get: (*see equation 2 below*)

Average time increases and both n and d increase, so we want n and d to be as small as possible. However, V is negative up to $d = 19$, so we still recommend 19 rows.

Requirement 3: Optimize location of stairways

Interpretation: This part asks us to find a stairway configuration that reduces average time to exit a station. We interpreted this as asking us to research general guidelines and use them to compile a report of optimal locations of one and multiple stairways.

Approach: While we assumed a single staircase at the end of a platform, stations such as Penn Station (PA), Central Station (NY), Union Station (DC), and Cardiff (UK) have staircases in the middle of the platform.

What if there is more than one staircase? We would pair staircases so they face opposite directions. Both Shibuya and Shinjuku in Japan do this and, considering each serves over a billion commuters a year, this placement should minimize traffic.

Equation 1

$$\frac{158767. + 1.37594\, d^6 + 0.0491012\, d^7 + d^3\,(395.77 - 251.827\, n) + d\,(-728253. - 4.64593\, n) + d^5\,(58.4606 + 24.8333\, n) + d^2\,(-534.166 + 59.5455\, n) + d^4\,(743.066 + 345.238\, n)}{-6082.42 + 7608.67\, d - 18.3589\, d^2 + 13.3202\, d^3 + 1.\, d^4}$$

Equation 2

$$\frac{79383.4 + 0.687968\, d^6 + 0.0245506\, d^7 + d^3\,(197.885 - 125.913\, n) + d\,(-364126. - 2.32297\, n) + d^5\,(29.2303 + 12.4166\, n) + d^2\,(-267.083 + 29.7727\, n) + d^4\,(371.533 + 172.619\, n)}{-6082.42 + 7608.67\, d - 18.3589\, d^2 + 13.3202\, d^3 + 1.\, d^4}$$

Requirement 4: Relationship between time and number of stairways

Interpretation: This part asks us to find a number of stairways to minimize average exit time. We interpreted this as asking us to solve for the optimal number of stairways (s) to minimize time as well as to find a relationship between time to street level and s.

Approach: In requirements 1 and 2, we analyzed a queue's effect on exit time. The fan that develops impacts people's ability to leave quickly. If more stairways were available, people would use the emptiest one, minimizing crowding at others.

Increasing the number of stairways reduces queue length until there is no queue, which is when the number of people leaving a train per second divided by the number of staircases is 2 since staircases can take 2 people/sec.

We found the number of people exiting a train per second in Requirement 1. Given that there are s staircases, the rate at which people enter each is:

$$\left(\frac{367d^3}{4800000}+\frac{4136497d^2}{1044000000}-\frac{1898321d}{261000000}+\frac{1411}{43500}\right)\frac{1}{s}$$

Setting this expression equal to 2 and not less than 2 allows the maximum number of people to arrive in each train. Once equality is reached, more staircases are unhelpful. Furthermore, adding staircases reduces platform space. This poses an injury risk that counteracts optimization of time. Thus, the only variable needed to optimize time is the number of people in a queue. As soon as a queue is eliminated, time is optimized.

Therefore, time to street level decreases as s approaches

$$\left(\frac{367d^3}{4800000}+\frac{4136497d^2}{1044000000}-\frac{1898321d}{261000000}+\frac{1411}{43500}\right)\frac{1}{2}$$

from the negative side, but does not change as s becomes greater than is expression. If $d = 20$, $s = 1.027$, meaning that one stai rway still causes queueing while two stairways eliminates Requirement 5: Relationship between time and number of people accommodated by stairways

Interpretation: We interpreted this as asking for a relationship between time walking to a stairway and number of people (k) a stairway accommodates. We chose to use average time to walk from car to stairs. We also defined k as the number of people that can stand side-by-side in a row of stairs since the number of people that can stand on stairs lengthwise should be constant. Thus, only k is related to time walking to street level.

Approach: If the number of people who can fit side by side on a stairway changes, queue length at the bottom of a stairway is affected. k is then equal to the number of people exiting or entering a staircase each second because according to our calculations, each second row of people exits and a new row enters the stairway. Thus, if queue length is reduced by a wider stairway, time to exit a station decreases.

As with requirement 4, widening stairways can reduce time up to a point, which is when stairways are wide enough so that no queue develops at the bottom. In Requirement 1, we found the number of people leaving a train per second (B), which is also the number of people entering a queue per second. Since k equals the number of people that exit/enter a stairway per second (V) the change in number of people in a queue is:

$$\frac{367d^3}{4800000}+\frac{4136497d^2}{1044000000}-\frac{1898321d}{261000000}+\frac{1411}{43500}-k.$$

We then plugged V above into our model, assigned $d = 20$ and $n = 20$, and found t as a function of k. Figure 3 is a graph of this function.

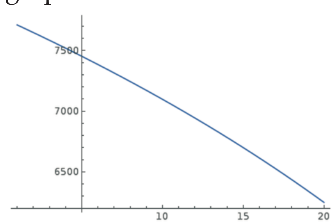

Figure 3: Graph of t as a function of k

As Figure 3 shows, the number of people accommodated in one row of a stairway increases (as stairway width increases), the average time to walk from a train to a stairway decreases nearly linearly because the queue shortens as more people exit.

3. Model Analysis

Strengths and weaknesses:

Our most significant strength is queue analysis. We considered a queue as two-dimensional. Using realistic constants, we modeled the speed of people in a queue. We also modeled queue growth as a function of the number of rows per car and cars per train. Small trains produce no queue; once ridership reaches a certain point, a queue develops. While a simpler model would conclude that the best way to optimize time is by shrinking the train to zero, our model finds a point at which trains can be as large as possible before a queue forms. Outside of distance, crowding and congestion are major determinants of exit time. Modeling the queue allows us to find a point that optimizes the most relevant variables: distance covered through a station and time moving through a queue. Our ability to find this optimum point is the most important strength.

Weaknesses of our model arise from values we negated. For example, when we calculated time to exit a train, we got about 3% of total time to egress from a station. Although 3% is small, it is not necessarily negligible.

Sensitivity Analysis
Our model is accurate even at large numbers of passengers. Given a train with 20 rows per car, we can fit the population of Bedford, MA (12,595 people) in a train with 131 cars. Our model gives 12.22 hours for all to exit the station, which is reasonable.

Problem B Paper: Chesterfield City Math Science High School

Team Members: Sean Brooks, Dylan Finch, Sarah Steele, Samantha Walmsley
Advisor: Pete Peterson

1: Variable Index

Table 1: Variable Index

Symbol	Explanation	Type of Data
I_0	Initial Infected Population	Discrete/Input Variable
C_0	Initial Contagious Population	Discrete/Input Variable
S_0	Initial Susceptible Population	Discrete/Input Variable
D_0	Initial Deceased Population	Discrete/Input Variable
R_0	Initial Removed Population	Discrete/Input Variable
I_n	Infected Population	Discrete/Output Variable
C_n	Contagious Population	Discrete/Output Variable
S_n	Susceptible Population	Discrete/Output Variable
D_n	Deceased Population	Discrete/Output Variable
R_n	Removed Population	Discrete/Output Variable
n	Timestamp in weeks	Constant
i	Chance Infected	Parameter
b	Probability of Contact	Parameter
μ	Death Probability	Parameter
α	Child and Elderly Probability	Parameter
μ_α	Death Probability for Children & Elderly	Parameter

2: Background Research & Problem Statement
A family of viruses that has become well known is Filoviridae, which causes hemorrhagic fever. Ebola is a Filoviridae that originated in West Africa and has spread across the world. Since 2000, there have been 1,290 confirmed Ebola cases; 784 (61%) of which resulted in death. Ebola was first identified during a 1976 African outbreak. Since its discovery, there have been 26 outbreaks, including the current one. All previous outbreaks have been contained. The latest, which began in late 2013, has infected about 14,000 people and is becoming a global threat. In areas without good medical infrastructure and with relatively poor health standards, the disease can spread rapidly and with greater severity than in countries with resources to treat the sick. In the U. S., the mortality rate is 20%.

The given data are from Indonesia, where half of a small village's 300 people had symptoms similar to those of Ebola. 10% of those with symptoms have died. We are tasked to create a model that determines the type and severity of the disease spread, and whether the epidemic is contained. The model also determines best containment measures. Using the model, recommendations can be made for further action.

3: Assumptions & Rationales

1. The virus in the case study is functionally equivalent to Ebola. Because the problem states that symptoms are similar to those of Ebola, it seems logical to model the behavior after Ebola.

2. Those who are removed or deceased do not infect the healthy. Ebola data predict that the virus cannot be passed on after a person recovers, except under unusual circumstances. We also assumed that the removal of bodies is done with care, and that no healthy people are infected by the deceased.

3. Death rates for children (age < 14 years) and elderly (age > 65 years) are equal. Both groups have a similarly increased chance of death from illness, as stated in the problem. We assumed that these probabilities equal for simplicity. In medical statistics that we found, children and elderly are grouped together.

4. Exactly 3 people travel to and from other villages (for trade) daily. We could not find data on inter-village trading in the Philippines. Due to the low number of people in the village, we used a 1% rate.

5. Nobody in the village can provide care for patients prior to the WHO personnel. Treatment of Ebola requires special materials such as IVs, medications for blood pressure, and transfusions.

6. WHO professionals can treat 15 people per day. We assumed that it takes 10 minutes to prepare to treat a victim, 15 minutes to treat a victim, 10 minutes for procedure after treatment, and 5 minutes to travel between houses, for a total of 40 minutes per victim.

7. The incubation period of the disease is 7 days.
This is based on Ebola's mean incubation time. The range is 19 days, with a minimum of 2 days, and a maximum of 21.

8. Medical professionals travel between houses to treat infected people. A town of 300 is unlikely to have suitable medical facilities.

9. 40% of those treated by a medical professional do not recover. In an African Ebola study about 70% died in a week, even with treatment. However, the problem says that 10% died in a week in Indonesia. We used an average.

10. We assume that everyone in the village is the same, except for the state of disease they are in (Susceptible, Infected, Removed, Deceased, Contagious).

Through this, we assume these four points:

a) Those who do not have Ebola are healthy; those who have it are otherwise healthy.

b) The probability that contact with an infected person transmits the disease is the same from person to person.

c) The probability of infection is the same for each person with whom a healthy individual comes into contact.

Although there are different probabilities of infection per person, we determined an average probability based on the number of carriers a person comes into contact with, and whether or not they themselves contracted the virus

d) The number of people in contact with a person who has Ebola is constant.

4: Model Design Process

The start of the modeling process began with a basic SIR model, which assumes that a person moves from being susceptible to being infected, and finally to being removed as immune.

In our version of SIR, a person moves from susceptible to infected, to contagious, and then to removed (immune), or the person dies. The dead are immediately buried in isolated graves, thus no longer infecting others. Our final model is a set of difference equations that take all these factors into account.

Our model in Table 2 has 5 equations. At any time, values from these equations should sum to the total population.

Population Part	Model
Susceptible	$S_{n+1} = S_n\left(1 - i\beta\right)^{C_n}$
Infected	$I_{n+1} = S_n\left(1 - \left(1 - i\beta\right)^{C_n}\right) - C_n$
Contagious	$C_{n+1} = I_n$
Deceased	$D_{n+1} = D_n + C_n\mu$
Removed	$R_{n+1} = R_n + C_n\left(1 - \mu\right)$

Table 2: List of Models

First we looked at the number of people who are still susceptible, modeled in:

$$S_{n+1} = S_n\left(1 - i\beta\right)^{C_n}$$

This equation takes the previous number of susceptible, S_n, and multiplies by the probability that a person comes into contact with and is infected by a contagious person, $(1 - i\beta)^{C_n}$. This is calculated by multiplying the probability a person comes into contact with someone, *ß*, by the probability of infection, *i*, reversing the probability by subtracting from 1. This is raised to the power of the number of contagious infected: $(1 - i\beta)^{C_n}$. We multiply this by S_n to get the number not infected.

Second, we looked at the number infected but not yet contagious. We assume that the period of being infected without being contagious is a single timestamp, or a single iteration of *n* (a week). This is modeled in:

$$I_{n+1} = I_n + S_n\left(1 - i\left(1 - \beta\right)^{C_n}\right) - C_n$$

We got this by taking the opposite of $(1 - i\beta)^{C_n}$ to get $(1 - (1 - i\beta)^{C_n})$, which gives the number infected in a given timestamp. From that we subtract the number of contagious, which equals the number infected from the previous timestamp.

As mentioned, the number of contagious equals the number of infected from the previous timestamp: $C_{n+1} = I_n$. We assume that it takes one timestamp for an infected person to become contagious. Similarly, the number of deceased equals the previous number of deceased added to the number of previously contagious who have died. This is found by multiplying the previous number of contagious by the probability of death, $C_n\mu$. The equation for D_{n+1} calculates the total deceased up to a point. Similarly, the number of recovered is calculated as the opposite.

5: Conclusion

n	S_n	I_n	C_n	D_n	R_n
0	150	120	15	15	0
1	69	81	120	21	9
2	0	69	81	69	81
3	0	0	69	101	130
4	0	0	0	129	171

Table 3: Population Output for Assumed Data

We assigned certain values (Table 3) in order to model disease spread:

1. For S_0, initial susceptible population, 150 is assigned. The problem says that half the population has symptoms. Thus, half has none and is susceptible.

2. For I_0, initial infected population, 85 is assigned.

3. For C_0, initial contagious population, 50 is assigned.

4. For D_0, initial deceased population, 15 is assigned, which is given in the problem.

5. For R_0, initial removed population, 0 is assigned. The problem does not specify any people who recovered.

6. For i, probability of infection, 1 is assigned. In other words everyone who comes into contact with an infected person becomes infected.

7. For b, probability of contact, 0.1 is assigned. To get this, we divided 30, the number contacted in a day, by 300, the starting population.

8. For μ, probability of death, 0.4 is assigned. This is the average of the death probability taken from the problem and that of a known strain of Ebola, 0.6.

In conclusion, the toll of this virus will be rapid and deadly. Table 3 shows that everyone is infected, with 171 recovering (immune) and 129 dying.

6: Classification of Epidemic

A disease is classified as epidemic if it spreads quickly. This disease is as aggressive as Ebola. If it spreads to Earth's entire population, about 2.8 billion would die even with treatment. If it is contained, it will not likely become a global problem. Figure 1 shows what would happen if it spreads to the world, and no action is taken.

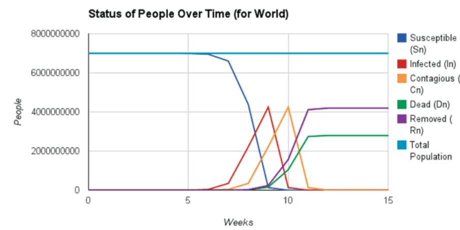

Figure 1: Graph for a world outbreak (given the initial population leaving the village)

7: Addressing Additional Information

1. Appears to spread through contact with bodily fluids of an infected person. Considering this type of spread, the model can be run again with small values for b.

2. The elderly and children are more likely to die if infected. Given that children and the elderly have an increased chance of death, and assuming that this increase is the same for each group (Assumption 1) the model needs to be adjusted for multiple death probabilities. We introduced the variable μ_α, the α signifying related to children and elderly.

The adjusted model accounts for different values for μ, death probability. It takes the proportion of children and elderly in the population. Before simplification, the model for Death and Removal (the two aspects changed by adjustment) was:

$$D_{n+1} = \alpha(D_n + \mu_\alpha C_n) + (1-\alpha)(D_n + \mu C_n)$$
$$R_{n+1} = \alpha(R_n + (1-\mu_\alpha)C_n) + (1-\alpha)(R_n + (1-\mu)C_n)$$

Population Part	Model
Susceptible	$S_{n+1} = S_n(1-i\beta)^{C_n}$
Infected	$I_{n+1} = I_n + S_n\left(1-(1-i\beta)^{C_n}\right) - C_n$
Contagious	$C_{n+1} = C_n + I_n - D_n - R_n$
Deceased	$D_{n+1} = D_n + C_n\left(\alpha(\mu_\alpha - \mu) + \mu\right)$
Removed	$R_{n+1} = R_n + C_n\left(1 - \mu + \alpha(\mu - \mu_\alpha)\right)$

Table 4: List of Adjusted Models

Table 4 shows how proportions of different constants were taken and added to form a whole. Figure 2 is a graph of the adjusted models with $\mu = 0.3$ and $\mu_\alpha = 0.8$.

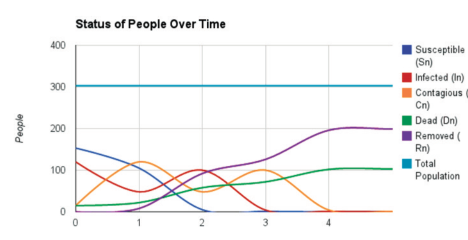

Figure 2: Graph for adjusted models (μ = 0.3 & μa = 0.8)

3. A nearby island is showing signs of infection.
   The island should be isolated. Only medical professionals should enter and leave. More medical professionals will be sent with equipment and supplies. The medical professionals will be quarantined for 21 days after visiting.

4. One of the researchers who returned appears infected.
   This person must be quarantined, along with everyone in his proximity. Also, any location the person traveled to after visiting the village must be isolated for 21 days. We chose 21 because that is the maximum incubation period of the disease.

8: Sensitivity & Error Analysis

Our model is relatively sensitive to change. The death rate is 40%, so 129 of 300 die in 6 weeks. If the rate decreases by 10%, the number of people dead in the same time is 101. If the rate increases by 10%, the number of people dead after 6 weeks is 159. We also tested a smaller change in death rate. If it increases by 1%, 131 people die; if it decreases by 1%, 126 die. If the model were applied to a larger scale, the change in death rate would cause a significant change in its prediction of the number surviving.

If the infection rate increases or decreases 10% from a base of 25%, the number dying stays constant. However, the time it takes the disease to reach everyone changes greatly. At an infection rate of 25%, it takes 6 weeks to reach everyone. If the rate decreases by 10% 3 of the original susceptible pool are never infected. If the rate increases by 10%, the disease reaches everyone in 3 weeks.

Lastly, if the probability of contact changes, it affects the number infected and death rate. If the original 10% decreases by 5%, 4 of the original susceptible pool never get sick, and 3 fewer die. Also, if the probability of contact increases by 5%, the disease reaches everyone by the 4th week, but the number dying stays constant.

The majority of the error in our models is from our many assumptions. Time constraints do not allow us to test every probability or circumstance. Given more time and resources there are a few changes we would like to research:

1. Make b, probability of contact, depend on the remaining susceptible population.Currently b can be changed per iteration of the model, but it remains constant during a run. In actuality, as the number of people dwindles, the probability decreases.

2. The models could be calculated continuously instead of discretely. This would allow for change over time to be monitored. It would also allow for immediate populations to be calculated rather than at weekly intervals.

3. We would like to go to Indonesia and learn more about the health care system. The situation described is in a different environment from one that we are familiar with. More information on the local level would allow for a better model.

4. We would like to add more variables.

9: Recommendations

1. Instruct people to minimize contact, even with apparently healthy people. A virus spreads by physical contact; thus, limiting contact is a way to slow disease spread. Certain symptoms (especially those of early stages) may not be evident.

2. Instruct people to stop travel to other towns and villages. Villages with which the affected village has traded since 3 weeks prior to the outbreak should also restrict travel.

3. Instruct people not to treat the sick.
   Limiting contact (except by medical professionals) is essential to disease control.

4. Instruct people to self-test for symptoms, including fever, fatigue, or muscle pain. Many early symptoms can be detected, and treating the disease early can help minimize severity and spread.

10: Questions

1. What is the actual probability of recovery for those treated by doctors? Knowing this would make the model more accurate. Without an accurate probability of recovery, it is impossible to know how helpful the doctors are.

2. What is the actual average number of people entering/leaving the island a day? Knowing this is important to determining how susceptible the outside world is. To better model the chance of infecting the outside world, this must be accurate.

3. At what stage of disease were the people with symptoms when medical professionals arrived, and how long were they sick? Knowing initial conditions is essential to understanding how a virus develops.